Universidade Federal
do Rio de Janeiro
Escola Politécnica

Geração de um Banco de Imagens e Vídeos Infravermelhos e Visíveis

Projeto de Graduação apresentado ao Curso de Engenharia Eletrônica e de Computação da Escola Politécnica, Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Engenheiro.

Orientadores:
Eduardo Antônio Barros da Silva
Andreas Ellmauthaler

DEL
Agosto de 2013

Universidade Federal do Rio de Janeiro

Escola Politécnica

Departamento de Engenharia Eletrônica e de Computação

Geração de um Banco de Imagens e Vídeos Infravermelhos e Visíveis

Autor:

Jonathan Nogueira Góis

Banca Examinadora

Orientador:

Eduardo Antônio Barros da Silva, Ph. D.

Orientador:

Andreas Ellmauthaler, M. Sc.

Examinador:

Sérgio Lima Netto , D. Sc.

Examinador:

Carla Liberal Pagliari, Ph. D.

DEL

Agosto de 2013

Gois, J.N.

Geração de um Banco de Imagens e Vídeos Infravermelhos e Visíveis - Rio de Janeiro: UFRJ/Escola Politécnica, 2013.

XI,56 p.: il.: 29,7 cm.

Orientador: Eduardo A. B. da Silva

Projeto de Graduação - UFRJ/Escola Politécnica/Curso de Engenharia Eletrônica e de Computação, 2013.

Referências Bibliográficas p.55-56

1. Banco de Vídeos e Imagens. 2. Fusão de Imagens. 3. Calibração. 4. Registro. I.Da Silva, Eduardo Antônio Barros II. Universidade Federal do Rio de Janeiro, Escola Politécnica, Curso de Engenharia Eletrônica e de Computação. III. Título.

UNIVERSIDADE FEDERAL DO RIO DE JANEIRO
Escola Politécnica - Departamento de Eletrônica e de Computação
Centro de Tecnologia, bloco H, sala H-217, Cidade Universitária
Rio de Janeiro - RJ CEP 21949-900

Este exemplar é de propriedade da Universidade Federal do Rio de Janeiro, que poderá incluí-lo em base de dados, armazená-lo em computador, microfilmá-lo ou arquivá-lo em qualquer outra forma que decidir adotar.

É permitida a menção, reprodução parcial ou integral e a transmissão entre bibliotecas deste trabalho, sem modificação de seu texto, em qualquer meio que esteja ou venha a ser fixado, para pesquisa acadêmica, comentários e citações, desde que sem finalidade comercial e que seja feita a referência bibliográfica completa.

Os conceitos expressos neste trabalho são de responsabilidade do(s) autor(es) e do(s) orientador(es).

Agradecimentos

Agradeço, primeiramente, a Deus por todas as oportunidades a mim concedidas. O que sou é graças a Ele. Agradeço também a minha família por todo apoio durante toda a minha vida e principalmente durante a graduação. Entre todas as dificuldades que passamos, a união da família sempre prevaleceu e prevalece até hoje. Agradeço pela criação e dedicação de meu pai, Ronaldo, ao meus estudos. Só sou o que sou hoje graças a toda sua dedicação e afeto. Muito obrigado, pai. Agradeço a minha mãe, Lourdes, que igualmente se dedicou na minha criação e por isso, muito obrigado.

Agradeço, em especial, a minha namorada Dayane. Pela compreensão, pelo apoio, pelo respeito e por todo o incentivo à este trabalho e a mim. Muito obrigado.

Agradeço aos amigos do LPS que contribuíram significantemente neste trabalho. Toda a experiência de vocês foi essencial no desenvolvimento deste trabalho. Agradeço pela paciência e pelo tempo dedicado à mim.

Agradeço a todos os professores que foram fundamentais na minha formação, em especial ao meu orientador Eduardo. Agradeço por toda paciência e pelo tempo dedicado a minha formação. Agradeço pelo exemplo, tanto como pessoa tanto como profissional.

Agradeço também ao meu coorientador Andreas Ellmauthaler. Em especial pela paciência e pela disponibilidade de sanar todas as minhas dúvidas. Agradeço pelo empenho em todos os dias de gravação e pelos meses de desenvolvimento dos algoritmos que sempre tinham em que melhorar (apesar de sempre estarem muito bons). Assim como o professor Eduardo, este também fica como exemplo.

Muito obrigado à professora Carla Pagliari, que também considero como orientadora. Obrigado pela igual dedicação em todos os dias de gravação, em todo o planejamento do trabalho. Obrigado pelos conselhos e pelas ideias. Muito obrigado por toda sua disponibilidade.

Resumo

Dada a diversidade de sensores de imagens existentes, é possível extrair de uma dada cena diferentes tipos de informação. No entanto, a combinação destas informações em uma única imagem pode ser mais significativa que cada uma separadamente. Técnicas de fusão de imagens são capazes de condensar esta informação e gerar uma imagem com mais informação.

Neste projeto propomos a criação de um banco de imagens e vídeos para esse propósito. Esse banco conterá pares de diferentes vídeos da mesma cena, compostos por uma imagem visível e uma imagem na faixa do infravermelho. Essas imagens são obtidas por sensores diferentes, estes têm diferentes resoluções e muitas outras características distintas, como ponto de vista e região espectral. Tal diferença nestas características exige um processamento nestas imagens de modo que elas tenham equivalência pixel a pixel.

Referimo-nos a esta transformação, como registro. O processamento necessário para implementar o registro das imagens e vídeos é a calibração da câmera, onde nós mensuramos os parâmetros intrínsecos da câmera, e a calibração estéreo, onde nós mensuramos os parâmetros extrínsecos. O registro é realizado após isto.

Neste trabalho, nós realizaremos todo o processamento e o procedimento de registro dos vídeos infravermelho e visível, e geramos o banco registrado.

Palavras-chave: Banco de Videos e Imagens, Fusão de Imagens, Calibração, Registro.

Abstract

Given the diversity of existing types of image sensors, it is possible to extract from a given scene different types of information. However, the combination of this information in a single image can be more meaningful than to have each of them separately. Image fusion techniques are capable of condensing this information and generate an image with more information.

In this project we propose the creation of a bank of images and videos for this purpose. This bank will contain pairs of different videos from the same scene, composed by a visible and an infrared image. These images are obtained from different sensors, that have different resolutions and several other different characteristics, such as point of view and spectral region. Such a difference in characteristics demands the processing of these images in order for them to have a pixel to pixel equivalence.

We refer to this processing as registration. The processing necessary to carry out image and video registration are the camera calibration, where we measure the intrinsic camera parameters, and the stereo calibration, where we measure the extrinsic camera parameters. The registration procedure is performed after this registration.

In this work we perform the entire processing and registration procedure of infrared and visible videos, and generate the registered bank.

Keywords: Video and Image Database, Image Fusion, Calibration, Registration

Sumário

Agradecimentos **v**

Resumo **vi**

Abstract **vii**

Lista de Figuras **x**

Lista de Tabelas **xi**

1 Introdução **1**

1.1 Fusão de Imagens 2

1.2 Objetivo 2

1.3 Organização 3

2 Calibração: Modelagem de uma câmera **5**

2.1 Modelo de Câmera - Pinhole 5

2.1.1 Modelo Básico 6

2.1.2 Deslocamento do ponto principal 7

2.1.3 Modelo Generalizado 8

2.2 Câmera Projetiva 8

2.3 Solução da Calibração da Câmera 9

2.3.1 Estimação Linear 9

2.3.2 Compensação das distorções não-lineares 12

2.4 Conclusão 13

3 Calibração: Detecção dos Pontos **14**

3.1 Algoritmo rápido de detecção dos pontos de calibração 16

3.2 Algoritmo robusto de detecção dos pontos de calibração 21
3.3 Conclusão 22

4 Alinhamento Temporal **23**
4.1 Algoritmo para Alinhamento Temporal 24
4.2 Conclusão 29

5 Registro das Imagens **30**
5.1 Geometria Epipolar 31
5.2 Matriz Fundamental F 32
5.2.1 Propriedades 32
5.3 Matriz Essencial 33
5.4 Homografia no Infinito 34
5.5 Calibração Estéreo 35
5.6 Retificação das Imagens 35
5.7 Conclusão 36

6 Banco de Vídeos **37**
6.1 Aquisição das Imagens 37
6.2 Gravação das Cenas 38
6.3 Descrição do Banco 39
6.4 Alinhamento Temporal 48
6.5 Calibração individual 48
6.6 Calibração Estéreo 49

7 Conclusões **50**
7.1 Trabalhos Futuros 50

Apêndices **50**

A Dados de Sincronismo do Banco de Vídeos **51**

B Dados de Calibração do Banco de Vídeos **53**

Bibliografia **55**

Lista de Figuras

1.1 Fusão em imagens militares. 2
1.2 Fusão em imagens de satélites. 3
1.3 Fusão em imagens médicas. 3

2.1 Geometria da câmera *pinhole* 6
2.2 Plano Focal e ponto principal 7

3.1 Padrão de Calibração usado nas gravações 15
3.2 Diferentes cenários de gravação 15
3.3 Algoritmo de Limiar Adaptativo 16
3.4 Limiar Adaptativo 17
3.5 Canny e Ajuste Elipsoidal 18
3.6 Resultado do algoritmo de detecção dos pontos de calibração 19
3.7 Método rápido de detecção dos pontos 20
3.8 Detecção de Áreas com Limiar Fixo 21
3.9 Método robusto de detecção dos pontos 22

4.1 Evolução quadro a quadro da posição vertical de cada lâmpada 25
4.2 Posição vertical das lâmpadas no tempo 26
4.3 Componente Vertical da velocidade de cada lâmpada 26
4.4 Resultado da similaridade em uma sequência de teste 27
4.5 Alinhamento Temporal 28

5.1 Linha Epipolar 31
5.2 Plano Epipolar 32

6.1 Equipamento de suporte para o par estéreo 38

Lista de Tabelas

6.1 Descrição do conteúdo das cenas gravadas no IME 39
6.2 Descrição do conteúdo das cenas gravadas no Forte São João 40
6.3 Descrição do conteúdo das cenas gravadas no IPqM (Internas) 40
6.4 Descrição do conteúdo das cenas gravadas no IPqM (Externas) 41
6.5 Cenas gravadas no IME 42
6.6 Cenas gravadas no IME (Cont.) 43
6.7 Cenas gravadas no Forte São João 43
6.8 Cenas gravadas no Forte São João (Cont.) 44
6.9 Cenas gravadas no IPqM (Internas) 45
6.10 Cenas gravadas no IPqM (Externas) 46
6.11 Cenas gravadas no IPqM (Externas - Cont.) 47

A.1 Descrição do sincronismo das cenas gravadas no IME 51
A.2 Descrição do sincronismo das cenas gravadas no Forte São João . . . 52
A.3 Descrição do sincronismo das cenas gravadas no IPqM 52

B.1 Descrição da calibração das cenas gravadas no IME 53
B.2 Descrição da calibração das cenas gravadas no Forte São João 54
B.3 Descrição da calibração das cenas gravadas no IPqM 54

Capítulo 1

Introdução

Com a evolução dos sensores de captura de imagens, cresceu a necessidade de combinar informações representadas por diferentes sensores. A fusão de imagens proporciona a integração de informações presentes em diferentes imagens a fim de que sua combinação seja uma imagem com mais informação do que cada uma separadamente.

Há alguns anos, os algoritmos de fusão de imagens se tornarem alvo de pesquisas em diversas áreas da ciência por terem diversas aplicações [6, 11]. O fato de apenas uma imagem apresentar de forma concisa diferentes informações, traz agilidade e facilidade em diversas tarefas.

A facilidade de obter estas informações é sempre crescente, tendo em vista que há inúmeros sensores disponíveis e acessíveis comercialmente. Apesar disto, são pouquíssimas as imagens disponíveis para o teste destes algoritmos, o que motiva este trabalho. As imagens para a realização destes testes são ditas registradas [1] por possuir equivalência pixel a pixel entre elas,ou seja, pixels na mesma posição nas duas imagens representam o mesmo ponto no mundo real, sendo o registro um dos passos para a criação do banco de vídeos.

O trabalho apresenta o método de criação de um banco de vídeos registrados para uso em fusão de imagens, tendo por objetivo a disponibilização deste banco ao público de interesse. Além disto abre-se uma nova possibilidade com este banco. Por se tratar de vídeos, há a informação temporal intrínseca a este e é possível fazer uso desta para aprimorar os atuais algoritmos de fusão.

O banco de vídeos é composto por 31 cenas diferentes gravadas em diversas

situações e em diversos cenários com a mudança do plano da cena, da iluminação e da movimentação presente.

1.1 Fusão de Imagens

Como dito, os algoritmos de fusão de imagens [6] são o alvo para a criação deste banco de vídeos que é interessante nas mais diversas áreas de pesquisa. Estas áreas podem ser sumarizadas em 4 grupos: militar, sensoriamento remoto, medicina e em aplicações industriais.

