

SISTEMA DE VISÃO TRIDIMENSIONAL PARA A INSPEÇÃO
SUBMARINA DE DUTOS E EQUIPAMENTOS

Lenildo Carqueija Silva

TESE SUBMETIDA AO CORPO DOCENTE DA COORDENAÇÃO
DOS PROGRAMAS DE PÓS-GRADUAÇÃO DE ENGENHARIA DA
UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO PARTE
DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU
DE MESTRE EM CIÊNCIAS EM ENGENHARIA ELÉTRICA.

Aprovada por:

Prof. Antonio Petraglia, Ph.D.

Prof^a. Mariane R. Petraglia, Ph.D.

Prof. Liu Hsu, Dr. d'Etat

Prof. Jacques Szczupak, Ph.D.

RIO DE JANEIRO, RJ – BRASIL
SETEMBRO DE 1998

SILVA, LENILDO CARQUELJA

Sistema de Visão Tridimensional para
a Inspeção Submarina de Dutos e Equipa-
mentos [Rio de Janeiro] 1998

ix, 87 p. 29,7 cm (COPPE/UFRJ,
M.Sc., Engenharia Elétrica, 1998)

Tese–Universidade Federal do Rio de
Janeiro, COPPE

1. Estereofotogrametria

2. Visão Computacional

I. COPPE/UFRJ II. Título (série)

Resumo da Tese apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Mestre em Ciências (M.Sc.)

SISTEMA DE VISÃO TRIDIMENSIONAL PARA A INSPEÇÃO
SUBMARINA DE DUTOS E EQUIPAMENTOS

Lenildo Carqueija Silva

Junho/1998

Orientador: Antonio Petraglia

Programa: Engenharia Elétrica

Este trabalho apresenta um sistema de visão tridimensional desenvolvido para ser utilizado em atividades de inspeção visual de instalações submarinas de petróleo por veículos remotamente operados. Um sistema de visão estéreo em tempo real é utilizado para a aquisição de pares estéreo de imagens que, depois de pré-processadas com o objetivo de melhorar a precisão dos parâmetros a serem extraídos (através de métodos de processamento de imagens como equalização de histograma e detecção de envoltória), são submetidas ao dimensionamento por meio de métodos estereofotogramétricos, com o uso de algoritmos *least squares* não-lineares para a obtenção de coordenadas tridimensionais.

Abstract of Thesis presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Master of Science (M.Sc.)

STEREO VISION SYSTEM FOR SUBMARINE INSPECTION OF
PIPELINES AND EQUIPMENTS

Lenildo Carqueija Silva

June/1998

Advisor: Antonio Petraglia

Program: Electrical Engineering

This work presents a three-dimensional measurement system for inspection activities of submarine petroleum installations by remotely operated vehicles. A real-time stereo vision system is used for the acquisition of stereo pairs of images that, after preprocessing, by applying methods of image processing such as histogram equalization and edge detection, are submitted for dimensioning of objects through stereophotogrammetric computations using non-linear least squares algorithms to obtain three-dimensional coordinates.

Índice

1	Introdução	1
1.1	Considerações Gerais	2
1.2	Fotogrametria: Motivos para seu Uso	3
1.3	Organização do Trabalho	4
2	Sistema de Visão Estéreo	6
2.1	Introdução	6
2.2	Arquitetura	7
2.2.1	Câmeras	7
2.2.2	Placas Digitalizadoras	8
2.2.3	Dispositivo de Visualização	9
2.2.4	Monitor	9
2.3	Operação do Sistema	9
3	Pré-processamento	11
3.1	Operações Pontuais	11
3.1.1	Imagem Negativa	11
3.1.2	Equalização de Histograma	12
3.2	Operações Espaciais	13
3.2.1	Filtragem por Média Espacial	13
3.2.2	Filtragem Passa-Baixa, Passa-Alta e Passa-Faixa	14
3.3	Detecção de Envoltória	16
3.3.1	Aproximações Discretas de Derivadas	18
3.3.2	Tipos de Operadores	19

3.3.3	Operador de Gradiente	21
3.3.4	Operador de Direção	22
3.3.5	Operadores de Laplace	26
3.3.6	Gradientes Estocásticos	28
4	Dimensionamento Tridimensional	31
4.1	Introdução	31
4.2	Orientação Exterior	36
4.3	Orientação Interior	41
4.4	Orientação Relativa	46
4.5	Orientação Absoluta	50
4.6	Triangulação Estéreo	54
4.7	Modelagem dos Erros Sistemáticos	59
5	Aplicação	60
5.1	Imagens Aquisitadas	60
5.2	Pré-processamento	61
5.3	Parâmetros de Orientação Exterior	61
5.4	Parâmetros de Orientação Relativa	67
5.5	Aplicação	68
6	Conclusão	79
A	Soluções <i>Least-Squares</i> Não-Lineares	81
B	A Decomposição do Valor Singular	84

Lista de Figuras

1.1	Diagrama em blocos do Sistema de Visão Tridimensional.	2
2.1	Arquitetura do Sistema de Visão Estéreo.	8
2.2	Exemplo de par estéreo aquistado.	10
3.1	Transformação negativa.	12
3.2	Transformação da equalização de histograma.	13
3.3	Exemplo de equalização de histograma	14
3.4	Máscaras espaciais $a(k, l)$	15
3.5	Filtros espaciais.	15
3.6	Diferentes fontes físicas de envoltórias em uma imagem.	17
3.7	Gradiente de $f(x, y)$ na direção r	19
3.8	Detecção de envoltórias por operador de gradiente.	22
3.9	Exemplos de detecção de envoltórias por operador de gradiente . .	24
3.10	Detecção de envoltórias por operador de direção.	26
3.11	Exemplo de detecção de envoltórias por operador de direção Kirsch. 27	
4.1	Referenciais global e da câmara.	35
4.2	Orientação exterior.	38
5.1	Imagens aquistadas para teste.	62
5.2	Detecção de envoltória aplicada às imagens da Fig. 5.1.	63
5.3	Mapa de envoltória obtido a partir das imagens da figura anterior. 64	
5.4	Pontos usados na determinação da orientação exterior.	66
5.5	Pontos usados na determinação da orientação relativa.	70
5.6	Pontos usados nos testes com o par estéreo 1.	75

5.7	Pontos usados nos testes com o par estéreo 2.	76
5.8	Pontos usados nos testes com o par estéreo 3.	77
5.9	Pontos usados nos testes com o par estéreo 4.	78

Lista de Tabelas

3.1	Operadores diferenciais discretos.	20
3.2	Operadores de gradiente.	23
3.3	Operadores de direção.	25
3.4	Operadores de direção tipo Kirsch.	29
3.5	Operadores de Laplace.	29
3.6	Gradientes estocásticos.	30
5.1	Determinação da orientação exterior	65
5.2	Parâmetros de orientação exterior calculados.	67
5.3	Determinação da orientação relativa.	69
5.4	Parâmetros de orientação relativa calculados.	71
5.5	Resultados dos testes – par estéreo 1.	71
5.6	Resultados dos testes – par estéreo 2.	72
5.7	Resultados dos testes – par estéreo 3.	73
5.8	Resultados dos testes – par estéreo 4.	74

Capítulo 1

Introdução

A atividade de inspeção de equipamentos tem como objetivos acompanhar as condições físicas dos equipamentos durante sua vida útil, investigar as causas de possíveis acidentes e propor soluções para evitar tais acidentes. O uso de métodos de inspeção torna-se necessário em atividades nas quais há riscos elevados à segurança dos equipamentos e de vidas humanas. As atividades voltadas para inspeção de equipamentos procuram atuar com os objetivos de controlar as condições físicas dos equipamentos, identificar causas de deteriorações, e controlar a qualidade dos reparos executados.

Este é o caso da inspeção submarina, onde é necessária a monitoração de equipamentos, tubulações e estruturas trabalhando sob condições hostis, por vezes em altas profundidades. Para a inspeção destas instalações, têm sido desenvolvidas diversas técnicas de ensaios não destrutivos, dentre as quais [1]:

- Ultra-som: a grande velocidade de propagação do som no aço, em torno de 5,9 km/s, permite a utilização desta técnica, onde uma alta velocidade de ensaio é mandatória.
- Medição de potencial eletroquímico: a proteção anticorrosiva de dutos é realizada principalmente pelos revestimentos de pintura e os sistemas de proteção catódica, sendo que neste caso o valor do potencial eletroquímico é o principal parâmetro a ser medido e controlado.

- Correntes parasitas: a inspeção por correntes parasitas é uma técnica de inspeção baseada na introdução da corrente elétrica no material a inspecionar e na observação da interação entre correntes e materiais.
- Partícula magnética: baseia-se na detecção de discontinuidades em materiais ferromagnéticos, com a aplicação de um campo magnético sobre uma estrutura coberta por partículas magnéticas.
- Inspeções visuais: consiste na utilização de aparelhos fotográficos e de filmagem para monitorar cabos, dutos e equipamentos, e para aumentar a eficiência das operações de intervenção submarina.

Destes métodos, o método de inspeções visuais foi o adotado no presente estudo. Sua metodologia consiste em, a partir de imagens da estrutura a avaliar, suprimir objetos que não são de interesse, e realizar um pré-processamento na imagem de forma a dar maior confiabilidade aos parâmetros a serem extraídos; em seguida, através de técnicas de fotogrametria analítica, conseguir o dimensionamento da falha.

Baseado nesta metodologia, foi desenvolvido um sistema modular integrado (*software e hardware*) de visão tridimensional para aquisição e processamento de imagens, a ser aplicado na inspeção visual submarina de dutos e equipamentos.

1.1 Considerações Gerais

O sistema de visão tridimensional abordado no presente estudo pode ser dividido em três estágios distintos, ilustrados na Fig. 1.1.



Figura 1.1: Diagrama em blocos do Sistema de Visão Tridimensional.

O **sistema de visão estéreo** consiste de um par de câmeras de vídeo conectadas a placas digitalizadoras, que adquirem imagens submarinas em tempo real para serem mostradas em três dimensões. Imagens de interesse são digitalizadas e gravadas em arquivo sob a forma de pares estéreo (duas imagens, cada qual de uma câmera, da mesma cena, obtidas no mesmo instante).

As imagens digitalizadas, antes de serem efetivamente utilizadas, são submetidas a um **pré-processamento**, de forma a adequá-las ao processo de dimensionamento, o que inclui técnicas de processamento de imagens tais como a equalização de histograma e a detecção de envoltórias, com o objetivo de melhorar a precisão dos parâmetros a serem extraídos das imagens.

Finalmente, as imagens processadas são submetidas ao **dimensionamento**, através de um algoritmo de dimensionamento que utiliza métodos fotogramétricos para a obtenção de coordenadas tridimensionais, a partir de pontos correspondentes no plano bidimensional das duas imagens do par estéreo. A partir das coordenadas tridimensionais, podem-se conseguir as dimensões de objetos (ou falhas) tridimensionais.

1.2 Fotogrametria: Motivos para seu Uso

A fotogrametria é uma técnica visual para a aquisição de medidas tridimensionais acuradas sem contato com o objeto-alvo a dimensionar. Não se trata de uma técnica recente, tem sido usada há várias décadas em compilação de mapas a partir de fotografias aéreas, inspeção topográfica [10], uso industrial, e construção naval [3].

Há alguns anos, tem aumentado o uso da fotogrametria na área de inspeção visual submarina como um método rápido e eficiente de obterem-se medições tridimensionais precisas. É necessário somente um par de imagens do objeto-alvo, aquisitadas a partir de uma câmera em duas posições diferentes, ou a partir de um par de câmeras, para permitir medições acuradas através destas imagens.

A fotogrametria na área de inspeção submarina foi usada inicialmente como uma ferramenta de inspeção, avaliação de corrosão, falhas de solda, pequenas

áreas de danos e tarefas similares. Depois, começou a ser aplicada também em atividades de vistoria de estruturas submarinas sofisticadas, com uma área de cobertura maior. Algumas das vantagens desta técnica são apresentadas a seguir [2]:

- é uma técnica sem contato ou interferência física com o objeto.
- permite medições de objetos que são difíceis de serem medidos por técnicas convencionais, devido à sua forma ou posição.
- usa fases separadas para a aquisição e processamento de dados, assim reduzindo o tempo de operação submarina, conseqüentemente reduzindo custos.
- o esforço total não aumenta proporcionalmente com o número de pontos a serem medidos.
- a informação extraída pode estar presente em diferentes formas: como medidas discretas, seções, contornos, obtidos a partir de coordenadas tridimensionais calculadas.

No sentido mais restrito, o termo “fotogrametria” refere-se ao processo de medir imagens em uma fotografia. Tradicionalmente, através do uso de pares de fotografias, medidas quantitativas de posições, distâncias e áreas têm sido definidas e aplicadas a mapas topográficos. Toda a informação disponível é então extraída de modelos fotogramétricos por modelos matemáticos simples, porém altamente acurados, usando as coordenadas dos vários pontos imagem (pontos no plano de projeção da imagem). Este método é chamado fotogrametria numérica ou analítica, e será o utilizado neste estudo.

1.3 Organização do Trabalho

Este trabalho é dividido em cinco capítulos e dois apêndices. O capítulo 2 descreve o sistema de visão estéreo adotado, e que é usado para a visualização e

aquisição de pares estéreo de imagens. O capítulo 3 aborda os métodos utilizados no pré-processamento das imagens aquiritadas, com a finalidade de se obter uma melhoria da qualidade da imagem, incluindo equalização de histograma e detecção de envoltórias. O capítulo 4 discute os métodos estereofotogramétricos utilizados para a obtenção de coordenadas tridimensionais, a partir de um par estéreo de imagens aquiritadas e pré-processadas, incluindo os problemas de orientação envolvidos, e o procedimento de triangulação estéreo. O capítulo 5 ilustra uma aplicação prática do sistema, no dimensionamento de estruturas submarinas.

Capítulo 2

Sistema de Visão Estéreo

A aquisição das imagens necessárias ao processo de dimensionamento é feita por um sistema de visão estéreo adequado para atividades submarinas, que possibilita a visualização em tempo real das imagens aquiritadas. O sistema, baseado em um computador pessoal, utiliza câmeras eletrônicas e placas digitalizadoras para permitir a visualização e gravação em arquivo de imagens. Este capítulo descreve brevemente o sistema de visão utilizado, e seus componentes principais.

2.1 Introdução

Diversas tecnologias para apresentação de imagens tridimensionais têm sido desenvolvidas, e dividem-se basicamente em dois tipos [4]

- as que utilizam dois canais independentes, direcionando para cada olho do observador as imagens correspondentes às câmeras esquerda e direita. Este processo tem o inconveniente de exigir equipamentos mais complexos, tendo entretanto qualidade e resolução melhores. Os *displays* podem utilizar dois monitores polarizados e óculos passivos.
- sistemas seqüenciais em que são utilizados apenas um *display* e um dispositivo de seleção (por exemplo, uma tela de cristal líquido operando como obturador) que é chaveado alternadamente, permitindo o direcionamento adequado das imagens esquerda e direita. A fim de se eliminar a percepção

da cintilação, as imagens para cada olho são apresentadas em cerca de 1/60 s, ou seja, a frequência de chaveamento deverá ser de 120 Hz, aproximadamente. Para isso, é necessário duplicar a frequência dos sinais de vídeo, que operam a 60 Hz. Isto é feito por meio de memórias digitais de vídeo, o que implica na digitalização das imagens por computador. A resolução e a qualidade destes sistemas dependem, basicamente, das características das câmeras e das placas digitalizadoras utilizadas.

O sistema de visão utilizado no presente estudo adotou a segunda tecnologia, principalmente pela maior flexibilidade e facilidade de manuseio, e pelo menor custo. Com este sistema, é possível controlar via *software* o alinhamento e a convergência das imagens, propiciando uma maior comodidade operacional, além de ampliar o controle de convergência das imagens, propiciando um aumento das faixas de distâncias utilizáveis das câmeras ao objeto. Além disso, tendo sido as imagens digitalizadas, a execução de medições tridimensionais por *software* próprio é mais facilitada.

2.2 Arquitetura

A arquitetura do sistema de visão estéreo [4] é ilustrada na Fig. 2.1. Os principais componentes desta arquitetura serão descritos a seguir.

2.2.1 Câmeras

A aquisição das imagens é feita por um par de câmeras eletrônicas digitais. Existem vários tipos de câmeras digitais, mas as mais usadas (e as adotadas neste sistema) são as que utilizam a tecnologia CCD (*Charge-Coupled Device*). Trata-se de um conjunto de receptores de luz enfileirados, capaz de detectar variações na intensidade e na frequência (cor) da luz. Estes receptores são densamente dispostos lado a lado: quanto maior o número deles, maior a resolução.

O uso de câmeras digitais representa uma vantagem em relação ao uso de câmeras fotográficas, utilizadas em alguns sistemas que utilizam técnicas de inspeções visuais, por permitir a visualização em tempo real das imagens, e uma

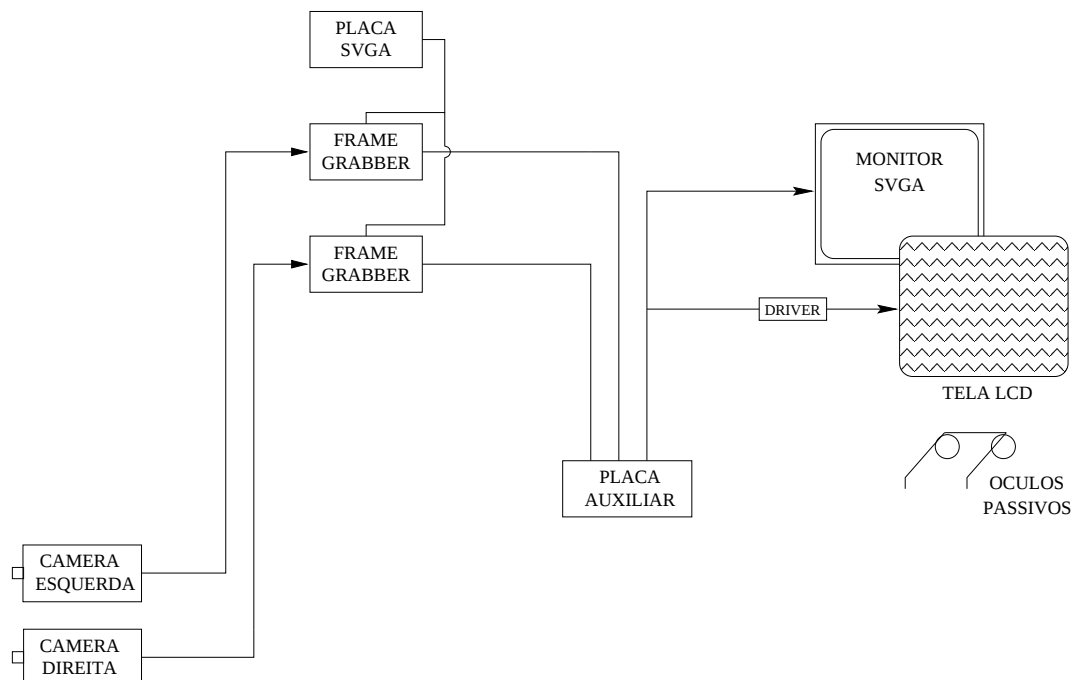


Figura 2.1: Arquitetura do Sistema de Visão Estéreo.

melhor qualidade, devido ao fato de a resolução de uma câmera CCD ser bem maior do que a resolução de uma fotografia comum.

As câmeras foram montadas sobre uma base fixa, alinhadas horizontalmente e espaçadas de uma distância que corresponde aproximadamente à separação entre os olhos de um ser humano adulto (tipicamente, cerca de 5,5 centímetros), para dar a sensação de visão estéreo real ao usuário do sistema. Cada câmera gera um sinal de vídeo em cores padrão NTSC, e corresponde a um canal (esquerdo ou direito) na arquitetura.

2.2.2 Placas Digitalizadoras

As placas digitalizadoras têm a finalidade de digitalizar o sinal de vídeo proveniente das câmeras CCD, transformando-o em um sinal que possa ser exibido no monitor SVGA do microcomputador. Constituem-se em duas placas tipo *Frame Grabber*, dispostas em paralelo, cada qual recebendo o sinal de uma câmera. Ambas as placas são ligadas a uma placa auxiliar, que realiza o chaveamento entre as imagens direita e esquerda. Sendo o chaveamento realizado externamente às

placas digitalizadoras, pode-se fazer alterações na configuração via *software*, o que não seria possível se o chaveamento fosse realizado internamente às placas digitalizadoras.

2.2.3 Dispositivo de Visualização

O dispositivo de visualização estereoscópico, constituído em uma tela de cristal líquido (LCD), e com o uso auxiliar de óculos passivos, permite a visualização de imagens tridimensionais a partir da composição de duas imagens bidimensionais, tendo-se com isso a noção de profundidade.

2.2.4 Monitor

Para a visualização das imagens, é utilizado um único monitor, tipo SVGA 17". Embora sejam gerados dois sinais de vídeo, ver-se-á mais adiante que apenas um monitor é necessário à operação do sistema.

2.3 Operação do Sistema

A função básica do sistema de visão é permitir a visualização das imagens em tempo real. As imagens, depois de captadas pelo par de câmeras digitais, são enviadas às duas placas digitalizadoras (cada câmera está ligada a uma placa). Cada placa converte a informação do sinal NTSC proveniente da câmera para um sinal VGA, que é enviado à placa auxiliar, que realiza o chaveamento entre as imagens dos canais esquerdo e direito, permitindo a visualização da duas imagens como se fossem uma só imagem tridimensional, com o auxílio da placa LCD e dos óculos passivos.

Como as placas digitalizadoras estão ligadas à placa de vídeo SVGA do computador, é possível a operação de *overlay* (o sinal proveniente das câmeras pode ser enviado diretamente à saída da placa de vídeo, permitindo a utilização de um único monitor tanto para a operação do computador quanto para a operação do *software* do sistema de visão).

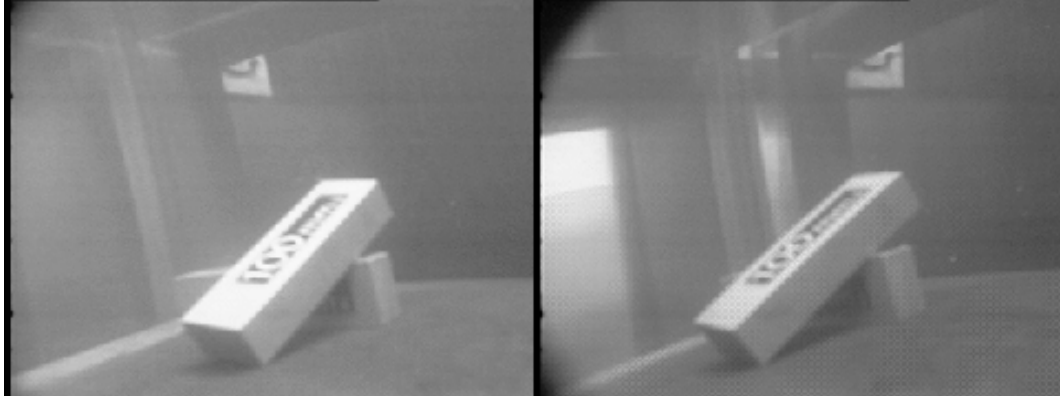


Figura 2.2: Exemplo de par estéreo adquirido.

As imagens são adquiridas pelo sistema de visão em pares estéreo e em seguida digitalizadas e gravadas em arquivo, em um formato gráfico com resolução de 640×512 pixels e profundidade de cores de 24 bits, para permitir uma melhor qualidade das imagens adquiridas. Além da visualização e aquisição de imagens, o *software* permite a realização de outras operações, como *zoom*, alinhamento horizontal, e convergência.

A Fig. 2.2 mostra um exemplo de par estéreo (em escala reduzida e em tons de cinza) adquirido pelo sistema descrito.

Capítulo 3

Pré-processamento

As imagens, depois de adquiridas pelo sistema de visão descrito no capítulo anterior, devem ser submetidas a um pré-processamento, que tem por finalidade obter uma melhoria de qualidade, e adequá-las ao processo de dimensionamento. O melhoramento da qualidade inclui técnicas como a equalização do histograma e a filtragem das imagens; já o uso de técnicas de detecção de envoltórias permite uma melhor caracterização dos limites de uma imagem, o que auxilia na determinação de coordenadas 2D nas imagens. Neste capítulo discutir-se-ão algumas destas técnicas, principalmente a detecção de envoltórias.

3.1 Operações Pontuais

Operações pontuais [6] são operações de memória zero onde um dado nível de cinza $u \in [0, L]$ é mapeado em um nível de cinza $v \in [0, L]$ de acordo com a transformação

$$v = f(u) \tag{3.1}$$

3.1.1 Imagem Negativa

Uma imagem negativa pode ser obtida escalando-se reversamente os níveis de cinza de acordo com a transformação a seguir, ilustrada na Fig. 3.1:

$$v = L - u \tag{3.2}$$

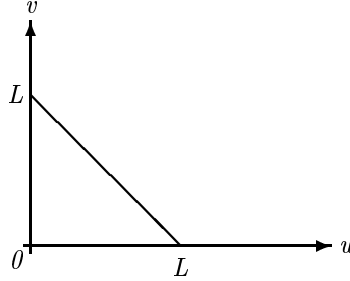


Figura 3.1: Transformação negativa.

3.1.2 Equalização de Histograma

O histograma de uma imagem representa a freqüência relativa da ocorrência de níveis de cinza em uma imagem. As técnicas de modelamento de histograma modificam uma imagem com a finalidade de dar ao seu histograma uma forma desejada. Isto é usual em estender os níveis de baixo contraste de imagens com histogramas com formato estreito.

