



**Manuel Eduardo Loaiza Fernandez**

**Calibração de Múltiplas Câmeras baseado em um Padrão  
Invariante**

**Tese de Doutorado**

Tese apresentada como requisito parcial para obtenção do título de Doutor pelo Programa de Pós-Graduação em Informática do Departamento de Informática da PUC-Rio.

Orientadores: Marcelo Gattass  
Co-Orientador: Alberto Barbosa Raposo

Rio de Janeiro, agosto de 2009



**Manuel Eduardo Loiza Fernandez**

## **Calibração de Múltiplas Câmeras baseado em um Padrão Invariante**

Tese apresentada como requisito parcial para obtenção do título de Doutor pelo Programa de Pós-Graduação em Informática do Departamento de Informática do Centro Técnico Científico da PUC-Rio. Aprovada pela Comissão Examinadora abaixo assinada.

**Prof. Marcelo Gattass**

Orientador  
Departamento de Informática PUC-Rio

**Prof. Alberto Barbosa Raposo**

Co-Orientador  
Departamento de Informática PUC-Rio

**Prof. Waldemar Celes Filho**

Departamento de Informática PUC-Rio

**Prof. Raul Queiroz Feitosa**

Departamento de Engenharia Elétrica PUC-Rio

**Prof. Paulo Cezar Pinto Carvalho**

Instituto Nacional de Matemática Pura e Aplicada IMPA

**Prof. Romero Tori**

Escola Politécnica da Universidade de São Paulo USP

**Prof. José Eugenio Leal**

Coordenador Setorial do Centro Técnico Científico - PUC-Rio

Rio de Janeiro, 20 de agosto de 2009

Todos os direitos reservados. É proibida a reprodução total ou parcial do trabalho sem autorização da universidade, do autor e do orientador.

### **Manuel Eduardo Loaiza Fernandez**

Graduou-se em Engenharia de Sistemas na Faculdade de Produção e Serviços da Universidad Nacional de San Agustín (UNSA), Arequipa - Perú em 2002 e mestrado em Informática pela Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio), Rio de Janeiro -Brasil em 2005. Atualmente pertence ao Grupo de Computação Gráfica - Tecgraf da Pontifícia Universidade Católica do Rio de Janeiro. Tem experiência na área de Ciência da Computação, com ênfase em Computação Gráfica, Visualização 3D, Realidade Virtual, Realidade Aumentada e Visão Computacional.

#### Ficha Catalográfica

Loaiza Fernandez, Manuel Eduardo

Calibração de múltiplas câmeras baseado em um padrão invariante / Manuel Eduardo Loaiza Fernandez ; orientador: Marcelo Gattass ; co-orientador: Alberto Barbosa Raposo. – 2009.

131 f. : il. (color.) ; 30 cm

Tese (Doutorado em Informática)–Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2009.

Incluí referências bibliográficas.

1. Informática – Teses. 2. Calibração de câmera. 3. Geometria epipolar. 4. Estrutura a partir do movimento. 5. Características projetivas invariante. 6. Processamento de imagem. 7. Rastreamento óptico. I. Gattass, Marcelo. II. Raposo, Alberto Barbosa. III. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Informática. IV. Título.

CDD: 004

À minha esposa, Mayra, ao meu filho, Diego e à minha família.

## **Agradecimentos**

A Deus, por ter me dado a vida e a força necessária para continuar cada dia com o mesmo entusiasmo e dedicação.

A minha esposa, Mayra e meu filho Diego, pela paciência tida neste últimos anos.

Aos meus pais, Eduardo e Violeta e meus avós, Clotilde, Edgar e Blanca, pelo amor e exemplo de vida que me deram e que sempre lembro em todos os momentos da minha vida.

Aos meus tios Oscar, Milar, Enrique, Cira, Frida, Marlene, Gloria, Digber e Gladis pelos conselhos e ajuda que recebi deles.

Aos meus irmãos, Luis e Marizze, por todo o carinho e ânimos que me deram durante estes dois anos.

Aos meus orientadores Marcelo Gattass e Alberto Barbosa Raposo, pela confiança depositada no meu trabalho e por todos os ensinamentos repassados ao longo dos anos. Agradeço-lhes profundamente por seu apoio e amizade.

Aos meus amigos e colegas do Tecgraf Felipe, Peter Hohl, Lucas, Luciano, Cesar, Cristina, Gustavo por terem me presenteado com suas amizades e apoio durante o desenvolvimento desta tese.

E finalmente, a todas as pessoas que não foram mencionadas, mas que foram importantes em algum momento da minha vida.

Ao CNPq e à PUC-Rio, pelos auxílios concedidos, sem os quais este trabalho não poderia ter sido realizado.

## Resumo

Loaiza Fernandez, Manuel Eduardo; Gattass, Marcelo; Raposo, Alberto B. **Calibração de Múltiplas Câmeras baseado em um Padrão Invariante.** Rio de Janeiro, 2009. 131p. Tese de Doutorado - Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro.

O processo de calibração de câmeras é uma etapa importante na instalação dos sistemas de rastreamento óptico. Da qualidade da calibração deriva o funcionamento correto e preciso do sistema de rastreamento. Diversos métodos de calibração têm sido propostos na literatura em conjunto com o uso de artefatos sintéticos definidos como padrões de calibração. Esses padrões, de forma e tamanho conhecidos, permitem a aquisição de pontos de referência que são utilizados para a determinação dos parâmetros das câmeras. Para minimizar erros, esta aquisição deve ser feita em todo o espaço de rastreamento. A fácil identificação dos pontos de referência torna o processo de aquisição eficiente. A quantidade e a qualidade das relações geométricas das feições do padrão influenciam diretamente na precisão dos parâmetros de calibração obtidos. É nesse contexto que esta tese se encaixa, propondo um novo método para múltipla calibração de câmeras, que é eficiente e produz resultados tão ou mais precisos que os métodos atualmente disponíveis na literatura. Nosso método também propõe um novo tipo de padrão de calibração que torna a tarefa de captura e reconhecimento de pontos de calibração mais robusta e eficiente. Deste padrão também derivam relações que aumentam a precisão do rastreamento. Nesta tese o processo de calibração de múltiplas câmeras é revisitado e estruturado de forma a permitir uma comparação das principais propostas da literatura com o método proposto. Esta estruturação também dá suporte a uma implementação flexível que permite a reprodução numérica de diferentes propostas. Finalmente, este trabalho apresenta resultados numéricos que permitem tirar algumas conclusões.

## Palavras-chave

Calibração de câmera; Geometria epipolar; Estrutura a partir do movimento; Características projetivas invariante; Processamento de imagem; Rastreamento óptico.

## **Abstract**

Manuel Eduardo Loaiza Fernandez; Gattass, Marcelo (Advisor); Raposo, Alberto B. (Co-Advisor). **Multiple Camera Calibration Based on Invariant Pattern.** Rio de Janeiro, 2009. 131p. DSc. Thesis – Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro.

The calibration of multiple cameras is an important step in the installation of optical tracking systems. The accuracy of a tracking system is directly related to the quality of the calibration process. Several calibration methods have been proposed in the literature in conjunction with the use of artifacts, called calibration patterns. These patterns, with shape and size known, allow the capture of reference points to compute camera parameters. To yield good results these points must be uniformly distributed over the tracking area. The determination of the reference points in the image is an expensive process prone to errors. The use of a good calibration pattern can reduce these problems. This thesis proposes a new multiple camera calibration method that is efficient and yields better results than previously proposed methods available in the literature. Our method also proposes the use of a new simple calibration pattern based on perspective invariant properties and useful geometric properties. This pattern yields robust reference point identification and more precise tracking. This thesis also revisits the multiple calibration process and suggests a framework to compare the existing methods including the one proposed here. This framework is used to produce a flexible implementation that allows a numerical evaluation that demonstrates the benefits of the proposed method. Finally the thesis presents some conclusions and suggestions for further work.

## **Keywords**

Camera Calibration; Epipolar Geometry; Structure from motion; Projective invariants; Image processing; Optical tracking.

# Sumário

Notações	15
1 Introdução	16
1.1. Motivação	17
1.2. Objetivo	19
1.3. Contribuições	20
1.4. Organização da tese	21
2 Múltipla calibração de câmeras	23
2.1. Primeira etapa: aquisição de dados para calibração	24
2.1.1. Padrões usados na calibração de câmeras	26
2.1.2. Padrão planar – 2D	27
2.1.3. Padrão unidimensional – 1D	29
2.1.4. Padrão adimensional	32
2.2. Segunda etapa: calibração inicial	35
2.2.1. Calibração fotogramétrica	35
2.2.1.1. Calibração baseada no cálculo de homografias	36
2.2.1.2. Cálculo da homografia	36
2.2.1.3. Determinação dos parâmetros intrínsecos e extrínsecos	38
2.2.1.4. Determinação dos coeficientes de distorção da lente	41
2.2.2. Auto-calibração	43
2.2.2.1. Geometria epipolar	44
2.2.2.2. Matriz fundamental	46
2.2.2.3. Cálculo da matriz fundamental	48
2.2.2.4. Determinação dos parâmetros extrínsecos	50
2.3. Terceira etapa: otimização da calibração	52
2.3.1. Funções de controle	53
2.3.1.1. Reprojeção 3D – 2D	53
2.3.1.2. Reprojeção 2D – 2D	54
2.3.1.3. Distância 3D entre pontos de referência reconstruídos	55



2.4. Distribuição das câmeras	57
2.4.1. Distribuição centralizada de câmeras	57
2.4.2. Distribuição descentralizada das câmeras	58
3 Trabalhos relacionados	61
3.1. Método proposto por Jean-Yves Bouguet	62
3.2. Método proposto por Borghese	64
3.3. Método proposto por Svoboda	67
3.4. Método proposto por Kurillo	69
4 Método proposto	72
4.1. Descrição do método	73
4.1.1. Aquisição de dados	77
4.1.1.1. Captura do padrão planar	77
4.1.1.2. Captura do padrão invariante	78
4.1.1.2.1. Treinamento do padrão invariante	78
4.1.1.2.2. Captura da amostra do padrão invariante	78
4.1.2. Calibração inicial	78
4.1.2.1. Calibração fotogramétrica	78
4.1.2.2. Auto-calibração	78
4.1.2.3. Otimização dos parâmetros	78
4.1.2.3.1. Definição da função de erro global	78
4.1.2.3.2. Colinearidade 2D e 3D	78
4.2. Sumarizando o método	78
5 Resultados	78
5.1. Equipamento e ambiente de teste	78
5.2. Metodologia dos testes	78
5.3. Descrição dos resultados	78
6 Conclusões	78
6.1. Trabalhos futuros	78
7 Referências Bibliográficas	78

8 Apêndice A: Modelo de câmera	78
8.1.1. Sistemas de coordenadas	78
8.1.2. Transformações	78
8.1.3. Parâmetros extrínsecos	78
8.1.4. Parâmetros intrínsecos	78
8.1.5. Coeficientes de distorção das lentes	78
9 Apêndice B: Projetivas Invariantes	78
9.1. Calculo do $P^2$ -Invariant para 4 pontos colineares	78

## Lista de figuras

Figura 1: Aplicações de realidade virtual e aumentada.	16
Figura 2: Sistemas de rastreamento óptico.	18
Figura 3: Sistema estéreo e de múltiplas câmeras.	23
Figura 4: Padrões planares usados para calibração de câmera em trabalhos como, (a) Bouguet [5], (b) Zhang [40] e Tsai [35][36] e (c) Heikkila [11].	27
Figura 5: Padrões planares usados em (a) Olsen e Hoover [25] e (b) Bouguet [5].	28
Figura 6: Padrão de calibração e captura proposto no trabalho de Borghese[4].	30
Figura 7: (a) Ambiente com uso de marcadores retrorreflexivos em Uematsu et al.[37], (b) marcadores coloridos usados como padrão em Mitchelson et al.[22].	31
Figura 8: Marcadores adimensionais propostos por (a)Svoboda et al. [32], (b)Pintaric et al. [28] e [13].	33
Figura 9: Elementos da geometria epipolar - linha epipolar, epipolos, linha base, plano epipolar.	45
Figura 10: Correspondência entre duas vistas de um sistema estéreo de câmeras.	46
Figura 11: Ambiguidade da definição das quatro possíveis soluções para $R$ e $t$ [10].	51
Figura 12: Exemplificação do erro definido pela distância da linha epipolar ao ponto correspondente.	55
Figura 13: Distribuição centralizada de câmeras.	58
Figura 14: Distribuição descentralizada de câmeras.	58
Figura 15: Calibração de um sistema estéreo de câmeras feita por Bouguet [5].	62
Figura 16: Padrões de calibração usados em [1][4][39]	65
Figura 17: O ponteiro (a), distribuição de 16 câmeras (b) e esquema da múltipla calibração de câmeras definido por Svoboda et al.[32]	65

Figura 18: Análise da correlação de um sistema composto por cinco câmeras usando teoria de grafos definida em [14].	70
Figura 19: Evolução dos padrões (a) planar, (b) unidimensional, (c) adimensional.	74
Figura 20: Padrões invariantes e processamento 2D para identificação individual.	75
Figura 21: Captura do padrão planar para calibração fotogramétrica.	77
Figura 22: Processo de treinamento feito simultaneamente num sistema de 4 câmeras.	78
Figura 23: Identificação grupal e etiquetagem individual dos marcadores que compõem os padrões invariantes.	78
Figura 24: Captura simultânea dois padrões invariantes num sistema de quatro câmeras.	78
Figura 25: Fluxograma geral do método de múltipla calibração de câmeras proposto nesta tese.	78
Figura 26: Fluxograma do processo de calibração fotogramétrica.	78
Figura 27: Fluxograma do processo de treinamento dos padrões invariantes.	78
Figura 28: Fluxograma do processo de captura dos padrões invariantes previamente treinados.	78
Figura 29: Fluxograma do processo de captura dos padrões invariantes previamente treinados (Continuação).	78
Figura 30: Fluxograma do processo de auto - calibração.	78
Figura 31: Fluxograma do processo de otimização final dos parâmetros das câmeras estimados na calibração inicial.	78
Figura 32: (a) Câmeras web Logitech, (b) e (c) Câmeras Fire-I.	78
Figura 33: Três sistemas estéreo de câmeras formados pelas seis câmeras usadas para os testes.	78
Figura 34: (a) Sistema X e (b) sistema Y de quatro câmeras.	78
Figura 35: Dois padrões invariantes usados nos testes.	78
Figura 36: Duas câmeras firewire conectadas em série.	78
Figura A1: Modelo de captura da câmera Pinhole [17].	78
Figura A2: Parâmetros extrínsecos na modelagem da	

transformação entre os SCM para o SCC [17].	78
Figura A3: Parâmetros intrínsecos na modelagem da transformação entre o SCC para o SCI [17].	78
Figura A4: Projeção de um ponto $P_c$ no SCC para um ponto $P'$ no SCI [17].	78
Figura A5: Transformação do SCI para o SCP[17].	78
Figura A6: Distorção causada por distintos tipos de lente - (a) Barrel,(b) Pincushion.	78

## Lista de tabelas

Tabela 1: Valores dos identificadores para cada padrão invariante usado nos testes.	78
Tabela 2: Resultados do erro médio medidos na amostra <i>B</i> , 1000 vistas, capturada para teste da calibração realizada do sistema <b>X</b> de quatro câmeras firewire.	78
Tabela 3: Resultados do erro médio medidos na amostra <i>B</i> , 1000 vistas, capturada para teste da calibração realizada do sistema <b>Y</b> de quatro câmeras, duas firewire e duas usb.	78
Tabela 4: Ganho percentual ao analisar os resultados apresentados na tabela 1.	78
Tabela 5: Ganho percentual ao analisar os resultados apresentados na tabela 2.	78
Tabela 6: Comparação valor do erro médio referente da medição da distância entre marcadores definidos em métodos acadêmicos e produtos comerciais.	78

## Notações

Para melhorar a leitura e interpretação das variáveis usadas para descrever algumas fórmulas apresentadas dentro dos capítulos desta tese, a seguir serão apresentadas algumas descrições básicas destas notações.

- $\mathbf{K}$  é a matriz que contém a informação dos parâmetros intrínsecos de uma câmera (matriz de dimensão  $3 \times 3$ ).
- $\mathbf{P}$  é a matriz de projeção da câmera (matriz de dimensão  $3 \times 4$ ).
- $\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3$  são os vetores unitários que compõem uma matriz de rotação  $\mathbf{R}$  (vetores de dimensão 3).
- $\mathbf{R}$  é a matriz de rotação (matriz de dimensão  $3 \times 3$ ).
- $\mathbf{T}, \mathbf{t}$  são representações do vetor translação (vetor de dimensão 3).
- $[\mathbf{u} \ \mathbf{v}]^T$  é o vetor que representa um ponto expresso em coordenadas da imagem.
- $[\mathbf{X}_w \ \mathbf{Y}_w \ \mathbf{Z}_w]^T$  é o vetor que representa um ponto expresso em coordenadas do mundo.
- $\mathbf{H}$  é uma matriz de homografia e  $h_{ij}$  são os componentes na fila  $i$  coluna  $j$  dessa matriz (matriz de dimensão  $3 \times 3$ ).
- $(\mathbf{u}_d, \mathbf{v}_d)$  coordenadas de um ponto na imagem onde está presente a distorção gerada pela lente da câmera.
- $(\mathbf{u}_u, \mathbf{v}_u)$  coordenadas de um ponto na imagem onde foi corrigida a distorção gerada pela lente da câmera.
- $\overrightarrow{AB}$  vetor definido entre os pontos  $A$  e  $B$ .
- $\mathbf{F}$  é a matriz fundamental (matriz de dimensão  $3 \times 3$ ).
- $\mathbf{E}$  é a matriz essencial (matriz de dimensão  $3 \times 3$ ).
- $\mathcal{E}$  é uma variável que representa um valor de erro.
- $\mathbf{M}$  é um ponto no mundo expresso em coordenadas homogêneas (vetor de dimensão 4).
- $\mathbf{m}$  é um ponto na imagem expresso em coordenadas homogêneas (vetor de dimensão 3).

# 1 Introdução

Dentro da área de realidade virtual e aumentada, uma ferramenta chave tem sido o sistema de rastreamento. É através deste tipo de sistema que se desenvolvem as melhores interfaces de interação entre o usuário e o cenário virtual. Diversas tecnologias têm sido usadas para implementá-los [23]: sonora, eletro-magnética, mecânica, óptica e híbrida (Figura 1). Porém, a tecnologia óptica é a mais utilizada por causa da disponibilidade e do baixo custo dos equipamentos necessários para montar um sistema desse tipo, além da sua característica de não poluição do campo de rastreamento com objetos como cabos e outros componentes eletrônicos.

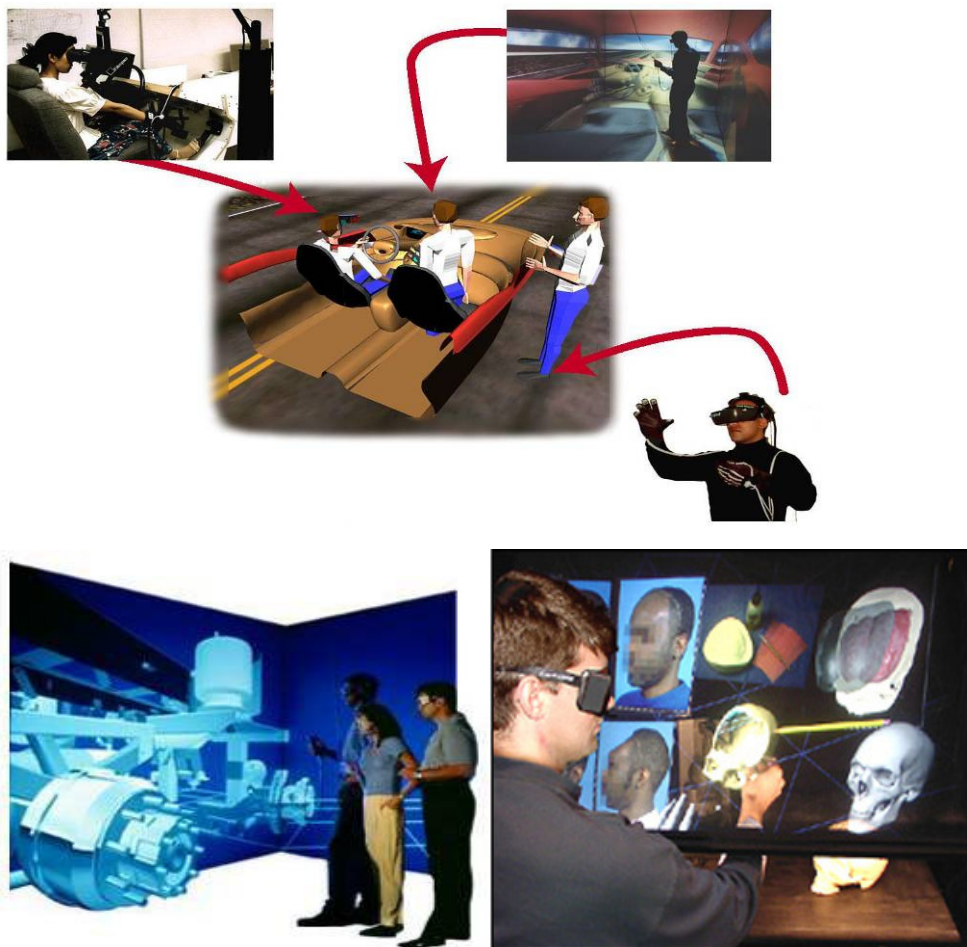


Figura 1: Aplicações de realidade virtual e aumentada.

(Fonte: <http://www.sgi.com/realitycenter/> e <http://www.vrl.umich.edu/intro/>)



Os sistemas de rastreamento óptico têm evoluído na direção de permitir o aumento do campo de rastreamento e, ao mesmo tempo, manter a precisão de captura e a recuperação da posição e da orientação 3D dos objetos rastreados. Para cumprir tal objetivo, em termos de equipamento, estes sistemas têm incrementado o número de câmeras que os compõem, com a intenção de capturar com nitidez um maior espaço real de rastreamento. Já na parte teórica, vêm-se aprofundado a pesquisa na direção de aprimorar cada vez mais o processo de calibração das câmeras. A otimização do processo de calibração permitiu melhorar a precisão do sistema, criando um vínculo direto entre esses dois aspectos: hardware e software. Por esse motivo, o processo de calibração é considerado o coração do funcionamento dos sistemas de rastreamento óptico, e a precisão alcançada no cálculo da posição 3D dos objetos rastreados está diretamente relacionada a uma correta e precisa calibração das câmeras em relação à área comum de rastreamento.