Na área militar (Figura 1.1), tem-se interesse em rastreamento de alvos, identificação e detecção de armas ocultas, detecção de minas terrestres, planejamento de situações táticas, entre outros. Por sensoriamento remoto (Figura 1.2) entende-se o uso de informações provenientes de satélites, aviões, ou qualquer outra fonte de aquisição para obtenções de informações sobre a superfície da terrestre. Além destes, a área médica (Figura 1.3) tem destaque por oferecer imagens provenientes de várias fontes que possibilitam um melhor diagnóstico ou planejamento clínico. E ainda a indústria faz uso dos métodos de fusão nas áreas da robótica e em controle de qualidade.

(a) Imagem Infravermelha. (b) Imagem Visível. (c) Imagem após a fusão.

Figura 1.1: Exemplo de Fusão em Imagens Militares.

1.2 Objetivo

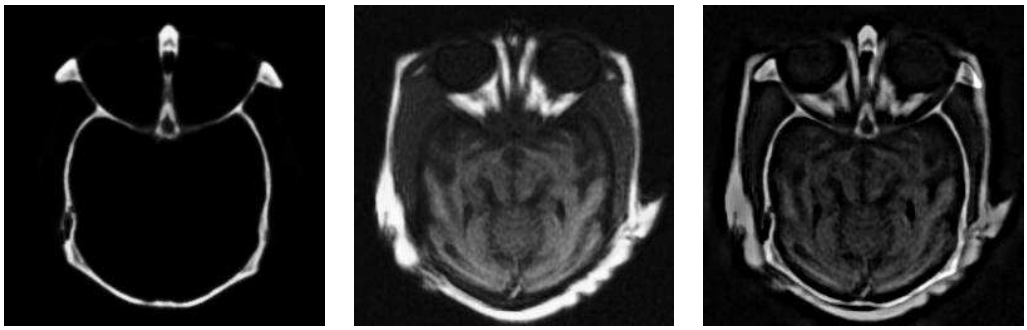
Com a apresentação do vasto uso das técnicas de fusão e tendo em vista a baixa disponibilidade de dados para este uso, o trabalho tem como objetivo a criação deste banco de imagens registradas. O trabalho teve foco em imagens de caráter militar, com o uso de uma câmera infravermelha e uma visível [11].

(a) Imagem Preto e Branco com alta resolução.

(b) Imagem Colorida com baixa resolução.

(c) Imagem após a fusão

Figura 1.2: Exemplo de Fusão em Sensoriamente Remoto.



(a) Tomografia computadorizada.

(b) Ressonância magnética.

(c) Imagem após a fusão.

Figura 1.3: Exemplo de Fusão em Imagens Médicas.

Este trabalho é um complemento à tese de doutoramento de Andreas Ellmauthaler, que disserta sobre técnicas de fusão de imagens usando decomposições em multiresolução, sendo esta a motivação inicial da criação do banco.

Tendo em vista a indisponibilidade de uma câmera que possua os dois sensores internos, foram usadas duas câmeras e após a etapa de gravação foi feito o registro.

1.3 Organização

Os capítulos estão organizados de forma que o Capítulo 1 é esta introdução, o Capítulo 2 trata da calibração de uma única vista, lidando principalmente sobre como calcular as homografias, a matriz de calibração e a matriz de caracterização da câmera. O Capítulo 3 diz respeito ao método de sincronismo temporal desenvolvido, tratando ainda da contextualização dos métodos existentes ao problema. O Capí-

tulo 4 é referente ao registro propriamente dito, tendo ainda uma introdução sobre geometria epipolar e como computar a matriz fundamental da câmera. O Capítulo 5 apresenta as características do banco desenvolvido, contém dos dados do registro e do conteúdo das gravações. Ao final, os apêndices contêm o detalhamento dos dados do banco de vídeos.

Capítulo 2

Calibração: Modelagem de uma câmera

Uma câmera tem como objetivo principal realizar o mapeamento entre pontos do mundo real ($\mathbb{R}^3$) e pontos da imagem ($\mathbb{R}^2$). Esta informação está diretamente ligada ao conhecimento prévio dos parâmetros da câmera. À determinação destes parâmetros damos o nome de calibração da câmera. Com isto e com a informação prévia da distância do plano onde se encontra o objeto imageado é possível realizar o registro das imagens (Capítulo 5) e com isso realizar as correspondências entre as duas imagens.

Para a associação entre os pontos das duas imagens (Infravermelha e Visível) deve-se usar um padrão aparente em ambas as câmeras. Com estes pontos é feito, então, o mapeamento entre $\mathbb{R}^3$ no mundo real e o $\mathbb{R}^2$ na imagem.

Neste Capítulo são expostos os conceitos de modelagem de câmera desde a simplificação da câmera *pinhole* até os conceitos referentes à câmera projetiva com a inserção das distorções da lente. É abordada ainda a maneira de, dadas as homografias, encontrar a matriz de calibração e com isso o modelo da câmera.

2.1 Modelo de Câmera - Pinhole

Antes de realizar a equivalência entre os pixels de duas imagens distintas, faz-se necessário encontrar a função de projeção, ou seja, a função que realiza a correspondência entre a imagem 2D resultante e a cena 3D. Porém, antes disso, são necessárias

a escolha e a modelagem das câmeras.

2.1.1 Modelo Básico

A modelagem mais simples é da câmera *pinhole*. Denotando um ponto no espaço $\mathbb{R}^3$ como $\mathbf{X} = \{X_m, Y_m, Z_m\}$, definido em relação aos eixos de coordenadas $\{C, X, Y\}$ e o ponto da imagem $\mathbf{x} = \{x_i, y_i\}$ definido nos eixos de coordenadas $\{P, x, y\}$ (Figura 2.1), procura-se a relação H ($\mathbb{R}^3 \mapsto \mathbb{R}^2$) tal que:

$$\begin{bmatrix} X_m \\ Y_m \\ Z_m \end{bmatrix} \xrightarrow{H} \begin{bmatrix} x_i \\ y_i \end{bmatrix} \tag{2.1}$$

Desta transformação, serão explicitados dois tipos de parâmetros: os parâmetros extrínsecos, tais como rotação e translação do eixo $\{P, x, y\}$ em relação a $\{C, X, Y, Z\}$, e os parâmetros intrínsecos como escalamento, centro óptico, entre outros.

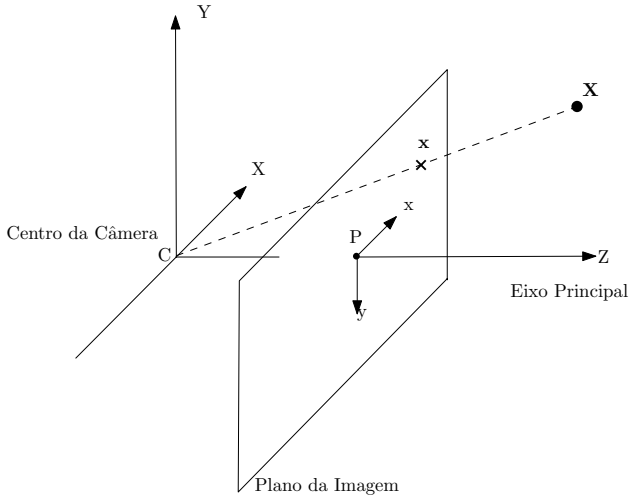


Figura 2.1: Geometria da câmera *pinhole*

Na Figura 2.1, para a modelagem *pinhole*, o ponto C é dito *Centro da Câmera* e o ponto P como *Ponto Principal* que é o ponto de intersecção entre o eixo principal e o plano da imagem (ou plano focal). A distância entre o centro da câmera e o plano focal é igual a f (distância focal) (Figura 2.2).

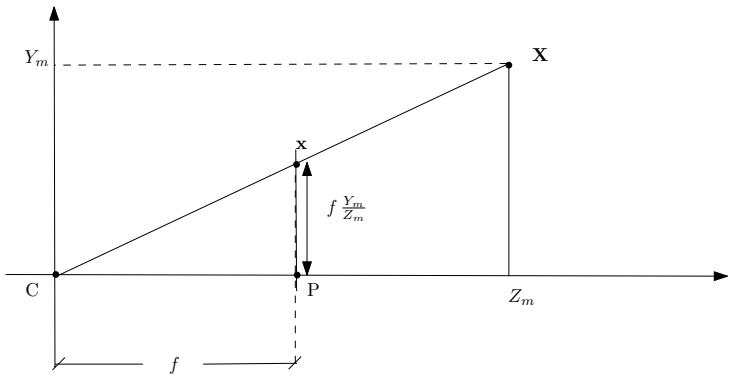


Figura 2.2: Plano Focal e ponto principal

Pode-se escrever facilmente, através da análise da Figura 2.2, que a matriz de projeção do ponto central da câmera, em coordenadas homogêneas, é:

$$\begin{bmatrix} X_m \\ Y_m \\ Z_m \\ 1 \end{bmatrix} \xrightarrow{P} \begin{bmatrix} x_i \\ y_i \\ Z \end{bmatrix} = \begin{bmatrix} f & & & 0 \\ & f & & 0 \\ & & 1 & 0 \end{bmatrix} \begin{bmatrix} X_m \\ Y_m \\ Z_m \\ 1 \end{bmatrix} \tag{2.2}$$

2.1.2 Deslocamento do ponto principal

Considerando ainda que o ponto central P da câmera não está localizado sobre o eixo Z e é expresso por $P = \{p_x, p_y\}$ então a transformação é dada por:

$$\begin{bmatrix} X_m \\ Y_m \\ Z_m \\ 1 \end{bmatrix} \xrightarrow{P} \begin{bmatrix} x_i + p_x \\ y_i + p_y \\ 1 \end{bmatrix} = \begin{bmatrix} f & & p_x & 0 \\ & f & p_y & 0 \\ & & 1 & 0 \end{bmatrix} \begin{bmatrix} X_m \\ Y_m \\ Z_m \\ 1 \end{bmatrix} \tag{2.3}$$

De forma simplificada:

$$x = K.\begin{bmatrix} I & | & 0 \end{bmatrix} X \tag{2.4}$$

onde a matriz K é dada por:

$$K = \begin{bmatrix} f & & p_x \\ & f & p_y \\ & & 1 \end{bmatrix} \tag{2.5}$$

2.1.3 Modelo Generalizado

Na equação (2.4) é representado o modelo simplificado da câmera *pinhole*. A matriz K contém os parâmetros intrínsecos da câmera e a segunda parte apresenta os parâmetros extrínsecos.

Para este modelo ser obtido, foi considerado que não houve translação nem rotação dos eixos que definem as coordenadas do mundo real e do plano da imagem. Para considerar estes efeitos sobre os eixos coordenados deve-se então generalizar a equação (2.4) como:

$$x = K.\begin{bmatrix} R & | & t \end{bmatrix} X \tag{2.6}$$

onde R é a matriz de rotação e t a translação.

2.2 Câmera Projetiva

Além de considerar a informação relativa de translação e rotação entre os eixos, pode-se adicionar a informação de cisalhamento entre os dois eixos. Pode-se ainda considerar que o pixel da imagem não é quadrado e com isso as distâncias focais nos eixos x e y não são iguais. Então, pode-se reescrever (2.5), que define os parâmetros intrínsecos da câmera como:

$$K = \begin{bmatrix} \alpha_x & s & p_x \\ 0 & \alpha_y & p_y \\ 0 & 0 & 1 \end{bmatrix} \tag{2.7}$$

onde α_x e α_y dizem respeito ao formato do pixel e s à informação de cisalhamento.

Então, explicitando os termos da equação (2.6):

$$x = \underbrace{\begin{bmatrix} \alpha_x & s & p_x \\ 0 & \alpha_y & p_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} R & | & t \end{bmatrix}}_{P} X \tag{2.8}$$

2.3 Solução da Calibração da Câmera

Com um modelo estabelecido e, com isso, conhecendo os parâmetros a serem estimados, podemos realizar a calibração da câmera. O processo de calibração tem início com os pontos de calibração sendo previamente definidos como será visto no Capítulo 3. Após isto, define-se a transformação H que mapeia os pontos do padrão nos pontos da imagem obtida. A esta transformação dá-se o nome de homografia. Em outras palavras, dado um ponto da imagem x_i deseja-se encontrar a homografia H tal que $x'_i = H.x_i$, onde x'_i é um ponto do padrão no mundo real em um dado plano (qualquer plano desde que seja o mesmo para todas as homografias).

Entretanto, para cada conjunto de pontos de calibração existe (ou não) uma homografia distinta para realizar esta operação. Isso decorre do fato que os pontos são estimados e as imagens são contaminadas por ruído e distorções provenientes de diversos fatores. Uma maneira robusta de estimar a homografia H de uma câmera é então utilizar uma estimação linear e após isso usar as informações de distorção para aprimorar o resultado.

O cálculo das homografias é o primeiro passo para a solução da calibração. Como será visto na próxima seção, são necessárias algumas homografias para a estimar a matriz de calibração. Para haver diversidade de informações nessa estimativa e, com isso, ser possível realizar um refinamento da matriz de calibração, deve-se haver uma variedade de homografias. Estas são provenientes de diferentes vistas, ou seja, quadros onde o padrão de calibração está presente em posições distintas. E por isso a movimentação do padrão durante a calibração é necessária.

2.3.1 Estimação Linear

Como dito, a primeira estimativa para a matriz de calibração é a linear. A solução obtida por esta estimação da matriz H desconsidera as distorções radial e tangencial

das lentes [3].