O objetivo da equalização de histograma é obter um histograma uniforme para a imagem resultante. Considere um valor de pixel $u \geq 0$ como uma variável aleatória com uma função densidade de probabilidade contínua $p_u(v)$ e distribuição de probabilidade cumulativa $F_u(v) \equiv P[u \leq v]$. Então, a variável aleatória

$$v \equiv F_u(u) \equiv \int_0^u p_u(v)dv \quad (3.3)$$

será uniformemente distribuída sobre $(0, 1)$. Para implementar esta transformação em imagens digitais, supõe-se que a entrada u tenha L níveis de cinza x_i , $i = 0, 1, \dots, L - 1$ com probabilidades $p_u(x_i)$. Estas probabilidades podem ser determinadas a partir do histograma da imagem que fornece $h(x_i)$, o número de pixels com um valor de nível de cinza x_i . Então,

$$p_u(x_i) = \frac{h(x_i)}{\sum_{i=0}^{L-1} h(x_i)}, \quad i = 0, 1, \dots, L - 1 \quad (3.4)$$

$$v \equiv \sum_{x_i}^u p_u(x_i) \quad (3.5)$$

A saída v' , que se assume que tenha também L níveis, é dada por

$$v' \equiv \text{int} \left[\frac{v - v_{\min}}{1 - v_{\min}} (L - 1) + 0.5 \right] \quad (3.6)$$

onde v_{min} é o menor valor positivo de v obtido de 3.5. Agora v' será uniformemente apenas aproximadamente distribuída devido ao fato de v não ser uma variável uniformemente distribuída. A Fig. 3.2 mostra o algoritmo de equalização de histograma para imagens digitais. De (3.5) nota-se que v é uma variável discreta que assume o valor

$$v_k = \sum_{i=0}^k p_u(x_i) \quad (3.7)$$

se $u = x_i$. A Eq. 3.6 requantiza uniformemente o conjunto v_k para v'_k . Nota-se que este passo de requantização é necessário devido ao fato de as probabilidades $p_u(x_k)$ e $p_v(v_k)$ serem as mesmas.

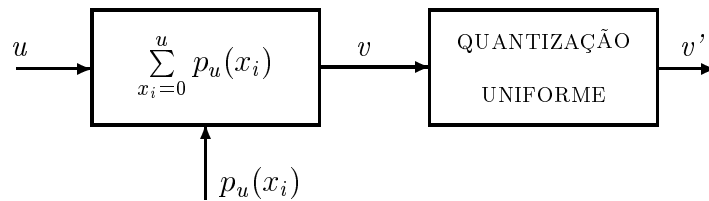


Figura 3.2: Transformação da equalização de histograma.

A Fig. 3.3 ilustra um exemplo de equalização de histograma de uma imagem monocromática. Observa-se que o histograma após a equalização torna-se aproximadamente uniforme.

3.2 Operações Espaciais

As operações espaciais em uma imagem são realizadas em vizinhanças locais dos pixels, através da convolução da imagem com um filtro de resposta ao impulso finita chamado máscara espacial [6].

3.2.1 Filtragem por Média Espacial

Aqui cada pixel é substituído por uma média ponderada dos pixels vizinhos:

$$v(m, n) = \sum_{k \in W} \sum_{l \in W} a(k, l) y(m - k, n - l) \quad (3.8)$$

onde $y(m, n)$ e $v(m, n)$ são as imagens de entrada e saída, respectivamente. W é uma janela escolhida adequadamente, e $a(k, l)$ são os pesos do filtro. No caso de

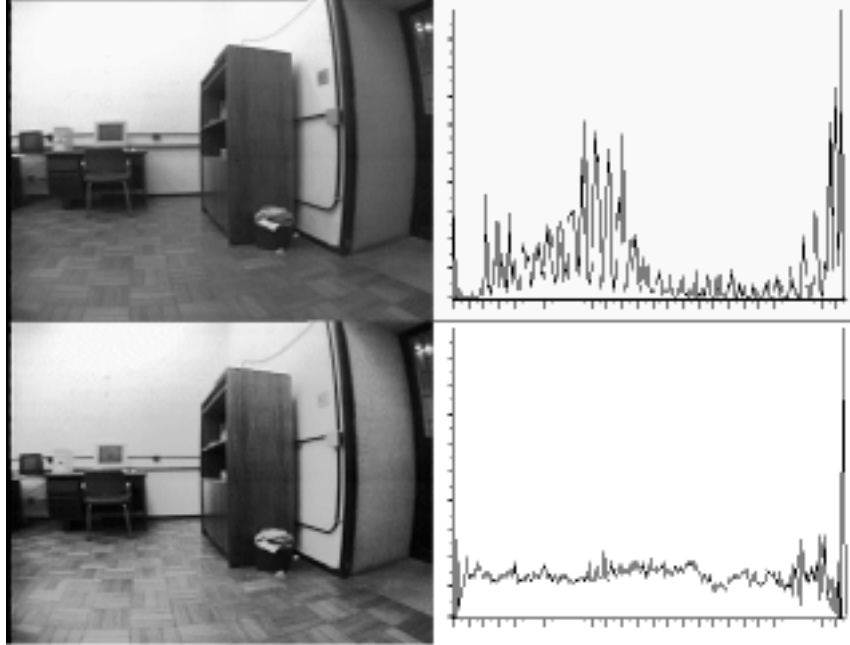


Figura 3.3: Equalização de histograma: acima, imagem original e seu histograma; abaixo, imagem e respectivo histograma após a equalização.

todos os pesos terem valores iguais,

$$v(m, n) = \frac{1}{N_W} \sum_{k \in W} \sum_{l \in W} y(m - k, n - l) \quad (3.9)$$

onde $a(k, l) = 1/N_W$ e N_W é o número de pixels na janela W . Um outro filtro usado é dado por

$$v(m, n) = \frac{1}{2}[y(m, n) + \frac{1}{4}[y(m-1, n) + y(m+1, n) + y(m, n-1) + y(m, n+1)]] \quad (3.10)$$

onde cada pixel é substituído pela sua média com os quatro pixels mais próximos. A Fig. 3.4 mostra algumas máscaras deste tipo. A filtragem por média espacial é usada para suavização de ruído e filtragem passa-baixa.

3.2.2 Filtragem Passa-Baixa, Passa-Alta e Passa-Faixa

Anteriormente, mencionou-se que uma operação de média espacial é uma filtragem passa-baixa. Se $h_{PB}(m, n)$ é um filtro FIR passa-baixa, então um filtro FIR passa-alta $h_{PA}(m, n)$ pode ser definido como

$$h_{PA}(m, n) = \delta(m, n) - h_{PB}(m, n) \quad (3.11)$$

$$\begin{array}{ccc}
 \begin{bmatrix} \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} \end{bmatrix} & \begin{bmatrix} \frac{1}{9} & \frac{1}{9} & \frac{1}{9} \\ \frac{1}{9} & \frac{1}{9} & \frac{1}{9} \\ \frac{1}{9} & \frac{1}{9} & \frac{1}{9} \end{bmatrix} & \begin{bmatrix} 0 & \frac{1}{8} & 0 \\ \frac{1}{8} & \frac{1}{4} & \frac{1}{8} \\ 0 & \frac{1}{8} & 0 \end{bmatrix} \\
 2 \times 2 & 3 \times 3 & 3 \times 3 \text{ ponderada}
 \end{array}$$

Figura 3.4: Máscaras espaciais $a(k, l)$.

Tal filtro pode ser implementado simplesmente subtraindo-se a saída do filtro passa-baixa de sua entrada.

Um filtro passa-faixa espacial pode ser caracterizado como

$$h_{PF}(m, n) = h_{B_1}(m, n) - h_{B_2}(m, n) \quad (3.12)$$

onde $h_{B_1}(m, n)$ e $h_{B_2}(m, n)$ são as respostas ao impulso finitas de filtros passa-baixa.

Os filtros passa-baixa são usuais para suavização de ruído. Os filtros passa-alta são usuais na extração de envoltórias. Os filtros passa-banda são usuais no melhoramento de contornos e outras características passa-alta na presença de ruído. A Fig. 3.5 ilustra os diagramas em blocos destes filtros.

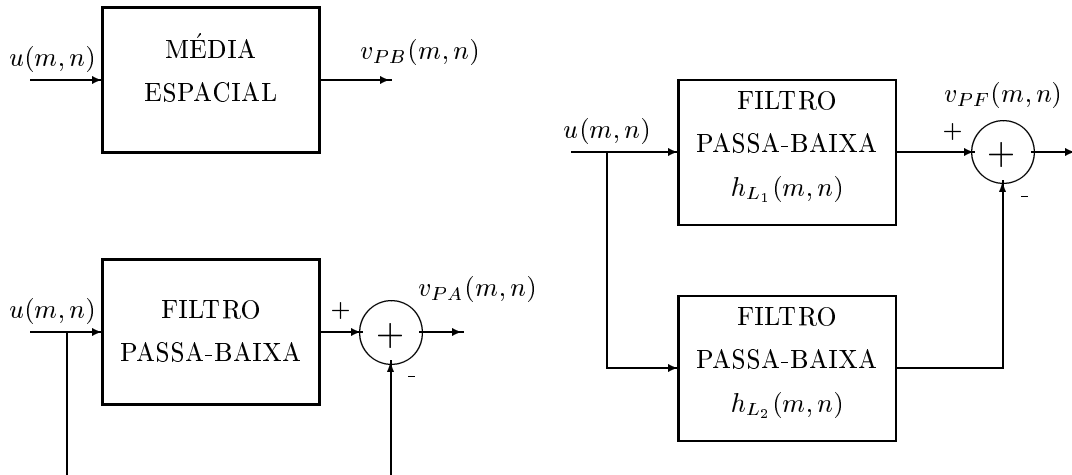


Figura 3.5: Filtros espaciais.

3.3 Detecção de Envoltória

Um problema de fundamental importância no processamento de uma imagem é a detecção de envoltórias. Envoltórias caracterizam fronteiras de objetos e portanto são usuais para segmentação, reconhecimento e identificação de objetos em imagens digitais, já que o processo de detecção de envoltórias permite a caracterização das fronteiras destes objetos.

Contudo, não há ainda uma definição matemática precisa de envoltória. Para um ser humano, é usualmente muito fácil encontrar os contornos de objetos em uma cena. Mas a correspondente operação é extremamente difícil para um computador. Há várias razões para isto [5]:

1. A primeira razão tem a ver com “linguagem”. Quando se fala sobre contornos de objetos, assume-se que a noção de objeto é bem entendida, mas de fato é um dos objetivos da visão computacional identificar objetos em cenas. Portanto, a ideia de contorno não pode ser amparada por uma definição da ideia de objeto uma vez que criar-se-ia um ciclo vicioso.
2. Mesmo sendo usada uma definição mais física de uma envoltória como uma descontinuidade de alguma espécie de função intensidade, ainda tem-se o problema de medição de ruído, que é a segunda razão pela qual a detecção de envoltórias em uma imagem é uma tarefa difícil. Realmente, a detecção de descontinuidades da imagem intensidade pode ser obtida matematicamente computando-se derivadas desta função. A detecção de envoltórias é difícil, contudo, devido ao fato de a intensidade ser uma medida física que é sujeita a ruído; além do mais, a operação de derivação é propensa a atenuar este ruído.
3. Uma terceira dificuldade relacionada a envoltórias é que existem diferentes fontes de envoltórias em uma cena, como pode ser visualizado na Fig. 3.6, que mostra uma cena composta de um objeto cilíndrico com uma sombra projetando-se sobre uma base poliedral. Algumas envoltórias podem surgir de sombras formadas por objetos (como em *s*), outras de variações na

reflectância de objetos (como em r), e outras de variações na profundidade (como em p). Todos estes processos distintos podem produzir envoltórias. Mas, se é preciso associar envoltórias de uma imagem a objetos de uma cena, é necessária a capacidade de identificar estes vários tipos de envoltórias.

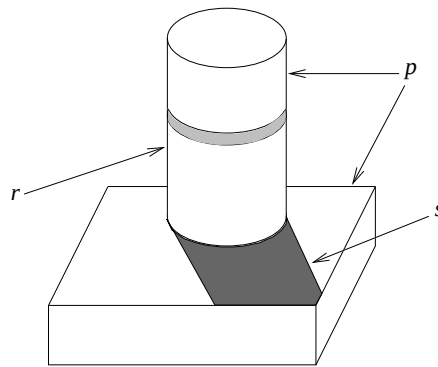


Figura 3.6: Diferentes fontes físicas de envoltórias em uma imagem.

Os primeiros detectores de envoltória tentavam contornar os problemas de computar algumas derivadas da imagem e de serem robustos ao ruído [5]. Estes dois requerimentos, contudo, revelam-se contraditórios. A idéia de usar derivadas da função intensidade da imagem para melhorar a detecção de envoltórias pode ser vista de várias maneiras:

- no domínio do tempo, uma envoltória em alguma direção corresponde a uma larga variação local da função intensidade em uma direção perpendicular à da envoltória.
- No domínio da frequência, um operador derivativo pode ser visto como um *booster* de alta frequência ($\sin \omega x$ torna-se $\omega \cos \omega x$), e portanto tal operador enfatizará mais as envoltórias e não as áreas planas.

Nota-se que estas simples idéias não levam em consideração o ruído. As seções seguintes apresentam detectores mais elaborados.

O paradigma geral é um método que consiste em dois passos. O primeiro consiste em melhorar a presença de envoltórias na imagem original $f(x, y)$, assim criando uma nova imagem $g(x, y)$, onde as envoltórias são mais destacadas. Valores maiores de $g(x, y)$ indicam a probabilidade da presença de uma envoltória.

O segundo passo consiste em aplicar um limiar a $g(x, y)$, de forma a realizar uma decisão “envoltória ou não-envoltória”, produzindo um mapa de envoltórias binário.

3.3.1 Aproximações Discretas de Derivadas

Alguns detectores de envoltórias usam aproximações discretas de derivadas contínuas [5]. Derivadas parciais $\frac{\partial f}{\partial x}$ e $\frac{\partial f}{\partial y}$ da função $f(x, y)$ podem ser aproximadas por meio de diferenças finitas:

$$\begin{aligned}\frac{\partial f(x, y)}{\partial x} &= \Delta_x f(x, y) = f(x + 1, y) - f(x, y) \\ \frac{\partial f(x, y)}{\partial y} &= \Delta_y f(x, y) = f(x, y + 1) - f(x, y)\end{aligned}\tag{3.13}$$

Sendo $f(x, y)$ uma superfície em $\mathbb{R}^3$ dada pela equação $z = f(x, y)$, o vetor tridimensional $\mathbf{G}_{\mathbf{xy}} = [\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, -1]^t$ é normal à superfície. Uma vez que a coordenada z de $\mathbf{G}_{\mathbf{xy}}$ é constante, sua projeção $\mathbf{g}_{\mathbf{xy}} = [\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}]^t$ no plano (x, y) carrega exatamente a mesma informação e aponta na direção da variação de intensidade máxima, enquanto sua norma é uma indicação da taxa de variação nesta direção e é chamada de imagem gradiente. Isto implica em particular que $\mathbf{g}_{xy}$ é ortogonal à direção das envoltórias. Na prática, $\mathbf{g}_{xy}$ é aproximada por $[\Delta_x, \Delta_y]^t$. Isto também permite computar a taxa de variações de intensidade em qualquer direção, não somente a direção de máxima variação, simplesmente projetando-a nesta direção: se $D_\theta f$ denota a derivada parcial na direção θ , tem-se $D_\theta f = \nabla f \cdot \mathbf{u}_\theta = \mathbf{g} \cdot \mathbf{u}_\theta$, onde $\mathbf{u}_\theta$ é o vetor unitário na direção θ .

Derivadas de ordem elevada são também usuais. Por exemplo, sabe-se que sempre que uma variação de intensidade ocorrer em alguma direção, há um pico correspondente na primeira derivada direcional ou, equivalentemente, uma passagem por zero na segunda derivada direcional. A tarefa de detectar variações de intensidade é então reduzida à tarefa de encontrar as passagens por zero da segunda derivada direcional na direção apropriada.

Derivadas da forma $\frac{\partial^{m+n} f}{\partial x^m \partial y^n}$ podem ser aproximadas concatenando-se os operadores diferenciais Δ_x e Δ_y :

$$\frac{\partial^2 f}{\partial x^2} \equiv \Delta_x^2 f = \Delta_x(\Delta_x f) = f(x+2, y) - 2f(x+1, y) + f(x, y) \quad (3.14)$$

Aproximações de derivadas com diferenças finitas são operadores lineares e invariantes no tempo, e podem ser consideradas como máscaras de convolução discreta. A Tab. 3.1 mostra alguns exemplos.

3.3.2 Tipos de Operadores

Em uma imagem, os pontos de envoltórias podem ser definidos como localidades onde ocorrem mudanças abruptas nas tonalidades de cinza da imagem. Por exemplo, podem ser definidos pontos de envoltória em uma imagem digital como pixels de tom preto com no mínimo um pixel de tom branco na vizinhança. Pode-se considerar também a envoltória como sendo as localidades entre duas regiões homogêneas da imagem onde há diferentes intensidades de iluminação. Esta definição implica que uma envoltória corresponde a uma variação local de iluminação, mas não necessariamente o contrário [7].

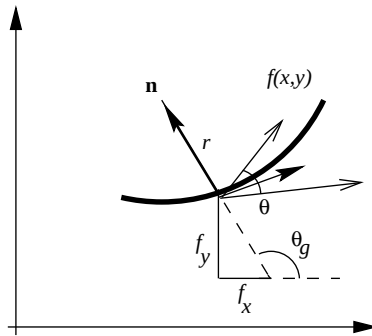


Figura 3.7: Gradiente de $f(x, y)$ na direção r .

Para uma imagem contínua $f(x, y)$ sua derivada assume um máximo local na direção da envoltória. Portanto, uma técnica de detecção de envoltória é medir o gradiente de f dentro de um raio r em uma direção θ , isto é,

$$\frac{\partial f}{\partial r} = \frac{\partial f}{\partial x} \frac{\partial x}{\partial r} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial r} = f_x \cos \theta + f_y \sin \theta$$

Tabela 3.1: Operadores diferenciais discretos.

Operador	resposta ao impulso
Δ_x	$\begin{bmatrix} 1 & -1 \end{bmatrix}$
Δ_y	$\begin{bmatrix} 1 \\ -1 \end{bmatrix}$
Δ_x^2	$\begin{bmatrix} 1 & -2 & 1 \end{bmatrix}$
Δ_y^2	$\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$
$\Delta_x^2 + \Delta_y^2$	$\begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}$

O máximo valor de $\frac{\partial f}{\partial r}$ é obtido quando $\frac{\partial}{\partial \theta} \frac{\partial f}{\partial r} = 0$. Daí,

$$-f_x \sin \theta_g + f_y \cos \theta_g \Rightarrow \theta_g = \arctan \left(\frac{f_y}{f_x} \right)$$

$$\left(\frac{\partial f}{\partial r} \right)_{\max} = \sqrt{f_x^2 + f_y^2}$$

onde θ_g é a direção da envoltória. Baseado nestes conceitos, dois tipos de operadores de detecção de envoltória são definidos: os operadores de gradiente e os operadores de direção. Para imagens digitais, estes operadores, também chamados de *máscaras*, representam aproximações de diferenças finitas dos gradientes ortogonais f_x, f_y ou do gradiente de direção $\frac{\partial f}{\partial r}$. Sendo $\mathbf{H}$ uma máscara $p \times p$ define-se, para uma imagem arbitrária $\mathbf{U}$, o produto interno em (m, n) como a correlação

$$\langle \mathbf{U}, \mathbf{H} \rangle_{m,n} \equiv \sum_i \sum_j h(i, j) u(i + m, j + n) = u(m, n) \odot h(-m, -n) \quad (3.15)$$

3.3.3 Operador de Gradiente

Os operadores de gradiente são representados por um par de máscaras $\mathbf{H}_1, \mathbf{H}_2$, que medem o gradiente da imagem $u(m, n)$ nas direções ortogonais (Fig. 3.8) [6]. Definindo-se os gradientes bidirecionais pelas expressões $g_1(m, n) \equiv \langle \mathbf{U}, \mathbf{H}_1 \rangle_{m,n}$ e $g_2(m, n) \equiv \langle \mathbf{U}, \mathbf{H}_2 \rangle_{m,n}$ o módulo e a fase do vetor gradiente são dados por

$$g(m, n) = \sqrt{g_1^2(m, n) + g_2^2(m, n)} \quad (3.16)$$

$$\theta_g(m, n) = \arctan \frac{g_2(m, n)}{g_1(m, n)} \quad (3.17)$$

O módulo pode ser também calculado através de

$$g(m, n) \equiv |g_1(m, n)| + |g_2(m, n)| \quad (3.18)$$

que é uma fórmula de mais fácil implementação, sendo a utilizada na prática.

Um pixel de coordenada (m, n) é declarado como um ponto de envoltória se $g(m, n)$ é maior ou igual a um limiar t [7]. Os pontos de envoltória constituem um mapa de envoltória $e(m, n)$, definido como

$$e(m, n) = \begin{cases} 1, & (m, n) \in I_g \\ 0, & (m, n) \notin I_g \end{cases} \quad \text{onde } I_g \equiv \{(m, n); g(m, n) > t\} \quad (3.19)$$

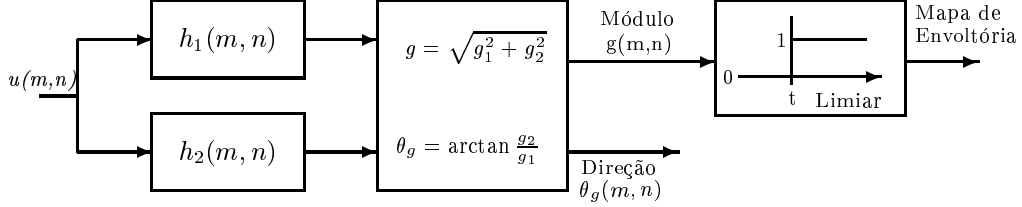


Figura 3.8: Detecção de envoltórias por operador de gradiente.

O valor 1 corresponde à tonalidade branca, e o valor 0, à tonalidade preta. Assim, o processo de geração do mapa de envoltória consiste em processar a imagem através de uma curva de transformação definida pelo limiar t , do tipo:

$$f(x) = \begin{cases} 1, & \text{se } x \geq t \\ 0, & \text{se } x < t \end{cases} \quad (3.20)$$

A Tab. 3.2 mostra alguns operadores de gradiente discretos mais comuns. Os operadores computam diferenças horizontais e verticais de somas locais, reduzindo assim o efeito de ruído nos dados. Observa-se que estes operadores têm a desejável propriedade de produzir zeros para regiões não uniformes. A Fig. 3.9 ilustra a aplicação destes operadores em uma imagem monocromática.

3.3.4 Operador de Direção

Os operadores de direção possuem a propriedade de medir o gradiente de uma imagem digital em uma determinada direção (Fig. 3.10) [6]. Se $g_k(m, n)$ é o gradiente de direção na direção $\theta_k = \frac{\pi}{2} + k\frac{\pi}{4}$, o gradiente em (m, n) é definido como

$$g(m, n) \equiv \max |g_k(m, n)| \quad (3.21)$$

Como no caso dos operadores de gradiente, pode-se aplicar um limiar para obter-se o mapa de envoltória correspondente.

A Tab. 3.3 mostra alguns tipos de máscaras utilizadas, para detecção de envoltória na direção Norte. Um deslocamento no sentido anti-horário dos oito elementos limítrofes destas máscaras dão uma rotação de 45° da direção do gradiente. Por exemplo, para o operador Kirsch da Tab. 3.3, as oito diferentes máscaras em oito diferentes direções são mostradas na Tab. 3.4. A Fig. 3.11 ilustra a aplicação destas máscaras a uma imagem monocromática.

Tabela 3.2: Operadores de gradiente.

Operador	Máscara $\mathbf{H}_1$	Máscara $\mathbf{H}_2$
Prewitt	$\begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}$
Isotropic	$\begin{bmatrix} -1 & 0 & 1 \\ -\sqrt{2} & 0 & \sqrt{2} \\ -1 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} -1 & -\sqrt{2} & -1 \\ 0 & 0 & 0 \\ 1 & \sqrt{2} & 1 \end{bmatrix}$
Sobel	$\begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}$

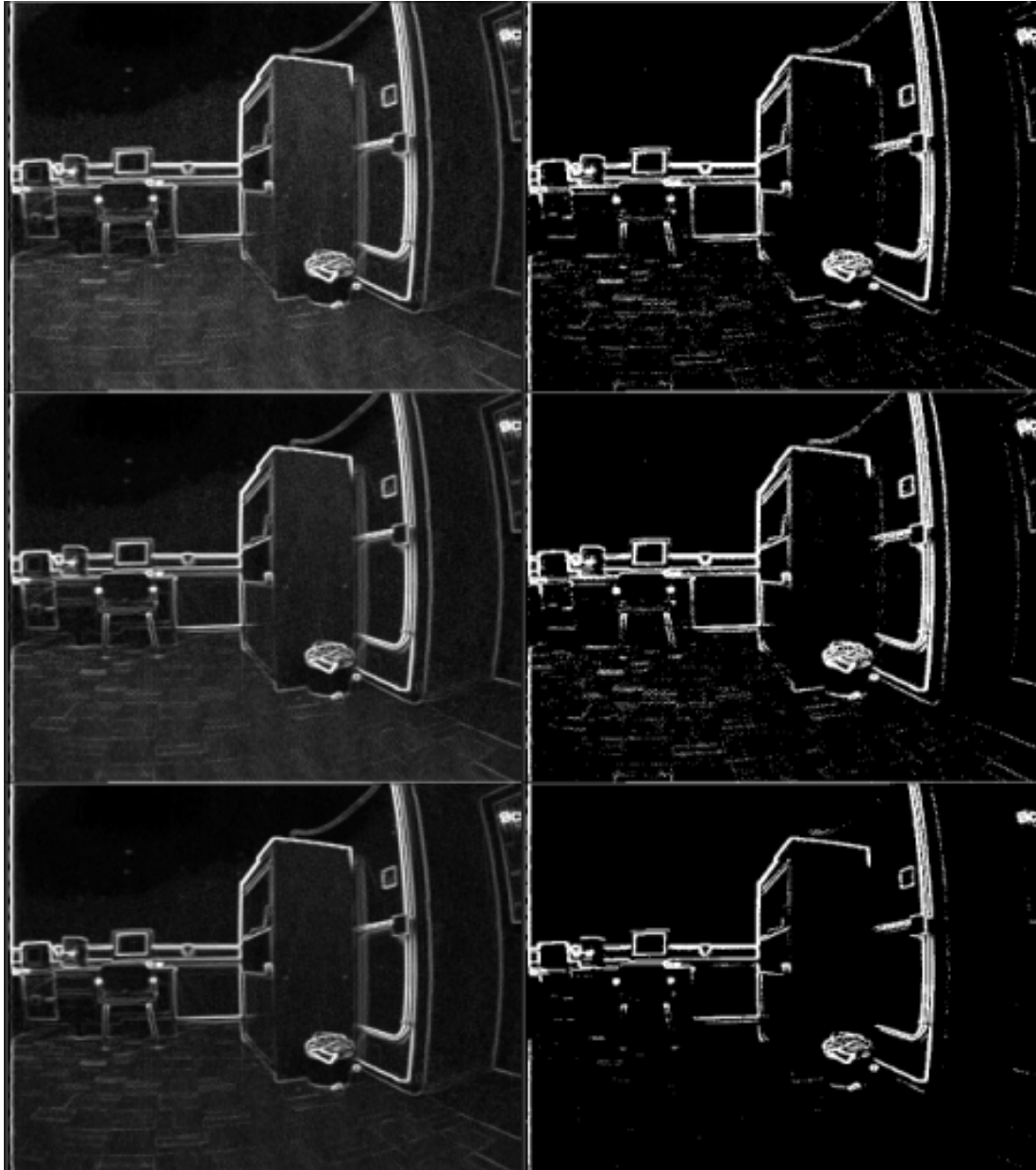


Figura 3.9: Exemplos de detecção de envoltórias por operador de gradiente. As imagens gradiente são mostradas à esquerda, e os respectivos mapas de envoltória à direita. De cima para baixo, cada par corresponde à aplicação, respectivamente, das máscaras Sobel, Prewitt e Isotropic.