### **1.1. Motivação**

A motivação deste trabalho de tese surgiu de uma pesquisa que começou analisando os diversos componentes e processos que determinam o funcionamento de um sistema de rastreamento óptico (Figura 2). Nossa intenção é montar um sistema de múltiplas câmeras sobre o qual implementaremos um sistema de rastreamento óptico. Esse tipo de sistema é muito utilizado na implementação de aplicações de realidade virtual e aumentada, e seu objetivo é o cálculo e a extração da posição 3D de pontos ou marcadores específicos que aparecem dentro do campo de visão de pelo menos duas câmeras. Com essa informação é possível criar artefatos, formados por marcadores reconstruídos no espaço de rastreamento, com os quais modelaremos eventos de interação entre o usuário e a aplicação.

Um aspecto fundamental para a precisão dos sistemas de rastreamento óptico é o método de calibração aplicado sobre as múltiplas câmeras. Com base no resultado da calibração é feita a reconstrução precisa da posição 3D dos marcadores rastreados. Foi analisando diversos métodos descritos na bibliografia sobre o tema de múltipla calibração de câmeras que encontramos alguns aspectos que poderiam ser melhorados, foi assim que surgiu a proposta de implementação de um novo método de múltipla calibração de câmeras que seja mais robusto que

os atuais e gere resultados melhores que os dos métodos encontrados na bibliografia.

A compilação inicial das informações sobre o processo de calibração de câmeras nos permitiu também entender e modelar melhor as etapas e os processos nucleares que compõem esse processo, a partir de um ponto de vista global.



Figura 2: Sistemas de rastreamento óptico.

(Fonte: <http://blog.digitalcontentproducer.com/briefingroom/2009/01/28/uc-merced-adopts-vicon-f40-motion-capture-system/>)

Como produto dessa compilação, conseguimos determinar alguns aspectos em comum presentes na maioria dos métodos, corroborando trabalhos prévios sobre o tema como o descrito por [19], e que neste trabalho será estendido para analisar o caso de múltipla calibração de câmeras. Esses aspectos semelhantes são modelados e definidos em três etapas básicas comumente encontradas na maioria dos métodos de calibração, tanto individual [19] como de múltiplas câmeras:

1. Aquisição de dados para calibração, definido pelo uso de diferentes tipos de padrões e técnicas de processamento de imagem usados na captura e recompilação de dados de entrada para o sistema de calibração.
2. Calibração inicial, correspondente ao cálculo inicial dos parâmetros que serão usados para modelar as características internas, a posição e a orientação das câmeras que formam nosso sistema.
3. Otimização dos valores calculados inicialmente para os parâmetros de calibração das câmeras.

A análise das vantagens e desvantagens dos diversos componentes de cada uma das três etapas acima permitiu conceituar melhor a estrutura do novo método de múltipla calibração de câmeras que propomos neste trabalho. Como produto da análise desses componentes, já na primeira etapa, propomos um novo tipo de padrão de calibração que traria algumas vantagens em comparação a outros padrões semelhantes. A proposta desse novo padrão surgiu de uma pesquisa feita sobre a modelagem de artefatos usados pelos sistemas de rastreamento óptico como padrões de rastreamento para interação [16], baseado nesse estudo trouxemos a ideia deste novo tipo de padrão de calibração descrito como um padrão invariante a uma projeção 3D para 2D, e o seu formato é definido por um grupo de quatro marcadores colocados colinearmente [16].

Finalmente, a nossa motivação se definiu melhor pelo objetivo de criar e propor um novo método para múltipla calibração de câmeras que trouxesse como contribuição direta o uso desse tipo de padrão de calibração, que já na primeira etapa de aquisição de dados propõe ser melhor em comparação a outros padrões encontrados na bibliografia. Percebemos também que o formato desse tipo de padrão contém informações adicionais que poderiam ser muito bem exploradas por outro componente do processo de calibração, relacionado à terceira etapa do mesmo processo: a otimização dos parâmetros de calibração. Com essas duas contribuições em mente, este trabalho pretende descrever um novo método de múltipla calibração de câmeras que será descrito ao longo desta tese.

## 1.2. Objetivo

O objetivo deste trabalho é a criação e implementação de um novo método de múltipla calibração de câmeras, que seja:

- Flexível em relação ao número e ao tipo de câmeras usadas no sistema.
- Rápido e simples de executar por parte do usuário.
- Robusto para conseguir sempre uma calibração válida entre as câmeras.
- Preciso na recuperação do posicionamento e das medidas de objetos rastreados.

- Capaz de obter como produto da calibração final resultados comparáveis, e em alguns casos melhores, aos dos sistemas acadêmicos e comerciais descritos na bibliografia.

Para cumprir esse objetivo, o método desenvolvido terá como novidade principal a proposta de um novo tipo de padrão de calibração que, apesar da sua uma estrutura simples, tem muita informação implícita no seu formato que pode ser facilmente extraída e explorada na busca de melhores resultados nas diferentes etapas do processo de múltipla calibração de câmeras.

### **1.3. Contribuições**

O trabalho apresentado nesta tese de doutorado traz como principal contribuição a proposta de um novo método de múltipla calibração de câmeras. O método propõe o uso de um novo tipo de padrão de calibração, que tem características inerentes a seu formato que permitem dar maior robustez ao processo de captura e reconhecimento do padrão, e isso está diretamente relacionado à etapa de aquisição de dados para calibração. Essas características são também decisivas para alcançar bons resultados, melhores que os de outros métodos propostos na bibliografia sobre múltipla calibração de câmeras.

Para poder mostrar as vantagens e contribuições específicas que nossa técnica oferece, este trabalho definirá um framework conceitual que nos ajudará a descrever e dividir claramente as etapas e os processos em comum encontrados na descrição dos diversos métodos de múltipla calibração de câmeras, além de facilitar a descrição de nosso método e de outros métodos analisados no capítulo de trabalhos relacionados. O framework também apoiará a especificação das comparações apresentadas no capítulo sobre resultados, no qual serão apresentadas quantitativamente as vantagens e contribuições do nosso método.

Uma contribuição específica é a proposta de um formato diferente de padrão de calibração, importante parte do nosso método, cujas características físicas são bem exploradas. Trata-se de um padrão invariante a projeção 3D-2D e apresenta uma estrutura simples definida por quatro marcadores distribuídos colinearmente. Apesar da simplicidade do formato, esse padrão contém muita informação inerente à sua estrutura e que é muito bem explorada no nosso método. Nós podemos citar duas vantagens principais fornecidas pelo uso de nosso padrão.

Uma das principais vantagens do nosso padrão está relacionada com seu uso na primeira etapa do processo de múltipla calibração: a aquisição de dados. Nessa etapa, as contribuições são:

- A possibilidade de captura simultânea de mais de um padrão invariante no mesmo espaço de rastreamento.
- Maior robustez no processo de captura e reconhecimento do padrão por causa do formato e da teoria sobre características projetivas invariantes, que pode ser aplicada aproveitando o formato colinear do padrão.

Algumas vantagens diretas da captura de mais de um padrão simultaneamente no mesmo espaço de calibração são a captura de uma amostra válida de pontos de referência em um intervalo de tempo menor, assim como uma maior quantidade de pontos referenciais e a possibilidade de distribuir melhor nossa amostra de pontos tanto no espaço de calibração como no plano da imagem das diferentes câmeras utilizadas.

Outra vantagem do uso de nosso padrão colinear está relacionada às diversas informações que podem ser extraídas do seu formato. Essas informações nos permitem modelar várias funções de controle que serão usadas para montar nossa função de erro global, especialmente quando realizamos o processo de otimização final dos parâmetros de calibração previamente calculados de forma básica entre as câmeras do sistema.

#### **1.4. Organização da tese**

A estrutura desta tese é definida pelos seguintes capítulos.

No capítulo 2, será feita a descrição de alguns conceitos preliminares muito úteis para descrever e entender melhor trabalhos relacionados à nossa proposta, assim como para explicar a concepção e a implementação do novo método de múltipla calibração de câmeras apresentada como contribuição deste trabalho.

No capítulo 3, serão descritos detalhadamente alguns trabalhos relacionados ao tema de múltipla calibração de câmeras que foram selecionados por causa da inovação apresentada nas suas propostas, ajudando a descrever melhor a evolução do tema de múltipla calibração na linha do tempo. Esses trabalhos relacionados

servirão de base para modelar e comparar características específicas do processo de calibração em relação ao método proposto neste trabalho de tese.

No capítulo 4, serão expostas a estrutura e a implementação da nossa proposta, definida como um novo método de múltipla calibração de câmeras baseada em um padrão invariante. Nesse capítulo serão apresentados os detalhes da concepção e a descrição das contribuições de nosso método.

No capítulo 5, será relatada uma rotina de testes realizados usando um cenário real de múltipla calibração. Descreveremos o tipo de equipamento disponível para a execução dos testes e os resultados alcançados com a intenção de sustentar as contribuições de nosso método. Esses testes ajudarão também a detalhar melhor alguns aspectos diretamente relacionados à implementação do método.

Finalizando, no capítulo 6, serão apresentadas conclusões e propostas de trabalhos futuros.

## 2 Múltipla calibração de câmeras

Neste capítulo descreveremos alguns conceitos relacionados ao processo de múltipla calibração de câmeras. Para organizar esses conceitos, propomos um framework conceitual que ajudará a esquematizar alguns aspectos e etapas comumente encontrados em diversos métodos de calibração, tanto para o caso de uma câmera como para o de múltiplas câmeras.

Para modelar o funcionamento de uma câmera, usaremos o clássico modelo de câmera Pinhole. No apêndice A, encontra-se a descrição detalhada dos diversos componentes e parâmetros usados para modelar teoricamente o processo de captura óptica, que é realizado internamente numa câmera real.

Define-se como calibração de câmera o processo realizado para calcular o valor dos parâmetros intrínsecos, extrínsecos e coeficientes de distorção da lente, que são comumente usados para modelar o funcionamento de uma câmera real. Logo, considera-se como múltipla calibração de câmeras o processo que permite a calibração de várias câmeras simultaneamente.

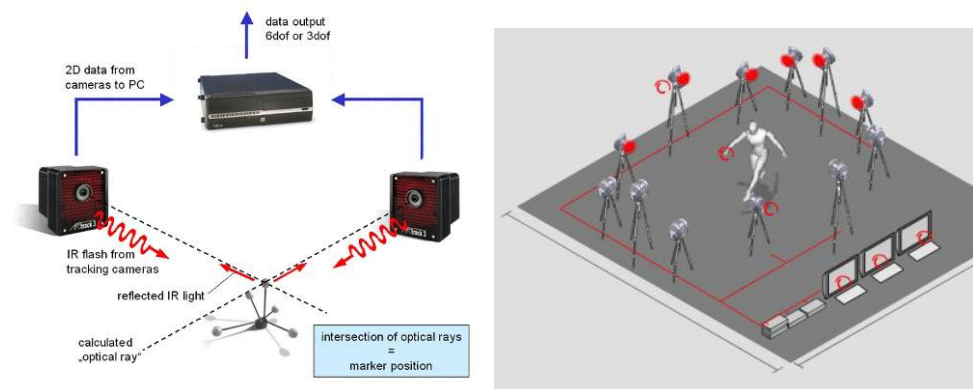


Figura 3: Sistema estéreo e de múltiplas câmeras.

(Fontes: <http://www.ar-tracking.de/General.144.0.html> e <http://www.digital-humans.org/Report2004/Documents/06-VirtualRealityForSantos.htm>)

O menor número de câmeras necessário para um sistema ser considerado de múltipla calibração é dois. Esse tipo de sistema é mais conhecido como sistema de câmeras em estéreo. Já o sistema composto por mais de duas câmeras ganha o nome de sistema de múltiplas câmeras (Figura 3).

Diferentes métodos têm sido definidos para afrontar esse tipo de problema, mas analisando a maioria deles, podemos estabelecer um framework conceitual que ajudará a descrever um conjunto de conceitos usado para resolver o problema específico da múltipla calibração de câmeras. Esse framework conceitual é composto de três etapas, as quais também têm sido definidas num estudo realizado sobre diferentes métodos de calibração de câmera individual, descrito no trabalho de Marques [19]. Nós adaptaremos e estenderemos a definição das três etapas, para um cenário onde analisaremos métodos de múltipla calibração de câmeras. Finalmente as etapas que compõem nosso framework conceitual são as seguintes:

1. Aquisição de dados para a calibração do sistema de múltiplas câmeras.
2. Calibração inicial, correspondente ao cálculo inicial dos parâmetros que usaremos para modelar as câmeras reais.
3. Processo de otimização dos valores calculados inicialmente para os parâmetros de calibração das câmeras.

A seguir, faremos uma descrição mais detalhada dos conceitos e processos de cada etapa do nosso framework conceitual.

### **2.1. Primeira etapa: aquisição de dados para calibração**

A primeira etapa do processo de múltipla calibração envolve diferentes métodos e técnicas utilizados na captura de dados de entrada que serão usados no posterior cálculo dos parâmetros que modelam teoricamente o processo de calibração das câmeras.

Esses métodos e técnicas são comumente compostos por diferentes algoritmos da área de processamento de imagem e reconhecimento de padrões, que, em conjunto com o uso e a definição de certos tipos de padrões sintéticos, ajudam a realizar a captura e o reconhecimento preciso de diferentes pontos de referência, conhecidos como marcadores, que são definidos pela estrutura de cada padrão.

Os algoritmos de processamento de imagem são utilizados com o objetivo de ressaltar características próprias do padrão, o que permite isolá-lo do resto de objetos presentes na mesma imagem. Depois, os algoritmos de reconhecimento de



padrão se encarregam de extrair e identificar individualmente os marcadores que um padrão contém.

O objetivo da utilização de um padrão é a captura de uma amostra relativamente grande de pontos de referência corretamente reconhecidos e distribuídos na maior parte do espaço de visão das câmeras, definido como espaço de rastreamento.

O formato dos diferentes tipos de padrões propostos na bibliografia sobre calibração de câmeras tenta cumprir alguns requisitos:

- O formato proposto deve estar definido por uma estrutura que permite sua nítida diferenciação em relação a outros objetos semelhantes que possam estar presentes no mesmo espaço de rastreamento. Isso ajuda a evitar possíveis erros, como falsos positivos no reconhecimento do padrão.
- O formato também deve permitir a sua fácil manipulação e movimentação no espaço de rastreamento. Isso permitirá que o campo visual enxergado pelas câmeras esteja corretamente preenchido pela captura de várias vistas válidas do padrão, o que representará um maior número de pontos de referência corretamente identificados e espalhados tanto no campo visual das câmeras como no espaço de rastreamento físico.
- A quantidade e a distribuição dos pontos de referência contidos no padrão devem permitir a recuperação de múltiplas informações em relação aos mesmos, como a identificação individual de cada ponto de referência em relação ao resto dos pontos, a identificação grupal de todos os pontos de referência que determinam o formato único do padrão, além de distâncias, ângulos e outras medidas que possam ser recuperadas a partir do conhecimento prévio da distribuição tridimensional desses pontos de referência que formam o padrão.

À medida que o formato de um padrão passa a conter muitos pontos de referência e detalhes na sua estrutura, também se encarece diretamente o processo de reconhecimento do padrão e, algumas vezes, a sua manipulação e movimentação em frente às câmeras torna-se limitada.

A falta de liberdade para manipular e movimentar o padrão diante das diferentes câmeras diminui a quantidade de vistas válidas do padrão capturadas

simultaneamente no campo visual compartilhado pelas várias câmeras que compõem o sistema. A observação acima será melhor exemplificada e entendida na descrição sobre diferentes tipos de padrão na seção seguinte.

### **2.1.1. Padrões usados na calibração de câmeras**

Os padrões de calibração têm como objetivo permitir a captura de pontos de referência a partir da detecção e do reconhecimento de algumas características implícitas do seu formato.

A amostra de pontos de referência extraída a partir da análise e do processamento de vistas capturadas do padrão é usada no processo de calibração tanto na etapa de calibração inicial como numa etapa final de otimização do valor dos parâmetros obtidos pela calibração inicial.

A captura de uma amostra de pontos de referência tem sido a principal ferramenta usada no processo de calibração de câmeras. O uso de padrões contendo esses pontos de referência tem sido estudado desde os métodos desenvolvidos e apresentados por [11][35][36] e [40], onde o padrão de calibração usado tem um formato planar e contém na sua superfície pontos de referência definidos por marcadores explícitos ou implícitos desenhados e/ou inseridos na superfície do padrão.

Outros padrões propostos para casos de estéreo e múltipla calibração de câmeras surgiram na última década, dentro de um grupo considerado como padrões 1D e muitas vezes usados em abordagens de auto-calibração de câmera descrita em trabalhos apresentados por [4] e [32]. Esses recentes trabalhos mostram a nova tendência de auto-calibração de câmeras, e têm como característica em comum a proposta de padrões mais simples, compostos por um ou dois marcadores, que trazem como principal vantagem a fácil manipulação, movimentação e reconhecimento nas diversas vistas capturadas do padrão dentro do espaço de rastreamento.

Também podemos citar outro tipo de padrão conhecido como padrão tridimensional composto por um conjunto de dois ou três planos colocados perpendicularmente um em relação ao outro, este tipo de padrão é pouco utilizado para a calibração de múltiplas câmeras principalmente pelos problemas de oclusão que pode gerar os planos que o conformam quando mais de uma câmera tenta

enxergar simultaneamente os marcadores inseridos na sua estrutura, por esse motivo não é considerado na nossa análise sobre múltipla calibração de câmera.

Tanto os padrões planares como os novos padrões 1D serão descritos a seguir, quando detalharemos um pouco mais as vantagens e desvantagens dos mesmos, tendo sempre como foco a calibração feita em duas ou mais câmeras simultaneamente.

### 2.1.2. Padrão planar – 2D

Dentro dessa tipificação de padrão, iremos considerar como padrão planar aquele cujos pontos de referência se encontram distribuídos sobre uma superfície plana 2D. Trabalhos apresentados por [11][35][36] e [40] mostram variações desse tipo de padrão, como os apresentados na Figura 4.

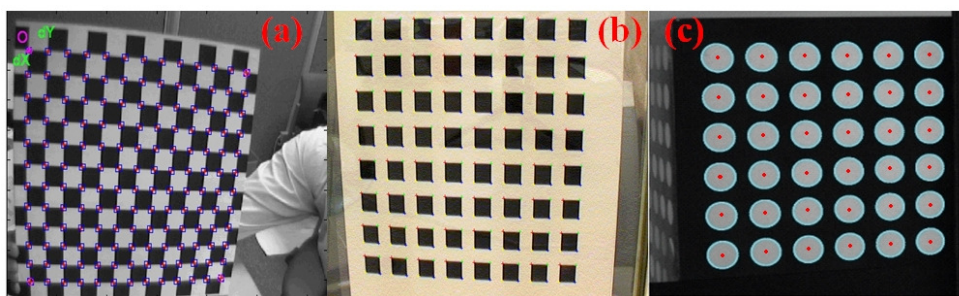


Figura 4: Padrões planares usados para calibração de câmera em trabalhos como, (a) Bouquet [5], (b) Zhang [40] e Tsai [35][36] e (c) Heikkila [11].

Como podemos observar, o aspecto dos padrões que define os pontos de referência varia em cada um dos diferentes trabalhos. Por exemplo, no trabalho de [35][36] e [40], o padrão é definido como um conjunto de quadrados pretos alinhados em um determinado número de filas e colunas, e os pontos de referência são determinados pelos cantos dos quadrados pretos impressos sobre um fundo branco. Essa arquitetura permite a captura de 4 pontos de referência por cada quadrado corretamente reconhecido. Já no trabalho de [11], a forma dos desenhos no padrão muda de um quadrado preto para círculos brancos num fundo preto, e os pontos de referência são definidos pelos centróides dos círculos brancos. Isso permite a captura de um ponto de referência por cada área circular corretamente reconhecida.

Ambos modelos de padrão planar compartilham a característica do formato em grade ou grelha, que permite a captura de um número de pontos de referências abundante por vista de padrão corretamente reconhecido e correlacionado.

Esse tipo de padrão planar também já foi utilizado na calibração de sistemas estéreo e de múltipla calibração de câmeras, como descrito em trabalhos realizados por [5] (estéreo) e [25] (múltiplas câmeras). Na figura 5, mostramos o tipo de padrão usado em cada trabalho.

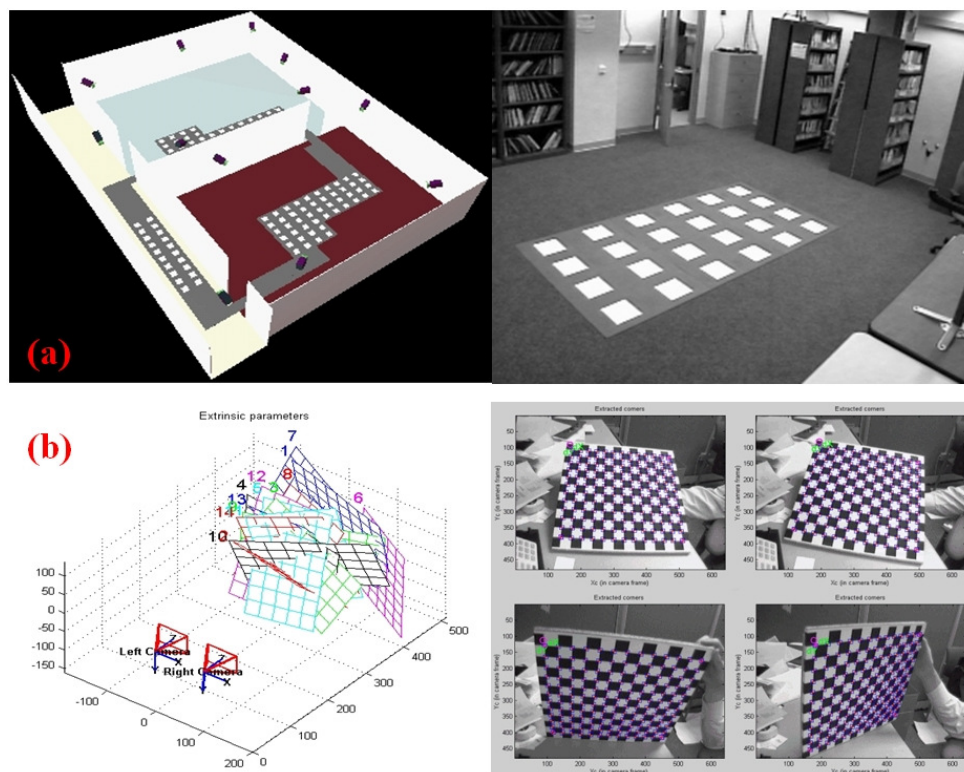


Figura 5: Padrões planares usados em (a) Olsen e Hoover [25] e (b) Bouguet [5].

