



COPPE/UFRJ

ESTIMAÇÃO DE MOVIMENTO DE UMA CÂMERA VIA GEOMETRIA
EPIPOLAR

Guilherme de Souza Aires

Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Engenharia Elétrica, COPPE, da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Mestre em Engenharia Elétrica.

Orientador: José Gabriel Rodriguez Carneiro
Gomes

Rio de Janeiro
Março de 2010

ESTIMAÇÃO DE MOVIMENTO DE UMA CÂMERA VIA GEOMETRIA
EPIPOLAR

Guilherme de Souza Aires

DISSERTAÇÃO SUBMETIDA AO CORPO DOCENTE DO INSTITUTO
ALBERTO LUIZ COIMBRA DE PÓS-GRADUAÇÃO E PESQUISA DE
ENGENHARIA (COPPE) DA UNIVERSIDADE FEDERAL DO RIO DE
JANEIRO COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A
OBTENÇÃO DO GRAU DE MESTRE EM CIÊNCIAS EM ENGENHARIA
ELÉTRICA.

Examinada por:

Prof. Antônio Petraglia, Ph.D.

Prof. José Gabriel Rodriguez Carneiro Gomes, Ph.D.

Prof. Raul Queiroz Feitosa, Dr.-Ing.

RIO DE JANEIRO, RJ – BRASIL

MARÇO DE 2010

Aires, Guilherme de Souza

Estimação de Movimento de uma Câmera via Geometria Epipolar/Guilherme de Souza Aires. – Rio de Janeiro: UFRJ/COPPE, 2010.

IX, 99 p. 29, 7cm.

Orientador: José Gabriel Rodriguez Carneiro Gomes

Dissertação (mestrado) – UFRJ/COPPE/Programa de Engenharia Elétrica, 2010.

Referências Bibliográficas: p. 82 – 84.

1. Movimento de Uma Câmera. 2. Geometria Epipolar. 3. Algoritmo de 8 Pontos. I. Gomes, José Gabriel Rodriguez Carneiro. II. Universidade Federal do Rio de Janeiro, COPPE, Programa de Engenharia Elétrica. III. Título.

*Às quatro mulheres de
minha vida:
Anette, Joyce, Jayne e Gisela,
que fazem de mim um homem
cada dia melhor.*

Agradecimentos

Agradeço a Deus, em primeiro lugar, pois com Ele tudo é possível.

Agradeço a toda minha família (irmão, pais, avós e tios), que sempre me deram todo suporte e incentivo.

Agradeço a minha amada Gisela, pelo seu carinho e força em todos os momentos.

Agradeço ao prof. José Gabriel Rodriguez Carneiro Gomes, pela sua amizade, conselhos e incentivos muito importantes no meu crescimento pessoal e profissional.

Agradeço ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo suporte financeiro.

Agradeço ao Laboratório de Processamento de Sinais Analógicos e Digitais (PADS), assim como todos alunos e professores, sendo verdadeiro sinônimo de uma grande família.

Agradeço ao doutorando Bruno Madeira e ao seu orientador prof. Luiz Velho, pela sua atenção e auxílio na pesquisa.

Agradeço ao pesquisador de pós-doutorado Gian Luca Mariottini, pela sua atenção via e-mail no uso da *toolbox* de Matlab (EGT).

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

ESTIMAÇÃO DE MOVIMENTO DE UMA CÂMERA VIA GEOMETRIA EPIPOLAR

Guilherme de Souza Aires

Março/2010

Orientador: José Gabriel Rodriguez Carneiro Gomes

Programa: Engenharia Elétrica

A estimação de movimento de uma câmera via correspondência entre imagens é uma das principais tarefas na área de Visão Computacional. A recuperação da posição e orientação de uma imagem relativa a outra pode ser usada em várias aplicações como visão robótica e realidade aumentada. A relação projetiva entre duas imagens é descrita pela Geometria Epipolar. Faremos um estudo dessa geometria e de suas técnicas lineares, como por exemplo, o Algoritmo de 8 Pontos, a fim de que possamos extrair os parâmetros do movimento de uma câmera. Analisaremos a eficiência e robustez dos algoritmos de modo a verificarmos suas viabilidades de uso. Concluimos que através de uma normalização dos dados de entrada do Algoritmo de 8 Pontos, podemos utilizá-lo como uma técnica para uma estimação inicial do movimento.

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

SINGLE CAMERA MOTION ESTIMATION THROUGH EPIPOLAR GEOMETRY

Guilherme de Souza Aires

March/2010

Advisor: José Gabriel Rodriguez Carneiro Gomes

Department: Electrical Engineering

A single camera motion estimation from image correspondence is one of the central tasks in Computer Vision. The recovery of the position and the orientation of one image relative to another can be used for many applications like robot vision and augmented reality. The projective relation between two views is described by Epipolar Geometry. We study this geometry and its linear techniques, like the 8-Point Algorithm, aiming at extracting the camera motion parameters. The algorithms' robustness and efficiency will be analysed, as well as its viability of use. We conclude that using the 8-point Algorithm with normalized input data is feasible for being used as an initial motion estimation technique.

Sumário

1	Introdução	1
1.1	Objetivo	1
1.2	Geometria Projetiva	2
1.3	Modelagem Matemática	4
1.3.1	Espaço-n Projetivo	4
1.3.2	Plano Projetivo	4
1.3.3	Pontos e Retas em \mathbb{P}^2	4
1.3.4	Transformação Projetiva no Plano	5
1.3.5	Espaço-3 Projetivo	6
1.3.6	Pontos e Planos em \mathbb{P}^3	6
1.3.7	Transformação Projetiva no Espaço-3	7
1.3.8	Transformação Projetiva do Espaço-3 no Plano	7
1.4	Estrutura da Dissertação	8
2	Modelagem de Câmeras	9
2.1	Modelo <i>Pin-hole</i>	9
2.1.1	Modelagem Física e Geométrica	10
2.2	Modelo do Sensor Óptico	14
2.3	Modelagem Matemática	14
2.3.1	Parâmetros Intrínsecos	15
2.3.2	Parâmetros Extrínsecos	17
2.3.3	Calibração	19
3	Geometria Epipolar	24
3.1	Definições	24
3.1.1	Elementos da Geometria Epipolar	25
3.2	Restrição Epipolar	26
3.3	Matriz Essencial	27
3.3.1	Propriedades da Matriz Essencial	29
3.4	Matriz Fundamental	32
3.4.1	Propriedades da Matriz Fundamental	33

3.5	Estimação das Matrizes Essencial e Fundamental	35
3.5.1	Algoritmo de 8 Pontos	35
3.5.2	Algoritmo de 8 Pontos Normalizado	39
4	Estimação de Movimento	46
4.1	Movimento de Translação Pura	46
4.2	Movimento de Translação com Rotação Conhecida	48
4.3	Movimento de Translação e Rotação	50
4.4	Cálculo do Fator de Escala do Vetor \mathbf{t}	54
5	Algoritmos, Testes e Resultados	57
5.1	Algoritmos	57
5.2	Testes e Resultados	59
5.2.1	Dados Sintéticos	59
5.2.2	Dados Reais	74
6	Conclusões	80
	Referências Bibliográficas	82
A	Provas de Teoremas, Fórmulas e Equações	85
B	Decomposição em Valores Singulares (SVD)	93
B.1	Definição	94
B.2	Propriedades	94
B.3	Otimização	95
B.3.1	Otimização Linear por Mínimos Quadrados ($\mathbf{Ax} = \mathbf{b}$)	95
B.3.2	Otimização Linear por Mínimos Quadrados ($\mathbf{Ax} = \mathbf{0}$)	96
C	Otimização com Restrição	97

Capítulo 1

Introdução

A estimação de movimento tem sido estudada em Visão Computacional desde o princípio desta ciência. É uma área crucial para várias aplicações nas áreas de processamento de imagens, computação gráfica e robótica. Por exemplo, o fluxo óptico, a incorporação de elementos gráficos em realidade virtual e a localização e navegação de um robô móvel são algumas típicas aplicações que necessitam dessa área de pesquisa.

Como muitas áreas do conhecimento no campo da Engenharia, a estimação de movimento em Visão Computacional procura em fenômenos da natureza a motivação para implementar sistemas baseados nesses conceitos. De acordo com este raciocínio, o fenômeno que usamos para nossa pesquisa foi a visão monocular humana. Brady [1] explica como pacientes com visão monocular podem obter novamente a visão 3D. Ele sugere que esses pacientes façam movimentos relativos de suas cabeças quando estiverem observando um objeto, logo, através dessas imagens em diferentes referenciais processadas pelo cérebro, é formada uma noção de 3D do cenário.

Embasados nessa idéia, veremos o conceito de Geometria Epipolar, que corresponde à matemática processada instintivamente pelo cérebro humano.

1.1 Objetivo

O objetivo do nosso trabalho consiste em estudarmos a Geometria Epipolar e suas peculiaridades. A partir desse estudo, analisaremos como, fazendo o uso de algoritmos lineares, podemos obter os parâmetros de rotação e translação do movimento de uma câmera. Analisaremos essas técnicas lineares de estimação em função da precisão de suas estimativas para verificarmos as suas limitações e a sua utilização para diferentes tipos de movimentos.

1.2 Geometria Projetiva

De modo que possamos entender a matemática por detrás das imagens capturadas por uma câmera, faremos inicialmente um estudo da Geometria Projetiva [2] e de sua origem.

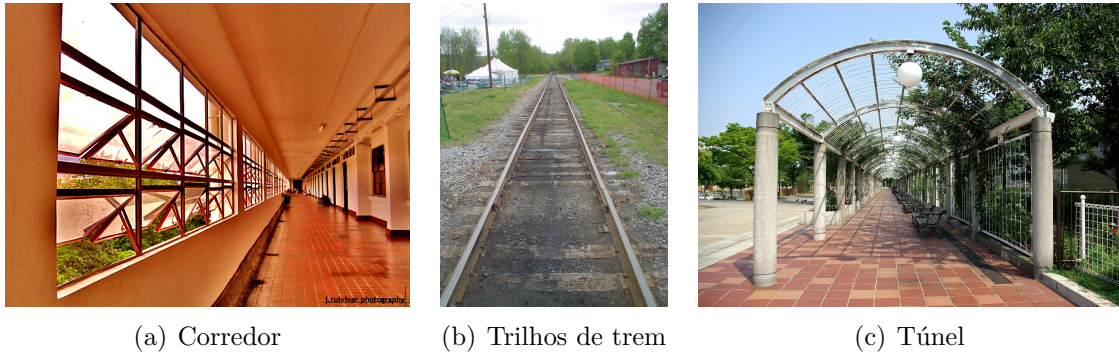


Figura 1.1: Exemplos de imagens capturadas por uma câmera (fonte: wikipedia)

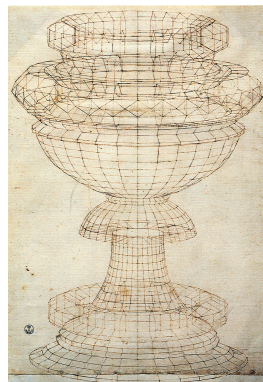
A geometria estudada nos cursos secundários, a que chamamos de Geometria Euclidiana, tem por principal objetivo cuidar das propriedades de áreas, comprimentos, ângulos e volumes das figuras e dos corpos geométricos. É, por assim dizer, uma geometria da medida, de acordo, aliás, com a sua origem entre os egípcios e com o sentido de seu próprio nome: medida (*metria*) da terra (*geo*).

Contudo, a geometria das imagens capturadas por câmeras não obedece tais propriedades. Facilmente podemos constatar isso, vendo as imagens na Figura 1.1. Os três exemplos correspondem a imagens compostas por retas paralelas no espaço euclidiano, ou seja, retas que não são concorrentes. Mas o que podemos verificar é que elas se encontram em um ponto nas imagens.

Observando esse fato, o primeiro estudo que abordou essa questão foi o estudo geométrico da perspectiva. Esse estudo é conhecido desde o período da Antiguidade no Egito e na Grécia, contudo, no período do Renascimento, teve seu auge [3]. Isso foi graças aos trabalhos de inúmeros artistas da época que começaram a utilizar-se das noções de perspectiva em suas artes. Podem ser vistos alguns exemplos na Figura 1.2.



(a) Pintura de Pietro Perugino (1481-82)
(fonte: wikipedia)



(b) Desenho de Paolo Uccello (séc. XV)
(fonte: wikipedia)

Figura 1.2: Exemplos do estudo da perspectiva na arte no período do Renascimento

Na Figura 1.3, podemos observar como funcionam as regras da perspectiva. As paralelas se encontram nos pontos pretos chamados de *pontos de fuga*, que por sua vez, estão todos contidos na *reta de fuga* (ou *reta do horizonte*). Podemos dizer que estas foram as bases matemáticas para a criação da Geometria Projetiva.

Regras principais da perspectiva [3]

1. a imagem de uma reta também corresponde a uma reta.
2. a imagem de retas paralelas são retas concorrentes.

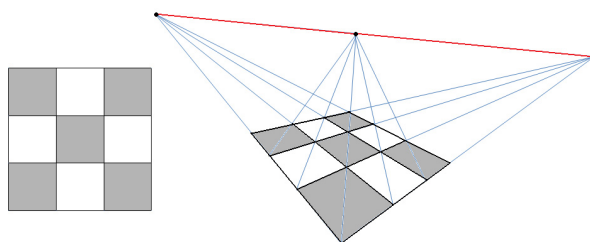


Figura 1.3: Objeto real e imagem em perspectiva

No século XIX, começou a surgir uma tendência nova: a idéia de que essa geometria seria independente do conceito de medida criando-se então, uma nova geometria. Através de Desargues, Monge e Carnot, o estudo da geometria projetiva teve seu início, mas quem a desenvolveu formalmente pela primeira vez, foi Jean Victor Poncelet. Ele fez um estudo sistemático das propriedades das figuras sem tomar a medida como base, nem como objetivo final. No ano de 1822, Poncelet evidenciou as propriedades das figuras que se mantinham inalteradas através das projeções e seções.

Para mais facilmente percebermos a distinção entre a Geometria Euclidiana e a Geometria Projetiva, convém que comecemos por considerar as duas espécies diferentes de operações (ou transformações) que aparecem predominantemente em cada uma delas, partindo dos conceitos já conhecidos da Geometria Euclidiana: as operações fundamentais da Geometria Euclidiana são os deslocamentos ou movimentos rígidos, e as operações fundamentais da Geometria Projetiva são as projeções e seções.

Após Poncelet, uma legião de matemáticos veio abraçar o estudo da Geometria Projetiva, entre eles: Gergonne, Brianchon, Plücker, Staudt, Ruje e Cremona.

Referências para um estudo mais aprofundado da Geometria Projetiva podem ser vistos em [4] e [5].

1.3 Modelagem Matemática

1.3.1 Espaço-n Projetivo

Um ponto no espaço projetivo de dimensão n (\mathbb{P}^n) é representado por um vetor $n + 1$ como $\mathbf{p} = [p_1, p_2, \dots, p_{n+1}]^T$. Nesse vetor ao menos um elemento p_i deve ser não nulo. Essas coordenadas são conhecidas como *coordenadas homogêneas*.

Nessa representação, dois pontos \mathbf{p}^A e \mathbf{p}^B são considerados iguais, se obedecem à relação indicada na Equação (1.1).

$$\mathbf{p}_i^A = \eta \mathbf{p}_i^B \quad \text{sendo } \eta \neq 0, i = 1, 2, \dots, n + 1 \quad (1.1)$$

1.3.2 Plano Projetivo

O espaço projetivo de dimensão 2 (\mathbb{P}^2) é chamado de plano projetivo. Esse é o modelo que utilizaremos para representar o plano da imagem de uma câmera.

1.3.3 Pontos e Retas em \mathbb{P}^2

Ponto - \mathbf{m}

Um ponto \mathbf{m} no plano projetivo é representado pelo vetor em coordenadas homogêneas $\mathbf{m} = [m_x, m_y, m_z]^T$. No plano da imagem, esse ponto corresponde a $[m_x/m_z, m_y/m_z]^T \in \mathbb{R}^2$.

Podemos dizer que $\mathbf{m}' = [\eta m_x, \eta m_y, \eta m_z]^T$ é igual a \mathbf{m} para qualquer valor real não nulo de η .

Reta - \mathbf{r}

Uma reta \mathbf{r} no plano projetivo é representada pela Equação (1.2).

$$r_x x + r_y y + r_z z = 0 \quad (1.2)$$

Logo, a representação em coordenadas homogêneas da reta é $\mathbf{r} = [r_x, r_y, r_z]^T$.

Podemos dizer que $\mathbf{r}' = [\eta r_x, \eta r_y, \eta r_z]^T$ é igual a \mathbf{r} para qualquer valor real não nulo de η .

A partir da Equação (1.2) podemos obter a relação $\mathbf{m}^T \mathbf{r} = \mathbf{r}^T \mathbf{m} = 0$ e com isso observar que não existem diferenças entre retas e pontos em \mathbb{P}^2 . Assim, se postula o *princípio da dualidade*: o ponto e a reta são dois elementos duais no plano, ou mais explicitamente, “o ponto é dual da reta no plano” e “a reta é dual do ponto no plano”.

A partir das definições feitas anteriormente, enunciaremos alguns teoremas que dizem respeito à pertinência entre pontos e retas. Suas provas podem ser vistas no Apêndice A.

Teorema 1.1 *Uma reta r corta um ponto \mathbf{p} quando o produto interno entre eles for nulo ($\mathbf{r} \cdot \mathbf{p} = 0$).*

Teorema 1.2 *O ponto \mathbf{p} determinado por duas retas distintas \mathbf{r}_1 e \mathbf{r}_2 pode ser calculado como produto vetorial entre as retas ($\mathbf{p} = \mathbf{r}_1 \times \mathbf{r}_2$).*

Teorema 1.3 *A reta \mathbf{r} determinada por dois pontos distintos \mathbf{p}_1 e \mathbf{p}_2 pode ser calculada como produto vetorial entre os pontos ($\mathbf{r} = \mathbf{p}_1 \times \mathbf{p}_2$).*

Ponto Ideal - \mathbf{m}_∞

O ponto ideal corresponde ao ponto localizado no infinito no espaço euclidiano, ponto de encontro entre as retas paralelas. No plano projetivo, corresponde ao encontro entre duas retas projetivas. O ponto ideal é da forma $\mathbf{m}_\infty = [a, b, 0]$ (Apêndice A).

Reta Ideal - \mathbf{r}_∞

A reta ideal corresponde à reta que liga todos os pontos ideais. A reta ideal é da forma $\mathbf{r}_\infty = [0, 0, c]$ (Apêndice A).

1.3.4 Transformação Projetiva no Plano

A transformação projetiva no plano (TPP) descreve um mapeamento linear $\mathbb{P}^2 \rightarrow \mathbb{P}^2$, sendo representada por uma matriz quadrada 3×3 não singular. A matriz

homogênea \mathbf{H} é expressa pela Equação(1.3):

$$\begin{bmatrix} m'_x \\ m'_y \\ m'_z \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{bmatrix} m_x \\ m_y \\ m_z \end{bmatrix} \longrightarrow \mathbf{m}' = \mathbf{H}\mathbf{m} \quad (1.3)$$

Podemos aplicar essa transformação para retas, onde obtemos que $\mathbf{r}' = \mathbf{H}^{-T}\mathbf{r}$ (Apêndice A).

1.3.5 Espaço-3 Projetivo

O espaço projetivo de dimensão 3 (\mathbb{P}^3) é chamado de espaço-3 projetivo. Esse é o modelo que utilizaremos para representar o cenário 3D.

1.3.6 Pontos e Planos em \mathbb{P}^3

Ponto - \mathbf{M}

Um ponto \mathbf{M} no espaço-3 projetivo é representado pelo vetor em coordenadas homogêneas $\mathbf{M} = [M_x, M_y, M_z, M_w]^T$. No cenário 3D, esse ponto corresponde a $\mathcal{M} = [M_x/M_w, M_y/M_w, M_z/M_w]^T \in \mathbb{R}^3$.

Podemos dizer que $\mathbf{M}' = [\eta M_x, \eta M_y, \eta M_z, \eta M_w]^T$ é igual a \mathbf{M} para qualquer valor real não nulo de η .

Plano - π

Um plano π no espaço-3 projetivo é representado pela Equação (1.4).

$$\pi_x x + \pi_y y + \pi_z z + \pi_w w = 0 \quad \rightarrow \quad \pi^T \mathbf{M} = 0 \quad (1.4)$$

Logo, a representação em coordenadas homogêneas do plano é dada por $\pi = [\pi_x, \pi_y, \pi_z, \pi_w]^T$.

Podemos dizer que $\pi' = [\eta \pi_x, \eta \pi_y, \eta \pi_z, \eta \pi_w]^T$ é igual a π para qualquer valor real não nulo de η .

Para definirmos um plano, precisamos de três pontos distintos não colineares descritos por M_1 , M_2 e M_3 . Como cada ponto satisfaz à Equação (1.4), obtemos a Equação (1.5):

$$\begin{bmatrix} \mathbf{M}_1^T \\ \mathbf{M}_2^T \\ \mathbf{M}_3^T \end{bmatrix} \pi = 0 \quad (1.5)$$

Da mesma forma, três planos distintos linearmente independentes entre si, definidos por π_1 , π_2 e π_3 , possuem como interseção um ponto \mathbf{M} . Podemos calcular o ponto segundo a Equação (1.6):

$$\begin{bmatrix} \pi_1^T \\ \pi_2^T \\ \pi_3^T \end{bmatrix} \mathbf{M} = 0 \quad (1.6)$$

Em suma, pontos e planos em \mathbb{P}^3 são elementos duais, assim como vimos os pontos e retas em \mathbb{P}^2 .

1.3.7 Transformação Projetiva no Espaço-3

A transformação projetiva no espaço-3 (TPE3) descreve um mapeamento linear $\mathbb{P}^3 \rightarrow \mathbb{P}^3$, sendo representada por uma matriz quadrada 4×4 não singular. A matriz homogênea \mathbf{W} é expressa pela Equação (1.7):

$$\begin{bmatrix} M'_x \\ M'_y \\ M'_z \\ M'_w \end{bmatrix} = \begin{bmatrix} w_{11} & w_{12} & w_{13} & w_{14} \\ w_{21} & w_{22} & w_{23} & w_{24} \\ w_{31} & w_{32} & w_{33} & w_{34} \\ w_{41} & w_{42} & w_{43} & w_{44} \end{bmatrix} \begin{bmatrix} M_x \\ M_y \\ M_z \\ M_w \end{bmatrix} \longrightarrow \mathbf{M}' = \mathbf{W}\mathbf{M} \quad (1.7)$$

Podemos aplicar essa transformação para planos, onde obtemos que $\pi' = \mathbf{W}^{-T}\pi$ (Apêndice A).

1.3.8 Transformação Projetiva do Espaço-3 no Plano

A transformação projetiva do espaço-3 no plano (TPE3P) descreve um mapeamento linear $\mathbb{P}^3 \rightarrow \mathbb{P}^2$, sendo representada por uma matriz 3×4 . A matriz homogênea \mathbf{G} é expressa pela Equação (1.8):

$$\begin{bmatrix} m_x \\ m_y \\ m_z \end{bmatrix} = \begin{bmatrix} g_{11} & g_{12} & g_{13} & g_{14} \\ g_{21} & g_{22} & g_{23} & g_{24} \\ g_{31} & g_{32} & g_{33} & g_{34} \end{bmatrix} \begin{bmatrix} M_x \\ M_y \\ M_z \\ M_w \end{bmatrix} \longrightarrow \mathbf{m} = \mathbf{G}\mathbf{M} \quad (1.8)$$

1.4 Estrutura da Dissertação

Este trabalho está estruturado da seguinte forma:

O Capítulo 2 apresenta um estudo da modelagem de câmeras incluindo os seus modelos físicos e matemáticos. A partir dessa modelagem, consideramos os principais parâmetros de uma câmera (intrínsecos e extrínsecos). Por fim, realizamos um estudo de uma técnica linear de calibração desses parâmetros.

O Capítulo 3 apresenta um estudo da Geometria Epipolar, que descreve a relação das projeções de pontos 3D em duas imagens. Inicialmente, analisamos os seus elementos principais e a restrição epipolar que correlaciona os pontos correspondentes nas projeções. Depois, estudamos as duas matrizes que descrevem essa restrição entre duas imagens: as matrizes essencial e fundamental. Finalmente, analisamos técnicas lineares para estimação dessas matrizes (Algoritmos de 8 Pontos e de 8 Pontos Normalizado).

O Capítulo 4 apresenta um estudo da estimação do movimento através da matriz essencial obtida a partir da Geometria Epipolar entre duas imagens. Dentro do capítulo, analisamos três diferentes casos de movimentos e verificamos como podemos estimar os parâmetros para cada caso.

O Capítulo 5 apresenta a estrutura básica dos algoritmos, a partir dos conceitos expostos nos Capítulos 2, 3 e 4. A partir desses algoritmos, realizamos testes de modo a avaliarmos as técnicas lineares de estimação de movimento de uma câmera. Os resultados dos testes são apresentados.

O Capítulo 6 apresenta as conclusões quanto a limitações das técnicas lineares para a estimação de movimento.

Capítulo 2

Modelagem de Câmeras

Os pontos do cenário 3D (\mathcal{M}) são mapeados nos pontos do plano 2D da imagem da câmera (x_{img}), conforme vemos na Figura 2.1.

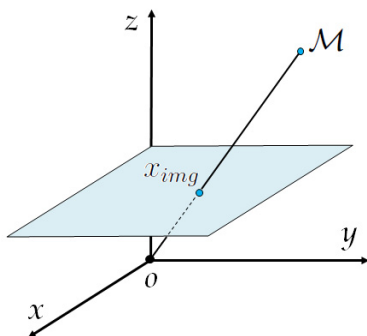


Figura 2.1: Exemplo do mapeamento dos pontos do cenário 3D no plano de projeção

Neste capítulo, introduziremos a modelagem projetiva de câmeras. Inicialmente, analisaremos os modelos físicos começando pelo modelo *pin-hole*. Depois, veremos um modelo mais específico para as nossas câmeras (modelo do sensor óptico). Por fim, juntando esses modelos, formularemos formalmente a modelagem matemática, estudando os parâmetros e a calibração de uma câmera.

2.1 Modelo *Pin-hole*

Esse modelo se baseia na análise física do comportamento óptico das câmeras escuras. A câmara escura consiste em uma caixa preta com um pequeno orifício em um dos seus lados. No lado oposto é formada a imagem invertida da cena à frente da pequena abertura. Todas câmeras fotográficas são baseadas nesse sistema. Uma revisão da história das câmeras escuras pode ser vista em [6] e [7].

2.1.1 Modelagem Física e Geométrica

Podemos ver como funciona a câmara escura na Figura 2.2.

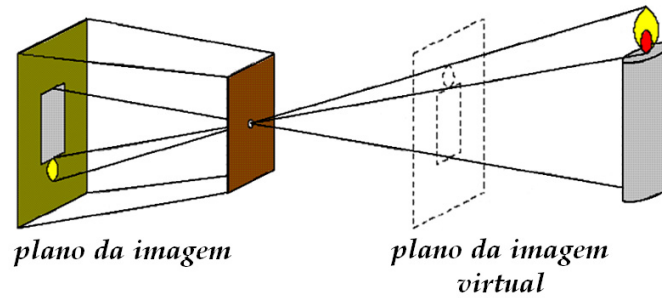


Figura 2.2: Imagens real e virtual projetadas pela câmara *pin-hole*

O modelo geométrico da câmara *pin-hole* pode ser visto na Figura 2.3. O centro de projeções C é sempre localizado na origem do sistema de coordenadas. A imagem projetada fica localizada atrás da origem ($z_c = -f$) no plano paralelo ao plano gerado pelos eixos x_c e y_c ($\widehat{x_c y_c}$), ficando essa imagem invertida em relação ao objeto do cenário (ver Figura 2.3).

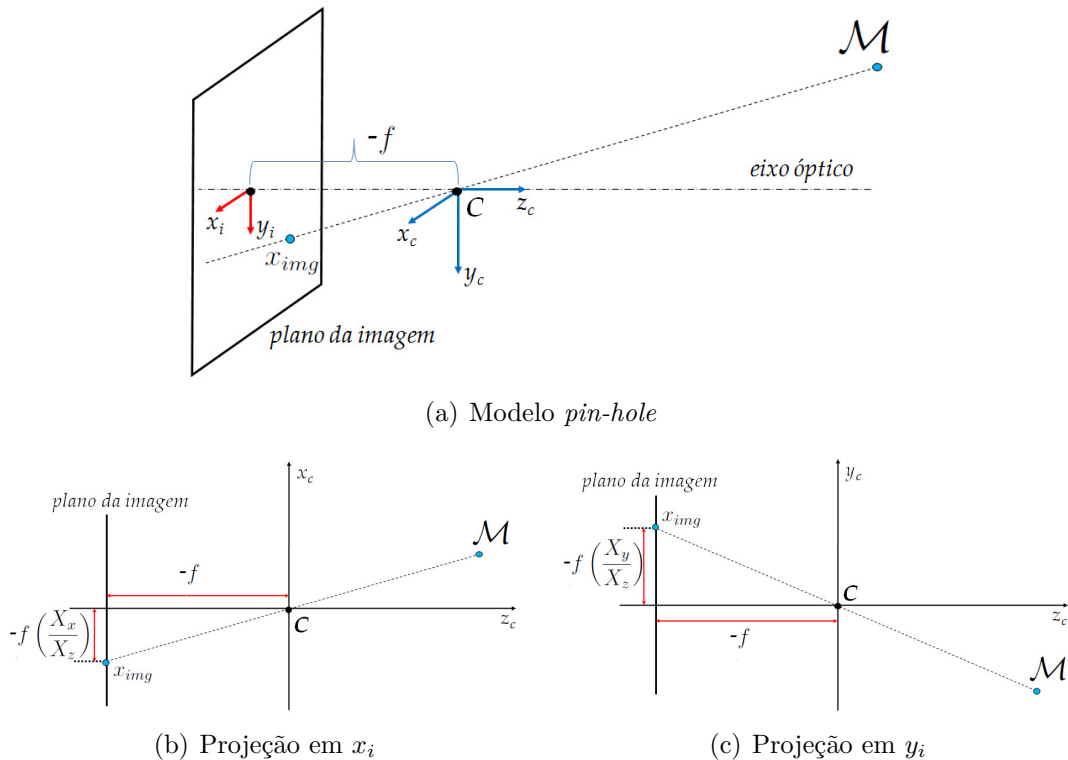
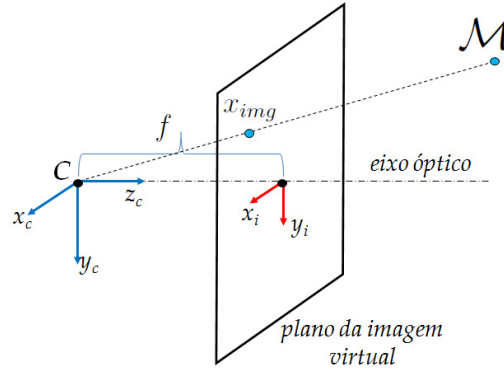


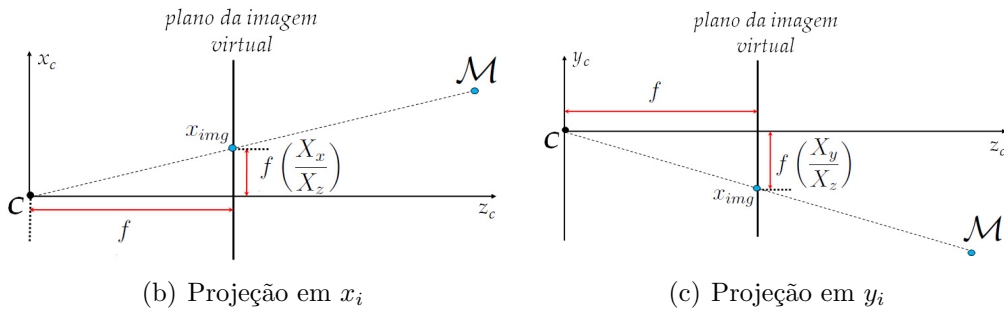
Figura 2.3: Modelagem geométrica do sistema *pin-hole* e análise de suas projeções em x_i e y_i

Existe um outro modelo (*pin-hole frontal*) onde analisamos a imagem no plano de projeção virtual (ver Figura 2.4). Nesse modelo, o plano da imagem é também

paralelo ao plano $\widehat{x_c y_c}$, mas posicionado na frente do centro de projeções em $z_c = f$. A vantagem desse modelo com relação ao anterior é que a imagem projetada não é invertida.



(a) Modelo *pin-hole* frontal



(b) Projecção em x_i

(c) Projecção em y_i

Figura 2.4: Modelagem geométrica do sistema *pin-hole* frontal e análise de suas projeções em x_i e y_i

A linha que liga o ponto $\mathcal{M} = [X_x, X_y, X_z]^T$ no sistema de coordenadas 3D no espaço euclidiano ao centro de projeções C (centro focal da câmera) intercepta o plano da imagem no ponto \mathbf{x}_{img} . Com uso da similaridade de triângulos, podemos afirmar que o ponto na imagem pode ser computado conforme a Equação (2.1):

$$\mathbf{x}_{img} = \left[f \left(\frac{X_x}{X_z} \right), f \left(\frac{X_y}{X_z} \right) \right]^T \quad (2.1)$$

Se representarmos os sistemas de coordenadas 3D no espaço euclidiano e do plano da imagem usando vetores homogêneos, a representação projetiva corresponde a:

$$\begin{bmatrix} f X_x \\ f X_y \\ X_z \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X_x \\ X_y \\ X_z \\ 1 \end{bmatrix} \quad (2.2)$$

Contudo, essa análise é feita sem a presença de lentes convergentes. O uso de lentes se faz necessário devido à limitação quanto ao tamanho de abertura da

câmara escura, que deve ser bem pequeno. Para diâmetros muito pequenos, ocorre o fenômeno da difração, que borra a imagem [8]. Para diâmetros grandes, a imagem também aparece borrada, mas devido ao fato de que cada ponto do objeto corresponder a uma mancha na imagem, conforme vemos ilustrado na Figura 2.5.

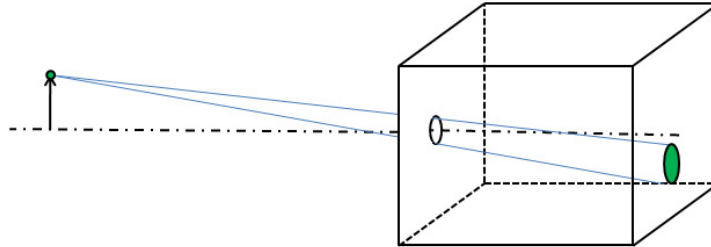


Figura 2.5: Ilustração do efeito de um orifício grande em uma câmara escura

Existe um diâmetro que oferece a perfeita projeção correspondendo a:

$$d = 2\sqrt{f\lambda} \quad \text{onde } \lambda: \text{ comprimento de onda da luz.}$$

O uso de lentes convergentes nas câmeras veio contornar esse problema da abertura, que limitava a entrada de luz a ser projetada. Contudo observou-se um novo fenômeno físico: o efeito do foco nas projeções das imagens.

Esse efeito pode ser observado na Figura 2.6, onde o foco é ajustado para o plano onde está localizado o tucano. Assim, as folhas localizadas à frente (região **A**) e as localizadas atrás desse plano (região **B**) têm a sua imagem projetada borrada e as outras na região no plano (região **C**) têm sua imagem nítida.

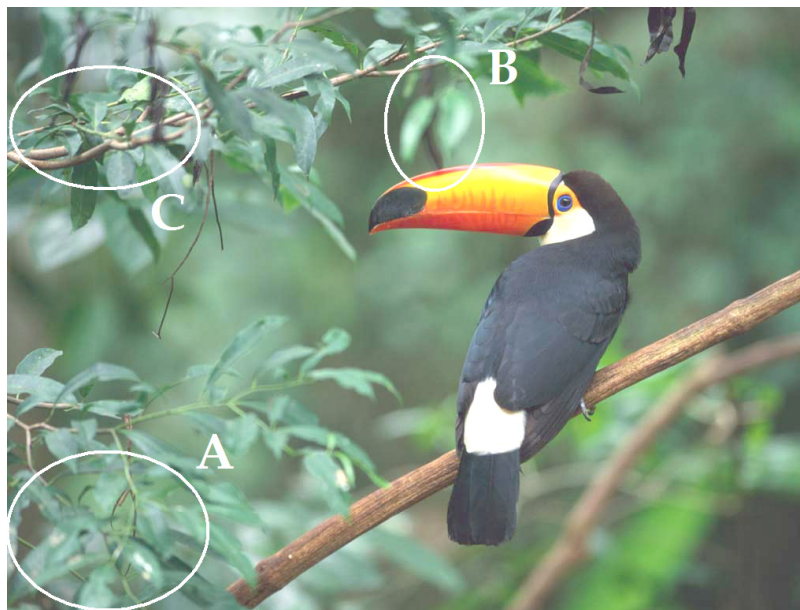


Figura 2.6: Exemplo para ilustrar objetos fora de foco (regiões **A** e **B**) e em foco (região **C**)

A explicação óptica para o fenômeno é que as lentes obedecem à *equação dos fabricantes de lentes* (Equação (2.3)), que relaciona a distância focal da lente (F) em função das distâncias do objeto à lente (d) e da lente ao plano de projeção (f) [8].

$$\frac{1}{F} = \frac{1}{d} + \frac{1}{f} \quad (2.3)$$

Quando essa relação é obedecida, um ponto do objeto é projetado como um ponto no plano da imagem, caso contrário, a imagem do ponto será espalhada na projeção. Nas Figuras 2.7, 2.8 e 2.9, podemos ver uma análise gráfica do comportamento dos raios de luz para esse fenômeno.