Tabela 3.3: Operadores de direção.

Operador	Máscara H (direção Norte)
Tipo I	$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix}$
Tipo II	$\begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$
Tipo III	$\begin{bmatrix} 1 & 1 & 1 \\ 1 & -2 & 1 \\ -1 & -1 & -1 \end{bmatrix}$
Kirsch	$\begin{bmatrix} 5 & 5 & 5 \\ -3 & 0 & -3 \\ -3 & -3 & -3 \end{bmatrix}$

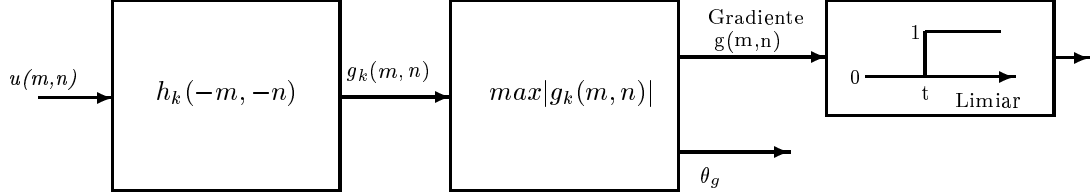


Figura 3.10: Detecção de envoltórias por operador de direção.

3.3.5 Operadores de Laplace

Os métodos anteriores de estimar os gradientes possuem uma melhor performance quando a transição entre os tons de cinza é bastante abrupta, como uma função degrau. Para imagens com transições mais suaves, é mais vantajoso aplicar derivadas de segunda ordem. Um operador freqüentemente utilizado é o operador de Laplace [6], definido por

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} \quad (3.22)$$

A Tab. 3.5 fornece três diferentes aproximações discretas para este operador, que podem ser usadas para a detecção de envoltória em imagens digitais. No caso do operador de tipo 1 mostrado na tabela, a Eq. 3.22 pode ser redefinida como

$$\nabla^2 f \approx 4f(x, y) - [f(x, y + 1) + f(x, y - 1) + f(x + 1, y) + f(x - 1, y)] \quad (3.23)$$

Devido às derivadas de segunda ordem, este operador tem a desvantagem de ser mais sensível ao ruído do que os outros anteriormente definidos. Além disso, o módulo de $\nabla^2 f$ pode produzir envoltórias duplas quando da geração do mapa de envoltória. Por isto, e também por não permitir a detecção de envoltórias em uma determinada direção, o operador de Laplace não é um bom operador para detecção de envoltória. Uma melhor utilização deste operador é usar suas passagens de zero para detectar os pontos de envoltória. Um operador de Laplace generalizado, que aproxima as funções Laplaciana e Gaussiana, é um poderoso detector de passagem de zero, sendo definido por

$$h(m, n) \equiv c \left[1 - \frac{(m^2 + n^2)}{\sigma^2} \right] \exp \left(- \frac{m^2 + n^2}{2\sigma^2} \right) \quad (3.24)$$

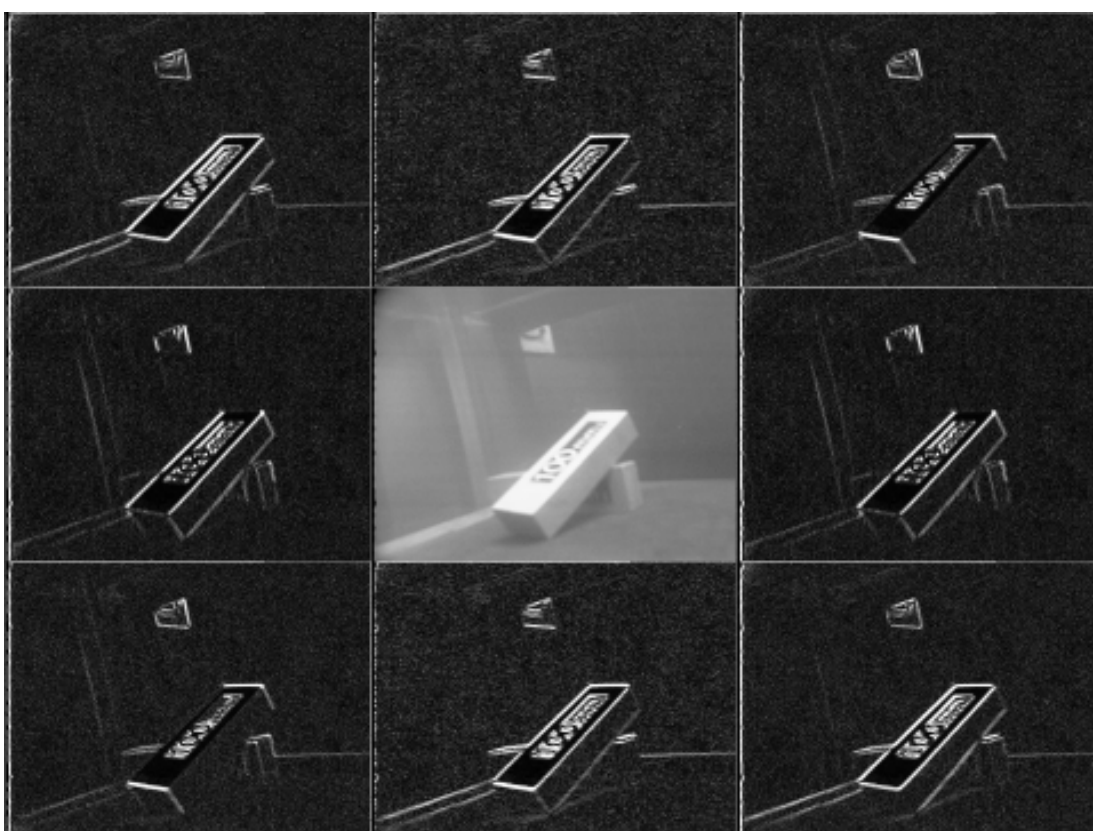


Figura 3.11: Exemplo de detecção de envoltórias por operador de direção usando o operador Kirsch. Ao centro, é mostrada a imagem original; no sentido horário, a contar do canto superior direito, são mostradas as imagens após a aplicação das máscaras da Tab. 3.4, respectivamente nas direções Nordeste, Leste, Sudeste, Sul, Sudoeste, Oeste, Noroeste e Norte.

As passagens de zero de uma dada imagem convoluída com $h(m, n)$ fornecem os pontos de envoltórias. Em uma matriz bidimensional, considera-se que ocorre uma passagem de zero quando há uma passagem de zero em pelo menos uma direção.

Na Eq. 3.24, a expressão $h(m, n)$ é a resposta ao impulso amostrada de um filtro passa-banda analógico cuja resposta em frequência é proporcional a $(\xi_1^2 + \xi_2^2) \exp[-2\sigma^2(\xi_1^2 + \xi_2^2)]$. Portanto, o detector de passagem de zero é equivalente a um filtro passa-baixa tendo uma resposta ao impulso Gaussiano seguido por um operador de Laplace. O filtro passa-baixa serve para atenuar a sensibilidade ao ruído do operador de Laplace.

3.3.6 Gradientes Estocásticos

Um problema no processo de detecção de envoltória surge quando se necessita trabalhar com imagens ruidosas, já que muitas vezes os algoritmos utilizados acabam acentuando o ruído na imagem, como ocorre com as máscaras de gradiente. Um modo de se resolver este problema é passar a imagem, depois de processada pelos detectores de envoltória, por um filtro passa-baixas, por um filtro de mediana, ou qualquer outro tipo de filtro que reduza a degradação da imagem. Uma alternativa mais eficiente é projetar máscaras para extração de envoltórias que levem em conta a presença de ruído na imagem de uma maneira controlada [6].

A Tab. 3.6 mostra operadores de gradiente estocástico H_1 , com

$$r_o(k, l) = 0.99^{\sqrt{k^2 + l^2}} \quad (3.25)$$

$$SNR = \frac{\sigma^2}{\sigma_\nu^2} \quad (3.26)$$

O operador $\mathbf{H}_2$ é obtido com a rotação horária de 90° , mas que devido às simetrias de $\mathbf{H}_1$, corresponde a $\mathbf{H}_1^t$.

Tabela 3.4: Operadores de direção tipo Kirsch.

5 5 5 -3 0 -3 -3 -3 -3 Norte	5 5 -3 5 0 -3 -3 -3 -3 Noroeste	5 -3 -3 5 0 -3 5 -3 -3 Oeste	-3 -3 -3 5 0 -3 5 5 -3 Sudoeste
-3 -3 -3 -3 0 -3 5 5 5 Sul	-3 -3 -3 -3 0 5 -3 5 5 Sudeste	-3 -3 5 -3 0 5 -3 -3 5 Leste	-3 5 5 -3 0 5 -3 -3 -3 Nordeste

Tabela 3.5: Operadores de Laplace.

Operador	Máscara H
Tipo 1	$\begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix}$
Tipo 2	$\begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$
Tipo 3	$\begin{bmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{bmatrix}$

Tabela 3.6: Gradientes estocásticos.

Máscara	SNR=1
3×3	$\begin{bmatrix} 0.97 & 0 & -0.97 \\ 1 & 0 & -1 \\ 0.97 & 0 & -0.97 \end{bmatrix}$
5×5	$\begin{bmatrix} 0.802 & 0.836 & 0 & -0.836 & -0.802 \\ 0.845 & 0.897 & 0 & -0.897 & -0.845 \\ 0.870 & 1 & 0 & -1 & -0.870 \\ 0.845 & 0.897 & 0 & -0.897 & -0.845 \\ 0.802 & 0.836 & 0 & -0.836 & -0.802 \end{bmatrix}$
7×7	$\begin{bmatrix} 0.641 & 0.672 & 0.719 & 0 & -0.719 & -0.672 & -0.641 \\ 0.656 & 0.719 & 0.781 & 0 & -0.781 & -0.719 & -0.656 \\ 0.688 & 0.781 & 0.875 & 0 & -0.875 & -0.781 & -0.688 \\ 0.703 & 0.813 & 1 & 0 & -0.813 & -1 & -0.703 \\ 0.688 & 0.781 & 0.875 & 0 & -0.875 & -0.781 & -0.688 \\ 0.656 & 0.719 & 0.781 & 0 & -0.781 & -0.719 & -0.656 \\ 0.641 & 0.672 & 0.719 & 0 & -0.719 & -0.672 & -0.641 \end{bmatrix}$

Capítulo 4

Dimensionamento Tridimensional

Depois de as imagens serem aquisitadas e posteriormente processadas, elas são submetidas a análise, para o dimensionamento de possíveis falhas e/ou defeitos na estrutura avaliada. O dimensionamento é feito através de um algoritmo de obtenção de coordenadas tridimensionais, baseado em métodos estereofotogramétricos. Este capítulo discute as técnicas de fotogrametria analítica usadas no algoritmo, dentre as quais se situam as soluções iterativas para os problemas de orientação exterior, interior e relativa, bem como os problemas de estimação de pose 2D–2D e 3D–3D e o procedimento de triangulação estéreo.

4.1 Introdução

A fotogrametria analítica [8] inclui o conjunto de técnicas pelas quais, a partir de coordenadas das projeções em perspectiva 2D, são feitas inferências sobre a posição 3D, orientação e dimensões de partes do objeto 3D observado em um referencial global. Estes problemas de inferência podem ser modelados como problemas *least-squares* não-lineares.

A orientação exterior de uma câmera é especificada por todos os parâmetros que determinam a pose da câmera no referencial global. Os parâmetros consistem da posição do centro de perspectividade e a direção do eixo óptico. A especificação da orientação exterior é feita a partir das coordenadas 3D de alguns pontos de controle cujas posições correspondentes na imagem são conhecidas, e resulta na

obtenção de três ângulos de rotação e três parâmetros de translação. A orientação interior de uma câmera é especificada por todos os parâmetros que determinam a geometria de um feixe de raios 3D das coordenadas da imagem medidas. Os parâmetros de orientação interior relacionam a geometria da projeção em perspectiva ideal à geometria física de uma câmera, e incluem a constante da câmera, o ponto principal, e a distorção da lente. A completa especificação da orientação de uma câmera é dada pela orientação interior e pela orientação exterior.

A orientação relativa de uma câmera em relação a outra constitui um modelo estéreo e é especificada por cinco parâmetros: três ângulos de rotação e dois parâmetros de translação. Quando duas câmeras estão em orientação relativa, cada par de raios correspondentes das duas câmeras intersectam-se no espaço 3D. O processo de determinação da orientação relativa assume que a orientação interior de cada câmera é conhecida. A orientação absoluta envolve a orientação de um modelo estéreo em um referencial global, e requer a obtenção de sete parâmetros: a escala, os três parâmetros de translação, e os três parâmetros de rotação. A orientação absoluta é feita a partir de coordenadas 3D de alguns pontos centrais cuja posição na imagem estéreo pode ser determinada. A completa especificação da orientação de um par de câmeras é dada pela especificação dos parâmetros determinados pela orientação relativa e pela orientação absoluta.

A relação entre os referenciais da câmera e o referencial global é dada por uma translação e uma rotação. Um ponto $\mathbf{x} = (x, y, z)^t$ no referencial global precisa primeiro ser expresso em relação à posição $\mathbf{t} = (x_0, y_0, z_0)^t$ da lente da câmera. Assim, em relação a uma origem que é a lente da câmera, o ponto $\mathbf{x}$ tem coordenadas

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} - \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix} \quad (4.1)$$

A projeção em perspectiva da câmera é obtida com respeito ao eixo z , cuja direção aponta para a frente do plano de visão. O eixo x é horizontalmente orientado, apontando para a direita, e o eixo y é verticalmente orientado, apontando para cima. Mas, as direções destes eixos no referencial da câmera podem

diferir das do referencial global. Representa-se a rotação pela qual o referencial da câmera é posicionado em correspondência com o referencial da câmera como uma seqüência de três rotações, cada qual em volta de cada um dos três eixos, e representadas por três ângulos ω , ϕ , e κ .

Portanto, se $\mathbf{R} = \mathbf{R}(\omega, \phi, \kappa)$ é a matriz de rotação de dimensões 3x3, tem-se $\mathbf{R}(\omega, \phi, \kappa) = \mathbf{R}(\omega)\mathbf{R}(\phi)\mathbf{R}(\kappa)$, onde $\mathbf{R}$ é dado por

$$\begin{pmatrix} \cos \phi \cos \kappa & \cos \omega \sin \kappa + \sin \omega \sin \phi \sin \kappa & \sin \omega \sin \kappa - \cos \omega \sin \phi \cos \kappa \\ -\cos \phi \sin \kappa & \cos \omega \cos \kappa - \sin \omega \sin \phi \sin \kappa & \sin \omega \cos \kappa + \cos \omega \sin \phi \sin \kappa \\ \sin \phi & -\sin \omega \cos \phi & \cos \omega \cos \phi \end{pmatrix} \quad (4.2)$$

Para representar relações envolvendo a matriz de rotação R e para evitar longas expressões trigonométricas, estabelece-se a convenção

$$\mathbf{R}(\omega, \phi, \kappa) = \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix} \quad (4.3)$$

Os ângulos ω , ϕ , e κ podem ser diretamente obtidos dos valores de r_{ij} :

$$\begin{aligned} \omega &= \arctan(-r_{32}/r_{33}) \\ \phi &= \arcsin(r_{31}) \\ \kappa &= \arctan(-r_{21}/r_{11}) \end{aligned} \quad (4.4)$$

Um ponto $\mathbf{x} = (x, y, z)^t$ do referencial do objeto é representado pelo ponto $\mathbf{x}' = (p, q, s)^t$ do referencial da câmera, onde

$$\begin{pmatrix} p \\ q \\ s \end{pmatrix} = \mathbf{R}(\omega, \phi, \kappa) \begin{pmatrix} x - x_0 \\ y - y_0 \\ z - z_0 \end{pmatrix} \quad (4.5)$$

Tendo uma representação para o ponto 3D no referencial da câmera, pode-se então tomar sua projeção em perspectiva. Uma vez que o plano de projeção da imagem esteja a uma distância f da lente da câmera, as coordenadas de projeção são dadas por

$$\begin{pmatrix} u \\ v \end{pmatrix} = \frac{f}{s} \begin{pmatrix} p \\ q \end{pmatrix} \quad (4.6)$$

Assume-se que o eixo óptico da câmera seja o eixo z .

As medições são feitas no plano de projeção da imagem. A origem do sistema de coordenadas do plano de medição da imagem não é necessariamente o ponto onde o eixo z da câmera intersecta o plano da imagem, ponto este chamado de ponto principal. Se sua localização é (u_0, v_0) no sistema de coordenadas do plano de medição da imagem, então as coordenadas do ponto projetado são escritas como

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} u_0 \\ v_0 \end{pmatrix} + \frac{f}{s} \begin{pmatrix} p \\ q \end{pmatrix} \quad (4.7)$$

Reescrevendo em termos das dependências rotacionais, tem-se

$$\frac{u - u_0}{f} = \frac{r_{11}(x - x_0) + r_{12}(y - y_0) + r_{13}(z - z_0)}{r_{31}(x - x_0) + r_{32}(y - y_0) + r_{33}(z - z_0)} \quad (4.8)$$

$$\frac{v - v_0}{f} = \frac{r_{21}(x - x_0) + r_{22}(y - y_0) + r_{23}(z - z_0)}{r_{31}(x - x_0) + r_{32}(y - y_0) + r_{33}(z - z_0)} \quad (4.9)$$

Este par de equações expressa uma razão de coordenadas da câmera em termos de uma razão envolvendo as coordenadas globais rotacionadas e translacionadas. É conhecido como a equação da projeção em perspectiva fundamental ou equação de colinearidade. E este par pode ser invertido de forma que uma razão de coordenadas globais possa ser expressa em termos de uma razão de coordenadas da câmera rotacionadas. Para ver isto, simplesmente precisa-se adicionar a identidade

$$1 = \frac{r_{31}(x - x_0) + r_{32}(y - y_0) + r_{33}(z - z_0)}{r_{31}(x - x_0) + r_{32}(y - y_0) + r_{33}(z - z_0)}$$

às duas equações anteriores e multiplicar ambos os lados do resultado por $\mathbf{R}^{-1}$.

Como $\mathbf{R}$ é uma matriz de rotação, $\mathbf{R}^{-1} = \mathbf{R}^t$. Desenvolvendo, chega-se a

$$\frac{x - x_0}{z - z_0} = \frac{r_{11}(u - u_0) + r_{21}(v - v_0) + r_{31}f}{r_{13}(u - u_0) + r_{23}(v - v_0) + r_{33}f} \quad (4.10)$$

$$\frac{y - y_0}{z - z_0} = \frac{r_{12}(u - u_0) + r_{22}(v - v_0) + r_{32}f}{r_{13}(u - u_0) + r_{23}(v - v_0) + r_{33}f} \quad (4.11)$$

As duas equações anteriores são também chamadas de equações de projeção em perspectiva fundamental geral, e mostram que a relação entre as coordenadas de projeção em perspectiva 2D medidas e as coordenadas 3D é uma função não-linear de $u_0, v_0, x_0, y_0, z_0, \omega, \phi$ e κ . Se forem fornecidos pares suficientes de

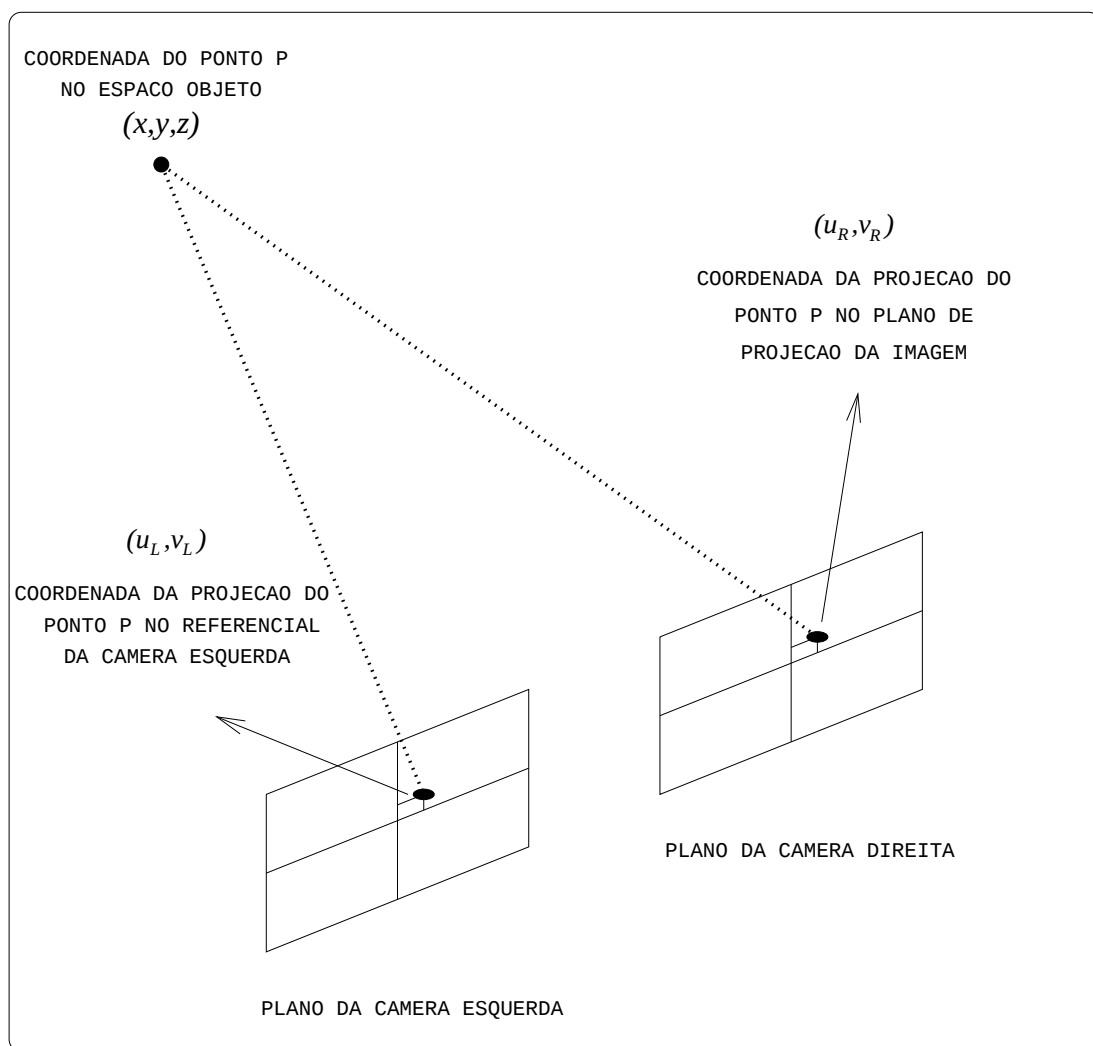


Figura 4.1: Referenciais global e da câmera.

pontos correspondentes 2D e 3D e uma solução aproximada inicial, pode-se obter os parâmetros desconhecidos $u_0, v_0, x_0, y_0, z_0, \omega, \phi$ e κ por uma técnica *least-squares* não-linear (descrita no Apêndice A).

4.2 Orientação Exterior

Em um problema de orientação exterior de uma câmera, um conjunto de N pontos 3D tendo posições conhecidas $(x_n, y_n, z_n)^t, n = 1, \dots, N$, em um referencial do objeto e o correspondente conjunto de pontos 2D tendo posições $(u_n, v_n)^t$ em um referencial da imagem são usados para a obtenção da rotação e translação desconhecidas que colocam o referencial da câmera no referencial global [8, 11].

Este problema é referido como o problema de calibração da câmera, quando se desconhece a posição da câmera no referencial do objeto, e como o problema de estimação de pose quando se desconhece a posição do objeto no referencial da câmera. A estimação da posição de pontos 3D de um modelo conhecido de suas orientações em perspectiva 2D é chamada de problema de resecção espacial.