Podemos reescrever [1] a igualdade que relaciona o ponto da imagem e o ponto do mundo real como:

$$Hx_i = \begin{bmatrix} h^{1T}x_i \\ h^{2T}x_i \\ h^{3T}x_i \end{bmatrix} = x'_i. \tag{2.9}$$

Sendo $x'_i = [x'_i, y'_i, w'_i]$ e fazendo o produto vetorial, então tem-se que:

$$x'_i \times Hx_i = \begin{bmatrix} y'_i h^{3T}x_i - w'_i h^{2T}x_i \\ w'_i h^{1T}x_i - x'_i h^{3T}x_i \\ x'_i h^{2T}x_i - y'_i h^{1T}x_i \end{bmatrix} = 0. \tag{2.10}$$

Como $h^{jT}x_i = x_i^T h^j$, pode-se reescrever:

$$\begin{bmatrix} y'_i x_i^T h^3 - w'_i x_i^T h^2 \\ w'_i x_i^T h^1 - x'_i x_i^T h^3 \\ x'_i x_i^T h^2 - y'_i x_i^T h^1 \end{bmatrix} = \begin{bmatrix} 0^T & -w'_i x_i^T & y'_i x_i^T \\ w'_i x_i^T & 0^T & -x'_i x_i^T \\ -y'_i x_i^T & x'_i x_i^T & 0^T \end{bmatrix} \begin{bmatrix} h^1 \\ h^2 \\ h^3 \end{bmatrix} = 0. \tag{2.11}$$

Tendo em vista que a terceira linha desta matriz é linearmente dependente (3ª linha é igual x'_i vezes a 1ª somada com y'_i vezes a segunda), então pode-se simplificar a equação anterior como:

$$\underbrace{\begin{bmatrix} 0^T & -w'_i x_i^T & y'_i x_i^T \\ w'_i x_i^T & 0^T & -x'_i x_i^T \end{bmatrix}}_{A_i} \begin{bmatrix} h^1 \\ h^2 \\ h^3 \end{bmatrix} = 0 \tag{2.12}$$

A solução então depende apenas da resolução da matriz A_i para cada vista que pode ser obtida através da SVD (*Singular Value Decomposition*)[1].

Após todas as homografias serem estimadas através deste processo conhecido como DLT (*Direct Linear Transformation*), foi utilizado o processo iterativo de Levenberg-Marquardt [1] para o refinamento da solução de cada homografia. Após isso, tem-se que estimar a matriz de calibração K. Lembrando que a matriz K, é definida como:

$$K = \begin{bmatrix} \alpha_x & s & p_x \\ 0 & \alpha_y & p_y \\ 0 & 0 & 1 \end{bmatrix} \tag{2.13}$$

Podemos definir [2] uma matriz B, tal que:

$$B = K^{-T}K^{-1} = \begin{bmatrix} B_{11} & B_{12} & B_{13} \\ B_{21} & B_{22} & B_{23} \\ B_{31} & B_{32} & B_{33} \end{bmatrix} \tag{2.14}$$

$$B = \begin{bmatrix} \frac{1}{\alpha_x} & -\frac{s}{\alpha_x^2\alpha_y} & \frac{p_y s - p_x\alpha_y}{\alpha_x^2\alpha_y} \\ -\frac{s}{\alpha_x^2\alpha_y} & \frac{s^2}{\alpha_x^2\alpha_y^2} + \frac{1}{\alpha_y^2} & \frac{-s(p_y s - p_x\alpha_y)}{\alpha_x^2\alpha_y^2} - \frac{p_y}{\alpha_y^2} \\ \frac{-(p_y s - p_x\alpha_y)}{\alpha_x^2\alpha_y} & -\frac{s(p_y s - p_x\alpha_y)}{\alpha_x^2\alpha_y^2} - \frac{p_y}{\alpha_y^2} & \frac{(p_y s - p_x\alpha_y)^2}{\alpha_x^2\alpha_y^2} + \frac{p_y^2}{\alpha_y} + 1 \end{bmatrix} \tag{2.15}$$

Como a matriz B é simétrica, podemos definir o vetor:

$$b = [B_{11}, B_{12}, B_{22}, B_{13}, B_{23}, B_{33}]^T \tag{2.16}$$

Tendo a *i-ésima* coluna da homografia como $h_i = [h_{i1}, h_{i2}, h_{i3}]^T$, pode-se escrever:

$$h_i^T B h_j = \underbrace{\begin{bmatrix} h_{i1}h_{j1} \\ h_{i1}h_{j2} + h_{i2}h_{j1} \\ h_{i2}h_{j2} \\ h_{i3}h_{j1} + h_{i1}h_{j3} \\ h_{i3}h_{j2} + h_{i2}h_{j3} \\ h_{i3}h_{j3} \end{bmatrix}^T}_{v_{ij}} b. \tag{2.17}$$

Sabendo que as homografias podem ser reescritas como :

$$[h_{i1} \quad h_{i2} \quad h_{i3}] = \lambda K[r_{i1} \quad r_{i2} \quad | \; t_i] \tag{2.18}$$

e que r_{i1} e r_{i2} são vetores ortonormais por definição e que por isso $h_{i1}^T K^{-T}K^{-1}h_{i2} = 0$ e $h_{i1}^T K^{-T}K^{-1}h_{i1} = h_{i2}^T K^{-T}K^{-1}h_{i2}$, podemos reduzir a expressão 2.17 como :

$$\underbrace{\begin{bmatrix} v_{12}^T \\ (v_{11} - v_{22})^T \end{bmatrix}}_{V} b = 0 \tag{2.19}$$

E com esse resultado, vemos que, dada uma homografia, temos duas equações linearmente independentes para a solução de B. Generalizando a expressão para n homografias diferentes, temos que a matriz V tem dimensão $2n \times 6$. Com esta dimensão, vê-se que são necessários no mínimo 3 homografias calculadas para encontrar uma solução para o sistema. E com esta é possível calcular a matriz de calibração K, cujo os parâmetros podem ser expressos por:

$$\alpha_x = \sqrt{\lambda B_{11}} \tag{2.20}$$

$$\alpha_y = \sqrt{\lambda B_{11}/(B_{11}B_{22} - B_{12}^2)} \tag{2.21}$$

$$p_x = s p_y/\alpha_y - B_{13}\alpha_x^2/\lambda \tag{2.22}$$

$$p_y = (B_{12}B_{13} - B_{11}B_{23})/(B_{11}B_{22} - B_{12}^2) \tag{2.23}$$

$$s = -\frac{B_{12}\alpha_x^2\alpha_y}{\lambda} \tag{2.24}$$

$$\lambda = B_{33} - [B_{13}^2 + p_y(B_{12}B_{13} - B_{11}B_{23})]/B_{11} \tag{2.25}$$

2.3.2 Compensação das distorções não-lineares

Após a primeira estimativa é necessário incluir as distorções radial e tangencial provenientes da lente. Estas distorções podem ser modeladas [1, 12] por:

$$F_D(\mathbf{x}_d, k, p) = \begin{bmatrix} \mathbf{x}_d(k_1 r^2 + k_2 r^4 + ...) + (2p_1\mathbf{x}_d\mathbf{y}_d + p_2(r_2 + 2\mathbf{x}_d^2)) \\ \mathbf{y}_d(k_1 r^2 + k_2 r^4 + ...) + (p_1(r^2 + 2\mathbf{y}_d^2) + 2p_2\mathbf{x}_d\mathbf{y}_d) \end{bmatrix} \tag{2.26}$$

onde, $\mathbf{x}_d = [x_d, y_d]^T$ são as coordenadas normalizadas e distorcidas, $k = k_1, k_2, ...$ os coeficientes da distorção radial e $p = p_1, p_2, ...$ da distorção tangencial e $r = \sqrt{x_d^2 + y_d^2}$. Este modelo de distorção das lentes é utilizado na nova reprojeção, ou seja, a distorção é adicionada ao modelo da câmera, com a finalidade de obter a

minimização da função custo dada por :

$$\sum_i \sum_j ||(x_{ij}^{sub} - x^{proj}(\mathbf{K}, k, p, \mathbf{R}_i, t_i, \mathbf{X}_j))||^2 \tag{2.27}$$

onde x_{ij}^{sub} é a posição do sub-pixel do j-ésimo ponto de calibração do i-ésimo quadro de calibração e $x^{proj}(\mathbf{K}, k, p, \mathbf{R}_i, t_i, \mathbf{X}_j))$ é a projeção do ponto do mundo real X_j.

Este modelo de distorções considera modelos polinomiais para a distorção. No trabalho utilizamos um modelo de distorções quadrático tanto para a distorção radial como para a distorção tangencial de ambas as câmeras. Depois de estimados os parâmetros das distorções, podemos adicioná-los ao modelo e realizar novamente a estimação linear até que a solução obtida para a calibração seja satisfatória (de acordo com a métrica escolhida).

2.4 Conclusão

Neste capítulo mostramos que, dado um modelo de câmera, é possível conhecer seus parâmetros através da calibração da câmera. Fazemos isso calculando homografias que correspondem ao mapeamento de um padrão conhecido do mundo real para a imagem criada. Tendo todas as homografias calculadas, podemos encontrar uma solução para a matriz de calibração sendo esta resultado de dois passos: a estimação linear e a compensação das distorções causada pelas lentes. Segue que, após a estimação linear, temos uma solução para a calibração dada. Porém, podemos refiná-la através da compensação das distorções. Retirando as distorções das imagens e recalculando as homografias, podemos refinar a matriz de calibração.

Após a definição da matriz de calibração, podemos realizar a calibração estéreo e por fim o registro. Antes de ser possível realizar os passos descritos, é necessário um método para a detecção dos pontos, tendo em vista o desafio da equivalência entre as imagens referentes ao espectro visível e ao espectro infravermelho. O próximo capítulo (Capítulo 3) traz os métodos utilizados para a realização desta tarefa.

Capítulo 3

Calibração: Detecção dos Pontos

Para computar os parâmetros extrínsecos, é necessário que a figura de comparação entre as duas imagens sejam igualmente notáveis nas duas câmeras. Normalmente esta comparação é feita usando um padrão quadriculado tendo como pontos de calibração cada interseção. Em outros casos, ainda é possível utilizar um padrão que usa pontos de controles circulares [4]. São inúmeros os métodos disponíveis para a calibração de câmeras de espectro visível, porém ao calibrar câmeras referentes ao espectro visível e infravermelho, os padrões são mais limitados.

Geralmente, para a calibração de câmeras infravermelhas são utilizados os mesmos padrões quadriculados impressos e aquecidos através de uma lâmpada disposta atrás do padrão. Esta solução é eficaz para a câmera de espectro visível, porém apresenta pouca acurácia dos pontos no infravermelho e apresenta problemas, como o pouco tempo de uso possível. Entretanto, este padrão pode ser adaptado utilizando a técnica sugerida por Vidas *et. al* [8], usando o mesmo padrão quadriculado impresso e com os quadrados recortados, tornando assim o padrão vazado. Com essa técnica é possível aumentar a acurácia dos pontos, não necessitando mais da lâmpada de aquecimento.

Entre outras técnicas, o uso de um padrão com pontos de controles circulares traz vantagens na calibração [4]. Para a utilização desta técnica no par estéreo IR-Visível, foram utilizadas pequenas lâmpadas dispostas em um padrão 9x9 com um espaçamento fixo conhecido.

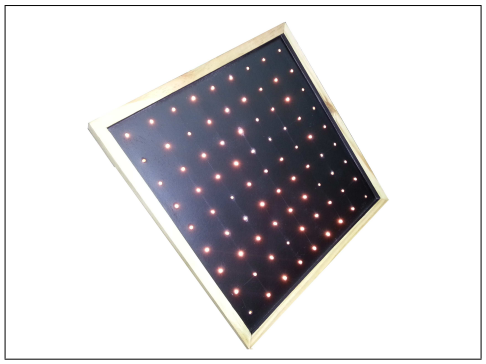


Figura 3.1: Padrão de Calibração usado nas gravações

Como os ambientes de gravação são adversos (Figura 3.2), há a necessidade de realizar diferentes processamentos de acordo com a situação (principalmente no caso das imagens no espectro visível).

(a)

(b)

Figura 3.2: Diferentes cenários de gravação

A seguir, estão descritos dois métodos propostos para a extração dos pontos de calibração, tendo como finalidade o uso em todas os ambientes e em todas as situações testadas para as imagens das duas câmeras.

3.1 Algoritmo rápido de detecção dos pontos de calibração

O primeiro método proposto é o método rápido. Este método é utilizado em cenas com menos iluminação e em cenas com menos áreas que possam ser confundidas como possíveis micro lâmpadas.

Essa consideração inicial de caso de uso é imposta principalmente pela primeira técnica utilizada neste método que é a binarização através de um limiar adaptativo. A binarização tem por objetivo tornar 1 as áreas onde, possivelmente, as lâmpadas estão localizadas e 0 nas demais áreas. O algoritmo para a definição deste limiar para binarização segue abaixo:

Passo 1: Inverter a imagem original e ajustar em 256 níveis de cinza.

Passo 2: Calcular a intensidade média μ da imagem e arredondar para o inteiro mais próximo [0,255]

Passo 3: Determinar o histograma da imagem usando 256 bins para a imagem anterior.

Passo 4: Extrair os μ primeiros bins e descartar os restantes.