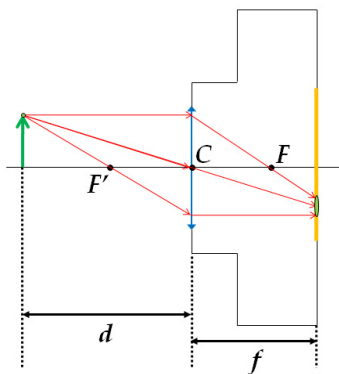


Figura 2.7: Imagem fora de foco (incidência dos raios de luz atrás do plano da imagem)

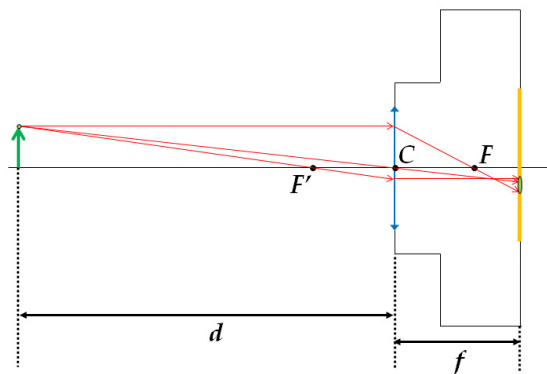


Figura 2.8: Imagem fora de foco (incidência dos raios de luz à frente do plano da imagem)

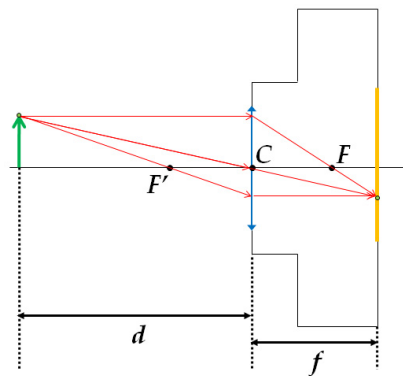


Figura 2.9: Imagem em foco

Dessa forma, se quisermos focalizar um objeto fora dessa distância, teremos que alterar ou a distância focal da lente ou a distância da lente ao plano da imagem.

2.2 Modelo do Sensor Óptico

O plano da imagem das câmeras digitais são compostas atualmente por dois tipos de sensores ópticos, os sensores CCD e CMOS. Apesar de os dois tipos de sensores terem diferenças entre si, isto não é levado em consideração em nosso trabalho, pois essencialmente eles possuem seus funcionamentos de captação de luz semelhantes.

Vimos no modelo *pin-hole* que as coordenadas x_i e y_i são valores contínuos, em unidades métricas. Nesse modelo, o plano da imagem é formado por elementos foto-sensíveis (sensores ópticos) com valores discretos chamados de *pixels*. Por isso, será necessário que façamos uma transformação desses elementos de modo que possamos juntar os dois modelos para gerarmos o modelo da câmera. Veremos, posteriormente, uma análise desses elementos do modelo do sensor dentro dos parâmetros intrínsecos da câmera.

2.3 Modelagem Matemática

Juntando os dois modelos descritos anteriormente e com base no conhecimento de Geometria Projetiva, podemos gerar um modelo matemático para um sistema de projeção de imagens de uma câmera.

$$\psi \mathbf{x} = \mathbf{P} \mathbf{X} \quad (2.4)$$

$$\psi \begin{bmatrix} x_u \\ x_v \\ 1 \end{bmatrix} = \begin{bmatrix} P_{11} & P_{12} & P_{13} & P_{14} \\ P_{21} & P_{22} & P_{23} & P_{24} \\ P_{31} & P_{32} & P_{33} & P_{34} \end{bmatrix} \begin{bmatrix} X_x \\ X_y \\ X_z \\ 1 \end{bmatrix}, \quad (2.5)$$

onde:

\mathbf{X} : ponto do cenário 3D normalizado em coordenadas homogêneas no \mathbb{P}^3

\mathbf{x} : ponto do plano da imagem 2D normalizado em coordenadas homogêneas no \mathbb{P}^2

\mathbf{P} : matriz de TPE3P

ψ : constante de profundidade

A matriz \mathbf{P} é chamada de matriz de projeção. Esta pode ser decomposta em duas matrizes em função das características de seus parâmetros.

$$\mathbf{P} = \mathbf{K} \mathcal{D}, \quad (2.6)$$

onde:

\mathbf{P} : matriz de projeção (3×4)

\mathbf{K} : matriz de calibração (3×3)

\mathcal{D} : matriz de orientação e posicionamento (3×4)

2.3.1 Parâmetros Intrínsecos

Parâmetros intrínsecos são os parâmetros da câmera que são associados à sua construção interna, logo quase todos não dependem da posição e orientação da câmera no espaço. O único parâmetro intrínseco que possui influência na posição e orientação é o parâmetro de projeção perspectiva. Isto porque o sistema de lentes obedece à Equação (2.3). Quando esse parâmetro é calibrado, o modelo funciona bem para uma região com profundidade fixa.

Esses parâmetros fornecerão as características ópticas, geométricas e digitais da câmera descritas a seguir:

1. a projeção perspectiva
 - f : comprimento do cone de projeção (distância entre a lente e o plano da imagem (em m))
2. a transformação de coordenadas do plano da imagem entre unidades contínuas métricas e discretas em pixels
 - s_u e s_v : tamanho do pixel nos eixos u e v (em $m \cdot pixels^{-1}$)
 - u_0 e v_0 : coordenadas do centro do sensor óptico (em $pixels$)
 - γ : fator de obliquidade (em $m \cdot pixels^{-1}$)
3. distorção geométrica radial introduzida pelas lentes
 - β_1 e β_2 : fatores de distorção

O modelo *pin-hole* será responsável pelo parâmetro f associado à projeção perspectiva, vide a Equação (2.1).

O modelo do sensor óptico estará associado aos parâmetros da transformação de coordenadas. A partir da Figura 2.10, podemos fazer uma análise dessa transformação.

Primeiramente, devemos fazer uma discretização dos pontos transformando as coordenadas métricas em coordenadas em pixel. Para isso, serão necessários os parâmetros do tamanho do pixel nos eixos u e v (s_u e s_v).

$$\begin{aligned}x_u &= \frac{x_{x_i}}{s_u} \\x_v &= \frac{x_{y_i}}{s_v}\end{aligned}\tag{2.7}$$

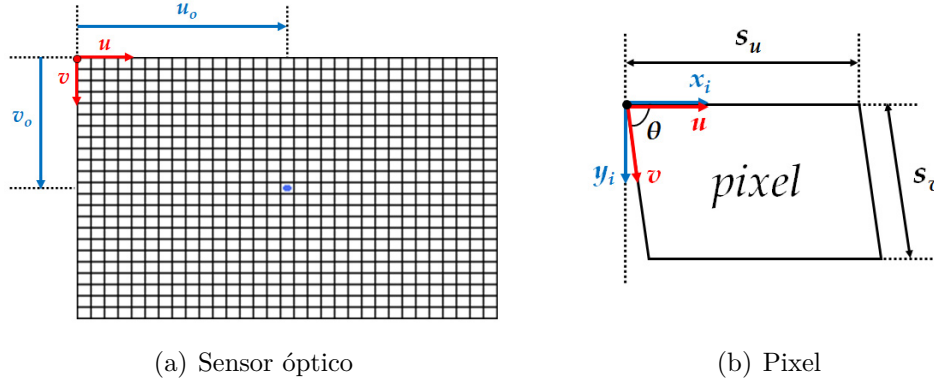


Figura 2.10: Imagens para análise dos parâmetros intrínsecos

Além disso, como vemos na Figura 2.10(a), a origem do sistema \widehat{uv} não fica no centro da imagem. Os parâmetros u_0 e v_0 são responsáveis pela translação, de modo que o sistema de coordenadas fique no centro da imagem do sensor óptico.

$$\begin{aligned} x_u &= \frac{x_{x_i}}{s_u} + u_0 \\ x_v &= \frac{x_{y_i}}{s_v} + v_0 \end{aligned} \quad (2.8)$$

Juntando o modelo do sensor com o *pin-hole*, obtemos as Equações (3.9). Podemos chamar estas equações de *modelo simples*.

$$\begin{aligned} x_u &= \frac{f}{s_u} \frac{X_x}{X_z} + u_0 \\ x_v &= \frac{f}{s_v} \frac{X_y}{X_z} + v_0 \end{aligned} \quad (2.9)$$

Um modelo mais geral leva em consideração o fato de que o sistema \widehat{uv} possa não ter base ortogonal como vemos na Figura 2.10(b). Devemos, então, fazer uma mudança de base [9], e para tal precisamos relacionar os sistemas vetoriais de modo a gerarmos as equações necessárias para obtermos a matriz de mudança de base [10].

Relação dos vetores \vec{u} e \vec{v} :

$$\begin{aligned} \vec{u} &= s_u \vec{x}_i \\ \vec{v} &= s_v [\vec{x}_i \cos \theta + \vec{y}_i \sin \theta] \end{aligned} \quad (2.10)$$

Relação dos vetores \vec{x}_i e \vec{y}_i :

$$\begin{aligned} \vec{x}_i &= \frac{1}{s_u} \vec{u} \\ \vec{y}_i &= -\frac{1}{s_u} \cot \theta \vec{u} + \frac{1}{s_v \sin \theta} \vec{v} \end{aligned} \quad (2.11)$$

Assim, a partir da Equações (2.11), podemos obter a base no \mathbb{R}^2 :

$$\begin{aligned} \hat{S} : \hat{u} &= (s_u, 0) \quad , \quad \hat{v} = (s_v \cos \theta, s_v \sin \theta) \\ \hat{I} : \hat{x}_i &= (1, 0) \quad , \quad \hat{y}_i = (0, 1) \end{aligned} \quad (2.12)$$

A matriz de mudança de base \mathbf{Q} da base não ortogonal \hat{S} para base ortogonal \hat{I} é dada por:

$$\mathbf{Q} = \begin{bmatrix} \frac{1}{s_u} & -\frac{1}{s_u} \cot \theta \\ 0 & \frac{1}{s_v \sin \theta} \end{bmatrix} \quad (2.13)$$

Conseqüentemente, o modelo geral será dado pelas Equações (2.14). Podemos notar que quando o ângulo $\theta = 90^\circ$, o modelo geral é equivalente ao simples.

$$\begin{aligned} x_u &= \frac{f}{s_u} \frac{X_x}{X_z} - \frac{f \cot \theta}{s_u} \frac{X_y}{X_z} + u_0 \\ x_v &= \frac{f}{s_v \sin \theta} \frac{X_y}{X_z} + v_0 \end{aligned} \quad (2.14)$$

Por definição, o fator de obliquidade γ será dado por:

$$\gamma = -\frac{f}{s_u} \cot \theta \quad (2.15)$$

Nesse trabalho, não faremos um estudo da distorção radial, pelo fato desta não possuir grande influência nos resultados das imagens. Contudo, mais informações podem ser encontradas em [11].

Por fim, podemos juntar esses parâmetros de forma a gerarmos uma matriz TPP. Essa matriz de transformação \mathbf{K} é chamada de *matriz de calibração*.

Sendo $\alpha_u = f/s_u$ e $\alpha_v = f/s_v$ os comprimentos do cone de projeção nas coordenadas dos eixos u e v da imagem (em *pixels*), vamos definir as matrizes de calibração em função dos modelos:

Modelo simples

$$\mathbf{K} = \begin{bmatrix} \alpha_u & 0 & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (2.16)$$

Modelo geral

$$\mathbf{K} = \begin{bmatrix} \alpha_u & -\alpha_u \cot \theta & u_0 \\ 0 & \frac{\alpha_v}{\sin \theta} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (2.17)$$

2.3.2 Parâmetros Extrínsecos

Parâmetros extrínsecos são os parâmetros da câmera que são associados à orientação e ao posicionamento do sistema de coordenadas da câmera centrado no centro óptico C em relação ao sistema de coordenadas global (conforme a Figura 2.11).

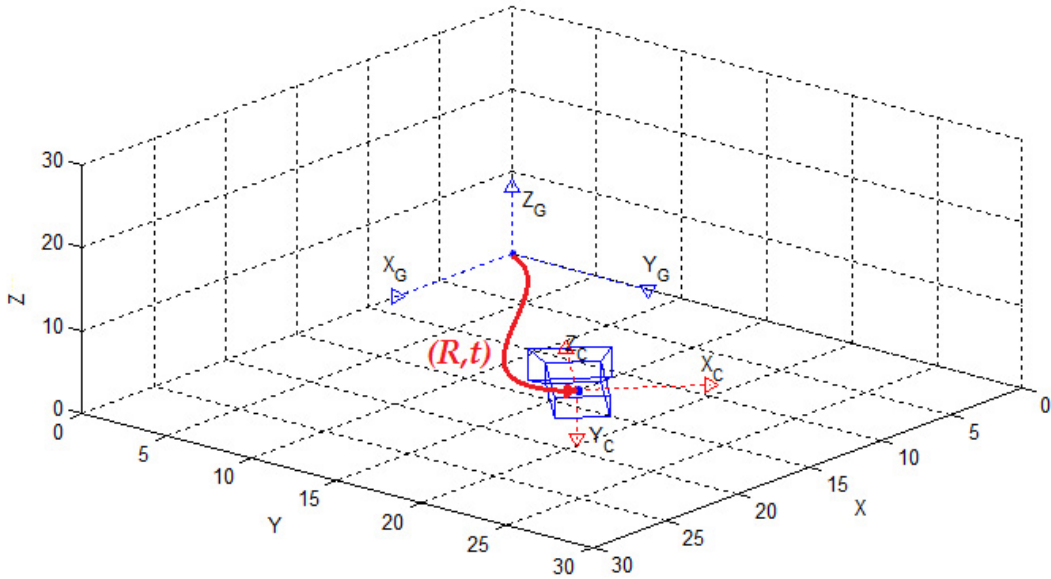


Figura 2.11: Parâmetros extrínsecos

Os parâmetros dividem-se em:

Parâmetros de translação: que estão associados ao posicionamento.

Correspondem ao deslocamento da origem do sistema de coordenadas da câmera para o sistema de coordenadas global, combinadas em um vetor de translação:

- t_x : deslocamento sobre o eixo x
- t_y : deslocamento sobre o eixo y
- t_z : deslocamento sobre o eixo z

Parâmetros de rotação: que estão associados à orientação.

Correspondem aos ângulos de rotação que deve ser aplicados aos eixos do sistema de coordenadas da câmera de modo que eles adquiram a mesma orientação do sistema de coordenadas global.

- ω : ângulo de rotação no eixo x no sentido anti-horário (*roll*)
- ϕ : ângulo de rotação no eixo y no sentido anti-horário (*pitch*)
- ρ : ângulo de rotação no eixo z no sentido anti-horário (*yaw*)

A representação da translação será dada pelo vetor \mathbf{t} conforme a Equação (2.18).

$$\mathbf{t} = \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix} \quad (2.18)$$

A representação da rotação será dada pela matriz \mathbf{R} . Utilizaremos a convenção dos ângulos de Euler Z-Y-X [12], conforme podemos ver a seguir.

A matriz de rotação de Givens é a matriz cuja rotação é associada a apenas um dos eixos do sistema referencial [13]. As matrizes de rotação de Givens de um sistema 3D são:

$$\mathbf{R}_x(\omega) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \omega & -\sin \omega \\ 0 & \sin \omega & \cos \omega \end{bmatrix} \rightarrow \text{eixo } x \quad (2.19)$$

$$\mathbf{R}_y(\phi) = \begin{bmatrix} \cos \phi & 0 & \sin \phi \\ 0 & 1 & 0 \\ -\sin \phi & 0 & \cos \phi \end{bmatrix} \rightarrow \text{eixo } y \quad (2.20)$$

$$\mathbf{R}_z(\rho) = \begin{bmatrix} \cos \rho & -\sin \rho & 0 \\ \sin \rho & \cos \rho & 0 \\ 0 & 0 & 1 \end{bmatrix} \rightarrow \text{eixo } z \quad (2.21)$$

Uma matriz de rotação em função das matrizes de rotação de Givens segundo a convenção dos ângulos de Euler Z-Y-X corresponde a:

$$\mathbf{R} = \mathbf{R}_z(\rho)\mathbf{R}_y(\phi)\mathbf{R}_x(\omega) \quad (2.22)$$

$$\mathbf{R} = \begin{bmatrix} \cos \rho \cos \phi & \cos \rho \sin \phi \sin \omega - \sin \rho \cos \omega & \cos \rho \sin \phi \cos \omega + \sin \rho \sin \omega \\ \sin \rho \cos \phi & \sin \rho \sin \phi \sin \omega + \cos \rho \cos \omega & \sin \rho \sin \phi \cos \omega - \cos \rho \sin \omega \\ -\sin \phi & \cos \phi \sin \omega & \cos \phi \cos \omega \end{bmatrix} \quad (2.23)$$

Por fim, podemos juntar esses parâmetros de forma a gerarmos uma matriz TPE3P. Essa matriz de transformação \mathcal{D} é chamada de *matriz de orientação e posicionamento*.

$$\mathcal{D} = [\mathbf{R} \quad \mathbf{t}] = \begin{bmatrix} \bar{\mathbf{r}}_1 & t_x \\ \bar{\mathbf{r}}_2 & t_y \\ \bar{\mathbf{r}}_3 & t_z \end{bmatrix} \quad \text{onde } \bar{\mathbf{r}}_n = [r_{n1}, r_{n2}, r_{n3}], \quad n = 1, 2, 3 \quad (2.24)$$

2.3.3 Calibração

O processo de estimar os parâmetros intrínsecos de uma câmera é chamado de calibração. Em certas técnicas de calibração podem ser estimados também os parâmetros extrínsecos. Essa estimativa é feita a partir de pontos referenciais controlados de modo a obtermos uma matriz de transformação dos pontos 3D para o

plano da imagem. Existem duas metodologias principais para calibração linear de câmeras:

Método de Tsai [14]: essa técnica requer um *grid* 3D como referência. É a técnica mais usada em calibração, pois, além de ser a mais robusta, com apenas uma imagem, podemos obter a estimação tanto dos parâmetros intrínsecos como dos extrínsecos.

Método de Zhang [15]: essa técnica requer um *grid* planar 2D como referência. Requer que façamos deslocamentos não necessariamente conhecidos da câmera ou do *grid* para fazermos a estimativa dos parâmetros intrínsecos. Ela é de mais fácil implementação que o outro método, mas é menos robusta.

As técnicas não-lineares utilizam-se muitas vezes das lineares como ponto de partida para seus algoritmos de otimização. As soluções lineares são, normalmente, muito próximas da solução não-linear do sistema [16]. Usaremos uma técnica linear baseada na metodologia de Tsai para estimar os parâmetros da câmera neste trabalho.

Estimação da matriz \mathbf{P}

Primeiramente, vamos escrever a matriz \mathbf{P} da Equação(2.5) da seguinte forma:

$$\mathbf{P} = \begin{bmatrix} \mathbf{q}_1^T & q_{14} \\ \mathbf{q}_2^T & q_{24} \\ \mathbf{q}_3^T & q_{34} \end{bmatrix} \quad (2.25)$$

Substituindo a Equação (2.25) na Equação (2.5), obtemos a relação entre as coordenadas dos pontos 3D do cenário e os pontos 2D da imagem. Definindo que $\mathcal{M} = [X_x, X_y, X_z]^T$, obtemos as seguintes equações para um ponto:

$$\begin{aligned} \psi x_x &= \mathbf{q}_1^T \mathcal{M} + q_{14} \\ \psi x_y &= \mathbf{q}_2^T \mathcal{M} + q_{24} \\ \psi &= \mathbf{q}_3^T \mathcal{M} + q_{34} \end{aligned}$$

Substituindo ψ nas equações anteriores, obtemos as duas equações lineares.

$$\begin{aligned} \mathbf{q}_1^T \mathcal{M}^{(i)} - x_x^{(i)} \mathbf{q}_3^T \mathcal{M}^{(i)} + q_{14} - x_x^{(i)} q_{34} &= 0 \\ \mathbf{q}_2^T \mathcal{M}^{(i)} - x_y^{(i)} \mathbf{q}_3^T \mathcal{M}^{(i)} + q_{24} - x_y^{(i)} q_{34} &= 0 \end{aligned}$$

Para N pontos, obtemos $2N$ equações compondo um sistema linear homogêneo:

$$\mathbf{A}\mathbf{q} = \mathbf{0}, \quad (2.26)$$

onde:

A: matriz ($2N \times 12$) dos pontos 2D da imagem e dos pontos 3D da referência no cenário

q: vetor (12×1) correspondendo a $\mathbf{q} = [\mathbf{q}_1^T, q_{14}, \mathbf{q}_2^T, q_{24}, \mathbf{q}_3^T, q_{34}]^T$

Podemos nesse caso, aplicar uma otimização de mínimos quadrados usando SVD (conforme explicação no Apêndice B) como algumas referências sugerem, como por exemplo em [17]. Contudo, Faugeras [16] sugere uma otimização com restrição sobre a norma de \mathbf{q}_3 , que será explicada a seguir.

Primeiramente, vamos obter a matriz **P** a partir dos parâmetros da câmera, ou seja, $\mathbf{P} = \mathbf{KD}$:

Modelo simples

$$\mathbf{P} = \begin{bmatrix} \alpha_u \bar{\mathbf{r}}_1 + u_0 \bar{\mathbf{r}}_3 & \alpha_u t_x + u_0 t_z \\ \alpha_v \bar{\mathbf{r}}_2 + v_0 \bar{\mathbf{r}}_3 & \alpha_v t_y + v_0 t_z \\ \bar{\mathbf{r}}_3 & t_z \end{bmatrix} \quad (2.27)$$

Modelo geral

$$\mathbf{P} = \begin{bmatrix} \alpha_u \bar{\mathbf{r}}_1 - \alpha_u \cot \theta \bar{\mathbf{r}}_2 + u_0 \bar{\mathbf{r}}_3 & \alpha_u t_x - \alpha_u \cot \theta t_y + u_0 t_z \\ \frac{\alpha_v}{\sin \theta} \bar{\mathbf{r}}_2 + v_0 \bar{\mathbf{r}}_3 & \frac{\alpha_v}{\sin \theta} t_y + v_0 t_z \\ \bar{\mathbf{r}}_3 & t_z \end{bmatrix} \quad (2.28)$$

Podemos notar que para qualquer um dos modelos $\mathbf{q}_3^T = \bar{\mathbf{r}}_3$. Como $\bar{\mathbf{r}}_3$ corresponde a uma linha de matriz de rotação (que é ortogonal), podemos afirmar que $\|\bar{\mathbf{r}}_3\|^2 = 1$ por ser um vetor de uma base ortonormal. Conseqüentemente, concluímos que

$$\|\mathbf{q}_3^T\|^2 = 1 \quad (2.29)$$

Na otimização com restrição sugerida por Faugeras, que é analisada no Apêndice C, usaremos essa restrição mais específica encontrada na Equação (2.29).

As matrizes **C** e **D** da otimização conforme o Apêndice C são definidas pelas Equações (2.30) e (2.31).

$$\mathbf{C} = \begin{bmatrix} (\mathcal{M}^{(1)})^T & 1 & \mathbf{0}_{1 \times 4} & -x_x^{(1)} \\ \mathbf{0}_{1 \times 4} & (\mathcal{M}^{(1)})^T & 1 & -x_y^{(1)} \\ \vdots & \vdots & \vdots & \vdots \\ (\mathcal{M}^{(N)})^T & 1 & \mathbf{0}_{1 \times 4} & -x_x^{(N)} \\ \mathbf{0}_{1 \times 4} & (\mathcal{M}^{(N)})^T & 1 & -x_y^{(N)} \end{bmatrix}_{2N \times 9} \quad (2.30)$$

$$\mathbf{D} = \begin{bmatrix} x_x^{(1)}(\mathcal{M}^{(1)})^T \\ x_y^{(1)}(\mathcal{M}^{(1)})^T \\ \vdots \\ x_x^{(N)}(\mathcal{M}^{(N)})^T \\ x_y^{(N)}(\mathcal{M}^{(N)})^T \end{bmatrix}_{2N \times 3} \quad (2.31)$$

A solução será composta pelos vetores \mathbf{y} e \mathbf{z} que corresponderão a:

$$\mathbf{y} = \begin{bmatrix} \mathbf{q}_1 \\ q_{14} \\ \mathbf{q}_2 \\ q_{24} \\ q_{34} \end{bmatrix}_{9 \times 1} \quad (2.32)$$

$$\mathbf{z} = \begin{bmatrix} \mathbf{q}_3 \end{bmatrix}_{3 \times 1} \quad (2.33)$$

Decomposição da matriz \mathbf{P}

Os parâmetros intrínsecos e extrínsecos são calculados a partir da matriz \mathbf{P} que é estimada pelo método explicado na seção anterior. A partir das Equações (2.25), (2.27) e (2.28), poderemos definir as equações de cada parâmetro. Cabe ressaltar que, além da solução do vetor \mathbf{q} para o sistema homogêneo $\mathbf{A}\mathbf{q} = \mathbf{0}$, a solução $-\mathbf{q}$ também é válida tanto para o sistema quanto para a restrição da Equação (2.29). A fim de resolver esse fator de escala $\epsilon = \pm 1$, mais uma vez, iremos analisar as Equações (2.27) e (2.28), e podemos ver que:

$$t_z = \epsilon q_{34} \quad (2.34)$$

O componente t_z sempre terá de ser sempre positivo, pois o *grid* estará sempre localizado à frente da câmera no processo de calibração, como podemos ver na Figura 2.11. Logo, a Equação (2.34) será a nossa referência quanto à escolha do fator de escala ϵ .

Podemos obter as relações dos nossos parâmetros conforme vemos nas Tabelas 2.1 e 2.2. Provas destas equações podem ser encontradas no Apêndice A.

Tabela 2.1: Equações dos parâmetros do modelo simples

Variáveis	Equações
u_0 e v_0	$u_0 = \mathbf{q}_1^T \mathbf{q}_3$ $v_0 = \mathbf{q}_2^T \mathbf{q}_3$
α_u e α_v	$\alpha_u = \sqrt{\mathbf{q}_1^T \mathbf{q}_1 - u_0^2}$ $\alpha_v = \sqrt{\mathbf{q}_2^T \mathbf{q}_2 - v_0^2}$
translação \mathbf{t}	$t_x = \epsilon \frac{(q_{14} - u_0 q_{34})}{\alpha_u}$ $t_y = \epsilon \frac{(q_{24} - v_0 q_{34})}{\alpha_v}$ $t_z = \epsilon q_{34}$
rotação \mathbf{R}	$\bar{\mathbf{r}}_1 = \epsilon \frac{(\mathbf{q}_1^T - u_0 \mathbf{q}_3^T)}{\alpha_u}$ $\bar{\mathbf{r}}_2 = \epsilon \frac{(\mathbf{q}_2^T - v_0 \mathbf{q}_3^T)}{\alpha_v}$ $\bar{\mathbf{r}}_3 = \epsilon \mathbf{q}_3^T$

Tabela 2.2: Equações dos parâmetros do modelo geral

Variáveis	Equações
u_0 e v_0	$u_0 = \mathbf{q}_1^T \mathbf{q}_3$ $v_0 = \mathbf{q}_2^T \mathbf{q}_3$
α_u e α_v	$\alpha_u = \sqrt{\mathbf{q}_1^T \mathbf{q}_1 - u_0^2} \sin \theta$ $\alpha_v = \sqrt{\mathbf{q}_2^T \mathbf{q}_2 - v_0^2} \sin \theta$
translação \mathbf{t}	$t_x = \epsilon [q_{14} + (q_{24} - v_0 q_{34}) \frac{\alpha_u}{\alpha_v} \cos \theta - u_0 q_{34}] \frac{1}{\alpha_u}$ $t_y = \epsilon (q_{24} - v_0 q_{34}) \frac{\sin \theta}{\alpha_v}$ $t_z = \epsilon q_{34}$
rotação \mathbf{R}	$\bar{\mathbf{r}}_1 = \epsilon [\mathbf{q}_1^T + (\mathbf{q}_2^T - v_0 \mathbf{q}_3^T) \frac{\alpha_u}{\alpha_v} \cos \theta - u_0 \mathbf{q}_3^T] \frac{1}{\alpha_u}$ $\bar{\mathbf{r}}_2 = \epsilon (\mathbf{q}_2^T - v_0 \mathbf{q}_3^T) \frac{\sin \theta}{\alpha_v}$ $\bar{\mathbf{r}}_3 = \epsilon \mathbf{q}_3^T$
ângulo θ	$\cos \theta = - \frac{(\mathbf{q}_1 \times \mathbf{q}_3) \cdot (\mathbf{q}_2 \times \mathbf{q}_3)}{\ \mathbf{q}_1 \times \mathbf{q}_3\ \ \mathbf{q}_2 \times \mathbf{q}_3\ }$

Capítulo 3

Geometria Epipolar

Duas imagens projetadas por uma câmera em movimento de um cenário 3D são relacionadas pela Geometria Epipolar, conforme vemos no exemplo da Figura 3.1.

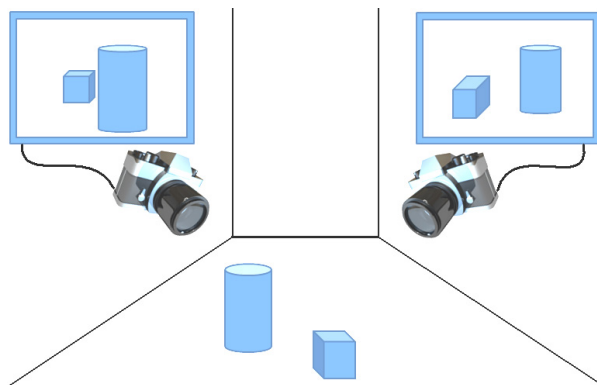


Figura 3.1: Exemplo de imagens para uso da Geometria Epipolar

Neste capítulo, daremos início com algumas definições sobre a Geometria Epipolar. Nas seções seguintes, descreveremos as matrizes essencial e fundamental que correlacionam os pares de pontos homólogos¹ de duas imagens de uma câmera em movimento e depois descreveremos como estimá-las de forma linear.

3.1 Definições

A Geometria Epipolar é a relação geométrica da interseção entre os planos das imagens com o conjunto de planos epipolares, tendo a reta da linha de base como interseção entre eles, conforme vemos na Figura 3.2. Na próxima seção, apresentamos seus principais elementos.

¹Um par de pontos homólogos corresponde às projeções de um ponto 3D do cenário nos planos de duas imagens.

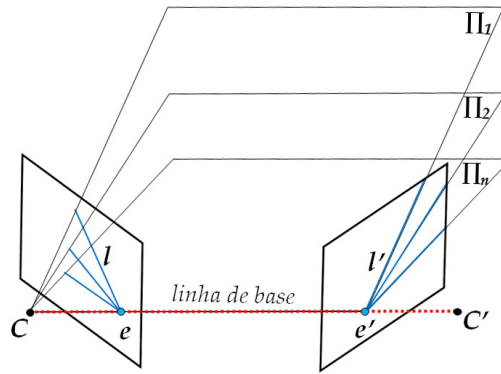


Figura 3.2: Ilustração da definição de Geometria Epipolar

3.1.1 Elementos da Geometria Epipolar

Os principais elementos da Geometria Epipolar são representados na Figura 3.3.

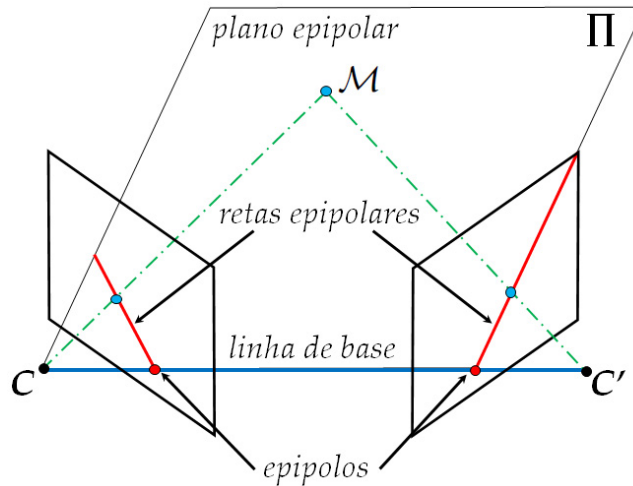


Figura 3.3: Elementos da Geometria Epipolar

Linha de Base - b

A linha de base (*baseline*) é a reta projetiva que passa pelos centros focais das duas câmeras, C e C' .

Plano Epipolar - Π

O plano epipolar é o plano gerado pelos três pontos não colineares: os dois centros focais C e C' e o ponto 3D do cenário \mathcal{M} .

Reta Epipolar - l

As retas epipolares são as retas projetivas geradas pela interseção entre o plano epipolar e os planos das imagens.

Epipolo - e

Os epipolos são os pontos gerados pela interseção entre as retas epipolares e a reta da linha de base.

3.2 Restrição Epipolar

A partir da Geometria Epipolar entre duas imagens de uma câmera, pode ser observado que: sendo o ponto projetado no plano da primeira imagem \mathbf{x}_1 conhecido, então a reta epipolar l_2 , formada por $\mathbf{e}_2 \times \mathbf{x}_2$ (Teorema 1.3), também será conhecida e possuirá o seu ponto homólogo, ou seja, a projeção do ponto 3D (\mathcal{M}) no plano da segunda imagem, o ponto \mathbf{x}_2 . Em outras palavras, a cada ponto projetado de um ponto 3D em uma imagem, está associada uma reta epipolar na outra imagem na qual situa-se o seu ponto homólogo, conforme os exemplos da Figura 3.4.

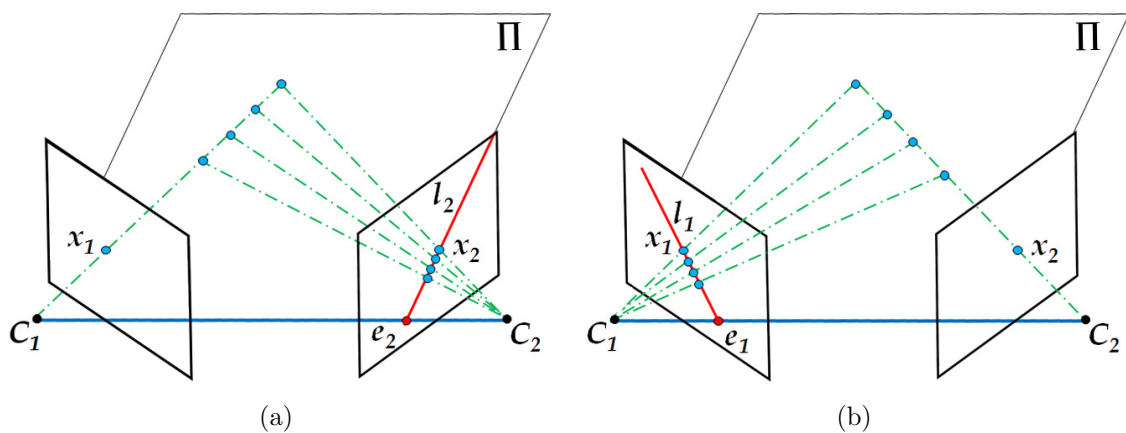


Figura 3.4: Exemplos gráficos da restrição epipolar

Isso corresponde à restrição epipolar que pontos homólogos devem satisfazer. Com isso, caso a restrição seja atendida em ambos os sentidos, pode-se afirmar que os dois pontos correspondem à projeção de um mesmo ponto 3D.

Um exemplo com duas imagens de uma câmera é apresentado na Figura 3.5, onde um ponto 3D é projetado em ambas as imagens nos pontos \mathbf{x}_1 e \mathbf{x}_2 respectivamente. Através dessa restrição, verificamos que ambos correspondem a um ponto de suas respectivas retas epipolares l_1 e l_2 .