Seja a translação desconhecida dada pelo vetor $\mathbf{t} = (x_0, y_0, z_0)^t$. O ponto $\mathbf{x}_n = (x_n, y_n, z_n)^t$ do referencial do objeto torna-se o ponto $\mathbf{x}'_n = (p_n, q_n, s_n)^t$ do referencial da câmera, onde

$$\begin{pmatrix} p_n \\ q_n \\ s_n \end{pmatrix} = \mathbf{R}(\omega, \phi, \kappa) \begin{pmatrix} x_n - x_0 \\ y_n - y_0 \\ z_n - z_0 \end{pmatrix}, \quad n = 1, \dots, N \quad (4.12)$$

A projeção em perspectiva observada é então dada por

$$\begin{pmatrix} u_n \\ v_n \end{pmatrix} = \frac{f}{s_n} \begin{pmatrix} p_n \\ q_n \end{pmatrix}, \quad n = 1, \dots, N \quad (4.13)$$

Este problema pode ser estabelecido como um problema *least-squares* não-linear, que é resolvido assumindo-se uma solução aproximada inicial com o objetivo de linearizar o problema. Soluções iniciais necessitam estar em torno de 10% da escala para os parâmetros de translação e em torno de 15° para os parâmetros de rotação. Assim sendo, e seguindo a notação usada no Apêndice A, definem-se inicialmente:

$$\boldsymbol{\beta} = (x_0, y_0, z_0, \omega, \phi, \kappa)^t$$

$$g_{2n-1}(\boldsymbol{\beta}) = u_n = fp_n/s_n \quad g_{2n}(\boldsymbol{\beta}) = v_n = fq_n/s_n \quad (4.14)$$

Para desenvolver-se a técnica padrão para a linearização far-se-á, por conveniência, uma pequena mudança na notação. Na equação da projeção em perspectiva, considera-se $(u_n, v_n)^t$ como uma função de $(x_0, y_0, z_0, \omega, \phi, \kappa)$, ou seja:

$$\begin{pmatrix} u_n \\ v_n \end{pmatrix} = \begin{pmatrix} u_n(x_0, y_0, z_0, \omega, \phi, \kappa) \\ v_n(x_0, y_0, z_0, \omega, \phi, \kappa) \end{pmatrix} = \frac{f}{s_n(x_0, y_0, z_0, \omega, \phi, \kappa)} \begin{pmatrix} p_n(x_0, y_0, z_0, \omega, \phi, \kappa) \\ q_n(x_0, y_0, z_0, \omega, \phi, \kappa) \end{pmatrix}$$

onde

$$\begin{pmatrix} p_n(x_0, y_0, z_0, \omega, \phi, \kappa) \\ q_n(x_0, y_0, z_0, \omega, \phi, \kappa) \\ s_n(x_0, y_0, z_0, \omega, \phi, \kappa) \end{pmatrix} = \mathbf{R}(\omega, \phi, \kappa) \begin{pmatrix} x_n - x_0 \\ y_n - y_0 \\ z_n - z_0 \end{pmatrix}$$

Então Δu_n e Δv_n indicam a variação no valor das funções u_n e v_n , devido a uma variação $(\Delta x_0, \Delta y_0, \Delta z_0, \Delta \omega, \Delta \phi, \Delta \kappa)$ no valor dos parâmetros $(x_0, y_0, z_0, \omega, \phi, \kappa)$.

Em cada iteração l a dificuldade principal refere-se à determinação do vetor $[\Delta g_1(\Delta \boldsymbol{\beta}; \boldsymbol{\beta}^l), \dots, \Delta g_{2N}(\Delta \boldsymbol{\beta}; \boldsymbol{\beta}^l)]^t = \mathbf{G}^l \Delta \boldsymbol{\beta}$. Devido ao fato de todas as componentes ímpares serem na realidade Δu_n para um n correspondente, e todas as componentes pares serem na realidade Δv_n para um n correspondente, tem-se:

$$\mathbf{G}^l \Delta \boldsymbol{\beta} = \begin{pmatrix} \Delta u_1 \\ \Delta v_1 \\ \Delta u_2 \\ \Delta v_2 \\ \dots \\ \Delta u_n \\ \Delta v_n \end{pmatrix} \quad (4.15)$$

Para encontrar a solução padrão, inicialmente aplica-se a regra da cadeia para a diferenciação parcial na equação 4.13 :

$$\Delta u_n = f \frac{s_n \Delta p_n - p_n \Delta s_n}{s_n^2} \quad \Delta v_n = f \frac{s_n \Delta q_n - q_n \Delta s_n}{s_n^2}$$

ou, em forma matricial:

$$\begin{pmatrix} \Delta u_n \\ \Delta v_n \end{pmatrix} = \frac{f}{s_n} \begin{pmatrix} 1 & 0 & -p_n/s_n \\ 0 & 1 & -q_n/s_n \end{pmatrix} \begin{pmatrix} \Delta p_n \\ \Delta q_n \\ \Delta s_n \end{pmatrix} \quad (4.16)$$

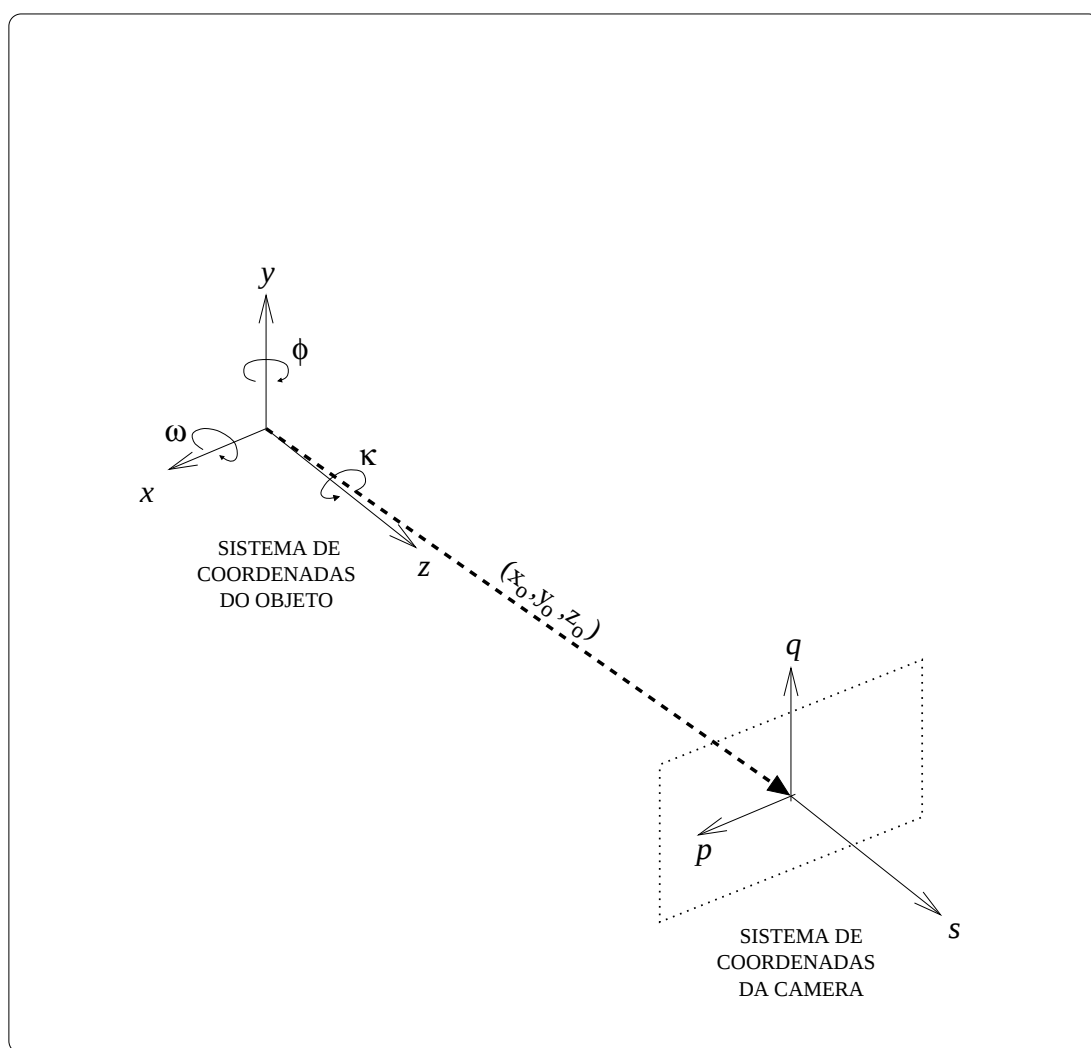


Figura 4.2: Orientação exterior.

Então,

$$\begin{pmatrix} \Delta p_n \\ \Delta q_n \\ \Delta s_n \end{pmatrix} = \mathbf{R}(\omega, \phi, \kappa) \begin{pmatrix} -\Delta x_0 \\ -\Delta y_0 \\ -\Delta z_0 \end{pmatrix} + \Delta \mathbf{R}(\omega, \phi, \kappa) \begin{pmatrix} x_n - x_0 \\ y_n - y_0 \\ z_n - z_0 \end{pmatrix} \quad (4.17)$$

onde

$$\Delta \mathbf{R}(\omega, \phi, \kappa) = \frac{\partial \mathbf{R}(\omega, \phi, \kappa)}{\partial \omega} \Delta \omega + \frac{\partial \mathbf{R}(\omega, \phi, \kappa)}{\partial \phi} \Delta \phi + \frac{\partial \mathbf{R}(\omega, \phi, \kappa)}{\partial \kappa} \Delta \kappa \quad (4.18)$$

Como $\mathbf{R}(\omega, \phi, \kappa) = \mathbf{R}(\kappa)\mathbf{R}(\phi)\mathbf{R}(\omega)$, pode-se escrever

$$\begin{aligned} \frac{\partial \mathbf{R}(\omega, \phi, \kappa)}{\Delta \omega} &= \mathbf{R}(\kappa)\mathbf{R}(\phi) \frac{\partial \mathbf{R}(\omega)}{\Delta \omega} \\ \frac{\partial \mathbf{R}(\omega, \phi, \kappa)}{\Delta \phi} &= \mathbf{R}(\kappa) \frac{\partial \mathbf{R}(\phi)}{\Delta \phi} \mathbf{R}(\omega) \\ \frac{\partial \mathbf{R}(\omega, \phi, \kappa)}{\Delta \kappa} &= \frac{\partial \mathbf{R}(\kappa)}{\Delta \kappa} i\mathbf{R}(\phi)i\mathbf{R}(\omega) \end{aligned}$$

Para a l -ésima iteração, tem-se

$$\begin{aligned} \begin{pmatrix} \Delta u_n \\ \Delta v_n \end{pmatrix} &= \frac{f}{s_n^l} \begin{pmatrix} 1 & 0 & -p_n^l/s_n^l \\ 0 & 1 & -q_n^l/s_n^l \end{pmatrix} \left[-R(\omega^l, \phi^l, \kappa^l) \begin{pmatrix} \Delta x_0 \\ \Delta y_0 \\ \Delta z_0 \end{pmatrix} \right. \\ &\quad \left. + \left(\frac{\partial \mathbf{R}}{\partial \omega}(\omega^l, \phi^l, \kappa^l) \Delta \omega + \frac{\partial \mathbf{R}}{\partial \phi}(\omega^l, \phi^l, \kappa^l) \Delta \phi + \frac{\partial \mathbf{R}}{\partial \kappa}(\omega^l, \phi^l, \kappa^l) \Delta \kappa \right) \begin{pmatrix} x_n - x_0^l \\ y_n - y_0^l \\ z_n - z_0^l \end{pmatrix} \right] \end{aligned}$$

onde

$$\begin{pmatrix} p_n^l \\ q_n^l \\ s_n^l \end{pmatrix} = \mathbf{R}(\omega^l, \phi^l, \kappa^l) \begin{pmatrix} x_n \\ y_n \\ z_n \end{pmatrix},$$

ou, em uma forma linear compacta,

$$\begin{pmatrix} \Delta u_n \\ \Delta v_n \end{pmatrix} = \mathbf{A}_n^l \mathbf{B}_n^l \begin{pmatrix} \Delta x_0 \\ \Delta y_0 \\ \Delta z_0 \\ \Delta \omega \\ \Delta \phi \\ \Delta \kappa \end{pmatrix} \quad (4.19)$$

onde

$$\mathbf{A}_n^l = \frac{f}{s_n^l} \begin{pmatrix} 1 & 0 & -p_n^l/s_n^l \\ 0 & 1 & -q_n^l/s_n^l \end{pmatrix} \quad (4.20)$$

$$\mathbf{B}_n^l = \begin{pmatrix} -\mathbf{R}(\omega^l, \phi^l, \kappa^l) & \mathbf{Q}(\omega^l, \phi^l, \kappa^l, x_0^l, y_0^l, z_0^l) \end{pmatrix} \quad (4.21)$$

sendo $\mathbf{Q}$ uma matriz 3 X 3 dada por

$$\mathbf{Q} = \begin{bmatrix} \frac{\partial \mathbf{R}}{\partial \omega} \begin{pmatrix} x_0 - x_0^l \\ y_0 - y_0^l \\ z_0 - z_0^l \end{pmatrix} & \frac{\partial \mathbf{R}}{\partial \phi} \begin{pmatrix} x_0 - x_0^l \\ y_0 - y_0^l \\ z_0 - z_0^l \end{pmatrix} & \frac{\partial \mathbf{R}}{\partial \kappa} \begin{pmatrix} x_0 - x_0^l \\ y_0 - y_0^l \\ z_0 - z_0^l \end{pmatrix} \end{bmatrix} \quad (4.22)$$

O procedimento acima conduz ao seguinte algoritmo para a solução do problema: na iteração l , a solução avaliável é $\boldsymbol{\beta}^l = (x_0^l, y_0^l, z_0^l, \omega_0^l, \phi_0^l, \kappa_0^l)$. O enésimo par de coordenadas observadas na imagem é $(u_n, v_n)^t$, e originam-se da projeção em perspectiva de $(x_n, y_n, z_n)^t$. Próximo à solução aproximada, poder-se-ia inferir que a projeção em perspectiva de $(x_n, y_n, z_n)^t$ teria coordenadas imagem $(u_n, v_n)^t$, onde

$$\begin{pmatrix} u_n^l \\ v_n^l \end{pmatrix} = \frac{f}{s_n^l} \begin{pmatrix} p_n^l \\ q_n^l \end{pmatrix}$$

$$\begin{pmatrix} p_n^l \\ q_n^l \\ s_n^l \end{pmatrix} = R(\omega^l, \phi^l, \kappa^l) \begin{pmatrix} x_n^l - x_0^l \\ y_n^l - y_0^l \\ z_n^l - z_0^l \end{pmatrix}$$

Os vetores $\boldsymbol{\alpha}$ e $\mathbf{g}^l$ da equação *least-squares* são então

$$\boldsymbol{\alpha} = \begin{pmatrix} u_1 \\ v_1 \\ \vdots \\ u_N \\ v_N \end{pmatrix} \quad \text{e} \quad \mathbf{g} = \begin{pmatrix} u_1^l \\ v_1^l \\ \vdots \\ u_N^l \\ v_N^l \end{pmatrix}$$

e a matriz $2N \times 6$ é dada por

$$\mathbf{G}^l = \begin{pmatrix} \mathbf{A}_1^l \mathbf{B}_1^l \\ \vdots \\ \mathbf{A}_N^l \mathbf{B}_N^l \end{pmatrix} \quad (4.23)$$

Para o cálculo de $\Delta\boldsymbol{\beta} = (\Delta x_0, \Delta y_0, \Delta z_0, \Delta\omega, \Delta\phi, \Delta\kappa)^t$ é usada a Eq. A.5 (do Apêndice A). A nova solução aproximada é então dada por

$$\begin{pmatrix} x_0^{l+1} \\ y_0^{l+1} \\ z_0^{l+1} \\ \omega^{l+1} \\ \phi^{l+1} \\ \kappa^{l+1} \end{pmatrix} = \begin{pmatrix} x_0^l \\ y_0^l \\ z_0^l \\ \omega^l \\ \phi^l \\ \kappa^l \end{pmatrix} + \begin{pmatrix} \Delta x_0 \\ \Delta y_0 \\ \Delta z_0 \\ \Delta\omega \\ \Delta\phi \\ \Delta\kappa \end{pmatrix} \quad (4.24)$$

4.3 Orientação Interior

A orientação interior de uma câmera [8] é especificada pela constante da câmera f , que é a distância entre o plano da imagem e a lente da câmera; pelo ponto principal $(u_p, v_p)^t$, que é a interseção do eixo óptico com o plano da imagem no referencial de medição localizado no plano da imagem; e pelas características de distorção geométrica da lente, que é assumida aqui ser isotrópica em volta do ponto principal.

Para determinar a orientação interior, precisa-se geralmente resolver o problema de orientação absoluta, com o problema adicional de determinar os parâmetros acima citados.

O modelo usado para a distorção geométrica é

$$(k_1 r_n^2 + k_2 r_n^4 + k_3 r_n^6) \begin{pmatrix} u_n - u_p \\ v_n - v_p \end{pmatrix} = \frac{f}{s_n} \begin{pmatrix} p_n \\ q_n \end{pmatrix}, \quad n = 1, \dots, N \quad (4.25)$$

onde

$$\begin{pmatrix} p_n(x_0, y_0, z_0, \omega, \phi, \kappa) \\ q_n(x_0, y_0, z_0, \omega, \phi, \kappa) \\ s_n(x_0, y_0, z_0, \omega, \phi, \kappa) \end{pmatrix} = \mathbf{R}(\omega, \phi, \kappa) \begin{pmatrix} x_n - x_0 \\ y_n - y_0 \\ z_n - z_0 \end{pmatrix}$$

e

$$r_n^2(u_p, v_p) = (u_n - u_p)^2 + (v_n - v_p)^2 \quad (4.26)$$

Os parâmetros que devem ser determinados são os parâmetros de orientação interior $f, u_p, v_p, k_1, k_2, k_3$, e os parâmetros de orientação exterior $x_0, y_0, z_0, \omega, \phi, \kappa$. Para determinar os valores dos parâmetros desconhecidos, resolve-se o sistema

não-linear da Eq. 4.25 linearizando iterativamente em volta da solução corrente aproximada e resolvendo o sistema linearizado para a próxima solução aproximada. O termo $k_1 r_n^2 + k_2 r_n^4 + k_3 r_n^6$ contribui somente para a distorção geométrica.

Linearizando inicialmente a Eq. 4.25, chega-se a

$$\begin{aligned} & [(k_1 + \Delta k_1) r_n^2 (u_p + \Delta u_p, v_p + \Delta v_p) + (k_2 + \Delta k_2) r_n^4 (u_p + \Delta u_p, v_p + \Delta v_p) + \\ & + (k_3 + \Delta k_3) r_n^6 (u_p + \Delta u_p, v_p + \Delta v_p)] \begin{bmatrix} u_n - (u_p + \Delta u_p) \\ v_n - (v_p + \Delta v_p) \end{bmatrix} = \\ & = (k_1^l r_n^{l^2} + k_2^l r_n^{l^4} + k_3^l r_n^{l^6}) \begin{pmatrix} u_n - u_p^l \\ v_n - v_p^l \end{pmatrix} + \mathbf{C}_n^l \begin{pmatrix} \Delta u_p \\ \Delta v_p \\ \Delta k_1 \\ \Delta k_2 \\ \Delta k_3 \end{pmatrix} \end{aligned}$$

onde

$$\mathbf{C}_n^l = \begin{pmatrix} c_{11_n}^l & c_{12_n}^l & c_{13_n}^l & c_{14_n}^l & c_{15_n}^l \\ c_{21_n}^l & c_{22_n}^l & c_{23_n}^l & c_{24_n}^l & c_{25_n}^l \end{pmatrix} \quad (4.27)$$

$$\begin{aligned} c_{11_n}^l &= -(k_1^l r_n^{l^2} + k_2^l r_n^{l^4} + k_3^l r_n^{l^6}) + (u_n - u_p^l)(2k_1^l r_n^l + 2k_2^l r_n^{l^3} + 2k_3^l r_n^{l^5}) \\ c_{12_n}^l &= (u_n - u_p^l)(2k_1^l r_n^l + 2k_2^l r_n^{l^3} + 2k_3^l r_n^{l^5}) \\ c_{13_n}^l &= (u_n - u_p^l) r_n^{l^2} \\ c_{14_n}^l &= (u_n - u_p^l) r_n^{l^4} \\ c_{15_n}^l &= (u_n - u_p^l) r_n^{l^6} \\ c_{21_n}^l &= (v_n - v_p^l)(2k_1^l r_n^l + 2k_2^l r_n^{l^3} + 2k_3^l r_n^{l^5}) \\ c_{22_n}^l &= -(k_1^l r_n^{l^2} + k_2^l r_n^{l^4} + k_3^l r_n^{l^6}) + (v_n - v_p^l)(2k_1^l r_n^l + 2k_2^l r_n^{l^3} + 2k_3^l r_n^{l^5}) \\ c_{23_n}^l &= (v_n - v_p^l) r_n^{l^2} \\ c_{24_n}^l &= (v_n - v_p^l) r_n^{l^4} \\ c_{25_n}^l &= (v_n - v_p^l) r_n^{l^6} \end{aligned}$$

Agora, lineariza-se o lado direito da Eq. 4.25. Na iteração l , tem-se

$$\begin{aligned} & \frac{(f^l + \Delta f) \begin{bmatrix} p_n(x_0^l + \Delta x_0, y_0 + \Delta y_0, z_0 + \Delta z_0, \omega_0 + \Delta \omega_0, \phi_0 + \Delta \phi_0, \kappa_0 + \Delta \kappa_0) \\ q_n(x_0^l + \Delta x_0, y_0 + \Delta y_0, z_0 + \Delta z_0, \omega_0 + \Delta \omega_0, \phi_0 + \Delta \phi_0, \kappa_0 + \Delta \kappa_0) \end{bmatrix}}{s_n(x_0^l + \Delta x_0, y_0 + \Delta y_0, z_0 + \Delta z_0, \omega_0 + \Delta \omega_0, \phi_0 + \Delta \phi_0, \kappa_0 + \Delta \kappa_0)} = \\ & = \frac{f^l \begin{bmatrix} p_n(x_0^l, y_0^l, z_0^l, \omega_0^l, \phi_0^l, \kappa_0^l) \\ q_n(x_0^l, y_0^l, z_0^l, \omega_0^l, \phi_0^l, \kappa_0^l) \end{bmatrix}}{s_n(x_0^l, y_0^l, z_0^l, \omega_0^l, \phi_0^l, \kappa_0^l)} + \mathbf{D}^l \begin{pmatrix} \Delta f \\ \Delta x_0 \\ \Delta y_0 \\ \Delta z_0 \\ \Delta \omega \\ \Delta \phi \\ \Delta \kappa \end{pmatrix} \end{aligned}$$

onde

$$\mathbf{D}_n^l = \begin{bmatrix} \frac{p_n(x_0^l, y_0^l, z_0^l, \omega^l, \phi^l, \kappa^l)}{s_n(x_0^l, y_0^l, z_0^l, \omega^l, \phi^l, \kappa^l)} \\ \frac{q_n(x_0^l, y_0^l, z_0^l, \omega^l, \phi^l, \kappa^l)}{s_n(x_0^l, y_0^l, z_0^l, \omega^l, \phi^l, \kappa^l)} \end{bmatrix} \mathbf{A}_n^l \mathbf{B}_n^l \quad (4.28)$$

e

$$\mathbf{A}_n^l = \frac{f^l}{s_n(x_0^l, y_0^l, z_0^l, \omega^l, \phi^l, \kappa^l)} \begin{bmatrix} 1 & 0 & -\frac{p_n(x_0^l, y_0^l, z_0^l, \omega^l, \phi^l, \kappa^l)}{s_n(x_0^l, y_0^l, z_0^l, \omega^l, \phi^l, \kappa^l)} \\ 0 & 1 & -\frac{q_n(x_0^l, y_0^l, z_0^l, \omega^l, \phi^l, \kappa^l)}{s_n(x_0^l, y_0^l, z_0^l, \omega^l, \phi^l, \kappa^l)} \end{bmatrix} \quad (4.29)$$

$$\mathbf{B}_n^l = \begin{bmatrix} -\mathbf{R}(\omega^l, \phi^l, \kappa^l) & \mathbf{Q}(\omega^l, \phi^l, \kappa^l, x_0^l, y_0^l, z_0^l) \end{bmatrix} \quad (4.30)$$

$$\mathbf{Q} = \begin{bmatrix} \frac{\partial \mathbf{R}}{\partial \omega} \begin{pmatrix} x_0 - x_0^l \\ y_0 - y_0^l \\ z_0 - z_0^l \end{pmatrix} & \frac{\partial \mathbf{R}}{\partial \phi} \begin{pmatrix} x_0 - x_0^l \\ y_0 - y_0^l \\ z_0 - z_0^l \end{pmatrix} & \frac{\partial \mathbf{R}}{\partial \kappa} \begin{pmatrix} x_0 - x_0^l \\ y_0 - y_0^l \\ z_0 - z_0^l \end{pmatrix} \end{bmatrix} \quad (4.31)$$

Juntando as formas linearizadas, tem-se

$$(k_1^l r_n^{l^2} + k_2^l r_n^{l^4} + k_3^l r_n^{l^6}) \begin{pmatrix} u_n - u_p^l \\ v_n - v_p^l \end{pmatrix} + \mathbf{C}_n^l \begin{pmatrix} \Delta u_p \\ \Delta v_p \\ \Delta k_1 \\ \Delta k_2 \\ \Delta k_3 \end{pmatrix} = \frac{f^l}{s_n^l} \begin{pmatrix} p_n^l \\ q_n^l \end{pmatrix} + \mathbf{D}_n^l \begin{pmatrix} \Delta f \\ \Delta x_0 \\ \Delta y_0 \\ \Delta z_0 \\ \Delta \omega \\ \Delta \phi \\ \Delta \kappa \end{pmatrix}$$

Rearranjando,

$$\mathbf{E}_n^l \begin{pmatrix} \Delta u_p \\ \Delta v_p \\ \Delta k_1 \\ \Delta k_2 \\ \Delta k_3 \\ \Delta f \\ \Delta x_0 \\ \Delta y_0 \\ \Delta z_0 \\ \Delta \omega \\ \Delta \phi \\ \Delta \kappa \end{pmatrix} = (k_1^l r_n^{l^2} + k_2^l r_n^{l^4} + k_3^l r_n^{l^6}) \begin{pmatrix} u_n - u_p^l \\ v_n - v_p^l \end{pmatrix} - \frac{f^l}{s_n^l} \begin{pmatrix} p_n^l \\ q_n^l \end{pmatrix} = h_n^l \quad (4.32)$$

onde

$$\mathbf{E}_n^l = \begin{pmatrix} -\mathbf{A}_n^l & \mathbf{D}_n^l \end{pmatrix} \quad (4.33)$$

O sistema *least-squares* total linearizado então torna-se o seguinte sistema linear de $2N$ equações:

$$\mathbf{E}^l \begin{pmatrix} \Delta u_p \\ \Delta v_p \\ \Delta k_1 \\ \Delta k_2 \\ \Delta k_3 \\ \Delta f \\ \Delta x_0 \\ \Delta y_0 \\ \Delta z_0 \\ \Delta \omega \\ \Delta \phi \\ \Delta \kappa \end{pmatrix} = \mathbf{h}^l \quad (4.34)$$

onde

$$\mathbf{E}^l = \begin{pmatrix} \mathbf{E}_1^l \\ \mathbf{E}_2^l \\ \vdots \\ \mathbf{E}_N^l \end{pmatrix} \quad \text{e} \quad \mathbf{h}^l = \begin{pmatrix} \mathbf{h}_1^l \\ \mathbf{h}_2^l \\ \vdots \\ \mathbf{h}_N^l \end{pmatrix}$$

Assumindo que as perturbações associadas com cada linha do sistema são não correlacionadas e têm variâncias idênticas, obtém-se

$$\begin{pmatrix} \Delta u_p \\ \Delta v_p \\ \Delta k_1 \\ \Delta k_2 \\ \Delta k_3 \\ \Delta f \\ \Delta x_0 \\ \Delta y_0 \\ \Delta z_0 \\ \Delta \omega \\ \Delta \phi \\ \Delta \kappa \end{pmatrix} = (\mathbf{E}^{l^t} \mathbf{E}^l)^{-1} \mathbf{E}^{l^t} \mathbf{h}^l$$

Então a solução aproximada é definida por

$$\begin{pmatrix} u_p^{t+1} \\ v_p^{t+1} \\ k_1^{t+1} \\ k_2^{t+1} \\ k_3^{t+1} \\ f^{t+1} \\ x_0^{t+1} \\ y_0^{t+1} \\ z_0^{t+1} \\ \omega^{t+1} \\ \phi^{t+1} \\ \kappa^{t+1} \end{pmatrix} = \begin{pmatrix} u_p^l \\ v_p^l \\ k_1^l \\ k_2^l \\ k_3^l \\ f^l \\ x_0^l \\ y_0^l \\ z_0^l \\ \omega^l \\ \phi^l \\ \kappa^l \end{pmatrix} + \begin{pmatrix} \Delta u_p \\ \Delta v_p \\ \Delta k_1 \\ \Delta k_2 \\ \Delta k_3 \\ \Delta f \\ \Delta x_0 \\ \Delta y_0 \\ \Delta z_0 \\ \Delta \omega \\ \Delta \phi \\ \Delta \kappa \end{pmatrix} \quad (4.35)$$

4.4 Orientação Relativa

A orientação relativa diz respeito à determinação da posição e orientação de uma fotografia com relação a outra, dado um conjunto de pontos imagem correspondentes. Depois de uma câmera estar em orientação relativa com outra, pontos correspondentes das duas fotografias intersectam-se no espaço 3D [8].