Apesar de esse tipo de padrão ser muito usado e conter muita informação em relação aos pontos de referência que podem ser recuperados por cada vista do padrão corretamente reconhecido, ele tem como principal desvantagem a complexidade do processamento da imagem e do reconhecimento do padrão requerido para poder ser identificado corretamente. Esse problema surge por causa de limitações em relação à distância e diferentes orientações em que uma vista do padrão pode ser capturada, limitando sua livre movimentação diante de uma ou várias câmeras do sistema.

Algumas restrições sobre a movimentação desse tipo de padrão já foram reconhecidas na implementação dos métodos de calibração individual. Tratam-se

principalmente de restrições na orientação do padrão planar que, em ângulos menores que  $45^\circ$  e distâncias muito longas da câmera, já apresentava perda no reconhecimento acertado. Essas restrições ficaram ainda mais notórias e restritivas quando esse tipo de padrão foi estendido para uso em sistemas com mais de uma câmera. Nesse novo ambiente de calibração, a interseção dos campos de visão das câmeras limitou muito mais as orientações mínimas nas quais o padrão pode ficar quando ele está sendo capturado pelas câmeras.

O trabalho de [5] mostra uma extensão no uso desse tipo de padrão para a calibração de um sistema estéreo. Nesse ambiente com duas câmeras, as limitações na movimentação e na orientação máxima de  $45^\circ$  do padrão ficam ainda mais notórias quando o ângulo entre as câmeras fica menor que  $180^\circ$  (Figura 5b). Já no trabalho de [25], o uso do padrão planar tenta contornar os problemas de movimentação com a geração de um padrão fracionado e de grandes dimensões que lhe permitem ser visível por várias câmeras bem distantes e com pouca interseção entre os campos de visão das câmeras (Figura 5a).

Em relação à abordagem apresentada por [5], essa última tem como principal diferença o número de pontos de referência que podem ser capturados. O método definido por [5] permite a captura de uma amostra muito maior de pontos de referência que o de [25]; em contrapartida, o método de [25] tenta quebrar a restrição do posicionamento e do uso de só duas câmeras imposta no trabalho de [5], e ampliá-la para suportar múltiplas câmeras distribuídas em um ambiente maior de rastreamento.

Por fim, as abordagens dos métodos acima têm como desvantagem comum a complexidade na aquisição dos pontos de referência. Devido a essa limitação foram criados novos tipos de padrões. A seguir serão descritas estas novas propostas de padrão com formato 1D e adimensional.

### **2.1.3. Padrão unidimensional – 1D**

O padrão considerado unidimensional (1D) foi proposto e descrito inicialmente no trabalho publicado por [4]. Ele é composto por dois marcadores esféricos colocados nos extremos de uma vara metálica (Figura 6). O trabalho de [4] propôs esse padrão como parte de um novo método de calibração de um

sistema de câmeras em estéreo. Esse método pode ser considerado já um método de auto-calibração, cujo conceito é descrito na seção 2.2.2.4.

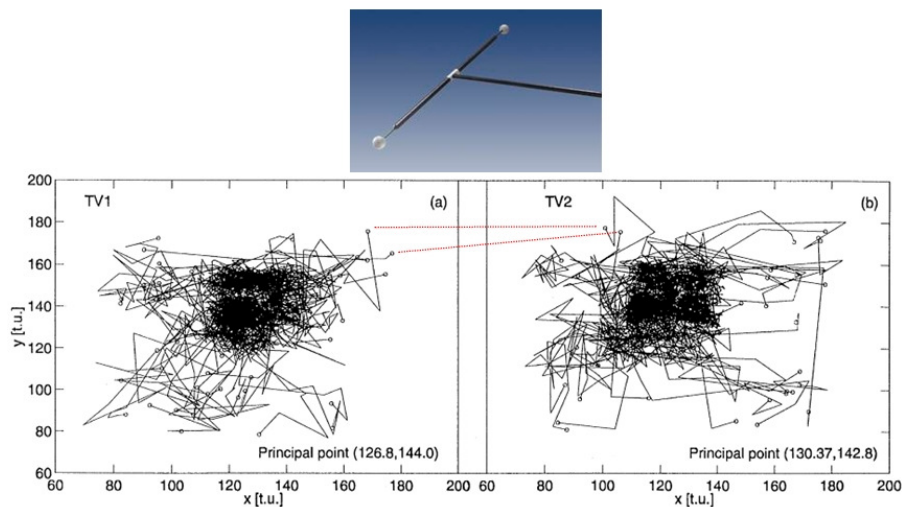


Figura 6: Padrão de calibração e captura proposto no trabalho de Borghese[4].

Uns anos mais tarde esse método foi estendido e usado nos trabalhos de [14] [22] e [37] para a calibração de um sistema formado por múltiplas câmeras, no qual se manteve o mesmo conceito do padrão 1D e só se mudou o procedimento de captura e reconhecimento.

Esse tipo de padrão 1D tem sido também muito utilizado como padrão de calibração em sistemas de rastreamento óptico comerciais como [1] e [39]. As principais vantagens desse tipo de padrão traz podem ser resumidas em:

- Liberdade na movimentação e manipulação do padrão.
- Fácil captura e visualização pelas diferentes câmeras dos sistemas
- Nenhuma restrição enquanto ao posicionamento das câmeras em relação a certas posições ou orientação do padrão.
- O processamento requerido para a captura e extração do padrão a partir das vistas capturadas pelas múltiplas câmeras é mais simples se comparado ao dos padrões planares.

Em relação à liberdade de movimentação e manipulação, esse novo formato permite que um usuário comum não tenha muitas restrições quando ao realizar o processo de captura de uma amostra de vistas do padrão, que é feito normalmente movimentando-o por toda a área de rastreamento. Essa simplicidade permite que o padrão seja colocado em áreas do espaço de rastreamento bem complicadas de capturar nos distintos campos de visão das câmeras.

Finalmente, o processamento da imagem e a extração das posições 2D dos marcadores que compõem esse padrão são bem flexíveis e menos complexos que o requerido para os padrões planares. Normalmente o processamento da imagem fica bem caracterizado pelo objetivo de encontrar alguma estratégia que ressalte os marcadores da vara em relação a outros objetos ou ao fundo do espaço de rastreamento. Um exemplo de duas adaptações são os trabalhos de [22] e [37] (Figura 7). No primeiro, a solução foi montar a vara com marcadores coloridos. As cores dos marcadores são facilmente distinguíveis em relação ao fundo do espaço de rastreamento, e só a aplicação de um algoritmo de *threshold* que use o valor da cor RGB dos marcadores é a única característica necessária para extraí-los de uma imagem capturada pelas câmeras. Já no trabalho de [37] segue uma estratégia mais próxima à usada em sistemas de rastreamento comerciais: usar iluminação infravermelha no espaço de rastreamento, junto a câmeras sensíveis a esse tipo de luz e marcadores revestidos com tecido retrorreflexivo. Essa arquitetura permite que os marcadores apareçam nas imagens das câmeras como os objetos mais brilhantes. Novamente, aplica-se um algoritmo de *threshold*, nesse caso, com um certo valor de brilho alto, perto da cor branca numa escala de tons de cinza, para extrair facilmente os marcadores nas imagens capturadas. Na figura 7 podemos observar padrões do tipo 1D utilizados tanto nos trabalhos de [37][22].

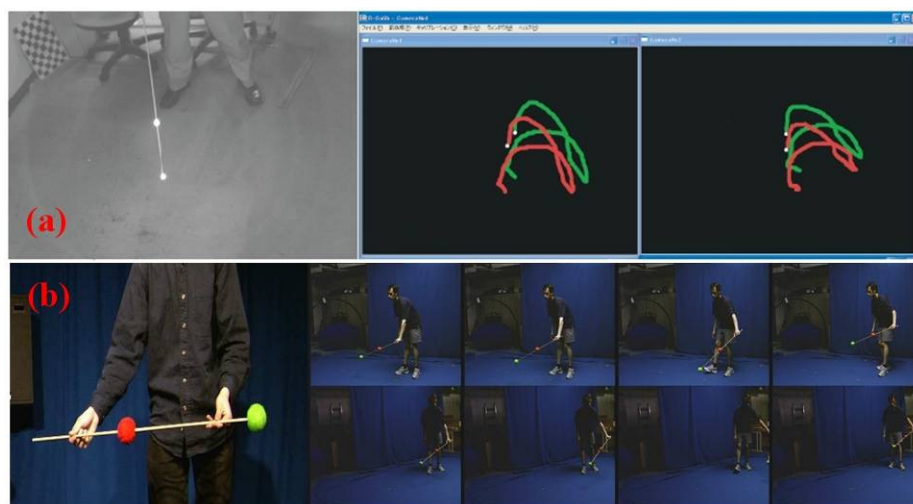


Figura 7: (a) Ambiente com uso de marcadores retrorreflexivos em Uematsu et al.[37], (b) marcadores coloridos usados como padrão em Mitchelson et al.[22].

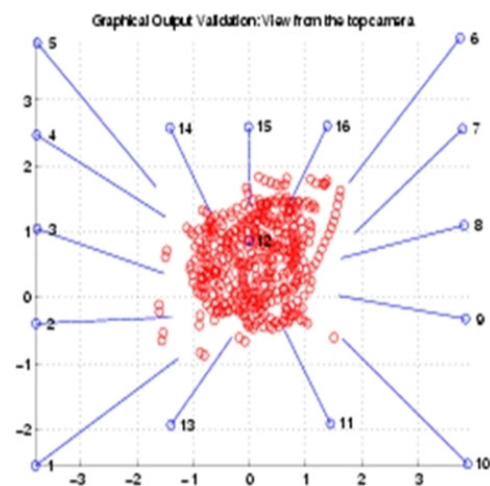
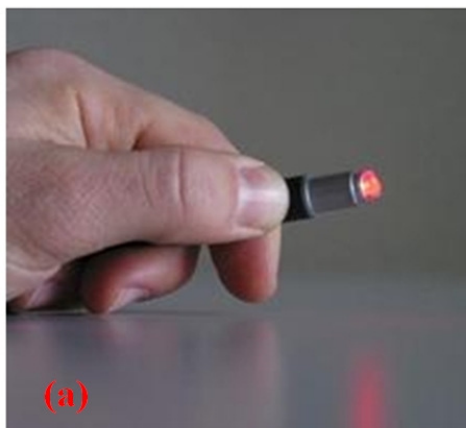
### 2.1.4. Padrão adimensional

O padrão considerado adimensional (0D) foi proposto e descrito recentemente no trabalho publicado por [32], que propôs um novo método de calibração de um sistema de múltiplas câmeras.

Esse padrão adimensional traz como vantagens as mesmas do padrão 1D: liberdade na movimentação do padrão, fácil visualização e processamento muito leve na captura e extração do padrão a partir das vistas capturadas pelas múltiplas câmeras. A principal diferença e a vantagem adicional é estar composto por um único marcador, o que o faz mais simples quanto ao seu desenho e às restrições para seu correto reconhecimento.

Uma versão modificada e mais restritiva do método apresentado por [32] tem sido usado por [28] na implementação de uma versão acadêmica de um sistema de rastreamento óptico. Esse sistema evoluiu para uma versão comercial chamado IOTracker [13].

Na figura 8, são mostradas algumas imagens que exemplificam a captura deste padrão e a posterior reconstrução 3D. No caso do trabalho publicado em [32] mostra-se como esse tipo de padrão foi implementado a partir de um ponteiro laser modificado (Figura 8a). Já nas propostas de [28] e [13] o padrão foi modelado como um marcador esférico coberto por tecido retrorreflexivo (Figura 8b).





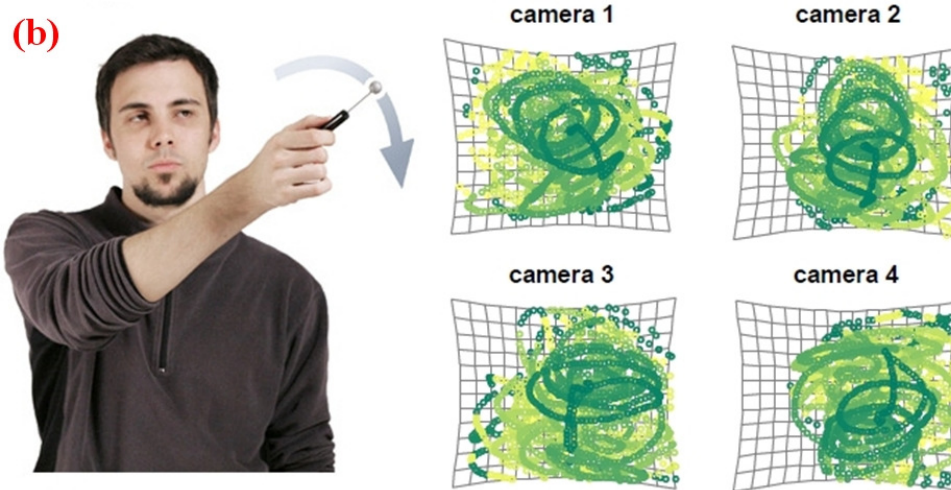


Figura 8: Marcadores adimensionais propostos por (a) Svoboda et al. [32], (b) Pintaric et al. [28] e [13].

Para poder achar os centróides desse tipo de marcador, o primeiro passo na captura do padrão é estabelecer uma estratégia de contraste que permita distinguir os marcadores dos outros objetos presentes no espaço de rastreamento. Por exemplo, para a calibração das câmeras no caso do padrão definido por [37] e os software comerciais [1] e [39] foi necessário montar uma infra-estrutura que usa:

- Iluminação de fontes de luz infravermelha invisível na faixa dos 850nm aproximadamente.
- Marcadores cobertos com fita ou tecido retroreflexivo.
- Câmeras com alta sensibilidade à luz infravermelha.

Essa infraestrutura permite que os marcadores sejam facilmente distinguidos nas imagens capturadas pelas câmeras por causa do brilho que eles apresentam em relação aos outros objetos no espaço de captura da amostra. No caso do trabalho apresentado por [32], esse contraste é conseguido pelo alto brilho que o ponteiro laser gera. Para isso tenta-se deixar esse espaço o mais escuro possível.

Uma vez criado o contraste, os próximos passos comumente definidos para extrair e identificar os marcadores nas imagens são:

- Conversão da imagem capturada para um formato em tons de cinzas.
- Aplicação de um filtro de *threshold* sobre a imagem em tons de cinza com a finalidade de binarizar e ressaltar mais ainda os marcadores que aparecem na imagem.

- Uma vez binarizada a imagem, os marcadores devem aparecer como áreas circulares dentro da imagem. É sobre essas áreas que se tenta calcular o centróide dos pixels.
- Após ter achado o centróide dessas áreas brancas, a coordenada 2D do centróide é considerada nosso ponto de referência relacionado à imagem do marcador correspondente.

O processamento da imagem descrito acima é feito para cada quadro capturado em cada câmera durante a movimentação do padrão 1D ou adimensional.

Tanto no caso do padrão unidimensional como no do adimensional, a simplicidade do processamento das imagens e a facilidade de movimentação tornam possível a captura de uma amostra de pontos de referência muito bem espalhada dentro da área de interseção dos campos de visão das câmeras, assim como a captura de um número grande de pontos de referência. Em contraposição, essa simplicidade do formato pode acarretar alguns problemas posteriores que podem desencadear em erros no processo de calibração das câmeras.

O principal problema desse tipo de padrão é a fácil contaminação da amostra de pontos de referência por causa de ruído ou de outros elementos erroneamente reconhecidos como possíveis marcadores válidos. Esses problemas são descritos nos trabalhos de [4] e [32], e algumas soluções e estratégias são propostas para descartar esses ruídos que podem gerar a identificação de falsos marcadores dentro das imagens processadas. As ferramentas mais usadas para descartar esses falsos marcadores variam desde o uso de algoritmos de predição do movimento até rotinas de correspondência só utilizadas quando é computada a matriz fundamental, que é definida na seção 2.2.2.4.

Nas soluções antes citadas, também se observam os diferentes momentos no processo de calibração nos quais elas podem ser aplicadas. Por exemplo, o uso de um algoritmo de predição de movimento seria considerado uma medida preventiva que pode ser utilizada logo ao início do processo de calibração, isto é, na etapa de captura do padrão, sem se estender para uma etapa posterior. Já a solução do uso da teoria por trás do cálculo da matriz fundamental para encontrar falsos marcadores é aplicada após a amostragem ter sido capturada. Essa solução é um pouco mais corretiva que preventiva e traz como ponto negativo a necessidade

de esperar até esse estágio avançado no processo de calibração para detectar os possíveis erros vindos da etapa de captura da amostra.

## **2.2. Segunda etapa: calibração inicial**

A segunda etapa tem como objetivo agrupar e descrever os diferentes métodos utilizados para calcular o valor inicial dos parâmetros intrínsecos, extrínsecos e da distorção das lentes das múltiplas câmeras que compõem nosso sistema. O processo do cálculo dos parâmetros das câmeras é definido como calibração de câmera.

Em geral, as diferentes técnicas de calibração de câmera podem ser agrupadas em duas categorias [7][8]: calibração fotogramétrica e auto-calibração.

Na implementação dos diversos métodos desenvolvidos para múltipla calibração de câmeras é muito comum dividir a tarefa de calibração de certos parâmetros usando métodos das duas categorias.

A seguir, definiremos essas duas categorias e as complementaremos com a descrição de dois algoritmos que pertencem a cada uma delas. Esses algoritmos são utilizados na implementação do método de múltipla calibração de câmeras proposto nesta tese.

### **2.2.1. Calibração fotogramétrica**

Neste tipo de método calculam-se os parâmetros de calibração de uma câmera usando a informação fornecida por um padrão cuja geometria é conhecida com precisão no espaço tridimensional onde ele é posicionado [7][8]. Um exemplo clássico desse tipo de método é o proposto por Tsai [35][36], que usa um padrão planar montado num sistema de trilhos que movimenta o padrão com precisão em frente à câmera. Outro trabalho semelhante, porém mais flexível, é proposto por Zhang [40], que também utiliza um padrão planar movimentado livremente diante da câmera.

Com base nas informações fornecidas pelos pontos referenciais colocados sobre o padrão planar, encontra-se uma correspondência direta entre a posição 3D dos marcadores nesse padrão e os pontos 2D correspondentes extraídos das imagens capturadas do padrão.

Este tipo de método gera resultados muito precisos, mas na prática, é pouco utilizado por causa do alto custo requerido para montar um padrão planar que ocupe o maior campo de visão da câmera, além de movimentá-lo, capturá-lo e reconhecê-lo com precisão em cada vista capturada do padrão. Outro inconveniente é o tempo despendido na captura de uma amostra de vistas válidas do padrão.

Este método está mais relacionado ao processo de calibração individual de câmera, e é muito utilizado especialmente para calcular os parâmetros intrínsecos e os coeficientes de distorção da lente. Os parâmetros extrínsecos também podem ser recuperados, porém, não são únicos e um valor diferente é calculado para cada uma das vistas do padrão planar capturadas e posteriormente processadas. Isso acontece porque os parâmetros extrínsecos calculam a posição e a orientação da câmera em relação a cada posição em que o padrão planar é capturado. Será descrito a seguir o método de calibração de câmera definido por Zhang [40].

#### **2.2.1.1. Calibração baseada no cálculo de homografias**

Para descrever esse tipo específico de calibração de câmera, incluído na categoria de métodos fotogramétricos, usaremos o modelo descrito no trabalho de Zhang [40], definido como: um método flexível para calibração de câmera que requer como dados de entrada vistas capturadas de um padrão planar que será movimentado em frente à câmera.

A calibração nesse método estima uma homógrafa  $H$  entre cada vista capturada de um padrão planar e a imagem projetada dele no plano da imagem da câmera. É assim que múltiplas homografias calculadas fornecem restrições suficientes para extrair os diferentes parâmetros da câmera.

O primeiro passo será calcular de forma robusta as homografias para cada vista do padrão capturada. Em seguida continuaremos com a extração dos parâmetros intrínsecos (únicos para a câmera) e extrínsecos contidos em cada homografia e finalmente faremos o cálculo dos coeficientes de distorção da lente.

#### **2.2.1.2. Cálculo da homografia**

Como foi apontado na seção 2.1.2, um padrão planar fornece um conjunto de pontos de referência distribuídos num formato de grade sobre sua superfície.

Sem perda de generalidade, podemos assumir que esse plano determina nosso sistema de coordenadas do mundo (SCM) e que cada ponto sobre este plano tem sua coordenada  $Z = 0$ . Com base nessa restrição, a matriz de projeção é definida na equação 1:

$$\begin{aligned} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} &\cong K \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \Rightarrow \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \cong P \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \\ \Rightarrow \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} &\cong [p_1 \quad p_2 \quad p_3 \quad p_4] \begin{bmatrix} X_w \\ Y_w \\ 0 \\ 1 \end{bmatrix} \cong [p_1 \quad p_2 \quad p_4] \begin{bmatrix} X_w \\ Y_w \\ 1 \end{bmatrix} \end{aligned} \quad (1)$$

Como podemos observar uma das colunas da matriz de projeção da câmera  $P$  é anulada devido à coordenada  $Z = 0$ . Se nós definíssemos a equação 1 mais explicitamente, determinando quais componentes dos parâmetros das câmeras estariam definidos pelas colunas  $p_1, p_2, p_4$  na equação 1, teríamos a seguinte equação:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \cong K [r_1 \quad r_2 \quad t] \begin{bmatrix} X_w \\ Y_w \\ 1 \end{bmatrix} \quad (2)$$

Na equação 2,  $r_1$  e  $r_2$  são as duas primeiras colunas da matriz de rotação  $R$ ,  $t$  é o vetor de translação que definem os parâmetros extrínsecos da câmera, e a matriz  $K$  determina os parâmetros intrínsecos da câmera.