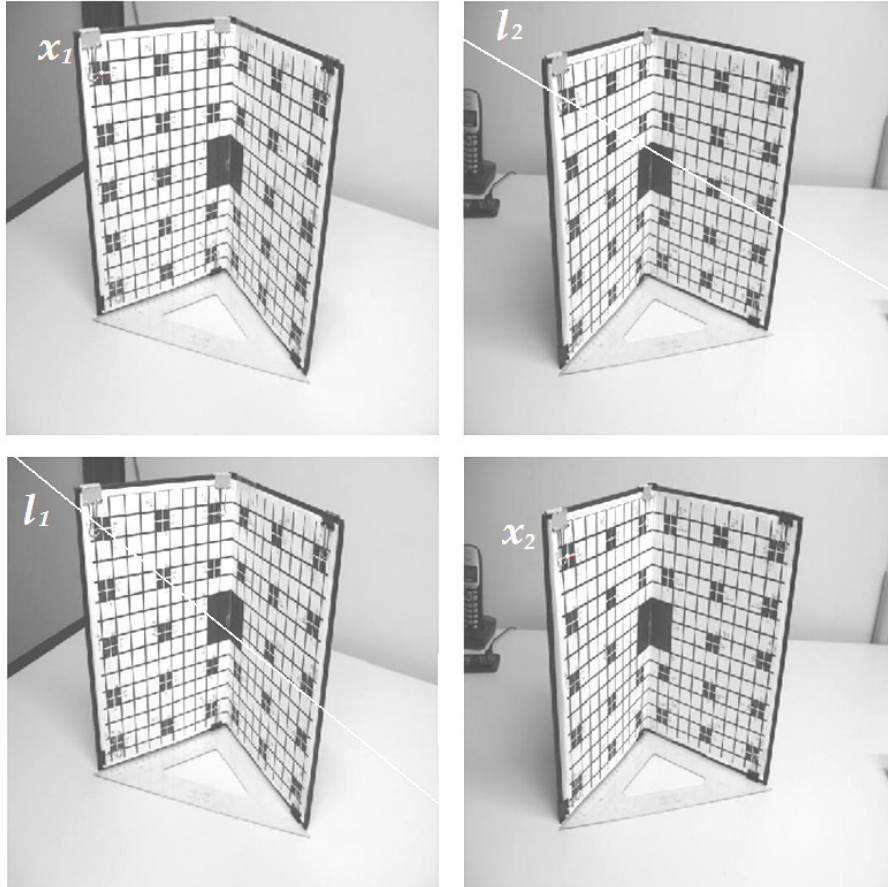


Figura 3.5: Exemplo real da restrição epipolar

A restrição epipolar pode ser descrita matematicamente segundo as matrizes essencial e fundamental associadas entre duas imagens da câmera, conforme veremos nas próximas seções.

3.3 Matriz Essencial

Antes de falarmos sobre a matriz essencial é necessário definirmos, primeiramente, o que vem a ser um ponto calibrado em coordenadas homogêneas (\check{x}). Este pode ser descrito como sendo um ponto do plano da imagem 2D, cuja matriz de calibração \mathbf{K} é igual à matriz identidade. Logo, sua matriz de projeção \mathbf{P} será composta somente pelos parâmetros extrínsecos da câmera.

A matriz essencial foi introduzida por Longuet-Higgins [18]. É a matriz que descreve a Geometria Epipolar dos pontos homólogos de duas imagens calibradas (quando $\mathbf{K} = \mathbf{I}$). Veremos a seguir que essa matriz correlaciona os pontos homólogos somente segundo a mudança de orientação e posicionamento entre as imagens.

Definição Matemática da Matriz Essencial

A partir da Figura 3.6, analisaremos como a matriz essencial correlaciona os pontos calibrados \check{x}_1 e \check{x}_2 entre as duas imagens.

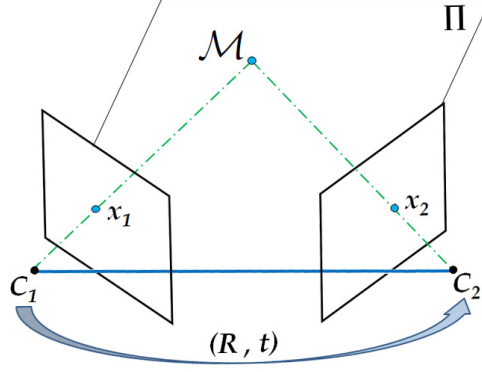


Figura 3.6: Exemplo de duas imagens com movimento da câmera de (\mathbf{R}, \mathbf{t})

Definindo \mathcal{M} como o ponto 3D do cenário, seu ponto em coordenadas homogêneas normalizadas em \mathbb{P}^3 é dado por:

$$\mathbf{X} = \begin{bmatrix} \mathcal{M} \\ 1 \end{bmatrix} \quad (3.1)$$

Agora, iremos fazer uma TPE3 (Seção 1.3.7), onde mapearemos o ponto \mathcal{M} no sistema de coordenadas da primeira imagem e na posição do espaço-3 projetivo sob o plano da imagem. Os parâmetros \mathbf{R}_0 e \mathbf{t}_0 correspondem aos parâmetros extrínsecos da câmera na primeira imagem.

$$\mathbf{X}_1 = \begin{bmatrix} \mathbf{R}_0 & \mathbf{t}_0 \\ \mathbf{0}_{3 \times 1} & 1 \end{bmatrix} \mathbf{X} = \begin{bmatrix} \psi_1 \check{x}_1 \\ 1 \end{bmatrix} \quad (3.2)$$

Faremos a TPE3P (Seção 1.3.8) do ponto 3D na projeção da primeira imagem \mathbf{X}_1 para a projeção da segunda imagem \check{x}_2 .

$$\psi_2 \check{x}_2 = \mathbf{P} \mathbf{X}_1 = \begin{bmatrix} \mathbf{R} & \mathbf{t} \end{bmatrix} \mathbf{X}_1 \quad (3.3)$$

Substituindo \mathbf{X}_1 pelo seu valor dado na Equação (3.2):

$$\psi_2 \check{x}_2 = \psi_1 \mathbf{R} \check{x}_1 + \mathbf{t} \quad (3.4)$$

Multiplicando pela esquerda pela matriz $[\mathbf{t}]_{\times}^2$:

$$\psi_2 [\mathbf{t}]_{\times} \check{x}_2 = \psi_1 [\mathbf{t}]_{\times} \mathbf{R} \check{x}_1 + \overbrace{[\mathbf{t}]_{\times} \mathbf{t}}^{=0} \quad (3.5)$$

² $[\cdot]_{\times}$ - operador matriz de produto vetorial (Apêndice A)

Multiplicando pela esquerda pelo vetor $\check{\mathbf{x}}_2^T$:

$$\psi_2 \overbrace{\check{\mathbf{x}}_2^T ([\mathbf{t}]_{\times} \check{\mathbf{x}}_2)}^{=0} = \psi_1 \left(\check{\mathbf{x}}_2^T [\mathbf{t}]_{\times} \mathbf{R} \check{\mathbf{x}}_1 \right) \quad (3.6)$$

Obtemos a equação da restrição epipolar que correlaciona os pontos das duas imagens:

$$\check{\mathbf{x}}_2^T [\mathbf{t}]_{\times} \mathbf{R} \check{\mathbf{x}}_1 = 0 \quad (3.7)$$

A matriz essencial é definida por:

$$\mathbf{E} = [\mathbf{t}]_{\times} \mathbf{R} \quad (3.8)$$

Com isso, vemos que a matriz essencial só depende da orientação e do posicionamento entre as imagens da câmera.

3.3.1 Propriedades da Matriz Essencial

Propriedade 3.3.1 (Restrição epipolar)

$$\check{\mathbf{x}}_2^T \mathbf{E} \check{\mathbf{x}}_1 = \check{\mathbf{x}}_1^T \mathbf{E}^T \check{\mathbf{x}}_2 = 0$$

Prova: Equação (3.7).

Propriedade 3.3.2 (Posto da matriz)

$$p(\mathbf{E}) = 2$$

Prova:

Como a matriz de translação gerada pelo operador $([\cdot]_{\times})$ possui posto 2 e a matriz de rotação \mathbf{R} possui posto completo, então a matriz essencial \mathbf{E} , formada pelo produto das duas matrizes (Equação 3.8), possuirá o menor posto, ou seja,

$$p(\mathbf{E}) = 2.$$

Propriedade 3.3.3 (Retas epipolares)

$$\check{\mathbf{l}}_2 = \mathbf{E}\check{\mathbf{x}}_1 \quad \check{\mathbf{l}}_1 = \mathbf{E}^T\check{\mathbf{x}}_2$$

Prova:

Sabemos que $\check{\mathbf{x}}_2$ é um ponto da reta projetiva $\check{\mathbf{l}}_2$, conseqüentemente:

$$\check{\mathbf{x}}_2^T \check{\mathbf{l}}_2 = 0$$

Logo, a partir da equação da restrição epipolar:

$$\check{\mathbf{x}}_2^T \overbrace{\mathbf{E}\check{\mathbf{x}}_1}^{=\check{\mathbf{l}}_2} = 0 \quad \rightarrow \quad \boxed{\check{\mathbf{l}}_2 = \mathbf{E}\check{\mathbf{x}}_1}$$

Sabemos que $\check{\mathbf{x}}_1$ é um ponto da reta projetiva $\check{\mathbf{l}}_1$, conseqüentemente:

$$\check{\mathbf{l}}_1^T \check{\mathbf{x}}_1 = 0$$

Logo, a partir da equação da restrição epipolar:

$$\overbrace{\check{\mathbf{x}}_2^T \mathbf{E}}^{=\check{\mathbf{l}}_1^T} \check{\mathbf{x}}_1 = 0 \quad \rightarrow \quad \boxed{\check{\mathbf{l}}_1 = \mathbf{E}^T\check{\mathbf{x}}_2}$$

Propriedade 3.3.4 (Epipolos)

$$\mathbf{E}\check{\mathbf{e}}_1 = \mathbf{0} \quad \mathbf{E}^T\check{\mathbf{e}}_2 = \mathbf{0}$$

Prova:

Para qualquer ponto $\check{\mathbf{x}}_1$ (diferente do epipolo $\check{\mathbf{e}}_1$), a reta epipolar $\check{\mathbf{l}}_2 = \mathbf{E}\check{\mathbf{x}}_1$ (Propriedade 3.3.3) contém o epipolo $\check{\mathbf{e}}_2$.

Logo,

$$\check{\mathbf{l}}_2^T \check{\mathbf{e}}_2 = 0 \quad \rightarrow \quad \overbrace{\check{\mathbf{x}}_1^T (\mathbf{E}^T \check{\mathbf{e}}_2)}^{=\check{\mathbf{l}}_2^T} = 0$$

ou seja, o epipolo $\check{\mathbf{e}}_2$ é o vetor do espaço nulo da matriz \mathbf{E}^T : $\boxed{\mathbf{E}^T\check{\mathbf{e}}_2 = \mathbf{0}}$.

A partir da equação: $\mathbf{E} = [\mathbf{t}]_{\times} \mathbf{R}$, obtemos $\mathbf{E}^T = \mathbf{R}^T [\mathbf{t}]_{\times}^T = \mathbf{R}^T [-\mathbf{t}]_{\times}$

Usando a Propriedade 3.3.3, obtemos:

$$\mathbf{R}^T([\mathbf{-t}]_{\times}\check{\mathbf{e}}_2) = \mathbf{R}^T(-\mathbf{t} \times \check{\mathbf{e}}_2) = 0$$

Concluimos que $\boxed{\check{\mathbf{e}}_2 = \mathbf{t}}$.

Para qualquer ponto $\check{\mathbf{x}}_2$ (diferente do epipolo $\check{\mathbf{e}}_2$), a reta epipolar $\check{\mathbf{I}}_1 = \mathbf{E}^T\check{\mathbf{x}}_2$ (Propriedade 3.3.3) contém o epipolo $\check{\mathbf{e}}_1$.

Logo,

$$\check{\mathbf{I}}_1^T\check{\mathbf{e}}_1 = 0 \quad \rightarrow \quad \overbrace{\check{\mathbf{x}}_2^T}^{\check{\mathbf{I}}_1^T}(\mathbf{E}\check{\mathbf{e}}_1) = 0$$

ou seja, o epipolo $\check{\mathbf{e}}_1$ é o vetor do espaço nulo da matriz \mathbf{E} : $\boxed{\mathbf{E}\check{\mathbf{e}}_1 = \mathbf{0}}$.

A partir da equação $\mathbf{E} = [\mathbf{t}]_{\times}\mathbf{R}$ e da Propriedade 3.3.3, obtemos:

$$[\mathbf{t}]_{\times}(\mathbf{R}\check{\mathbf{e}}_1) = \mathbf{t} \times (\mathbf{R}\check{\mathbf{e}}_1) = 0$$

Concluimos que $\boxed{\check{\mathbf{e}}_1 = \mathbf{R}^T\mathbf{t}}$.

Propriedade 3.3.5 (Matriz dos valores singulares)

$$\boxed{\mathbf{E} = \mathbf{U}\mathbf{S}_E\mathbf{V}^T \quad \rightarrow \quad \mathbf{S}_E = \begin{bmatrix} s_{\alpha} & & \\ & s_{\alpha} & \\ & & 0 \end{bmatrix}, \forall s_{\alpha} \in \mathbb{R}^+}$$

Prova:

Essa prova é baseada nos resultados do artigo [19].

Sabemos que: $\boxed{\mathbf{E} = [\mathbf{t}]_{\times}\mathbf{R}}$.

Existe uma matriz de rotação \mathbf{R}_* , que transforma o vetor de translação \mathbf{t} em

$$\mathbf{a} = \mathbf{R}_*\mathbf{t} = [0, 0, \|\mathbf{t}\|]^T.$$

A partir da propriedade do operador matriz de produto vetorial [20], que diz:

Para um vetor $\mathbf{v} \in \mathbb{R}^3$ e uma matriz $\mathbf{H} \in \mathbb{R}^{3 \times 3}$, se $\det(\mathbf{H}) = +1$ e $\mathbf{v}' = \mathbf{H}\mathbf{v}$, então $[\mathbf{v}]_{\times} = \mathbf{H}^T[\mathbf{v}']_{\times}\mathbf{H}$.

Podemos afirmar que:

$$[\mathbf{t}]_{\times} = \mathbf{R}_*^T[\mathbf{a}]_{\times}\mathbf{R}_*$$

Então,

$$\mathbf{E}\mathbf{E}^T = [\mathbf{t}]_{\times} \overbrace{\mathbf{R}\mathbf{R}^T}^{=I} [\mathbf{t}]_{\times}^T = [\mathbf{t}]_{\times} [\mathbf{t}]_{\times}^T = \mathbf{R}_*^T [\mathbf{a}]_{\times} \overbrace{\mathbf{R}_* \mathbf{R}_*^T}^{=I} [\mathbf{a}]_{\times}^T \mathbf{R}_* = \mathbf{R}_*^T [\mathbf{a}]_{\times} [\mathbf{a}]_{\times}^T \mathbf{R}_*$$

Calculemos $[\mathbf{a}]_{\times} [\mathbf{a}]_{\times}^T$:

$$[\mathbf{a}]_{\times} [\mathbf{a}]_{\times}^T = \begin{bmatrix} 0 & -\|\mathbf{t}\| & 0 \\ \|\mathbf{t}\| & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & \|\mathbf{t}\| & 0 \\ -\|\mathbf{t}\| & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} \|\mathbf{t}\|^2 & 0 & 0 \\ 0 & \|\mathbf{t}\|^2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Podemos concluir que $[\mathbf{a}]_{\times} [\mathbf{a}]_{\times}^T$ corresponde a matriz diagonal Λ dos autovalores da matriz $\mathbf{E}\mathbf{E}^T$. A partir dos conhecimentos de SVD do Apêndice B, podemos concluir que a matriz diagonal dos valores singulares de \mathbf{E} será dada por:

$$\mathbf{S}_E = \begin{bmatrix} \|\mathbf{t}\| & 0 & 0 \\ 0 & \|\mathbf{t}\| & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Por causa do posto incompleto da matriz operador produto vetorial, o vetor \mathbf{t} é definido exceto por um fator de escala. Conseqüentemente, podemos definir: $s_{\alpha} = \eta \|\mathbf{t}\|$, sendo $\eta \in \mathbb{R}^+$.

Dessa forma obtemos

$$\mathbf{S}_E = \begin{bmatrix} s_{\alpha} & 0 & 0 \\ 0 & s_{\alpha} & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

3.4 Matriz Fundamental

Da mesma forma como iniciamos o estudo da matriz essencial na Seção 3.3, é necessário definirmos um ponto não-calibrado em coordenadas homogêneas ($\tilde{\mathbf{x}}$) como sendo um ponto do plano da imagem 2D cuja matriz de calibração \mathbf{K} é diferente da matriz identidade.

A relação entre os pontos calibrados e não-calibrados é dada por uma matriz TPP (Seção 1.3.4) representada pela matriz de calibração (\mathbf{K}).

$$\tilde{\mathbf{x}} = \mathbf{K}\check{\mathbf{x}} \tag{3.9}$$

A Geometria Epipolar que relaciona os pontos homólogos das duas imagens não-calibradas é descrita pela matriz fundamental \mathbf{F} , a ser definida em seguida.

A partir da equação da restrição epipolar da Propriedade 3.3.1 e da Equação (3.9), temos:

$$\tilde{\mathbf{x}}_2^T \mathbf{K}_2^{-T} \mathbf{E} \mathbf{K}_1^{-1} \tilde{\mathbf{x}}_1 = 0 \quad (3.10)$$

A matriz \mathbf{F} depende da orientação e do posicionamento entre as imagens da câmera como a matriz \mathbf{E} , mas também dos parâmetros intrínsecos das imagens inseridos nas matrizes de calibração associadas a cada imagem:

$$\mathbf{F} = \mathbf{K}_2^{-T} \mathbf{E} \mathbf{K}_1^{-1} \quad (3.11)$$

3.4.1 Propriedades da Matriz Fundamental

Propriedade 3.4.1 (Restrição epipolar)

$$\tilde{\mathbf{x}}_2^T \mathbf{F} \tilde{\mathbf{x}}_1 = \tilde{\mathbf{x}}_1^T \mathbf{F}^T \tilde{\mathbf{x}}_2 = 0$$

Prova: Equação (3.10).

Propriedade 3.4.2 (Posto da matriz)

$$p(\mathbf{F}) = 2$$

Prova:

Como a matriz essencial \mathbf{E} possui posto 2 (Propriedade 3.3.2) e as matrizes de calibração das imagens \mathbf{K}_1 e \mathbf{K}_2 possuem posto completo, então a matriz fundamental \mathbf{F} , formada pelo produto entre elas (Equação (3.11)), possuirá o menor posto, ou seja, $p(\mathbf{F}) = 2$.

Propriedade 3.4.3 (Retas epipolares)

$$\tilde{\mathbf{l}}_2 = \mathbf{F} \tilde{\mathbf{x}}_1 \quad \tilde{\mathbf{l}}_1 = \mathbf{F}^T \tilde{\mathbf{x}}_2$$

Prova:

Sabemos que $\tilde{\mathbf{x}}_2$ é um ponto da reta projetiva $\tilde{\mathbf{l}}_2$, conseqüentemente:

$$\tilde{\mathbf{x}}_2^T \tilde{\mathbf{l}}_2 = 0$$

Logo, a partir da equação da restrição epipolar:

$$\tilde{\mathbf{x}}_2^T \overbrace{\mathbf{F} \tilde{\mathbf{x}}_1}^{=\tilde{\mathbf{l}}_2} = 0 \quad \rightarrow \quad \boxed{\tilde{\mathbf{l}}_2 = \mathbf{F} \tilde{\mathbf{x}}_1}$$

Sabemos que $\tilde{\mathbf{x}}_1$ é um ponto da reta projetiva $\tilde{\mathbf{I}}_1$, conseqüentemente:

$$\tilde{\mathbf{I}}_1^T \tilde{\mathbf{x}}_1 = 0$$

Logo, a partir da equação da restrição epipolar:

$$\overbrace{\tilde{\mathbf{x}}_2^T \mathbf{F}}^{=\tilde{\mathbf{I}}_1^T} \tilde{\mathbf{x}}_1 = 0 \quad \rightarrow \quad \boxed{\tilde{\mathbf{l}}_1 = \mathbf{F}^T \tilde{\mathbf{x}}_2}$$

Propriedade 3.4.4 (Epipolos)

$\mathbf{F} \tilde{\mathbf{e}}_1 = \mathbf{0}$	$\mathbf{F}^T \tilde{\mathbf{e}}_2 = \mathbf{0}$
--	--

Prova:

Para qualquer ponto $\tilde{\mathbf{x}}_1$ (diferente do epipolo $\tilde{\mathbf{e}}_1$), a reta epipolar $\tilde{\mathbf{l}}_2 = \mathbf{F} \tilde{\mathbf{x}}_1$ (Propriedade 3.4.3) contém o epipolo $\tilde{\mathbf{e}}_2$. Logo,

$$\tilde{\mathbf{l}}_2^T \tilde{\mathbf{e}}_2 = 0 \quad \rightarrow \quad \overbrace{\tilde{\mathbf{x}}_1^T (\mathbf{F}^T \tilde{\mathbf{e}}_2)}^{=\tilde{\mathbf{l}}_2^T} = 0$$

ou seja, o epipolo $\tilde{\mathbf{e}}_2$ é o vetor do espaço nulo da matriz \mathbf{F}^T : $\boxed{\mathbf{F}^T \tilde{\mathbf{e}}_2 = \mathbf{0}}$. A partir da equação $\mathbf{F} = \mathbf{K}_2^{-T} \mathbf{E} \mathbf{K}_1^{-1} = \mathbf{K}_2^{-T} [\mathbf{t}]_{\times} \mathbf{R} \mathbf{K}_1^{-1}$, obtemos:

$$\mathbf{F}^T = \mathbf{K}_1^{-T} \mathbf{R}^T [\mathbf{t}]_{\times}^T \mathbf{K}_2^{-1} = \mathbf{K}_1^{-T} \mathbf{R}^T [-\mathbf{t}]_{\times} \mathbf{K}_2^{-1}$$

Usando a Propriedade 3.4.3, obtemos:

$$\mathbf{K}_1^{-T} \mathbf{R}^T ([-\mathbf{t}]_{\times} \mathbf{K}_2^{-1} \tilde{\mathbf{e}}_2) = \mathbf{K}_1^{-T} \mathbf{R}^T (-\mathbf{t} \times (\mathbf{K}_2^{-1} \tilde{\mathbf{e}}_2)) = 0$$

Concluimos que $\boxed{\tilde{\mathbf{e}}_2 = \mathbf{K}_2 \mathbf{t}}$.

Para qualquer ponto $\tilde{\mathbf{x}}_2$ (diferente do epipolo $\tilde{\mathbf{e}}_2$), a reta epipolar $\tilde{\mathbf{l}}_1 = \mathbf{F}^T \tilde{\mathbf{x}}_2$ (Propriedade 3.4.3) contém o epipolo $\tilde{\mathbf{e}}_1$. Logo,

$$\tilde{\mathbf{l}}_1^T \tilde{\mathbf{e}}_1 = 0 \quad \rightarrow \quad \overbrace{\tilde{\mathbf{x}}_2^T (\mathbf{F} \tilde{\mathbf{e}}_1)}^{=\tilde{\mathbf{l}}_1^T} = 0$$

ou seja, o epipolo $\tilde{\mathbf{e}}_1$ é o vetor do espaço nulo da matriz \mathbf{F} : $\boxed{\mathbf{F}\tilde{\mathbf{e}}_1 = \mathbf{0}}$.

Usando a equação $\mathbf{F} = \mathbf{K}_2^{-T}[\mathbf{t}]_{\times}\mathbf{R}\mathbf{K}_1^{-1}$ e a Propriedade 3.4.3, obtemos:

$$\mathbf{K}_2^{-T}[\mathbf{t}]_{\times}(\mathbf{R}\mathbf{K}_1^{-1}\tilde{\mathbf{e}}_1) = \mathbf{K}_2^{-T}\mathbf{t} \times (\mathbf{R}\mathbf{K}_1^{-1}\tilde{\mathbf{e}}_1) = 0$$

Concluimos que $\boxed{\tilde{\mathbf{e}}_1 = \mathbf{K}_1\mathbf{R}^T\mathbf{t}}$.

Propriedade 3.4.5 (Matriz dos valores singulares)

$$\mathbf{F} = \mathbf{U}\mathbf{S}_F\mathbf{V}^T \quad \rightarrow \quad \mathbf{S}_F = \begin{bmatrix} s_{\alpha} & & \\ & s_{\beta} & \\ & & 0 \end{bmatrix}, \quad \forall s_{\alpha} \text{ e } \forall s_{\beta} \in \mathbb{R}^+$$

Prova:

Como a matriz fundamental \mathbf{F} possui posto 2 (Propriedade 3.4.2), isso implica que na SVD (Apêndice B) haverá dois valores singulares (s_{α} e s_{β}).

3.5 Estimação das Matrizes Essencial e Fundamental

Existem várias formas de estimarmos essas matrizes [21]. Neste trabalho, iremos optar pelo uso de técnicas lineares em função de serem mais rápidas e de mais fácil implementação e análise.

3.5.1 Algoritmo de 8 Pontos

O Algoritmo de 8 Pontos foi introduzido por Longuet-Higgins [18] para o cálculo da matriz essencial. Nesse artigo, a matriz essencial é usada para calcular a estrutura de um cenário via duas imagens de câmeras calibradas.

Iremos notar que, na sua formulação, [18] utiliza a equação da restrição epipolar para obter o sistema de equações lineares. Sabendo que a restrição epipolar é igual tanto para matriz essencial quanto para a fundamental (Propriedades 3.3.1 e 3.4.1), esse algoritmo pode ser utilizado também para estimar a matriz fundamental, conforme proposto por Hartley e Zisserman [13].

Se possuímos ao menos 8 pontos homólogos nas duas imagens, podemos determinar os nove parâmetros da matriz \mathbf{F} (ou \mathbf{E} , caso os pontos sejam calibrados), exceto por um fator de escala. Sejam os pontos homólogos $\mathbf{x}_1 = [u_1, v_1, 1]^T$

e $\mathbf{x}_2 = [u_2, v_2, 1]^T$. Como eles são homólogos, a equação da restrição epipolar é satisfeita:

$$\mathbf{x}_2^T \mathbf{F} \mathbf{x}_1 = 0$$

$$\begin{bmatrix} u_2 & v_2 & 1 \end{bmatrix} \begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{bmatrix} \begin{bmatrix} u_1 \\ v_1 \\ 1 \end{bmatrix} = 0$$

Efetuada essa multiplicação matricial, obtemos a seguinte equação linear:

$$u_2 u_1 f_{11} + u_2 v_1 f_{12} + u_2 f_{13} + v_2 u_1 f_{21} + v_2 v_1 f_{22} + v_2 f_{23} + u_1 f_{31} + v_1 f_{32} + f_{33} = 0$$

Dessa forma, formamos um conjunto de equações lineares para N pares de pontos homólogos ($N \geq 8$):

$$\begin{bmatrix} u_1^{(1)} u_2^{(1)} & u_2^{(1)} v_1^{(1)} & u_2^{(1)} & v_2^{(1)} u_1^{(1)} & v_1^{(1)} v_2^{(1)} & v_2^{(1)} & u_1^{(1)} & v_1^{(1)} & 1 \\ u_1^{(2)} u_2^{(2)} & u_2^{(2)} v_1^{(2)} & u_2^{(2)} & v_2^{(2)} u_1^{(2)} & v_1^{(2)} v_2^{(2)} & v_2^{(2)} & u_1^{(2)} & v_1^{(2)} & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ u_1^{(N)} u_2^{(N)} & u_2^{(N)} v_1^{(N)} & u_2^{(N)} & v_2^{(N)} u_1^{(N)} & v_1^{(N)} v_2^{(N)} & v_2^{(N)} & u_1^{(N)} & v_1^{(N)} & 1 \end{bmatrix} \begin{bmatrix} f_{11} \\ f_{12} \\ f_{13} \\ f_{21} \\ f_{22} \\ f_{23} \\ f_{31} \\ f_{32} \\ f_{33} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\mathbf{A} \mathbf{f} = \mathbf{0} \tag{3.12}$$

Como vemos na Equação (3.12), trata-se de achar a solução de um sistema homogêneo. Resolveremos o problema através da otimização linear de mínimos quadrados através da SVD, conforme o Apêndice B. A solução de \mathbf{F} será definida exceto por um fator de escala.

A inacurácia na determinação dos pontos homólogos e o ruído de suas coordenadas implica no posto da matriz ser $p(\mathbf{A}) = 9$. Por sua vez, a matriz \mathbf{F} (ou \mathbf{E}) estimada será não singular, ou seja, a solução não atende à propriedade do posto de \mathbf{F} (ou \mathbf{E}) ser igual 2, o que corresponderia ao caso em que o posto da matriz fosse $p(\mathbf{A}) = 8$.

Uma solução para este problema consiste em substituir \mathbf{F} pela matriz singular \mathbf{F}^* que minimize a norma de Frobenius: $\|\mathbf{F} - \mathbf{F}^*\|$ [22], fazendo o mesmo para matriz \mathbf{E} [20].

Fazemos a SVD da matriz \mathbf{F} e obtemos:

$$\mathbf{F} = \mathbf{U}_F \mathbf{S}_F \mathbf{V}_F^T \quad (3.13)$$

onde \mathbf{S}_F é a matriz diagonal $[s_a, s_b, s_c]$ com valores singulares de \mathbf{F} , sendo $s_a \geq s_b \geq s_c$. A solução que minimiza a norma de Frobenius será dada por [22]:

$$\mathbf{F}^* = \mathbf{U}_F \mathbf{S}_{F^*} \mathbf{V}_F^T, \quad (3.14)$$

onde \mathbf{S}_{F^*} é a matriz diagonal $[s_a, s_b, 0]$ com valores singulares de F^* , sendo $s_a \geq s_b > 0$. No caso da estimativa da matriz \mathbf{E} , prova-se [20] que esta correção do problema da singularidade leva ao resultado

$$\mathbf{E}^* = \mathbf{U}_E \mathbf{S}_{E^*} \mathbf{V}_E^T, \quad (3.15)$$

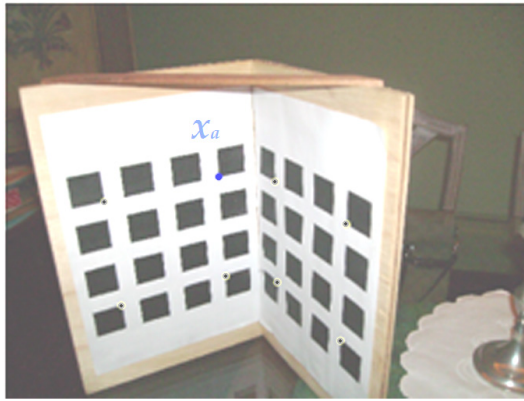
onde \mathbf{S}_{E^*} é a matriz diagonal $[\frac{s_a+s_b}{2}, \frac{s_a+s_b}{2}, 0]$ com valores singulares de \mathbf{E}^* .

Entretanto, esse ajuste da singularidade implica em um aumento do erro do algoritmo, conforme podemos ver na análise com as figuras a seguir.

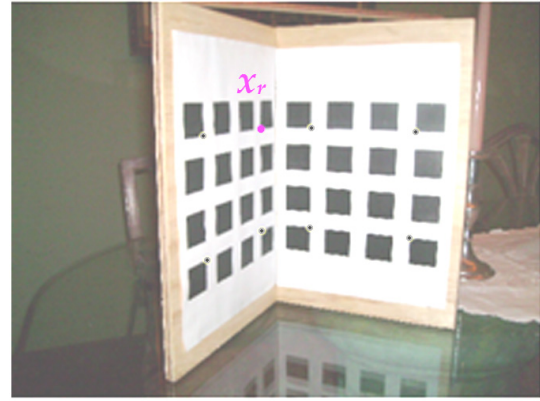
Usando as quatro fotos de um *grid* em posições e orientações da câmera diferentes (Figura 3.7), através do Algoritmo de 8 Pontos, estimamos as matrizes \mathbf{F} para cada caso em função de uma quinta imagem comum às quatro. Usamos os oito pontos 3D demarcados nas imagens. A partir de um mesmo ponto 3D em comum entre as imagens, obtivemos, na quinta imagem, suas retas epipolares $(\mathbf{l}_a, \mathbf{l}_r, \mathbf{l}_{vd}, \mathbf{l}_{vm})$ geradas a partir dos pontos $(\mathbf{x}_a, \mathbf{x}_r, \mathbf{x}_{vd}, \mathbf{x}_{vm})$ e suas respectivas matrizes \mathbf{F} (Propriedade 3.4.3).

No primeiro caso (Figura 3.8), as matrizes \mathbf{F} não sofreram o ajuste de singularidade no algoritmo. No segundo caso (Figura 3.9), as matrizes \mathbf{F} sofreram o ajuste de singularidade no algoritmo.

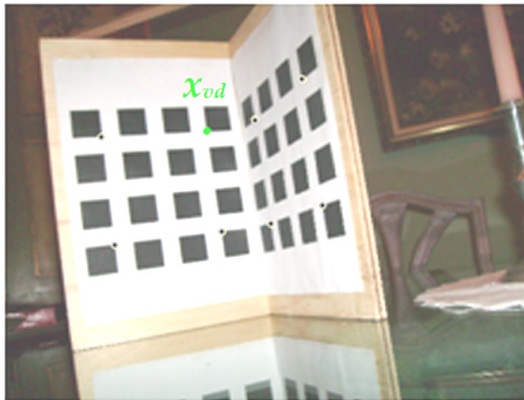
Podemos observar que depois do ajuste de singularidade do algoritmo, as retas não mais convergiram para o mesmo ponto hómologo. Uma solução para corrigir essa instabilidade será estudada na seção seguinte, no Algoritmo de 8 Pontos Normalizado.



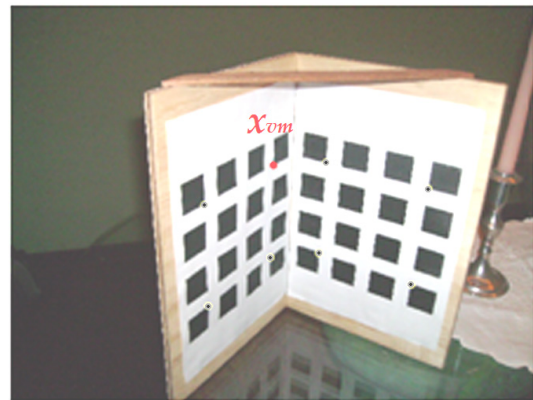
(a) Ponto x_a



(b) Ponto x_r



(c) Ponto x_{vd}



(d) Ponto x_{vm}

Figura 3.7: Imagens iniciais para os Algoritmos de 8 Pontos

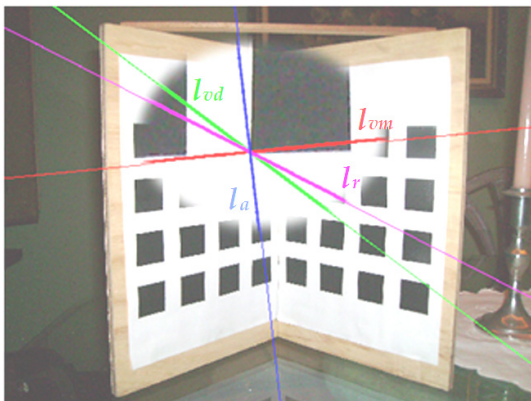


Figura 3.8: Algoritmo de 8 Pontos sem ajuste de singularidade

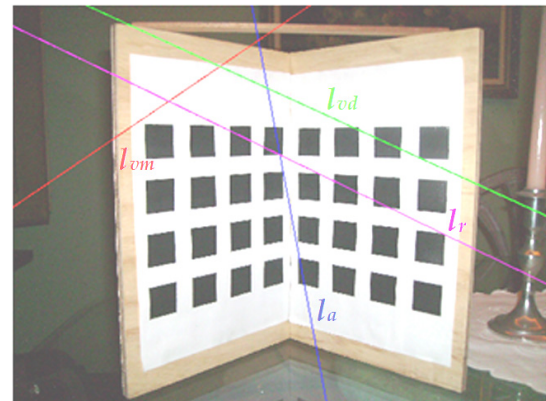


Figura 3.9: Algoritmo de 8 Pontos com ajuste de singularidade

3.5.2 Algoritmo de 8 Pontos Normalizado

Apesar da simplicidade do Algoritmo de 8 Pontos, ele sempre foi muito criticado pela sua elevada sensibilidade ao ruído gerado na escolha dos pontos homólogos. Contudo, [23] sugere que uma transformação (translação e escalamento) dos pontos homólogos do par de imagens antes da formulação das equações lineares do Algoritmo de 8 Pontos (Equação (3.12)), acarretaria em uma melhora da estabilidade da solução. Analisaremos essa estabilidade de modo a verificarmos qual transformação implicará numa maior robustez da solução.

Análise de Estabilidade do Algoritmo de 8 Pontos

O número de condição $\kappa = \lambda_{\max}/\lambda_{\min}$ é um fator importante na análise de estabilidade de um problema linear, pois ele corresponde a uma medida da sensibilidade da solução do sistema [24] às variações dos dados. A solução mais robusta possível corresponde a $\kappa \rightarrow 1$. Se κ for grande, pequenas variações nos dados (no nosso caso, posições dos pontos homólogos) podem causar mudanças significativas na solução do sistema.

O número de condição para avaliar a estabilidade de nosso problema será dado por $\kappa_{8pts} = \lambda_1/\lambda_8$, pois a matriz \mathbf{A} deve ter posto 8, e não 9. A razão principal para o fraco condicionamento da solução obtida da matriz $\mathbf{A}^T \mathbf{A}$ é a falta de homogeneidade nas coordenadas da imagem.

Suponha que os pontos homólogos \mathbf{x}_1 e \mathbf{x}_2 correspondam a $[100, 100, 1]^T$ em uma imagem de dimensão 200×200 .

Então, a linha correspondente da matriz \mathbf{A} será igual a:

$$\mathbf{l}_x = [10^4, 10^4, 10^2, 10^4, 10^4, 10^2, 10^2, 10^2, 1]. \quad (3.16)$$

A contribuição para matriz $\mathbf{A}^T \mathbf{A}$ será da forma $\mathbf{l}_x^T \mathbf{l}_x$ contendo valores entre 10^8 e 1. Assumindo que os elementos das diagonais da matriz $\mathbf{A}^T \mathbf{A}$ sejam iguais a:

$$\text{diag}(\mathbf{A}^T \mathbf{A}) = [10^8, 10^8, 10^4, 10^8, 10^8, 10^4, 10^4, 10^4, 1], \quad (3.17)$$

fazemos uso da *Propriedade de Entrelaçamento* [25], e com essa propriedade obtaremos um valor limite para o número de condição de nosso interesse.