Passo 5: Calcular a média μ' de todos os pixels do histograma modificado. Este será o limiar inicial T

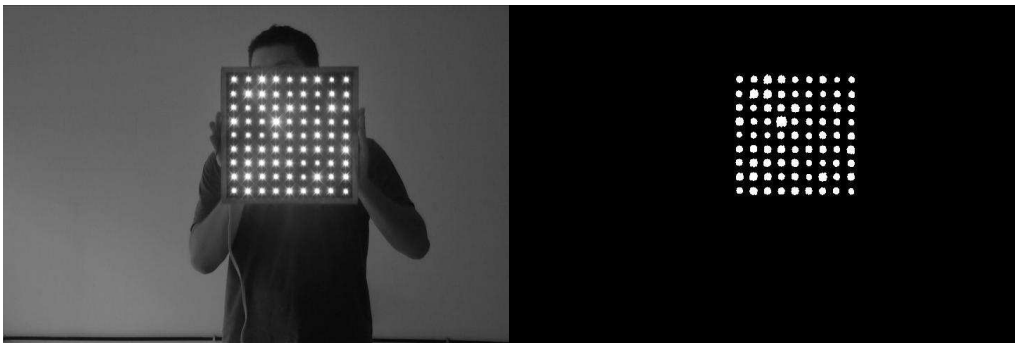
Passo 6: Calcular as médias $\mu 1$ e $\mu 2$ dos pixels acima e abaixo do limiar T.

Passo 7: O limiar T deve ser atualizado com a média entre $\mu 1$ e $\mu 2$

Passo 8: Repetir desde o passo 6 até que o limiar T não varie em relação à iteração anterior dentro de um valor definido anteriormente.

Figura 3.3: Algoritmo de Limiar Adaptativo

O resultado do algoritmo pode ser visto na Figura 3.4



(a) Imagem Original (b) Imagem binária

Figura 3.4: Limiar Adaptativo

Encontrado o limiar, notamos que as áreas onde as lâmpadas estão localizadas são disformes e incoerentes com o esperado da irradiação esférica de cada lâmpada. Isso se dá por diversos motivos mas principalmente pelas reflexões da lâmpada no próprio padrão e pelo efeito do movimento nas imagens. Para a retirar estes efeitos são utilizadas técnicas morfológicas que tem por objetivo reduzi-los na imagem binária.

Após essas técnicas de pré-processamento ditas anteriormente, as áreas de interesse ainda não tem o padrão esperado. Com o conhecimento prévio do formato da lâmpada, é esperado que a saída da imagem binarizada seja idealmente um círculo ou ainda, de forma mais geral, uma elipse.

E para isso é feito um ajuste elipsoidal para ser utilizado na determinação de um ponto representativo da posição do centro da micro lâmpada.

Para o ajuste da elipse, primeiramente é aplicado um filtro de Canny [7] às imagens. O conjunto de pontos formados pelas bordas de cada micro lâmpada é agrupado e é estimada a elipse representativa da micro lâmpada [4].

Com a elipse estimada, uma nova imagem binária é formada. Os pontos interiores à elipse são representados por 1 e as demais áreas, por 0.

(a) Imagem após Detector de Arestas (b) Imagem após o ajuste elipsoidal

Figura 3.5: Imagem após o filtro Canny e Imagem após o ajuste elipsoidal

Após essa imagem encontrada (Figura 3.5), temos que estimar a melhor posição para o ponto de calibração. Para isso é utilizado que o melhor ponto para a elipse representativa à lâmpada é o seu centro de gravidade.

Para o cálculo do centro de gravidade são utilizadas a imagem original em nível de cinza e a imagem binarizada.

O método usa os valores dos pixels originais e da imagem binária. Realizando o produto pixel a pixel destas imagens, é feita a separação morfológica de cada área de interesse e com esta área é calculada o centro de gravidade tendo como peso o valor do pixel original da imagem.

Além da informação da forma da microlâmpada (que foi utilizada até agora para a determinação do centro de gravidade), outra informação útil que pode se retirar do padrão de calibração é a disposição das lâmpadas.

Tendo o ponto $X'_i = (x'_i, y'_i)$ como um ponto aproximado e o ponto $X_i = (x_i, y_i)$ um ponto real, podemos considerar que existe uma transformação H que leva todos os pontos aproximados para os pontos reais que minimiza a função custo dada por :

$$p(H) = \sum_{i=1}^{N} (||X_i - H.X'_i||^2) \tag{3.1}$$

Então, achando a transformação H que minimiza a função custo que representa o somatório da distância entre os pontos reais e os pontos transformados,

tem-se uma melhor aproximação dos pontos que serão usados no processo de calibração.

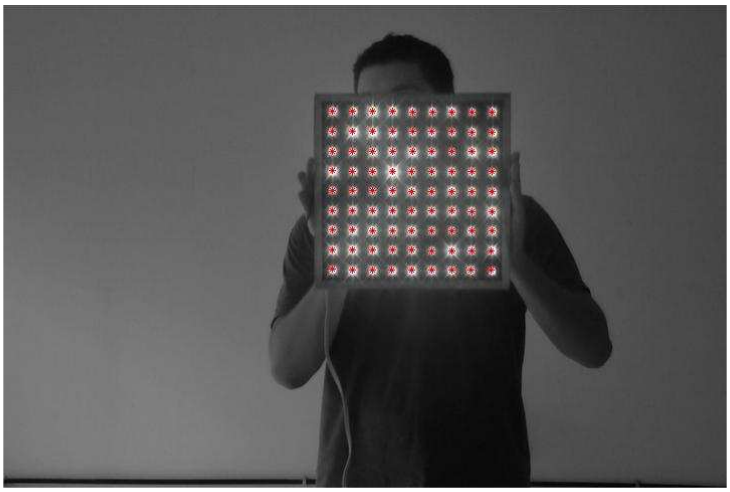


Figura 3.6: Resultado do algoritmo de detecção dos pontos de calibração

À parte das técnicas mostradas até o momento, deve-se levar em conta a retirada da marca d'água das imagens infravermelhas. A técnica da clusterização *k-means* [15] é usada para retirar estas áreas (*a priori* consideradas de interesse) que estão muito afastadas das áreas reais. Essa técnica é usada ainda para retirar possíveis pontos mais quentes da cena que se confundem às lâmpadas.

O *k-means* é uma técnica usada para agrupar as células (onde cada célula contém as áreas de interesse previamente definidas) e classificá-las de acordo com a distância média da célula ao centro do grupo. Mais detalhes sobre este método podem ser encontrados na literatura em [15].

Sumarizando o método rápido, temos o diagrama de blocos a seguir:

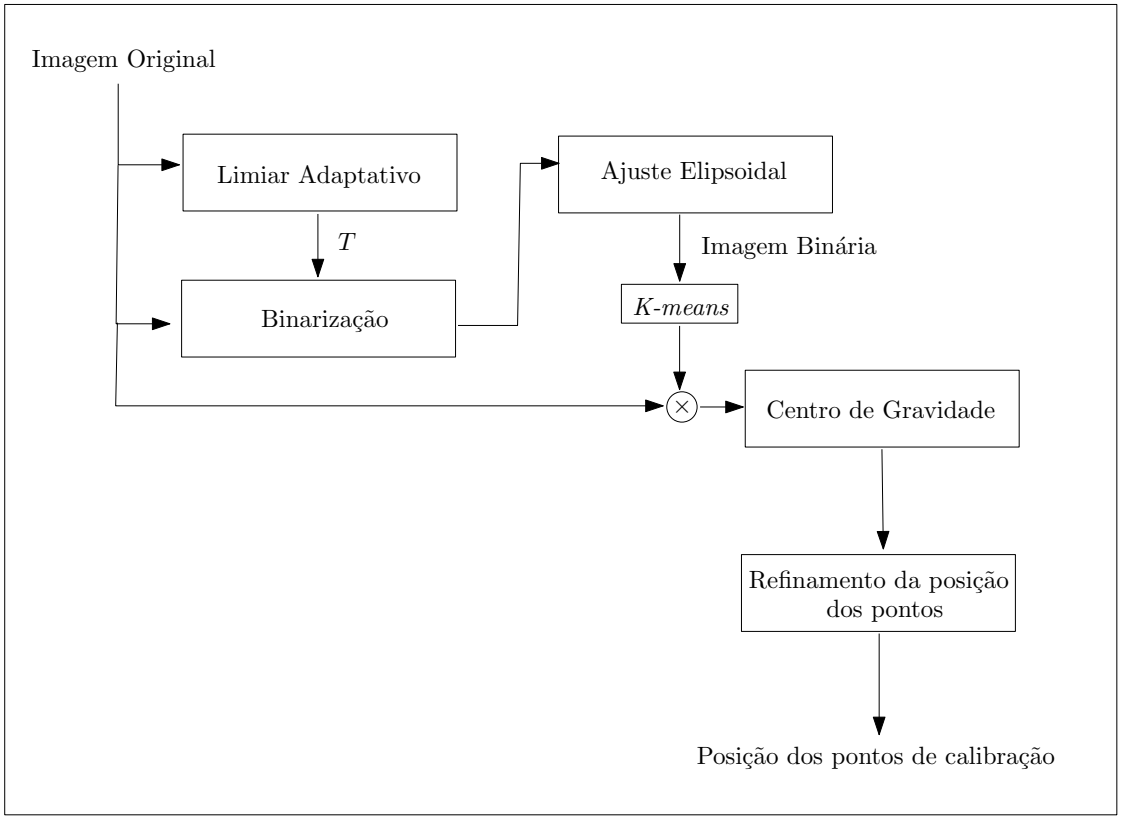


Figura 3.7: Diagrama de Blocos do método rápido para a detecção dos pontos

Em gravações externas, onde há muita informação no fundo da imagem ou ainda existem muitas regiões claras (principalmente no vídeo visível), o método pode não ser muito eficiente principalmente pelo uso do limiar adaptativo na binarização. Para a resolução desse problema, podemos utilizar o algoritmo apenas na região do padrão previamente segmentado da imagem original [16].

Em outros casos, onde esta solução para a detecção não é suficiente, podemos utilizar o método a seguir.

3.2 Algoritmo robusto de detecção dos pontos de calibração

Esta solução começa com a possibilidade de muitos falsos positivos, como por exemplo em ambientes muito iluminados. A principal diferença entre o método rápido e o método robusto de detecção dos pontos de calibração é a binarização das imagens.

Em alguns casos, o uso do limiar adaptativo pode trazer muitos falsos positivos e/ou muito falso negativos. A solução encontrada nesses casos é o uso de um limiar fixo combinado ao processo de iteração na procura dos pontos, como é descrito no algoritmo abaixo:

Passo 1: Definir um limiar inicial T e um passo fixo $\Delta\ t$.

Passo 2: Realizar a comparação da imagem com o limiar e computar a imagem binária resultante.

Passo 3: Contar o número de áreas dos pontos de interesse e comparar ao número esperado.

Passo 4: Caso o número exceda o esperado, aumentar o limiar T e voltar ao **Passo 2**. Caso ocorra o contrário, voltar ao limiar anterior e retirar os falsos positivos com outro método.

Figura 3.8: Detecção de Áreas com Limiar Fixo

O algoritmo de limiar fixo mostrado anteriormente é aplicado e é utilizado um número de áreas maior do que as 81 esperadas (tipicamente 83). Com estas áreas são aplicadas as mesmas técnicas morfológicas e o *K-means* como no algoritmo rápido. Destes 83 pontos são calculados os centros de gravidade, escolhidos conjuntos de 81 e aplicando o refinamento como do método anterior, é escolhido o melhor conjunto de 81 pontos. Este conjunto é dado pelo menor erro médio quadrático dada na equação (3.1).