Então, a homografia  $H$  achada para cada vista do nosso padrão planar é determinada pelo produto  $K [r_1, r_2, t]$  mais um fator de escala. A homografia permite relacionar um ponto de referência  $m$  detectado numa imagem capturada e processada do padrão planar, com seu ponto correspondente  $M$ , definido em coordenadas do mundo  $(X_w, Y_w, 1)$  definido no modelo real do padrão planar. Essa correspondência é exemplificada na equação 3.

$$\begin{aligned} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} &\cong K [r_1 \quad r_2 \quad t] \begin{bmatrix} X_w \\ Y_w \\ 1 \end{bmatrix} \Rightarrow \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \cong H \begin{bmatrix} X_w \\ Y_w \\ 1 \end{bmatrix} \Rightarrow m \cong H M \\ H &= \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \end{aligned} \quad (3)$$

Baseados na relação descrita na equação 3, nós podemos extrair duas equações para cada ponto correspondente  $m_i \leftrightarrow M_i$ :

$$\begin{aligned} u(h_{31}X_w + h_{32}Y_w + h_{33}) &= h_{11}X_w + h_{12}Y_w + h_{13} \\ v(h_{31}X_w + h_{32}Y_w + h_{33}) &= h_{21}X_w + h_{22}Y_w + h_{23} \end{aligned} \quad (4)$$

Com essas duas equações relacionando os pontos  $m_i$  e  $M_i$  em cada vista do padrão planar capturado, podemos montar um sistema de equações que nos permitirá calcular os componentes  $h_{ij}$  da nossa matriz de homografia  $H$ . O sistema é definido na equação 5.

$$\begin{bmatrix} X_w & Y_w & 1 & 0 & 0 & 0 & -uX_w & -uY_w & -u \\ 0 & 0 & 0 & X_w & Y_w & 1 & -vX_w & -vY_w & -v \end{bmatrix} h = 0 \quad (5)$$

$$h = [h_{11} \quad h_{12} \quad h_{13} \quad h_{21} \quad h_{22} \quad h_{23} \quad h_{31} \quad h_{32} \quad h_{33}]^T$$

Se tivermos em nosso padrão  $n$  pontos correspondentes  $m_i \leftrightarrow M_i$ , poderemos formar um sistema de equações definido pela multiplicação da matriz  $L$ , de dimensão  $2n \times 9$ , e o vetor  $h$ , que é formado pelos componentes da nossa matriz de homografia  $H$ . Esse sistema de equações é definido por  $Lh = 0$  e tem como solução o autovetor associado ao menor autovalor da matriz  $L^T L$ . Para calcular a solução para o sistema  $Lh = 0$  aplicaremos o método chamado de “decomposição em valores singulares” (SVD - *singular value decomposition*). Devido às diferentes escalas dos componentes da matriz  $L$ , esta fica numericamente mal condicionada, e é recomendável fazer uma normalização dos dados  $m_i$  e  $M_i$  antes de montar nosso sistema de equações.

Por fim, uma vez aplicado o método de SVD, determinaremos todos os componentes  $h_{ij}$  que definem nossa matriz de homografia  $H$ , lembrando que esse cálculo da homografia  $H$  é feito para cada vista do padrão planar capturada, o que produzirá ao final,  $n$  matrizes de homografia.

### 2.2.1.3. Determinação dos parâmetros intrínsecos e extrínsecos

Uma vez calculadas as  $n$  homografias, correspondentes às  $n$  vistas válidas capturadas de nosso padrão planar, continuaremos com a extração da matriz de parâmetros intrínsecos  $K$  baseada na análise das múltiplas homografias calculadas.

Para realizar a extração dessa matriz  $K$ , utilizaremos o conceito sobre cônica absoluta, cuja descrição pode ser encontrada no livro de Hartley e Zisserman [10], nele a cônica absoluta é definida como:

$$K^{-T}K^{-1} = \begin{bmatrix} \frac{1}{f_x^2} & -\frac{s}{f_x^2 f_y} & \frac{sv_0 - u_0 f_y}{f_x^2 f_y} \\ -\frac{s}{f_x^2 f_y} & \frac{s^2}{f_x^2 f_y^2} + \frac{1}{f_y^2} & -\frac{s(sv_0 - u_0 f_y)}{f_x^2 f_y^2} - \frac{v_0}{f_y^2} \\ \frac{sv_0 - u_0 f_y}{f_x^2 f_y} & -\frac{s(sv_0 - u_0 f_y)}{f_x^2 f_y^2} - \frac{v_0}{f_y^2} & \frac{(sv_0 - u_0 f_y)^2}{f_x^2 f_y^2} + \frac{v_0^2}{f_y^2} + 1 \end{bmatrix} \quad (6)$$

A matriz  $K$  contém seis incógnitas definidas pelos parâmetros intrínsecos  $(f_x, f_y, s, u_0, v_0)$  mais um fator de escala  $\lambda$ . Já a cônica absoluta é definida como a matriz simétrica  $K^{-T}K^{-1}$  contendo seis entradas. Como pode ser observado na equação 6, as incógnitas da matriz  $K$  determinam as entradas da matriz  $K^{-T}K^{-1}$ , que por sua vez é a definição da cônica absoluta. Essa relação permite calcular os componentes da matriz  $K^{-T}K^{-1}$  e extrair os parâmetros intrínsecos a partir desses componentes.

A imagem de uma cônica absoluta é obtida a partir das seguintes propriedades baseadas na definição da homografia  $H$ :

$$H = \lambda K \begin{bmatrix} r_1 & r_2 & t \end{bmatrix} \quad (7)$$

As propriedades são determinadas pela multiplicação de  $HH^T$ , no qual  $\lambda$  é o fator de escala, assim temos:

$$\begin{aligned} H^T H &= \lambda^2 (K \begin{bmatrix} r_1 & r_2 & t \end{bmatrix})^T K \begin{bmatrix} r_1 & r_2 & t \end{bmatrix} \\ H^T H &= \lambda^2 \begin{bmatrix} r_1 & r_2 & t \end{bmatrix}^T K^T K \begin{bmatrix} r_1 & r_2 & t \end{bmatrix} \\ H^T K^{-T} K^{-1} H &= \lambda^2 \begin{bmatrix} r_1 & r_2 & t \end{bmatrix}^T \begin{bmatrix} r_1 & r_2 & t \end{bmatrix} \end{aligned} \quad (8)$$

Considerando o lado direito da equação 8, determinamos duas propriedades a partir da ortonormalidade da matriz de rotação definidas pelas suas colunas  $r_1$  e  $r_2$ .

1. A coluna  $r_1$  da matriz de rotação é perpendicular a  $r_2$ , o que faz do seu produto zero, sendo esse resultado inalterado pela escala  $\lambda$ .

$$r_1^T r_2 = 0 \quad (9)$$

2. Outra restrição que vem da ortonormalidade é que cada coluna da matriz de rotação é um vetor unitário, o que pode ser expresso como  $r_1^T r_1 = 1$ .

$$r_1^T r_1 = r_2^T r_2 \quad (10)$$

Se a matriz de homografia  $H$  é formada pelos vetores das colunas  $h_1, h_2$  e  $h_3$ , então podemos escrever as restrições dadas nas equações 9 e 10 dentro da equação 8, assim teremos:

$$h_1^T K^{-T} K^{-1} h_2 = 0 \quad (11)$$

$$h_1^T K^{-T} K^{-1} h_1 = h_2^T K^{-T} K^{-1} h_2 \quad (12)$$

Agora se definimos uma matriz  $B$  como:

$$B = K^{-T} K^{-1} = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix} \quad (13)$$

Notamos que esta matriz  $B$  representa a imagem da cônica absoluta, descrita como uma matriz simétrica, o que permite defini-la com seis incógnitas. Reescreveremos as equações em 11 e 12 usando essa matriz.

$$h_1^T B h_2 = 0 \quad (14)$$

$$h_1^T B^{-1} h_1 - h_2^T B h_2 = 0 \quad (15)$$

Expandindo as equações 14 e 15, podemos montar um sistema de equações que permita calcular os valores dos componentes  $b_{ij}$  da matriz simétrica  $B$ . Essa expansão é definida na equação embaixo.

$$\begin{bmatrix} h_1 h_{12} & h_2 h_{21} + h_1 h_{22} & h_2 h_{22} & h_2 h_{31} + h_1 h_{32} & h_2 h_{31} + h_2 h_{32} & h_3 h_{32} \\ h_1^2 - h_2^2 & 2(h_1 h_{21} - h_2 h_{22}) & h_2^2 - h_2^2 & 2(h_1 h_{31} - h_2 h_{32}) & 2(h_2 h_{31} - h_2 h_{32}) & (h_3^2 - h_2^2) \end{bmatrix} b = 0 \quad (16)$$

Na qual  $b$  é definida pelos componentes de matriz simétrica  $B$ :

$$b = [b_{11} \quad b_{12} \quad b_{22} \quad b_{13} \quad b_{23} \quad b_{33}]^T$$

Se nós temos  $n$  imagens de nosso padrão planar, podemos determinar  $n$  equações semelhantes à equação 16 e formar uma matriz  $V$  de dimensão  $2n \times 6$  que defina o sistema:

$$Vb = 0 \quad (17)$$

Esse novo sistema de equações precisa de pelo menos três vistas de nosso padrão planar. Se  $n \geq 3$  podemos achar uma solução utilizando novamente o método de decomposição de valores singulares (SVD).

Com os componentes de nossa matriz  $B$  já calculados, podemos extrair os valores dos parâmetros intrínsecos definidos na matriz  $K$ , utilizando as seguintes equações apontadas no trabalho de Zhang [40].

$$v_0 = \frac{b_{12} b_{13} - b_{11} b_{23}}{b_{11} b_{22} - b_{12}^2} \quad \lambda = b_{33} - \frac{b_{13}^2 + v_0 (b_{12} b_{13} - b_{11} b_{23})}{b_{11}}$$

$$f_x = \sqrt{\frac{\lambda}{b_{11}}} \quad f_y = \sqrt{\frac{\lambda b_{11}}{b_{11} b_{22} - b_{12}^2}}$$



$$s = \frac{-b_{12}f_x^2 f_y}{\lambda} \quad u_0 = \frac{sv_0}{f_y} - \frac{b_{13}f_x^2}{\lambda}$$

Uma vez calculado o valor da matriz  $K$ , podemos prosseguir estimando os parâmetros extrínsecos, isto é, a matriz de rotação e o vetor de translação. Para isso usaremos a equação 15, que define a nossa matriz da homografia  $H$ :

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = H \begin{bmatrix} X_w \\ Y_w \\ 1 \end{bmatrix} \Rightarrow \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \lambda K \begin{bmatrix} r_1 & r_2 & t \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ 1 \end{bmatrix}$$

$$H = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} = [h_1 \quad h_2 \quad h_3]$$

$$r_1 = \lambda K^{-1} h_1$$

$$r_2 = \lambda K^{-1} h_2$$

$$r_3 = r_1 \times r_2$$

$$t = \lambda K^{-1} h_3$$

Na qual  $r_1, r_2, r_3$  são as colunas da matriz de rotação  $R$  e  $t$  é o vetor de translação.

Finalmente, tendo completado a extração dos parâmetros intrínsecos e extrínsecos para nossa câmera, faremos a extração dos coeficientes de distorção da lente.

#### 2.2.1.4. Determinação dos coeficientes de distorção da lente

Uma vez calculados os parâmetros intrínsecos e extrínsecos para a câmera, podemos achar uma solução para definir os valores dos coeficientes de distorção radial ( $k_1, k_2$ ) considerados no trabalho de Zhang [40]. Esses coeficientes modelam bem o problema gerado pela distorção da lente.

Para calcular o valor dos coeficientes utilizaremos a informação sobre a posição dos pontos referenciais  $m_i = (u_d, v_d)$  processados e extraídos a partir das vistas capturadas de nosso padrão planar. Consideraremos essas posições como as posições reais dos pontos na imagem. Porém, também podemos obter uma aproximação dessas posições reprojando os pontos 3D  $M_i = (X_w, Y_w, I)$  determinados por nosso padrão real usando a homografia  $H$  previamente calculada. Isso permite obter pontos reprojados sobre o plano da imagem  $m'_i =$

$(u_u, v_u)$  considerados pontos ideais sem distorção. O resíduo da diferença entre os valores dos pontos reais  $(u_d, v_d)$  e os pontos ideais  $(u_u, v_u)$  será usado para encontrar o valor dos nossos coeficientes de distorção. Essa relação entre pontos distorcidos ou reais e os sem distorção ou ideais é definida nas equações abaixo:

$$u_d = u_u + (u_u - u_0)(k_1 r^2 + k_2 r^4) \quad (18)$$

$$v_d = v_u + (v_u - v_0)(k_1 r^2 + k_2 r^4) \quad (19)$$

O valor de  $r$  é o raio determinado pelos valores normalizados do ponto sem distorção  $(u_u, v_u)$  em relação ao centro da imagem  $(u_0, v_0)$ . Esse novo ponto é definido como  $(x, y)$  e raio é calculado na equação embaixo.

$$r^2 = x^2 + y^2 \quad (20)$$

Com as equações 18, 19 e 20 podemos determinar o seguinte sistema de equações:

$$\begin{bmatrix} (u_u - u_0)r^2 & (u_u - u_0)r^4 \\ (v_u - v_0)r^2 & (v_u - v_0)r^4 \end{bmatrix} \begin{pmatrix} k_1 \\ k_2 \end{pmatrix} = \begin{pmatrix} u_d - u_u \\ v_d - v_u \end{pmatrix} \quad (21)$$

Como temos  $m$  pontos referenciais  $(u_d, v_d)$  capturados por cada vista do padrão planar e  $n$  vistas do padrão, podemos montar um sistema de equações de dimensão  $2mn$ , definido da seguinte forma:

$$Dk = d$$

$$D = \begin{bmatrix} (u_u^1 - u_0)r^2 & (u_u^1 - u_0)r^4 \\ (v_u^1 - v_0)r^2 & (v_u^1 - v_0)r^4 \\ \vdots & \vdots \\ (u_u^{mn} - u_0)r^2 & (u_u^{mn} - u_0)r^4 \\ (v_u^{mn} - v_0)r^2 & (v_u^{mn} - v_0)r^4 \end{bmatrix}, k = \begin{pmatrix} k_1 \\ k_2 \end{pmatrix}, d = \begin{pmatrix} u_d^1 - u_u^1 \\ v_d^1 - v_u^1 \\ \vdots \\ u_d^{mn} - u_u^{mn} \\ v_d^{mn} - v_u^{mn} \end{pmatrix} \quad (22)$$

Usando o método de mínimos quadrados, podemos achar uma solução para esse sistema., que pode ser definida pela pseudo-inversa da matriz  $D$ :

$$k = (D^T D)^{-1} D^T d \quad (23)$$

Finalmente, termina-se o processo de calibração de câmera, usando como únicas informações as vistas de um padrão planar.

Um resumo do método apresentado nesta seção é descrito a seguir:

### Método de calibração baseado no cálculo de homografias

- **Definição:** Calibração de câmera usando  $n$  vistas de um padrão planar

- **Dados de entrada:** Pontos 2D extraídos de cada vista do padrão linear
- **Fluxo do processo:**
  1. Calcular as homografias  $H_i$  para cada uma das  $n$  vistas capturadas do padrão planar, são necessárias  $n \geq 3$  vistas.
  2. Calcular a matriz  $V$  de dimensão  $2n \times 6$  determinada pelas restrições de ortonormalidade definidas nas equações 14, 15 e 16.
  3. Calcular uma solução para o sistema  $Vb = 0$ , no qual  $b$  contém os componentes da matriz simétrica  $B$ , que representa a imagem de um cônica absoluta definida a partir da matriz de parâmetros intrínsecos  $K$  (equações 6 e 16).
  4. Extrair os parâmetros intrínsecos, e depois os parâmetros extrínsecos usando as equações apontadas na seção 2.2.1.3.
  5. Estimar o valor dos coeficientes de distorção das lentes, conforme seção 2.2.1.4.
- **Dados de saída:** Solução inicial para os valores dos parâmetros intrínsecos, extrínsecos e coeficientes de distorção da lente para uma câmera.

Com a descrição desse método fotogramétrico tentamos explicar o processo que permite calibrar todos os parâmetros de uma câmera. Na seção seguinte descreveremos outro tipo de método de calibração, chamado de auto-calibração. Este tipo de método foi proposto para resolver a calibração de mais de uma câmera, o que já o diferencia dos métodos fotogramétricos.

### 2.2.2. Auto-calibração

Neste tipo de método, a calibração dos parâmetros intrínsecos e extrínsecos é feita a partir das informações fornecidas por pontos correspondentes corretamente identificados em imagens capturadas pela(s) câmera(s). Nesse sentido, este tipo de método não precisa de padrões complexos com uma estrutura bem conhecida, como os métodos fotogramétricos. Pelo contrário, os pontos correspondentes podem ser capturados usando padrões unidimensionais ou adimensionais, como os descritos na seção 2.1.3 e 2.1.4.

No caso de usar diferentes câmeras é preciso que os pontos referenciais estejam devidamente correlacionados e identificados entre as diversas câmeras

que visualizam um mesmo cenário. Pode-se afirmar que esses métodos assumem que a cena visualizada é estática, e é só baseado na informação 2D dos pontos referenciais capturados e corretamente correlacionados, que podemos extrair e calcular os parâmetros de calibração das câmeras. Esse tipo de método é também muito utilizado em aplicações que realizam reconstrução 3D.

A área de pesquisa que explora essa nova estratégia de calibração de câmera e define esse tipo de método é conhecida como “estrutura a partir do movimento” (*structure from motion*). A base da implementação deste tipo de método se encontra na teoria da geometria epipolar, representada principalmente pelo cálculo da matriz fundamental.

Nesta seção, descreveremos alguns conceitos referentes à geometria epipolar, ao cálculo da matriz fundamental e à extração dos parâmetros extrínsecos das câmeras baseada na decomposição da matriz fundamental. Esse processo de extração dos parâmetros extrínsecos é utilizado na implementação do método proposto neste trabalho.

### 2.2.2.1. Geometria epipolar

Geometria epipolar é a geometria intrínseca projetiva existente entre duas vistas de um mesmo cenário, capturadas a partir de diferentes perspectivas. As vistas podem ser capturadas por uma mesma câmera, sendo necessário apenas mudar a posição da câmera no momento da captura, ou as vistas podem vir da captura de várias câmeras enxergando um mesmo cenário simultaneamente.

A geometria epipolar é independente da estrutura da cena e só depende dos parâmetros internos das câmeras e da posição relativa entre elas [10]. Para definir a posição relativa entre as câmeras utiliza-se uma transformação composta de uma matriz de rotação e um vetor de translação, que permite levar o sistema de coordenadas da câmera (SCC) da primeira câmera para o da segunda. A matriz de rotação e o vetor de translação são considerados os parâmetros extrínsecos que relacionam os sistemas das duas câmeras, mais conhecido como sistema estéreo de câmeras.

Essa geometria é geralmente requerida devido à necessidade de encontrar uma correspondência entre pontos que aparecem nas diferentes imagens das vistas de um mesmo cenário. Assim, suponhamos que temos um ponto  $M$  no mundo, que

é projetado sobre o plano da imagem de duas câmeras como sendo os pontos  $m$  e  $m'$  respectivamente (Figura 9). Para mostrar a correspondência que existe entre os pontos  $m$  e  $m'$ , primeiro definiremos o plano  $\pi$  formado pelos vetores  $\overrightarrow{CM}$ ,  $\overrightarrow{C'M}$  e  $\overrightarrow{C'C}$ , nos quais estão contidos nossos pontos  $m$  e  $m'$ . Este plano é definido como o plano epipolar.

Agora, se projetamos o vetor  $\overrightarrow{CM}$ , que passa pelo ponto  $m$  no plano da imagem da câmera da esquerda sobre o plano da imagem da direita, descendo pelo plano epipolar, nós vemos que esse vetor é projetado na forma de uma linha  $l'$  no plano da imagem da direita. Essa linha deve conter o ponto  $m'$  que é a projeção do mesmo ponto  $M$  sobre o plano da imagem da direita. Esse processo permite demonstrar a correspondência entre os pontos  $m$  e  $m'$ , que são as projeções de um ponto no espaço  $M$  sobre os planos da imagem das câmeras.

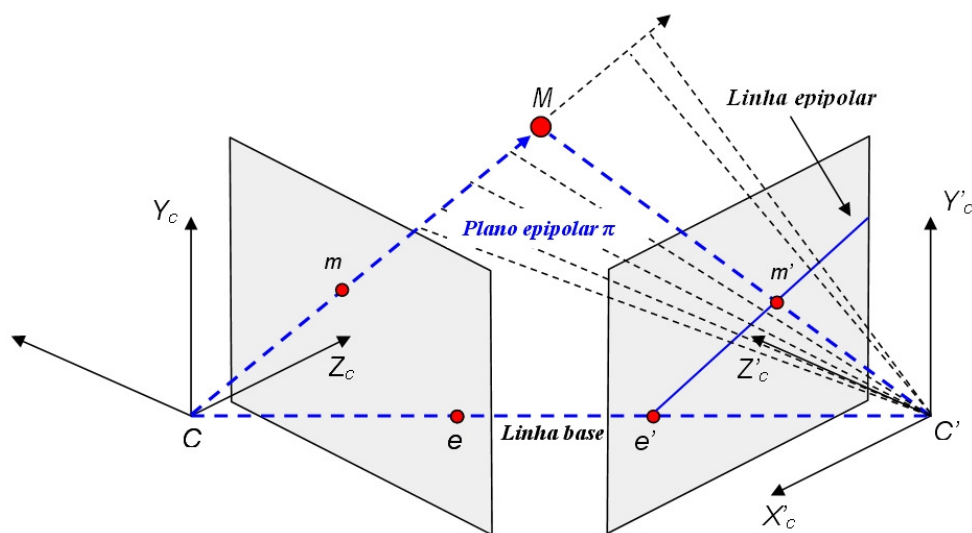


Figura 9: Elementos da geometria epipolar - linha epipolar, epipolos, linha base, plano epipolar.

Os pontos de interseção da linha base, determinada pelo vetor  $\overrightarrow{C'C}$ , com cada plano da imagem são definidos como os epipolos  $e$  e  $e'$ . Já as linhas  $l$  e  $l'$  formadas pelas projeções dos vetores definidos por  $\overrightarrow{CM}$ ,  $\overrightarrow{C'M}$  sobre os planos da imagem das câmeras, são descritas pelas linhas epipolares (Figura 10).

Como foi observado, a geometria epipolar ajuda a entender a geometria projetiva implícita existente entre duas vistas de um mesmo cenário, mas um elemento chave que formula essa correspondência formalmente é a matriz

fundamental. Na seção seguinte, descreveremos como essa matriz também relaciona os parâmetros das câmeras.

### 2.2.2.2. Matriz fundamental

Como foi definido, a matriz fundamental é um componente que permite expressar matematicamente as propriedades da geometria epipolar usando a álgebra linear.