Vamos indicar \mathbf{H}_w como submatriz principal $w \times w$ da matriz $\mathbf{A}^T \mathbf{A}$ (esquema apresentado na Figura 3.10) e $\lambda_i(H_w)$ como seu i -ésimo maior autovalor.

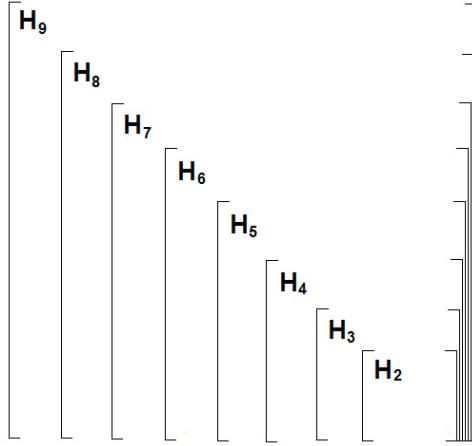


Figura 3.10: Esquema das submatrizes de $\mathbf{A}^T \mathbf{A}$ para análise de estabilidade

Logo, $\mathbf{H}_9 = \mathbf{A}^T \mathbf{A}$, e o número de condição que queremos saber para análise da estabilidade será

$$\kappa_{8pts} = \frac{\lambda_1(H_9)}{\lambda_8(H_9)}. \quad (3.18)$$

Como a soma dos autovalores será igual ao traço da matriz \mathbf{H}_2 ($\lambda_1(H_2) + \lambda_2(H_2) = 10^4 + 1$) e a matriz \mathbf{H}_2 é semidefinida positiva, podemos concluir que

$$\lambda_1(H_2) \leq 10^4 + 1. \quad (3.19)$$

A partir da *Propriedade de Entrelaçamento*, os autovalores são ordenados da seguinte forma:

$$\lambda_8(H_9) \leq \lambda_7(H_8) \leq \dots \leq \lambda_1(H_2) \leq 10^4 + 1. \quad (3.20)$$

Dessa forma, o maior autovalor de \mathbf{H}_9 não será menor do que o maior valor dentre os elementos da diagonal, ou seja,

$$\lambda_1(H_9) \geq 10^8. \quad (3.21)$$

A partir das Equações (3.20) e (3.21), podemos verificar que o número de condição para análise da estabilidade torna-se um valor alto, correspondendo a

$$\kappa_{8pts} = \frac{\lambda_1(H_9)}{\lambda_8(H_9)} \geq \frac{10^8}{10^4 + 1}. \quad (3.22)$$

Em geral, o valor de $\lambda_8(H_9)$ será muito menor do que $10^4 + 1$, implicando que o número de condição (κ_{8pts}) seja maior ainda.

Análise da Matriz de Transformação da Normalização

De modo a contornar o problema ilustrado pela Equação (3.22), os pontos das imagens devem sofrer uma transformação para diminuir o valor do número de condição κ_{spts} , e com isso melhorar a robustez do algoritmo.

Transformação de Translação

Primeiramente, vamos supor que as coordenadas dos pontos da primeira imagem (no eixo u) sejam, por exemplo: $[1001, 1002, 998, \dots]$. Caso apliquemos uma translação do valor 1000 nessas coordenadas, seus números serão agora: $[1, 2, -2, \dots]$.

Podemos verificar que, nos valores não transladados, o valor significativo na análise do algoritmo é escondido por um “*offset*” das coordenadas, no nosso exemplo de valor 1000. Pode-se notar que a parte significativa está localizada muitas vezes no terceiro ou quarto algarismo significativo do valor original da coordenada. A remoção desse “*offset*”, dado por exemplo, pela média das coordenadas no eixo correspondente (centróide dos pontos), implica numa melhora da robustez.

A translação dos pontos deslocando o centróide para a origem do sistema de coordenadas é definida por:

$$T_{trans.} = \begin{bmatrix} 1 & 0 & -\mu_u \\ 0 & 1 & -\mu_v \\ 0 & 0 & 1 \end{bmatrix},$$

onde μ_u é a média das coordenadas ao longo do eixo u

$$\mu_u = \frac{1}{N} \sum_{i=1}^N u^{(i)}$$

e μ_v é a média das coordenadas ao longo do eixo v

$$\mu_v = \frac{1}{N} \sum_{i=1}^N v^{(i)}.$$

Transformação de Escalamento

Uma outra transformação muito útil na normalização será a transformação de escalamento. Pela análise da estabilidade do Algoritmo de 8 Pontos, pode-se verificar que um escalamento das coordenadas de modo que as coordenadas homogêneas estejam na média iguais à unidade, acarretará em uma melhora significativa do número de condição da matriz $\mathbf{A}^T \mathbf{A}$. Na referência [23] há dois tipos de escalamento (isotrópico e não-isotrópico) para normalização.

- **Escalamento isotrópico:** os pontos sofrem uma mudança de escala igual para as duas coordenadas associadas aos eixos u e v , de modo que a distância média dos pontos ao centróide (desvio padrão médio) seja $\sqrt{2}$.

$$T_{esc. iso.} = \begin{bmatrix} \frac{\sqrt{2}}{\hat{\sigma}} & 0 & 0 \\ 0 & \frac{\sqrt{2}}{\hat{\sigma}} & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

onde $\hat{\sigma}$ é a média dos desvios padrão das coordenadas dos eixos u e v :

$$\hat{\sigma} = \sqrt{\frac{1}{2N} \sum_{i=1}^N (u^{(i)} - \mu_u)^2 + (v^{(i)} - \mu_v)^2}.$$

- **Escalamento não-isotrópico:** os pontos sofrem uma mudança de escala independente para cada uma das duas coordenadas associadas aos eixos u e v , de modo que o desvio padrão dos pontos ao centróide seja unitário em cada eixo.

$$T_{esc. \tilde{n}-iso.} = \begin{bmatrix} \frac{1}{\sigma_u} & 0 & 0 \\ 0 & \frac{1}{\sigma_v} & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

onde σ_u é o desvio padrão das coordenadas ao longo do eixo u ,

$$\sigma_u = \sqrt{\frac{1}{N} \sum_{i=1}^N (u^{(i)} - \mu_u)^2}$$

e σ_v é o desvio padrão das coordenadas ao longo do eixo v ,

$$\sigma_v = \sqrt{\frac{1}{N} \sum_{i=1}^N (v^{(i)} - \mu_v)^2}.$$

Devemos lembrar também que o ajuste da singularidade da matriz \mathbf{F} (ou \mathbf{E}) do Algoritmo de 8 Pontos implica num ajuste da matriz em questão, levando em consideração uma influência equivalente para todas as entradas, as coordenadas dos pares de pontos homólogos das imagens, independente de suas magnitudes. Com isso, entradas de dados com pequenas magnitudes sofrem muito mais perturbação em relação às entradas maiores. A normalização tem como objetivo equilibrar essas coordenadas na estimação de modo que o ajuste de singularidade não cause divergência da solução do sistema.

Normalização de Hartley

A normalização será dada pelo produto das matrizes de transformação $T_{Hartley} = (T_{esc} \cdot T_{trans.})$.

- Normalização isotrópica:

$$T_{Hartley\ iso.} = \begin{bmatrix} \frac{\sqrt{2}}{\hat{\sigma}} & 0 & -\frac{\sqrt{2}\mu_u}{\hat{\sigma}} \\ 0 & \frac{\sqrt{2}}{\hat{\sigma}} & -\frac{\sqrt{2}\mu_v}{\hat{\sigma}} \\ 0 & 0 & 1 \end{bmatrix}$$

- Normalização não-isotrópica:

$$T_{Hartley\ \tilde{n}-iso.} = \begin{bmatrix} \frac{1}{\sigma_u} & 0 & -\frac{\mu_u}{\sigma_u} \\ 0 & \frac{1}{\sigma_v} & -\frac{\mu_v}{\sigma_v} \\ 0 & 0 & 1 \end{bmatrix}$$

Na Figura 3.11, vemos um exemplo com os pontos de referência obtidos de uma câmera de 2 megapixels. Vemos esses pontos normalizados segundo a transformação isotrópica (Figura 3.12) e segundo a transformação não-isotrópica (Figura 3.13) de Hartley e podemos observar que após as normalizações os dados ficam concentrados na origem.

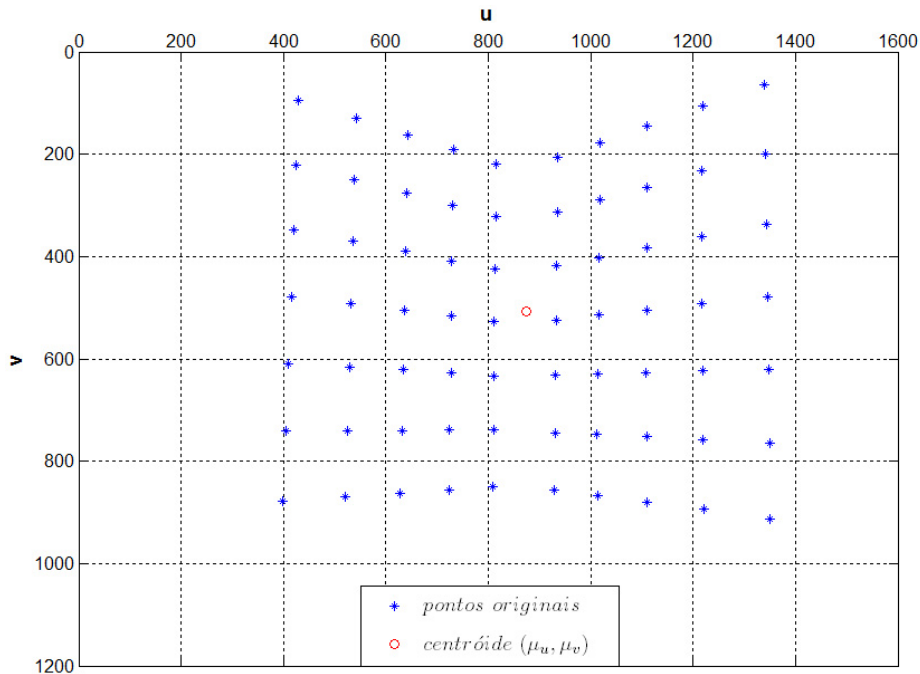


Figura 3.11: Pontos medidos de uma imagem de 2 megapixels

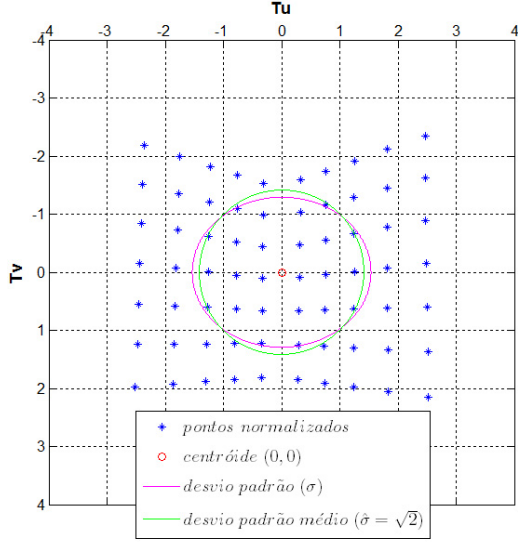


Figura 3.12: Normalização isotrópica dos pontos da Figura 3.11

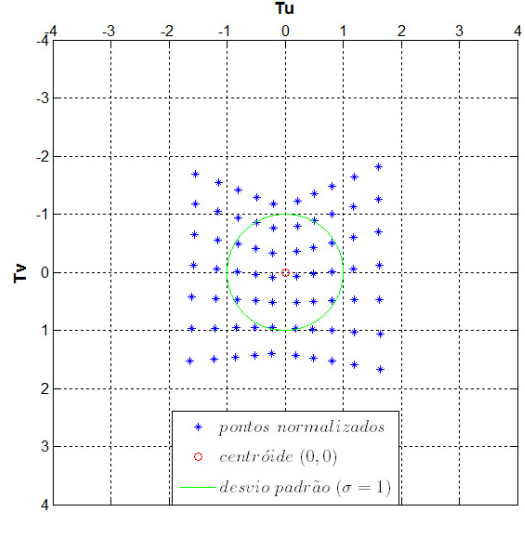


Figura 3.13: Normalização não-isotrópica dos pontos da Figura 3.11

Estimação da matriz \mathbf{F} (ou \mathbf{E}) do sistema

Na rotina do algoritmo, inicialmente, normalizamos os pontos da primeira imagem através de uma transformação de Hartley (isotrópica ou não-isotrópica) gerada a partir do conjunto desses pontos, e analogamente, realizamos o mesmo procedimento para os pontos da segunda imagem. Definimos esses pontos normalizados da seguinte forma:

$$\check{\mathbf{x}}_1 = \mathbf{T}_1 \mathbf{x}_1 \quad \check{\mathbf{x}}_2 = \mathbf{T}_2 \mathbf{x}_2 \quad (3.23)$$

Sabemos pela equação da restrição epipolar (para \mathbf{F} - Propriedade 3.4.1 e para \mathbf{E} - Propriedade 3.3.1):

$$\mathbf{x}_2^T \mathbf{F} \mathbf{x}_1 = 0.$$

Substituindo pelos pontos normalizados das equações obtemos

$$\check{\mathbf{x}}_2^T \mathbf{T}_2^{-T} \mathbf{F} \mathbf{T}_1^{-1} \check{\mathbf{x}}_1 = 0. \quad (3.24)$$

A relação da matriz \mathbf{F} (ou \mathbf{E}) com a estimada através dos pontos normalizados é dada por:

$$\check{\mathbf{F}} = \mathbf{T}_2^{-T} \mathbf{F} \mathbf{T}_1^{-1} \quad (3.25)$$

Ou seja, após realizarmos a estimação da matriz fundamental (ou essencial) do Algoritmo de 8 Pontos com os dados normalizados segundo uma normalização de Hartley, devemos fazer uma desnormalização da matriz estimada de modo a obter-

mos a matriz \mathbf{F} (ou \mathbf{E}) do sistema, além de realizarmos o ajuste final de singularidade:

$$\mathbf{F} = \mathbf{T}_2^T \check{\mathbf{F}} \mathbf{T}_1. \quad (3.26)$$

Para observarmos a melhoria do algoritmo, vejamos o exemplo mostrado no final da seção do Algoritmo de 8 Pontos. Realizamos o mesmo procedimento usando o Algoritmo de 8 Pontos Normalizado com transformação isotrópica de Hartley. Podemos ver que agora, as retas epipolares com ajuste de singularidade convergem para próximo do ponto correto.

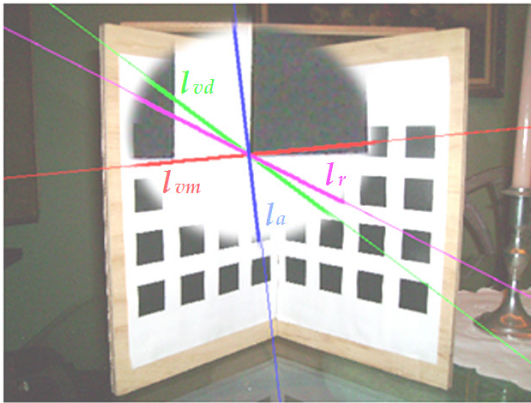


Figura 3.14: Algoritmo de 8 Pontos Normalizado sem ajuste de singularidade

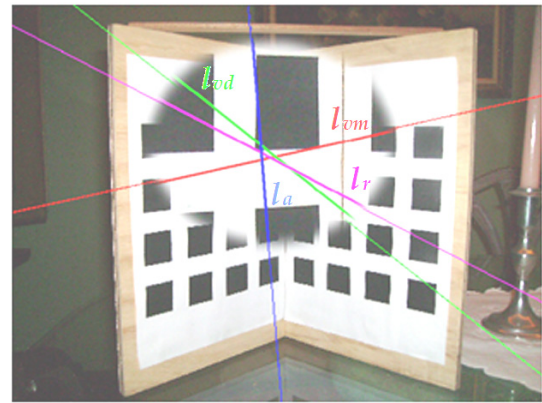


Figura 3.15: Algoritmo de 8 Pontos Normalizado com ajuste de singularidade

Capítulo 4

Estimação de Movimento

Como vimos no capítulo anterior, a matriz essencial \mathbf{E} é descrita somente em função dos parâmetros orientação e posicionamento entre as duas imagens da câmera (\mathbf{R}, \mathbf{t}) conforme a Equação (3.8). Nesse capítulo, iremos estimar esses parâmetros para diferentes tipos de movimento de uma câmera.

4.1 Movimento de Translação Pura

O primeiro tipo de movimento que iremos analisar corresponde ao movimento de translação pura [13], ou seja, matriz de rotação $\mathbf{R} = \mathbf{I}$ ($\omega = 0^\circ, \phi = 0^\circ, \rho = 0^\circ$), conforme a Figura 4.1.

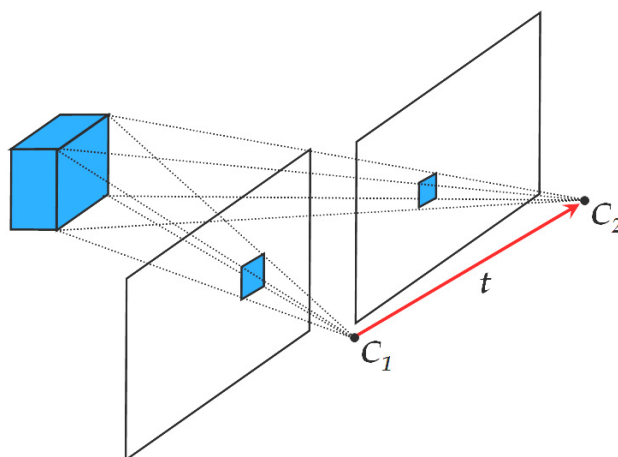


Figura 4.1: Movimento de translação pura de uma câmera

Podemos observar, na Figura 4.2, que a situação de uma câmera estacionária capturando os pontos de referência do cenário 3D e se deslocando segundo uma translação pura $-\mathbf{t}$, é semelhante ao movimento de translação pura da câmera de \mathbf{t} na Figura 4.1.

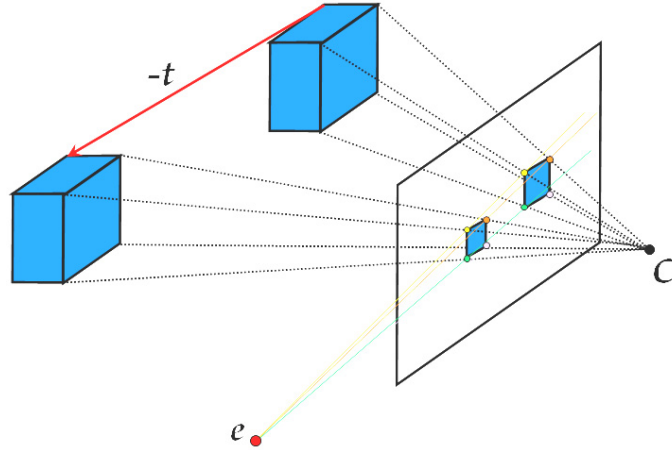


Figura 4.2: Movimento de translação pura do cenário 3D capturado por uma câmera estacionária

Na situação da câmera estacionária, podemos observar que os pontos de referência no espaço 3D formam retas paralelas ao vetor de translação \mathbf{t} e que a projeção de retas paralelas corresponde a retas concorrentes, tendo o ponto de fuga como o ponto de encontro entre si.

- A partir da propriedade dos epípolos (Propriedade 3.3.4), obtemos

$$\check{\mathbf{e}}_1 = \mathbf{t} = \mathbf{E}_{t.p.}$$

$$\check{\mathbf{e}}_2 = \overbrace{\mathbf{R}^T}^{=I} \mathbf{t} = \mathbf{t} = \mathbf{E}_{t.p.}^T$$

- A partir da propriedade das retas epipolares (Propriedade 3.3.3), obtemos

$$\check{\mathbf{l}}_2 = \mathbf{E}_{t.p.} \check{\mathbf{x}}_1 \quad \therefore \quad \check{\mathbf{l}}_2 = \check{\mathbf{e}}_1 \times \check{\mathbf{x}}_1 \quad \rightarrow \quad \check{\mathbf{l}}_2 = \check{\mathbf{l}}_1$$

$$\check{\mathbf{l}}_1 = \mathbf{E}_{t.p.}^T \check{\mathbf{x}}_2 \quad \therefore \quad \check{\mathbf{l}}_1 = \check{\mathbf{e}}_2 \times \check{\mathbf{x}}_2 \quad \rightarrow \quad \check{\mathbf{l}}_1 = \check{\mathbf{l}}_2$$

Conseqüentemente, esse ponto de fuga corresponderá ao epípolo de ambas as imagens do movimento da câmera ($\mathbf{e} = \check{\mathbf{e}}_1 = \check{\mathbf{e}}_2$), assim como essas retas projetivas serão as retas epipolares de ambas as imagens ($\mathbf{l} = \check{\mathbf{l}}_1 = \check{\mathbf{l}}_2$).

4.2 Movimento de Translação com Rotação Conhecida

O movimento que iremos analisar nessa seção corresponde ao movimento de translação com rotação previamente conhecida. Essa análise é baseada no algoritmo de Sun [26].

Sabemos que os pontos homólogos obedecem à restrição epipolar para cada caso:

Caso calibrado:

$$\check{\mathbf{x}}_2^T \mathbf{E} \check{\mathbf{x}}_1 = 0 \quad \therefore \quad \check{\mathbf{x}}_2^T ([\mathbf{t}]_{\times} \mathbf{R}) \check{\mathbf{x}}_1 = 0 \quad (4.1)$$

onde: $\check{\mathbf{x}}_1 = [u_1, v_1, 1]^T$ e $\check{\mathbf{x}}_2 = [u_2, v_2, 1]^T$.

Caso não-calibrado:

$$\tilde{\mathbf{x}}_2^T \mathbf{F} \tilde{\mathbf{x}}_1 = 0 \quad \therefore \quad \tilde{\mathbf{x}}_2^T \mathbf{K}_2^{-T} ([\mathbf{t}]_{\times} \mathbf{R}) \mathbf{K}_1^{-1} \tilde{\mathbf{x}}_1 = 0 \quad (4.2)$$

onde: $\mathbf{K}_1^{-1} \tilde{\mathbf{x}}_1 = [u_1, v_1, 1]^T$ e $\mathbf{K}_2^{-1} \tilde{\mathbf{x}}_2 = [u_2, v_2, 1]^T$.

Obtemos então a seguinte equação:

$$\begin{bmatrix} u_2 & v_2 & 1 \end{bmatrix} \begin{bmatrix} 0 & -t_z & t_y \\ t_z & 0 & -t_x \\ -t_y & t_x & 0 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \begin{bmatrix} u_1 \\ v_1 \\ 1 \end{bmatrix} = 0 \quad (4.3)$$

Desenvolvendo a equação, obtemos:

$$A_x t_x + A_y t_y + A_z t_z = 0$$

onde:

$$\begin{cases} A_x = u_1 r_{21} + v_1 r_{22} + r_{23} - u_1 v_2 r_{31} - v_1 v_2 r_{32} - v_2 r_{33} \\ A_y = u_1 u_2 r_{31} + v_1 u_2 r_{32} + u_2 r_{33} - u_1 r_{11} - v_1 r_{12} - r_{13} \\ A_z = u_1 v_2 r_{11} + v_1 v_2 r_{12} + v_2 r_{13} - u_1 u_2 r_{21} - v_1 u_2 r_{22} - u_2 r_{23} \end{cases}$$

Formemos um conjunto de equações lineares para N pares de pontos homólogos ($N \geq 2$):

$$\begin{bmatrix} A_x^{(1)} & A_y^{(1)} & A_z^{(1)} \\ \vdots & \vdots & \vdots \\ A_x^{(N)} & A_y^{(N)} & A_z^{(N)} \end{bmatrix} \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$\mathbf{A}\mathbf{t} = \mathbf{0} \quad (4.4)$$

A solução da Equação (4.4) corresponde à de um sistema homogêneo. Resolveremos o problema via a otimização linear de mínimos quadrados através da SVD, conforme o Apêndice B. Conseqüentemente, a solução \mathbf{t} será obtida exceto por um fator de escala, para o qual será aplicada à restrição $\|\mathbf{t}\|^2 = 1$.

Podemos verificar o sinal de \mathbf{t} através do cálculo da constante de profundidade da primeira imagem da estimativa de movimento (ψ_1). Este valor tem que ser obrigatoriamente positivo, pois os pontos do cenário 3D estão localizados à frente da câmera, ou seja, o componente no eixo z do sistema de coordenadas da câmera é maior do que zero.

A transformação entre os pontos em coordenadas homogêneas dos planos 2D das imagens (Equação (3.4)) é dada pela equação:

$$\psi_2 \mathbf{x}_2 = \mathbf{R}(\psi_1 \mathbf{x}_1) + \mathbf{t}, \text{ ou seja:}$$

$$\psi_2 \begin{bmatrix} u_2 \\ v_2 \\ 1 \end{bmatrix} = \psi_1 \begin{bmatrix} \bar{\mathbf{r}}_1 \\ \bar{\mathbf{r}}_2 \\ \bar{\mathbf{r}}_3 \end{bmatrix} \mathbf{x}_1 + \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix}$$

Esta equação pode ser separada em três equações lineares:

$$\psi_2 u_2 = \psi_1 (\bar{\mathbf{r}}_1 \mathbf{x}_1) + t_x \quad (4.5)$$

$$\psi_2 v_2 = \psi_1 (\bar{\mathbf{r}}_2 \mathbf{x}_1) + t_y \quad (4.6)$$

$$\psi_2 = \psi_1 (\bar{\mathbf{r}}_3 \mathbf{x}_1) + t_z \quad (4.7)$$

Dividindo a Equação (4.5) pela Equação (4.7), obtemos:

$$u_2 = \frac{\psi_1 (\bar{\mathbf{r}}_1 \mathbf{x}_1) + t_x}{\psi_1 (\bar{\mathbf{r}}_3 \mathbf{x}_1) + t_z}$$

$$\psi_1 u_2 (\bar{\mathbf{r}}_3 \mathbf{x}_1) + u_2 t_z = \psi_1 (\bar{\mathbf{r}}_1 \mathbf{x}_1) + t_x$$

$$\psi_1 (\bar{\mathbf{r}}_1 - u_2 \bar{\mathbf{r}}_3) \mathbf{x}_1 = u_2 t_z - t_x$$

$$\psi_1 = \frac{u_2 t_z - t_x}{(\bar{\mathbf{r}}_1 - u_2 \bar{\mathbf{r}}_3) \mathbf{x}_1} \quad (4.8)$$

Dividindo a Equação (4.6) pela Equação (4.7), obtemos:

$$\psi_1 = \frac{v_2 t_z - t_y}{(\bar{\mathbf{r}}_2 - v_2 \bar{\mathbf{r}}_3) \mathbf{x}_1}. \quad (4.9)$$

Com isso, devemos escolher uma das duas equações de ψ_1 e verificar qual dos dois vetores de translação ($\pm \mathbf{t}$) corresponde a $\psi_1 > 0$.

4.3 Movimento de Translação e Rotação

Esse é o último e mais completo movimento de nosso trabalho, movimento de rotação e translação desconhecidos [20]. Na prova da Propriedade 3.3.5, dissemos que existe uma matriz de rotação \mathbf{R}_* , que transforma o vetor de translação \mathbf{t} em

$$\mathbf{a} = [0, 0, \|\mathbf{t}\|]^T. \quad (4.10)$$

Além disso, dissemos que a partir do propriedade do operador matriz de produto vetorial [20], obtemos:

$$[\mathbf{t}]_{\times} = \mathbf{R}_*^T [\mathbf{a}]_{\times} \mathbf{R}_*. \quad (4.11)$$

Dada uma matriz de rotação de Givens com ângulo $\rho = +\frac{\pi}{2}$ no eixo z :

$$\mathbf{R}_z \left(+\frac{\pi}{2}\right) = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (4.12)$$

podemos renomear $[\mathbf{a}]_{\times}$ de modo a obtermos a Equação (4.13):

$$\begin{aligned} [\mathbf{a}]_{\times} &= \mathbf{R}_z \left(+\frac{\pi}{2}\right) \underbrace{\mathbf{R}_z^T \left(+\frac{\pi}{2}\right) [\mathbf{a}]_{\times}} \\ [\mathbf{a}]_{\times} &= \mathbf{R}_z \left(+\frac{\pi}{2}\right) \begin{bmatrix} \|\mathbf{t}\| & & \\ & \|\mathbf{t}\| & \\ & & 0 \end{bmatrix}. \end{aligned} \quad (4.13)$$

Na equação da matriz essencial, $\mathbf{E} = [\mathbf{t}]_{\times} \mathbf{R}$, podemos substituir $[\mathbf{t}]_{\times}$ pelas Equações (4.10) e (4.11) e obter

$$\mathbf{E} = [\mathbf{t}]_{\times} \mathbf{R} = \mathbf{R}_*^T \mathbf{R}_z \left(+\frac{\pi}{2}\right) \begin{bmatrix} \|\mathbf{t}\| & & \\ & \|\mathbf{t}\| & \\ & & 0 \end{bmatrix} \mathbf{R}_* \mathbf{R}. \quad (4.14)$$

Aplicando a SVD de \mathbf{E} , obtemos:

$$\mathbf{E} = \mathbf{U}\mathbf{S}_E\mathbf{V}^T, \quad (4.15)$$

onde:

$$\mathbf{U} = \mathbf{R}_*^T \mathbf{R}_z \left(+\frac{\pi}{2}\right), \quad \mathbf{S}_E = \begin{bmatrix} \|\mathbf{t}\| & & \\ & \|\mathbf{t}\| & \\ & & 0 \end{bmatrix}, \quad \mathbf{V}^T = \mathbf{R}_* \mathbf{R}.$$

Sendo assim, primeiramente, vamos extrair a matriz de rotação \mathbf{R} .

Sabemos que $\mathbf{V}^T = \mathbf{R}_* \mathbf{R}$. Daqui, temos que extrair a matriz \mathbf{R}_* , logo

$$\begin{aligned} \mathbf{R} &= \overbrace{\mathbf{R}_*^T \mathbf{R}_z \left(+\frac{\pi}{2}\right)}^{\mathbf{U}} \mathbf{R}_z^T \left(+\frac{\pi}{2}\right) \overbrace{\mathbf{R}_* \mathbf{R}}^{\mathbf{V}^T} \\ \mathbf{R} &= \mathbf{U} \mathbf{R}_z^T \left(+\frac{\pi}{2}\right) \mathbf{V}^T. \end{aligned} \quad (4.16)$$

Agora para obtermos a matriz do vetor de translação $[\mathbf{t}]_\times$, devemos extrair \mathbf{R} da Equação (4.14) referente à matriz essencial \mathbf{E} :

$$[\mathbf{t}]_\times = \mathbf{R}_*^T \mathbf{R}_z \left(+\frac{\pi}{2}\right) \mathbf{S}_E \mathbf{R}_*, \quad (4.17)$$

$$[\mathbf{t}]_\times = \overbrace{\mathbf{R}_*^T \mathbf{R}_z \left(+\frac{\pi}{2}\right)}^{\mathbf{U}} \mathbf{S}_E \mathbf{R}_z \left(+\frac{\pi}{2}\right) \overbrace{\mathbf{R}_*^T \left(+\frac{\pi}{2}\right) \mathbf{R}_*}_{\mathbf{U}^T},$$

$$[\mathbf{t}]_\times = \mathbf{U} \mathbf{S}_E \mathbf{R}_z \left(+\frac{\pi}{2}\right) \mathbf{U}^T. \quad (4.18)$$

Existe uma solução para a matriz de valores singulares de \mathbf{E} da forma:

$$\mathbf{S}_E = \begin{bmatrix} -\|\mathbf{t}\| & & \\ & -\|\mathbf{t}\| & \\ & & 0 \end{bmatrix}. \quad (4.19)$$

Essa solução corresponde à mesma análise vista anteriormente, só que com uma matriz de rotação de Givens com ângulo $\rho = -\frac{\pi}{2}$ no eixo z :

$$\mathbf{R}_z \left(-\frac{\pi}{2}\right) = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (4.20)$$

A solução do par $(\mathbf{R}, [\mathbf{t}]_\times)$ corresponde a

$$\mathbf{R} = \mathbf{U} \mathbf{R}_z^T \left(-\frac{\pi}{2}\right) \mathbf{V}^T. \quad (4.21)$$

$$[\mathbf{t}]_{\times} = \mathbf{U} \mathbf{S}_E \mathbf{R}_z \left(-\frac{\pi}{2}\right) \mathbf{U}^T. \quad (4.22)$$

Cabe observarmos a seguinte afirmação:

Toda matriz de rotação é ortogonal, contudo, nem toda matriz ortogonal é de rotação, ou seja, para uma matriz ortogonal ser de rotação, o seu determinante deve ser igual a +1.

Logo, devemos fazer essa verificação nas matrizes \mathbf{U} e \mathbf{V} da SVD de \mathbf{E} . Caso alguma delas tenha o seu determinante igual a -1 , temos de obter as matrizes \mathbf{U} e \mathbf{V} da SVD de $-\mathbf{E}$.

Como a matriz \mathbf{E} proporciona as 2 soluções, assim também, devido à propriedade do posto incompleto, a matriz $-\mathbf{E}$ proporciona mais 2 soluções. Logo, como não é possível uma matriz de rotação $-\mathbf{R}$, pois o determinante seria igual a -1 , a solução corresponde a $-\mathbf{t}]_{\times}$. A solução para isso corresponde a inverter o ângulo da matriz de rotação \mathbf{R}_z mantendo \mathbf{U} como sendo uma matriz de rotação. Isso implica que a multiplicação:

$$\mathbf{S}_E \mathbf{R}_z \left(\pm\frac{\pi}{2}\right) \mathbf{R}_z^T \left(\mp\frac{\pi}{2}\right) = -\mathbf{S}_E.$$

Caso 1:

$$\mathbf{U} = \mathbf{R}_*^T \mathbf{R}_z \left(+\frac{\pi}{2}\right), \quad \mathbf{S}_E = \begin{bmatrix} \|\mathbf{t}\| & & \\ & \|\mathbf{t}\| & \\ & & 0 \end{bmatrix}, \quad \mathbf{V}^T = \mathbf{R}_* \mathbf{R},$$

$$\mathbf{R} = \mathbf{U} \mathbf{R}_z^T \left(+\frac{\pi}{2}\right) \mathbf{V}^T, \quad (4.23)$$

$$[\mathbf{t}]_{\times} = \mathbf{U} \mathbf{S}_E \mathbf{R}_z \left(-\frac{\pi}{2}\right) \mathbf{U}^T. \quad (4.24)$$

Caso 2:

$$\mathbf{U} = \mathbf{R}_*^T \mathbf{R}_z \left(-\frac{\pi}{2}\right), \quad \mathbf{S}_E = \begin{bmatrix} -\|\mathbf{t}\| & & \\ & -\|\mathbf{t}\| & \\ & & 0 \end{bmatrix}, \quad \mathbf{V}^T = \mathbf{R}_* \mathbf{R},$$

$$\mathbf{R} = \mathbf{U} \mathbf{R}_z^T \left(-\frac{\pi}{2}\right) \mathbf{V}^T, \quad (4.25)$$

$$[\mathbf{t}]_{\times} = \mathbf{U} \mathbf{S}_E \mathbf{R}_z \left(+\frac{\pi}{2}\right) \mathbf{U}^T. \quad (4.26)$$

Tabela 4.1: Soluções de $(\mathbf{R}, [\mathbf{t}]_{\times})$	
matriz de rotação \mathbf{R}	matriz de translação $[\mathbf{t}]_{\times}$
$\mathbf{UR}_z^T(+\frac{\pi}{2})\mathbf{V}^T$	$\mathbf{US}_E\mathbf{R}_z(+\frac{\pi}{2})\mathbf{U}^T$
$\mathbf{UR}_z^T(-\frac{\pi}{2})\mathbf{V}^T$	$\mathbf{US}_E\mathbf{R}_z(-\frac{\pi}{2})\mathbf{U}^T$
$\mathbf{UR}_z^T(+\frac{\pi}{2})\mathbf{V}^T$	$\mathbf{US}_E\mathbf{R}_z(-\frac{\pi}{2})\mathbf{U}^T$
$\mathbf{UR}_z^T(-\frac{\pi}{2})\mathbf{V}^T$	$\mathbf{US}_E\mathbf{R}_z(+\frac{\pi}{2})\mathbf{U}^T$

Logo, teremos quatro possíveis soluções de $(\mathbf{R}, [\mathbf{t}]_{\times})$, listadas na Tabela 4.1.

A Figura 4.3 apresenta a representação gráfica dessas quatro soluções, onde cada configuração é composta dos dois modelos *pin-hole* frontal da câmera, antes e após o movimento com os dados de cada solução.

Podemos verificar na Figura 4.3 que somente uma das configurações corresponde a resposta do problema, pois ela possui as constantes de profundidade ψ_1 e ψ_2 positivas (configuração (1)). Para obter essas constantes, podemos usar as mesmas fórmulas das Equações (4.8) e (4.9).