Sejam $(x_L, y_L, z_L)^t$ e $(x_R, y_R, z_R)^t$ as posições 3D das lentes das câmeras da direita e da esquerda, respectivamente. Sejam $(\omega_L, \phi_L, \kappa_L)^t$ e $(\omega_R, \phi_R, \kappa_R)^t$ os ângulos de rotação especificando a orientação exterior dos referenciais das câmeras da esquerda e da direita, respectivamente. Sejam dados um conjunto de N pontos-imagem $\{(u_{Ln}, v_{Ln})^t\}_{n=1}^N$ da imagem esquerda e o conjunto de N pontos correspondentes $\{(u_{Rn}, v_{Rn})^t\}_{n=1}^N$ da imagem direita. O ponto $(u_{Ln}, v_{Ln})^t$ na imagem esquerda corresponde ao ponto $(u_{Rn}, v_{Rn})^t$ na imagem direita se houver algum ponto 3D $(x_n, y_n, z_n)^t$ tal que $(u_{Ln}, v_{Ln})^t$ seja a projeção em perspectiva de $(x_n, y_n, z_n)^t$ na imagem esquerda e $(u_{Rn}, v_{Rn})^t$ seja a projeção em perspectiva de $(x_n, y_n, z_n)^t$ na imagem direita.

Considera-se a separação das lentes das duas câmeras ao longo do eixo x como uma constante que controla a escala, não sendo por isso um parâmetro na orientação relativa. Portanto, a orientação relativa é especificada pelos cinco parâmetros $(y_R - y_L)$, $(z_R - z_L)$, $(\omega_R - \omega_L)$, $(\phi_R - \phi_L)$ e $(\kappa_R - \kappa_L)$. A solução assume que a orientação interior das câmeras é conhecida e que todas as posições imagem são expressas em escala idêntica e com respeito aos pontos principais.

Para resolver o problema de orientação relativa, considera-se inicialmente

$$\mathbf{Q}_L^t = \mathbf{R}(\omega_L, \phi_L, \kappa_L) \quad (4.36)$$

como sendo a matriz de rotação associada com a orientação exterior da imagem esquerda e

$$\mathbf{Q}_R^t = \mathbf{R}(\omega_R, \phi_R, \kappa_R) \quad (4.37)$$

como sendo a matriz de rotação associada com a orientação exterior da imagem direita. Seja f_R a distância entre o plano da imagem direita e a lente direita, e f_L a distância entre o plano da imagem esquerda e a lente esquerda. Então, da equação de colinearidade da projeção em perspectiva,

$$\begin{pmatrix} u_{Ln} \\ v_{Ln} \\ f_L \end{pmatrix} = \frac{1}{\lambda_{Ln}} \mathbf{Q}_L^t \begin{pmatrix} x_n - x_L \\ y_n - y_L \\ z_n - z_L \end{pmatrix} \quad \text{e} \quad \begin{pmatrix} u_{Rn} \\ v_{Rn} \\ f_R \end{pmatrix} = \frac{1}{\lambda_{Rn}} \mathbf{Q}_R^t \begin{pmatrix} x_n - x_R \\ y_n - y_R \\ z_n - z_R \end{pmatrix} \quad (4.38)$$

Portanto,

$$\begin{pmatrix} x_n \\ y_n \\ z_n \end{pmatrix} = \begin{pmatrix} x_L \\ y_L \\ z_L \end{pmatrix} + \lambda_{Ln} \mathbf{Q}_L \begin{pmatrix} u_{Ln} \\ v_{Ln} \\ f_L \end{pmatrix} = \begin{pmatrix} x_R \\ y_R \\ z_R \end{pmatrix} + \lambda_{Rn} \mathbf{Q}_R \begin{pmatrix} u_{Rn} \\ v_{Rn} \\ f_R \end{pmatrix} \quad (4.39)$$

onde $\lambda_{Ln} = \frac{s_{Ln}}{f_L}$ e $\lambda_{Rn} = \frac{s_{Rn}}{f_R}$. Subtraindo uma equação da outra, tem-se

$$0 = \begin{pmatrix} x_R \\ y_R \\ z_R \end{pmatrix} - \begin{pmatrix} x_L \\ y_L \\ z_L \end{pmatrix} + \lambda_{Rn} \mathbf{Q}_R \begin{pmatrix} u_{Rn} \\ v_{Rn} \\ f_R \end{pmatrix} - \lambda_{Ln} \mathbf{Q}_L \begin{pmatrix} u_{Ln} \\ v_{Ln} \\ f_L \end{pmatrix} \quad (4.40)$$

Como o vetor $\mathbf{Q}_R \begin{pmatrix} u_{Rn} \\ v_{Rn} \\ f_R \end{pmatrix} \times \mathbf{Q}_L \begin{pmatrix} u_{Ln} \\ v_{Ln} \\ f_L \end{pmatrix}$ é ortogonal tanto a $\mathbf{Q}_R \begin{pmatrix} u_{Rn} \\ v_{Rn} \\ f_R \end{pmatrix}$ quanto a $\mathbf{Q}_L \begin{pmatrix} u_{Ln} \\ v_{Ln} \\ f_L \end{pmatrix}$, obtém-se a equação de coplanaridade, na qual cada par

de pontos imagem correspondentes precisa satisfazer a

$$\begin{pmatrix} x_n - x_R \\ y_n - y_R \\ z_n - z_R \end{pmatrix}^t \left[\mathbf{Q}_R \begin{pmatrix} u_{Rn} \\ v_{Rn} \\ f_R \end{pmatrix} \times \mathbf{Q}_L \begin{pmatrix} u_{Ln} \\ v_{Ln} \\ f_L \end{pmatrix} \right] = 0, \quad n = 1, \dots, N \quad (4.41)$$

Uma vez que a orientação relativa é o referencial da imagem direita especificado com respeito ao referencial da imagem esquerda, tem-se $(x_L, y_L, z_L)^t = 0$ e $\mathbf{Q}_L = \mathbf{I}$, a matriz identidade. Além disso, x_R é conhecido. Disto é claro que há no mínimo cinco pontos para estabelecer equações de coplanaridade para resolver o problema para os cinco parâmetros y_R , z_R , ω_R , ϕ_R e κ_R . Nesta formulação o sistema não-linear que tem de ser resolvido no critério *least-squares* para y_R , z_R , ω_R , ϕ_R e κ_R é

$$\begin{pmatrix} x_R \\ y_R \\ z_R \end{pmatrix}^t \left[\mathbf{Q}_R \begin{pmatrix} u_{Rn} \\ v_{Rn} \\ f_R \end{pmatrix} \times \mathbf{Q}_L \begin{pmatrix} u_{Ln} \\ v_{Ln} \\ f_L \end{pmatrix} \right] = 0, \quad n = 1, \dots, N \quad (4.42)$$

Para resolver-se este sistema, assume-se uma solução aproximada inicial y_R^0 , z_R^0 , ω_R^0 , ϕ_R^0 , κ_R^0 . Em cada iteração linearizam-se as equações não-lineares em volta da solução corrente e resolve-se o sistema linearizado no critério *least-squares*.

Fazendo-se $\mathbf{R}(\omega_R, \phi_R, \kappa_R) = \mathbf{Q}_R$, tem-se em cada iteração l

$$\begin{aligned}
 0 &= \begin{pmatrix} x_R \\ y_R^l + \Delta y \\ z_R^l + \Delta z \end{pmatrix}^t \left[\mathbf{R}(\omega_R^l + \Delta\omega, \phi_R^l + \Delta\phi, \kappa_R^l + \Delta\kappa) \begin{pmatrix} u_{Ln} \\ v_{Ln} \\ f_L \end{pmatrix} \times \begin{pmatrix} u_{Ln} \\ v_{Ln} \\ f_L \end{pmatrix} \right] \\
 &= \begin{pmatrix} x_R \\ y_R^l \\ z_R^l \end{pmatrix}^t \left[\mathbf{R}(\omega_R^l, \phi_R^l, \kappa_R^l) \begin{pmatrix} u_{Ln} \\ v_{Ln} \\ f_L \end{pmatrix} \times \begin{pmatrix} u_{Ln} \\ v_{Ln} \\ f_L \end{pmatrix} \right] + \\
 &\quad \begin{pmatrix} x_R \\ y_R^l \\ z_R^l \end{pmatrix}^t \left[\Delta\mathbf{R}(\omega_R^l, \phi_R^l, \kappa_R^l) \begin{pmatrix} u_{Ln} \\ v_{Ln} \\ f_L \end{pmatrix} \times \begin{pmatrix} u_{Ln} \\ v_{Ln} \\ f_L \end{pmatrix} \right] + \\
 &\quad \begin{pmatrix} 0 \\ \Delta y \\ \Delta z \end{pmatrix}^t \left[\mathbf{R}(\omega_R^l, \phi_R^l, \kappa_R^l) \begin{pmatrix} u_{Ln} \\ v_{Ln} \\ f_L \end{pmatrix} \times \begin{pmatrix} u_{Ln} \\ v_{Ln} \\ f_L \end{pmatrix} \right]
 \end{aligned}$$

Como

$$\Delta\mathbf{R}(\omega_R^l, \phi_R^l, \kappa_R^l) = \frac{\partial\mathbf{R}}{\partial\omega}(\omega_R^l, \phi_R^l, \kappa_R^l)\Delta\omega + \frac{\partial\mathbf{R}}{\partial\phi}(\omega_R^l, \phi_R^l, \kappa_R^l)\Delta\phi + \frac{\partial\mathbf{R}}{\partial\kappa}(\omega_R^l, \phi_R^l, \kappa_R^l)\Delta\kappa \quad (4.43)$$

pode-se reescrever o sistema acima como o seguinte sistema linear

$$\begin{pmatrix} a_{11}^l & a_{21}^l & a_{31}^l & a_{41}^l & a_{51}^l \\ & & \vdots & & \\ a_{1N}^l & a_{2N}^l & a_{3N}^l & a_{4N}^l & a_{5N}^l \end{pmatrix} \begin{pmatrix} \Delta\omega \\ \Delta\phi \\ \Delta\kappa \\ \Delta y \\ \Delta z \end{pmatrix} = -x_R \begin{pmatrix} a_{41}^l \\ \vdots \\ a_{4N}^l \end{pmatrix} - y_R \begin{pmatrix} a_{51}^l \\ \vdots \\ a_{5N}^l \end{pmatrix} - z_R \begin{pmatrix} a_{61}^l \\ \vdots \\ a_{6N}^l \end{pmatrix} \quad (4.44)$$

onde

$$a_{1n}^l = \begin{pmatrix} x_R \\ y_R^l \\ z_R^l \end{pmatrix}^t \left[\frac{\Delta \mathbf{R}}{\Delta \omega}(\omega_R^l, \phi_R^l, \kappa_R^l) \begin{pmatrix} u_{Rn} \\ v_{Rn} \\ f_R \end{pmatrix} \times \begin{pmatrix} u_{Ln} \\ v_{Ln} \\ f_L \end{pmatrix} \right] \quad (4.45)$$

$$a_{2n}^l = \begin{pmatrix} x_R \\ y_R^l \\ z_R^l \end{pmatrix}^t \left[\frac{\Delta \mathbf{R}}{\Delta \phi}(\omega_R^l, \phi_R^l, \kappa_R^l) \begin{pmatrix} u_{Rn} \\ v_{Rn} \\ f_R \end{pmatrix} \times \begin{pmatrix} u_{Ln} \\ v_{Ln} \\ f_L \end{pmatrix} \right] \quad (4.46)$$

$$a_{3n}^l = \begin{pmatrix} x_R \\ y_R^l \\ z_R^l \end{pmatrix}^t \left[\frac{\Delta \mathbf{R}}{\Delta \kappa}(\omega_R^l, \phi_R^l, \kappa_R^l) \begin{pmatrix} u_{Rn} \\ v_{Rn} \\ f_R \end{pmatrix} \times \begin{pmatrix} u_{Ln} \\ v_{Ln} \\ f_L \end{pmatrix} \right] \quad (4.47)$$

$$\begin{pmatrix} a_{4n}^l \\ a_{5n}^l \\ a_{6n}^l \end{pmatrix} = \mathbf{R}(\omega_R^l, \phi_R^l, \kappa_R^l) \begin{pmatrix} u_{Rn}^l \\ v_{Rn}^l \\ f_R \end{pmatrix} \times \begin{pmatrix} u_{Ln}^l \\ v_{Ln}^l \\ f_L \end{pmatrix}, \quad n = 1, \dots, N \quad (4.48)$$

Este sistema linear é resolvido no critério *least-squares*, e os valores ajustados para a próxima iteração são dados por

$$\begin{pmatrix} \omega_R^{t+1} \\ \phi_R^{t+1} \\ \kappa_R^{t+1} \\ y_R^{t+1} \\ z_R^{t+1} \end{pmatrix} = \begin{pmatrix} \omega_R^l \\ \phi_R^l \\ \kappa_R^l \\ y_R^l \\ z_R^l \end{pmatrix} + \begin{pmatrix} \Delta \omega_R \\ \Delta \phi_R \\ \Delta \kappa_R \\ \Delta y_R \\ \Delta z_R \end{pmatrix} \quad (4.49)$$

4.5 Orientação Absoluta

Supõe-se que foi determinada a localização de N pontos 3D $p_1, \dots, p_N$ relativos ao referencial da câmera. Supõe-se também que no referencial do objeto estes mesmos N pontos 3D tenham localizações $q_1, \dots, q_N$. Para responder à questão “Onde está o objeto?”, precisa-se determinar onde o referencial do modelo está com respeito ao referencial da câmera. Para responder à questão “Onde está a câmera?”, precisa-se determinar onde o referencial da câmera está com respeito

ao referencial do modelo. Portanto, precisa-se determinar uma matriz de rotação R e um vetor de translação t que satisfaçam

$$\mathbf{p} = \mathbf{R}\mathbf{q} + \mathbf{t} \quad (4.50)$$

O problema de orientação absoluta 3D-3D, por vezes chamado o problema de estimação de pose [8], é inferir R e T de $q_1, \dots, q_N$ e $p_1, \dots, p_N$. Uma generalização deste problema é assumir uma escala desconhecida. Então,

$$\mathbf{p}_n = s\mathbf{R}\mathbf{q}_n + t, \quad n = 1, \dots, N \quad (4.51)$$

Assume-se que s é conhecido e igual a 1.

Para determinar $\mathbf{R}$ e $\mathbf{T}$, estabelece-se um problema *least-squares*. Minimizar-se-á $\sum_{n=1}^N \mathbf{w}_n \|\mathbf{p}_n - (\mathbf{R}\mathbf{q}_n + \mathbf{t})\|^2$ sujeito à condição de $\mathbf{R}$ ser uma matriz de rotação, isto é, $\mathbf{R}^t = \mathbf{R}^{-1}$. Para expressar estas condições usando multiplicadores de Lagrange, faz-se

$$\mathbf{R} = \begin{pmatrix} \mathbf{r}_1^t \\ \mathbf{r}_2^t \\ \mathbf{r}_3^t \end{pmatrix} \quad (4.52)$$

onde cada $\mathbf{r}_i$ é um vetor 3 x 1. A condição $\mathbf{R}^t = \mathbf{R}^{-1}$ então incorpora as seis equações

$$\begin{aligned} \mathbf{r}_1^t \mathbf{r}_1 &= \mathbf{r}_2^t \mathbf{r}_2 = \mathbf{r}_3^t \mathbf{r}_3 = 1 \\ \mathbf{r}_1^t \mathbf{r}_2 &= \mathbf{r}_1^t \mathbf{r}_3 = \mathbf{r}_2^t \mathbf{r}_3 = 0 \end{aligned} \quad (4.53)$$

O problema *least-squares* com estas condições pode ser escrito como a minimização de ϵ^2 , onde

$$\epsilon^2 = \sum_{n=1}^N \sum_{k=1}^3 \mathbf{w}_n (\mathbf{p}_{nk} - \mathbf{r}_k^t \mathbf{q}_n - \mathbf{t}_k)^2 + \sum_{k=1}^3 \lambda_k (\mathbf{r}_k^t \mathbf{r}_k - 1) + 2\lambda_4 \mathbf{r}_1^t \mathbf{r}_2 + 2\lambda_5 \mathbf{r}_1^t \mathbf{r}_3 + 2\lambda_6 \mathbf{r}_2^t \mathbf{r}_3 \quad (4.54)$$

$$\mathbf{p}_n = \begin{pmatrix} p_{n1} \\ p_{n2} \\ p_{n3} \end{pmatrix} \quad \mathbf{q}_n = \begin{pmatrix} q_{n1} \\ q_{n2} \\ q_{n3} \end{pmatrix} \quad \mathbf{t} = \begin{pmatrix} t_1 \\ t_2 \\ t_3 \end{pmatrix} \quad (4.55)$$

Tomando-se a derivada parcial de ϵ^2 com relação a t_n ,

$$\frac{\partial \epsilon^2}{\partial t_k} = \sum_{n=1}^N 2\mathbf{w}_n (\mathbf{p}_{nk} - \mathbf{r}_k^t \mathbf{q}_n - \mathbf{t}_k)(-1), \quad k = 1, 2, 3 \quad (4.56)$$

igualando estas derivadas parciais a zero,

$$\sum_{n=1}^N \mathbf{w}_n (\mathbf{p}_n - \mathbf{R}\mathbf{q}_n - \mathbf{t}) = 0 \quad (4.57)$$

Rearranjando, obtém-se

$$\mathbf{t} = \bar{\mathbf{p}} - \mathbf{R}\bar{\mathbf{q}} \quad (4.58)$$

onde

$$\bar{\mathbf{q}} = \frac{\sum_{n=1}^N \mathbf{w}_n \mathbf{q}_n}{\sum_{n=1}^N \mathbf{w}_n} \quad \text{e} \quad \bar{\mathbf{p}} = \frac{\sum_{n=1}^N \mathbf{w}_n \mathbf{p}_n}{\sum_{n=1}^N \mathbf{w}_n} \quad (4.59)$$

Assumindo que $\mathbf{R}$ seja conhecido, $\mathbf{t}$ é rapidamente determinado. Substituindo $\mathbf{t}$ por $\bar{\mathbf{p}} - \mathbf{R}\bar{\mathbf{q}}$ na definição de ϵ^2 , resulta em

$$\epsilon^2 = \sum_{n=1}^N \mathbf{w}_n \sum_{k=1}^3 [\mathbf{p}_{nk} - \bar{\mathbf{p}}_n - \mathbf{r}_k^t (\mathbf{q}_n - \bar{\mathbf{q}})]^2 + \sum_{k=1}^3 \lambda_k (\mathbf{r}_k^t \mathbf{r}_k - 1) + 2\lambda_4 \mathbf{r}_1^t \mathbf{r}_1 + 2\lambda_5 \mathbf{r}_1^t \mathbf{r}_3 + 2\lambda_6 \mathbf{r}_2^t \mathbf{r}_3 \quad (4.60)$$

onde

$$\bar{\mathbf{p}} = \begin{pmatrix} \bar{p}_1 \\ \bar{p}_2 \\ \bar{p}_3 \end{pmatrix} \quad \text{e} \quad \bar{\mathbf{q}} = \begin{pmatrix} \bar{q}_1 \\ \bar{q}_2 \\ \bar{q}_3 \end{pmatrix} \quad (4.61)$$

Agora tomam-se derivadas parciais de ϵ^2 com respeito às componentes de cada $\mathbf{r}_n$. Então,

$$\begin{aligned} \frac{\partial \epsilon^2}{\partial \mathbf{r}_1} &= \sum_{n=1}^N 2\mathbf{w}_n [\mathbf{p}_{n1} - \bar{\mathbf{p}}_1 - \mathbf{r}_1^t (\mathbf{q}_n - \bar{\mathbf{q}})] (\mathbf{q}_n - \bar{\mathbf{q}}) (-1) + 2\lambda_1 \mathbf{r}_1 + 2\lambda_4 \mathbf{r}_2 + 2\lambda_5 \mathbf{r}_3 \\ \frac{\partial \epsilon^2}{\partial \mathbf{r}_2} &= \sum_{n=1}^N 2\mathbf{w}_n [\mathbf{p}_{n2} - \bar{\mathbf{p}}_2 - \mathbf{r}_2^t (\mathbf{q}_n - \bar{\mathbf{q}})] (\mathbf{q}_n - \bar{\mathbf{q}}) (-1) + 2\lambda_2 \mathbf{r}_2 + 2\lambda_4 \mathbf{r}_1 + 2\lambda_6 \mathbf{r}_3 \\ \frac{\partial \epsilon^2}{\partial \mathbf{r}_3} &= \sum_{n=1}^N 2\mathbf{w}_n [\mathbf{p}_{n3} - \bar{\mathbf{p}}_3 - \mathbf{r}_3^t (\mathbf{q}_n - \bar{\mathbf{q}})] (\mathbf{q}_n - \bar{\mathbf{q}}) (-1) + 2\lambda_3 \mathbf{r}_3 + 2\lambda_5 \mathbf{r}_1 + 2\lambda_6 \mathbf{r}_2 \end{aligned} \quad (4.62)$$

Igualando estas derivadas a zero e rearranjando, obtém-se

$$\begin{aligned} \sum_{n=1}^N \mathbf{w}_n (\mathbf{q}_n - \bar{\mathbf{q}}) (\mathbf{q}_n - \bar{\mathbf{q}})^t \mathbf{r}_1 + \lambda_1 \mathbf{r}_1 + \lambda_4 \mathbf{r}_2 + \lambda_5 \mathbf{r}_3 &= \sum_{n=1}^N \mathbf{w}_n (\mathbf{p}_{n1} - \bar{\mathbf{p}}_1) (\mathbf{q}_n - \bar{\mathbf{q}}) \\ \sum_{n=1}^N \mathbf{w}_n (\mathbf{q}_n - \bar{\mathbf{q}}) (\mathbf{q}_n - \bar{\mathbf{q}})^t \mathbf{r}_2 + \lambda_4 \mathbf{r}_1 + \lambda_2 \mathbf{r}_2 + \lambda_6 \mathbf{r}_3 &= \sum_{n=1}^N \mathbf{w}_n (\mathbf{p}_{n2} - \bar{\mathbf{p}}_2) (\mathbf{q}_n - \bar{\mathbf{q}}) \\ \sum_{n=1}^N \mathbf{w}_n (\mathbf{q}_n - \bar{\mathbf{q}}) (\mathbf{q}_n - \bar{\mathbf{q}})^t \mathbf{r}_3 + \lambda_5 \mathbf{r}_1 + \lambda_6 \mathbf{r}_2 + \lambda_3 \mathbf{r}_3 &= \sum_{n=1}^N \mathbf{w}_n (\mathbf{p}_{n3} - \bar{\mathbf{p}}_3) (\mathbf{q}_n - \bar{\mathbf{q}}) \end{aligned} \quad (4.63)$$

Faz-se

$$\mathbf{a} = \sum_{n=1}^N (\mathbf{q}_n - \bar{\mathbf{q}})(\mathbf{q}_n - \bar{\mathbf{q}})^t \quad \mathbf{\Lambda} = \begin{pmatrix} \lambda_1 & \lambda_4 & \lambda_5 \\ \lambda_4 & \lambda_2 & \lambda_6 \\ \lambda_5 & \lambda_6 & \lambda_3 \end{pmatrix} \quad (4.64)$$

e

$$\mathbf{b} = \begin{pmatrix} b_1 & b_2 & b_3 \end{pmatrix}^t \quad \text{quando} \quad b_k = \sum_{n=1}^N \mathbf{w}_n (\mathbf{p}_{nk} - \bar{\mathbf{p}}_k)(\mathbf{q}_n - \bar{\mathbf{q}}) \quad (4.65)$$

a equação acima pode ser simplesmente escrita como

$$\mathbf{a}\mathbf{R}^t + \mathbf{R}^t\mathbf{\Lambda} = \mathbf{b}$$

Multiplicando ambos os lados da equação acima por $\mathbf{R}$, tem-se

$$\mathbf{R}\mathbf{a}\mathbf{R}^t + \mathbf{\Lambda} = \mathbf{R}\mathbf{b}$$

Como $\mathbf{a} = \mathbf{a}^t$, tem-se $(\mathbf{R}\mathbf{a}\mathbf{R}^t)^t = \mathbf{R}\mathbf{a}\mathbf{R}^t$. Agora, tanto $\mathbf{R}\mathbf{a}\mathbf{R}^t$ quanto $\mathbf{\Lambda}$ são simétricos, então $\mathbf{R}\mathbf{b}$ precisa ser também simétrico. Isto significa que

$$\mathbf{R}\mathbf{b} = (\mathbf{R}\mathbf{b})^t$$

A solução para $\mathbf{R}$ torna-se agora imediata. Seja a decomposição do valor singular [12] de $\mathbf{b}$ dada por

$$\mathbf{b} = \mathbf{U}\mathbf{S}\mathbf{V}$$

onde $\mathbf{U}$ e $\mathbf{V}$ são ortonormais e $\mathbf{S}$ é diagonal. Então,

$$\mathbf{R}\mathbf{U}\mathbf{S}\mathbf{V} = (\mathbf{U}\mathbf{S}\mathbf{V})^t\mathbf{R}^t = \mathbf{V}^t\mathbf{S}\mathbf{U}^t\mathbf{R}^t$$

Daí,

$$\mathbf{R} = \mathbf{V}^t\mathbf{U}^t \quad (4.66)$$

A matriz $\mathbf{R}$ é garantidamente ortonormal, mas não se pode afirmar que tenha determinante $+1$. Pode ter determinante -1 . Em este último caso, significa que $\mathbf{R}$ incorpora uma operação de reflexão além da equação de rotação. Tal resultado pode ocorrer se os pontos observados forem poucos em número e forem muito ruidosos.