O método robusto pode ser descrito pelo diagrama em blocos a seguir:

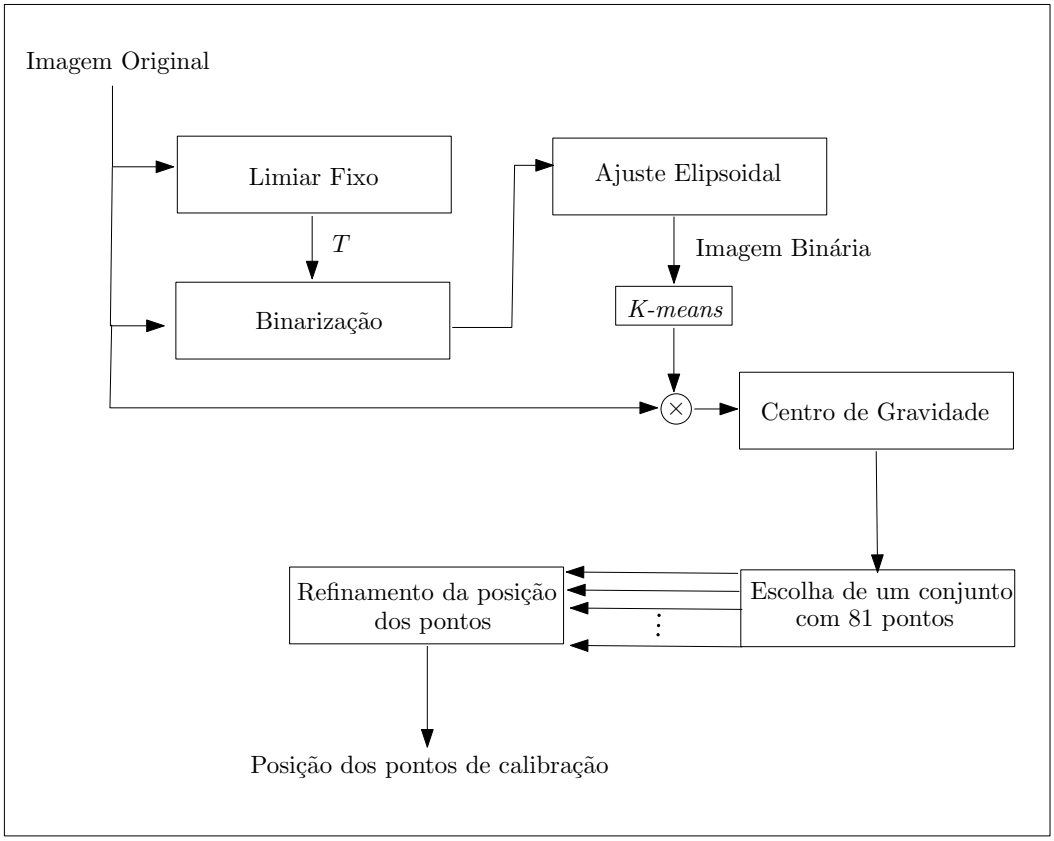


Figura 3.9: Diagrama de Blocos do método robusto para a detecção dos pontos

3.3 Conclusão

Tendo em vista as técnicas apresentadas neste capítulo, notamos que é possível realizar a detecção dos pontos de maneiras distintas visando tanto a complexidade computacional quanto a eficácia do método (Seções 3.1 e 3.2). Estas soluções podem ser utilizadas ainda de forma iterativa de modo que possamos otimizar a escolha da matriz de calibração como visto na referência [12].

Capítulo 4

Alinhamento Temporal

O alinhamento temporal entre sequências de vídeos [13] é importante em inúmeras aplicações onde um par de câmeras estéreo é utilizado, como por exemplo gravações de vídeos 3D.

Em alguns casos é possível realizar o casamento usando métodos comparativos, que são aplicáveis nos casos de alta correlação entre as informações presentes em ambos os vídeos. Em outros casos é possível realizar uma adequação do hardware de captura para realizar o acionamento de ambos os dispositivos de captura, tornando, com isso, o sincronismo desnecessário.

Entretanto, esta solução trivial não é possível quando os acionamentos das câmeras são distintos ou ainda quando a taxa de captura entre elas é diferente.

A relevância de um bom método de sincronismo se dá pois as cenas a serem registradas necessitam de correspondência pixel a pixel, ou seja um pixel da imagem infravermelha corresponda ao mesmo pixel da imagem visível. As fidelidades temporal e espacial são fundamentais para os testes dos algoritmos de fusão de imagens [6].

O desenvolvimento do método de sincronismo fez-se necessário pois as informações em cada quadro são diferentes, ou ainda não há (ou há poucas) correspondências entre as imagens provenientes de cada câmera. O padrão de calibração desenvolvido ajuda a aumentar as características em comum, facilitando assim obter pontos comparativos nas duas sequências de vídeos.

Além disso, o método pode ser aperfeiçoado com o uso de artifícios na gravação da parte de calibração de cada cena. Para realizar a calibração de cada câmera

é necessária a diversificação de posições dos pontos do padrão de modo que estes sejam não-coplanares (o padrão deve estar em diferentes planos a cada vista), permitindo com isso a estimação mais acurada das matrizes de rotação e translação (Capítulo 2). Por este lado, realizar translações, tanto verticais quanto horizontais, ou rotações no mesmo plano não acrescentam dados para a calibração. Esta parte da cena é então destinada ao alinhamento temporal.

Tendo em vista que há disparidade entre o par estéreo utilizado (acoplamento horizontal entre as câmeras), o proposto foi fazer uso do movimento vertical. Na gravação das cenas foi utilizada a repetição dos movimentos de sincronismo (movimentação vertical do padrão) e de calibração (mudança do plano do padrão de calibração) ao menos duas vezes. Não há necessidade de se utilizarem tantos quadros para realizar o sincronismo, porém isso aumenta ainda mais a robustez do método trazendo alta correlação temporal dos vídeos sincronizados.

Neste capítulo 4 será apresentada a forma de determinar o deslocamento temporal existente entre os vídeos e com isso realizar o casamento temporal das sequências de vídeo.

4.1 Algoritmo para Alinhamento Temporal

Dada a condição de acoplamento entre as câmeras, dispostas lado a lado (como dito anteriormente), devemos encontrar uma relação entre o movimento vertical dos pontos de interesse dado que há uma disparidade imposta nessa montagem impossibilitando uma fácil comparação do movimento horizontal. A posição y de cada ponto de interesse é mostrada na figura abaixo (Figura 4.1).

Nesta figura é possível perceber que há um deslocamento entre a sequência infravermelha e a visível. Disso temos que, dada as duas sequências de vídeo S_V e S_I, estamos interessados no deslocamento $\Delta\hat{t}$ que maximiza a função de similaridade s entre as duas sequências.

$$\Delta\hat{t} = \arg\max_{\Delta t}(\ s(S_V, S_I, \Delta t)\ ) \tag{4.1}$$

O método proposto [13] faz uso da velocidade vertical na função de similaridade. Dados os pontos extraídos (como visto no Capítulo 3), para calcular essa

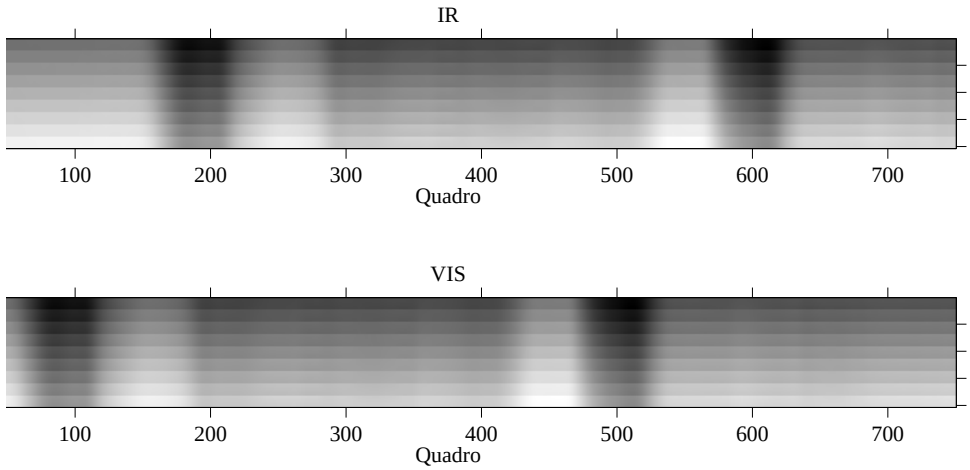


Figura 4.1: Evolução quadro a quadro da posição vertical de cada lâmpada. Os valores mais escuros representam um menor valor da componente y.

componente temos que subtrair a posição y de dois quadros seguintes, podendo assim calcular uma matriz com as velocidades de cada ponto. Sejam as matrizes que representam as coordenadas y de cada quadro n dadas por :

$$Y_{IR}(n) = \begin{bmatrix} y_{11}^{IR}(n) & y_{12}^{IR}(n) & ... & y_{19}^{IR}(n) \\ y_{21}^{IR}(n) & y_{22}^{IR}(n) & ... & y_{29}^{IR}(n) \\ . & . & . & . \\ . & . & . & . \\ . & . & . & . \\ y_{91}^{IR}(n) & y_{92}^{IR}(n) & ... & y_{99}^{IR}(n) \end{bmatrix}, Y_{VIS}(n) = \begin{bmatrix} y_{11}^{VIS}(n) & y_{12}^{VIS}(n) & ... & y_{19}^{VIS}(n) \\ y_{21}^{VIS}(n) & y_{22}^{VIS}(n) & ... & y_{29}^{VIS}(n) \\ . & . & . & . \\ . & . & . & . \\ . & . & . & . \\ y_{91}^{VIS}(n) & y_{92}^{VIS}(n) & ... & y_{99}^{VIS}(n) \end{bmatrix}, \tag{4.2}$$

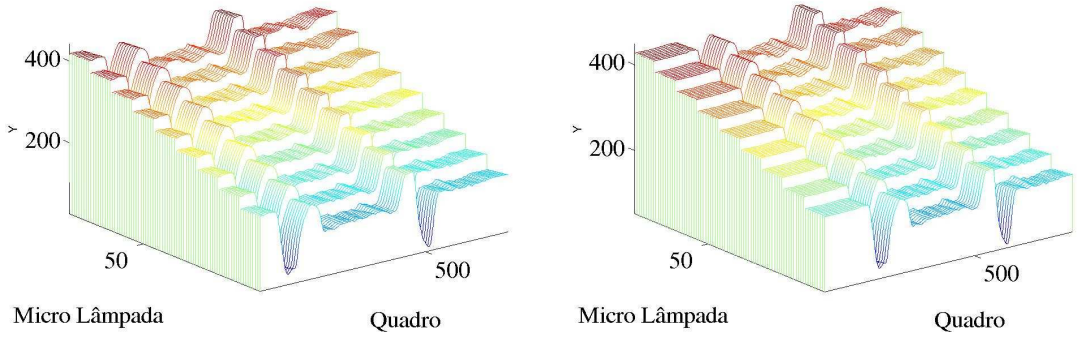
E o movimento sendo expresso por :

$$M_V(n) = Y_{VIS}(n) - Y_{VIS}(n-1), n = 2, 3..., N \tag{4.3}$$

$$M_{IR}(n) = Y_{IR}(n) - Y_{IR}(n-1), n = 2, 3..., N \tag{4.4}$$

De uma sequência de teste, podemos notar o comportamento da posição das lâmpadas a cada quadro n (Figura 4.2) e seu comportamento em relação a velocidade (Figura 4.3). Como esperado, a coordenada vertical de cada linha é semelhante

(pois as câmeras são pouco rotacionadas entre elas) e com isso podemos ainda usar um ponto (ou a média entre a posição vertical) de cada linha para representar o deslocamento do padrão.



(a) Visível (b) Infravermelho

Figura 4.2: Evolução quadro a quadro da posição vertical de cada lâmpada

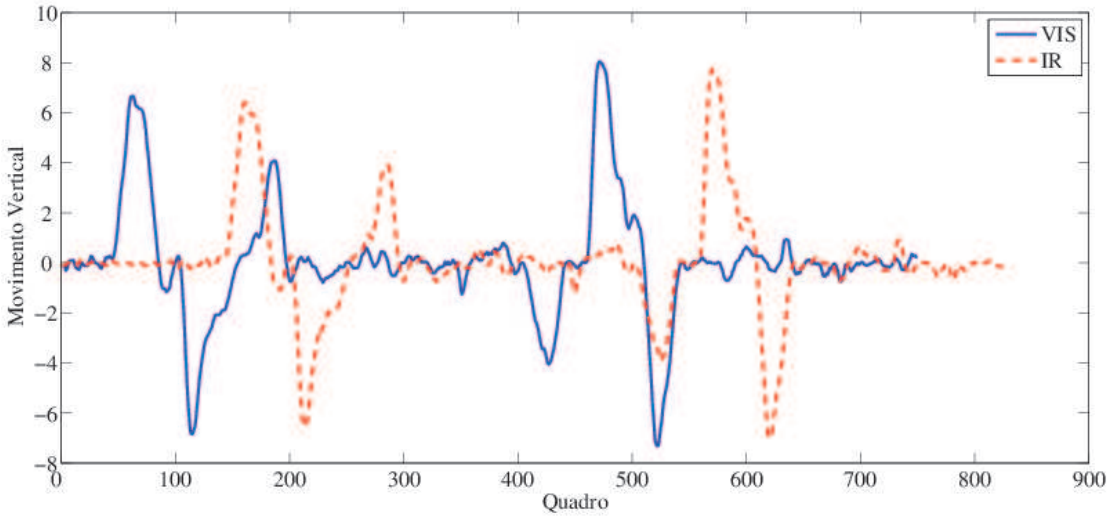


Figura 4.3: Evolução quadro a quadro da componente vertical da velocidade de uma dada lâmpada

Porém, utilizando as velocidades de todas as micro lâmpadas, e dadas as duas matrizes de velocidade M_V e M_I, deve-se definir a matriz de similaridade elas. A similaridade então pode ser escrita como o produto interno normalizado entre as

duas sequências M_V e M_I.

$$s(S_v, S_t, \Delta t) = \frac{\sum_{m=1}^{M} \sum_{n \in \mathbb{N}} M_V(m, n - \Delta t) M_I(m, n)}{\sqrt{\sum_{m=1}^{M} \sum_{n \in \mathbb{N}} (M_V(m, n - \Delta t))^2 \sum_{k=1}^{K} \sum_{l \in \mathbb{N}} (M_I(k, l))^2}} \tag{4.5}$$

Aplicando a similaridade a uma das sequência de teste (Figura 4.4), podemos ver um fator de correlação de 0.997 entre as sequências após o deslocamento Δt de 98 quadros.