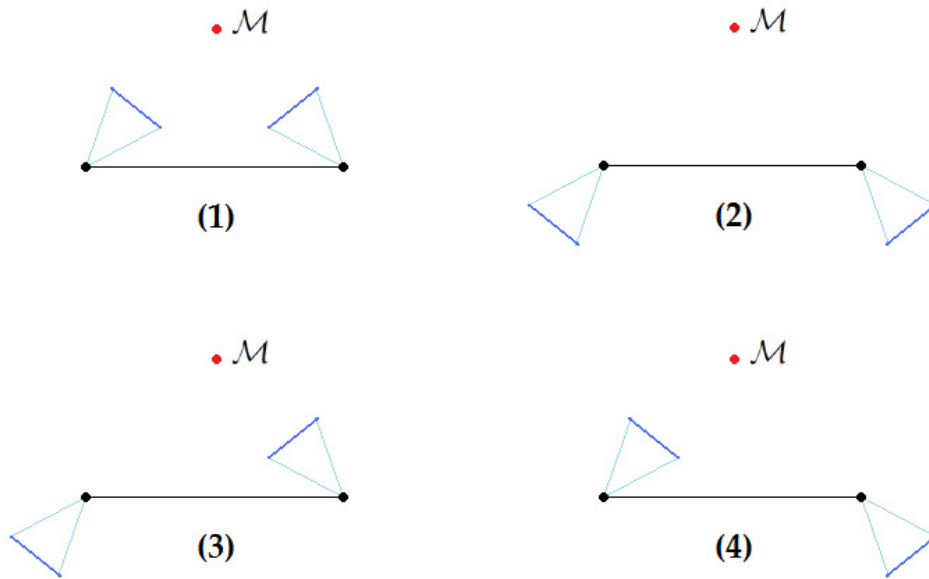


Figura 4.3: Configurações obtidas da estimativa de (\mathbf{R}, \mathbf{t})

4.4 Cálculo do Fator de Escala do Vetor \mathbf{t}

Como podemos observar pelo estudo dos movimentos nas Seções 4.1, 4.2 e 4.3, em todos os casos, a estimativa do vetor de translação é obtida exceto por um fator de escala. Demonstraremos a seguir uma forma linear desenvolvida em nosso trabalho para a obtenção desse fator de escala.

Primeiramente, devemos fazer um cálculo de reconstrução 3D de dois pontos do cenário 3D, associados aos seus respectivos pares de pontos homólogos das duas imagens. Para tal, aplicaremos a técnica de triangulação linear [13].

Sejam os pontos \mathbf{x}_1 e \mathbf{x}_2 as projeções nos dois planos das imagens de um ponto \mathbf{X} do cenário 3D normalizados em coordenadas homogêneas.

$$\mathbf{x}_1 = [u_1, v_1, 1]^T \quad \mathbf{x}_1 = \mathbf{P}_1 \mathbf{X} \quad (4.27)$$

$$\mathbf{x}_2 = [u_2, v_2, 1]^T \quad \mathbf{x}_2 = \mathbf{P}_2 \mathbf{X} \quad (4.28)$$

Podemos afirmar com relação à primeira imagem que:

$$\mathbf{x}_1 \times (\mathbf{P}_1 \mathbf{X}) = \mathbf{0} \quad \rightarrow \quad [\mathbf{x}_1]_{\times} (\mathbf{P}_1 \mathbf{X}) = \mathbf{0}$$

$$\begin{bmatrix} 0 & -1 & v_1 \\ 1 & 0 & -u_1 \\ -v_1 & u_1 & 0 \end{bmatrix} \begin{bmatrix} \bar{\mathbf{P}}_1^1 X \\ \bar{\mathbf{P}}_1^2 X \\ \bar{\mathbf{P}}_1^3 X \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \text{onde } \bar{\mathbf{P}}_1^n = [\mathbf{P}_1(n1), \mathbf{P}_1(n2), \mathbf{P}_1(n3), \mathbf{P}_1(n4)]$$

Assim, obtemos as seguintes equações lineares:

$$\begin{cases} u_1(\bar{\mathbf{P}}_1^3 \mathbf{X}) - (\bar{\mathbf{P}}_1^1 \mathbf{X}) = 0 \\ v_1(\bar{\mathbf{P}}_1^3 \mathbf{X}) - (\bar{\mathbf{P}}_1^2 \mathbf{X}) = 0 \\ u_1(\bar{\mathbf{P}}_1^2 \mathbf{X}) - v_1(\bar{\mathbf{P}}_1^1 \mathbf{X}) = 0 \end{cases} \quad (4.29)$$

Analogamente, podemos obter as equações similares para análise da segunda imagem:

$$\mathbf{x}_2 \times (\mathbf{P}_2 \mathbf{X}) = \mathbf{0} \quad \rightarrow \quad [x_2]_{\times} (\mathbf{P}_2 \mathbf{X}) = \mathbf{0}$$

$$\begin{cases} u_2(\bar{\mathbf{P}}_2^3 \mathbf{X}) - (\bar{\mathbf{P}}_2^1 \mathbf{X}) = 0 \\ v_2(\bar{\mathbf{P}}_2^3 \mathbf{X}) - (\bar{\mathbf{P}}_2^2 \mathbf{X}) = 0 \\ u_2(\bar{\mathbf{P}}_2^2 \mathbf{X}) - v_2(\bar{\mathbf{P}}_2^1 \mathbf{X}) = 0 \end{cases} \quad (4.30)$$

Juntando as equações linearmente independentes de (4.29) e (4.30) referentes às

duas imagens, formulamos o sistema:

$$\begin{bmatrix} u_1 \bar{\mathbf{P}}_1^3 - \bar{\mathbf{P}}_1^1 \\ v_1 \bar{\mathbf{P}}_1^3 - \bar{\mathbf{P}}_1^2 \\ u_2 \bar{\mathbf{P}}_2^3 - \bar{\mathbf{P}}_2^1 \\ v_2 \bar{\mathbf{P}}_2^3 - \bar{\mathbf{P}}_2^2 \end{bmatrix} \begin{bmatrix} X_x \\ X_y \\ X_z \\ X_w \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad (4.31)$$

$$\mathbf{A}\mathbf{X} = 0 \quad (4.32)$$

Agrupando essas equações teremos um sistema linear homogêneo conforme a Equação (4.32), que será resolvido através de otimização de mínimos quadrados via SVD, conforme o Apêndice B. A solução obtida de \mathbf{X} será normalizada de tal modo que a coordenada X_w seja igual a 1 e, então, as outras três coordenadas corresponderão à estimativa das coordenadas do ponto 3D.

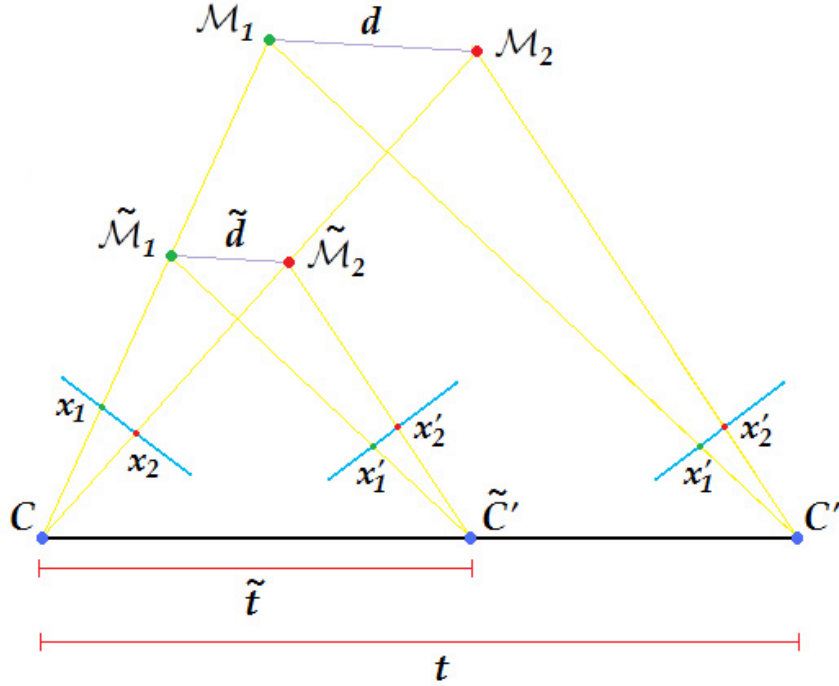


Figura 4.4: Esquema gráfico do cálculo do fator de escala

Podemos ver na Figura 4.4 os três modelos *pin-hole* frontal da câmera: câmera antes do movimento (centro óptico: C), câmera com movimento com o vetor de translação exceto por um fator de escala (centro óptico: \tilde{C}') e câmera com movimento com o vetor de translação com fator de escala (centro óptico: C'). A partir das projeções nas imagens de dois pontos do cenário 3D (\mathcal{M}_1 e \mathcal{M}_2), obtemos os pontos $\tilde{\mathcal{M}}_1$ e $\tilde{\mathcal{M}}_2$ através da triangulação linear com o vetor de translação exceto por um fator de escala.

Podemos observar na Figura 4.4 que os pontos $\mathcal{M}_1\mathcal{M}_2CC'$ e $\tilde{\mathcal{M}}_1\tilde{\mathcal{M}}_2C\tilde{C}'$ formam

dois quadriláteros semelhantes. Conseqüentemente, a razão entre os módulos dos vetores de translação (t por \tilde{t}) é igual à razão entre as distâncias euclidianas dos pontos 3D (d por \tilde{d}). Dessa forma, o fator de escala corresponderá a:

$$\zeta = \frac{d}{\tilde{d}} \quad (4.33)$$

Capítulo 5

Algoritmos, Testes e Resultados

Nesse capítulo, exporemos as idéias contidas nos Capítulos 2, 3 e 4. Com isso, será gerada uma estrutura em diagramas de blocos dos algoritmos. Estes serão utilizados na estimação de movimentos de uma câmera em testes com dados sintéticos e reais.

5.1 Algoritmos

Inicialmente, como sabemos que qualquer imagem capturada por uma câmera corresponde a uma imagem não calibrada, o primeiro algoritmo que iremos mostrar corresponde ao algoritmo de calibração, mostrado no diagrama da Figura 5.1.

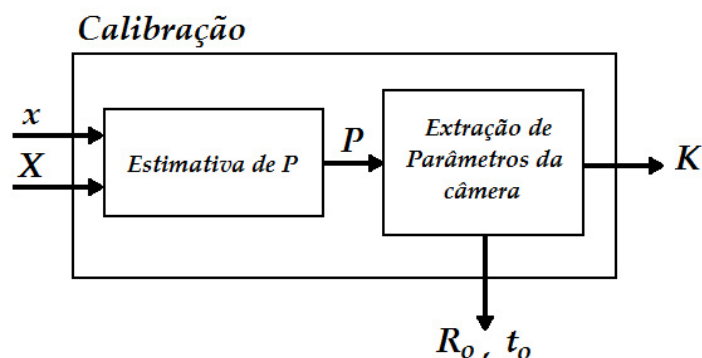


Figura 5.1: Algoritmo de calibração

Este algoritmo é baseado nos conceitos expostos no Capítulo 2. Os pontos do *grid* 3D de referência \mathbf{X} e suas respectivas projeções no plano da imagem \mathbf{x} correspondem às variáveis de entrada do algoritmo. Empregando a técnica de otimização linear com restrição exposta no Apêndice C, obtemos a estimativa da matriz de projeção \mathbf{P} e através das equações das Tabelas 2.1 (para um modelo simples) ou 2.2 (para um modelo geral), obtemos os parâmetros da câmera, cujos parâmetros intrínsecos formarão a matriz de calibração \mathbf{K} .

O próximo algoritmo que iremos expor corresponde ao algoritmo de estimação da matriz essencial \mathbf{E} , mostrado no diagrama da Figura 5.2.

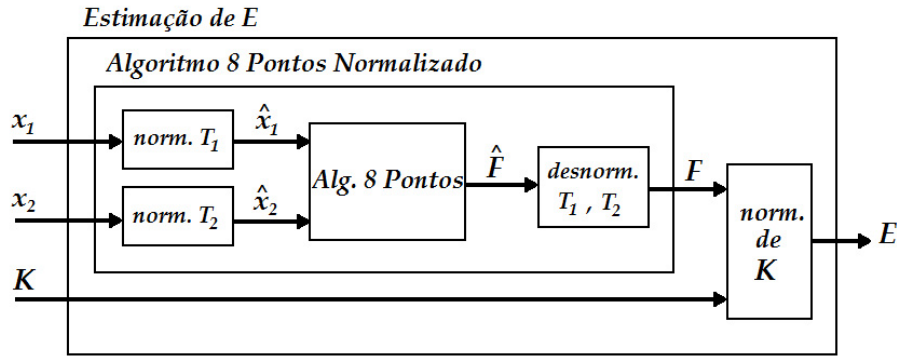


Figura 5.2: Algoritmo de estimação da matriz essencial \mathbf{E}

Este algoritmo é baseado nos conceitos expostos no Capítulo 3. De posse de duas imagens com uma mudança de posicionamento entre si, os pares de pontos homólogos das imagens \mathbf{x}_1 e \mathbf{x}_2 correspondem às variáveis de entrada do algoritmo. Conforme veremos nos primeiros testes, utilizaremos a técnica linear do Algoritmo de 8 Pontos com normalização para estimarmos a matriz fundamental \mathbf{F} . De posse da matriz de calibração \mathbf{K} , podemos transformar a matriz fundamental \mathbf{F} na matriz essencial \mathbf{E} , segundo a relação:

$$\mathbf{E} = \mathbf{K}^T \mathbf{F} \mathbf{K}$$

Finalmente, o último algoritmo corresponde ao da estimação dos parâmetros do movimento da câmera, conforme o diagrama da Figura 5.3.

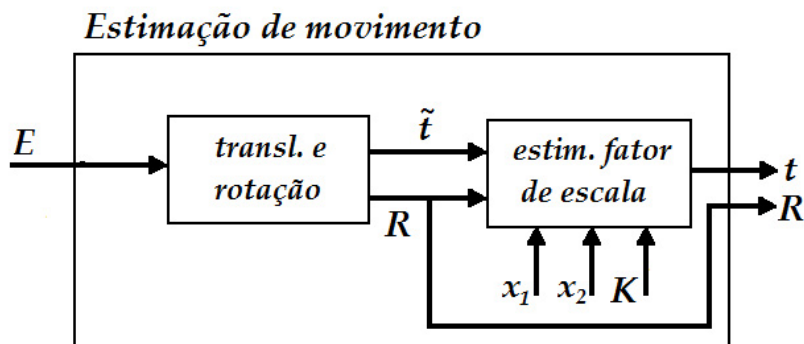


Figura 5.3: Algoritmo de estimação dos parâmetros de movimento

O algoritmo da Figura 5.3 é baseado nos conceitos expostos no Capítulo 4. De posse da matriz essencial \mathbf{E} , extraímos os parâmetros de rotação e translação segundo a Tabela 4.1. Depois, obtemos a estimativa do fator de escala do vetor de translação através da reconstrução 3D dos pontos e do valor da medida da distância euclidiana entre estes pontos.

5.2 Testes e Resultados

Após estruturados os algoritmos, iremos testá-los de modo a provar sua eficiência. Contudo, esses testes devem ser feitos com cautela, de modo que possamos identificar as deficiências do algoritmo. Nossos primeiros testes foram feitos com dados sintéticos, ou seja, a câmera e os pontos são todos virtuais. Podemos dessa forma, controlar todos os parâmetros dos testes de modo a verificarmos as falhas dos algoritmos. Depois disso, partimos para os testes com imagens reais.

5.2.1 Dados Sintéticos

Fizemos uso da *toolbox* EGT [27] de Matlab para auxiliarmos na visualização destes testes. Através dela, podemos observar os movimentos dessas câmeras virtuais em 3D. Na Figura 5.4, pode ser visto um exemplo do uso da *toolbox*, onde a câmera de referência C_0 é movimentada para câmera C_1 segundo os parâmetros: $\mathbf{t} = [30, 0, 0]^T$ (cm), $\omega = 0^\circ$, $\phi = -25^\circ$, $\rho = 0^\circ$. A Figura 5.5 mostra a projeção dos pontos 3D do *grid* virtual no plano da imagem da câmera C_0 , e a Figura 5.6 a projeção dos mesmos na imagem da câmera C_1 .

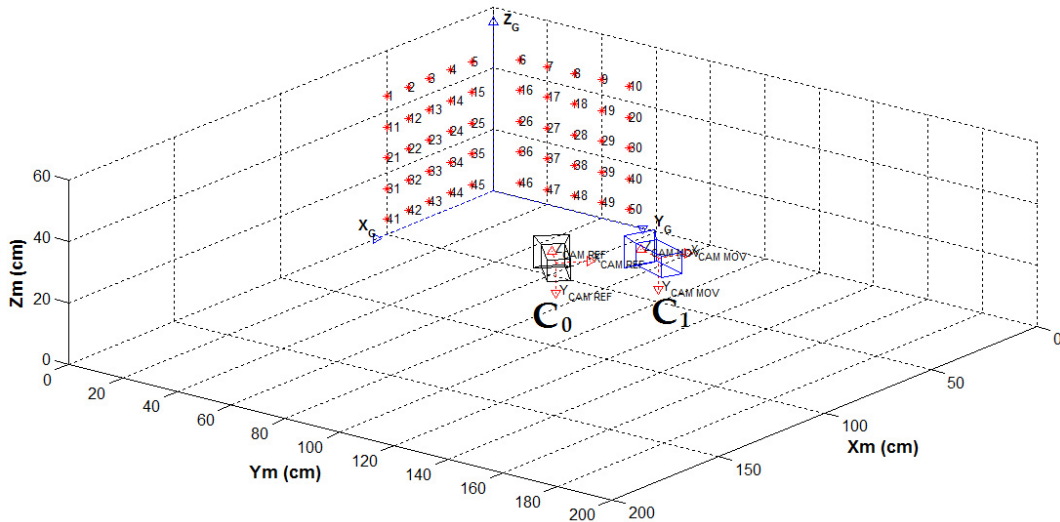


Figura 5.4: Exemplo de uso da *toolbox* EGT

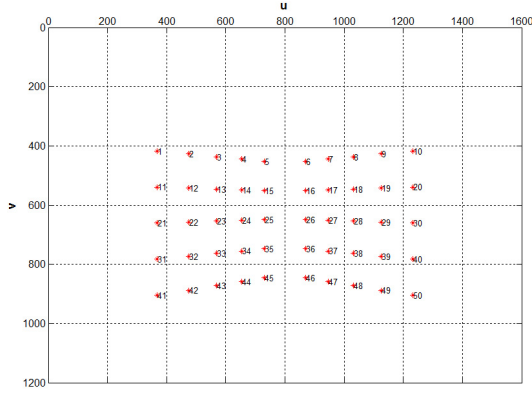


Figura 5.5: Imagem da câmera de referência \mathbf{C}_0

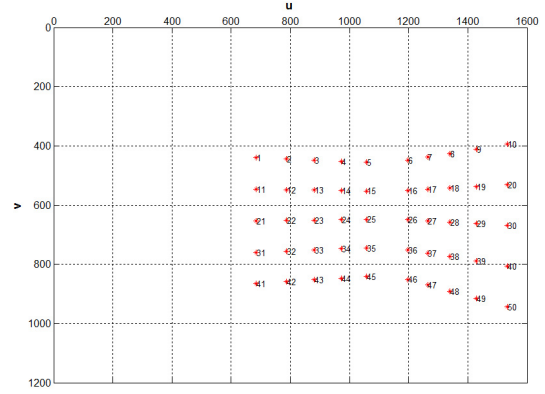


Figura 5.6: Imagem da câmera em movimento \mathbf{C}_1

De modo a ficarem o mais próximo possível dos testes com imagens reais, o que será visto posteriormente neste capítulo, os testes com dados sintéticos apresentam as seguintes características:

- **Resolução das imagens:** 1600×1200 pixels (câmera com sensor de 2 megapixels),

- **Matriz de calibração:** $\mathbf{K} = \begin{bmatrix} 1400 & 0 & 800 \\ 0 & 1400 & 600 \\ 0 & 0 & 1 \end{bmatrix}$,

- **Distância da câmera aos pontos 3D:** aproximadamente 150 cm.

Avaliação da Estimação da Matriz \mathbf{F}

Conforme analisamos matematicamente no Capítulo 4, foi feita uma comparação de desempenho dos Algoritmos de 8 Pontos e de 8 Pontos Normalizado através das duas normalizações propostas por Hartley [23] em função da estimativa da matriz fundamental \mathbf{F} .

Nas três análises dessa seção, foram utilizadas aproximadamente 50 pares de imagens virtuais com movimentos de translação e rotação distintos.

Na primeira análise, referente à Figura 5.7, estimamos a matriz fundamental \mathbf{F} de todos os pares utilizando as três técnicas lineares com 8, 18, 32, 50, 72 e 98 pontos homólogos por par. A partir da matriz \mathbf{A} dos sistemas homogêneos gerados pelos algoritmos, obtivemos os números de condição (κ_{8pts}) médios formado pelos autovalores da matriz $\mathbf{A}^T \mathbf{A}$, conforme descrito na Seção 3.2 referente à análise de estabilidade.

Podemos observar na Figura 5.7, que o uso da normalização dos dados de entrada implica em uma redução significativa do valor do número de condição (κ_{8pts}), e que o aumento do número de pontos não influencia na variação desse valor.

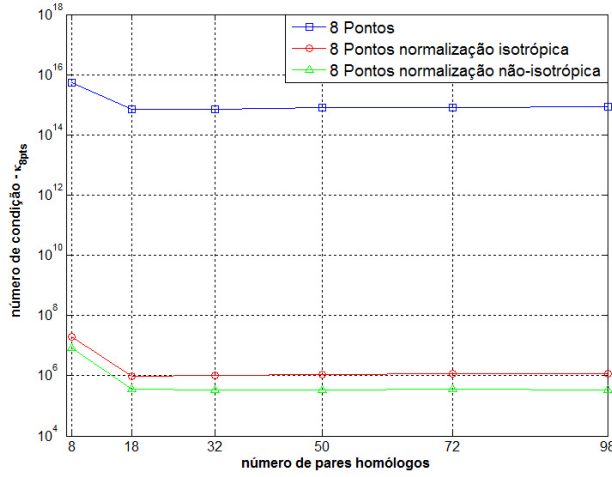


Figura 5.7: Análise do valor do número de condição κ_{8pts}

No gráfico da Figura 5.8, verificamos a influência do aumento do número de pontos homólogos usados na estimativa da matriz fundamental \mathbf{F} sobre o erro médio da distância geométrica dos pontos das duas imagens, associadas às suas respectivas retas epipolares formadas por:

$$\begin{aligned} \mathbf{x}_1 &\rightarrow \mathbf{l}_1 = \mathbf{F}^T \mathbf{x}_2 \\ \mathbf{x}_2 &\rightarrow \mathbf{l}_2 = \mathbf{F} \mathbf{x}_1 \end{aligned}$$

Esse erro corresponde à menor distância do ponto à reta e é definida como:

$$d(\mathbf{x}, \mathbf{l}) = \frac{\|\mathbf{x} \cdot \mathbf{l}\|}{\sqrt{l_x^2 + l_y^2}}$$

Podemos constatar a partir da Figura 5.8, que a normalização dos dados de entrada reduz o erro médio dos pontos às retas epipolares, e que o aumento do número de pontos não influencia na variação desse erro.

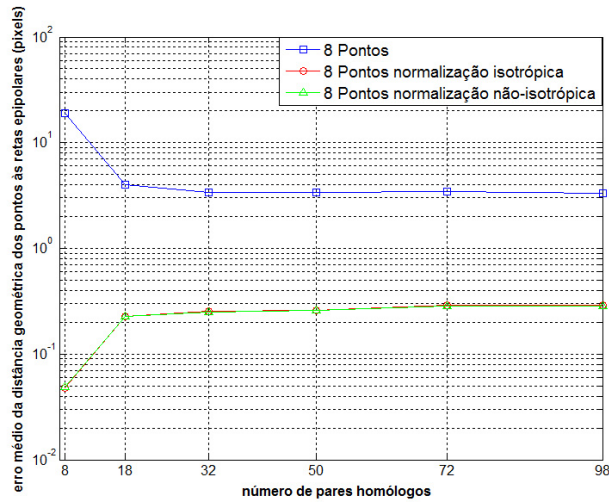


Figura 5.8: Análise da influência do número de pontos

De modo a verificarmos a eficiência do Algoritmo de 8 Pontos (Figura 5.9), fixamos as imagens com 50 pontos homólogos e inserimos uma componente de ruído do tipo gaussiana com média nula e desvio padrão variando de 0 a 5 pixels nas coordenadas dos pontos das imagens antes da discretização. Estimamos a matriz fundamental \mathbf{F} para cada nível de ruído e medimos o erro médio da distância geométrica dos pontos a suas referentes retas epipolares, conforme mostrado no teste anterior.

Podemos verificar na Figura 5.9, que o aumento do ruído nas coordenadas dos pontos implica, para as estimações através do Algoritmo de 8 Pontos, em um elevado crescimento dos erros associados aos pontos e, para as estimações através dos algoritmos normalizados, em um crescimento mais suave desses erros.

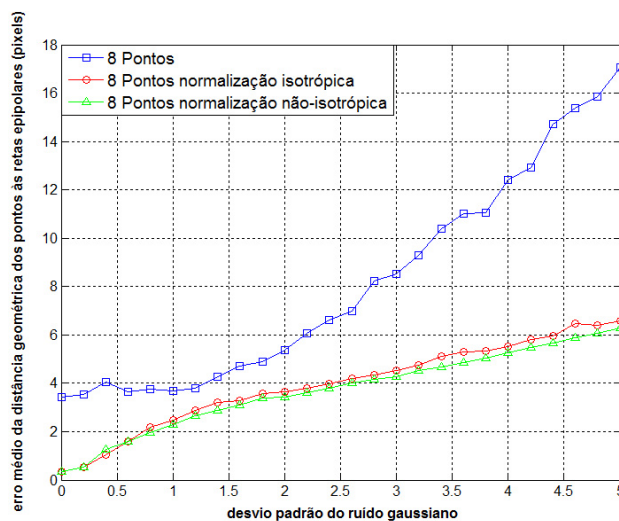


Figura 5.9: Análise da influência do ruído

Avaliação da Estimação de Movimento

Definindo 50 pontos de referência no cenário 3D, fixamos a câmera em uma condição inicial de modo que os pontos de referência fiquem em média a 150 cm de profundidade do centro óptico e os pontos projetados fiquem centralizados no plano da imagem. A partir dessa condição inicial, realizamos movimentos de translação pura ($\mathbf{R} = \mathbf{I}$) em 26 direções diferentes referentes aos vetores unitários de translação conforme a Tabela 5.1.

Para cada direção, analisamos o movimento com o módulo do vetor de translação variando de 1 a 30 cm. Nessa análise, primeiramente, estimamos no algoritmo o vetor de translação unitário (sem fator de escala) e, depois, estimamos o fator de escala correspondendo ao módulo do vetor de translação do movimento da câmera.

Tabela 5.1: Vetores unitários de translação (em cm)

Translação	eixo x	eixo y	eixo z	Translação	eixo x	eixo y	eixo z
t_1	1	0	0	t_{14}	$\frac{\sqrt{2}}{2}$	0	$-\frac{\sqrt{2}}{2}$
t_2	0	1	0	t_{15}	0	$\frac{\sqrt{2}}{2}$	$-\frac{\sqrt{2}}{2}$
t_3	0	0	1	t_{16}	$-\frac{\sqrt{2}}{2}$	$-\frac{\sqrt{2}}{2}$	0
t_4	-1	0	0	t_{17}	$-\frac{\sqrt{2}}{2}$	0	$-\frac{\sqrt{2}}{2}$
t_5	0	-1	0	t_{18}	0	$-\frac{\sqrt{2}}{2}$	$-\frac{\sqrt{2}}{2}$
t_6	0	0	-1	t_{19}	$\frac{\sqrt{3}}{3}$	$\frac{\sqrt{3}}{3}$	$\frac{\sqrt{3}}{3}$
t_7	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{2}}{2}$	0	t_{20}	$-\frac{\sqrt{3}}{3}$	$\frac{\sqrt{3}}{3}$	$\frac{\sqrt{3}}{3}$
t_8	$\frac{\sqrt{2}}{2}$	0	$\frac{\sqrt{2}}{2}$	t_{21}	$\frac{\sqrt{3}}{3}$	$-\frac{\sqrt{3}}{3}$	$\frac{\sqrt{3}}{3}$
t_9	0	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{2}}{2}$	t_{22}	$\frac{\sqrt{3}}{3}$	$\frac{\sqrt{3}}{3}$	$-\frac{\sqrt{3}}{3}$
t_{10}	$-\frac{\sqrt{2}}{2}$	$\frac{\sqrt{2}}{2}$	0	t_{23}	$-\frac{\sqrt{3}}{3}$	$-\frac{\sqrt{3}}{3}$	$\frac{\sqrt{3}}{3}$
t_{11}	$-\frac{\sqrt{2}}{2}$	0	$\frac{\sqrt{2}}{2}$	t_{24}	$-\frac{\sqrt{3}}{3}$	$\frac{\sqrt{3}}{3}$	$-\frac{\sqrt{3}}{3}$
t_{12}	$\frac{\sqrt{2}}{2}$	$-\frac{\sqrt{2}}{2}$	0	t_{25}	$\frac{\sqrt{3}}{3}$	$-\frac{\sqrt{3}}{3}$	$-\frac{\sqrt{3}}{3}$
t_{13}	0	$-\frac{\sqrt{2}}{2}$	$\frac{\sqrt{2}}{2}$	t_{26}	$-\frac{\sqrt{3}}{3}$	$-\frac{\sqrt{3}}{3}$	$-\frac{\sqrt{3}}{3}$

Na Figura 5.10, analisamos a estimação do vetor unitário de translação do movimento de translação pura. Comparamos nos gráficos da figura os erros médios dos vetores unitários estimados (correspondendo ao valor médio dos módulos dos vetores erros obtidos através da subtração vetorial entre o vetor unitário estimado da translação e o vetor unitário da translação do movimento) em função do módulo do vetor de translação do movimento da câmera para cada uma das três técnicas.

Podemos observar que com o aumento do módulo do vetor de translação do movimento para todas as técnicas, o erro médio tende sempre a diminuir. Além disso, os algoritmos normalizados apresentaram sempre os menores erros sendo que, a partir do módulo do vetor de translação do movimento de aproximadamente 10 cm, esses erros ficaram abaixo de 0,1 cm.

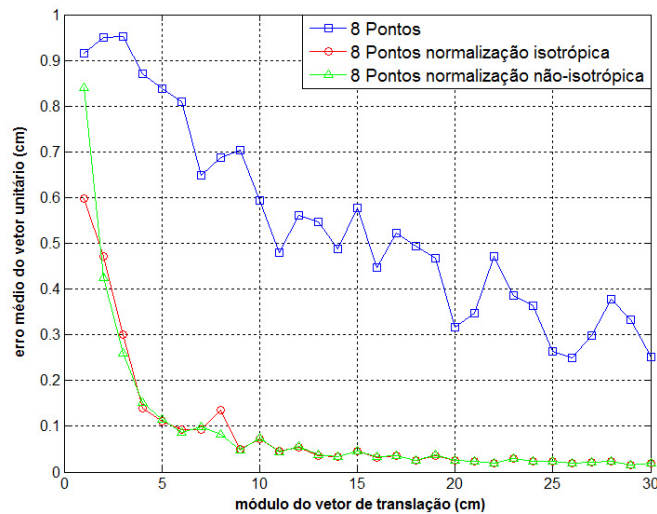


Figura 5.10: Análise da estimação do vetor unitário de translação da câmera em função dos módulos do vetor t

Na Figura 5.11, é feita a análise da estimação do fator de escala do vetor de translação do movimento de translação pura. Nos gráficos da figura são feitos um comparativo entre os erros médios dos fatores de escala estimados (definido como sendo o valor médio dos valores absolutos da subtração entre os módulos dos vetores de translação com fator de escala estimado e do movimento) e o módulo do vetor de translação do movimento da câmera para cada uma das três técnicas.

Podemos verificar que o aumento do módulo do vetor de translação do movimento implica no aumento do erro do fator de escala para o Algoritmo de 8 Pontos. Contudo, nas técnicas com uso da normalização, os erros são praticamente constantes e abaixo de 1 cm.

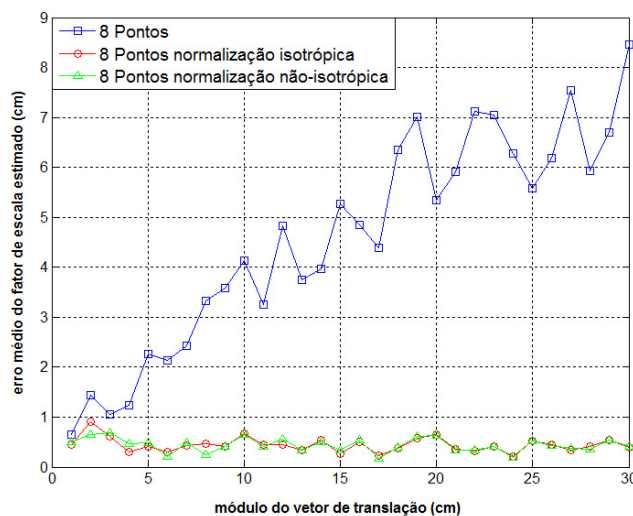


Figura 5.11: Análise da estimação do fator de escala da translação da câmera em função dos módulos do vetor \mathbf{t}

Utilizando a câmera na mesma condição inicial do teste descrito anteriormente, iremos analisar os movimentos de translação com rotação da câmera.

Nesse teste, fixamos o módulo do vetor de translação em 15 cm. Rotacionamos a câmera segundo os seguintes ângulos $[0^\circ \pm 1^\circ \pm 2,5^\circ \pm 5^\circ \pm 10^\circ \pm 15^\circ \pm 20^\circ]$ no seu eixo y e transladamos a mesma em 18 direções diferentes, sendo 9 direções para os ângulos não-positivos (Tabela 5.2) e as outras para os não-negativos (Tabela 5.3), de modo que todos os 50 pontos foram projetados na imagem após o movimento.

Tabela 5.2: Vetores unitários de translação para as rotações não-positivas (em cm)

Translação	eixo x	eixo y	eixo z	Translação	eixo x	eixo y	eixo z
\mathbf{t}_1^-	1	0	0	\mathbf{t}_6^-	$\frac{\sqrt{3}}{3}$	$\frac{\sqrt{3}}{3}$	$\frac{\sqrt{3}}{3}$
\mathbf{t}_2^-	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{2}}{2}$	0	\mathbf{t}_7^-	$\frac{\sqrt{3}}{3}$	$-\frac{\sqrt{3}}{3}$	$\frac{\sqrt{3}}{3}$
\mathbf{t}_3^-	$\frac{\sqrt{2}}{2}$	0	$\frac{\sqrt{2}}{2}$	\mathbf{t}_8^-	$\frac{\sqrt{3}}{3}$	$\frac{\sqrt{3}}{3}$	$-\frac{\sqrt{3}}{3}$
\mathbf{t}_4^-	$\frac{\sqrt{2}}{2}$	$-\frac{\sqrt{2}}{2}$	0	\mathbf{t}_9^-	$\frac{\sqrt{3}}{3}$	$-\frac{\sqrt{3}}{3}$	$-\frac{\sqrt{3}}{3}$
\mathbf{t}_5^-	$\frac{\sqrt{2}}{2}$	0	$-\frac{\sqrt{2}}{2}$				

Tabela 5.3: Vetores unitários de translação para as rotações não-negativas (em cm)

Translação	eixo x	eixo y	eixo z	Translação	eixo x	eixo y	eixo z
\mathbf{t}_1^+	-1	0	0	\mathbf{t}_6^+	$-\frac{\sqrt{3}}{3}$	$\frac{\sqrt{3}}{3}$	$\frac{\sqrt{3}}{3}$
\mathbf{t}_2^+	$-\frac{\sqrt{2}}{2}$	$\frac{\sqrt{2}}{2}$	0	\mathbf{t}_7^+	$-\frac{\sqrt{3}}{3}$	$-\frac{\sqrt{3}}{3}$	$\frac{\sqrt{3}}{3}$
\mathbf{t}_3^+	$-\frac{\sqrt{2}}{2}$	0	$\frac{\sqrt{2}}{2}$	\mathbf{t}_8^+	$-\frac{\sqrt{3}}{3}$	$\frac{\sqrt{3}}{3}$	$-\frac{\sqrt{3}}{3}$
\mathbf{t}_4^+	$-\frac{\sqrt{2}}{2}$	$-\frac{\sqrt{2}}{2}$	0	\mathbf{t}_9^+	$-\frac{\sqrt{3}}{3}$	$-\frac{\sqrt{3}}{3}$	$-\frac{\sqrt{3}}{3}$
\mathbf{t}_5^+	$-\frac{\sqrt{2}}{2}$	0	$-\frac{\sqrt{2}}{2}$				

Na Figura 5.12, analisamos a estimação do vetor unitário de translação do movimento de translação com rotação. Fazemos um comparativo entre os erros médios dos vetores unitários estimados (correspondendo ao valor médio dos módulos dos vetores erros obtidos através da subtração vetorial entre o vetor unitário estimado da translação e o vetor unitário da translação do movimento) e o módulo dos ângulos de rotação do eixo y da câmera para cada uma das três técnicas.