4.6 Triangulação Estéreo

Discutir-se-á a seguir o procedimento de triangulação estéreo no qual é baseada a inferência da posição de um ponto 3D $(x, y, z)^t$ dadas as projeções em perspectiva $(u_L, v_L)^t$ e $(u_R, v_R)^t$ nas imagens esquerda e direita, respectivamente, de um par estéreo [8]. O procedimento de triangulação é um caso especial da determinação de um ponto 3D da intersecção de mais de dois raios. Assume-se que $(u_L, v_L)^t$ e $(u_R, v_R)^t$ são referenciados ao ponto principal e que qualquer distorção geométrica das lentes tem sido compensada. O procedimento de triangulação faz uso do paralaxe, que é o deslocamento na projeção em perspectiva de um ponto causada por uma variação translacional na posição de observação.

Para ilustrar este procedimento de triangulação, primeiro considera-se o caso no qual as posições das lentes das câmeras esquerda e direita estão, respectivamente, em

$$\begin{pmatrix} -b_x/2 \\ 0 \\ 0 \end{pmatrix} \quad \text{e} \quad \begin{pmatrix} b_x/2 \\ 0 \\ 0 \end{pmatrix}$$

Assume-se que o plano da imagem esteja a uma distância f da frente de cada câmera e que ambas as câmeras são orientadas identicamente, com o eixo x do referencial da câmera orientado ao longo da linha definida pela posição das lentes da câmera. Sendo $(x, y, z)^t$ um ponto 3D e sendo $(u_L, v_L)^t$ e $(u_R, v_R)^t$ sua projeção em perspectiva nas imagens esquerda e direita, respectivamente,

$$\begin{pmatrix} u_L \\ v_L \end{pmatrix} = \frac{f}{z} \begin{pmatrix} x + b_x/2 \\ y \end{pmatrix} \quad \text{e} \quad \begin{pmatrix} u_R \\ v_R \end{pmatrix} = \frac{f}{z} \begin{pmatrix} x - b_x/2 \\ y \end{pmatrix}$$

Nesta situação, $v_L = v_R$, tal que o paralaxe y seja zero.

A solução para $(x, y, z)^t$, dados $(u_L, v_L)^t$ e $(u_R, v_R)^t$, pode ser obtida da diferença $u_L - u_R$, que é chamada paralaxe x . Agora,

$$u_L - u_R = \frac{f}{z} \left[x + \frac{b_x}{L} - \left(x - \frac{b_x}{L} \right) \right] = \frac{f}{z} (b_x)$$

Portanto,

$$z = \frac{f b_x}{u_L - u_R} \quad (4.67)$$

Logo que a profundidade z seja determinada, as coordenadas $(x, y)^t$ são facilmente determinadas das equações de projeção em perspectiva:

$$\begin{aligned} \begin{pmatrix} x \\ y \end{pmatrix} &= \frac{z}{f} \begin{pmatrix} u_L \\ v_L \end{pmatrix} - \begin{pmatrix} b_x/2 \\ 0 \end{pmatrix} \\ &= \frac{z}{f} \begin{pmatrix} u_L \\ v_L \end{pmatrix} + \begin{pmatrix} b_x/2 \\ 0 \end{pmatrix} \end{aligned} \quad (4.68)$$

A relação $z = fb_x/(u_L - u_R)$ para determinar a profundidade do paralaxe x é uma clássica relação que em situações do mundo real é geralmente próxima à da usual, por três razões:

- As projeções em perspectiva observadas estão sujeitas a erros de medição, tal que $v_L \neq v_R$ para pontos correspondentes;
- os referenciais da câmera para as imagens esquerda e direita podem frequentemente ter orientações levemente diferentes;
- quando há duas diferentes câmeras que tomam as imagens esquerda e direita, é quase sempre o caso em que $f_L \neq f_R$.

Isto significa que mesmo se a relação $z = fb_x/(u_L - u_R)$ for generalizada para incluir os efeitos de $f_R \neq f_L$ e para usar u_L e u_R , onde eles agora representam coordenadas x depois da rotação das coordenadas imagem esquerda e direita medidas para corrigir as diferentes orientações dos referenciais das câmeras esquerda e direita, ainda seria um problema pois, com relação à medição e ao erro de orientação, $v_L \neq v_R$. Isto significa que o raio

$$L = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \middle| \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} b_x/2 \\ 0 \\ 0 \end{pmatrix} + \lambda_L \begin{pmatrix} u_L \\ v_L \\ f_L \end{pmatrix} \right\}$$

associado com o ponto $(u_L, v_L)^t$ da imagem esquerda, e o raio

$$R = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \middle| \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -b_x/2 \\ 0 \\ 0 \end{pmatrix} + \lambda_R \begin{pmatrix} u_R \\ v_R \\ f_R \end{pmatrix} \right\}$$

associados com o correspondente ponto $(u_R, v_R)^t$ da imagem direita não interseção no espaço 3D. Em este caso o melhor que pode-se fazer é encontrar este ponto 3D cuja distância a L e R seja minimizada.

No referencial comum que a orientação relativa tem sido computada no caso de ambas as imagens, faz-se $(x_L, y_L, z_L)^t$ ser a posição da lente da câmera da imagem esquerda, e $\mathbf{R}(\omega_L, \phi_L, \kappa_L)$ ser a rotação que gira o referencial da imagem esquerda no referencial comum. Similarmente, faz-se $(x_R, y_R, z_R)^t$ ser a posição da lente da câmera da imagem direita, e $\mathbf{R}(\omega_R, \phi_R, \kappa_R)$ ser a rotação que gira o referencial da imagem direita no referencial comum. Faz-se f_L e f_R serem as constantes da câmera associadas com as imagens esquerda e direita, respectivamente. Então, considerando-se que não haja ruído de medição, o ponto $(x, y, z)^t$ precisa permanecer na interseção dos dois raios:

$$\left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \middle| \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x_L \\ y_L \\ z_L \end{pmatrix} + \lambda_L \mathbf{Q}_L \begin{pmatrix} u_L \\ v_L \\ f_L \end{pmatrix} \right\} \quad (4.69)$$

e

$$\left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \middle| \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x_R \\ y_R \\ z_R \end{pmatrix} + \lambda_R \mathbf{Q}_R \begin{pmatrix} u_R \\ v_R \\ f_R \end{pmatrix} \right\} \quad (4.70)$$

Desde que haja algum ruído de medição, não espera-se que exista um λ_L e um λ_R de forma que

$$\begin{pmatrix} x_L \\ y_L \\ z_L \end{pmatrix} + \lambda_L \mathbf{Q}_L \begin{pmatrix} u_L \\ v_L \\ f_L \end{pmatrix} = \begin{pmatrix} x_R \\ y_R \\ z_R \end{pmatrix} + \lambda_R \mathbf{Q}_R \begin{pmatrix} u_R \\ v_R \\ f_R \end{pmatrix} \quad (4.71)$$

mantenha-se exatamente. Espera-se, contudo, que seja encontrado um λ_L e um λ_R tal que o erro quadrático ϵ^2 representando o quadrado do módulo da diferença entre os dois lados da equação anterior seja minimizado [8], ou seja,

$$\epsilon^2 = \left\| \begin{pmatrix} x_L \\ y_L \\ z_L \end{pmatrix} + \lambda_L \begin{pmatrix} p_L \\ q_L \\ s_L \end{pmatrix} - \begin{pmatrix} x_R \\ y_R \\ z_R \end{pmatrix} + \lambda_R \begin{pmatrix} p_R \\ q_R \\ s_R \end{pmatrix} \right\|^2$$

seja minimizado, onde

$$\begin{pmatrix} p_L \\ q_L \\ s_L \end{pmatrix} = \mathbf{Q}_L \begin{pmatrix} u_L \\ v_L \\ f_L \end{pmatrix} \quad \text{e} \quad \begin{pmatrix} p_R \\ q_R \\ s_R \end{pmatrix} = \mathbf{Q}_R \begin{pmatrix} u_R \\ v_R \\ f_R \end{pmatrix}$$

Para determinar-se os minimizantes λ_L e λ_R , tomam-se as derivadas parciais de ϵ^2 com respeito a λ_L e λ_R , que em seguida são igualadas a zero, e então resolvem-se os sistemas resultantes. Daí,

$$\begin{pmatrix} \lambda_L \\ \lambda_R \end{pmatrix} = \mathbf{A}_\lambda \mathbf{b}_\lambda \quad (4.72)$$

onde

$$\mathbf{A}_\lambda = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \quad \mathbf{b}_\lambda = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$

e

$$a_{11} = \frac{1}{k} \begin{pmatrix} p_R \\ q_R \\ s_R \end{pmatrix}^t \begin{pmatrix} p_R \\ q_R \\ s_R \end{pmatrix} \quad a_{22} = \frac{1}{k} \begin{pmatrix} p_L \\ q_L \\ s_L \end{pmatrix}^t \begin{pmatrix} p_L \\ q_L \\ s_L \end{pmatrix}$$

$$a_{12} = a_{21} = -\frac{1}{k} \begin{pmatrix} p_L \\ q_L \\ s_L \end{pmatrix}^t \begin{pmatrix} p_R \\ q_R \\ s_R \end{pmatrix}$$

$$b_1 = - \left[\begin{pmatrix} x_L \\ y_L \\ z_L \end{pmatrix} - \begin{pmatrix} x_R \\ y_R \\ z_R \end{pmatrix} \right]^t \begin{pmatrix} p_L \\ q_L \\ s_L \end{pmatrix} \quad b_2 = - \left[\begin{pmatrix} x_L \\ y_L \\ z_L \end{pmatrix} - \begin{pmatrix} x_R \\ y_R \\ z_R \end{pmatrix} \right]^t \begin{pmatrix} p_R \\ q_R \\ s_R \end{pmatrix}$$

$$k = \left[\begin{pmatrix} p_R \\ q_R \\ s_R \end{pmatrix}^t \begin{pmatrix} p_R \\ q_R \\ s_R \end{pmatrix} \right] \left[\begin{pmatrix} p_L \\ q_L \\ s_L \end{pmatrix}^t \begin{pmatrix} p_L \\ q_L \\ s_L \end{pmatrix} \right] - \left[\begin{pmatrix} p_L \\ q_L \\ s_L \end{pmatrix}^t \begin{pmatrix} p_R \\ q_R \\ s_R \end{pmatrix} \right]^2$$

Uma outra forma de obter-se estes minimizantes é feita a partir de (4.71) [9].

Expandindo esta igualdade,

$$\begin{pmatrix} x_L \\ y_L \\ z_L \end{pmatrix} + \lambda_L \begin{pmatrix} r_{11}^L & r_{21}^L & r_{31}^L \\ r_{12}^L & r_{22}^L & r_{32}^L \\ r_{13}^L & r_{23}^L & r_{33}^L \end{pmatrix} \begin{pmatrix} u_L \\ v_L \\ f_L \end{pmatrix} = \begin{pmatrix} x_R \\ y_R \\ z_R \end{pmatrix} + \lambda_R \begin{pmatrix} r_{11}^R & r_{21}^R & r_{31}^R \\ r_{12}^R & r_{22}^R & r_{32}^R \\ r_{13}^R & r_{23}^R & r_{33}^R \end{pmatrix} \begin{pmatrix} u_R \\ v_R \\ f_R \end{pmatrix} \quad (4.73)$$

onde r_{ij}^L são os elementos da matriz de rotação da imagem esquerda, e r_{ij}^R são os elementos da matriz de rotação da imagem direita.

Sejam as expressões

$$\begin{aligned} U_L &= r_{11}^L u_L + r_{21}^L v_L + r_{31}^L f_L & U_R &= r_{11}^R u_R + r_{21}^R v_R + r_{31}^R f_R \\ V_L &= r_{12}^L u_L + r_{22}^L v_L + r_{32}^L f_L & V_R &= r_{12}^R u_R + r_{22}^R v_R + r_{32}^R f_R \\ W_L &= r_{13}^L u_L + r_{23}^L v_L + r_{33}^L f_L & W_R &= r_{13}^R u_R + r_{23}^R v_R + r_{33}^R f_R \end{aligned}$$

Então, simplifica-se a equação 4.73:

$$\begin{aligned} x_L + \lambda_L U_L &= x_R + \lambda_R U_R \\ y_L + \lambda_L V_L &= y_R + \lambda_R V_R \\ z_L + \lambda_L W_L &= z_R + \lambda_R W_R \end{aligned}$$

Resolvendo, obtêm-se as seguintes equações:

$$\lambda_L = \frac{(x_L - x_R)V_R - (y_L - y_R)U_R}{V_L U_R - U_R V_L} \quad \text{e} \quad \lambda_R = \frac{(x_L - x_R)V_L - (y_L - y_R)U_L}{V_L U_R - U_R V_L} \quad (4.74)$$

O cálculo de λ_L e λ_R pela Eq. 4.72 revela-se mais exato, por envolver a minimização da diferença entre os dois lados da Equ. 4.71. Já o cálculo pela Eq. 4.74 parte do princípio de que os dois lados desta equação sejam iguais em valor, o que pode conduzir a erros maiores na obtenção das coordenadas tridimensionais. Por isso, será utilizada, para fins de cálculo de λ_L e λ_R , a Eq. 4.72, apenas.

Finalmente, a coordenada do ponto tridimensional é obtida a partir ou da Equ. 4.69 ou da Equ. 4.70. Idealmente, os resultados obtidos por estas duas equações deveriam ser iguais, mas podem haver pequenas diferenças, pois o cálculo de λ_L e λ_R não é exato.

4.7 Modelagem dos Erros Sistemáticos

O modelo descrito anteriormente para o cálculo da coordenada tridimensional representa uma situação ideal, onde não existiriam erros sistemáticos, como a distorção das lentes. Como isto não ocorre, torna-se necessária uma modificação, com a finalidade de se introduzir parâmetros que modelem erros sistemáticos. Estes parâmetros são obtidos no processo de determinação da orientação interior.

As coordenadas imagem $(u_L, v_L)^t$ e $(u_R, v_R)^t$, que representam as projeções em perspectiva do ponto $(x, y, z)^t$ nas imagens esquerda e direita, respectivamente, precisam ser alteradas para incluir os a correção dos erros sistemáticos.

As novas coordenadas imagem $(u_c, v_c)^t$, obtidas a partir das coordenadas imagem medidas $(u, v)^t$ são dadas por:

$$\begin{pmatrix} u_c \\ v_c \end{pmatrix} = (1 + \delta_r) \begin{pmatrix} u - u_p \\ v - v_p \end{pmatrix} \quad (4.75)$$

onde

$$\delta_r = k_1 r^2 + k_2 r^4 + k_3 r^6$$

representa a distorção geométrica radial, e r é dado pela Eq. 4.25. Já k_1 , k_2 , k_3 , u_p e v_p são obtidos durante a determinação da orientação interior.

Finalmente, estas novas coordenadas imagem substituem na Eq. 4.71 as coordenadas imagem medidas, e a partir daí o cálculo da coordenada tridimensional é efetuado como já visto na seção anterior.

Capítulo 5

Aplicação

Este capítulo ilustra uma aplicação do sistema de dimensionamento tratado neste estudo, no dimensionamento de peças metálicas submarinas. Para tal foram realizados alguns testes com imagens de peças aquisitadas em um tanque de prova simulando a condição real de uso. Os testes consistem no cálculo de coordenadas tridimensionais, e o posterior cálculo de dimensões de arestas das peças metálicas utilizadas. Os resultados obtidos serão apresentados, bem como o erro obtido nas medidas. A título de ilustração, também serão apresentados os resultados obtidos do cálculo de parâmetros na solução dos problemas de orientação exterior e relativa.

5.1 Imagens Aquisitadas

As imagens a serem utilizadas para a execução dos testes com os algoritmos desenvolvidos para o dimensionamento de objetos 3D (vide capítulo anterior) são inicialmente aquisitadas com o uso do sistema de visão estéreo apresentado no capítulo 2. Os testes foram feitos em um tanque de prova cheio d'água, no qual foram submersas algumas peças metálicas retangulares, em posições variadas. Foram coletadas imagens várias pelo par de câmeras CCD, e para o presente estudo foram selecionados quatro pares de imagens estéreo.

As imagens selecionadas, que são mostradas na Fig. 5.1 originalmente possuem resolução de 640×512 pixels; contudo, são apresentadas em escala reduzida, e em

tons de cinza. Para cada par estéreo, respectivamente são mostradas as imagens esquerda e direita. A peça metálica visualizada na Fig. 5.1(b) possui dimensões (em milímetros) de $59,10 \times 80,10 \times 19,92$, enquanto as outras peças metálicas possuem, cada, dimensões de $10 \times 2 \times 3$.

5.2 Pré-processamento

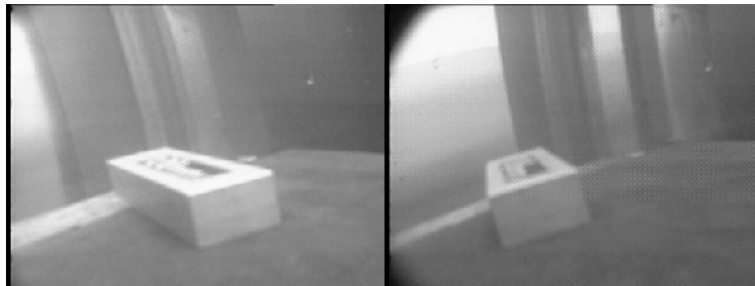
O pré-processamento, como já visto no capítulo 3, consiste em, através de técnicas de processamento de imagens, obter uma melhora da qualidade da imagem, de forma a adequá-la ao processo de dimensionamento tridimensional. No caso das imagens aquisitionadas para os testes, será necessária a aplicação de técnicas de detecção de envoltórias, que visam a obter uma maior nitidez na visualização das bordas das peças metálicas utilizadas nos testes, de forma a obter-se uma maior acuracidade na aquisição de coordenadas bidimensionais das imagens.

Dentre os vários operadores descritos no capítulo 3, os operadores de gradiente apresentaram os melhores resultados, sendo por isso os utilizados na obtenção das envoltórias das imagens aquisitionadas. Os operadores de direção também apresentaram bons resultados na detecção de envoltórias em determinada direção, o que pode ser útil em certos casos, quando se desejar realçar os limites em uma direção apenas.

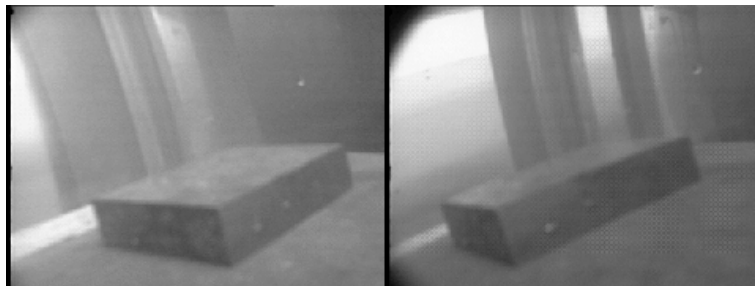
Os resultados da aplicação de operadores de gradiente (no caso, usando a máscara Sobel) às imagens da Fig. 5.1, e os respectivos mapas de envoltória (em negativo), são visualizados respectivamente nas Fig. 5.2 e 5.3.

5.3 Parâmetros de Orientação Exterior

Para o cálculo das coordenadas tridimensionais usando os algoritmos de fotogrametria apresentados no capítulo anterior, é necessária antes a solução dos problemas de orientação exterior. A orientação exterior representa a posição da origem do referencial imagem e sua orientação em relação ao mundo real, e consiste na obtenção de seis parâmetros: três parâmetros de translação e três



(a) Par estéreo 1.



(b) Par estéreo 2.

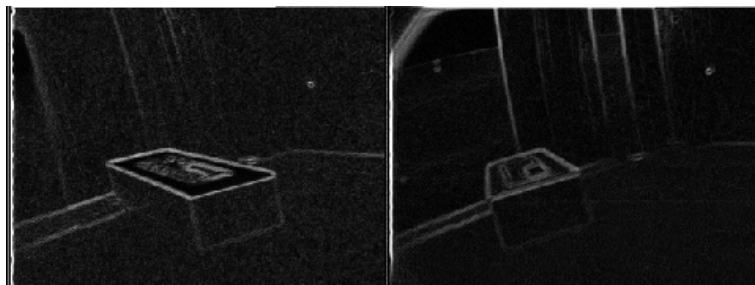


(c) Par estéreo 3.

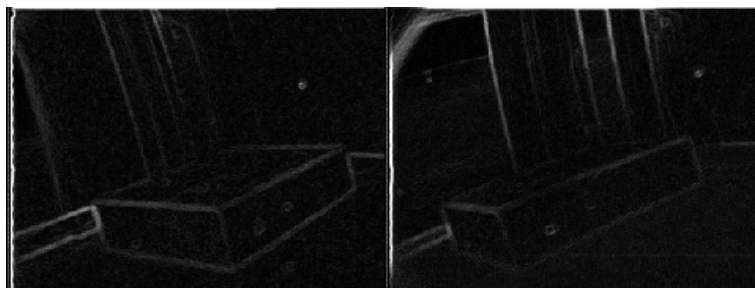


(d) Par estéreo 4.

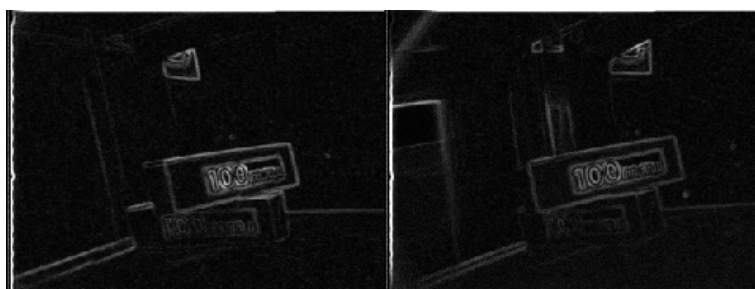
Figura 5.1: Imagens adquiridas para teste.



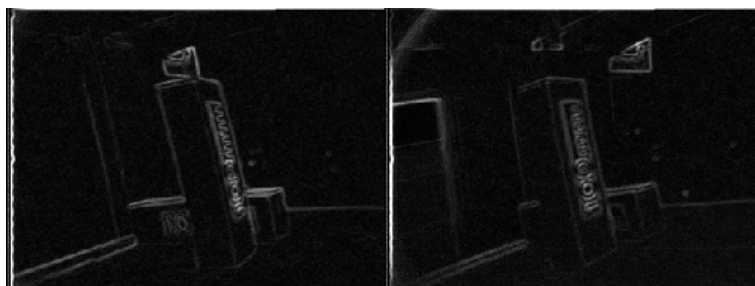
(a) Par estéreo 1.



(b) Par estéreo 2.