Essa matriz encapsula a geometria intrínseca que descreve a geometria epipolar. Internamente essa matriz pode ser determinada pelos parâmetros intrínsecos e extrínsecos que correlacionam nossas câmeras.

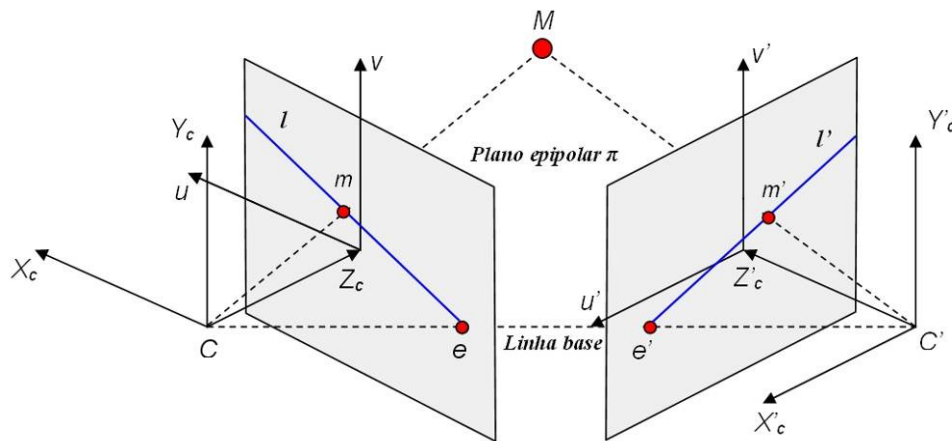


Figura 10: Correspondência entre duas vistas de um sistema estéreo de câmeras.

Tendo como referência a Figura 10, observarmos que um ponto  $M$  ( $M' = M$ ,  $M'$  expressado no SCC da direita) junto à origem das câmeras  $C$  e  $C'$  define três vetores coplanares,  $\overrightarrow{CM}$ ,  $\overrightarrow{C'C}$  e  $\overrightarrow{C'M'}$ . Essa relação pode ser expressa como:

$$\overrightarrow{C'M'}^T \left( \overrightarrow{C'C} \times \left( \overrightarrow{CM} \right)' \right) = 0 \quad (24)$$

Na equação 24,  $\left( \overrightarrow{CM} \right)'$  é o vetor  $\overrightarrow{CM}$  expresso no SCC da direita. Considerando  $m$  a imagem do ponto no espaço  $M$  expresso no SCC da esquerda e  $m'$  a imagem do ponto no espaço  $M'$  expresso no SCC da direita, no qual  $M' = M$ , nós podemos expressar essa relação com as seguintes equações:

$$m = K \begin{bmatrix} I_3 & | 0 \end{bmatrix} M = KM = K \overrightarrow{CM} \quad (25)$$

$$m' = K' \begin{bmatrix} I_3 & | 0 \end{bmatrix} M' = K'M' = K' \overrightarrow{C'M'} \quad (26)$$

O vetor  $\overrightarrow{CM}$  está localizado no SCC da esquerda cuja origem é  $C$ . Então podemos rodar o SCC da esquerda para ficar alinhado com o SCC da direita usando uma matriz de rotação  $R$ .

$$\left(\overrightarrow{CM}\right)' = R \overrightarrow{CM} \quad (27)$$

Substituindo a equação 27 na equação 24, temos:

$$\overrightarrow{C'M'}^T \left(\overrightarrow{C'C} \times R \overrightarrow{CM}\right) = 0 \quad (28)$$

Se resolvermos a equação 25 e 26 em relação aos vetores 3D  $\overrightarrow{CM}$  e  $\overrightarrow{C'M'}$ , definimos  $\overrightarrow{CM} = K^{-1}m$  e  $\overrightarrow{C'M'} = K'^{-1}m'$ . Inserindo essas expressões na equação 28, podemos redefini-la como:

$$\left(K'^{-1}m'\right)^T \left(\overrightarrow{C'C} \times R K^{-1}m\right) = 0 \quad (29)$$

Na equação 29, podemos determinar que nossa linha base  $\overrightarrow{C'C}$  é o vetor de translação  $t$  ( $\overrightarrow{C'C} = t$ ), que levará a origem da nossa câmera à esquerda para a posição da origem da câmera à direita. Nós podemos usar a versão “skew symmetric matrix” do vetor  $t$  definida como  $\tilde{t}$  para representar o produto vetorial do vetor  $\overrightarrow{C'C}$  e a matriz de rotação  $R$  na equação 29.

$$m'^T K'^{-T} \tilde{t} R K^{-1}m = 0 \quad (30)$$

Na qual o vetor  $\tilde{t}$  é definido como:

$$\tilde{t} = \begin{bmatrix} 0 & -t_z & t_y \\ t_z & 0 & -t_x \\ -t_y & t_x & 0 \end{bmatrix} \Rightarrow t = \overrightarrow{C'C} = \begin{pmatrix} t_x \\ t_y \\ t_z \end{pmatrix}$$

É na transformação algébrica descrita na equação 30 que podemos expressar a correspondência entre os pontos  $m$  e  $m'$  nas duas vistas de nosso sistema estéreo de câmeras. Então a matriz fundamental é definida assim:

$$\begin{aligned} m'^T K'^{-T} \tilde{t} R K^{-1}m &= 0 \\ m'^T F m &= 0 \\ F &= K'^{-T} \tilde{t} R K^{-1} \end{aligned} \quad (31)$$

Como podemos observar, é a matriz fundamental que representa a álgebra definida pela geometria epipolar, internamente relacionando os parâmetros

intrínsecos e extrínsecos de nossas câmeras. A projeção do ponto  $m$  no plano da esquerda sobre o plano da imagem da câmera da direita fica definida como a linha epipolar  $l'$  e também pode ser expressa em função da matriz fundamental:

$$l' = Fm \quad (32)$$

$$l = F^T m' \quad (33)$$

Na equação 32, define-se a projeção do ponto  $m'$  no plano da direita sobre o plano da imagem da câmera da esquerda que determina a linha epipolar  $l$ . Essa projeção também é expressa em função da matriz fundamental (Figura 10).

Na seção a seguir será descrito o processo de cálculo da matriz fundamental baseado na informação fornecida por um conjunto de pontos correspondentes identificados corretamente entre as imagens capturadas por um sistema estéreo de câmeras.

### 2.2.2.3. Cálculo da matriz fundamental

Existem diferentes técnicas para o cálculo da matriz fundamental, como os apresentados no livro de Hartley e Zisserman [10], mas um método muito utilizado por sua simplicidade é o algoritmo dos 8 pontos, que precisa de 8 pontos correspondentes  $m \leftrightarrow m'$  entre duas imagens para realizar esse cálculo. As informações sobre o conjunto de pontos correspondentes e a correspondência baseada na definição da matriz fundamental  $m^T F m = 0$ , permitirão construir um sistema de equações cuja solução achará o valor dos componentes de nossa matriz fundamental.

Para iniciar, nós teremos como dados de entrada um número,  $n \geq 8$ , de pontos correspondentes  $m \leftrightarrow m'$  que são relacionados pela matriz fundamental baseados na equação  $m^T F m = 0$ . Nesta equação representaremos em coordenadas homogêneas nossos pontos  $m = (u, v, 1)^T$  e  $m' = (u', v', 1)^T$ . Usando as coordenadas homogêneas de  $m$  e  $m'$  podemos expandir a equação 31 e redefini-la numa única equação:

$$m'^T F m = 0$$

$$\begin{pmatrix} u' \\ v' \\ 1 \end{pmatrix}^T \begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{bmatrix} \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = 0$$



$$uu'f_{11} + uv'f_{21} + uf_{31} + vu'f_{12} + vv'f_{22} + vf_{32} + u'f_{13} + v'f_{23} + f_{33} = 0 \quad (34)$$

É assim que cada ponto correspondente gerará uma equação como a apresentada acima. A equação 34 também pode ser modelada como o produto de dois vetores:

$$\begin{pmatrix} uu' & vu' & u' & uv' & vv' & v' & u & v & 1 \end{pmatrix} f = 0 \quad (35)$$

Na qual o vetor  $f$  contém os componentes que determinam a matriz fundamental  $F$ . Se nós consideramos um conjunto de  $n$  pontos correspondentes, poderemos modelar um sistema de equações da seguinte forma:

$$Af = \begin{bmatrix} u_1u'_1 & v_1u'_1 & u'_1 & u_1v'_1 & v_1v'_1 & v'_1 & u_1 & v_1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ u_nu'_n & v_nu'_n & u'_n & u_nv'_n & v_nv'_n & v'_n & u_n & v_n & 1 \end{bmatrix} f = 0 \quad (36)$$

Para poder achar uma solução para  $f$  baseada na matriz  $A$ , ela deve ter posto 8. A fim de evitar uma solução trivial  $f = 0$ , adicionamos algumas restrições, como obrigar a norma da solução para  $|f| = 1$  e definir o último componente de  $f$  igual  $f_{33} = 1$ . Com essas duas restrições, podemos resolver o sistema de equações  $Af = 0$  usando mínimos quadrados. Para calcular uma solução para  $Af = 0$  podemos usar o método SVD, que determina a solução para  $f$  pelo autovetor relacionado ao menor autovalor definido para a matriz  $A$ . Com o vetor  $f$  calculado, nós determinamos os valores dos componentes da matriz fundamental  $F$ .

É importante garantir que a matriz fundamental seja singular e tenha posto 2, mas comumente a solução da matriz fundamental  $F$  achada com o método SVD não garante que a solução achada para os valores da matriz  $F$  tenha essas propriedades. Porém, existe um método simples que permite achar uma nova matriz fundamental  $F'$  que tenha posto 2 e esteja baseada na matriz  $F$  previamente calculada. Com base na minimização da norma de Frobenius  $|F - F'|$  sujeita à restrição  $\det(F') = 0$ , o método consiste em realizar a SVD de nossa matriz  $F$ , que é definida como  $F = UDV^T$ , na qual  $D = (r,s,t)$  é a matriz que contém na diagonal os dados dos autovalores para nossa matriz  $F$ . Os autovalores estão em ordem decrescente  $r \geq s \geq t$ . Então nossa nova matriz  $F'$  será definida como  $F' = U \text{diag}(r,s,0) V^T$ , onde o menor autovalor é zerado. Essa nova matriz  $F'$  minimiza a norma definida por Frobenius e estabelecida como nossa nova matriz fundamental.

Finalmente, com a matriz fundamental já calculada, o passo seguinte será entender como extrair os parâmetros extrínsecos para o sistema estéreo de câmeras. O cálculo dos parâmetros intrínsecos e da distorção da lente de cada câmera fica por conta do método fotogramétrico descrito na seção 2.2.1.3.

#### 2.2.2.4. Determinação dos parâmetros extrínsecos

Como se observa na equação 31, a matriz fundamental pode ser descrita usando os parâmetros intrínsecos e extrínsecos das câmeras de um sistema estéreo. Até este momento temos já calculados os parâmetros intrínsecos ( $K$ ,  $K'$ ) e a matriz fundamental  $F$  para esse sistema, mas ainda falta calcular os parâmetros extrínsecos definidos por  $R$  e  $t$  na equação 31.

Os parâmetros extrínsecos determinam um novo tipo de matriz conhecida como a matriz essencial  $E$ . Essa matriz contém todas as informações que permitem posicionar uma câmera em relação à câmera vizinha e é definida na equação 37:

$$\begin{aligned} F &= K'^{-T} \tilde{t} R K^{-1} \\ F &= K'^{-T} E K^{-1} \\ E &= \tilde{t} R \end{aligned} \tag{37}$$

Uma forma de calcular os componentes  $R$  e  $t$  dessa matriz essencial será apresentada a seguir:

Primeiro, com base na definição da equação 31 e sabendo os valores das matrizes de parâmetros intrínsecos das câmeras, podemos extrair a matriz essencial de nossa matriz fundamental, previamente calculada, da seguinte forma:

$$E = K'^T F K \tag{38}$$

Segundo, vamos assumir que o SCC da esquerda é nosso sistema de coordenadas base. Isso significa que nossa câmera à direita estará posicionada em relação ao SCC da esquerda, o que nos permite criar duas novas matrizes de projeção  $P$  e  $P'$  simplificadas. Estas são definidas unicamente pela matriz de rotação  $R$  e pelo vetor de translação  $t$ . A definição de  $P$  e  $P'$  exemplifica o processo que permite levar o SCC da direita para o SCC da esquerda:

$$\begin{aligned} P &= [I_3 \quad | \quad 0] \\ P' &= [R \quad | \quad t] \end{aligned}$$

Agora suponhamos que a SVD da nossa matriz essencial  $E$  é definida da seguinte forma:

$$E = U \text{diag} (1 \quad 1 \quad 0) V^T \quad (39)$$

Baseados na SVD da matriz essencial  $E$ , nós podemos determinar a matriz  $R$  e o vetor  $t$  como segue:

$$R = U W V^T \quad \text{ou} \quad R = U W^T V^T \quad (40)$$

$$t = u_3 \quad \text{ou} \quad t = -u_3 \quad (41)$$

Onde  $u_3$  é a última coluna de nossa matriz  $U$ , e a matriz  $W$  é definida como:

$$W = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Essa estratégia para achar possíveis soluções para a matriz  $R$  e o vetor  $t$  baseada na fatorização das matrizes componentes da SVD da matriz  $E$ , gera quatro possíveis soluções produto das combinações das duas possíveis soluções para  $R$  e  $t$  determinadas nas equações 40 e 41. Essas soluções ficam definidas como:

$$\begin{aligned} P' &= [U W V^T \mid u_3] & P' &= [U W V^T \mid -u_3] \\ P' &= [U W^T V^T \mid u_3] & P' &= [U W^T V^T \mid -u_3] \end{aligned}$$

A ambiguidade criada pela definição dessas quatro possíveis soluções é ilustrada na figura 11.

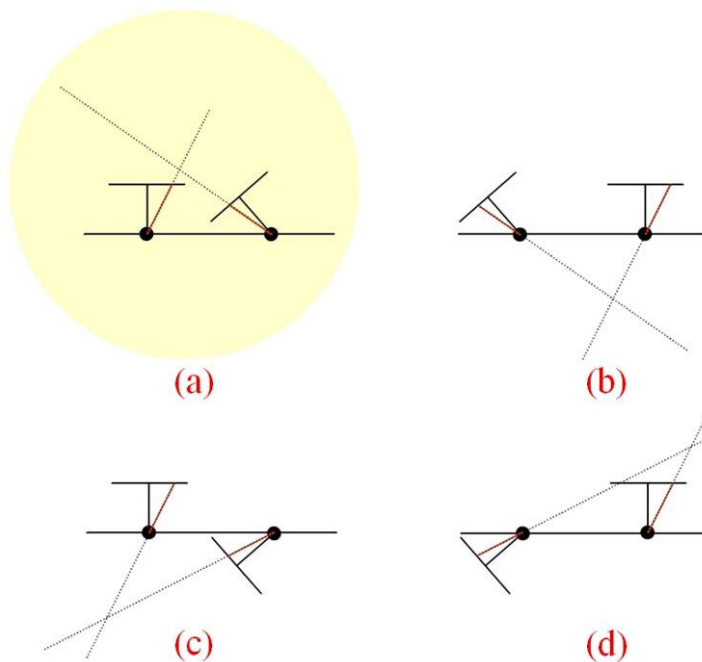


Figura 11: Ambiguidade da definição das quatro possíveis soluções para  $R$  e  $t$  [10].

Como podemos observar na Figura 11, só uma das soluções gera um sistema de coordenadas válido, onde o espaço de visão das câmeras ficará definido em frente às duas câmeras (Figura 11a). Para poder encontrar qual solução das quatro soluções propostas tem a característica visualizada na Figura 11a é realizado um teste simples, que consiste em:

1. Escolher um ponto da amostra de pontos  $m \leftrightarrow m'$ .
2. Usando os valores para R e t determinados em cada solução, reconstruir esse ponto no espaço.
3. Determinar a profundidade desse ponto em relação a cada câmera.
4. Escolher a solução na qual a profundidade encontrada para o ponto reconstruído é positiva em relação às duas câmeras.

Finalmente, terminamos o processo de cálculo de todos os parâmetros das câmeras que formam um sistema estéreo. Os métodos descritos nesta seção têm como objetivo achar uma solução inicial para esses valores. No entanto, quando há um sobredimensionamento de informações de entrada – como um número maior de vistas do padrão planar, no caso do método de calibração fotogramétrico ou mais do que oitos pontos referenciais correspondentes no cálculo da matriz fundamental – talvez seja necessário algum método de otimização que permita melhorar essa estimativa inicial. Na seção a seguir será descrito esse tipo de método, que define a última etapa do nosso framework conceitual definido para o processo de múltipla calibração de câmeras.

### **2.3. Terceira etapa: otimização da calibração**

Esta terceira etapa tem como objetivo agrupar os diferentes métodos usados para otimizar o valor inicial dos parâmetros de calibração, obtidos com os métodos anteriormente descritos. As soluções obtidas para os parâmetros das câmeras, tanto com o método fotogramétrico como com o de auto-calibração, foram calculadas com base na minimização algébrica de um sistema de equações que modela a correspondência entre pontos 3D-2D (fotogramétrica) e pontos 2D-2D (auto-calibração). Essas soluções satisfazem uma aproximação algébrica linear para um problema que não é linear, como o processo de calibração de câmera. A minimização algébrica pode também não levar em conta outras propriedades físicas do modelo de câmera real definidas pelas características dos pontos de

referência usados no processo de calibração e por alguns componentes internos das câmeras, como a lente. Por esse motivo e pela definição de não-linearidade do processo de calibração, é requerida a aplicação de algoritmos de otimização que minimizem o erro produzido na avaliação das características físicas implícitas, que vem junto as informações fornecidas pelos pontos referenciais capturados a partir dos nossos padrões de calibração.

Os algoritmos de minimização não-linear tentam melhorar a estimativa inicial obtida para os parâmetros intrínsecos, extrínsecos e coeficientes de distorção da lente das múltiplas câmeras incluídas no sistema. Para poder realizar esse aprimoramento dos valores dos parâmetros, os algoritmos precisam da definição de uma função global de erro que avalie, a cada iteração, se a nova solução estimada a partir de alguma variação da solução inicial gera melhores resultados. Muitas vezes essa função de erro global não é única e é definida por um conjunto de funções de controle especificamente criadas e relacionadas a propriedades e características físicas presentes tanto no processo de calibração como no formato definido entre os pontos referenciais capturados a partir dos padrões de calibração. Algumas das funções de controle comumente utilizadas serão descritas na próxima seção.

### **2.3.1. Funções de controle**

Entende-se como função de controle aquela função que avalia o resultado definido por alguma característica ou propriedade física obtida a partir dos componentes incluídos no processo de calibração de câmera. Descreveremos a seguir algumas das funções mais utilizadas.

#### **2.3.1.1. Reprojção 3D – 2D**

Este tipo de função é utilizada quando temos informação tanto da posição 2D na imagem como da posição 3D no espaço dos pontos referenciais capturados em nossa amostra para calibração. Com essa informação mais a estimativa dos parâmetros de calibração, nós podemos reprojeter sobre o plano da imagem os pontos 3D, que devem ficar muito próximos aos pontos 2D correspondentes aos mesmos pontos referenciais, mas que foram extraídos diretamente da imagem

baseados nos algoritmos de processamento de imagem como os descritos na seção sobre padrões de calibração.

O somatório da diferença entre o ponto 2D extraído a partir da imagem e o calculado com base no processo de reprojeção é o valor de erro determinado para esse tipo de função de controle. O objetivo da minimização é aproximar este de zero. Essa função é utilizada tanto no método fotogramétrico quanto no método de auto-calibração e é definida como:

$$\mathcal{E}_{Reproj. 3D-2D} = \sum_{i=1}^n \sum_{j=1}^m \left\| m_{ij} - m_{Reproj}(K, R, t, M_j) \right\|^2 \quad (42)$$

Na equação 42, o erro de reprojeção é estabelecido pelo somatório da norma das distâncias entre o ponto de referência extraído a partir da imagem  $m_{ij}$  e o ponto correspondente  $m_{Reproj}$  – que é a reprojeção do ponto  $M_j$  no espaço usando os parâmetros da câmera  $K, R, t$ . Nós podemos ter  $i=1\dots n$  número de vistas do padrão de calibração capturadas que internamente fornecem  $j = 1..m$  pontos de referência por vista.

### 2.3.1.2. Reprojeção 2D – 2D

Este tipo de função é definida como a distância da linha epipolar ao ponto correspondente num sistema estéreo de câmeras, como exposto na seção 2.2.2.4. Para poder modelar essa função precisamos das informações sobre os pontos correspondentes 2D  $m_i \leftrightarrow m_i'$ , extraídos a partir de duas imagens de um mesmo cenário. Para ambas, já calculamos o valor da matriz fundamental que as relaciona.

A relação estabelecida pela matriz fundamental permite a projeção de um ponto  $m$  numa câmera  $A$  sobre o plano da imagem de uma câmera  $B$ . A matriz fundamental  $F$  determina a relação entre as câmeras  $A$  e  $B$ . A projeção do ponto  $m$  então cria uma linha  $l'$  no plano da imagem da câmera  $B$  (Figura 12) e sobre essa linha é que deve estar posicionado o ponto correspondente  $m'$ . No caso ideal, a distância do ponto  $m$  à linha  $l'$  deve ser zero quando o ponto  $m$  cai exatamente sobre a linha  $l'$ . No entanto, por causa do ruído e do valor inicial dos parâmetros da câmera, essa distância fica perto de zero.

Então essa função é definida como uma projeção 2D-2D, na qual o erro estabelecido é medido pela distância de um ponto  $m'$  à sua linha correspondente  $l'$  e pela distância de um ponto  $m$  à sua linha correspondente  $l$ .