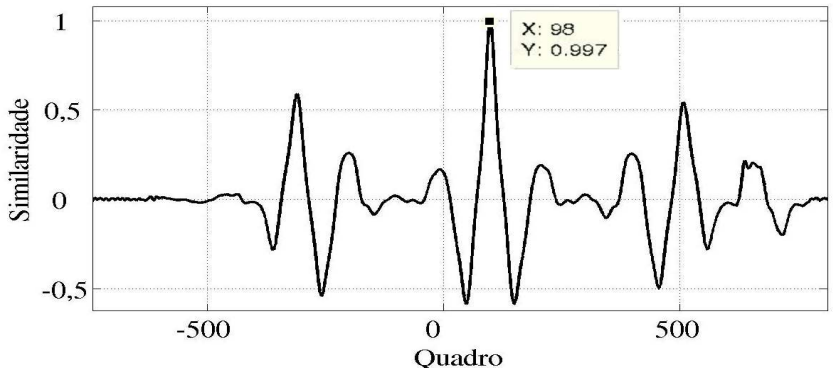
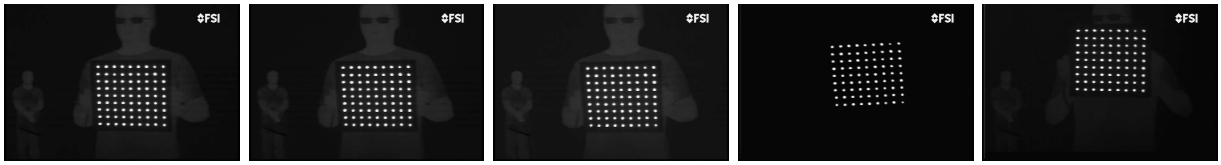
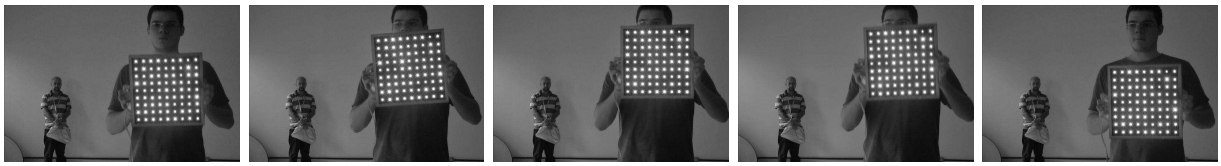


Figura 4.4: Resultado da similaridade em uma sequência de teste

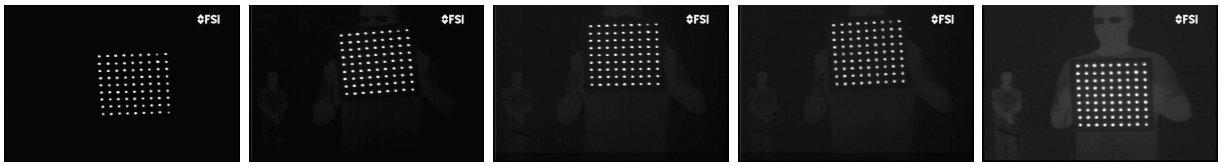
Temporalmente, o resultado obtido pode ser visto na sequência de imagens abaixo. Os quadros mostrados estão espaçados em 10 quadros, em ambos os casos.



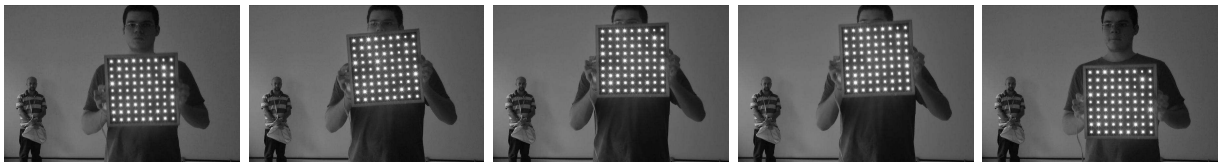
(a) Quadros IR desalinhados



(b) Quadros VIS desalinhados



(c) Quadros IR alinhados



(d) Quadros VIS alinhados

Figura 4.5: Alinhamento Temporal

4.2 Conclusão

Dado o método de sincronismo mostrado, podemos estimar um deslocamento temporal entre as duas sequências com a similaridade dada pela equação (4.5). Com isso é possível realizar a equivalência entre quadros das diferentes sequências permitindo, com isso, a realização do registro, como segue no próximo capítulo.

Os resultados da aplicação deste algoritmo em todas as sequências gravadas estão disponíveis na Seção 6.3.

Capítulo 5

Registro das Imagens

No Capítulo 2 foi mostrada a modelagem da matriz da câmera P e como é possível, a partir das homografias, conhecer a matriz de calibração e os demais parâmetros. Esta calibração tem o intuito de, dados todos os parâmetros, podermos realizar o registro das imagens obedecendo o proposto, ou seja, fazer corresponder pixel a pixel as imagens do par estéreo.

Este objetivo pode ser simplificado caso a cena esteja muito distante das câmeras. Isso torna a aproximação da *Homografia no infinito* válida (Seção 5.4).

Um fato importante na gravação das cenas que possibilita o registro é que toda a cena acontece em um mesmo plano. Ou seja, a aproximação citada anteriormente é estendida nos casos em que haja mais que um plano, porém a cena está suficientemente longe para ser considerada estando toda no mesmo plano. Essa distância é mensurada a partir da distância entre as câmeras. Em outras palavras, esta aproximação é válida quando, dada a disparidade das câmeras, a distância relativa entre o plano de aquisição e o plano da cena pode ser considerada muito maior que aquela. Na prática, foi utilizada a disparidade mínima entre as câmeras (câmeras lado a lado), e após testes, foi utilizada uma distância de cerca de 15 metros para a realização das cenas.

Mesmo com esta consideração há a impossibilidade de realizar o casamento de todos os planos presentes na imagem, mesmo estes sendo distantes. O possível é realizar o casamento de um plano específico, o plano da cena.

Neste capítulo serão apresentados os métodos usados para o registro e ainda o fundamento teórico para a criação do banco de vídeos (Seções 5.1, 5.2).

5.1 Geometria Epipolar

Um conceito importante no registro é a Geometria Epipolar. A geometria epipolar tem por objetivo estabelecer as relações entre as imagens capturadas de dois ou mais pontos de vistas distintos. Em outras palavras, dadas duas câmeras dispostas a capturar a mesma cena, onde o par de imagens representa duas perspectivas diferentes, a geometria epipolar tenta estabelecer as relações entre as imagens.

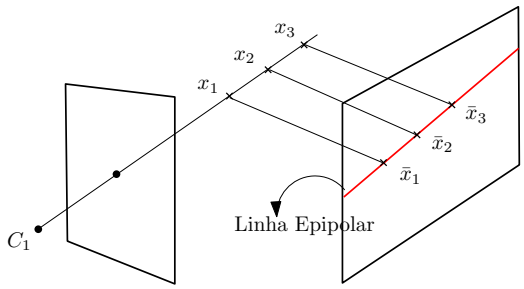


Figura 5.1: Linha Epipolar

Dado o problema citado, a pergunta é como estabelecer esta relação. Dos conceitos de modelagem de câmeras (Capítulo 2) surge um conceito importante. Dado os pontos centrais das duas câmeras, a linha que passa por estes dois pontos é chamada *linha base*. Tendo esta linha e o plano da cena, é possível definir o *epipolo*. O *epipolo* é então definido como o ponto de interseção entre a linha de base e o plano da imagem. Além destes conceitos, existem os conceitos de linha (Figura 5.1) e plano epipolares (Figura 5.2). O plano epipolar é definido pelos pontos centrais das duas câmeras e por um ponto do mundo real. E dado um plano epipolar, a linha epipolar é determinada pela interseção do plano da imagem com o plano epipolar.

Qualitativamente, vê-se que dado um ponto em uma das imagens, o resultado deste processo é uma linha em outra imagem. Esta linha representa as posições possíveis do ponto da primeira imagem na segunda. Esta conclusão traz a necessidade de se conhecer a distância do plano da cena. Com a linha epipolar e com a distância ao plano é possível obter a equivalência ponto a ponto. O mapeamento entre os pontos em uma imagem e as retas epipolares em outra é definida pela matriz fundamental F.

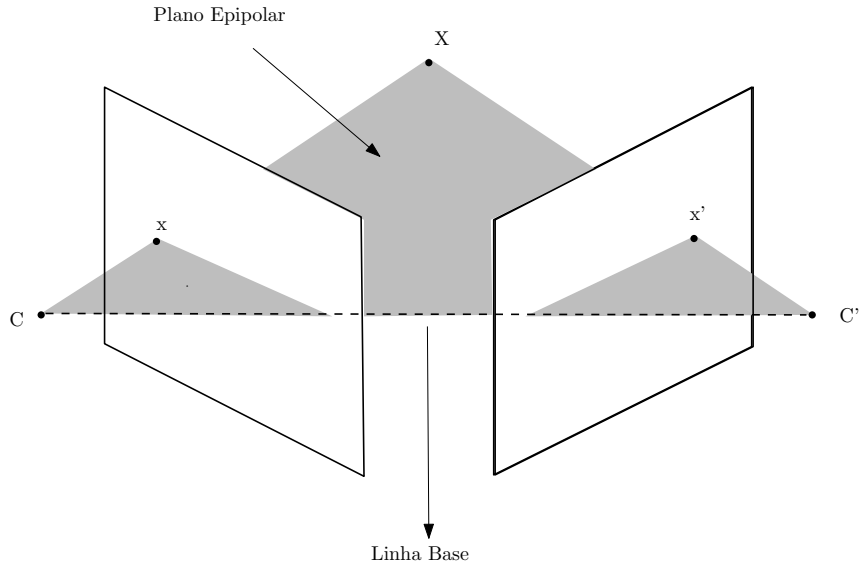


Figura 5.2: Plano Epipolar

5.2 Matriz Fundamental F

A matriz fundamental é a representação algébrica da geometria epipolar e representa o mapeamento entre os pontos de uma vista em linhas epipolares na outra. É importante ressaltar que a matriz fundamental independe da cena [1], depende apenas dos parâmetros das câmeras, como visto no Capítulo 2.

5.2.1 Propriedades

Dada a matriz fundamental, seguem algumas propriedades importantes relacionadas à geometria epipolar.

- F tem dimensão 3×3, 7 graus de liberdade e posto 2.
- Transposta:
  F é a matriz fundamental de um par de câmeras (P, P') e sua transposta representa as câmeras na ordem oposta (P', P).
- Linhas Epipolares :
  Para qualquer ponto x (representado em coordenadas homogêneas) em uma das imagens, a linha epipolar correspondente na segunda imagem é dada por

$l' = Fx$. Da mesma maneira, qualquer linha epipolar em uma imagem é mapeada em um ponto x' dado por $l = F^T x'$.

- Epipolo:
  Para qualquer ponto x, a linha epipolar $l' = Fx$ contém o epipolo e'. Então e' satisfaz $e'^T(Fx) = (e'^T F)x = 0$.

Tipicamente, para se calcular a matriz fundamental é utilizado um número de correspondências conhecidas nas duas vistas. Tendo os pontos q de uma imagem mapeados em q' na outra imagem, a matriz fundamental é definida como:

$$q'Fq = 0 \tag{5.1}$$

Para a solução deste sistema existem técnicas como o *Algoritmo dos oito pontos* presentes na literatura [1].

5.3 Matriz Essencial

A matriz essencial é uma simplificação da matriz fundamental através das normalizações das coordenadas. Algebricamente:

$$E = [t]_\times R \tag{5.2}$$

onde $[t]_\times$ é a matriz cujo produto com um ponto em coordenadas homogêneas é igual ao produto vetorial entre o vetor de translação t e as coordenadas homogêneas normalizadas dele e R é a matriz de rotação dada na calibração (Ver Equação 2.6).

Pode-se computar a matriz fundamental através da matriz essencial:

$$F = K'^{-T} E K^{-1} \tag{5.3}$$

onde o par (K', K) são as matrizes da calibração das duas câmeras.

Com os resultados obtidos na calibração de cada câmera é possível determinar as matrizes R e com isso obter a matriz fundamental F e/ou a matriz essencial E.

5.4 Homografia no Infinito

Utilizando os conceitos de matriz fundamental é possível realizar a correspondência entre um ponto em uma imagem e uma reta em outra imagem. Isto é resultado de uma das propriedades da matriz fundamental que diz que ela sempre tem posto 2. Porém, o que se deseja é realizar o casamento ponto a ponto entre as duas imagens e para isso é necessário o conhecimento da distância do plano da cena.

Dada as matrizes das câmeras:

$$P = K[I \mid 0] \tag{5.4}$$

$$P' = K'[R \mid t] \tag{5.5}$$

$$\tag{5.6}$$

E os epipolos [1] como :

$$e \ = P \begin{bmatrix} -R^T t \\ 1 \end{bmatrix} = KR^T t \tag{5.7}$$

$$e' \ = P' \begin{bmatrix} 0 \\ 1 \end{bmatrix} = K't \tag{5.8}$$

a matriz fundamental é escrita como :

$$F = [e']_\times K'RK^{-1} = K'^{-T}RK^T[e]_\times \tag{5.9}$$

E como :

$$x'^T F x = 0 \tag{5.10}$$

pode-se escrever a relação entre os pontos em função das matrizes de calibração e da distância do plano da cena Z:

$$x' = \underbrace{K'RK^{-1}}_{H_\infty} x \quad + \quad K't/Z \tag{5.11}$$

Pode-se observar que se quanto maior a distância Z menor a influência do segundo termo na transformação. Fazendo $Z \rightarrow \infty$, tem-se a *Homografia no infinito*, que é definida como a transformação que leva os pontos de uma imagem de uma câmera para outra considerando o plano da cena no infinito.