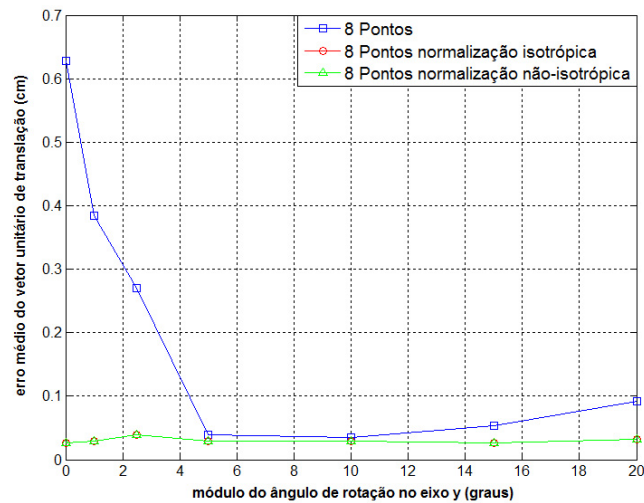


Figura 5.12: Análise da estimação do vetor unitário de translação da câmera em função das rotações no eixo y

Na Figura 5.13, analisamos a estimação do fator de escala do vetor de translação do movimento de translação com rotação. Nos gráficos da figura, comparamos os erros médios dos fatores de escala estimados (correspondendo ao valor médio dos valores absolutos da subtração entre os módulos dos vetores de translação com fator de escala estimado e do movimento) e o módulo dos ângulos de rotação do eixo y da câmera para cada uma das três técnicas.

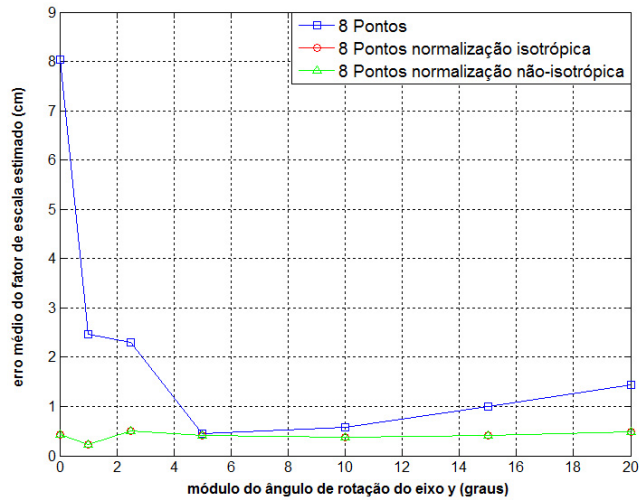


Figura 5.13: Análise da estimação do fator de escala da translação da câmera em função das rotações no eixo y

Nas Figuras 5.14, 5.15 e 5.16, fazemos a análise da estimação dos ângulos de rotação nos eixos x , y e z da câmera após o movimento de translação com rotação. É feita, para cada figura, uma comparação dos erros médios da estimativa dos ângulos de rotação referentes a cada um dos três eixos da câmera (correspondendo ao valor médio dos valores absolutos da subtração entre o ângulo estimado e o do movimento referentes a cada eixo em análise) em função do módulo dos ângulos de rotação do eixo y da câmera para cada uma das três técnicas.

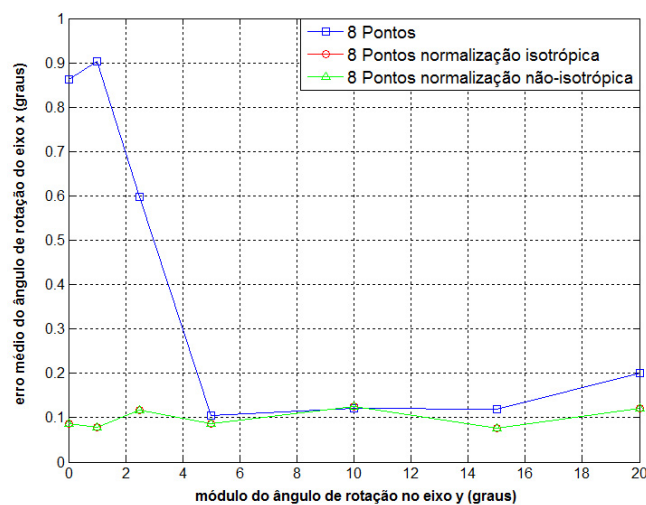


Figura 5.14: Análise do erro médio do ângulo de rotação no eixo x em função das rotações no eixo y

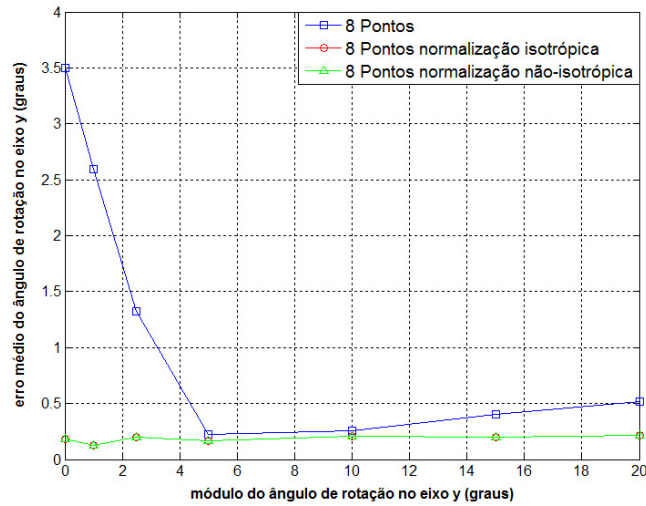


Figura 5.15: Análise do erro médio do ângulo de rotação no eixo y em função das rotações no eixo y

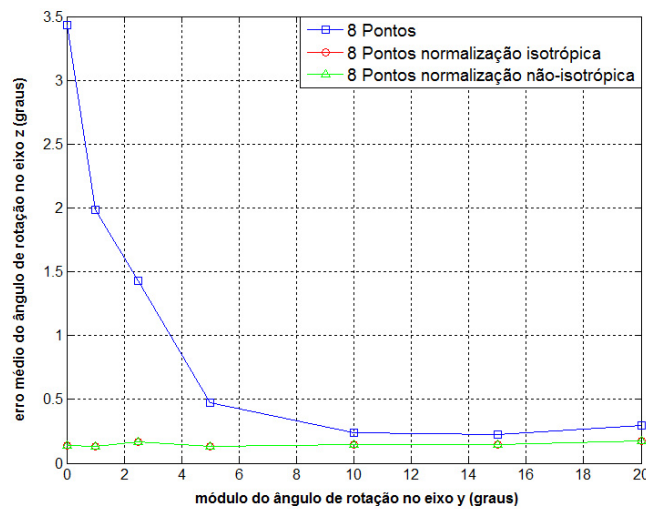


Figura 5.16: Análise do erro médio do ângulo de rotação no eixo z em função das rotações no eixo y

Podemos constatar, a partir dos gráficos das Figuras 5.12 e 5.13 (referentes a estimação da translação do movimento) e das Figuras 5.14, 5.15 e 5.16 (referentes a estimação da rotação do movimento), que as estimações utilizando as técnicas com normalização possuem erros pequenos e constantes para quaisquer rotações. Contudo, a técnica utilizando o Algoritmo de 8 Pontos possui erros mais elevados para movimentos com menores rotações de 0 a 5 graus.

Avaliação da Estimação de Trajetória

Neste teste, realizamos a estimação da posição de uma câmera sobre uma trajetória pré-definida.

No caso, essa trajetória corresponde a um movimento senoidal no plano \widehat{xz} do sistema referencial da câmera na posição inicial. Essa corresponde a mesma utilizada na condição inicial dos testes da estimação do movimento, sendo utilizados os mesmos 50 pontos de referência do cenário 3D para nossas estimações.

A fórmula da trajetória é dada por:

$$x = 50 \sin\left(\pi \frac{z}{60}\right) (\text{cm}), \text{ com } z \text{ variando de } 0 \text{ a } -120 \text{ cm e } y = 0 \text{ cm.}$$

Capturamos as imagens da câmera nas posições marcadas na Figura 5.17.

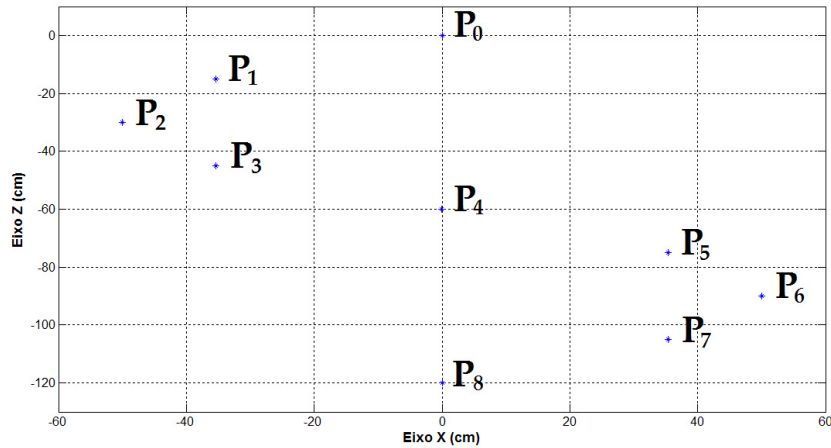


Figura 5.17: Posições da trajetória de referência

Nessas posições, aplicamos rotações no eixo y da câmera em relação a sua orientação na posição inicial (P_0), em função de uma constante angular ΔR . As rotações das posições obedecem à Tabela 5.4.

Tabela 5.4: Rotações das posições

Posições	Rotações no eixo y
P_0	0
P_1	ΔR
P_2	$2\Delta R$
P_3	ΔR
P_4	0
P_5	$-\Delta R$
P_6	$-2\Delta R$
P_7	$-\Delta R$
P_8	0

Nos testes foram utilizados os seguintes valores de $\Delta R = [0^\circ \ 1^\circ \ 2,5^\circ \ 5^\circ \ 10^\circ \ 15^\circ]$, e cada valor correspondendo a uma trajetória estimada.

Nos gráficos das Figuras 5.18, 5.19 e 5.20, vemos as estimativas das trajetórias utilizando o Algoritmo de 8 Pontos nas vistas XZ, XY e YZ, respectivamente.

Nos gráficos das Figuras 5.21, 5.22 e 5.23, vemos as estimativas das trajetórias utilizando o Algoritmo de 8 Pontos Normalizado com normalização isotrópica nas vistas XZ, XY e YZ, respectivamente.

Nos gráficos das Figuras 5.24, 5.25 e 5.26, vemos as estimativas das trajetórias utilizando o Algoritmo de 8 Pontos Normalizado com normalização não-isotrópica nas vistas XZ, XY e YZ, respectivamente.

Podemos observar que para as trajetórias estimadas utilizando o Algoritmo de 8 Pontos (Figuras 5.18, 5.19 e 5.20), os erros das estimações das posições tendem a aumentar, sendo esse crescimento mais intensificado para as trajetórias com menores valores de ΔR (ΔR menores que 5°).

As trajetórias estimadas utilizando os algoritmos normalizados (isotrópica: Figuras 5.21, 5.22 e 5.23; não-isotrópica: Figuras 5.24, 5.25 e 5.26) possuem as estimações praticamente iguais e essas sempre muito próximas da posição real da câmera na trajetória senoidal de referência.

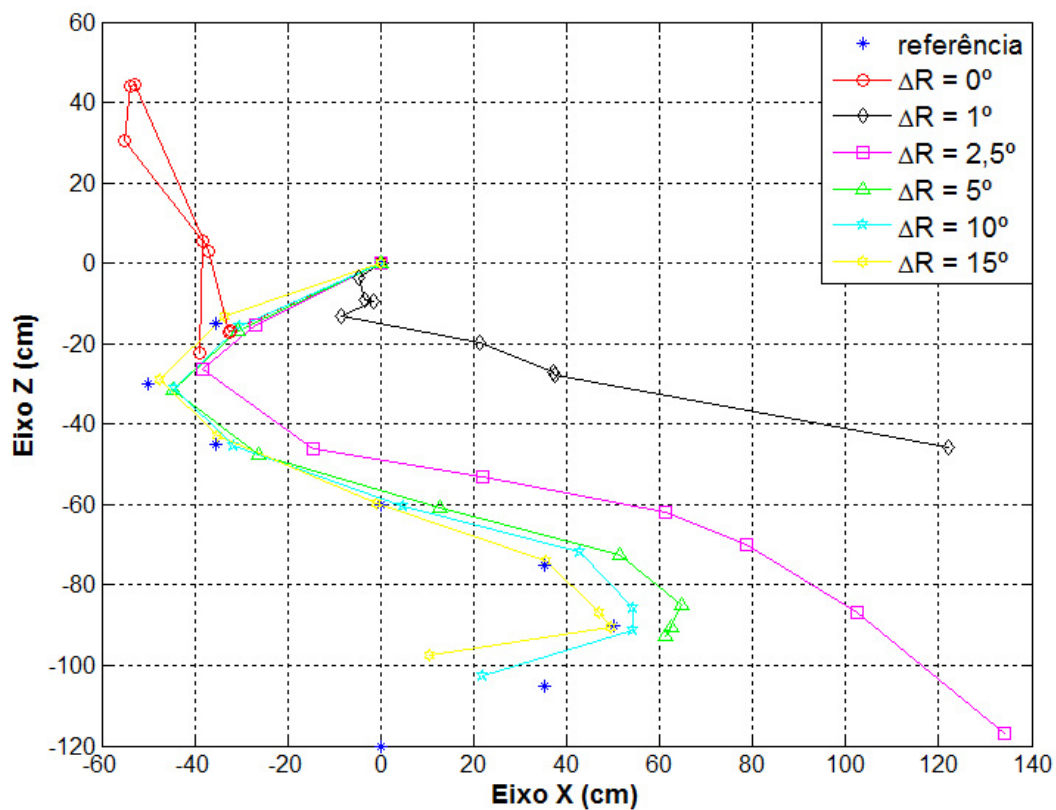


Figura 5.18: Estimativas das trajetórias com Algoritmo de 8 Pontos (vista XZ)

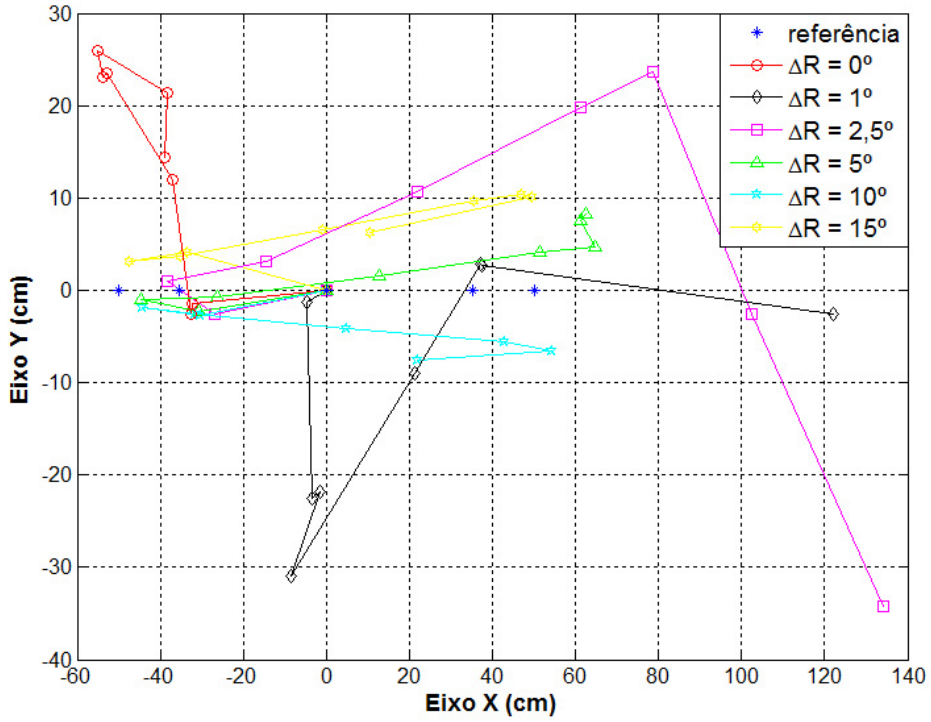


Figura 5.19: Estimativas das trajetórias com Algoritmo de 8 Pontos (vista XY)

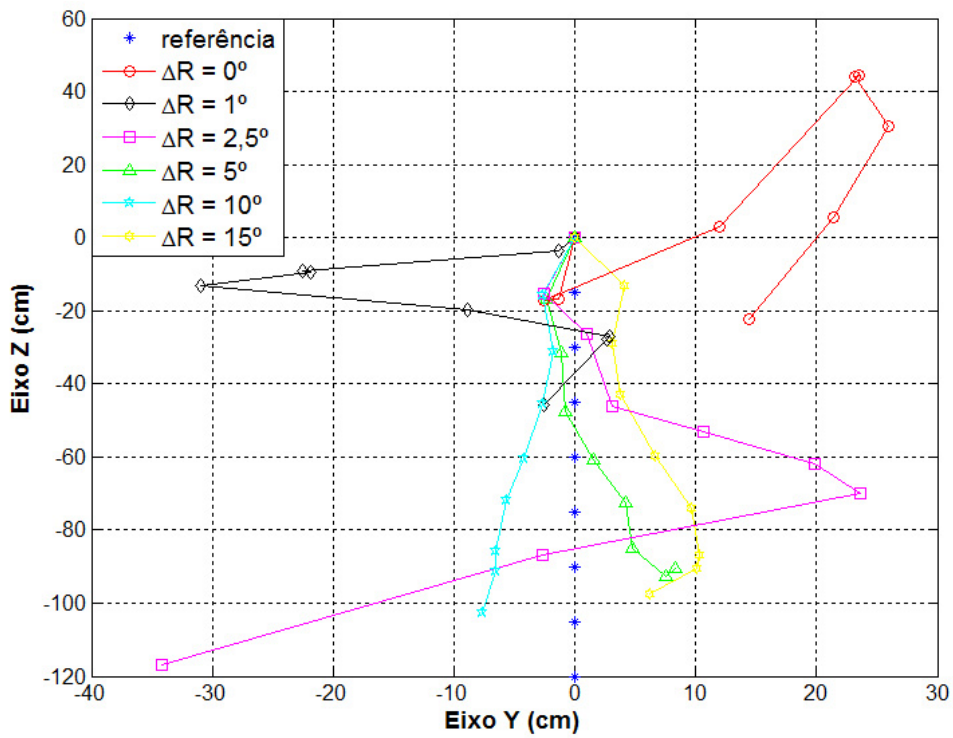


Figura 5.20: Estimativas das trajetórias com Algoritmo de 8 Pontos (vista YZ)

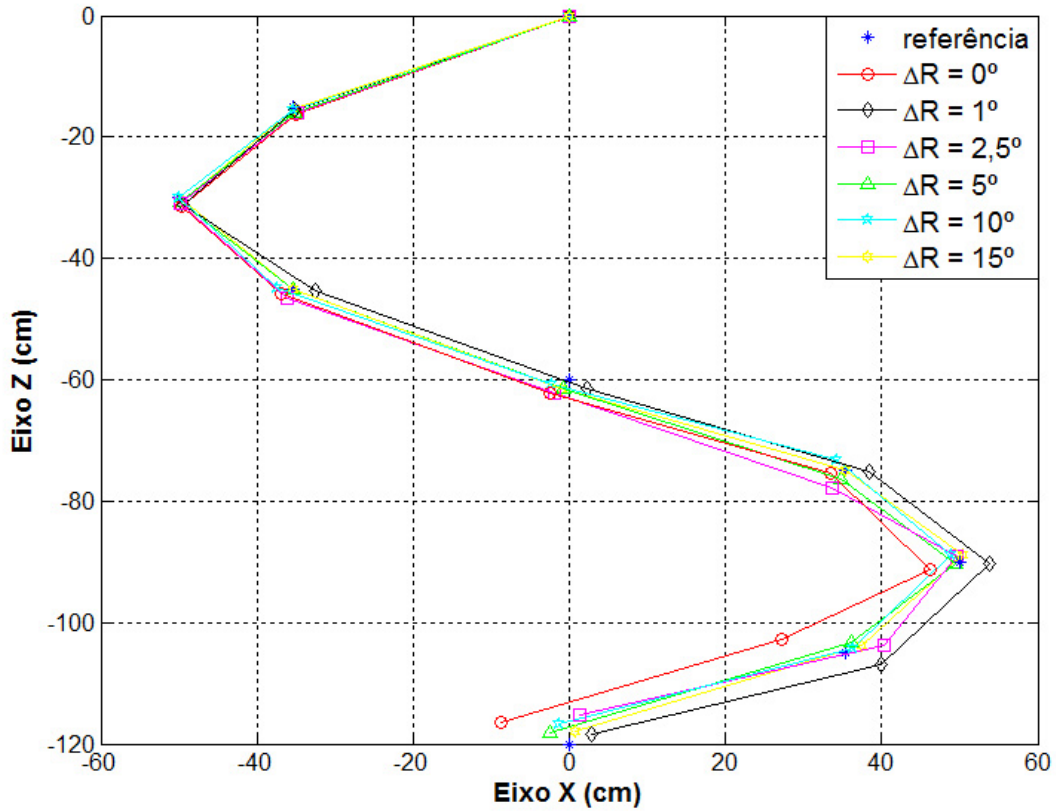


Figura 5.21: Estimativas das trajetórias com Algoritmo de 8 Pontos Normalizado - isotrópica (vista XZ)

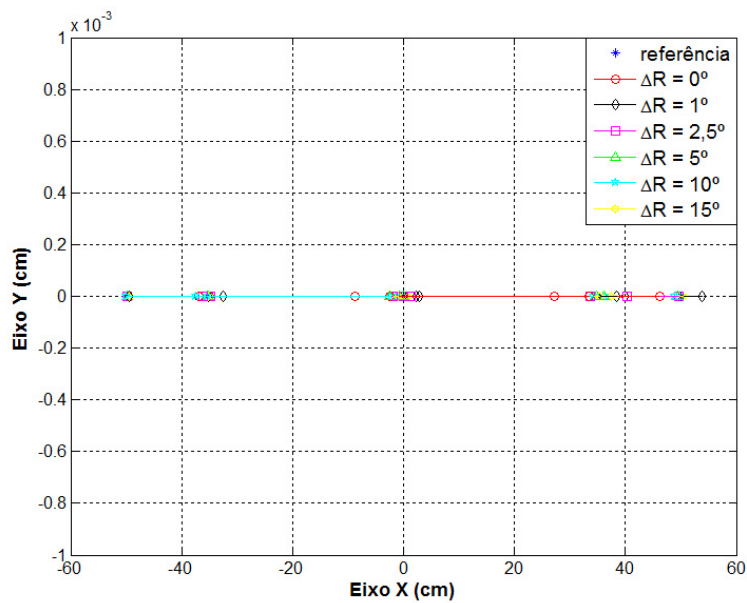


Figura 5.22: Estimativas das trajetórias com Algoritmo de 8 Pontos Normalizado - isotrópica (vista XY)

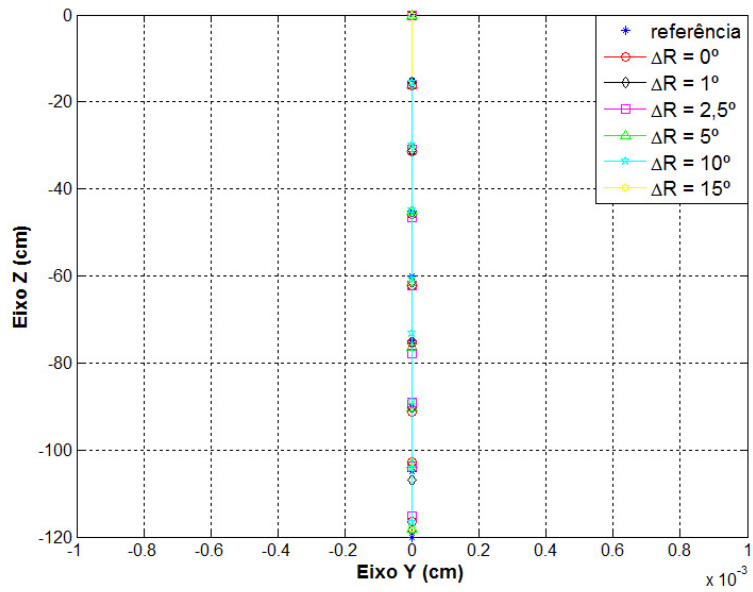


Figura 5.23: Estimativas das trajetórias com Algoritmo de 8 Pontos Normalizado - isotrópica (vista YZ)

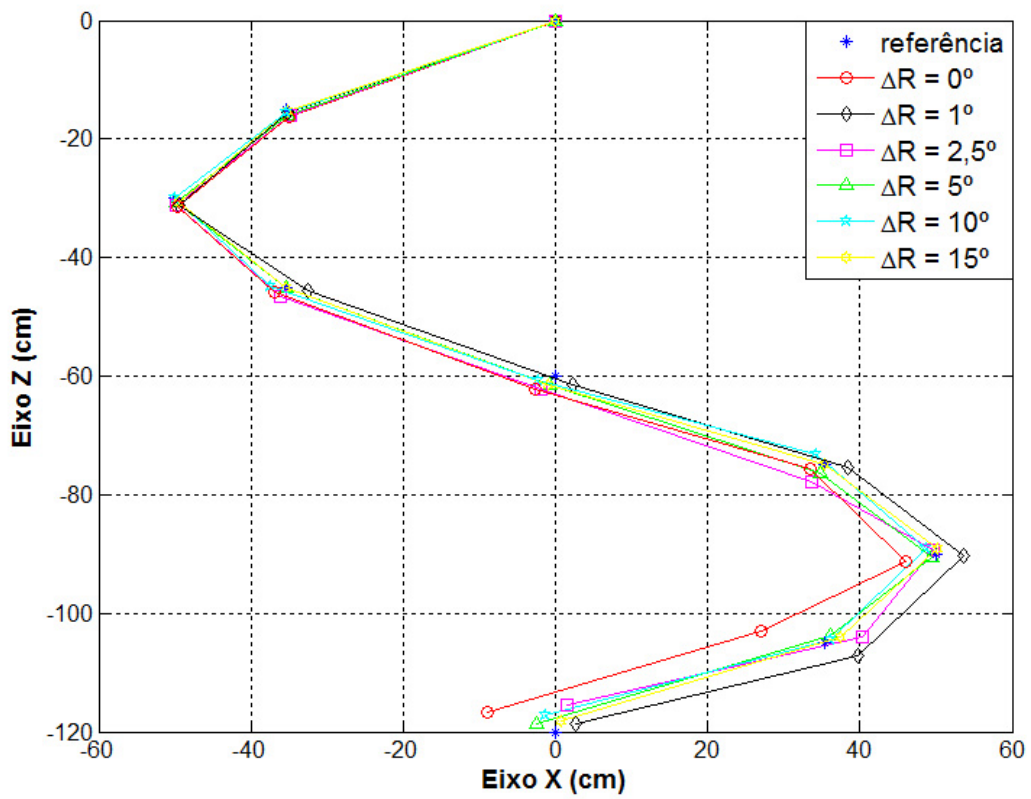


Figura 5.24: Estimativas das trajetórias com Algoritmo de 8 Pontos Normalizado - não-isotrópica (vista XZ)

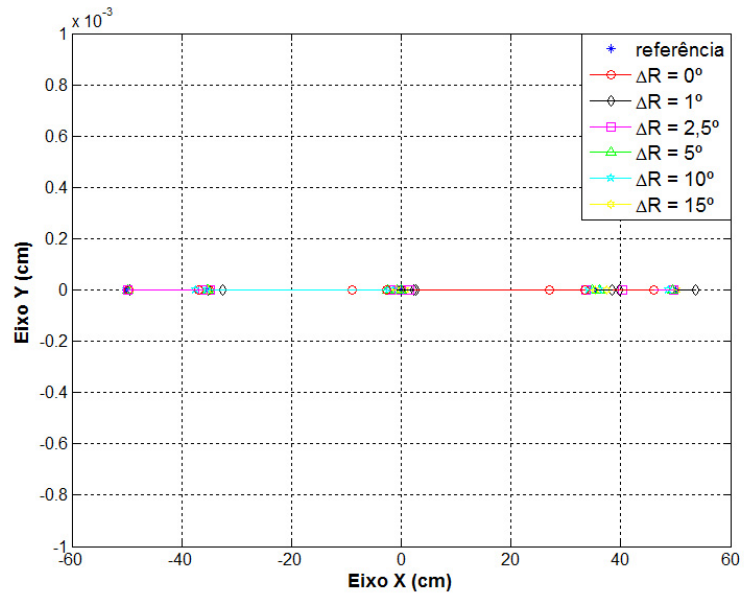


Figura 5.25: Estimativas das trajetórias com Algoritmo de 8 Pontos Normalizado - não-isotrópica (vista XY)

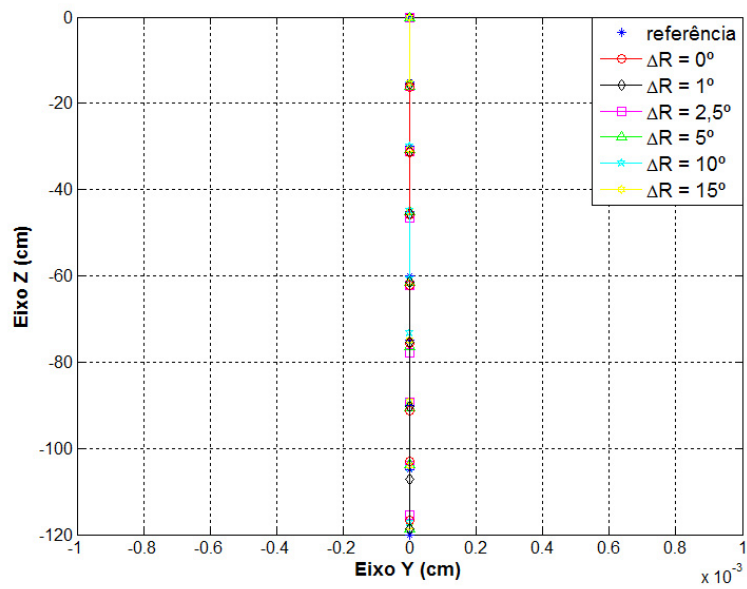


Figura 5.26: Estimativas das trajetórias com Algoritmo de 8 Pontos Normalizado - não-isotrópica (vista YZ)

5.2.2 Dados Reais

Finalmente, depois de avaliados os algoritmos com dados sintéticos, vamos testá-los com imagens reais. Utilizamos uma câmera digital da Logitech modelo Quickcam Orbit AF. Ela é composta por um sensor óptico de 2 megapixels (resolução 1600×1200) e um conjunto de lentes Carl Zeiss ajustável (distância focal variável de 2mm a 3,2 mm).

Inicialmente, precisamos obter a matriz de calibração \mathbf{K} das imagens. Calibramos a câmera antes de começar os experimentos usando como referência 3D as bordas dos azulejos de um canto entre duas paredes, conforme podemos ver na Figura 5.27.

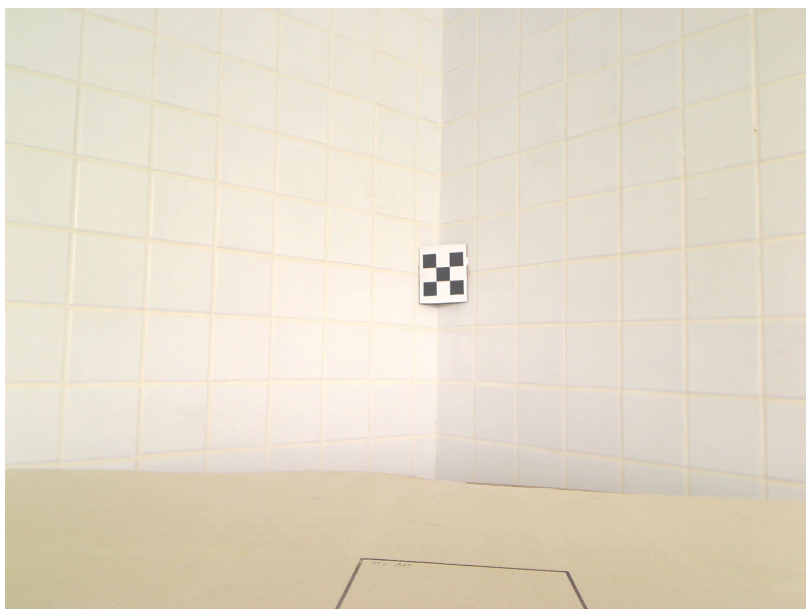


Figura 5.27: Exemplo de imagem usada para calibração da câmera

Nesses testes, utilizamos um mapa de posições conforme vemos na Figura 5.28. Calibramos a câmera usando a imagem capturada na posição de referência \mathbf{O} do mapa da Figura 5.28. Utilizamos 70 pontos correspondentes aos azulejos que medimos previamente com uma régua e obtivemos um erro médio absoluto de 1,17 pixels na reprojeção dos mesmos com a matriz estimada. Na Tabela 5.5 seguem os parâmetros de calibração estimados.

Tabela 5.5: Parâmetros da câmera estimados na calibração

Parâmetros	Valor	Unidade
u_0	797,69	pixels
v_0	598,25	pixels
α_u	1427,5	(pixels)cm ⁻²
α_v	1410,1	(pixels)cm ⁻²
θ	90,23	graus
\mathbf{t}	$\begin{bmatrix} 5,2503 \\ 24,3141 \\ 138,5536 \end{bmatrix}$	cm
\mathbf{R}	$\begin{bmatrix} -0,7582 & 0,6517 & 0,0195 \\ -0,0629 & -0,0433 & -0,9971 \\ -0,6490 & -0,7572 & 0,0738 \end{bmatrix}$	

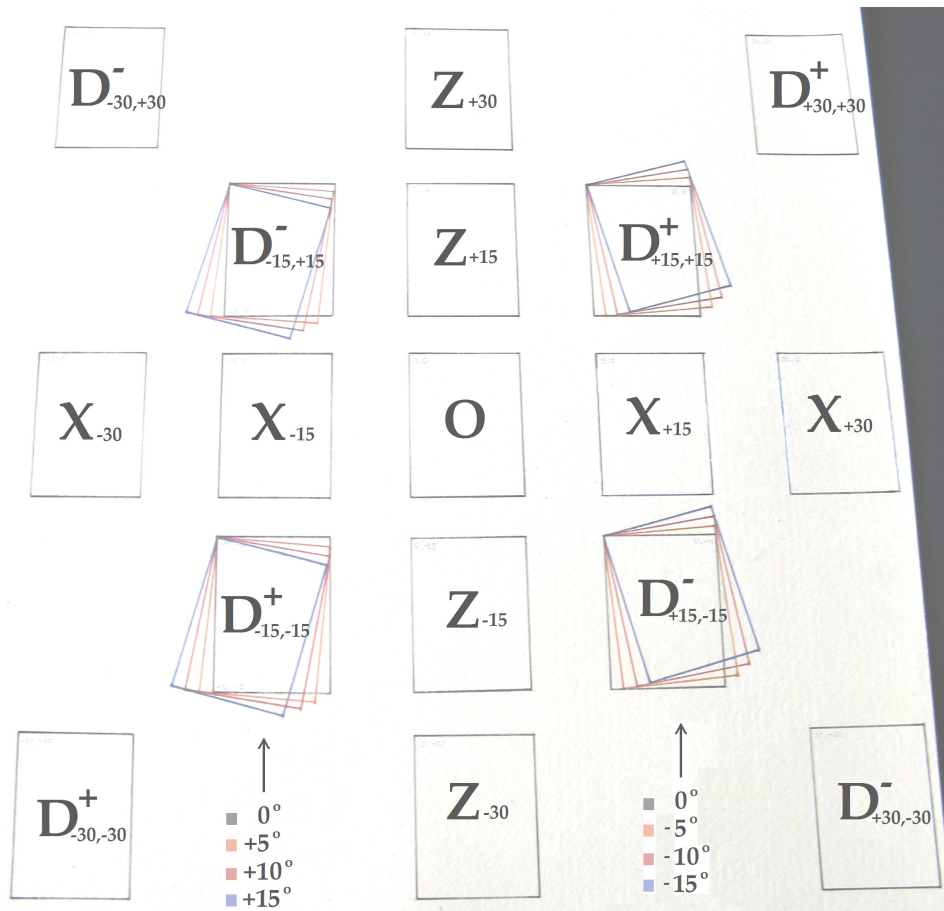


Figura 5.28: Mapa dos testes com imagens reais

Depois de calibrada a câmera, realizamos os testes segundo o mapa de posições exposto na Figura 5.28. Na imagem capturada em cada posição, medimos manualmente as coordenadas associadas aos cantos dos alvos em forma de tabuleiros de xadrez, conforme vemos na Figura 5.30. Totalizam-se 80 pontos de referência por imagem.