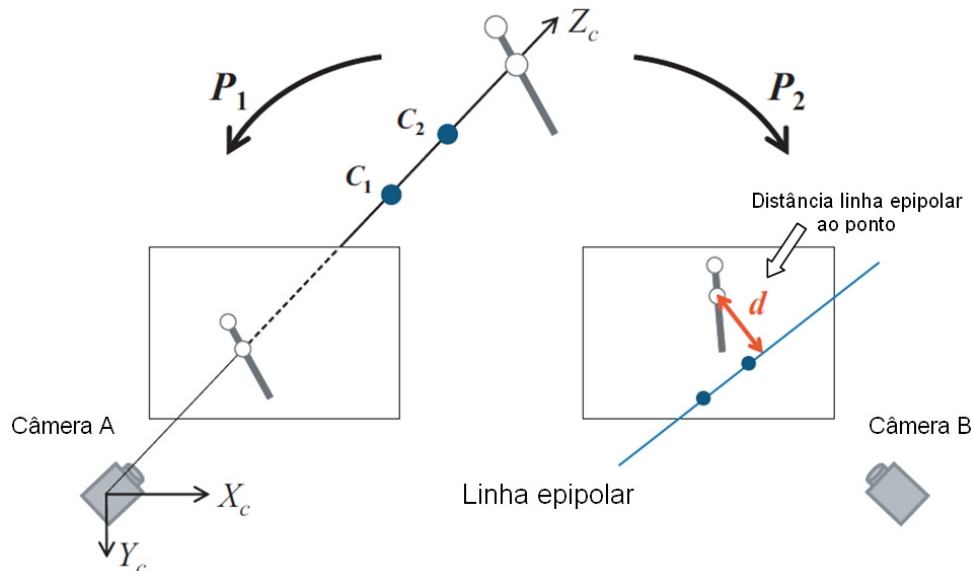


Figura 12: Exemplificação do erro definido pela distância da linha epipolar ao ponto correspondente.

O valor do erro estabelecido pelo somatório das diferenças entre as linhas  $l$  e  $l'$  a partir da projeção dos pontos correspondentes  $m_i \leftrightarrow m_i'$  é o que o algoritmo de otimização minimizará. Esse tipo de função é utilizado nos métodos de auto-calibração, especialmente nos que calculam o valor da matriz fundamental. Essa função é definida como:

$$\mathcal{E}_{\text{Reproj. 2D-2D}} = \sum_{i=1}^n \text{dist}(m_i, F^T m_i') + \text{dist}(m_i', F m_i) \quad (43)$$

A equação 43 expressa o erro de reprojeção 2D- 2D, estabelecido pelo somatório das distâncias entre o ponto  $m$  ou  $m'$  e a linha correspondente  $l = F^T m'$  e  $l' = F m$ . Nós podemos ter  $i=1 \dots n$  número de pontos correspondentes  $m_i \leftrightarrow m_i'$ .

### 2.3.1.3. Distância 3D entre pontos de referência reconstruídos

Este tipo de função é definida como a distância 3D que pode ser recuperada quando temos nossos pontos correspondentes  $m_i \leftrightarrow m_i'$  reconstruídos tridimensionalmente no espaço determinado pelo campo de visão das câmeras. Essa reconstrução 3D deve ser feita em escala semelhante a alguma unidade de medida métrica (metros, centímetros, milímetros, etc). Um exemplo de aplicação

dessa função de controle é quando usamos padrões unidimensionais, como os descritos na seção 2.1.3.

O processo de reconstrução 3D dos pontos correspondentes  $mi \leftrightarrow mi'$  pode ser feito usando um algoritmo de triangulação como o apontado em [29][30], que usa todos os parâmetros da câmera para calcular a posição 3D de um marcador  $M$  com informações dadas por  $mi$  e  $mi'$ . Com esses dados poderemos comparar as distâncias medidas a partir dos marcadores reconstruídos com o valor real medido manualmente diretamente entre os marcadores físicos do padrão. Este tipo de função é utilizada nos métodos fotogramétricos e no de auto-calibração e é definida como:

$$\mathcal{E}_{dist.3D} = \sum_{i=1}^m \left( \left| d_0 - \|X_0^i - X_1^i\| + \dots + \left| d_n - \|X_n^i - X_{n+1}^i\| \right| \right| \right) \quad (44)$$

A equação 44 expressa o somatório do erro gerado pela diferença das distâncias  $\|X_n - X_{n+1}\|$  entre marcadores reconstruídos  $X_{i=1\dots n+1}$  e a medida real da distância definida como  $d_{i=1\dots n}$ , tudo isso para  $m$  vistas reconstruídas do padrão.

Assim estabelecemos quais são as funções de controle mais utilizadas para determinar a função de erro global que nosso algoritmo de otimização usará para aprimorar nossa solução inicial. A função de erro global pode ser expressa da seguinte maneira:

$$\mathcal{E}_{\text{Erro Global}} = \mathcal{E}_{\text{Reproj. 3D-2D}} + \mathcal{E}_{\text{Reproj. 2D-2D}} + \mathcal{E}_{\text{dist. 3D}} \quad (45)$$

Essa função de erro pode mudar dependendo do método aplicado e do padrão a ser usado na aquisição de dados, onde deve se garantir que haja informação de entrada suficiente para modelar as funções de controle antes descritas. Um exemplo simples que representa a observação anterior é o uso de um padrão adimensional, que só gera informações suficientes para implementar o erro de reprojeção 3D-2D e reprojeção 2D-2D. Já os padrões planar e unidimensional podem garantir informações para definir uma função de erro da forma explicitada na equação 45.

Por fim, uma vez determinados os dois componentes básicos para rodar um algoritmo de otimização – a amostra de entrada correspondente aos pontos referenciais capturados com diversos formatos de padrões e a definição das funções de controle e função de erro global – poderemos executar o algoritmo de



otimização. No caso do processo de calibração o comportamento dos parâmetros que definem o processo não é linear. Um dos algoritmos muito utilizados para esse processo é o descrito por Levenberg Marquart [2] e uma boa implementação pode ser encontrada em [6]. Outro método também utilizado para a otimização de parâmetros no processo de calibração é o Downhill Simplex [24], que foi usado nos trabalhos de [37]. Na implementação do método proposto neste trabalho foram aplicados os dois métodos para otimizar nossos parâmetros e realizar algumas medidas de comparação no capítulo de resultados.

A seção a seguir descreverá mais uma característica importante para estabelecer como podem ser montados os sistemas de múltiplas câmeras: a distribuição e posicionamento das câmeras em relação ao espaço de rastreamento. Esta característica gera restrições adicionais no momento de escolher qual o método de calibração a ser utilizado para um determinado tipo de distribuição.

## **2.4. Distribuição das câmeras**

Entre os formatos existentes para distribuir e posicionar múltiplas câmeras que formarão parte de um sistema único – como o definido quando resolvemos o problema de múltipla calibração de câmeras – nós encontramos dois tipos bem diferentes: distribuição centralizada e descentralizada de câmeras, que serão descritos nas seções seguintes.

### **2.4.1. Distribuição centralizada de câmeras**

Neste tipo de distribuição e posicionamento de múltiplas câmeras, encontram-se classificados os sistemas que se caracterizam pelo fato de todas as câmeras conseguirem enxergar um mesmo ponto central de referência, definido como centro do sistemas de coordenadas do mundo (SCM). Trabalhos apresentados em [4],[32] e [37] propõem e testam algoritmos de estéreo e múltipla calibração de câmeras em sistemas com esse tipo de distribuição. O posicionamento das câmeras costuma ter uma estrutura circular, como o apresentado na figura 13.

Para determinar onde fica o centro do sistema de coordenadas do mundo é usado algum tipo de padrão geralmente planar, como os descritos na seção 2.1.2. Usando o padrão como referência calculam-se a posição e a orientação das

câmeras em relação ao centro do padrão. Esse tipo de cálculo pode ser feito usando um método de calibração fotogramétrico que permite calcular os parâmetros extrínsecos de cada câmera em relação ao padrão .

A principal vantagem desse tipo de distribuição é que um mesmo padrão pode ser visto simultaneamente por todas as câmeras, especialmente na captura de uma amostra para calibração. A desvantagem fica por conta da área restrita que as câmeras podem cobrir ou enxergar como área real de calibração. Um exemplo desse tipo de distribuição é mostrado na Figura 13.



Figura 13: Distribuição centralizada de câmeras.

(Fonte: <http://blogs.ign.com/Blogs/BlogPage.aspx?blog=EA-FIFA-Street&year=2008&month=01>)

Porém, existe outro grupo de distribuição de câmeras cujo objetivo é cobrir áreas amplas de captura e rastreamento que não podem ser enxergadas simultaneamente por todas as câmeras do sistema. Essa outra forma de distribuição é conhecida como descentralizada e será descrita na seção a seguir.

#### **2.4.2. Distribuição descentralizada das câmeras**

Neste outro tipo de classificação de distribuição das câmeras, encontram-se os sistemas que não têm uma área de visão comum entre todas as câmeras, porém, garante-se que para cada câmera há uma outra que tenha um campo de visão compartilhado. Esse tipo de distribuição é mais utilizado em sistemas que

precisam cobrir ou enxergar áreas amplas de rastreamento ou captura. Trabalhos como os apresentados em [22] e [14] descreveram esse tipo de distribuição de câmeras.

Neste tipo de distribuição, as câmeras são posicionadas uma em relação a outra de modo de criar uma estrutura semelhante a um grafo conexo, na qual algumas câmeras serão denominadas de câmeras referenciais ou pivôs. Uma câmera pivô é selecionada dependendo de certas características, como compartilhar um maior campo de visão com outras câmeras. Conforme descrito em [14] uma estratégia para posicionar as câmeras é montar um grafo cujos vértices são nossas câmeras e cujas arestas têm um determinado peso gerado por alguma função de erro como as apresentadas na seção 2.1.3.1. O valor do erro permite escolher entre quais arestas temos o menor caminho que relacione todas as câmeras. Esse caminho ajuda a conseguir um menor erro de propagação gerado pelo posicionamento concatenado das câmeras. Essa abordagem mostra mais flexibilidade e tem a vantagem de cobrir um maior campo útil de visão em relação à distribuição das câmeras. Um exemplo desses sistemas numa aplicação de captura de movimento para jogos é mostrada na Figura 14.



Figura 14: Distribuição descentralizada de câmeras.

(Fonte: <http://blogs.ign.com/Blogs/BlogPage.aspx?blog=EA-FIFA-Street&year=2008&month=01>)

As duas distribuições de câmeras descritas acima geram diferenças e problemáticas específicas que nos obrigam a diferenciá-las, especialmente quando é feita a escolha do método certo para a calibração das câmeras.

Terminadas a definição das três etapas que compõem nosso framework conceitual, proposto como parte deste trabalho, e da caracterização da distribuição e do posicionamento de câmeras, também termina a exposição dos conceitos preliminares. Neste capítulo, fizemos a descrição e a revisão de vários conceitos importantes que nos ajudaram a modelar o problema que surge quando necessitamos calibrar um sistema formado por múltiplas câmeras. Esses conceitos serão citados nos próximos capítulos, que descreverão a concepção e a implementação do método de múltipla calibração de câmeras proposto nesta tese.

### 3 Trabalhos relacionados

Neste capítulo, serão apresentados e descritos alguns métodos propostos na última década para realizar a tarefa de múltipla calibração de câmeras. Esses métodos, que têm o objetivo de calcular tanto os parâmetros intrínsecos para cada câmera como a posição e a orientação delas no espaço de rastreamento, vêm aprimorando suas estratégias em duas frentes distintas:

- A primeira é a simplificação do processo de aquisição e captura dos pontos de referência, de tal forma que seja executado por um usuário qualquer. Como prova disso, podemos citar a evolução e a proposta, nos últimos anos, dos diferentes tipos de padrões descritos na seção 2.1.1.
- A segunda é a proposta de novos métodos que internamente misturam um método fotogramétrico e um de auto-calibração. Esses métodos considerados híbridos tentam aproveitar a robustez do cálculo dos métodos fotogramétricos e a flexibilidade dos de auto-calibração para calibrar sistemas compostos por múltiplas câmeras.

Dentre as diversas pesquisas relacionadas ao tema de múltipla calibração, escolhemos os seguintes trabalhos, que tipificam bem aspectos em comum que o processo de múltipla calibração de câmeras deve levar em conta, assim como as estratégias antes mencionadas:

- Calibração de câmeras em estéreo, apresentado por Bouguet [5] e Borghese et al. [4].
- Múltipla calibração de câmeras centralizada, apresentado por Uematsu et al. [37], Svoboda et al. [32] e Pintaric et al. [28].
- Múltipla calibração de câmeras descentralizada, apresentado por Kurillo et al. [14].

Considerando os trabalhos acima mencionados, podemos ressaltar que a proposta dos diferentes métodos de múltipla calibração começou com testes num sistema estéreo de câmeras. A partir desse sistema básico, deu-se início ao desenvolvimento de outro tipo de sistema mais complexo composto por mais de duas câmeras, mas com o mesmo problema: a calibração simultânea das mesmas.

Na seções seguintes, descreveremos alguns detalhes referentes aos trabalhos de Bouguet [5], Borghese et al. [4], Svoboda et al. [32] e Kurillo et al. [14], que inseriram inovações importantes na evolução dos métodos de múltipla calibração de câmeras. Os detalhes descritos servirão de guia para a posterior descrição das contribuições do método proposto neste trabalho.

### 3.1. Método proposto por Jean-Yves Bouguet

O método de calibração de câmeras em estéreo apresentado por Jean-Yves Bouguet [5] é um trabalho muito interessante porque serve como exemplo para entender e mostrar como o uso de um método de calibração individual pode ser estendido para um cenário onde são calibradas mais de uma câmera.

Este método contém os mesmos elementos usados numa calibração fotogramétrica, descrita na seção 2.2.1.3; porém, os resultados do processo da calibração inicial são correlacionados de tal forma que permitem posicionar as câmeras em relação a um mesmo ponto de referência, que é reconhecido como centro de coordenadas do mundo. Por causa dessa característica, pode-se dizer que o método dá suporte a um tipo de distribuição e posicionamento de câmera centralizado (Figura 15).

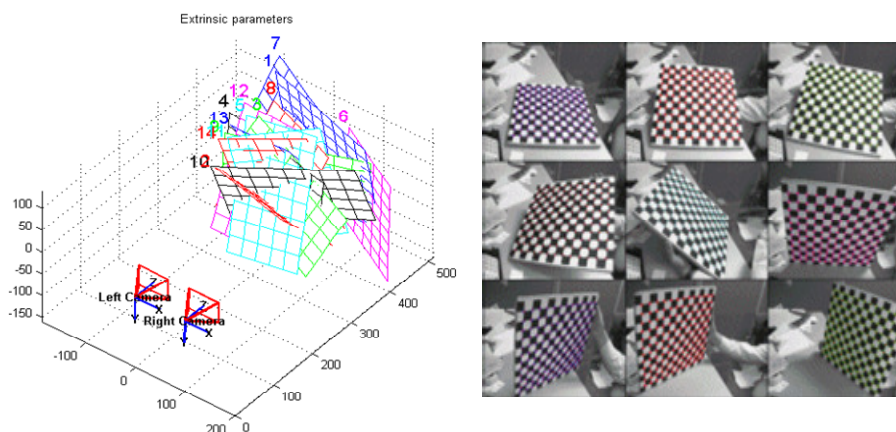


Figura 15: Calibração de um sistema estéreo de câmeras feita por Bouguet [5].

O modelo de calibração estéreo de câmeras apresentado por Jean-Yves Bouguet [5] será definido usando as etapas de nosso framework conceitual a seguir:

- **Aquisição de dados:** este método usa um único padrão de calibração que é o clássico padrão planar, semelhante a um tabuleiro de xadrez . Usando esse padrão, são capturadas simultaneamente  $n$

vistas corretamente identificadas. A amostra de  $n$  vistas fornecerá informações suficientes para realizar tanto a calibração inicial individual de cada câmera como a posterior otimização dessa calibração. Por causa das restrições derivadas da captura e reconhecimento desse tipo de padrão, o posicionamento das câmeras é restrito de tal forma que todas elas possam enxergar corretamente o padrão planar. Esse tipo de restrição no posicionamento comumente exige que as câmeras sejam posicionadas formando ângulos de  $\sim 135^\circ$  a  $180^\circ$  entre elas. Definimos  $\sim 135^\circ$  por causa da limitação que os padrões planares têm quando vistos em ângulos menores que  $45^\circ$  ( $180^\circ - 45^\circ = 135^\circ$ ).

- **Calibração inicial:** este método realiza uma calibração fotogramétrica com o conjunto de vistas capturadas individualmente em cada câmera. Como resultado, temos uma solução inicial única para o valor dos parâmetros intrínsecos e o coeficientes de distorção da lente em cada câmera, e múltiplas soluções para os parâmetros extrínsecos, definidas para cada vista do padrão planar. Para relacionar as duas câmeras, deve-se posicionar uma em relação a outra, o que é representado pelo cálculo de uma matriz de rotação  $R$  e um vetor de translação  $T$  que leve o SCC à esquerda para ficar paralelo ao SCC à direita. Isso é possível compondo os parâmetros extrínsecos ( $R_l, T_l$  e  $R_r, T_r$ ) calculados nas duas câmeras em relação a uma mesma vista específica do padrão [17] (Equações 46 e 47). Dessa forma, ocorre também a calibração inicial dos parâmetros extrínsecos que relacionam nossas câmeras.

$$R = R_r R_l^T \quad (46)$$

$$T = T_r - R T_l \quad (47)$$

- **Otimização da calibração:** este método realiza duas otimizações. A primeira é feita com a amostra de calibração capturada individualmente em cada câmera, onde são aprimorados principalmente os valores dos parâmetros intrínsecos e os coeficientes de distorção da lente nas câmeras. A segunda etapa utiliza a mesma amostra de calibração, porém, usando-a em conjunto

para aprimorar os valores dos novos parâmetros extrínsecos definidos pelas equações 46 e 47. A função de erro usada para a otimização é baseada principalmente na reprojeção 3D-2D, descrita na seção 2.3.1.1.

A principal vantagem deste método é a robustez em relação ao cálculo dos parâmetros intrínsecos e dos coeficientes de distorção da lente. Este tipo de método também permite que câmeras cujos parâmetros são desconhecidos possam ser calibrados e usados pelo sistema.

As desvantagens podem ser resumidas como:

- Difícil captura de uma amostra de calibração usando um padrão planar, o que traz muitas complicações por causa das restrições de movimento.
- Precisa de um intervalo de tempo grande para poder capturar uma amostra válida para calibração.
- Apesar de ter muita informação implícita no formato do padrão, o método de otimização só modela a função de erro principal baseado na função que mede a reprojeção 2D dos pontos da amostragem para cada vista do padrão.
- A mais importante desvantagem é a dificuldade para estender este método para uso em sistemas de múltipla calibração, principalmente pelas restrições relacionadas às possíveis distribuições e posicionamentos das câmeras no espaço de rastreamento – derivadas das restrições de captura e reconhecimento do padrão planar usado pelo método.

### **3.2. Método proposto por Borghese**

O trabalho apresentado por Borghese et al. [4] é definido como um método de auto-calibração. Na época em que foi apresentado, mostrou-se como uma evolução no processo de calibração de sistemas estéreo de câmeras. As contribuições desse trabalho são a proposta de um novo tipo de padrão de calibração e a possibilidade de calcular tanto os parâmetros intrínsecos como os extrínsecos. Esse novo padrão usado é um tipo de padrão unidimensional, como o descrito na seção 2.1.3. (Figura 16), o que tornou o método muito atraente pela



facilidade de manipulação por parte do usuário na captura da amostra de calibração.

Este método foi também muito bem recebido pelas empresas comerciais que produzem soluções relacionadas a sistemas de rastreamento óptico, como A.R.T. GmbH[1] e VICON[39], que adotaram o método para realizar a calibração dos seus sistemas.



Figura 16: Padrões de calibração usados em [1][4][39]

Para descrever as suas principais características usaremos as etapas de nosso framework conceitual:

- **Aquisição de dados:** este método usa um padrão de unidimensional com o qual captura simultaneamente uma amostra grande de  $n$  vistas corretamente identificadas nas diferentes câmeras do sistema estéreo. Por causa do formato do padrão, foi necessário implementar algumas estratégias para evitar a captura de marcadores errados, como o uso de um algoritmo de predição para rastrear na imagem a movimentação dos marcadores e garantir a identificação individual deles nas diferentes imagens capturadas pelo sistema estéreo de câmeras. Uma outra restrição para garantir isso foi a necessidade de manter a vara com os marcadores paralela à vertical do SCI, a fim de evitar oclusão entre os marcadores.
- **Calibração inicial:** este método utiliza o mesmo processo do método de auto-calibração descrito na seção 2.2.2.4, porém, adiciona a característica de também recuperar uma estimativa inicial para

alguns parâmetros intrínsecos como:  $f_x$ ,  $f_y$ ,  $u_0$ ,  $v_0$ ,  $s$  (Apêndice A). Para extrair esses valores, o método faz algumas suposições e restrições, como:  $u_0$  e  $v_0$  serão iguais à metade da dimensão da resolução da imagem capturada pela câmera,  $s = 0$ , só um valor para os componentes da distância focal  $f$  será extraído ( $f = f_x = f_y$ ) e se assume que a distorção da lente é desprezível e não é levada em conta.

- **Otimização da calibração:** este método realiza uma única otimização, que leva em conta todos os parâmetros intrínsecos e extrínsecos e define a função de erro baseada nas funções de controle de reprojeção 2D-2D e distância 3D entre os pontos de referência reconstruídos a partir das imagens dos marcadores do padrão, semelhantes ao descrito nas seções 2.3.1.2 e 2.3.1.3. A novidade do método foi explorar as medidas de distância entre os marcadores como mais uma função de controle inserida no processo de otimização.

As principais vantagens que este método traz podem ser descritas como:

- A possibilidade de extrair tanto a distância focal como os parâmetros extrínsecos a partir da análise e decomposição da matriz fundamental.
- Proposta de uma nova forma de captura rápida, numerosa e simples de executar de uma amostra para calibração graças à simplicidade do padrão.
- Capacidade de estender este método para seu uso em sistemas de múltiplas câmeras, como o demonstrado na evolução e uso deste método em produtos comerciais como [1] e [39].

As principais desvantagens são:

- A simplicidade do padrão, que pode gerar fácil contaminação da amostra com falsos marcadores.
- Falta de robustez no cálculo dos parâmetros intrínsecos, o que pode provocar erros no processo de rastreamento.

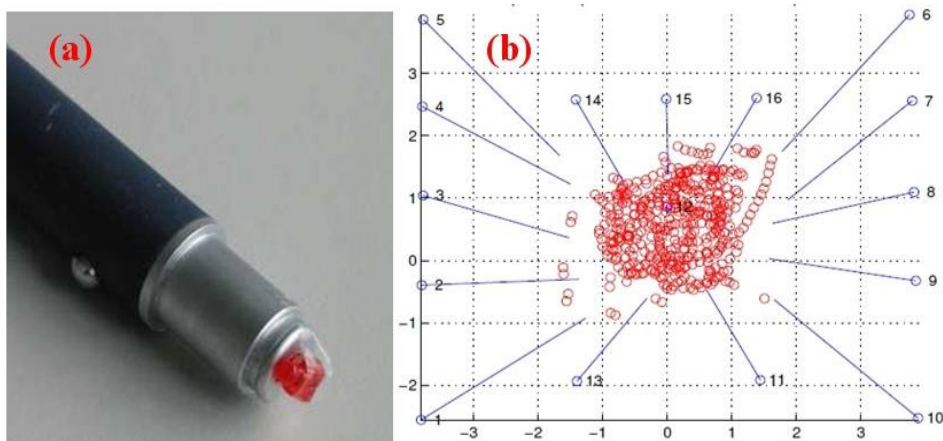
- Impossibilidade de usar câmeras com grau de distorção de lente alto porque o método não recupera os valores dos coeficientes de distorção da lente.