5.5 Calibração Estéreo

No Capítulo 2 foi definido o modelo de câmera como mostrado abaixo:

$$x = \underbrace{\begin{bmatrix} \alpha_x & s & p_x \\ 0 & \alpha_y & p_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} R & | & t \end{bmatrix}}_{P} X \tag{5.12}$$

Conhecida a matriz K, ainda resta ao modelo os parâmetros extrínsecos, ou seja as informações de rotação e translação entre as câmeras. Como foi utilizado o mesmo sistemas de coordenadas do mundo real para ambas as calibrações, os vetores de rotação e translação podem ser deduzidos através das homografias calculadas de cada vista das duas câmeras [1].

5.6 Retificação das Imagens

Como visto anteriormente, as matrizes de calibração têm por objetivo realizar a transformação $x \rightarrow x'$. Contudo, é mais apropriado realizar a transformação em ambas as imagens [1], levando-as a um plano em comum. A este artifício dá-se o nome de *Retificação*.

Como resultado da retificação, as linhas epipolares calculadas ficam paralelas ao eixo x. E por isso, ainda há um passo adicional que é a corrieção da suposição em que o plano da cena está no infinito. Isto é feito através da translação horizontal de uma das imagens até que o plano da cena esteja superposto ao mesmo plano na outra imagem.

Foi dito que, um ponto em uma imagem gera uma reta em outra imagem. Isto é verdade quando não há distorções como as geradas, por exemplo, pelas lentes. Na prática, um ponto em uma imagem gera uma curva em outra imagem de acordo com o modelo de distorção calculado. Isto foi levado em conta nos métodos de retificação apresentados em [1] para a criação do banco.

5.7 Conclusão

Neste capítulo foram apresentados os fundamentos sobre geometria epipolar o que possibilita o entendimento de como realizar o cálculo da homografia no infinito. Juntamente com a retificação, foi possível então a criação do banco de vídeos. Realizando a calibração em cada câmera, o sincronismo temporal e a retificação das imagens, as sequências com os vídeos casados puderam ser geradas. No próximo capítulo mostraremos os dados do resultado final do trabalho.

Capítulo 6

Banco de Vídeos

Após a gravação, edição e o registro dos vídeos obtidos, foi organizado o banco de vídeos e imagens. Neste capítulo segue a descrição do processo de gravação e descrição dos arquivos e das cenas.

Os dados estarão disponíveis em uma página contando com os resultados de todas as etapas de calibração. A página para acesso é www.lps.ufrj.br/ fusion.

Este Capítulo apresenta o detalhamento sobre as condições de gravação das cenas (Seção 6.1 e 6.2), seguido pelos detalhes das cenas, tais como o conteúdo (Seção 6.3), os dados de sincronismo temporal (Seção 6.4) e calibração (Seção 6.5).

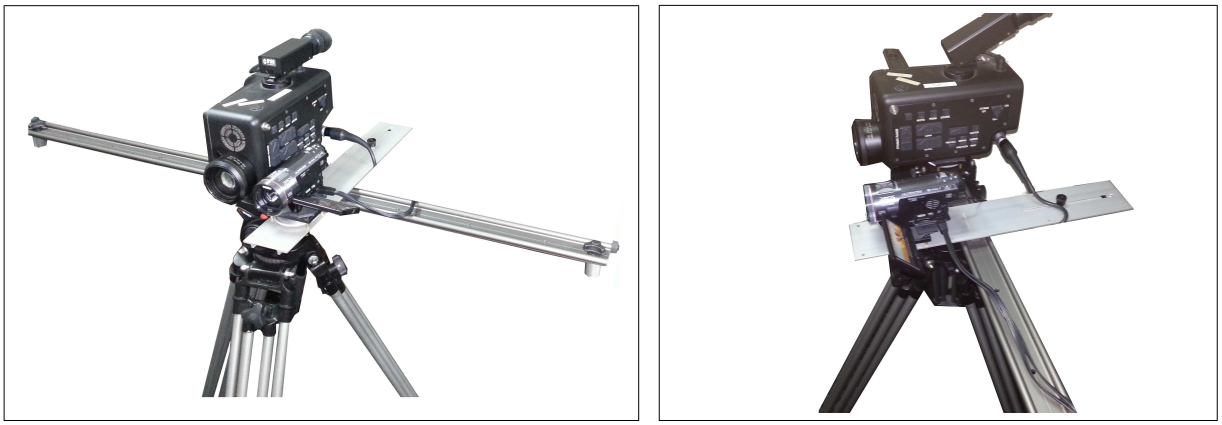
6.1 Aquisição das Imagens

O resultado do trabalho é o banco de imagens co-registradas. Estas imagens são gravadas a partir de uma câmera visível e uma infravermelha. A câmera visível é a *Panasonic HDC-TM700* com acionamento remoto e gravação interna. Foi utilizado uma taxa de quadros de gravação de 30 fps em todas as sequências. Os vídeos são codificados em H-264 com resolução de 1920 por 1080 pixels. O campo de visão desta câmera é ajustado ao campo de visão da câmera infravermelha. A câmera infravermelha é a *FLIR Prism DS* com faixa espectral de 3.6 a 5.0μm. As imagens são armazenadas no padrão NTSC com resolução 720 por 480 pixels. O campo de visão da câmera é de 17°(horizontal) por 13° (vertical) e a faixa de temperatura é de $-10\,°$C a $450\,°$C. A saída da câmera é de 320 por 244 pixels e é armazenada no formato anteriormente citado com o auxílio da placa de aquisição DeckLink HD

Extreme 3D e o software BlackMagic Media Express.

6.2 Gravação das Cenas

Para a gravação das cenas foi utilizado um tripé com encaixe para as duas câmeras e ajuste de distância entre elas adaptável. Além do ajuste horizontal é possível o ajuste de altura e inclinação das câmeras independentemente(Figura 6.1).



(a) (b)

Figura 6.1: Equipamento de suporte para o par estéreo

As cenas foram gravadas em 3 lugares e em diversas condições. Os lugares de gravação foram o Instituto Militar de Engenharia (IME), o Instituto de Pesquisas da Marinha e o Forte São João. No IME foram gravadas cenas internas com luz/pouca luz e na área externa, com iluminação natural tendo como cenário o jardim local. No IPqM foram gravadas cenas externas com alvos distantes e na área interna com iluminação variável. No Forte São João foram gravadas cenas externas com mais exposição a luz solar comparativamente às gravadas no IME.

6.3 Descrição do Banco

O banco foi separado em dois tipos de sequências: internas e externas. As cenas internas foram gravadas no Instituto Militar de Engenharia e no Instituto de Pesquisas da Marinha. As demais, gravadas em todos os outros ambientes, incluindo os dois anteriores.

As descrições sobre o conteúdo da cena e duração estão dispostas nas tabelas a seguir:

Local: **Instituto Militar de Engenharia**	
Cena	**Descrição da cena**
Interna 1	Duas pessoas transportando armas (uma pistola e uma faca cerâmica) em sacolas de materiais diferentes (plástica e tecido). Movimentação lateral das duas pessoas e revelação das armas ao final.
Interna 2	Igual à cena *Interna 1* porém com movimentação individual independente.
Interna 3	Depósito em uma cadeira de uma sacola contendo uma arma por uma pessoa que cruza a cena.
Interna 4	Igual à cena *Interna 3* porém com a presença de outra pessoa na cena (sem movimentação).
Interna 5	Ocultação de uma arma atrás de um jornal, sem movimentações e com baixa ocultação da cena visível por conta de fumaça.
Interna 6	Igual à cena *Interna 5* com mais ocultação por fumaça.
Externa 1	Gravação de ambiente externo com a iluminação natural com inúmeras pessoas em trânsito. Atentar a presença de um maquinário fixo na cena.
Externa 2	Igual à cena *Externa 1*.

Tabela 6.1: Descrição do conteúdo das cenas gravadas no IME

Local: **Forte São João**	
Cena	**Descrição da cena**
Externa 1	Três soldados com roupas camufladas escondidos na vegetação. Movimentação da vegetação e dos soldados ao final.
Externa 2	Igual à cena *Externa 1*
Externa 3	Dois soldados ocultos na vegetação que saem após serem encobertos por fumaça. Presença de um soldado oculto por fumaça e um exposto.
Externa 4	Igual à cena *Externa 3*
Externa 5	Uma pessoa oculta na vegetação portando uma arma. Pouca ocultação causada pela fumaça.
Externa 6	Cena com maior movimentação e diferentes armas.
Externa 7	Igual à cena *Externa 5* com muita ocultação causada pela fumaça.

Tabela 6.2: Descrição do conteúdo das cenas gravadas no Forte São João

Local: **Instituto de Pesquisas da Marinha**	
Cena	**Descrição da cena**
Interna 1	Ambiente escuro com duas pessoas em cena com movimentação em um mesmo plano. Mudança de cenário com iluminação de certas áreas em tempos distintos.
Interna 2	Ambiente com iluminação artificial e quatro pessoas em cena. Uma das pessoas está ocultando uma arma. Mudança da iluminação do ambiente com o mesmo padrão anterior.
Interna 3	Ambiente com iluminação artificial e movimentação de cinco pessoas durante a cena. Presença de fumaça no decorrer da sequência e alteração da iluminação como nos casos anteriores.
Interna 4	Presença de fumaça muito densa com ocultação total do cenário. Aumento de visibilidade gradual da movimentação de até quatro pessoas. Número de pessoas variável durante a cena. Iluminação artificial constante.
Interna 5	Uma pessoa em movimentação com ocultação parcial ou total no decorrer da cena. Presença de um objeto quente oculto à imagem visível.

Tabela 6.3: Descrição do conteúdo das cenas gravadas no IPqM (Internas)

Local: **Instituto de Pesquisas da Marinha**	
Cena	**Descrição da cena**
Externa 1	Iluminação natural com duas pessoas ocultas na vegetação em movimentação (Oculto-Visível).
Externa 2	Igual à *Externa 1* com mudança na movimentação (Visível-Oculto).
Externa 3	Iguais as anteriores com presença de um carro retirando um dos integrantes da cena.
Externa 4	Mudança de cenário em relação aos anteriores mantendo a iluminação natural. Apenas uma pessoa em cena.
Externa 5	Igual à *Externa 4*
Externa 6	Gravação da Baia de Guanabara ao entardecer. Pouca movimentação dos navios.
Externa 7	Igual à anterior com mudança nas configurações das câmeras.
Externa 8	Igual à *Externa 6* com a mudança da lente.
Externa 9	Gravação da Baia de Guanabara durante a noite.
Externa 10	Igual à anterior com mudança nas configurações das câmeras.
Externa 11	Igual à *Externa 10* com a mudança da lente.

Tabela 6.4: Descrição do conteúdo das cenas gravadas no IPqM (Externas)

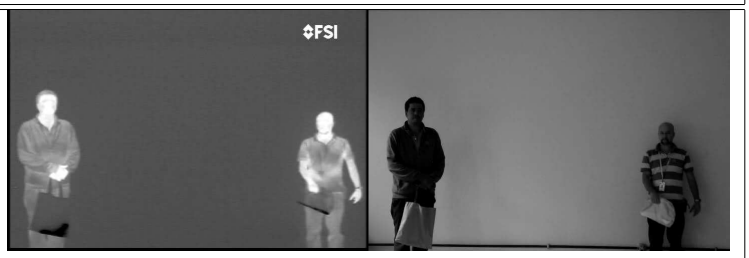
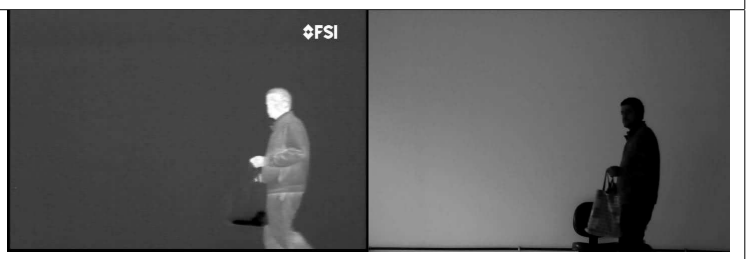
Local: **Instituto Militar de Engenharia**		
Cena	Duração	***Screenshot***
Interna 1	0m46s	
Interna 2	0m44s	
Interna 3	0m19s	
Interna 4	0m19s	
Interna 5	0m47s	
Interna 6	0m46s	

Tabela 6.5: Cenas gravadas no IME

Local: **Instituto Militar de Engenharia**		
Cena	Duração	***Screenshot***
Externa 1	2m49s	
Externa 2	3m42s	

Tabela 6.6: Cenas gravadas no IME (Cont.)