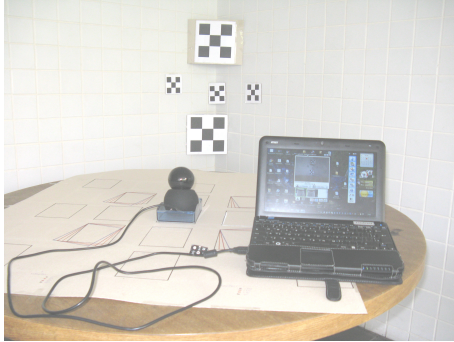


Figura 5.29: Câmera na posição de referência **O**

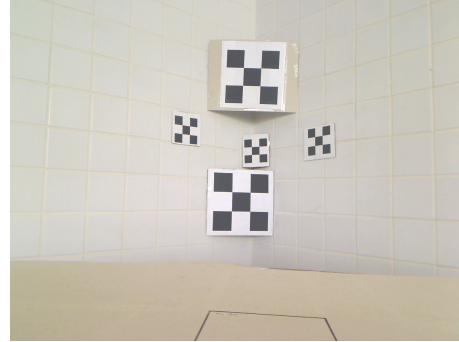


Figura 5.30: Imagem capturada na posição de referência **O**

A fotografia da câmera na posição de referência **O** é mostrada na Figura 5.29, e a imagem capturada pela câmera nessa posição aparece na Figura 5.30. Utilizando o Algoritmo de 8 Pontos com normalização não-isotrópica e a câmera na posição de referência **O** como posição inicial do movimento, estimamos os seguintes movimentos:

Movimentos de Translação Pura

Nesse teste, avaliamos, inicialmente, o movimento de translação pura sobre os eixos x e z da câmera. As translações no mapa (Figura 5.28), correspondem, para a translação sobre o eixo x , às posições **X**, e sobre o eixo z , às posições **Z**.

Tabela 5.6: Deslocamento sobre o eixo x

Mapa	Translação	Trans. Estimada	Erro da translação	Unidade
X ₊₃₀	$\begin{bmatrix} 30,0 \\ 0,0 \\ 0,0 \end{bmatrix}$	$\begin{bmatrix} 29,54 \\ 1,14 \\ -2,21 \end{bmatrix}$	2,52	cm
X ₊₁₅	$\begin{bmatrix} 15,0 \\ 0,0 \\ 0,0 \end{bmatrix}$	$\begin{bmatrix} 14,85 \\ 0,63 \\ -0,49 \end{bmatrix}$	0,81	cm
X ₋₁₅	$\begin{bmatrix} -15,0 \\ 0,0 \\ 0,0 \end{bmatrix}$	$\begin{bmatrix} -14,69 \\ -0,57 \\ 1,42 \end{bmatrix}$	1,56	cm
X ₋₃₀	$\begin{bmatrix} -30,0 \\ 0,0 \\ 0,0 \end{bmatrix}$	$\begin{bmatrix} -29,51 \\ -1,07 \\ 3,47 \end{bmatrix}$	3,67	cm

A partir da Tabela 5.6, que analisa o deslocamento da câmera no eixo x , podemos observar que quanto mais a câmera se afasta da posição de referência, maior é o erro na estimativa do vetor de translação, pois o uso da matriz de calibração obtida na posição de referência começa a não ser mais válido sendo mais uma fonte de erro na estimativa. Em média, esse erro corresponde a 10% do valor do módulo do vetor de translação do movimento.

Tabela 5.7: Deslocamento sobre o eixo z

Mapa	Translação	Trans. Estimada	Erro da translação	Unidade
\mathbf{Z}_{+30}	$\begin{bmatrix} 0,0 \\ 0,0 \\ 30,0 \end{bmatrix}$	$\begin{bmatrix} 0,96 \\ 2,19 \\ 29,55 \end{bmatrix}$	2,44	cm
\mathbf{Z}_{+15}	$\begin{bmatrix} 0,0 \\ 0,0 \\ 15,0 \end{bmatrix}$	$\begin{bmatrix} 1,19 \\ 1,77 \\ 14,72 \end{bmatrix}$	2,16	cm
\mathbf{Z}_{-15}	$\begin{bmatrix} 0,0 \\ 0,0 \\ -15,0 \end{bmatrix}$	$\begin{bmatrix} -1,12 \\ -1,62 \\ -14,65 \end{bmatrix}$	2,01	cm
\mathbf{Z}_{-30}	$\begin{bmatrix} 0,0 \\ 0,0 \\ -30,0 \end{bmatrix}$	$\begin{bmatrix} -2,96 \\ -4,43 \\ -29,26 \end{bmatrix}$	5,38	cm

Na Tabela 5.7 (análise do movimento no eixo z), podemos ver que novamente o fenômeno do afastamento da posição de referência da câmera, influencia no aumento do erro da estimativa do vetor de translação, assim como vimos na análise do deslocamento do eixo x .

No segundo teste, avaliamos o movimento de translação pura sobre os dois eixos xz da câmera. No mapa (Figura 5.28), estas translações correspondem às posições \mathbf{D}^- (diagonal com ângulo de -45°) e \mathbf{D}^+ (diagonal com ângulo de $+45^\circ$).

Tabela 5.8: Deslocamento na diagonal em -45°

Mapa	Translação	Trans. Estimada	Erro da translação	Unidade
$\mathbf{D}_{-30,+30}^-$	$\begin{bmatrix} -30,0 \\ 0,0 \\ 30,0 \end{bmatrix}$	$\begin{bmatrix} -22,87 \\ -0,16 \\ 27,34 \end{bmatrix}$	7,60	cm
$\mathbf{D}_{-15,+15}^-$	$\begin{bmatrix} -15,0 \\ 0,0 \\ 15,0 \end{bmatrix}$	$\begin{bmatrix} -11,90 \\ 0,77 \\ 14,62 \end{bmatrix}$	3,20	cm
$\mathbf{D}_{+15,-15}^-$	$\begin{bmatrix} 15,0 \\ 0,0 \\ -15,0 \end{bmatrix}$	$\begin{bmatrix} 11,93 \\ -0,85 \\ -14,47 \end{bmatrix}$	3,22	cm
$\mathbf{D}_{+30,-30}^-$	$\begin{bmatrix} 30,0 \\ 0,0 \\ -30,0 \end{bmatrix}$	$\begin{bmatrix} 22,32 \\ -2,86 \\ -27,76 \end{bmatrix}$	8,49	cm

Tabela 5.9: Deslocamento na diagonal em $+45^\circ$

Mapa	Translação	Trans. estimada	Erro da translação	Unidade
$D_{+30,+30}^+$	$\begin{bmatrix} 30,0 \\ 0,0 \\ 30,0 \end{bmatrix}$	$\begin{bmatrix} 24,41 \\ 1,13 \\ 25,95 \end{bmatrix}$	6,98	cm
$D_{+15,+15}^+$	$\begin{bmatrix} 15,0 \\ 0,0 \\ 15,0 \end{bmatrix}$	$\begin{bmatrix} 11,19 \\ 1,77 \\ 14,72 \end{bmatrix}$	3,40	cm
$D_{-15,-15}^+$	$\begin{bmatrix} -15,0 \\ 0,0 \\ -15,0 \end{bmatrix}$	$\begin{bmatrix} -14,59 \\ -1,91 \\ -11,65 \end{bmatrix}$	3,87	cm
$D_{-30,-30}^+$	$\begin{bmatrix} -30,0 \\ 0,0 \\ -30,0 \end{bmatrix}$	$\begin{bmatrix} -29,13 \\ -5,36 \\ -19,99 \end{bmatrix}$	11,39	cm

A partir das Tabelas 5.8 e 5.9, podemos constatar que os erros das estimações dos vetores de translação são um pouco maiores os das Tabelas 5.6 e 5.7, haja vista que o deslocamento em um único eixo faz com que o erro das estimativas seja menor em relação ao deslocamento em dois eixos (x e z).

Movimentos de Translação e Rotação

Nesse teste avaliamos o movimento de translação com rotação sobre o eixo y da câmera.

Tabela 5.10: Rotações na posição: $D_{-15,-15}^+$

Translação	Trans. Estimada	Erro da translação	Unidade	Rotação	Rot. estimada	Erro da rotação	Unidade
$\begin{bmatrix} -15,0 \\ 0,0 \\ -15,0 \end{bmatrix}$	$\begin{bmatrix} -14,59 \\ -1,91 \\ -11,65 \end{bmatrix}$	3,87	cm	$\begin{bmatrix} 0,0 \\ 0,0 \\ 0,0 \end{bmatrix}$	$\begin{bmatrix} -0,50 \\ 0,19 \\ -0,88 \end{bmatrix}$	$\begin{bmatrix} 0,50 \\ 0,19 \\ 0,88 \end{bmatrix}$	graus
$\begin{bmatrix} -15,0 \\ 0,0 \\ -15,0 \end{bmatrix}$	$\begin{bmatrix} -11,02 \\ -0,65 \\ -15,19 \end{bmatrix}$	4,04	cm	$\begin{bmatrix} 0,0 \\ 5,0 \\ 0,0 \end{bmatrix}$	$\begin{bmatrix} -0,79 \\ 5,18 \\ 0,47 \end{bmatrix}$	$\begin{bmatrix} 0,79 \\ 0,18 \\ 0,47 \end{bmatrix}$	graus
$\begin{bmatrix} -15,0 \\ 0,0 \\ -15,0 \end{bmatrix}$	$\begin{bmatrix} -10,34 \\ -0,84 \\ -15,64 \end{bmatrix}$	4,77	cm	$\begin{bmatrix} 0,0 \\ 10,0 \\ 0,0 \end{bmatrix}$	$\begin{bmatrix} -0,46 \\ 10,35 \\ 1,28 \end{bmatrix}$	$\begin{bmatrix} 0,46 \\ 0,35 \\ 1,28 \end{bmatrix}$	graus
$\begin{bmatrix} -15,0 \\ 0,0 \\ -15,0 \end{bmatrix}$	$\begin{bmatrix} -10,58 \\ -0,93 \\ -15,48 \end{bmatrix}$	4,54	cm	$\begin{bmatrix} 0,0 \\ 15,0 \\ 0,0 \end{bmatrix}$	$\begin{bmatrix} 0,55 \\ 15,71 \\ -2,15 \end{bmatrix}$	$\begin{bmatrix} 0,55 \\ 0,71 \\ 2,15 \end{bmatrix}$	graus

Na Tabela 5.10, podemos observar que os erros das estimações dos vetores de translação são praticamente iguais, e as estimativas dos ângulo de rotação de 0 a 15 graus, possuem erros de até no máximo 2,5 graus. Podemos ver na análise em deslocamentos equidistantes, nas Tabelas 5.11, 5.12 e 5.13, que obtemos as mesmas respostas para as estimações de translação e rotação.

Tabela 5.11: Rotações na posição: $D_{-15,15}^-$

Translação	Trans. Estimada	Erro da translação	Unidade	Rotação	Rot. estimada	Erro da rotação	Unidade
$\begin{bmatrix} -15,0 \\ 0,0 \\ 15,0 \end{bmatrix}$	$\begin{bmatrix} -11,90 \\ 0,77 \\ 14,62 \end{bmatrix}$	3,20	cm	$\begin{bmatrix} 0,0 \\ 0,0 \\ 0,0 \end{bmatrix}$	$\begin{bmatrix} 0,86 \\ -0,76 \\ 0,85 \end{bmatrix}$	$\begin{bmatrix} 0,86 \\ 0,76 \\ 0,85 \end{bmatrix}$	graus
$\begin{bmatrix} -15,0 \\ 0,0 \\ 15,0 \end{bmatrix}$	$\begin{bmatrix} -13,44 \\ 2,50 \\ 13,01 \end{bmatrix}$	3,55	cm	$\begin{bmatrix} 0,0 \\ 5,0 \\ 0,0 \end{bmatrix}$	$\begin{bmatrix} -0,41 \\ 5,09 \\ 0,21 \end{bmatrix}$	$\begin{bmatrix} 0,41 \\ 0,09 \\ 0,21 \end{bmatrix}$	graus
$\begin{bmatrix} -15,0 \\ 0,0 \\ 15,0 \end{bmatrix}$	$\begin{bmatrix} -12,66 \\ 2,53 \\ 13,76 \end{bmatrix}$	3,65	cm	$\begin{bmatrix} 0,0 \\ 10,0 \\ 0,0 \end{bmatrix}$	$\begin{bmatrix} 0,48 \\ 10,36 \\ -1,09 \end{bmatrix}$	$\begin{bmatrix} 0,48 \\ 0,36 \\ 1,09 \end{bmatrix}$	graus
$\begin{bmatrix} -15,0 \\ 0,0 \\ 15,0 \end{bmatrix}$	$\begin{bmatrix} -12,49 \\ 2,61 \\ 13,90 \end{bmatrix}$	3,77	cm	$\begin{bmatrix} 0,0 \\ 15,0 \\ 0,0 \end{bmatrix}$	$\begin{bmatrix} 1,77 \\ 16,64 \\ -2,49 \end{bmatrix}$	$\begin{bmatrix} 1,77 \\ 1,64 \\ 2,49 \end{bmatrix}$	graus

Tabela 5.12: Rotações na posição: $D_{15,15}^+$

Translação	Trans. Estimada	Erro da translação	Unidade	Rotação	Rot. estimada	Erro da rotação	Unidade
$\begin{bmatrix} 15,0 \\ 0,0 \\ 15,0 \end{bmatrix}$	$\begin{bmatrix} 11,90 \\ 0,77 \\ 14,62 \end{bmatrix}$	3,20	cm	$\begin{bmatrix} 0,0 \\ 0,0 \\ 0,0 \end{bmatrix}$	$\begin{bmatrix} 0,89 \\ 0,55 \\ -0,31 \end{bmatrix}$	$\begin{bmatrix} 0,89 \\ 0,55 \\ 0,31 \end{bmatrix}$	graus
$\begin{bmatrix} 15,0 \\ 0,0 \\ 15,0 \end{bmatrix}$	$\begin{bmatrix} 15,93 \\ 2,49 \\ 9,80 \end{bmatrix}$	5,84	cm	$\begin{bmatrix} 0,0 \\ -5,0 \\ 0,0 \end{bmatrix}$	$\begin{bmatrix} 0,65 \\ -6,41 \\ -0,40 \end{bmatrix}$	$\begin{bmatrix} 0,65 \\ 1,41 \\ 0,40 \end{bmatrix}$	graus
$\begin{bmatrix} 15,0 \\ 0,0 \\ 15,0 \end{bmatrix}$	$\begin{bmatrix} 15,93 \\ 2,70 \\ 9,74 \end{bmatrix}$	5,98	cm	$\begin{bmatrix} 0,0 \\ -10,0 \\ 0,0 \end{bmatrix}$	$\begin{bmatrix} 0,19 \\ -9,87 \\ -1,23 \end{bmatrix}$	$\begin{bmatrix} 0,19 \\ 0,13 \\ 1,23 \end{bmatrix}$	graus
$\begin{bmatrix} 15,0 \\ 0,0 \\ 15,0 \end{bmatrix}$	$\begin{bmatrix} 15,92 \\ 2,86 \\ 9,73 \end{bmatrix}$	6,06	cm	$\begin{bmatrix} 0,0 \\ -15,0 \\ 0,0 \end{bmatrix}$	$\begin{bmatrix} -0,51 \\ -13,25 \\ 2,14 \end{bmatrix}$	$\begin{bmatrix} 0,51 \\ 1,75 \\ 2,14 \end{bmatrix}$	graus

Tabela 5.13: Rotações na posição: $D_{15,-15}^-$

Translação	Trans. Estimada	Erro da translação	Unidade	Rotação	Rot. estimada	Erro da rotação	Unidade
$\begin{bmatrix} 15,0 \\ 0,0 \\ -15,0 \end{bmatrix}$	$\begin{bmatrix} -11,90 \\ 0,77 \\ 14,62 \end{bmatrix}$	3,20	cm	$\begin{bmatrix} 0,0 \\ 0,0 \\ 0,0 \end{bmatrix}$	$\begin{bmatrix} 0,27 \\ -0,96 \\ 0,42 \end{bmatrix}$	$\begin{bmatrix} 0,27 \\ 0,96 \\ 0,42 \end{bmatrix}$	graus
$\begin{bmatrix} 15,0 \\ 0,0 \\ -15,0 \end{bmatrix}$	$\begin{bmatrix} 11,12 \\ -0,17 \\ -15,25 \end{bmatrix}$	3,88	cm	$\begin{bmatrix} 0,0 \\ -5,0 \\ 0,0 \end{bmatrix}$	$\begin{bmatrix} 0,31 \\ -5,38 \\ -0,09 \end{bmatrix}$	$\begin{bmatrix} 0,31 \\ 0,38 \\ 0,09 \end{bmatrix}$	graus
$\begin{bmatrix} 15,0 \\ 0,0 \\ -15,0 \end{bmatrix}$	$\begin{bmatrix} 10,79 \\ -0,16 \\ -15,48 \end{bmatrix}$	4,23	cm	$\begin{bmatrix} 0,0 \\ -10,0 \\ 0,0 \end{bmatrix}$	$\begin{bmatrix} -1,02 \\ -11,07 \\ 0,05 \end{bmatrix}$	$\begin{bmatrix} 1,02 \\ 1,07 \\ 0,04 \end{bmatrix}$	graus
$\begin{bmatrix} 15,0 \\ 0,0 \\ -15,0 \end{bmatrix}$	$\begin{bmatrix} 10,51 \\ -0,03 \\ -15,68 \end{bmatrix}$	4,54	cm	$\begin{bmatrix} 0,0 \\ -15,0 \\ 0,0 \end{bmatrix}$	$\begin{bmatrix} -1,81 \\ -16,89 \\ 0,49 \end{bmatrix}$	$\begin{bmatrix} 1,81 \\ 1,89 \\ 0,49 \end{bmatrix}$	graus

Capítulo 6

Conclusões

A partir dos primeiros testes, vemos a melhora na robustez da estimativa da matriz fundamental, onde o número de condição diminuiu de 10^{15} para 10^5 em média, conforme o gráfico da Figura 5.7. Através do gráfico da Figura 5.9, vemos como os erros geométricos das estimativas crescem rapidamente no Algoritmo de 8 Pontos, enquanto os algoritmos com normalização mostraram uma maior robustez ao ruído. Mesmo com o aumento do número de pontos de entrada, o erro permanece em média constante, como mostrado na Figura 5.8.

Com relação às estimações de movimento, podemos constatar que para movimentos com pequenas translações, todos os métodos mostram erros elevados, como podemos ver nas Figuras 5.10 e 5.11. Para movimento com rotação, a partir das Figuras 5.12, 5.13, 5.14, 5.15 e 5.16, podemos verificar que o Algoritmo de 8 Pontos apresenta maiores erros, principalmente para pequenas rotações, em comparação com os algoritmos normalizados.

Com relação às estimações de trajetórias, podemos constatar que os algoritmos normalizados apresentam erros menores, conforme as Figuras 5.21, 5.22 e 5.23 (associadas a normalização isotrópica) e 5.24, 5.25 e 5.26 (associadas a normalização não-isotrópica). Na trajetória dos testes, observamos que as estimativas através do Algoritmo de 8 Pontos tendem a divergir da trajetória, conforme vemos nas Figuras 5.18, 5.19 e 5.20.

O Algoritmo de 8 Pontos é mal condicionado, e essa técnica só se torna viável quando é usada a normalização dos dados de entrada como sugere Hartley [23].

Nos testes com imagens reais não obtivemos resultados tão bons quanto os sintéticos, mesmo com a normalização de Hartley. As estimações mostraram-se sensíveis a variações dos dados de entrada. Observamos que, à medida em que nos afastávamos da região de calibração da câmera, o erro das estimativas aumentava. Uma rotina de autocalibração, como por exemplo [28], poderia ser uma solução para diminuir esse erro.

Ambas as técnicas de normalização do Algoritmo de 8 Pontos podem ser utilizadas como técnicas para uma estimativa inicial rápida do movimento. Com um ajuste não-linear usando técnicas como Levenberg-Marquardt ou Gradiente Descendente (exemplos podem ser vistos em [13] e [20], respectivamente), pode ser possível vir a melhorar a exatidão dos resultados do Algoritmo de 8 Pontos Normalizado.

Em trabalhos futuros, pode ser considerada a implementação de uma rotina de autocalibração e a avaliação da melhora dos algoritmos com um ajuste não-linear.

Referências Bibliográficas

- [1] BRADY, F. B. *A Singular View: The Art of Seeing With One Eye*. Medical Economics Company, 2004.
- [2] MOHR, R., TRIGGS, B. “Projective Geometry for Image Analysis”, *Int. Symp. of Photogrammetry and Remote Sensing*, Julho 1996.
- [3] TYLER, C. W. “Perspective Studies”. , data de acesso - Março 2010. Disponível em: www.ski.org/CWTyler_lab/CWTyler/Art%20Investigations/Perspective/.
- [4] CARVALHO, M. M. S. *Introdução à Geometria Projetiva*. Publicação, Departamento de Matemática, Faculdade Nacional de Filosofia - Universidade do Brasil, 1953.
- [5] SEMPLE, J. G., KNEEBONE, G. T. *Algebraic Projective Geometry*. Oxford University Press, 1952.
- [6] WIKIPEDIA. “Camera Obscura”. , data de acesso - Março 2010. Disponível em: http://en.wikipedia.org/wiki/Camera_obscura.
- [7] WILGUS, J., WILGUS, B. “The Magic Mirror of Life: An Appreciation of the Camera Obscura - What is a Camera Obscura?” , data de acesso - Março 2010. Disponível em: <http://brightbytes.com/cosite/what.html>.
- [8] ALONSO, M., FINN, E. J. *Física*. Addison-Wesley, 1999.
- [9] LIPSCHUTZ, S. *Álgebra Linear: Teoria e Problemas*. Makron Books, 1994.
- [10] FAUGERAS, O., LUONG, Q.-T., PAPADOPOULOU, T. *The Geometry of Multiple Images: The Laws That Govern The Formation of Images of A Scene and Some of Their Applications*. MIT Press, 2001.
- [11] SILVA, L. C. *Método Robusto para a Calibração de Câmeras em Estereofotogrametria*. Tese de doutorado, COPPE - Instituto Alberto Luiz Coimbra de Pós-Graduação e Pesquisa de Engenharia, Universidade Federal do Rio de Janeiro, 2003.

- [12] CRAIG, J. J. *Introduction to Robotics: Mechanics and Control*. Prentice Hall, 2005.
- [13] HARTLEY, R., ZISSERMAN, A. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2003.
- [14] TSAI, R. Y. “An Efficient and Accurate Camera Calibration Technique for 3D Machine Vision”, *IEEE Conf. Computer Vision and Pattern Recognition*, pp. 364–374, 1986.
- [15] ZHANG, Z. “A Flexible New Technique for Camera Calibration”, *IEEE Trans. Pattern Anal. Mach. Intell.*, v. 22, n. 11, pp. 1330–1334, 2000.
- [16] FAUGERAS, O. *Three-dimensional Computer Vision: a Geometric Viewpoint*. MIT Press, 1993.
- [17] TRUCCO, E., VERRI, A. *Introductory Techniques for 3-D Computer Vision*. Prentice Hall, 1998.
- [18] LONGUET-HIGGINS, H. C. “A Computer Algorithm for Reconstructing a Scene from Two Projections”, *Nature*, v. 293, pp. 133–135, Setembro 1981.
- [19] HUANG, T. S., FAUGERAS, O. D. “Some Properties of the E Matrix in Two-View Motion Estimation”, *IEEE Trans. Pattern Anal. Mach. Intell.*, v. 11, n. 12, pp. 1310–1312, 1989.
- [20] MA, Y., SOATTO, S., KOSECKA, J., et al. *An Invitation to 3-D Vision: From Images to Geometric Models*. Springer-Verlag, 2003.
- [21] LUONG, Q.-T., DERICHE, R., FAUGERAS, O., et al. *On Determining the Fundamental Matrix: Analysis of Different Methods and Experimental Results*. Relatório técnico, INRIA, Abril 1993.
- [22] ZHANG, Z. “Determining the Epipolar Geometry and its Uncertainty: A Review”, *Int. Journal of Computer Vision*, v. 27, n. 2, pp. 161–195, 1998.
- [23] HARTLEY, R. I. “In Defense of the Eight-Point Algorithm”, *IEEE Trans. Pattern Anal. Mach. Intell.*, v. 19, n. 6, pp. 580–593, 1997.
- [24] STRANG, G. *Linear Algebra and Its Applications*. Thomson Learning, 1988.
- [25] GOLUB, G. H., VAN LOAN, C. F. *Matrix Computations*. 1996.
- [26] SUN, C. “2-Point linear algorithm for camera translation vector estimation with known rotation”, *Robotica*, v. 18, n. 5, pp. 557–561, 2000.

- [27] MARIOTTINI, G. L., PRATTICHIZZO, D. “EGT for multiple view geometry and visual servoing: robotics vision with pinhole and panoramic cameras”, *IEEE Robotics and Automation Magazine*, v. 12, n. 4, pp. 26–39, Dezembro 2005.
- [28] MAYBANK, S. J., FAUGERAS, O. D. “A Theory of Self-calibration of a Moving Camera”, *Int. Journal of Computer Vision*, v. 8, n. 2, pp. 123–151, 1992.

Apêndice A

Provas de Teoremas, Fórmulas e Equações

Teorema 1.1 (p. 5)

Sendo $\mathbf{r} = [r_x, r_y, r_z]^T$ e $\mathbf{p} = [p_x, p_y, p_z]^T$.

A equação da reta \mathbf{r} conforme Equação (1.2):

$$r_x x + r_y y + r_z z = 0$$

Substituindo as coordenadas x , y e z pelo ponto \mathbf{p} :

$$r_x p_x + r_y p_y + r_z p_z = \mathbf{r}^T \mathbf{p} = \mathbf{p}^T \mathbf{r} = \mathbf{r} \cdot \mathbf{p}$$

Logo, $\mathbf{r} \cdot \mathbf{p} = 0$ corresponderá a um ponto \mathbf{p} na equação da reta \mathbf{r} .

Teorema 1.2 (p. 5)

Sendo $\mathbf{r}^A = [r_x^A, r_y^A, r_z^A]^T$, $\mathbf{r}^B = [r_x^B, r_y^B, r_z^B]^T$ e $\mathbf{p} = [p_x, p_y, p_z]^T$.

As equações das retas \mathbf{r}^A e \mathbf{r}^B conforme Equação (1.2) ficam:

$$\begin{aligned} r_x^A x + r_y^A y + r_z^A z &= 0 \\ r_x^B x + r_y^B y + r_z^B z &= 0 \end{aligned}$$

O ponto \mathbf{p} é o ponto de encontro entre as duas retas \mathbf{r}^A e \mathbf{r}^B , logo obtemos o seguinte sistema linear:

$$\begin{aligned} r_x^A p_x + r_y^A p_y + r_z^A p_z &= 0 \\ r_x^B p_x + r_y^B p_y + r_z^B p_z &= 0 \end{aligned}$$

O resultado de $\mathbf{r}^A \times \mathbf{r}^B$ obtemos:

$$\mathbf{r}^A \times \mathbf{r}^B = \begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ r_x^A & r_y^A & r_z^A \\ r_x^B & r_y^B & r_z^B \end{vmatrix} = \left[(r_y^A r_z^B - r_z^A r_y^B), (r_z^A r_x^B - r_x^A r_z^B), (r_x^A r_y^B - r_y^A r_x^B) \right]^T$$

Se substituirmos $\mathbf{r}^A \times \mathbf{r}^B$ no sistema linear acima, veremos que $\mathbf{r}^A \times \mathbf{r}^B$ é solução para o sistema. Logo, $\boxed{\mathbf{p} = \mathbf{r}^A \times \mathbf{r}^B}$.

Teorema 1.3 (p. 5)

Sejam $\mathbf{p}^A = [p_x^A, p_y^A, p_z^A]^T$, $\mathbf{p}^B = [p_x^B, p_y^B, p_z^B]^T$ e $\mathbf{r} = [r_x, r_y, r_z]^T$.

Os pontos \mathbf{p}^A e \mathbf{p}^B fazem parte da reta \mathbf{r} , logo obtemos o seguinte sistema linear:

$$\begin{aligned} p_x^A r_x + p_y^A r_y + p_z^A r_z &= 0 \\ p_x^B r_x + p_y^B r_y + p_z^B r_z &= 0 \end{aligned}$$

O resultado de $\mathbf{p}^A \times \mathbf{p}^B$ é:

$$\mathbf{p}^A \times \mathbf{p}^B = \begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ p_x^A & p_y^A & p_z^A \\ p_x^B & p_y^B & p_z^B \end{vmatrix} = \left[(p_y^A p_z^B - p_z^A p_y^B), (p_z^A p_x^B - p_x^A p_z^B), (p_x^A p_y^B - p_y^A p_x^B) \right]^T$$

Se substituirmos $\mathbf{p}^A \times \mathbf{p}^B$ no sistema linear acima, veremos que $\mathbf{p}^A \times \mathbf{p}^B$ é solução para o sistema. Logo, $\boxed{\mathbf{r} = \mathbf{p}^A \times \mathbf{p}^B}$.

Fórmula do ponto ideal (p. 5)

Duas retas paralelas em \mathbb{P}^2 possuem a relação entre as suas duas primeiras componentes iguais segundo um fator de escala. Sendo assim, consideremos as retas paralelas $\mathbf{r}^A = [r_x, r_y, k_1]^T$, $\mathbf{r}^B = [r_x, r_y, k_2]^T$, onde $k_1 \neq k_2$.

Pelo Teorema 1.2, podemos dizer que $\mathbf{m}_\infty = \mathbf{r}^A \times \mathbf{r}^B$. Logo,

$$\mathbf{m}_\infty = \begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ r_x & r_y & k_1 \\ r_x & r_y & k_2 \end{vmatrix} = (k_2 - k_1) [-r_y, r_x, 0]^T$$

Conseqüentemente, podemos considerar $\boxed{\mathbf{m}_\infty = [a, b, 0]^T}$

Fórmula da reta ideal (p. 5)

Consideremos os pontos ideais $\mathbf{m}_\infty^A = [p_x^A, p_y^A, 0]^T$, $\mathbf{m}_\infty^B = [p_x^B, p_y^B, 0]^T$.

Pelo Teorema 1.3, podemos dizer que $\mathbf{r}_\infty = \mathbf{m}_\infty^A \times \mathbf{m}_\infty^B$. Logo,

$$\mathbf{r}_\infty = \begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ p_x^A & p_y^A & 0 \\ p_x^B & p_y^B & 0 \end{vmatrix} = [0, 0, (p_x^A p_y^B - p_y^A p_x^B)]^T$$

Conseqüentemente, podemos considerar $\boxed{\mathbf{r}_\infty = [0, 0, c]^T}$.

Fórmula TPP para retas (p. 6)

Sabemos pelo Teorema 1.1, que $\mathbf{r}'^T \mathbf{p}' = 0$ e $\mathbf{r}^T \mathbf{p} = 0$, logo, $\mathbf{r}'^T \mathbf{p}' = \mathbf{r}^T \mathbf{p}$.

Sabemos pela Equação (1.3), que $\mathbf{p}' = \mathbf{H}\mathbf{p}$.

Substituindo \mathbf{p}' na equação temos:

$$\begin{aligned} (\mathbf{r}'^T \mathbf{H})\mathbf{p} &= (\mathbf{r}^T)\mathbf{p} \quad \rightarrow \quad \mathbf{r}'^T \mathbf{H} = \mathbf{r}^T \\ \mathbf{H}^T \mathbf{r}' &= \mathbf{r} \\ \boxed{\mathbf{r}' = \mathbf{H}^{-T} \mathbf{r}} \end{aligned}$$

Fórmula TPE3 para planos (p. 7)

Podemos pegar o Teorema 1.1 no \mathbb{P}^2 para os elementos duais ponto e reta, e expressá-lo de forma semelhante em \mathbb{P}^3 para os elementos duais ponto e plano, obtendo $\pi'^T \mathbf{M}' = 0$ e $\pi^T \mathbf{M} = 0$, logo, $\pi'^T \mathbf{M}' = \pi^T \mathbf{M}$.

Sabemos pela Equação (1.7), que $\mathbf{M}' = \mathbf{H}\mathbf{M}$.