### 3.3. Método proposto por Svoboda

O trabalho apresentado por Svoboda et al. [32] também é definido como um método de auto-calibração, cuja característica mais ressaltante é o fato de ter sido desenhado para calibrar sistemas com mais de duas câmeras, saindo do convencional sistema estéreo de câmeras. Outra característica é a distribuição centralizada das câmeras, que deve ser garantida por causa de restrições impostas na concepção do método.

Uma das novidades que este método ofereceu foi proposta de um novo tipo de padrão de calibração, o que é definido como adimensional (seção 2.1.4). Esse novo padrão, assim como o padrão unidimensional, traz como característica muito atraente a facilidade de manipulação por parte do usuário na captura da amostra de calibração.

Este tipo de método foi também muito bem recebido pelas empresas comerciais que trabalham com soluções relacionadas a sistemas de rastreamento óptico, como IOTracker [13] e OptiTrack [26], que adotaram este método para calibrar seus sistemas.



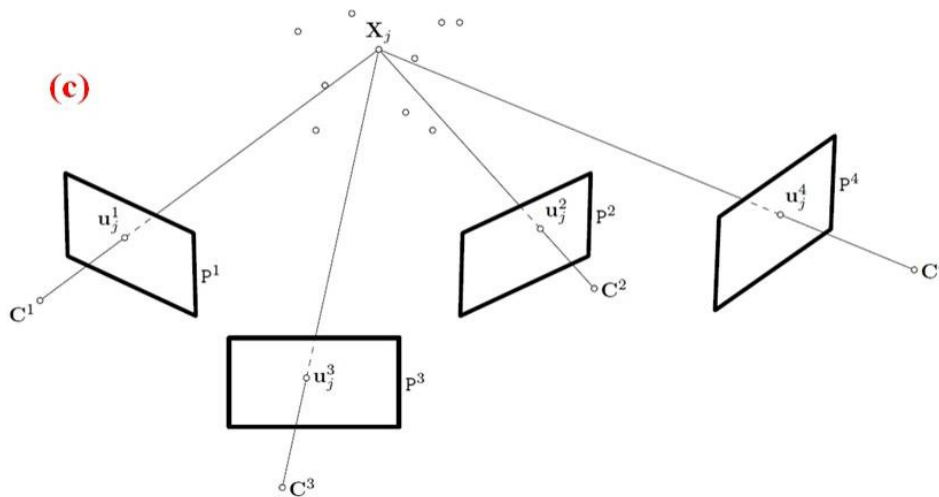


Figura 17: O ponteiro (a), distribuição de 16 câmeras (b) e esquema da múltipla calibração de câmeras definido por Svoboda et al.[32].

Seguindo as etapas de nosso framework conceitual, descreveremos a seguir as principais características deste método:

- **Aquisição de dados:** este método usa um padrão adimensional e captura simultaneamente uma amostra de  $n$  vistas do mesmo. Por causa do formato do padrão, foram implementadas algumas estratégias para a captura do mesmo, como: um algoritmo de predição para rastrear na imagem a movimentação do marcador e garantir a identificação individual nas imagens capturadas pelas múltiplas câmeras, e o cenário de captura teve que ser adaptado para ressaltar o padrão do resto de objetos.

- **Calibração inicial:** é proposto um novo método de calibração inicial baseado em dois conceitos chamados de “Rank – 4 factorization” e “Euclidean stratification”, mais detalhes dessas técnicas podem ser encontrados em [32]. Para calcular os coeficientes de distorção da lente, este método propõe o uso de algum método fotogramétrico, o que o caracteriza como um método híbrido. Algumas variações do método, como a apresentada por Pintaric et al.[28], adicionam a possibilidade de usá-lo ainda num sistema estéreo de câmeras, mas para conseguir isso, é necessário diminuir o número de parâmetros intrínsecos que serão recuperados, assumindo algumas restrições como:  $u_0$  e  $v_0$  serão iguais à metade da dimensão da resolução da imagem capturada pela câmera,  $s = 0$ , só um valor para os componentes da distância focal  $f$  será extraído ( $f = f_x = f_y$ ) e essa distância é assumida como igual para todas as câmeras.

• **Otimização da calibração:** este método realiza uma única otimização a que leva em conta todos os parâmetros intrínsecos e principalmente os extrínsecos. A função de erro é baseada na função de controle de reprojeção 3D-2D, devido às poucas informações que podem ser extraída do formato do padrão adimensional.

As principais vantagens que este método propõe podem ser descritas como:

- Propor um padrão de captura muito simples que permite uma captura numerosa de pontos de referência.
- Suportar sistemas de múltiplas câmeras formados por  $n \geq 3$  câmeras.

As principais desvantagens são descritas abaixo:

- Mais uma vez, a simplicidade do padrão pode gerar a fácil contaminação da amostra com falsos marcadores que podem ser considerados válidos.
- O cálculo dos parâmetros intrínsecos não é robusto, o que pode gerar erro no processo de rastreamento.
- Não é possível usar câmeras com grau de distorção de lente alto porque o método não recupera os valores dos coeficientes de distorção da lente.

### 3.4. Método proposto por Kurillo

O trabalho apresentado por Kurillo et al. [14] também é definido como um método híbrido, isto é, mistura o uso de um método de calibração fotogramétrica e um de auto-calibração para calcular os valores para dos parâmetros intrínsecos, extrínsecos e distorção da lente. A sua característica mais ressaltante é a possibilidade de calibrar de sistemas de câmeras descentralizados, nos quais a área de cobertura das câmeras deve ser maior. Esse tipo de sistema vem ganhando espaço nos últimos anos, especialmente em aplicações destinadas a segurança, robótica e captura de movimento.

Para descrever as suas principais características usaremos as etapas de nosso framework conceitual:

- **Aquisição de dados:** por ser um método híbrido, ele faz uso de dois tipos de padrão: o planar (calibração fotogramétrica) e unidimensional (auto-calibração).

- **Calibração inicial:** aplica-se o método de calibração fotogramétrica para a recuperação dos parâmetros intrínsecos e coeficientes de distorção das lentes como foi descrito na seção 2.2.1.4, a fim de recuperar os parâmetros extrínsecos, utiliza-se uma estratégia semelhante à descrita na seção 2.2.1.3
- **Otimização da calibração:** este método realiza uma otimização em duas partes. A primeira é feita com o objetivo de otimizar o valor de todos os parâmetros das câmeras usando uma função de erro global baseada nas funções de controle: reprojeção 2D-2D (seção 2.1.3.1.2) e distância 3D entre pontos de referência reconstruídos (seção 2.1.3.1.3). A segunda parte é baseada no uso da teoria de grafos; com o objetivo de encontrar a melhor correlação de duas a duas câmeras que permita relacionar todas as câmeras do sistema. Kurillo et al. [14] provaram que nem todas as calibrações de duas a duas câmeras entre todas as câmeras do sistema geram bons resultados, só aquelas que compartilham um campo de visão comum amplo.

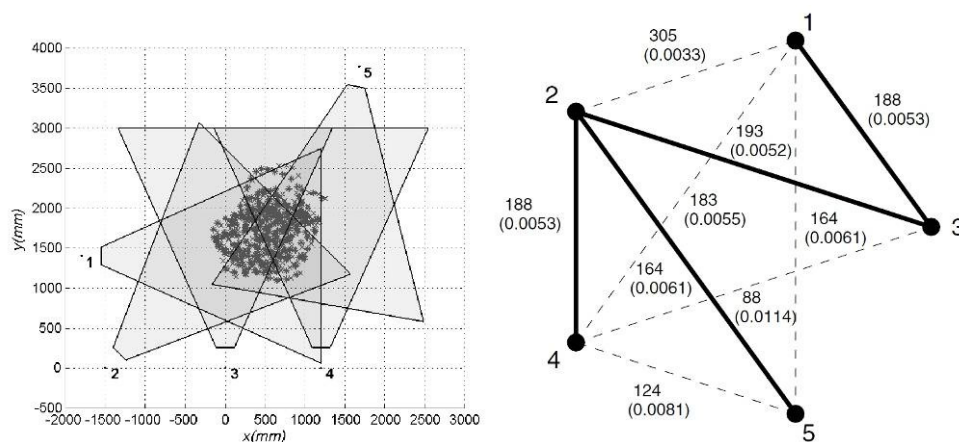


Figura 18: Análise da correlação de um sistema composto por cinco câmeras usando teoria de grafos definida em [14].

A maior contribuição do trabalho apresentado por Kurillo et al. [14] foi modelar a múltipla calibração de câmeras distribuídas em forma descentralizada como um problema da teoria de grafos (Figura 18), no qual as câmeras são os nós e as arestas definem a vizinhança de uma câmera em relação a outra. O objetivo ao analisar o grafo é encontrar quais arestas, que correlacionam duas câmeras, geram menor valor na função de erro global definida na primeira parte da etapa de otimização. Com o valor da função de erro define-se um peso para cada aresta do grafo, e o problema de escolher a melhor correlação de câmeras fica resumido a

encontrar o caminho com valor de erro menor que una todos os nós de nosso grafo. Esse caminho não precisa ser cíclico, só deve permitir que todos os nós estejam correlacionados. Maiores detalhes podem ser encontrados no artigo proposto por Kurillo et al. [14] .

A principal vantagem deste método é a capacidade de dar suporte à calibração de sistemas de múltiplas câmeras distribuídas em forma descentralizada, e as desvantagens que pode ter este método são poucas, especialmente por sua natureza híbrida, mas podemos citar problemas na captura da amostra de vistas do padrão unidimensional semelhantes às descritas no método proposto por Borghese et al. [4] na seção 3.2.

Neste capítulo, apresentamos a descrição de quatro métodos muito importantes na evolução da múltipla calibração de câmeras. Como podemos observar, as propostas de múltipla calibração começaram analisando sistemas estéreo de câmeras, e evoluíram até dar suporte a sistemas formados por múltiplas câmeras que podem ser posicionadas de tal forma que todas elas não enxerguem o mesmo campo de visão simultaneamente.

Uma observação importante a ser ressaltada é a tendência de os métodos serem concebidos cada vez mais como híbridos, isto é, misturar na etapa de calibração inicial o uso de um método fotogramétrico e um de auto-calibração para definir os diferentes parâmetros das câmeras. Essa tendência na concepção dos métodos parece prover flexibilidade e, ao mesmo tempo, robustez ao processo de calibração.

Finalmente, todos os métodos descritos neste capítulo compartilharam mais uma característica tecnológica: a sincronização eletrônica que existe entre todas as câmeras do sistema. Isso significa que as câmeras conseguem garantir que as imagens sejam capturadas exatamente no tempo “ $t$ ” em todas as câmeras. Segundo a bibliografia sobre o tema, essa característica ajuda a melhorar a precisão do sistema. Porém, mostraremos na implementação do método proposto nesta tese, que conseguimos resultados semelhantes aos descritos em especificações de sistemas comerciais, sem garantir uma sincronização muito precisa.

A seguir será descrita a concepção e a implementação do método de múltipla calibração de câmeras proposto nesta tese de doutorado.

## 4 Método proposto

Este capítulo tem por objetivo descrever detalhadamente os conceitos e diversos processos aplicados na concepção e modelagem do um novo método de múltipla calibração de câmeras baseado num padrão invariante. A proposta deste método é o tema principal desta tese.

O método apresentado neste trabalho pode ser definido como híbrido, o que significa que é composto por dois tipos de métodos de calibração básicos: um método fotogramétrico e um de auto-calibração, que já foram descritos na seção 2.1.2.

A mistura desses dois tipos de métodos de calibração básicos é feita com os seguintes propósitos:

- O uso de um método fotogramétrico permite calcular uma estimativa inicial confiável dos parâmetros intrínsecos e dos coeficientes de distorção da lente para cada câmera do sistema. Esse tipo de método é muito útil porque permite nos recuperar uma estimativa inicial dos parâmetros intrínsecos e dos coeficientes de distorção da lente de câmeras cujas informações específicas eram desconhecidas. Um outro cenário no qual esta opção também é aplicável é quando temos câmeras que mudam constantemente o tipo de lente para poder cobrir um campo de visão maior ou menor.
- A utilização de um método de auto-calibração permite realizar o cálculo dos parâmetros extrínsecos que relacionam a posição de uma câmera com a de outra. A ideia deste tipo de método num sistema com múltiplas câmeras é realizar a calibração dos parâmetros extrínsecos de todas as possíveis combinações de duas câmeras que podem ser geradas para nosso sistema de  $n$  câmeras e, com esses resultados, encontrar a melhor calibração, duas a duas câmeras, que permita relacionar todas as câmeras do sistema, criando uma concatenação entre elas. A ideia de concatenação das calibrações dos diversos sistemas estéreo foi proposta no trabalho de [14], que se



mostrou muito útil quando foi preciso calibrar múltiplas câmeras que cobriam uma área de rastreamento muito ampla e bem distribuída.

A combinação do uso desses dois tipos de calibração define a estrutura base de um método híbrido. Esta estrutura permite uma maior flexibilidade e robustez no cálculo de todos os parâmetros de calibração, especialmente quando se tem equipamentos muito variados ou se precisamos cobrir áreas de rastreamento amplas ou pequenas.

Nas seções a seguir, serão descritos os diferentes processos que compõem a estrutura de nosso método.

#### **4.1. Descrição do método**

O método proposto neste trabalho surgiu da iniciativa de implementar uma técnica que permita calibrar simultaneamente um sistema formado por múltiplas câmeras. No começo, ao analisar o problema da múltipla calibração e os distintos métodos encontrados na literatura sobre o tema, percebemos que os métodos mais novos adotam uma estrutura híbrida, definida pelo uso simultâneo de dois tipos de calibração básica: a fotogramétrica e a de auto-calibração. Cada um desses dois tipos de calibração é utilizado para o cálculo de um tipo de parâmetro específico das câmeras. O método fotogramétrico é comumente aplicado para calcular os parâmetros intrínsecos e os coeficientes de distorção de cada câmera, e o de auto-calibração é útil para calcular os parâmetros extrínsecos.

Levando em consideração a flexibilidade e a robustez que uma estrutura híbrida oferece, decidimos modelar o nosso método seguindo esse conceito.

Outro fator que também chamou a nossa atenção foi a evolução do formato dos diferentes padrões de calibração usados em trabalhos referentes ao processo de múltipla calibração, os quais foram especialmente propostos por métodos de auto-calibração (Figura 19). Nós observamos que tanto os padrões unidimensionais como adimensionais descritos nas seções 2.1.3 e 2.1.4 têm sido os mais utilizados na proposta de novos métodos de múltipla calibração [4][7][8][14][22][28][32] e [37], assim como em possíveis evoluções desses métodos que são implementadas em produtos comerciais [1][39][13] e [26]. Percebemos que o motivo principal do uso desse tipo de padrão é sua ampla

visibilidade e fácil identificação quando capturado por múltiplas câmeras simultaneamente. Essa característica é importantíssima especialmente quando temos configurações de câmeras cujos campos de visão têm ângulos de interseção que impossibilitariam a captura simultânea de várias vistas válidas do padrão, como acontece, por exemplo, com padrões mais complexos como os clássicos padrões planares.



Figura 19: Evolução dos padrões (a) planar, (b) unidimensional, (c) adimensional.

Nessa evolução, especialmente do padrão unidimensional para o adimensional, algumas informações úteis foram perdidas, como a distância existente entre os diversos marcadores que compõem um padrão unidimensional. A redução no número de marcadores pode gerar um padrão de formato mais simples e fácil de capturar, porém, a diminuição das informações implícitas que o padrão pode fornecer afeta diretamente a última etapa do processo de múltipla calibração, que é a otimização dos parâmetros de calibração. As informações implícitas no formato do padrão ajudam a modelar algumas funções de controle que compõem a função de erro global, definida como um dos componentes básicos do algoritmo de otimização. Quanto mais informações temos em relação às características implícitas do padrão, melhor estará definida a função de erro. Outro problema que surge quando simplificamos o formato do padrão de calibração é a diminuição da robustez no reconhecimento do mesmo como padrão único. Isso é especialmente ressaltado na etapa de aquisição de dados, onde é feita a captura de várias vistas do padrão nas quais teremos que identificá-lo corretamente e correlacionar individualmente os marcadores que o compõem.

Como foi relatado nas seções 2.1.3 e 2.1.4 uma das desvantagens em comum, dos padrões unidimensional e adimensional, é a fácil contaminação da amostra de vistas capturadas com falsos marcadores, devido ao reconhecimento de objetos com formatos semelhantes que podem ser facilmente confundidos com os marcadores. Esse falso reconhecimento pode estar relacionado à falta de uma estrutura robusta que permita identificar o conjunto de marcadores que compõe o

padrão como um objeto único. Um exemplo é o caso do padrão planar, cuja estrutura definida pela distribuição dos pontos de referência na sua superfície pode servir como método de reconhecimento robusto, capaz de filtrar e descartar possíveis objetos muito semelhantes ao padrão.

Com base na análise das características e problemática acima descritas em relação à definição e ao reconhecimento de um padrão de calibração simples e robusto, nós propomos como parte de nosso método de múltipla calibração de câmeras, um novo tipo de padrão unidimensional definido como padrão invariante cuja estrutura é formada por um conjunto de quatro marcadores (Figura 20).

Este novo tipo de padrão tenta manter as vantagens dos padrões unidimensional e adimensional em relação à visibilidade e ao fácil reconhecimento e, ao mesmo tempo, procura dar robustez ao processo de reconhecimento do padrão como objeto único. A robustez vem da aplicação de um novo método de reconhecimento criado e testado especialmente para este tipo de padrão nos trabalhos de [16][33][20][21]. A principal contribuição desse novo processo de reconhecimento do padrão é a capacidade de identificar o padrão como um objeto único, o que depois permitirá a identificação individual dos marcadores que o compõem. Outra vantagem é poder usar mais de um desses padrões simultaneamente no mesmo espaço de rastreamento quando realizamos a captura da amostra para calibração. O processo de reconhecimento desse tipo de padrão é implementado apenas com a análise das informações 2D das posições dos marcadores em coordenadas da imagem, processo feito só na etapa inicial de aquisição de dados (Figura 20).

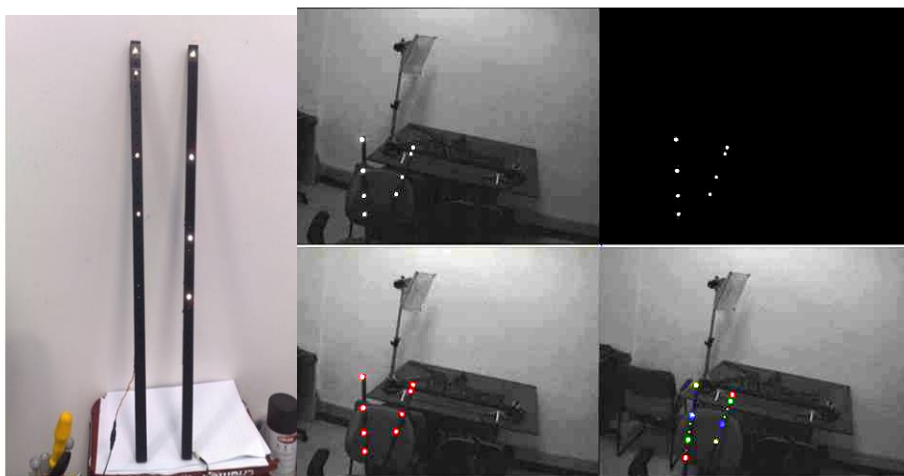


Figura 20: Padrões invariantes e processamento 2D para identificação individual.

Finalmente, a proposta desse novo tipo de padrão invariante gera informações implícitas, devido a seu formato, que permitirão usar todas as funções de controle definidas na seção 2.3.1, com as quais podemos modelar uma função de erro global robusta na etapa de otimização. No entanto, além dessas funções, propomos modelar a característica de colinearidade de nosso padrão, como mais duas funções de controle para o processo de otimização. Definimos duas funções de controle porque consideraremos que o teste de colinearidade pode ser efetuado tanto com as posições 2D dos marcadores extraídas da imagem, como também com a colinearidade dos marcadores após serem reconstruídos em 3D no espaço de rastreamento. Isso pode ser feito usando um algoritmo de triangulação, a estimativa dos parâmetros das câmeras e as posições 2D dos pontos correspondentes da amostra de calibração.

Para descrever um resumo dos processos que compreendem o método proposto neste trabalho, nós usaremos como guia o framework conceitual definido na seção 2.3:

- **Aquisição de dados:** sendo um método híbrido, nosso método usa dois tipos de padrão de calibração - um padrão planar, aplicado para realizar a calibração fotogramétrica, e um padrão unidimensional, baseado no padrão invariante proposto utilizado para executar o método de auto-calibração.
- **Calibração inicial:** aplicamos o método descrito na seção 2.2.1.3 para calcular uma estimativa inicial dos parâmetros intrínsecos e dos coeficientes de distorção da lente para cada câmera do sistema. Além disso, usamos o método de auto-calibração descrito na seção 2.2.2.4 para calcular uma estimativa inicial dos parâmetros extrínsecos de cada par de câmeras definidos como produto da combinação das  $n$  câmeras que formam o sistema de múltiplas câmeras.
- **Otimização da calibração:** devido à informação fornecida pelo padrão invariante, nosso método permite que a função de erro global seja definida usando todas as funções de controle descritas na seção 2.1.3, às quais é possível acrescentar duas funções de controle adicionais propostas com base na colinearidade dos marcadores que compõem o padrão invariante: colinearidade 2D e 3D.

Nas seções a seguir, serão descritas mais detalhadamente as etapas do framework conceitual acima.

#### 4.1.1. Aquisição de dados

Nesta primeira etapa do processo de múltipla calibração de câmeras, descreveremos como é feito o processo de captura de cada um dos padrões usados pelo método proposto.

##### 4.1.1.1. Captura do padrão planar

O primeiro processo a ser realizado nesta etapa de calibração é a captura de um número entre 100 e 200 vistas de um padrão planar semelhante ao formato de um tabuleiro de xadrez mostrado da Figura 21.