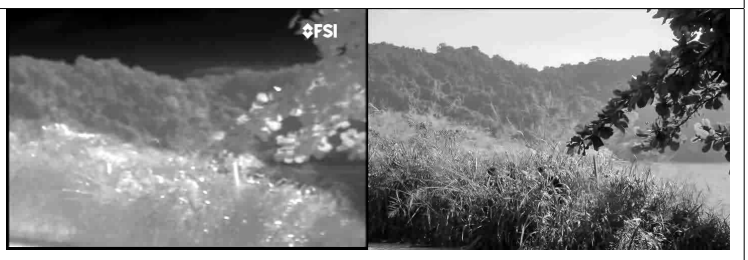
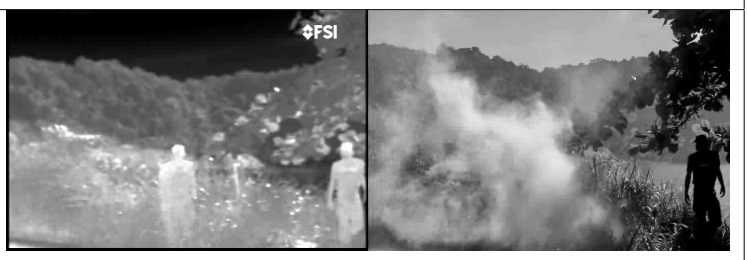
Local: **Forte São João**		
Cena	Duração	***Screenshot***
Externa 1	1m05s	
Externa 2	1m02s	
Externa 3	0m41s	

Tabela 6.7: Cenas gravadas no Forte São João

Local: **Forte São João**		
Externa 4	0m41s	
Externa 5	0m52s	
Externa 6	1m30s	
Externa 7	0m47s	

Tabela 6.8: Cenas gravadas no Forte São João (Cont.)

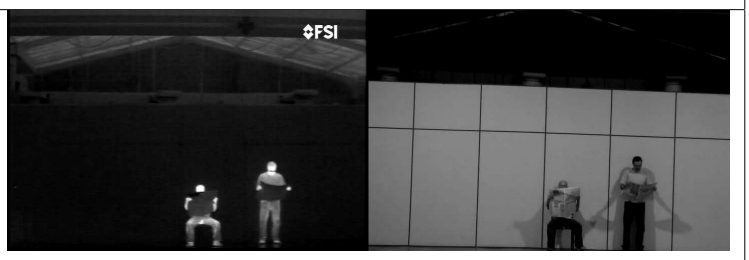
Local: **Instituto de Pesquisas da Marinha**		
Cena	Duração	***Screenshot***
Interna 1	2m03s	
Interna 2	1m37s	
Interna 3	5m13s	
Interna 4	5m32s	
Interna 5	5m51s	

Tabela 6.9: Cenas gravadas no IPqM (Internas)

Local: **Instituto de Pesquisas da Marinha**		
Cena	Duração	***Screenshot***
Externa 1	1m00s	
Externa 2	0m54s	
Externa 3	5m25s	
Externa 4	0m43s	
Externa 5	0m43s	
Externa 6	0m59s	

Tabela 6.10: Cenas gravadas no IPqM (Externas)

Local: **Instituto de Pesquisas da Marinha**		
Cena	Duração	***Screenshot***
Externa 7	1m56s	
Externa 8	1m58s	
Externa 9	4m00s	
Externa 10	2m45s	
Externa 11	4m09s	

Tabela 6.11: Cenas gravadas no IPqM (Externas - Cont.)

6.4 Alinhamento Temporal

A partir da formulação mostrada no Capítulo 4 podemos escrever a similaridade como :

$$s(S_v, S_t, \Delta t) = \frac{\sum_{m=1}^{M} \sum_{n \in \mathbb{N}} M_V(m, n - \Delta t) M_I(m, n)}{\sqrt{\sum_{m=1}^{M} \sum_{n \in \mathbb{N}} (M_V(m, n - \Delta t))^2 \sum_{k=1}^{K} \sum_{l \in \mathbb{N}} (M_I(k, l))^2}} \tag{6.1}$$

A partir disto, as tabelas presentes no Apêndice A contém os dados do sincronismo resultantes desta formulação, sendo eles a similaridade e o deslocamento temporal da cena gravada, para todas as sequências presentes no banco.

Em suma, os dados da similaridade tem média 0.99644 e variância 0.00002. O resultado da similaridade é a métrica utilizada para mensurar o sincronismo entre as sequências, tendo como valor máximo 1 (como pode ser visto na equação 6.1).

6.5 Calibração individual

Como métrica para a calibração individual foi utilizado o erro médio de reprojeção. O MRE (*Mean Reprojection Error*) é calculado através da reprojeção dos pontos de calibração nos pontos do mundo real através da matriz de calibração encontrada (Equação 2.27).

Os dados referentes ao erro de médio de reprojeção estão disponíveis no Apêndice B e os demais dados estarão disponíveis na página do projeto.

6.6 Calibração Estéreo

Como dito, a calibração estéreo apresenta os parâmetros extrínsecos relativos às duas câmeras, sendo estes a rotação e a translação relativa entre elas. Durante a gravação das sequências em um dado ambiente, não houve mudanças na montagem das câmeras e nenhum contanto nas câmeras (pois o acionamento de gravação é remoto). Com isso foram utilizados as mesmas configurações de rotação e translação em um conjunto de sequências.

Estes dados estarão presentes na página juntamente com o restante dos dados para consulta.

Capítulo 7

Conclusões

O trabalho tem como produto final o banco de vídeos IR/Visível registrados. Para facilitar a aquisição deste banco, foi desenvolvido uma página na internet onde este é disponibilizado. As versões disponíveis são os vídeos sem compressão e os arquivos com cada quadro em formato *jpg*.

Nesta página de internet há a disponibilidade da consulta de todos os artigos resultantes deste trabalho e os arquivos de simulação para a repetição dos passos de sincronismo, calibração e retificação.

7.1 Trabalhos Futuros

Como trabalho futuro tem-se o uso do banco nos algoritmos de fusão já desenvolvidos, o aprimoramento destas técnicas a partir de um banco com mais informações e realizar os testes nos algoritmos de fusão de imagens em vídeos fazendo uso da informação temporal.

Apêndice A

Dados de Sincronismo do Banco de Vídeos

Os dados contidos neste apêndice são resultado da aplicação do método descrito no Capítulo 4 sendo a similaridade dada pela equação 4.5.

Local: **Instituto Militar de Engenharia**		
Cena	Deslocamento Temporal (Quadros)	**Similaridade**
Interna 1	54.0	0.995144
Interna 2	38.0	0.998834
Interna 3	7.0	0.999019
Interna 4	39.0	0.998925
Interna 5	30.0	0.997871
Interna 6	22.0	0.998804
Externa 1	-1.0	0.978544
Externa 2	20.0	0.997034

Tabela A.1: Descrição do sincronismo das cenas gravadas no IME

Local: **Forte São João**		
Cena	Deslocamento Temporal (Quadros)	**Similaridade**
Externa 1	99.0	0.998863
Externa 2	49.0	0.997651
Externa 3	48.0	0.985379
Externa 4	69.0	0.995536
Externa 5	80.0	0.997811
Externa 6	56.0	0.999359
Externa 7	73.0	0.999743

Tabela A.2: Descrição do sincronismo das cenas gravadas no Forte São João

Local: **IPqM**		
Cena	Deslocamento Temporal (Quadros)	**Similaridade**
Interna 1	59.0	0.999021
Interna 2	36.0	0.997495
Interna 3	79.0	0.997415
Interna 4	79.0	0.997042
Interna 5	98.0	0.997717
Externa 1	109.0	0.998747
Externa 2	96.0	0.998523
Externa 3	89.0	0.998534
Externa 4	40.0	0.996802
Externa 5	828.0	0.994227
Externa 6	93.0	0.996662
Externa 7	82.0	0.998372
Externa 8	40.0	0.998056
Externa 9	43.0	0.997986
Externa 10	18.0	0.997377
Externa 11	16.0	0.998870

Tabela A.3: Descrição do sincronismo das cenas gravadas no IPqM

Apêndice B

Dados de Calibração do Banco de Vídeos

O MRE (*Mean Reprojection Error* ou Erro médio de Reprojeção) é a métrica utilizada para mensurar a qualidade da calibração. Tendo em vista um plano determinado comum às duas câmeras, foi descrita a posição do padrão no mundo real. Então, com essa descrição, foi aplicada a calibração obtida com os pontos detectados e mensurado o erro médio entre a posição real e a obtida após a calibração. Os resultados desta mensuração estão disponíveis nas tabelas abaixo.

Local: **Instituto Militar de Engenharia**		
Cena	MRE_{VIS}	MRE_{IR}
Interna 1	0.02694	0.02029
Interna 2	0.02330	0.03719
Interna 3	0.02654	0.04443
Interna 4	0.02342	0.03054
Interna 5	0.03365	0.03842
Interna 6	0.02534	0.03241
Externa 1	0.01756	0.03206
Externa 2	0.02185	0.04551

Tabela B.1: Descrição da calibração das cenas gravadas no IME

Local: **Forte São João**		
Cena	MRE_{VIS}	MRE_{IR}
Externa 1	0.02440	0.03653
Externa 2	0.02106	0.03253
Externa 3	0.03414	0.04131
Externa 4	0.02078	0.02575
Externa 5	0.03123	0.03213
Externa 6	0.02047	0.03555
Externa 7	0.01840	0.02938

Tabela B.2: Descrição da calibração das cenas gravadas no Forte São João

Local: **Instituto de Pesquisas da Marinha**		
Cena	MRE_{VIS}	MRE_{IR}
Interna 1	0.03665	0.045549
Interna 2	0.03443	0.04658
Interna 3	0.02865	0.03372
Interna 4	0.02407	0.03586
Interna 5	0.02563	0.04094
Externa 1	0.02024	0.05018
Externa 2	0.02307	0.03453
Externa 3	0.02456	0.03422
Externa 4	0.01952	0.04166
Externa 5	0.02075	0.05430
Externa 6	0.02747	0.04407
Externa 7	0.03290	0.04662
Externa 8	0.04233	0.04785
Externa 9	0.02805	0.04161
Externa 10	0.03431	0.04429
Externa 11	0.02277	0.03131

Tabela B.3: Descrição da calibração das cenas gravadas no IPqM

Referências Bibliográficas

[1] HARTLEY, R. and ZISSERMAN, A. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2nd edition, 2003.

[2] ZHANG, Z., "A Flexible New Technique for Camera Calibration", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 22, n. 10, pp. 1066–1077, 2000.

[3] HEIKKILA, J., " A Four-step Camera Calibration Procedure with Implicit Image Correction", *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp.1106,1112, 17-19 June 1997.

[4] HEIKKILA, J., "Geometric Camera Calibration using Circular Control Points", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 10, pp. 1066–1077, 2000.

[5] PRAKASH, C. D.,KARAM, L. J. "Camera Calibration using Adaptive Segmentation and Ellipse Fitting for Localizing Control Points", *Proceedings of the 2012 IEEE International Conference on Image Processing*, October, 2012.

[6] ELLMAUTHALER, A., DA SILVA, E. A. B., PAGLIARI, C. L.,NEVES, S. R., "Infrared-visible image fusion using the undecimated wavelet transform with spectral factorization and target extraction", *2012 19th IEEE International Conference on Image Processing*, pp. 2661–2664, September. 2012.

[7] CANNY, J. , "A computacional approach to edge detection", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 8, no. 6, pp.679-698, November 1986.

[8] VIDAS, S., LAKEMOND, R., DENMAN, S., *et al.*, "A Mask-Based Approach for the Geometric Calibration of Thermal-Infrared Cameras", *IEEE Transactions on Instrumentation and Measurement*, v. 61, n. 6, pp. 1625–1635, June 2012.

[9] FITZGIBBON, A., PILU, M., FISHER, R., "Direct least square fitting of ellipses",*IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, no. 5, pp. 476-480, May 1999.

[10] DATTA, A., KIM, J.-S., KANADE, T., "Accurate camera calibration using iterative refinement of control points", *2009 IEEE 12th International Conference on Computer Vision Workshops, ICCV Workshops*, pp. 1201–1208, Sep. 2009.

[11] SCHELLE, S.R., CHAN, A. L., "Enhanced Target Tracking Through Infrared-Visible Image Fusion", *14th International Conference on Information Fusion*, July 2011.

[12] ELLMAUTHALER, A., DA SILVA, E. A. B., PAGLIARI, C. L., GOIS, J.N., NEVES, S. R., "A Novel Iterative Calibration Approach for Thermal Infrared Cameras ", *IEEE International Conference on Image Processing* , 2013.

[13] ELLMAUTHALER, A., DA SILVA, E. A. B., PAGLIARI, C. L., GOIS J. N., "A Robust Temporal Alignment Technique for Infrared and Visible-Light Video Sequences", *XXXI Simpósio Brasileiro de Telecomunicações*, September 2013.

[14] GONZALEZ, R. C. , WOODS, R.E.,EDDINS, S. L., "Digital Image Processing Using MATLAB" *Gatesmark Publishing*, 2009.

[15] SEBER, G. A. F., "Multivariate Observations" *Hoboken, NJ: Jonh Wiley and Sons,Inc.*, 1984.

[16] OPENCV, "OpenCV v2.11 C Reference",
http://opencv.willowgarage.com/documentation/c/index.html,
(último acesso em Maio de 2013).