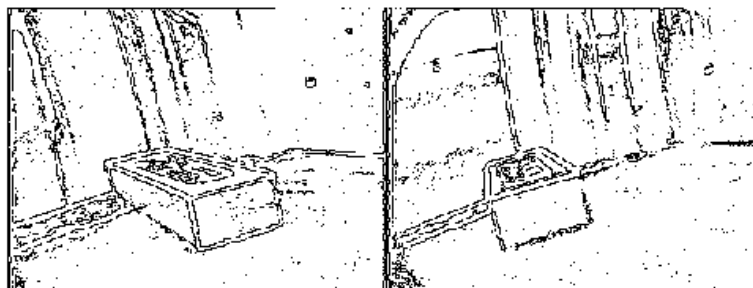


(c) Par estéreo 3.

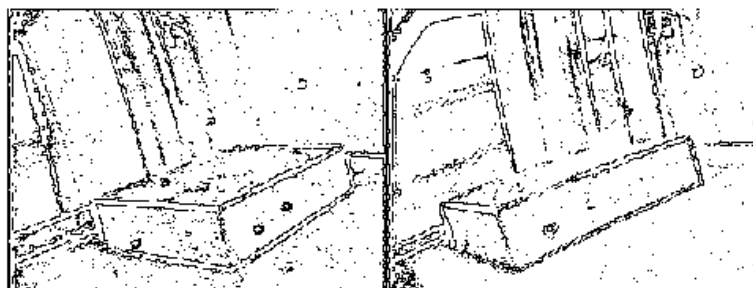


(d) Par estéreo 4.

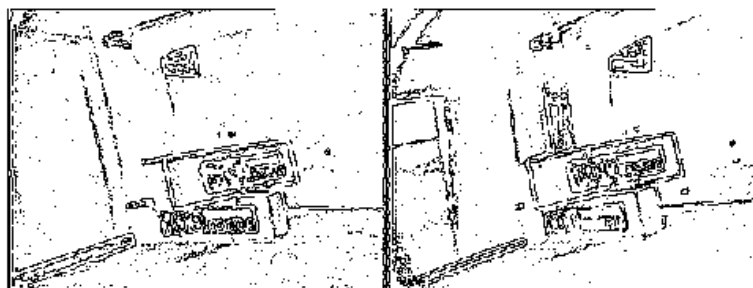
Figura 5.2: Detecção de envoltória aplicada às imagens da Fig. 5.1.



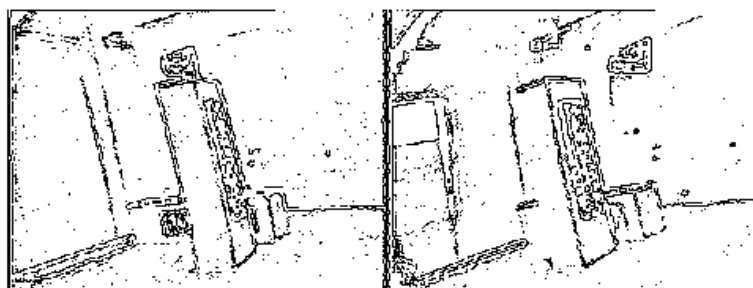
(a) Par estéreo 1.



(b) Par estéreo 2.



(c) Par estéreo 3.



(d) Par estéreo 4.

Figura 5.3: Mapa de envoltória obtido a partir das imagens da figura anterior.

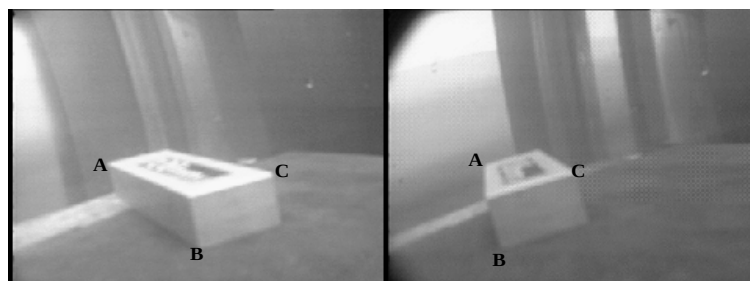
ângulos de rotação. Para o cálculo destes parâmetros, são inicialmente obtidos de cada imagem do par estéreo um conjunto de coordenadas bidimensionais de pontos correspondentes no referencial da imagem, e o conjunto correspondente de coordenadas tridimensionais no referencial do objeto.

As coordenadas bidimensionais, em pixels, são obtidas diretamente de cada imagem (esquerda e direita), sendo considerada a origem o canto superior esquerdo da imagem. Já as coordenadas tridimensionais são coordenadas cartesianas de um sistema tridimensional arbitrário no espaço tridimensional.

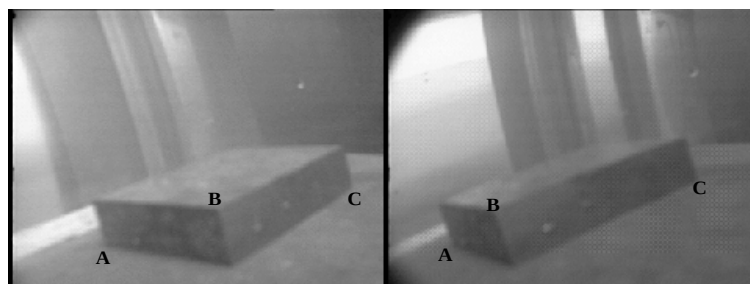
Tabela 5.1: Coordenadas dos pontos usados na determinação dos parâmetros de orientação exterior para cada imagem.

Par Estéreo	x	y	z	u_L	v_L	u_R	v_R
1	0	20	205	168	281	172	283
	0	0	105	327	427	200	426
	30	20	105	455	300	329	295
2	0	0	189,10	169	424	122	425
	0	19,92	130	360	358	189	360
	80,10	0	130	589	319	537	311
3	0	0	180	205	352	222	353
	80	30	140	266	271	239	274
	100	-30	200	478	393	470	386
4	0	30	180	206	352	225	354
	70	0	130	322	455	283	455
	90	100	160	358	142	330	137

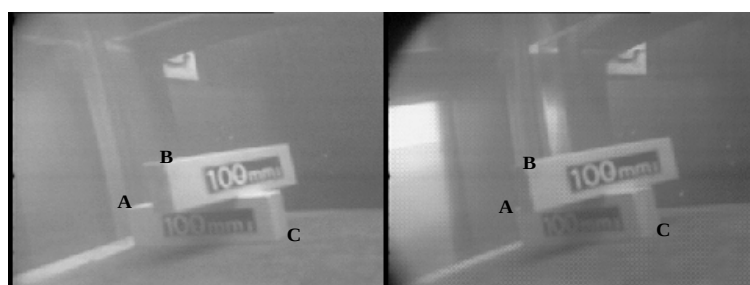
Os três pontos usados no processo de obtenção dos parâmetros de orientação exterior são indicados na Fig. 5.4 (são os pontos rotulados A , B e C). A Tab. 5.3 mostra as coordenadas bidimensionais destes pontos (u_L e v_L para a imagem esquerda, e u_R e v_R para a imagem direita), expressas em pixels, bem como as coordenadas tridimensionais correspondentes (x , y , z), expressas em milímetros.



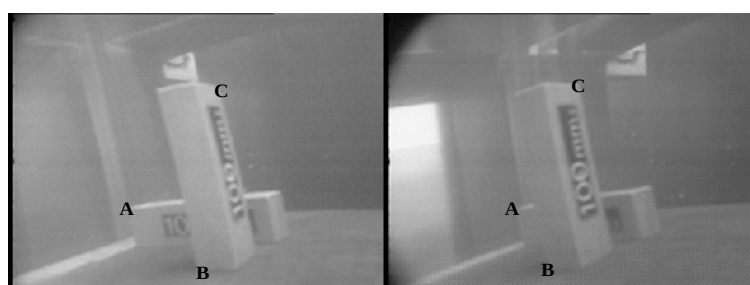
(a) Par estéreo 1.



(b) Par estéreo 2.



(c) Par estéreo 3.



(d) Par estéreo 4.

Figura 5.4: Pontos usados na determinação da orientação exterior.

Tabela 5.2: Parâmetros de orientação exterior calculados.

Par Estéreo	Parâmetros					
	x_0	y_0	z_0	ω	ϕ	κ
1 (L)	-13,499	16,620	-52,669	0,202	0,220	-0,313
(R)	3,911	10,152	-53,880	0,113	-0,012	-0,319
2 (L)	-565,717	248,463	-359,247	0,387	0,817	-0,526
(R)	-270,562	138,735	-599,791	0,143	0,397	-0,427
3 (L)	63,503	3,060	-39,485	-0,004	-0,781	-0,189
(R)	71,461	1,617	-31,038	-0,031	-0,920	-0,264
4 (L)	468,192	126,207	-358,137	0,110	-0,655	-0,174
(R)	546,831	79,340	-274,672	-6,255	-0,826	-6,494

A partir destas coordenadas, e com o auxílio das fórmulas apresentadas na seção (4.2), são obtidos os parâmetros de orientação exterior, para cada imagem de cada par estéreo, e que são mostrados na Tab. 5.2. Os parâmetros de translação x_0 , y_0 e z_0 são expressos em milímetros, enquanto os parâmetros de rotação são ângulos, expressos em radianos. Na coluna “**Par Estéreo**”, (L) indica que os parâmetros são da imagem esquerda, enquanto (R) indica que os parâmetros são da imagem direita.

5.4 Parâmetros de Orientação Relativa

A orientação relativa, conforme visto no capítulo 4, diz respeito à determinação da posição e orientação de uma imagem com relação à outra, dado um conjunto de pontos correspondentes. A determinação da orientação relativa de um par de câmeras requer a determinação de cinco parâmetros: dois parâmetros de translação ($y_R - y_L$ e $z_R - z_L$) e três ângulos de rotação ($\omega_R - \omega_L$, $\phi_R - \phi_L$, $\kappa_R - \kappa_L$). Como a orientação relativa é o referencial da imagem direita com respeito ao referencial da imagem esquerda, tem-se y_L , z_L , ω_L , ϕ_L e κ_L nulos. Então, é necessária

apenas a determinação dos parâmetros y_L , z_L , ω_L , ϕ_L e κ_L . Como a separação entre as câmeras ao longo do eixo x é considerada uma constante que controla a escala, não constitui um parâmetro da orientação relativa. Esta constante, no caso do par de câmeras usadas neste sistema, tem o valor de $x_R = 5,5\text{cm}$.

Para a determinação da orientação exterior, são necessários dois conjuntos de cinco pontos-imagem correspondentes, ou seja, pontos do plano bidimensional de cada imagem do par estéreo: $\{(u_{Ln}, v_{Ln}), n = 1, \dots, 5\}$ para a imagem esquerda, e $\{(u_{Rn}, v_{Rn}), n = 1, \dots, 5\}$ para a imagem direita. Estes pontos são indicados na Fig. 5.5 pelos vértices rotulados A , B , C , D , e E , e suas coordenadas (em pixels) mostradas na Tab. 5.3. Feitos os cálculos com as fórmulas da seção (4.4), obtêm-se para cada par estéreo os cinco parâmetros, que são mostrados na Tab. 5.4.

5.5 Aplicação

Serão mostrados, nesta seção, os resultados dos testes realizados nas imagens aquiritadas. Os resultados de diversas medidas feitas de arestas das peças metálicas das imagens são mostrados nas tabelas 5.5 a 5.8, sendo comparadas as dimensões reais e calculadas de cada aresta, bem como o erro obtido (absoluto e relativo) em cada medida. Os pontos utilizados em cada imagem para os cálculos são vértices das peças metálicas, e são indicados nas figuras 5.6 a 5.9.

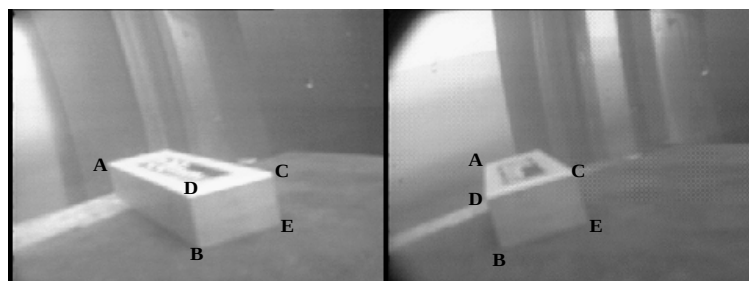
A medida de cada aresta foi feita a partir das coordenadas tridimensionais dos vértices que delimitam a aresta. Por exemplo, no cálculo da dimensão da aresta $\overline{\mathbf{AB}}$, inicialmente foram calculadas as coordenadas dos vértices $\mathbf{A}$ e $\mathbf{B}$ a partir das fórmulas das seções (4.6) e (4.7), e então foi calculada a distância entre as coordenadas através de uma fórmula de geometria analítica:

$$\overline{\mathbf{AB}} = \sqrt{(X_A - X_B)^2 + (Y_A - Y_B)^2 + (Z_A - Z_B)^2}$$

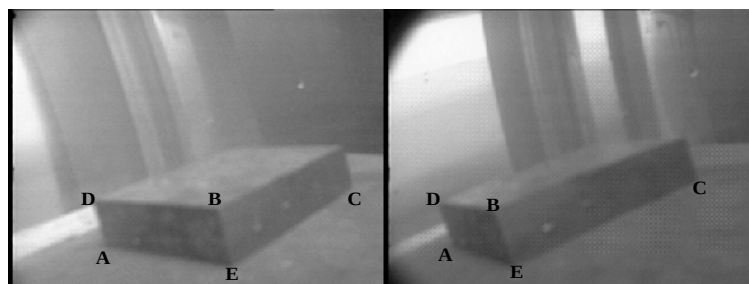
onde $\mathbf{A} = (X_A, Y_A, Z_A)^t$ e $\mathbf{B} = (X_B, Y_B, Z_B)^t$ representam as coordenadas de cada vértice, e $\overline{\mathbf{AB}}$ a distância entre as coordenadas.

Tabela 5.3: Coordenadas dos pontos usados na determinação dos parâmetros de orientação relativa para cada par estéreo.

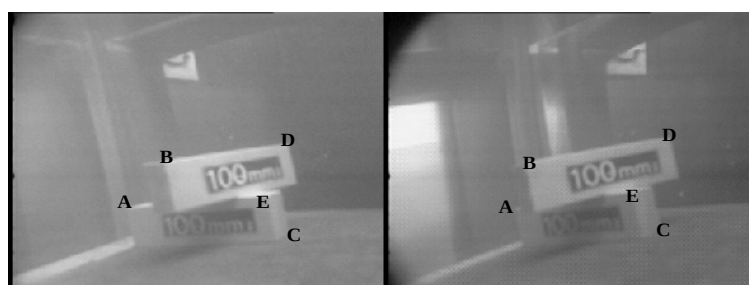
Par Estéreo	u_L	v_L	u_R	v_R
1	168	281	172	283
	327	427	200	426
	455	300	329	295
	309	340	174	343
	468	384	350	380
2	169	424	122	425
	360	358	189	360
	589	319	537	311
	146	344	95	349
	384	458	225	457
3	205	352	222	353
	266	271	239	274
	478	393	470	386
	443	333	423	326
	476	241	486	231
4	206	352	225	354
	322	455	283	455
	358	142	330	137
	252	145	223	151
	466	323	456	316



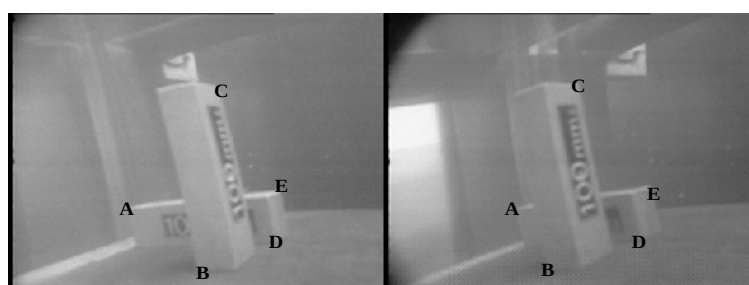
(a) Par estéreo 1.



(b) Par estéreo 2.



(c) Par estéreo 3.



(d) Par estéreo 4.

Figura 5.5: Pontos usados na determinação da orientação relativa.

Tabela 5.4: Parâmetros de orientação relativa calculados.

Par Estéreo	Parâmetros				
	y_R	z_R	ω_R	ϕ_R	κ_R
1	3.461	19.554	0.001	-0.652	-0.064
2	1.434	-47.138	-0.376	-1.415	-0.402
3	0.565	-24.492	-0.006	-0.764	-0.105
4	3.946	-21.477	-0.018	-0.757	-0.099

Tabela 5.5: Resultados dos testes – par estéreo 1.

Aresta	Dimensão		Erro Obtido	
	Calculada	Real	Absoluto	Relativo
$\overline{AB}$	101,980	101,980	0,000	0,000 %
$\overline{AC}$	104,403	104,403	0,000	0,000 %
$\overline{AD}$	100,789	100,000	0,789	0,789 %
$\overline{AE}$	104,622	106,301	0,168	1,580 %
$\overline{BC}$	101,980	101,980	0,000	0,000 %
$\overline{BD}$	19,345	20,000	0,655	3,273 %
$\overline{BE}$	27,621	30,000	2,379	7,931 %
$\overline{CD}$	28,282	30,000	1,718	5,728 %
$\overline{CE}$	19,296	20,000	0,704	3,520 %
$\overline{DE}$	31,676	36,056	4,380	12,146 %

Tabela 5.6: Resultados dos testes – par estéreo 2.

Aresta	Dimensão		Erro Obtido	
	Calculada	Real	Absoluto	Relativo
$\overline{AB}$	62,367	62,367	0,000	0,000 %
$\overline{AC}$	99,543	99,543	0,000	0,000 %
$\overline{AD}$	22,261	19,920	2,341	11,754 %
$\overline{AE}$	60,521	59,100	1,421	2,404 %
$\overline{AF}$	103,147	101,517	1,630	1,606 %
$\overline{BC}$	82,540	82,540	0,000	0,000 %
$\overline{BD}$	61,782	59,100	2,682	4,538 %
$\overline{BF}$	83,886	80,100	3,786	4,726 %
$\overline{CD}$	100,849	101,517	0,668	0,658 %
$\overline{CE}$	83,295	80,100	3,195	3,989 %
$\overline{CF}$	20,221	19,920	0,301	1,510 %
$\overline{DE}$	69,095	62,367	6,728	10,788 %
$\overline{DF}$	100,009	99,543	0,466	0,468 %
$\overline{EF}$	90,756	82,540	8,216	9,954 %

Tabela 5.7: Resultados dos testes – par estéreo 3.

Aresta	Dimensão		Erro Obtido	
	Calculada	Real	Absoluto	Relativo
$\overline{AC}$	106,301	106,301	0,000	0,000 %
$\overline{AE}$	106,712	100,000	0,671	6,711 %
$\overline{AF}$	28,695	30,000	1,305	4,348 %
$\overline{AG}$	98,078	101,980	3,902	3,826 %
$\overline{AH}$	112,459	104,403	8,056	7,716 %
$\overline{BJ}$	35,575	36,056	0,480	1,332 %
$\overline{BK}$	109,953	104,403	5,550	5,316 %
$\overline{CF}$	98,119	101,980	3,861	3,786 %
$\overline{CG}$	29,562	30,000	0,438	1,460 %
$\overline{CH}$	20,678	20,000	0,678	3,391 %
$\overline{DK}$	33,629	30,000	3,629	12,097 %
$\overline{EF}$	104,562	100,440	0,159	0,152 %
$\overline{EH}$	28,735	30,000	1,265	1,212 %
$\overline{IJ}$	21,214	20,000	1,214	6,072 %
$\overline{JK}$	106,648	101,980	4,668	4,577 %

Tabela 5.8: Resultados dos testes – par estéreo 4.

Aresta	Dimensão		Erro Obtido	
	Calculada	Real	Absoluto	Relativo
$\overline{AE}$	92,773	101,980	9,207	9,028 %
$\overline{AF}$	30,377	30,000	0,377	1,257 %
$\overline{AG}$	98,078	101,980	3,902	3,826 %
$\overline{BC}$	106,301	106,301	0,000	0,000 %
$\overline{BI}$	21,496	20,000	1,496	7,480 %
$\overline{BK}$	106,457	101,980	4,477	4,390 %
$\overline{BL}$	104,579	100,000	4,579	4,579 %
$\overline{CH}$	102,132	100,000	2,132	2,132 %
$\overline{CI}$	105,117	104,403	0,713	0,683 %
$\overline{DE}$	29,751	36,056	6,304	17,485 %
$\overline{DF}$	31,922	30,000	1,922	6,407 %
$\overline{DG}$	100,516	100,000	0,516	0,516 %
$\overline{EF}$	20,922	20,000	0,922	4,610 %
$\overline{EG}$	108,152	106,301	1,851	1,741 %
$\overline{EJ}$	27,999	30,000	2,001	6,671 %
$\overline{GJ}$	96,913	101,980	5,068	4,969 %
$\overline{HI}$	27,628	30,000	2,372	7,906 %
$\overline{HK}$	106,896	104,403	2,493	2,388 %
$\overline{HL}$	107,145	106,301	0,843	0,793 %
$\overline{KL}$	23,206	20,000	3,206	16,031 %

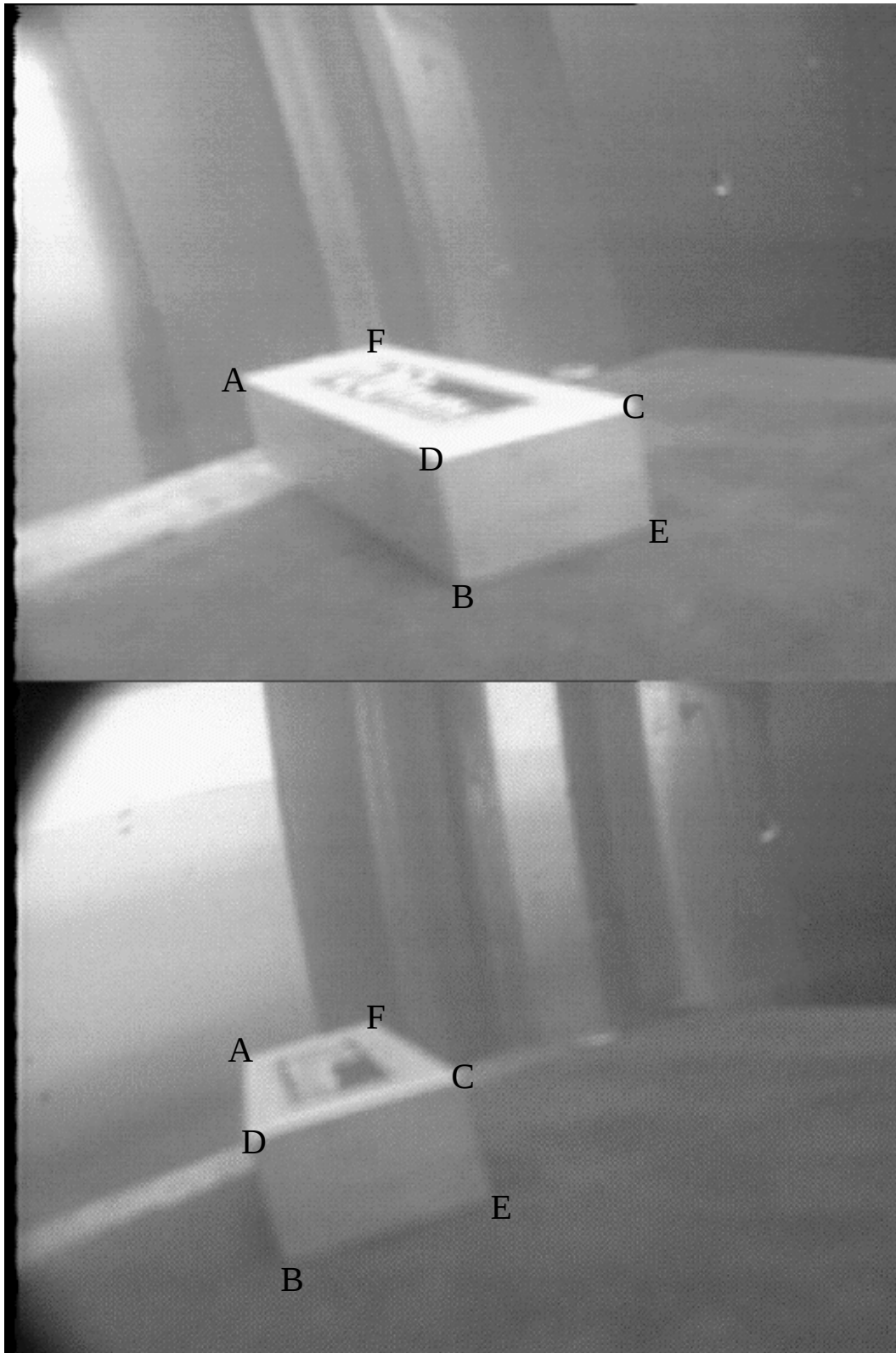


Figura 5.6: Pontos usados nos testes com o par estéreo 1.