Substituindo \mathbf{M}' na equação temos:

$$\begin{aligned} (\pi'^T \mathbf{H})\mathbf{M} &= (\pi^T)\mathbf{M} \quad \rightarrow \quad \pi'^T \mathbf{H} = \pi^T \\ \mathbf{H}^T \pi' &= \pi \\ \boxed{\pi' = \mathbf{H}^{-T} \pi} \end{aligned}$$

Equações dos parâmetros do modelo simples - Tabela 2.1 (p. 23)

Matriz de projeção:

$$P = \begin{bmatrix} \epsilon \mathbf{q}_1^T & \epsilon q_{14} \\ \epsilon \mathbf{q}_2^T & \epsilon q_{24} \\ \epsilon \mathbf{q}_3^T & \epsilon q_{34} \end{bmatrix} = \begin{bmatrix} \alpha_u \bar{\mathbf{r}}_1 + u_0 \bar{\mathbf{r}}_3 & \alpha_u t_x + u_0 t_z \\ \alpha_v \bar{\mathbf{r}}_2 + v_0 \bar{\mathbf{r}}_3 & \alpha_v t_y + v_0 t_z \\ \bar{\mathbf{r}}_3 & t_z \end{bmatrix}$$

Podemos, logo de início, obter os seguintes parâmetros: $\bar{\mathbf{r}}_3 = \epsilon \mathbf{q}_3^T$ e $t_z = \epsilon q_{34}$

1. Coordenadas do centro do sensor óptico

- u_0 :

$$\begin{aligned} \begin{cases} \alpha_u \bar{\mathbf{r}}_1 + u_0 \bar{\mathbf{r}}_3 = \epsilon \mathbf{q}_1^T \\ \bar{\mathbf{r}}_3 = \epsilon \mathbf{q}_3^T \end{cases} \\ \alpha_u \overbrace{(\bar{\mathbf{r}}_1 \cdot \bar{\mathbf{r}}_3)}{=0} + u_0 \overbrace{(\bar{\mathbf{r}}_3 \cdot \bar{\mathbf{r}}_3)}{=1} = \overbrace{\epsilon^2}^{=1} \mathbf{q}_1^T \mathbf{q}_3 \\ \boxed{u_0 = \mathbf{q}_1^T \mathbf{q}_3} \end{aligned}$$

- v_0 :

$$\begin{cases} \alpha_v \bar{\mathbf{r}}_2 + v_0 \bar{\mathbf{r}}_3 = \epsilon \mathbf{q}_2^T \\ \bar{\mathbf{r}}_3 = \epsilon \mathbf{q}_3^T \end{cases}$$

$$\alpha_v \overbrace{(\bar{\mathbf{r}}_2 \cdot \bar{\mathbf{r}}_3)}{=0} + v_0 \overbrace{(\bar{\mathbf{r}}_3 \cdot \bar{\mathbf{r}}_3)}{=1} = \overbrace{\epsilon^2}^{=1} \mathbf{q}_2^T \mathbf{q}_3$$

$$v_0 = \mathbf{q}_2^T \mathbf{q}_3$$

2. Comprimento da altura do cone de projeção nos eixos u e v

- α_u :

$$\alpha_u \bar{\mathbf{r}}_1 + u_0 \bar{\mathbf{r}}_3 = \epsilon \mathbf{q}_1^T$$

$$(\alpha_u \bar{\mathbf{r}}_1 + u_0 \bar{\mathbf{r}}_3)^2 = (\epsilon \mathbf{q}_1^T)^2$$

$$\alpha_u^2 \overbrace{(\bar{\mathbf{r}}_1 \cdot \bar{\mathbf{r}}_1)}{=1} + 2\alpha_u u_0 \overbrace{(\bar{\mathbf{r}}_1 \cdot \bar{\mathbf{r}}_3)}{=0} + u_0^2 \overbrace{(\bar{\mathbf{r}}_3 \cdot \bar{\mathbf{r}}_3)}{=1} = \overbrace{\epsilon^2}^{=1} \mathbf{q}_1^T \mathbf{q}_1$$

$$\alpha_u = \sqrt{\mathbf{q}_1^T \mathbf{q}_1 - u_0^2}$$

- α_v :

$$\alpha_v \bar{\mathbf{r}}_2 + v_0 \bar{\mathbf{r}}_3 = \epsilon \mathbf{q}_2^T$$

$$(\alpha_v \bar{\mathbf{r}}_2 + v_0 \bar{\mathbf{r}}_3)^2 = (\epsilon \mathbf{q}_2^T)^2$$

$$\alpha_v^2 \overbrace{(\bar{\mathbf{r}}_2 \cdot \bar{\mathbf{r}}_2)}{=1} + 2\alpha_v v_0 \overbrace{(\bar{\mathbf{r}}_2 \cdot \bar{\mathbf{r}}_3)}{=0} + v_0^2 \overbrace{(\bar{\mathbf{r}}_3 \cdot \bar{\mathbf{r}}_3)}{=1} = \overbrace{\epsilon^2}^{=1} \mathbf{q}_2^T \mathbf{q}_2$$

$$\alpha_v = \sqrt{\mathbf{q}_2^T \mathbf{q}_2 - v_0^2}$$

3. Vetor de translação \mathbf{t}

- t_z :

$t_z = \epsilon q_{34}$
- t_x :

$$\begin{cases} \alpha_u t_x + u_0 t_z = \epsilon q_{14} \\ t_z = \epsilon q_{34} \end{cases}$$

$$t_x = \epsilon \frac{q_{14} - u_0 q_{34}}{\alpha_u}$$
- t_y :

$$\begin{cases} \alpha_v t_y + v_0 t_z = \epsilon q_{24} \\ t_z = \epsilon q_{34} \end{cases}$$

$$t_y = \epsilon \frac{q_{24} - v_0 q_{34}}{\alpha_v}$$

4. Matriz de rotação \mathbf{R}

- $\bar{\mathbf{r}}_3$:

$\bar{\mathbf{r}}_3 = \epsilon \mathbf{q}_3^T$

- $\bar{\mathbf{r}}_1$:

$$\begin{cases} \alpha_u \bar{\mathbf{r}}_1 + u_0 \bar{\mathbf{r}}_3 = \epsilon \mathbf{q}_1^T \\ \bar{\mathbf{r}}_3 = \epsilon \mathbf{q}_3^T \end{cases}$$

$$\bar{\mathbf{r}}_1 = \epsilon \frac{\mathbf{q}_1^T - u_0 \mathbf{q}_3^T}{\alpha_u}$$

- $\bar{\mathbf{r}}_2$:

$$\begin{cases} \alpha_v \bar{\mathbf{r}}_2 + v_0 \bar{\mathbf{r}}_3 = \epsilon \mathbf{q}_2^T \\ \bar{\mathbf{r}}_3 = \epsilon \mathbf{q}_3^T \end{cases}$$

$$\bar{\mathbf{r}}_2 = \epsilon \frac{\mathbf{q}_2^T - v_0 \mathbf{q}_3^T}{\alpha_v}$$

Equações dos parâmetros do modelo geral - Tabela 2.2 (p. 23)

Matriz de projeção:

$$P = \begin{bmatrix} \epsilon \mathbf{q}_1^T & \epsilon q_{14} \\ \epsilon \mathbf{q}_2^T & \epsilon q_{24} \\ \epsilon \mathbf{q}_3^T & \epsilon q_{34} \end{bmatrix} = \begin{bmatrix} \alpha_u \bar{\mathbf{r}}_1 - \alpha_u \cot \theta \bar{\mathbf{r}}_2 + u_0 \bar{\mathbf{r}}_3 & \alpha_u t_x - \alpha_u \cot \theta t_y + u_0 t_z \\ \frac{\alpha_v}{\sin \theta} \bar{\mathbf{r}}_2 + v_0 \bar{\mathbf{r}}_3 & \frac{\alpha_v}{\sin \theta} t_y + v_0 t_z \\ \bar{\mathbf{r}}_3 & t_z \end{bmatrix}$$

Podemos, logo de início, obter os seguintes parâmetros: $\bar{\mathbf{r}}_3 = \epsilon \mathbf{q}_3^T$ e $t_z = \epsilon q_{34}$

1. Coordenadas do centro do sensor óptico

- u_0 :

$$\begin{cases} \alpha_u \bar{\mathbf{r}}_1 - \alpha_u \cot \theta \bar{\mathbf{r}}_2 + u_0 \bar{\mathbf{r}}_3 = \epsilon \mathbf{q}_1^T \\ \bar{\mathbf{r}}_3 = \epsilon \mathbf{q}_3^T \end{cases}$$

$$\alpha_u \overbrace{(\bar{\mathbf{r}}_1 \cdot \bar{\mathbf{r}}_3)}^{=0} - \alpha_u \cot \theta \overbrace{(\bar{\mathbf{r}}_2 \cdot \bar{\mathbf{r}}_3)}^{=0} + u_0 \overbrace{(\bar{\mathbf{r}}_3 \cdot \bar{\mathbf{r}}_3)}^{=1} = \overbrace{\epsilon^2}^{=1} \mathbf{q}_1^T \mathbf{q}_3$$

$$u_0 = \mathbf{q}_1^T \mathbf{q}_3$$

- v_0 :

$$\begin{cases} \frac{\alpha_v}{\sin \theta} \bar{\mathbf{r}}_2 + v_0 \bar{\mathbf{r}}_3 = \epsilon \mathbf{q}_2^T \\ \bar{\mathbf{r}}_3 = \epsilon \mathbf{q}_3^T \end{cases}$$

$$\frac{\alpha_v}{\sin \theta} \overbrace{(\bar{\mathbf{r}}_2 \cdot \bar{\mathbf{r}}_3)}^{=0} + v_0 \overbrace{(\bar{\mathbf{r}}_3 \cdot \bar{\mathbf{r}}_3)}^{=1} = \overbrace{\epsilon^2}^{=1} \mathbf{q}_2^T \mathbf{q}_3$$

$$v_0 = \mathbf{q}_2^T \mathbf{q}_3$$

2. Comprimento da altura do cone de projeção nos eixos u e v

- α_v :

$$\frac{\alpha_v}{\sin \theta} \bar{\mathbf{r}}_2 + v_0 \bar{\mathbf{r}}_3 = \epsilon \mathbf{q}_2^T$$

$$\left(\frac{\alpha_v}{\sin \theta} \bar{\mathbf{r}}_2 + v_0 \bar{\mathbf{r}}_3 \right)^2 = (\epsilon \mathbf{q}_2^T)^2$$

$$\frac{\alpha_v^2}{\sin^2 \theta} \overbrace{(\bar{\mathbf{r}}_2 \cdot \bar{\mathbf{r}}_2)}^{=1} + 2 \frac{\alpha_v}{\sin \theta} v_0 \overbrace{(\bar{\mathbf{r}}_2 \cdot \bar{\mathbf{r}}_3)}^{=0} + v_0^2 \overbrace{(\bar{\mathbf{r}}_3 \cdot \bar{\mathbf{r}}_3)}^{=1} = \overbrace{\epsilon^2}^{=1} \mathbf{q}_2^T \mathbf{q}_2$$

$$\alpha_v = \sqrt{\mathbf{q}_2^T \mathbf{q}_2 - v_0^2 \sin^2 \theta}$$

- α_u :

$$\alpha_u \bar{\mathbf{r}}_1 - \alpha_u \cot \theta \bar{\mathbf{r}}_2 + u_0 \bar{\mathbf{r}}_3 = \epsilon \mathbf{q}_1^T$$

$$[\alpha_u (\bar{\mathbf{r}}_1 - \cot \theta \bar{\mathbf{r}}_2) + u_0 \bar{\mathbf{r}}_3]^2 = (\epsilon \mathbf{q}_1^T)^2$$

$$\alpha_u^2 (\bar{\mathbf{r}}_1 - \cot \theta \bar{\mathbf{r}}_2)^2 + 2\alpha_u u_0 \overbrace{[(\bar{\mathbf{r}}_1 \cdot \bar{\mathbf{r}}_3) - \cot \theta (\bar{\mathbf{r}}_2 \cdot \bar{\mathbf{r}}_3)]}^{=0} + u_0^2 \overbrace{(\bar{\mathbf{r}}_3 \cdot \bar{\mathbf{r}}_3)}^{=1} = \overbrace{\epsilon^2}^{=1} \mathbf{q}_1^T \mathbf{q}_1$$

$$\alpha_u^2 \overbrace{[(\bar{\mathbf{r}}_1 \cdot \bar{\mathbf{r}}_1)]}^{=1} - 2 \cot \theta \overbrace{[(\bar{\mathbf{r}}_1 \cdot \bar{\mathbf{r}}_2)]}^{=0} + \cot^2 \theta \overbrace{[(\bar{\mathbf{r}}_2 \cdot \bar{\mathbf{r}}_2)]}^{=1} = \mathbf{q}_1^T \mathbf{q}_1 - u_0^2$$

$$\alpha_u^2 \overbrace{[(\sin^2 \theta + \cos^2 \theta) / \sin^2 \theta]}^{=1} = \mathbf{q}_1^T \mathbf{q}_1 - u_0^2$$

$$\boxed{\alpha_u = \sqrt{\mathbf{q}_1^T \mathbf{q}_1 - u_0^2} \sin \theta}$$

3. Vetor de translação \mathbf{t}

- t_z : $\boxed{t_z = \epsilon q_{34}}$

- t_y :

$$\begin{cases} \frac{\alpha_v}{\sin \theta} t_y + v_0 t_z = \epsilon q_{24} \\ t_z = \epsilon q_{34} \end{cases}$$

$$\boxed{t_y = \epsilon (q_{24} - v_0 q_{34}) \frac{\sin \theta}{\alpha_v}}$$

- t_x :

$$\begin{cases} \alpha_u t_x - \alpha_u \cot \theta t_y + u_0 t_z = \epsilon q_{14} \\ t_y = \epsilon (q_{24} - v_0 q_{34}) \frac{\sin \theta}{\alpha_v} \\ t_z = \epsilon q_{34} \end{cases}$$

$$\alpha_u (t_x - \cot \theta t_y) = \epsilon (q_{14} - u_0 q_{34})$$

$$\boxed{t_x = \epsilon [q_{14} + (q_{24} - v_0 q_{34}) \frac{\alpha_u}{\alpha_v} \cos \theta - u_0 q_{34}] \frac{1}{\alpha_u}}$$

4. Matriz de rotação \mathbf{R}

- $\bar{\mathbf{r}}_3$: $\boxed{\bar{\mathbf{r}}_3 = \epsilon \mathbf{q}_3^T}$

- $\bar{\mathbf{r}}_2$:

$$\begin{cases} \frac{\alpha_v}{\sin \theta} \bar{\mathbf{r}}_2 + v_0 \bar{\mathbf{r}}_3 = \epsilon \mathbf{q}_2^T \\ \bar{\mathbf{r}}_3 = \epsilon \mathbf{q}_3^T \end{cases}$$

$$\boxed{\bar{\mathbf{r}}_2 = \epsilon (\mathbf{q}_2^T - v_0 \mathbf{q}_3^T) \frac{\sin \theta}{\alpha_v}}$$

- $\bar{\mathbf{r}}_1$:

$$\begin{cases} \alpha_u \bar{\mathbf{r}}_1 - \alpha_u \cot \theta \bar{\mathbf{r}}_2 + u_0 \bar{\mathbf{r}}_3 = \epsilon \mathbf{q}_1^T \\ \bar{\mathbf{r}}_2 = \epsilon (\mathbf{q}_2^T - v_0 \mathbf{q}_3^T) \frac{\sin \theta}{\alpha_v} \\ \bar{\mathbf{r}}_3 = \epsilon \mathbf{q}_3^T \end{cases}$$

$$\alpha_u (\bar{\mathbf{r}}_1 - \cot \theta \bar{\mathbf{r}}_2) = \epsilon (\mathbf{q}_1^T - u_0 \mathbf{q}_3^T)$$

$$\boxed{\bar{\mathbf{r}}_1 = \epsilon [\mathbf{q}_1^T + (\mathbf{q}_2^T - v_0 \mathbf{q}_3^T) \frac{\alpha_u}{\alpha_v} \cos \theta - u_0 \mathbf{q}_3^T] \frac{1}{\alpha_u}}$$

5. Ângulo entre os eixos da imagem u e v

• θ :

Como: $\|\mathbf{q}_3\|^2 = 1$

$$\text{Fórmula: } \boxed{\|a \times b\|^2 + \|a \cdot b\|^2 = \|a\|^2 \|b\|^2}$$

Logo,

$$\alpha_u = \sqrt{\mathbf{q}_1^T \mathbf{q}_1 - u_0^2} \sin \theta$$

$$\alpha_u = \sqrt{\mathbf{q}_1^T \mathbf{q}_1 - (\mathbf{q}_1^T \mathbf{q}_3)^2} \sin \theta$$

$$\alpha_u = \sqrt{\|\mathbf{q}_1\|^2 \|\mathbf{q}_3\|^2 - \|\mathbf{q}_1 \cdot \mathbf{q}_3\|^2} \sin \theta$$

$$\boxed{\alpha_u = \|\mathbf{q}_1 \times \mathbf{q}_3\| \sin \theta}$$

$$\alpha_v = \sqrt{\mathbf{q}_2^T \mathbf{q}_2 - v_0^2} \sin \theta$$

$$\alpha_v = \sqrt{\mathbf{q}_2^T \mathbf{q}_2 - (\mathbf{q}_2^T \mathbf{q}_3)^2} \sin \theta$$

$$\alpha_v = \sqrt{\|\mathbf{q}_2\|^2 \|\mathbf{q}_3\|^2 - \|\mathbf{q}_2 \cdot \mathbf{q}_3\|^2} \sin \theta$$

$$\boxed{\alpha_v = \|\mathbf{q}_2 \times \mathbf{q}_3\| \sin \theta}$$

$$\text{Fórmula de Lagrange: } \boxed{a \times (b \times c) = b(a \cdot c) - c(a \cdot b)}$$

$$\bar{\mathbf{r}}_2 = \epsilon (\mathbf{q}_2^T - v_0 \mathbf{q}_3^T) \frac{\sin \theta}{\alpha_v}$$

$$\bar{\mathbf{r}}_2 = \epsilon (\mathbf{q}_2^T - \mathbf{q}_2^T \mathbf{q}_3 \mathbf{q}_3^T) \frac{\sin \theta}{\alpha_v}$$

$$\bar{\mathbf{r}}_2^T = \epsilon [\mathbf{q}_2 (\underbrace{\mathbf{q}_3 \cdot \mathbf{q}_3}_{\|\mathbf{q}_3\|^2=1}) - \mathbf{q}_3 (\mathbf{q}_3 \cdot \mathbf{q}_2)] \frac{\sin \theta}{\alpha_v}$$

$$\boxed{\bar{\mathbf{r}}_2^T = \epsilon \left[\mathbf{q}_3 \times \frac{(\mathbf{q}_2 \times \mathbf{q}_3)}{\|\mathbf{q}_2 \times \mathbf{q}_3\|} \right]}$$

$$\bar{\mathbf{r}}_1 = \epsilon [\mathbf{q}_1^T + (\mathbf{q}_2^T - v_0 \mathbf{q}_3^T) \frac{\alpha_u}{\alpha_v} \cos \theta - u_0 \mathbf{q}_3^T] \frac{1}{\alpha_u}$$

$$\bar{\mathbf{r}}_1 = \epsilon \left[(\mathbf{q}_1^T - \mathbf{q}_1^T \mathbf{q}_3 \mathbf{q}_3^T) \frac{1}{\alpha_u} + (\mathbf{q}_2^T - \mathbf{q}_2^T \mathbf{q}_3 \mathbf{q}_3^T) \frac{1}{\alpha_v} \cos \theta \right]$$

$$\bar{\mathbf{r}}_1^T = \epsilon \left[(\mathbf{q}_1 - \mathbf{q}_3 \mathbf{q}_3^T \mathbf{q}_1) \frac{1}{\alpha_u} + (\mathbf{q}_2 - \mathbf{q}_3 \mathbf{q}_3^T \mathbf{q}_2) \frac{1}{\alpha_v} \cos \theta \right]$$

$$\bar{\mathbf{r}}_1^T = \epsilon \left[(\mathbf{q}_1 (\underbrace{\mathbf{q}_3 \cdot \mathbf{q}_3}_{\|\mathbf{q}_3\|^2=1}) - \mathbf{q}_3 (\mathbf{q}_3 \cdot \mathbf{q}_1)) \frac{1}{\alpha_u} + (\mathbf{q}_2 (\underbrace{\mathbf{q}_3 \cdot \mathbf{q}_3}_{\|\mathbf{q}_3\|^2=1}) - \mathbf{q}_3 (\mathbf{q}_3 \cdot \mathbf{q}_2)) \frac{1}{\alpha_v} \cos \theta \right]$$

$$\boxed{\bar{\mathbf{r}}_1^T = \frac{\epsilon}{\sin \theta} \left[\mathbf{q}_3 \times \frac{(\mathbf{q}_1 \times \mathbf{q}_3)}{\|\mathbf{q}_1 \times \mathbf{q}_3\|} + \mathbf{q}_3 \times \frac{(\mathbf{q}_2 \times \mathbf{q}_3)}{\|\mathbf{q}_2 \times \mathbf{q}_3\|} \cos \theta \right]}$$

Como: $\bar{\mathbf{r}}_1^T \cdot \bar{\mathbf{r}}_2^T = 0$

$$\left(\mathbf{q}_3 \times \frac{(\mathbf{q}_1 \times \mathbf{q}_3)}{\|\mathbf{q}_1 \times \mathbf{q}_3\|} \right) \cdot \left(\mathbf{q}_3 \times \frac{(\mathbf{q}_2 \times \mathbf{q}_3)}{\|\mathbf{q}_2 \times \mathbf{q}_3\|} \right) + \overbrace{\left(\mathbf{q}_3 \times \frac{(\mathbf{q}_2 \times \mathbf{q}_3)}{\|\mathbf{q}_2 \times \mathbf{q}_3\|} \right) \cdot \left(\mathbf{q}_3 \times \frac{(\mathbf{q}_2 \times \mathbf{q}_3)}{\|\mathbf{q}_2 \times \mathbf{q}_3\|} \right)}^{=1} \cos \theta = 0$$

$$\cos \theta = -\frac{1}{\|\mathbf{q}_1 \times \mathbf{q}_3\| \|\mathbf{q}_2 \times \mathbf{q}_3\|} \left[\mathbf{q}_3 \times (\mathbf{q}_1 \times \mathbf{q}_3) \cdot \mathbf{q}_3 \times (\mathbf{q}_2 \times \mathbf{q}_3) \right]$$

Fórmula: $\boxed{a \times b \cdot c = a \cdot b \times c}$

$$\cos \theta = -\frac{1}{\|\mathbf{q}_1 \times \mathbf{q}_3\| \|\mathbf{q}_2 \times \mathbf{q}_3\|} \left[\mathbf{q}_3 \cdot (\mathbf{q}_1 \times \mathbf{q}_3) \times \mathbf{q}_3 \times (\mathbf{q}_2 \times \mathbf{q}_3) \right]$$

Usando a fórmula de Lagrange:

$$\cos \theta = -\frac{1}{\|\mathbf{q}_1 \times \mathbf{q}_3\| \|\mathbf{q}_2 \times \mathbf{q}_3\|} \left[\mathbf{q}_3 \cdot \left(\mathbf{q}_3 [(\mathbf{q}_1 \times \mathbf{q}_3) \cdot (\mathbf{q}_2 \times \mathbf{q}_3)] - (\mathbf{q}_2 \times \mathbf{q}_3) [(\mathbf{q}_1 \times \mathbf{q}_3) \cdot \mathbf{q}_3] \right) \right]$$

$$\cos \theta = -\frac{1}{\|\mathbf{q}_1 \times \mathbf{q}_3\| \|\mathbf{q}_2 \times \mathbf{q}_3\|} \left[\underbrace{\|\mathbf{q}_3\|^2}_{=1} \mathbf{q}_3 \cdot [(\mathbf{q}_1 \times \mathbf{q}_3) \cdot (\mathbf{q}_2 \times \mathbf{q}_3)] - \underbrace{\mathbf{q}_3 \cdot (\mathbf{q}_2 \times \mathbf{q}_3)}_{=0} [(\mathbf{q}_1 \times \mathbf{q}_3) \cdot \mathbf{q}_3] \right]$$

$$\boxed{\cos \theta = -\frac{(\mathbf{q}_1 \times \mathbf{q}_3) \cdot (\mathbf{q}_2 \times \mathbf{q}_3)}{\|\mathbf{q}_1 \times \mathbf{q}_3\| \|\mathbf{q}_2 \times \mathbf{q}_3\|}}$$

Fórmula do operador matriz de produto vetorial - $[\cdot]_{\times}$

Seja $\mathbf{a} = [a_x, a_y, a_z]^T$ e $\mathbf{b} = [b_x, b_y, b_z]^T$.

Podemos calcular o produto vetorial $\mathbf{a} \times \mathbf{b}$ da seguinte forma.

$$\mathbf{a} \times \mathbf{b} = \begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ a_x & a_y & a_z \\ b_x & b_y & b_z \end{vmatrix} = [(a_y b_z - a_z b_y), (a_z b_x - a_x b_z), (a_x b_y - a_y b_x)]^T$$

Se colocarmos esse produto na forma matricial do operador $([\cdot]_{\times})$ obtemos:

$$[\mathbf{a}]_{\times} \mathbf{b} = [(-a_z b_y + a_y b_z), (a_z b_x - a_x b_z), (-a_y b_x + a_x b_y)]^T$$

$$\boxed{[\mathbf{a}]_{\times} = \begin{bmatrix} 0 & -a_z & a_y \\ a_z & 0 & -a_x \\ -a_y & a_x & 0 \end{bmatrix}}$$

Apêndice B

Decomposição em Valores Singulares (SVD)

A decomposição em valores singulares (SVD) é uma das mais importantes decomposições de matrizes usadas em Visão Computacional. Como sabemos do conhecimento básico de Álgebra Linear, uma matriz quadrada \mathbf{A} pode ser diagonalizável em uma matriz diagonal onde sua diagonal principal são os autovalores de \mathbf{A} . Isto pode ser escrito como:

$$\mathbf{A} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^T \quad (\text{B.1})$$

onde:

\mathbf{Q} é uma matriz ortogonal dos autovetores, ou seja, a matriz com todos os vetores colunas (ou linhas) possuindo norma unitária e gerando um sistema de bases, ortogonais entre si

$\mathbf{\Lambda}$ é uma matriz diagonal cuja diagonal principal $[\lambda_1, \lambda_2, \dots, \lambda_n]$ são os autovalores de \mathbf{A}

A decomposição em autovalores só é válida para matrizes quadradas, ou seja, $\mathbf{A} \in \mathbb{R}^{n \times n}$. A SVD surgiu como uma decomposição, similar a de autovalores, contudo mais geral, de forma a incluir o caso de matrizes não quadradas, ou seja, $\mathbf{A} \in \mathbb{R}^{m \times n}$.

A seguir, iremos dar uma definição formal da SVD, descrever algumas de suas propriedades e por fim, explicar como e porque podemos usá-la para otimização.

B.1 Definição

Para qualquer matriz $\mathbf{A} \in \mathbb{R}^{m \times n}$ existe a seguinte decomposição:

$$\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^T \quad (\text{B.2})$$

onde:

\mathbf{U} é uma matriz ortogonal ($m \times m$), cujos vetores coluna são os *vetores singulares à esquerda* de \mathbf{A} .

\mathbf{S} é uma matriz diagonal ($m \times n$) com valores não negativos, cujos valores da diagonal não nula de \mathbf{S} são chamados de *valores singulares* de \mathbf{A} . Os valores singulares de \mathbf{S} são decrescentes, ou seja, $s_1 \geq s_2 \geq \dots \geq s_r > 0$.

\mathbf{V} é uma matriz ortogonal ($n \times n$), cujos vetores coluna são os *vetores singulares à direita* de \mathbf{A} .

B.2 Propriedades

A matriz diagonal \mathbf{S} é composta somente por elementos não-negativos. Os valores singulares de \mathbf{A} serão $[s_1, s_2, \dots, s_r]$ e preencherão r primeiros lugares da matriz diagonal principal de \mathbf{S} . Logo, r corresponderá ao posto da matriz \mathbf{A} .

As colunas da matriz ortogonal \mathbf{U} serão os autovetores da matriz quadrada $\mathbf{A}\mathbf{A}^T$.

$$\mathbf{A}\mathbf{A}^T = (\mathbf{U}\mathbf{S}\mathbf{V}^T)(\mathbf{V}\mathbf{S}^T\mathbf{U}^T) = \mathbf{U}\mathbf{S}\mathbf{S}^T\mathbf{U}^T \quad (\text{B.3})$$

As colunas da matriz ortogonal \mathbf{V} serão os autovetores da matriz quadrada $\mathbf{A}^T\mathbf{A}$.

$$\mathbf{A}^T\mathbf{A} = (\mathbf{V}\mathbf{S}^T\mathbf{U}^T)(\mathbf{U}\mathbf{S}\mathbf{V}^T) = \mathbf{V}\mathbf{S}^T\mathbf{S}\mathbf{V}^T \quad (\text{B.4})$$

Os autovalores da matriz quadrada $\mathbf{S}\mathbf{S}^T$ ($m \times m$) corresponderão a $s_1^2, s_2^2, \dots, s_r^2$ e $(m - r)$ zeros. Os autovalores da matriz quadrada $\mathbf{S}^T\mathbf{S}$ ($n \times n$) corresponderão a $s_1^2, s_2^2, \dots, s_r^2$ e $(n - r)$ zeros. Conseqüentemente, os valores singulares de \mathbf{S} corresponderão à raiz quadrada dos autovalores não nulos de $\mathbf{A}\mathbf{A}^T$ e $\mathbf{A}^T\mathbf{A}$ (por isso da não negatividade dos elementos de \mathbf{S}).

As colunas de \mathbf{U} e \mathbf{V} fornecerão as bases ortonormais para os quatro sub-espços fundamentais:

- primeiras r colunas de \mathbf{U} : espaço coluna de \mathbf{A}
- últimas $m - r$ colunas de \mathbf{U} : espaço nulo à esquerda de \mathbf{A}
- primeiras r colunas de \mathbf{V} : espaço linha de \mathbf{A}
- últimas $n - r$ colunas de \mathbf{V} : espaço nulo à direita de \mathbf{A}

B.3 Otimização

A SVD é excelente para cálculos computacionais estáveis. Uma das razões principais é o fato de decompor a matriz em \mathbf{U} e \mathbf{V} matrizes ortogonais. A multiplicação por uma matriz ortogonal nunca muda o comprimento de um vetor (propriedade de preservação da norma [24]: $\|\mathbf{Q}\mathbf{x}\| = \|\mathbf{x}\|$), como, por exemplo, em:

$$\|\mathbf{U}\mathbf{x}\|^2 = \mathbf{x}^T \mathbf{U}^T \mathbf{U} \mathbf{x} = \mathbf{x}^T \mathbf{x} = \|\mathbf{x}\|^2 \quad (\text{B.5})$$

A outra razão se deve ao fato da fácil obtenção da noção de grandeza dos valores singulares da matriz \mathbf{S} .

B.3.1 Otimização Linear por Mínimos Quadrados ($\mathbf{Ax} = \mathbf{b}$)

Podemos usar a SVD para solucionar sistemas de equações lineares do tipo $\mathbf{Ax} = \mathbf{b}$, sendo $\mathbf{A} \in \mathbb{R}^{m \times n}$ e $m \geq n$. Isto será equivalente a minimização da norma quadrada $\|\mathbf{Ax} - \mathbf{b}\|^2$, que corresponde a otimização linear por mínimos quadrados.

$$\min \|\mathbf{Ax} - \mathbf{b}\|^2 \Leftrightarrow \min \|\mathbf{Ax} - \mathbf{b}\|$$

$$\min \|\mathbf{Ax} - \mathbf{b}\| = \min \|\mathbf{USV}^T \mathbf{x} - \mathbf{b}\|$$

Usando a propriedade da preservação da norma, multiplicando por \mathbf{U}^T pela esquerda

$$\min \|\mathbf{Ax} - \mathbf{b}\| = \min \|\mathbf{SV}^T \mathbf{x} - \mathbf{U}^T \mathbf{b}\|$$

Substituindo $\mathbf{y} = \mathbf{V}^T \mathbf{x}$ e $\mathbf{b}^* = \mathbf{U}^T \mathbf{b}$, teremos o sistema $\mathbf{S}\mathbf{y} = \mathbf{b}^*$:

$$\begin{bmatrix} s_1 & & & \\ & \ddots & & \\ & & s_n & \\ \hline & & & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} b_1^* \\ \vdots \\ b_n^* \\ b_{n+1}^* \\ \vdots \\ b_m^* \end{bmatrix} \quad (\text{B.6})$$

A solução desse sistema será simplesmente $y_i = b_i^*/s_i$. Obviamente, $s_i \neq 0$, assumindo que o posto de \mathbf{A} deve ser n .

Por fim, a solução da otimização em mínimos quadrados será:

$$\mathbf{x} = \mathbf{V}\mathbf{y} \quad (\text{B.7})$$

B.3.2 Otimização Linear por Mínimos Quadrados ($\mathbf{Ax} = \mathbf{0}$)

Contudo, algumas vezes as equações do sistema linear $\mathbf{Ax} = \mathbf{b}$ estão na forma homogênea, sendo $\mathbf{b} = \mathbf{0}$. Então, nosso problema de minimização se torna $\min \|\mathbf{Ax}\|$. Devemos prestar atenção que pode ser obtida a solução trivial, ou seja, $x_i = 0$. Assim, vamos impor uma condição em \mathbf{x} para evitá-la. Enunciaremos o problema da seguinte forma:

Encontrar \mathbf{x} que $\min \|\mathbf{Ax}\|^2$ sendo $\|\mathbf{x}\|^2 = 1$ ou

Encontrar \mathbf{x} que $\min \|\mathbf{Ax}\|$ sendo $\|\mathbf{x}\| = 1$

$$\min \|\mathbf{Ax}\| = \min \|\mathbf{USV}^T \mathbf{x}\|$$

Usaremos, novamente, a propriedade da preservação da norma, multiplicando por \mathbf{U}^T pela esquerda

$$\min \|\mathbf{Ax}\| = \min \|\mathbf{SV}^T \mathbf{x}\|$$

Substituindo $\mathbf{y} = \mathbf{V}^T \mathbf{x}$, nosso problema será $\min \|\mathbf{Sy}\|$ sendo $\|\mathbf{y}\| = 1$.

Como a matriz diagonal \mathbf{S} possui seus valores singulares em ordem decrescente, então a solução mínima será $\mathbf{y} = [0, 0, \dots, 1]^T$.

Finalmente, sabendo que $\mathbf{x} = \mathbf{Vy}$, então a solução para \mathbf{x} será a última coluna da matriz \mathbf{V} . Em outras palavras, podemos dizer que a solução corresponde ao autovetor associado ao menor autovalor da matriz quadrada $\mathbf{A}^T \mathbf{A}$.

Apêndice C

Otimização com Restrição

A referência [16] mostra como realizar a otimização linear de mínimos quadrados de um sistema homogêneo do tipo $\mathbf{Ax} = \mathbf{0}$ dada uma restrição do tipo $\|\mathbf{Bx}\|^2 = 1$, sendo $m \times n$ a dimensão de \mathbf{A} (onde $m > n$), $n \times 1$ a dimensão de \mathbf{x} e $p \times n$ a dimensão de \mathbf{B} (onde $p < n$). Isso implica que selecionaremos somente alguns elementos de \mathbf{x} para compormos a restrição. É uma minimização mais restritiva do que aquela feita com SVD no Apêndice B.

Seja $\mathbf{z} = \mathbf{Bx}$ de dimensão $p \times 1$ e o vetor \mathbf{y} de dimensão $n - p \times 1$ composto pelas $n - p$ coordenadas restantes de \mathbf{x} . Temos:

$$\mathbf{Ax} = \mathbf{Cy} + \mathbf{Dz} \quad (\text{C.1})$$

onde:

\mathbf{C} é uma sub-matriz de \mathbf{A} ($m \times n - p$)

\mathbf{D} é uma sub-matriz de \mathbf{A} ($m \times p$)

O problema da otimização é representado agora da seguinte forma:

$$\min_{z,y} \|\mathbf{Cy} + \mathbf{Dz}\|^2 \quad \text{sendo } \|\mathbf{z}\|^2 = 1 \quad (\text{C.2})$$

Utilizando-se da técnica de multiplicadores de Lagrange, o problema é equivalente a:

$$\min_{z,y} \mathbf{H} = \|\mathbf{Cy} + \mathbf{Dz}\|^2 + \lambda(1 - \|\mathbf{z}\|^2) \quad (\text{C.3})$$

Calculando as derivadas parciais de \mathbf{H} em relação aos vetores \mathbf{z} e \mathbf{y}

$$\frac{\partial \mathbf{H}}{\partial \mathbf{y}} = 2(\mathbf{C}^T \mathbf{Cy} + \mathbf{C}^T \mathbf{Dz}) \quad (\text{C.4})$$

$$\frac{\partial \mathbf{H}}{\partial \mathbf{z}} = 2(\mathbf{D}^T \mathbf{Dz} + \mathbf{D}^T \mathbf{Cy} - \lambda \mathbf{z}) \quad (\text{C.5})$$

Igualando a zero as derivadas parciais encontramos os seguintes resultados:

Na Equação (C.4):

$$\mathbf{C}^T \mathbf{C} \mathbf{y} + \mathbf{C}^T \mathbf{D} \mathbf{z} = \mathbf{0} \quad \rightarrow \quad \mathbf{y} = -(\mathbf{C}^T \mathbf{C})^{-1} \mathbf{C}^T \mathbf{D} \mathbf{z}. \quad (\text{C.6})$$

Na Equação (C.5), e também substituindo \mathbf{y} da Equação (C.6):

$$\begin{aligned} \mathbf{D}^T \mathbf{D} \mathbf{z} + \mathbf{D}^T \mathbf{C} \mathbf{y} - \lambda \mathbf{z} &= \mathbf{0} \\ \mathbf{D}^T \mathbf{D} \mathbf{z} + \mathbf{D}^T \mathbf{C} [-(\mathbf{C}^T \mathbf{C})^{-1} \mathbf{C}^T \mathbf{D} \mathbf{z}] &= \lambda \mathbf{z} \\ (\mathbf{D}^T \mathbf{D} - \mathbf{D}^T \mathbf{C} (\mathbf{C}^T \mathbf{C})^{-1} \mathbf{C}^T \mathbf{D}) \mathbf{z} &= \lambda \mathbf{z} \end{aligned} \quad (\text{C.7})$$

A Equação (C.7) pode ser reescrita na forma:

$$\mathbf{D}^T (\mathbf{I} - \mathbf{C} (\mathbf{C}^T \mathbf{C})^{-1} \mathbf{C}^T) \mathbf{D} \mathbf{z} = \mathbf{E} \mathbf{z} = \lambda \mathbf{z} \quad (\text{C.8})$$

Se substituirmos \mathbf{y} da Equação (C.6) na Equação (C.2), obtemos:

$$\begin{aligned} &\| \mathbf{C} [-(\mathbf{C}^T \mathbf{C})^{-1} \mathbf{C}^T \mathbf{D} \mathbf{z}] + \mathbf{D} \mathbf{z} \|^2 \\ &= \| (\mathbf{I} - \mathbf{C} (\mathbf{C}^T \mathbf{C})^{-1} \mathbf{C}^T) \mathbf{D} \mathbf{z} \|^2 \\ &= \mathbf{z}^T \mathbf{D}^T (\mathbf{I} - \mathbf{C} (\mathbf{C}^T \mathbf{C})^{-1} \mathbf{C}^T)^T (\mathbf{I} - \mathbf{C} (\mathbf{C}^T \mathbf{C})^{-1} \mathbf{C}^T) \mathbf{D} \mathbf{z} \end{aligned}$$

Logo, quando a expressão a cima for mínima, temos $(\mathbf{C} (\mathbf{C}^T \mathbf{C})^{-1} \mathbf{C}^T) \rightarrow \mathbf{0}$, implicando que

$$\rightarrow \mathbf{z}^T \overbrace{\mathbf{D}^T (\mathbf{I} - \mathbf{C} (\mathbf{C}^T \mathbf{C})^{-1} \mathbf{C}^T) \mathbf{D} \mathbf{z}}^{=\lambda \mathbf{z}} = \lambda \overbrace{\| \mathbf{z} \|^2}^{=1} = \lambda$$

Conseqüentemente, o erro mínimo corresponderá ao menor autovalor da matriz \mathbf{E} . Como esse critério requer que esse mínimo seja não-negativo, teremos de provar que os autovalores de \mathbf{E} são todos não-negativos.

Para isso, teremos de provar que \mathbf{E} é uma matriz semidefinida positiva. Podemos observar na Equação (C.8), que caso provemos que $\mathbf{F} = (\mathbf{I} - (\mathbf{C}^T \mathbf{C})^{-1} \mathbf{C}^T)$ é uma matriz semidefinida positiva, então provamos que \mathbf{E} também será.

Então, seja \mathbf{v} um autovetor associado a um autovalor não-nulo de \mathbf{F} .

$$(\mathbf{I} - \mathbf{C} (\mathbf{C}^T \mathbf{C})^{-1} \mathbf{C}^T) \mathbf{v} = \lambda_F \mathbf{v} \quad (\text{C.9})$$

Multiplicando por \mathbf{C}^T a esquerda, obtemos:

$$\lambda_F \mathbf{C}^T \mathbf{v} = \mathbf{C}^T (\mathbf{I} - \mathbf{C} (\mathbf{C}^T \mathbf{C})^{-1} \mathbf{C}^T) \mathbf{v} = [\mathbf{C}^T - \overbrace{(\mathbf{C}^T \mathbf{C}) (\mathbf{C}^T \mathbf{C})^{-1} \mathbf{C}^T}] \mathbf{v} = \mathbf{0}$$

$$\lambda_F \mathbf{C}^T \mathbf{v} = 0 \tag{C.10}$$

Então, temos duas soluções possíveis $\lambda_F = 0$ e $\mathbf{C}^T \mathbf{v} = 0$.

Abrindo a Equação C.9 e considerando o caso da solução $\mathbf{C}^T \mathbf{v} = 0$ para a Equação C.10, obtemos:

$$\begin{aligned} \mathbf{v} - \mathbf{C}(\mathbf{C}^T \mathbf{C})^{-1} \overbrace{\mathbf{C}^T \mathbf{v}}^{=0} &= \lambda_F \mathbf{v} \\ \mathbf{v} &= \lambda_F \mathbf{v}. \end{aligned} \tag{C.11}$$

Como sabemos que o vetor \mathbf{v} é não-nulo por ser um autovetor, com isso a outra solução corresponde a $\lambda_F = 1$. Com isso provamos que os autovalores de \mathbf{F} são não-negativos (0 e +1), implicando em \mathbf{F} ser semidefinida positiva e, por sua vez, \mathbf{E} também.