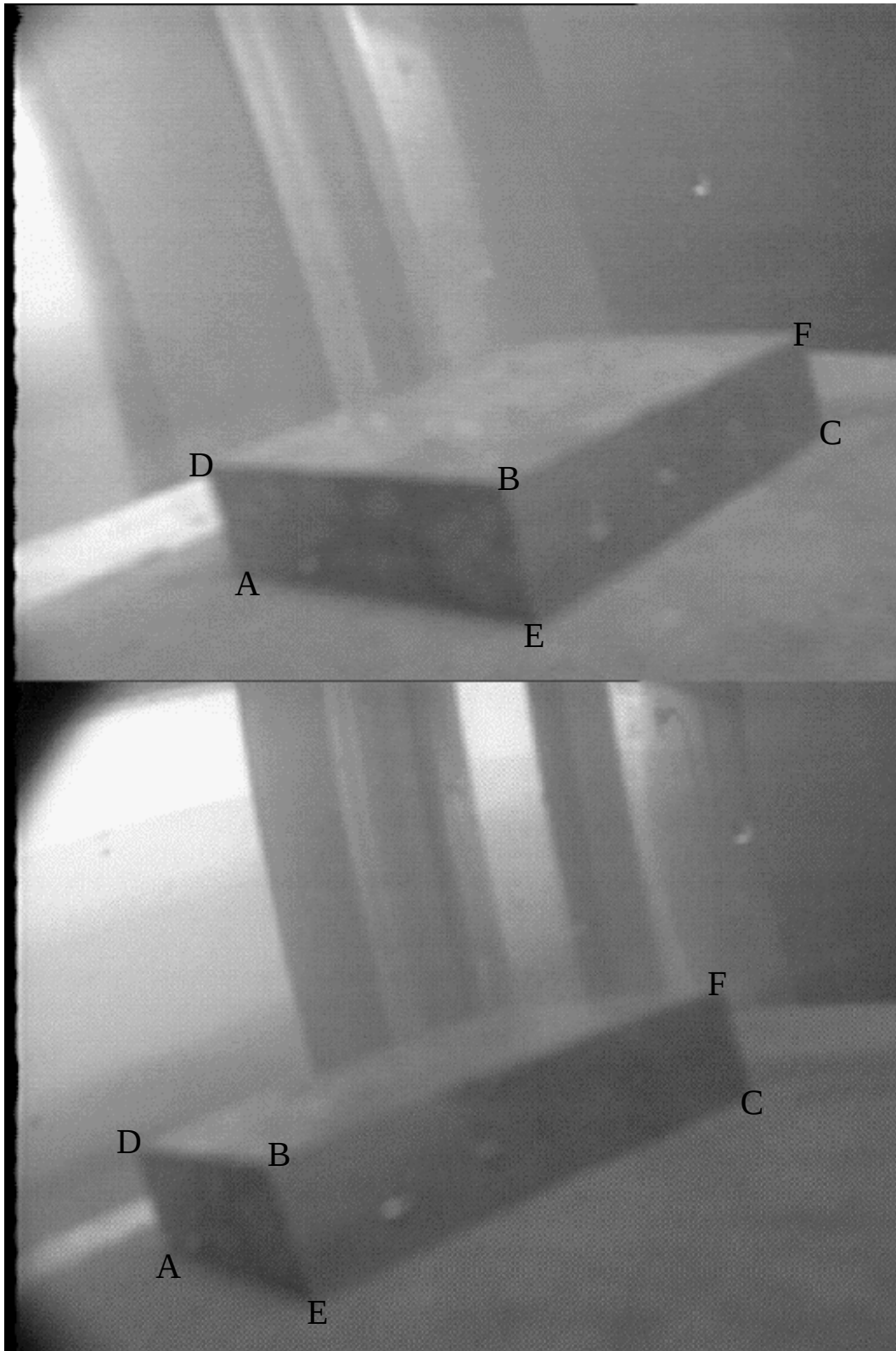


Figura 5.7: Pontos usados nos testes com o par estéreo 2.

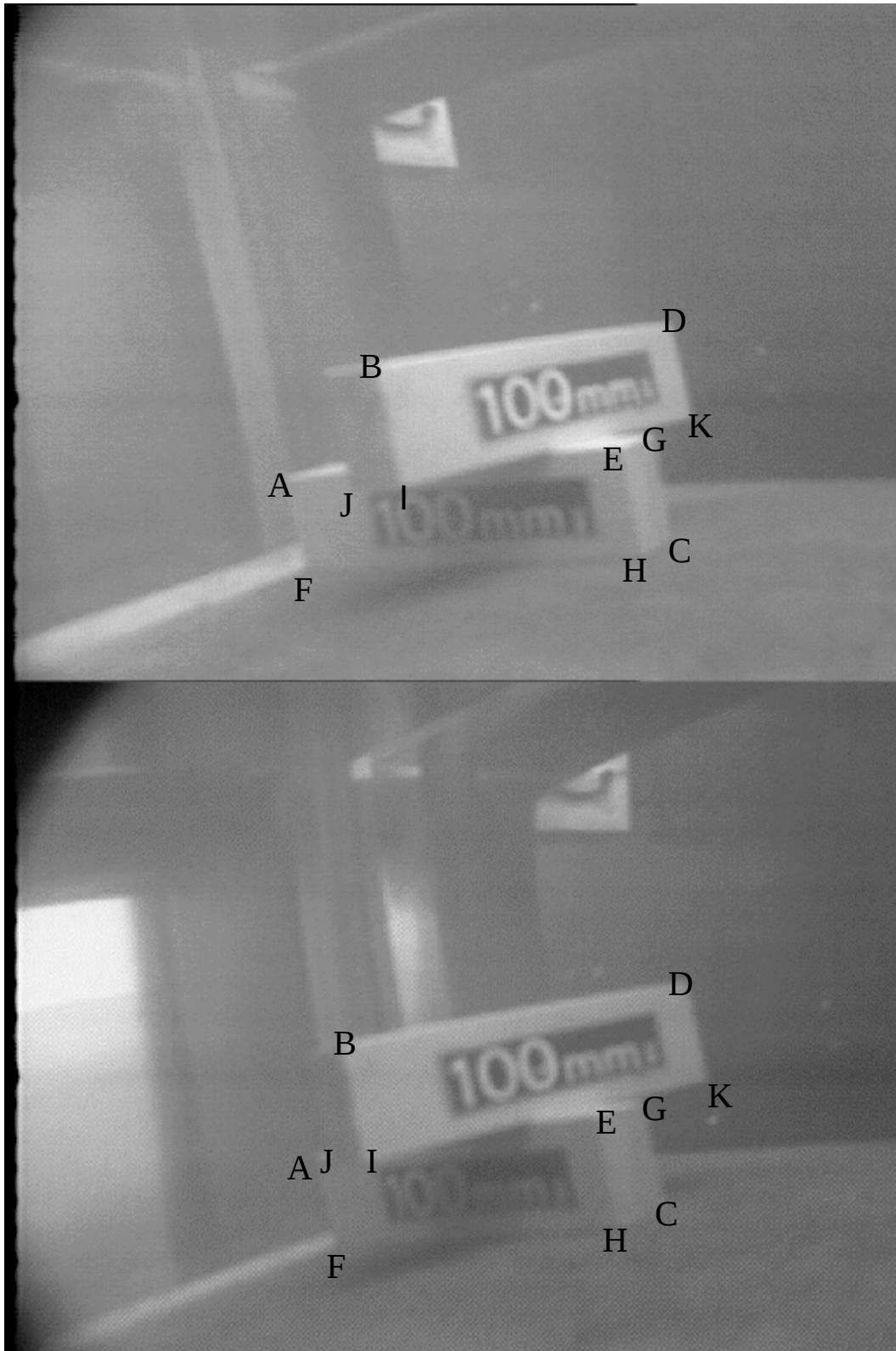


Figura 5.8: Pontos usados nos testes com o par estéreo 3.

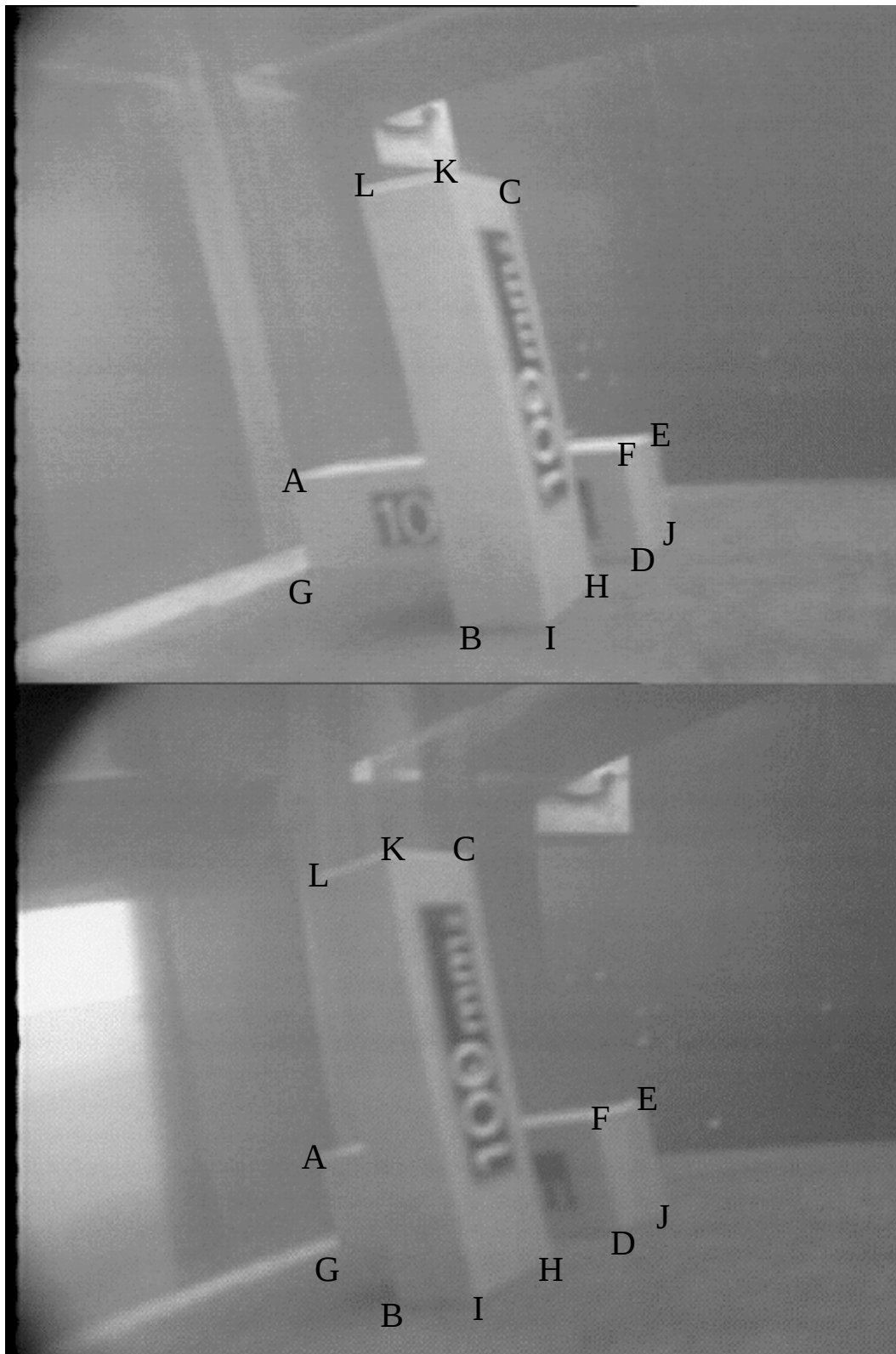


Figura 5.9: Pontos usados nos testes com o par estéreo 4.

Capítulo 6

Conclusão

O presente estudo abordou, nos capítulos anteriores, um sistema de medição tridimensional a ser utilizado em atividades de inspeção submarina de tubos e equipamentos de petróleo. Como visto no capítulo 1, o sistema divide-se em três blocos: visão estéreo, pré-processamento, e dimensionamento tridimensional.

O sistema de visão estéreo utilizado permitiu a visualização das imagens em tempo real, com noção de profundidade, através de um *software* de interface amigável.

O pré-processamento, através de técnicas de processamento de imagens, tem o objetivo de efetuar uma melhoria na qualidade da imagem, de forma a adequá-la ao processo de dimensionamento. A técnica de detecção de envoltórias aplicada às imagens aquisitadas obteve uma melhoria na visualização dos limites dos objetos presentes na imagem, permitindo uma melhor definição na obtenção dos pontos necessários ao processo de dimensionamento. Para a detecção de envoltórias, foram utilizados operadores de gradiente, que apresentaram melhores resultados na definição dos limites.

O dimensionamento tridimensional permitiu a obtenção de coordenadas tridimensionais de pontos no espaço 3D, a partir de coordenadas bidimensionais aquisitadas do plano de cada imagem de um par estéreo, através de um programa desenvolvido em linguagem C, com o uso de algoritmos de fotogrametria analítica. A partir destas coordenadas tridimensionais foram feitas (para efeito de teste do algoritmo desenvolvido) medições de distâncias entre pontos nas imagens (no caso

dos testes, medições das dimensões de arestas de peças metálicas presentes nas imagens), sendo comparadas as dimensões reais de cada aresta com as dimensões calculadas, bem como o erro obtido em cada medição.

Os erros obtidos nos testes, em geral foram pequenos, na média ficando entre 3% e 6 % com relação à dimensão real de cada aresta. Os erros maiores se devem em parte ao fato de que as imagens terem sido adquiridas uma única vez em laboratório do CENPES/Petrobrás sem que tivesse havido o objetivo de procurar imagens com iluminação adequada. Na operação em tempo real, contudo, o veículo remoto ao qual serão acopladas as câmeras transportará também um sistema de iluminação, de forma que será possível procurar o ângulo mais adequado ao dimensionamento do objeto de interesse.

O cálculo das coordenadas apresenta ainda uma certa sensibilidade à variação de alguns pixels no valor das coordenadas bidimensionais. A escolha dos pontos necessários ao processo de determinação da orientação exterior também pode influenciar no cálculo das coordenadas. O dimensionamento de profundidades em imagens essencialmente planas, por exemplo, pode ocasionar erros maiores.

Apêndice A

Soluções *Least-Squares* Não-Lineares

Para a obtenção dos diversos parâmetros envolvidos nas soluções dos problemas de orientação exterior, interior e relativa apresentados no capítulo 4, utiliza-se uma técnica *least-squares* não-linear [8]. Para resolver o problema *least-squares* não-linear, inicia-se com uma solução aproximada dada inicialmente, e então linearizam-se iterativamente as funções não-lineares em volta da solução aproximada corrente, e enfim resolvendo-se o problema linearizado para os ajustes da solução corrente.

Supõe-se que $\beta_1, \dots, \beta_M$ são os parâmetros desconhecidos controlando cada uma das transformações não-lineares $g_1, \dots, g_k$ e que $\alpha_1, \dots, \alpha_k$ são os valores observados de $g_1, \dots, g_k$. O modelo ruidoso é dado por

$$\alpha_k = g_K(\beta_1, \dots, \beta_M) + \xi_k, \quad k = 1, \dots, K \quad (\text{A.1})$$

onde $\xi_1, \dots, \xi_k$ são variáveis aleatórias gaussianas aditivas com média nula tendo matriz de covariância $\mathcal{R}$. Sobre estas assunções, deve-se minimizar o critério *least-squares* dado por

$$\epsilon^2 = (\boldsymbol{\alpha} - \mathbf{g})^t \mathcal{R} (\boldsymbol{\alpha} - \mathbf{g}) \quad (\text{A.2})$$

onde

$$\boldsymbol{\alpha} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_k \end{pmatrix} \quad \text{e} \quad \mathbf{g} = \begin{pmatrix} g_1(\beta_1, \dots, \beta_M) \\ g_2(\beta_1, \dots, \beta_M) \\ \vdots \\ g_k(\beta_1, \dots, \beta_M) \end{pmatrix}$$

Se há uma solução aproximada dada $\boldsymbol{\beta}^0 = (\beta_1^0, \dots, \beta_M^0)^t$, começa-se linearizando as transformações não-lineares em volta de $\boldsymbol{\beta}^0$ e resolve-se para obter os ajustes $\Delta\boldsymbol{\beta} = (\Delta\beta_1, \dots, \Delta\beta_M)^t$, que quando adicionados a $\boldsymbol{\beta}^l$ (solução aproximada corrente $\boldsymbol{\beta}$ após l iterações) constituem uma solução aproximada melhor.

Na l -ésima iteração $\boldsymbol{\beta}^l = (\beta_1^l, \dots, \beta_M^l)^t$ é a solução linear corrente. A linearização procede representando-se cada $g_k(\boldsymbol{\beta}^l + \Delta\boldsymbol{\beta})$ por uma expansão de série de Taylor de 1ª ordem de g_k em volta de $\boldsymbol{\beta}^l$:

$$g_k(\boldsymbol{\beta}^l + \Delta\boldsymbol{\beta}) = g_k(\boldsymbol{\beta}^l) + \Delta g_k(\Delta\boldsymbol{\beta}; \boldsymbol{\beta}^l) \quad (\text{A.3})$$

onde Δg_k , a derivada total de g_k , é uma função linear do vetor de ajustes $\Delta\boldsymbol{\beta}$ dado por

$$\Delta g_k(\Delta\boldsymbol{\beta}; \boldsymbol{\beta}^l) = \left(\frac{\partial g_k}{\partial \beta_1}(\boldsymbol{\beta}^l) \dots \frac{\partial g_k}{\partial \beta_M}(\boldsymbol{\beta}^l) \right) \Delta\boldsymbol{\beta}$$

A derivada total Δg_k tem uma interpretação direta da Eq. A.2. Se o parâmetro $\boldsymbol{\beta}$ é alterado por uma importância $\Delta\boldsymbol{\beta}$, então o valor resultante de g_k será alterado por um valor de Δg_k . Substituindo as expressões linearizadas no critério *least-squares*, obtém-se:

$$\epsilon^2 = (\boldsymbol{\alpha} - \mathbf{g}^l - \mathbf{G}^l \Delta\boldsymbol{\beta})^t \mathcal{R} (\boldsymbol{\alpha} - \mathbf{g}^l - \mathbf{G}^l \Delta\boldsymbol{\beta})$$

onde

$$\mathbf{g}^l = \begin{pmatrix} g_1(\beta_1^l, \dots, \beta_M^l) \\ g_2(\beta_1^l, \dots, \beta_M^l) \\ \vdots \\ g_k(\beta_1^l, \dots, \beta_M^l) \end{pmatrix}$$

e $\mathbf{G}^l$ é o jacobiano, dado por

$$\mathbf{G}^l = \begin{pmatrix} \frac{\partial g_1}{\partial \beta_1} & \frac{\partial g_1}{\partial \beta_2} & \cdots & \frac{\partial g_1}{\partial \beta_M} \\ \frac{\partial g_2}{\partial \beta_1} & \frac{\partial g_2}{\partial \beta_2} & \cdots & \frac{\partial g_2}{\partial \beta_M} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g_k}{\partial \beta_1} & \frac{\partial g_k}{\partial \beta_2} & \cdots & \frac{\partial g_k}{\partial \beta_M} \end{pmatrix}$$

Nota-se que o problema *least-squares* não linear tornou-se um problema *least-squares* linear nos ajustes de $\Delta\boldsymbol{\beta}$. Para resolvê-lo, tomam-se derivadas parciais de ϵ^2 com respeito a cada $\Delta\beta_m$ e zera-se cada derivada parcial. Assim, sendo

$$\frac{\partial \epsilon^2}{\partial \Delta\boldsymbol{\beta}} = 0 = -2\mathbf{G}^{lt}\mathcal{R}^{-1}(\alpha - \mathbf{g}^l - \mathbf{G}^l\Delta\boldsymbol{\beta})$$

resulta que

$$\Delta\boldsymbol{\beta} = (\mathbf{G}^{tl}\mathcal{R}\mathbf{G}^l)^{-1}\mathbf{G}^{lt}(\alpha - \mathbf{g}^l) \quad (\text{A.4})$$

e a equação de atualização de $\boldsymbol{\beta}$ é dada por

$$\boldsymbol{\beta}^{l+1} = \boldsymbol{\beta}^l + \Delta\boldsymbol{\beta}. \quad (\text{A.5})$$

Apêndice B

A Decomposição do Valor Singular

A decomposição do valor singular (SVD – *Singular Value Decomposition*) é uma poderosa ferramenta computacional para analisar matrizes e problemas envolvendo matrizes que tem aplicações em diferentes campos [12, 13, 14].

Teorema 1 (Decomposição do Valor Singular) *Seja $\mathbf{A}$ uma matriz real $m \times n$. Existem matrizes ortogonais*

$$\mathbf{U} = [u_1, \dots, u_m] \in \mathbb{R}^{m \times n} \quad e \quad \mathbf{V} = [v_1, \dots, v_m] \in \mathbb{R}^{n \times n}$$

tal que

$$\mathbf{U}^t \mathbf{A} \mathbf{V} = \text{diag}(\sigma_1, \dots, \sigma_p) \in \mathbb{R}^{m \times n} \quad p = \min\{m, n\}$$

onde $\sigma_1 \geq \sigma_2 \geq \dots \sigma_p \geq 0$.

Prova. Faz-se $x \in \mathbb{R}^n$ e $y \in \mathbb{R}^m$ serem vetores normais unitários que satisfaçam $\mathbf{A}x = \sigma y$ com $\sigma = \|\mathbf{A}\|_2$. Devido ao fato de qualquer conjunto ortonormal poder ser estendido para formar uma base ortonormal para o espaço inteiro, é possível encontrar $\mathbf{V}_1 \in \mathbb{R}^{n \times (n-1)}$ e $\mathbf{U}_1 \in \mathbb{R}^{m \times (m-1)}$ tal que $\mathbf{V} = [x \quad \mathbf{V}_1] \in \mathbb{R}^{n \times n}$ e $\mathbf{U} = [y \quad \mathbf{U}_1] \in \mathbb{R}^{m \times m}$ são ortonormais. Não é difícil mostrar que $\mathbf{U}^t \mathbf{A} \mathbf{V}$ tem a seguinte estrutura:

$$\mathbf{U}^t \mathbf{A} \mathbf{V} = \begin{bmatrix} \sigma & \omega' \\ 0 & \mathbf{B} \end{bmatrix} \equiv \mathbf{A}_1.$$

Desde que

$$\left\| \mathbf{A}_1 \begin{pmatrix} \sigma \\ \omega \end{pmatrix} \right\|_2^2 \geq (\sigma^2 + \omega' \omega)^2$$

tem-se $\|\mathbf{A}_1\|_2^2 \geq (\sigma^2 + \omega' \omega)$. Mas $\sigma^2 = \|\mathbf{A}\|_2^2 = \|\mathbf{A}_1\|_2^2$, e assim precisa-se ter $\omega = 0$.

Um óbvio argumento de indução completa a prova do teorema. $\square$

Os σ_i são os valores singulares de $\mathbf{A}$ e os vetores u_i e v_i são o i -ésimo vetor singular esquerdo e o i -ésimo vetor singular direito, respectivamente. É fácil verificar comparando-se as colunas nas equações $\mathbf{A}\mathbf{V} = \Sigma\mathbf{U}$ e $\mathbf{A}^t\mathbf{U} = \Sigma^t\mathbf{V}$ que

$$\left. \begin{array}{l} \mathbf{A}v_i = \sigma_i u_i \\ \mathbf{A}^t u_i = \sigma_i v_i \end{array} \right\} i = 1 : \min\{m, n\}$$

É conveniente ter a seguinte notação para designar os valores singulares:

$$\begin{aligned} \sigma_1(\mathbf{A}) &= \text{o } i\text{-ésimo valor singular de } \mathbf{A}, \\ \sigma_{\max}(\mathbf{A}) &= \text{o maior valor singular de } \mathbf{A}, \\ \sigma_{\min}(\mathbf{A}) &= \text{o menor valor singular de } \mathbf{A}. \end{aligned}$$

Os valores singulares de $\mathbf{A}$ são precisamente os comprimentos dos semi-eixos do hiperelipsóide E definido por $E = \{\mathbf{A}x : \|x\|_2 = 1\}$.

Bibliografia

- [1] PETRAGLIA, A., PETRAGLIA, M. R., *Sistema de Aquisição e Processamento de Sinais para Ensaios não Destrutivos*. COPPETEC, 1995.
- [2] TURNER, J., YULE, D. J., ZANRE, J., “Real Time Photogrammetry – A Technique for Today or Tomorrow?”. *Subtech'91*, vol. 27, *Society for Underwater Techonlogy*, pp. 319-331, 1991.
- [3] HICKS, R. S., GUNN, M. J., “Industrial Photogrammetry – Its Application to Shipbuilding”. *Marine Technology*, vol. 25, no. 3, pp. 229-236, Julho 1988.
- [4] HENRIQUES, P. C. F., SPERANDIO, A. G., BORGES, J. L. P., et al, *Sistema de Medição Tridimensional para Aplicações Submarinas – Sistema TV3D Submarino sem Cintilação (TV3D System Flickerless)*. 1º Relatório Parcial, CENPES, 1994.
- [5] FAUGERAS, O., *Three-Dimensional Computer Vision: a Geometric Viewpoint*. The MIT Press, 1993.
- [6] JAIN, K. A., *Fundamentals of Digital Image Processing*. Prentice-Hall Inc., 1989.
- [7] MENDES Fº, J. R., BONILAURI, L.M.C., GUIMARÃES, M.R.C., *Laboratório de Processamento Digital de Imagens*. Projeto Final, DEL/UFRJ, 1989.
- [8] HARALICK, R. M., SHAPIRO, L. G., *Computer and Robot Vision - Volume II*. Addison-Wesley Publishing Co., 1993.

-
- [9] TOMMASELLI, A. M. G., *Intersecção Fotogramétrica - Desenvolvimento de Funções em C*. UNESP, 1995.
- [10] MOFFITT, F. H., MIKHAIL, E. M., *Photogrammetry*. 12^a ed. Har & Row, Publishers, 1980.
- [11] LU, C. P., MJOLSNESS, E., HAGER, G. D., "Online Computation of Exterior Orientation with Application to Hand-Eye Calibration". *Mathl. Comput. Modelling*, Vol. 24, n. 5/6, pp. 121-143, 1996.
- [12] GOLUB, G. H., VAN LOAN, C. F., *Matrix Computations*. 2^a ed. The Johns Hopkins University Press, 1989.
- [13] GOLUB, G. H., REINSCH, C. "Singular Value Decomposition and Least Squares Solutions". In: *Numerische Mathematik, vol. 14, Handbook Series Linear Algebra*, pp. 403-20, 1970.
- [14] BUSINGER, P. A., GOLUB, G. H., "Algorithm 358: Singular Value Decomposition of a Complex Matrix". In: *Communications of the Association of Computers and Machines, vol. 12*, pp. 564-65, 1969.