Figura 21: Captura do padrão planar para calibração fotogramétrica.

O processo de captura desse padrão planar é realizado individualmente em cada câmera do sistema. Para capturar e reconhecer este tipo de padrão a partir das imagens capturadas, nós usamos a implementação em C++ fornecida pela biblioteca OpenCV[12].

#### 4.1.1.2. Captura do padrão invariante

O segundo passo desta etapa de aquisição de dados é a captura de uma amostra de vistas do novo padrão invariante, que é constituído por quatro marcadores colineares. Esse tipo de padrão foi inicialmente proposto dentro da área de reconhecimento de padrões, como definido nos trabalhos de Meer et al. [20] [21]. Acharmos a teoria por trás do reconhecimento desse padrão interessante, e a analisamos e testamos num ambiente onde a tarefa seria modelar e usar esse tipo de padrão como padrões usados na calibração de câmeras. Alguns trabalhos prévios definidos sobre esses testes foram descritos em Loaiza et al.[16][33].

A teoria sobre características projetivas invariantes propõe que a relação definida pelo *cross ratio* calculado a partir de quatro pontos colineares gera um tipo de identificador único que permite diferenciá-lo de outro grupo de quatro pontos colineares cujas distâncias entre os pontos que o compõem sejam totalmente diferentes das distâncias do primeiro conjunto. A relação definida pelo *cross ratio* para um grupo de quatro pontos colineares foi estudada por Meer et al. [20] e, baseado nessa relação, ele propôs uma técnica denominada “*P<sup>2</sup>-invariant*” (Apêndice B), que propõe uma forma de gerar um identificador único para este tipo de padrão colinear. O identificador proposto por Meer et al. [20] foi testado por nós em Loaiza et al. [16] e ficou sendo definido não como sendo um valor único, e sim como um intervalo contendo um valor mínimo e um máximo.

O motivo pelo qual o valor do identificador é melhor definido como um intervalo vem de imprecisões geradas no processo de extração das posições 2D dos pontos que formam o padrão colinear na imagem. Leves variações do valor da posição 2D recuperado a partir de uma imagem geram pequenos ruídos no cálculo do valor do *cross ratio* e afetam o processo de geração do valor do identificador “*P<sup>2</sup>-invariant*”. No entanto, como foi testado em [16], essas variações não são muito distantes e podem ser mensuradas dentro de um intervalo totalmente único em comparação a outro padrão colinear.

A característica de poder criar um identificador único (*ProjInvID*) para esse tipo padrão permite que este seja mais robusto se comparado aos padrões unidimensional e adimensional mostrados na Figura 19.

Para poder então calcular o valor do *ProjInvID* de cada padrão invariante, precisa-se realizar uma etapa de treinamento onde calcularemos os valores do intervalo único definido como identificador único para cada padrão.

Portanto, o processo de captura da amostra de vistas para o padrão invariante é dividido em duas fases: treinamento e captura real.

#### 4.1.1.2.1. Treinamento do padrão invariante

O treinamento do padrão invariante consiste na captura de várias vistas do padrão colinear onde será calculado o valor do identificador “*P<sup>2</sup>-invariant*” para cada vista analisada. Obtido este resultado, armazenaremos o menor e o maior valor calculado para o identificador “*P<sup>2</sup>-invariant*”, que será usado como valores mínimo e máximo de um intervalo definido como o identificador único do padrão “*ProjInvID*”.

Nesta primeira parte do treinamento, é importante fazer uma observação adicional: no trabalho inicial de [20] onde foi definido o identificador “*P<sup>2</sup>-invariant*” não foram levados em conta os problemas que podem surgir por causa da distorção da lente da câmera, que por exemplo no caso de lentes grande angular podem distorcer significativamente a característica colinear dos pontos. Nos trabalhos [20] e [21], a distorção das lentes era baixa e não deteriorou o processo de treinamento e reconhecimento de nosso padrão colinear. Porém, propomos neste trabalho um caso singular ao usarmos lentes com grande distorção, como por exemplo uma lente grande angular. Nesse caso nossa proposta é executar apenas a etapa de treinamento e a posterior captura do padrão se tivermos a informação necessária para corrigir a distorção da lente. Isso é possível se considerarmos que podemos obter o valor dos coeficientes de distorção da lente com a calibração inicial desses parâmetros feita a partir da amostra de vistas do padrão planar descrita na seção anterior. Com essa consideração continuaremos a relatar como é o processo de treinamento do padrão invariante, que é descrito abaixo e visualmente exemplificado na Figura 22:

1. Em cada vista do padrão extrair e calcular as coordenadas 2D de cada marcador que compõe o padrão colinear. Esse processo é semelhante ao definido no final da seção 2.1.1 (Figura 22).
2. Avaliar se há quatro pontos 2D identificados na imagem.
3. Testar se os quatro pontos têm certo grau de colinearidade. Para isso, é usado um algoritmo de aproximação de linha (*fitting line*) e definido como limiar uma distância máxima permitida de três pixels da linha calculada a qualquer marcador do padrão.
4. Calcular o valor do identificador “ $P^2$ -invariant” para o padrão colinear.
5. Armazenar o menor e o maior valor achado no cálculo do valor em cada vista ( $P^2$ -invariant<sub>min</sub>,  $P^2$ -invariant<sub>max</sub>).
6. Testar todas as vistas, e definir como identificador do padrão colinear o intervalo definido pelos valores mínimo e máximo do identificador “ $P^2$ -invariant” calculados nas  $n$  vistas capturadas do padrão ( $ProjInvID = [ P^2$ -invariant<sub>min</sub>,  $P^2$ -invariant<sub>max</sub> ]).

Por fim, repetimos o processo anterior para cada câmera e para cada novo padrão colinear a ser utilizado. O número mínimo de vistas capturadas para o treinamento em nossa implementação foi de 100 vistas para cada padrão analisado. Na Figura 22, mostramos algumas vistas capturadas durante processo de treinamento de um mesmo padrão invariante feito simultaneamente num sistema formado por quatro câmeras.





Figura 22: Processo de treinamento feito simultaneamente num sistema de 4 câmeras.

Um último detalhe no processo de treinamento do padrão invariante é definir o reconhecimento individual de cada marcador que o compõe. Para realizar essa identificação, definimos uma heurística simples baseada na proporção da distância entre os dois primeiros e dois últimos marcadores em relação à distância de todo o padrão, tudo em 2D. Como é possível observar na Figura 23, montamos nossos marcadores de tal forma que eles não tenham a mesma distância entre eles. Além disso, também foi feito um distanciamento menor entre os primeiros marcadores ( $A, B$ ) em relação aos dois últimos ( $C, D$ ). Esse distanciamento diferenciado permite definir como primeiro marcador aquele que fica no extremo do padrão que faz parte dos dois marcadores que definem a menor distância. Isso gera como consequência uma proporção menor dessa distância em relação à distância total definida pelos quatro marcadores. Nós escolhemos não trabalhar

comparando as distâncias porque, como é bem conhecido nesse tipo de projeção 3D-2D, as distâncias não são preservadas numa projeção em perspectiva, porém quando testadas as proporções geradas pelas distâncias  $AB/AD < DC/AD$  (Figura 24), estas se mostraram como indicadores referenciais muito úteis para definir o primeiro e último marcador do padrão. Então comprovou-se que nossa heurística gera resultados confiáveis quando aplicada depois da identificação grupal acertada do arranjo de marcadores colineares usando o valor do *ProjInvID*. Só depois de os marcadores estarem bem identificados como grupo procede-se a calcular a heurística descrita acima para identificar e etiquetar cada marcador do padrão, tal como é mostrado na Figura 23.

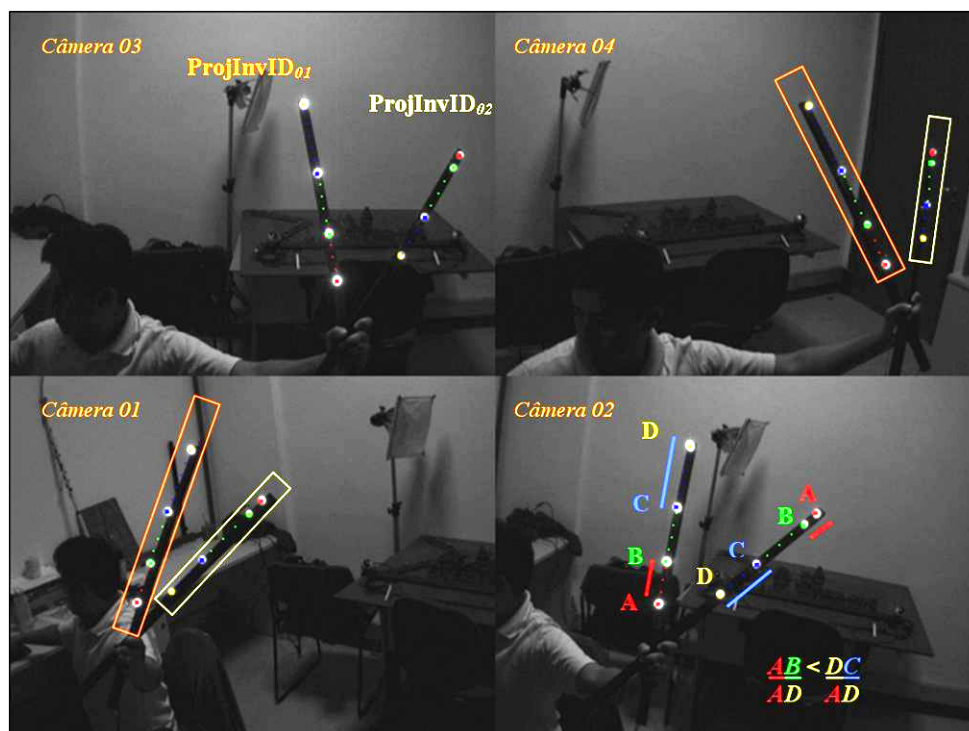


Figura 23: Identificação grupal e etiquetagem individual dos marcadores que compõem os padrões invariantes.

Após termos treinado todos os padrões colineares a serem utilizados como padrões de calibração, e termos definido como é feita a identificação individual de cada marcador do padrão, podemos prosseguir para a etapa seguinte, referente à captura da amostra de calibração simultânea em todas as câmeras que nos permita executar o método de auto-calibração.

#### 4.1.1.2.2. Captura da amostra do padrão invariante

Diferentemente da etapa de treinamento, a captura da amostra de vistas do padrão ou padrões invariantes é feita de forma simultânea por todas as câmeras. Isso significa que movimentaremos nossos padrões pelo espaço de rastreamento e, à medida que eles são reconhecidos simultaneamente pelas diferentes câmeras, serão preenchidos buffers contendo a informação de pontos correspondentes definidos pelo reconhecimento simultâneo de um mesmo padrão nas diferentes câmeras. O processo de captura é descrito a seguir:

1. Carregar os valores dos identificadores  $ProjInvID_i$  ( $i=1\dots n\underline{u}mero\ de\ padr\u00f5es$ ) de cada padrão colinear previamente treinado em cada câmara  $CameraID_j$  ( $j=1\dots n\underline{u}mero\ de\ c\u00e2meras$ ).
2. Extrair e calcular as coordenadas 2D de cada marcador que comp\u00f5e um poss\u00edvel padr\u00e3o colinear, processo \u00e9 semelhante ao mostrado na Figura 22.
3. Avaliar se temos  $n\underline{u}mero\ de\ marcadores\ 2D \geq 4$  corretamente identificados na imagem.
4. Gerar combina\u00e7\u00f5es de quatro marcadores a partir dos  $N$  pontos reconhecidos na imagem  $C_4^N$ .
5. Para cada uma das combina\u00e7\u00f5es:
  - a. Testar se os quatro marcadores 2D da combina\u00e7\u00e3o t\u00eam certo grau de colinearidade. Para isso, usamos um algoritmo de aproxima\u00e7\u00e3o de linha (*fitting line*) e definimos como limiar uma dist\u00e2ncia m\u00e1xima permitida de tr\u00eas pixels entre a linha calculada pelo algoritmo e qualquer marcador do padr\u00e3o.
  - b. Se o padr\u00e3o for colinear, calcular o valor do identificador “ $P^2$ -invariant”, se n\u00e3o for, continuar com a pr\u00f3xima combina\u00e7\u00e3o.
  - c. Calcular e comparar o valor do identificador “ $P^2$ -invariant” do padr\u00e3o colinear com cada um dos identificadores  $ProjInvID_i$  ( $i=1\dots n\underline{u}mero\ de\ padr\u00f5es$ ) dos padr\u00f5es previamente treinados.
  - d. Se o padr\u00e3o colinear tiver um identificador considerado semelhante a algum identificador treinado, armazenar as

posições 2D dos marcadores que o compõem e ativar um indicador  $Match_{ij}$  que detalhe que o padrão  $ProjInvID_i$  foi identificado na câmera  $CameraID_j$ . Se o identificador não for semelhante aos  $ProjInvID_i$  carregados, continuar com a próxima combinação.

6. Após analisadas todas as vistas capturadas simultaneamente nas  $m$  câmeras num mesmo tempo  $t$ , realizaremos as correspondências entre padrões corretamente identificados entre todas as câmeras.
  - a. Para cada par de câmeras  $CameraID_x$  e  $CameraID_y$ , onde  $x \neq y$ , verificar se o mesmo indicador  $Match_{ix}$  e  $Match_{iy}$  para o padrão  $i$  está ativo nas duas câmeras.
  - b. Se o indicador  $Match_{ij}$  estiver ativo nas duas câmeras,  $x$  e  $y$ , então armazenar os valores das coordenadas 2D dos marcadores do padrão colinear identificado em ambas as câmeras num buffer de pontos correspondentes  $CorrespondencePoints_{x,y}$ .
7. Após completar um número de  $n$  vistas de padrões invariantes colineares devidamente correlacionadas entre as câmeras  $x$  e  $y$ , terminar o processo de captura da amostra entre esse par de câmeras.
8. Após capturadas as  $n$  vistas entre todas as combinações de duas câmeras de nosso sistema, terminar o processo de captura da amostra entre todas as câmeras.

Finalmente, com isso capturamos nossa amostra de calibração para executar o método de auto-calibração entre cada par de câmeras. Cabe ressaltar que em nossa implementação foram utilizados dois tipos de padrões invariantes (Figura 24), com os quais foi capturada essa amostra de calibração. O motivo principal de usar só dois padrões invariantes é o custo computacional gerado quando se usam mais de dois padrões, o que equivale a usar (*número de padrões*  $\times 4 = N$ ) marcadores. Se analisarmos especialmente o passo 4 da descrição do processo de captura acima, observaremos que a geração de combinações  $C_4^N$  - definida para encontrar os conjuntos de quatro marcadores colineares a partir dos  $n$  marcadores detectados na imagem - pode crescer exponencialmente à medida que aumentamos o número de padrões invariantes. Por esse motivo, preferimos limitar o uso de dois padrões invariantes simultâneos para capturar a amostra de

calibração. O número de vistas capturadas dos padrões invariantes entre os pares de câmeras variou de 50 até 2500 vistas do padrão.



Figura 24: Captura simultânea dois padrões invariantes num sistema de quatro câmeras.

#### 4.1.2. Calibração inicial

Nesta segunda etapa do processo de múltipla calibração de câmeras, descreveremos quais algoritmos foram usados para realizar os dois tipos de calibração básica escolhidos pelo método proposto. Assim, dividiremos esta etapa em duas partes, correspondentes às descrições dos métodos de calibração fotogramétrica e de auto-calibração.

##### 4.1.2.1. Calibração fotogramétrica

Para implementar o método de calibração fotogramétrica, utilizamos como referência o método de Zhang [40], cuja descrição teórica foi detalhada na seção 2.2.1.3 Para tal usamos a implementação deste método fornecida pela biblioteca em OpenCV[12], e os parâmetros recuperados são:

- Parâmetros intrínsecos:  $(f_x, f_y)$  e centro da imagem  $(u_0, v_0)$  definidos no modelo de câmera descrito no apêndice A.
- Coeficientes de distorção da lente: distorção radial  $(k1, k2)$  e tangencial  $(p1, p2)$  também descritas no apêndice A.

A informação necessária para executar o método de calibração fotogramétrico é fornecida pela amostra de pontos referenciais obtidos a partir da captura e do processamento das vistas do padrão planar (Figura 21). Nós usamos amostras contendo entre 50 e 200 vistas do padrão planar, que foram capturadas entre 30 e 60 cm de distância das câmeras.

#### 4.1.2.2. Auto-calibração

O método de auto-calibração usado neste trabalho é o mesmo descrito por Dorfmüller-Ulhaas em [7][8], e cuja descrição teórica foi feita na seção 2.2.2.4. Este processo pode ser definido como a calibração dos  $m$  sistemas estéreos definidos pelo número de combinação  $C_2^n$  onde  $n$  é o número de câmeras que formam o sistema. Os parâmetros extrínsecos recuperados são:

- Matriz de rotação  $R_i$ , onde  $i = 1...m$  e  $m = C_2^n$ . Esta matriz pode ser modelada pela composição de três ângulos de rotação de Euler definidos como  $(\alpha, \beta, \gamma)$ . Essa simplificação é muito usada especialmente para diminuir o número de parâmetros na etapa de otimização.
- Vetor translação  $T_i$ , onde  $i = 1...m$ .

A informação necessária para a execução do método de auto-calibração é fornecida pela amostra de pontos referenciais obtidos a partir da captura e do processamento das vistas do padrão invariante e da estimativa inicial dos parâmetros intrínsecos e coeficientes de distorção da lente, calculados com o método de calibração fotogramétrica.

Um processo importante dentro no método de auto-calibração é o cálculo da matriz fundamental. O algoritmo utilizado para descrever esse cálculo foi o clássico algoritmo dos 8 pontos, considerado como um método linear que produz uma boa solução para o valor da matriz fundamental na presença de 8 pontos correspondentes. Porém, nós temos um sistema de equações sobredimensionado por causa do grande número de pontos correspondentes contidos na amostra

capturada com o padrão invariante, por esse motivo, é necessário usar um método mais robusto para o cálculo da matriz fundamental.

Uma estratégia usada para dar robustez ao cálculo da matriz fundamental é usar um algoritmo iterativo em conjunto com um método linear, como o definido pelo algoritmo de 8 pontos. Um exemplo dessa estratégia é mostrada nos trabalhos [34] e [31], onde utiliza-se o algoritmo de RANSAC [9] em conjunto com o algoritmo de 8 pontos. O novo algoritmo aplicado para realizar o cálculo da matriz fundamental mistura o uso do algoritmo dos 8 pontos e o insere na rotina de otimização iterativa definida como parte da execução do algoritmo de RANSAC. Uma vantagem dessa estratégia é que o algoritmo de RANSAC determinará a cada iteração, possíveis pontos erroneamente correspondidos (*outliers*), o que é muito útil. A implementação deste algoritmo pode ser encontrada na biblioteca OpenCV [12] e foi a utilizada em nosso cálculo da matriz fundamental no método proposto neste trabalho.

Uma vez calculada a matriz fundamental, o passo seguinte é a extração da matriz essencial usando os parâmetros intrínsecos calculados na calibração fotogramétrica. É a partir da matriz essencial que recuperaremos a orientação e a posição de uma câmera em relação a seu par correspondente. O processo de auto-calibração é realizado em cada um dos  $m$  ( $= C_2^n$ ) sistemas de câmera em estéreo definidos em nosso sistema. Na implementação, todos os processos referentes à recuperação da posição e da orientação das câmeras foram implementados seguindo os algoritmos e os detalhes descritos em [7][8][10][18] e cuja explicação detalhada foi já descrita na seção 2.2.2.4 desta tese.

Por fim, terminado o cálculo de uma solução inicial para todos os parâmetros das câmeras, procederemos com a próxima etapa do processo de calibração, que é a otimização desses parâmetros, descrita na seção seguinte.

#### **4.1.2.3. Otimização dos parâmetros**

Nesta terceira etapa do processo de calibração, descreveremos como o método proposto neste trabalho usa toda a informação implícita relacionada às características do formato do padrão invariante para definir e modelar distintas funções de controle com as quais montaremos uma função de erro global, usada

na otimização dos parâmetros das câmeras. Como foi descrito na seção 2.1.3.1, as funções de controle são modeladas com base em:

- Estimativa ou medição de certas características extraídas do formato do padrão, como, por exemplo medição e comparação de distâncias 3D entre marcadores.
- Estimativa ou medição de certas características relacionadas ao processo de calibração e à informação extraída da amostra de calibração, como: reprojeção 3D-2D, reprojeção 2D-2D e reconstrução 3D dos marcadores.

Para realizar a estimativa ou medição das características relacionadas ao padrão, são usadas funções matemáticas que internamente são compostas e dependem dos valores dos parâmetros das câmeras que desejamos otimizar. Quanto mais informações podemos extrair do formato do padrão, melhor será a modelagem matemática das características nele implícitas.

Portanto, as funções de controle então são criadas para quantificar e controlar o valor correto dessas características quando são testadas novas soluções para os parâmetros das câmeras geradas durante a execução do algoritmo de otimização.

#### **4.1.2.3.1. Definição da função de erro global**

A maioria dos algoritmos de otimização precisa de uma função de erro global para guiar e salvar as melhores mudanças nos valores dos parâmetros que estão sendo otimizados. Os trabalhos apresentados no capítulo 3 definem suas funções de erro baseados nas funções de controle expostas na seção 2.1.3.1. Nem todas as funções de controle são usadas simultaneamente em alguns desses trabalhos, isso porque, em alguns casos, o padrão de calibração não permite modelar todas elas. Por exemplo, no caso do padrão adimensional usado por Svoboda et al. [32], não é possível usar a função objetivo relacionada à extração de medidas, uma vez que só temos um único marcador como referência. Já no caso do padrão unidimensional usado por Borghese [4], por ser constituído por dois marcadores, podemos extrair a distância entre os marcadores e portanto modelar a função objetivo de recuperação de medidas.



Em nosso método, as características do padrão invariante permitem que todas as funções de controle da seção 2.1.3.1 sejam modeladas: re-projeção 3D-2D, extração de medidas reais 3D entre marcadores do padrão e re-projeção 2D-2D. Essas três funções objetivo costumam ser as mais utilizadas para montar a função de erro global no algoritmo de otimização. No entanto, como foi definido pelo método proposto, nós usamos um tipo de padrão de calibração que, além das características necessárias para modelar as funções objetivo comumente usadas, contém mais uma informação adicional entre os marcadores que o compõem: a colinearidade dos marcadores.

Esse tipo de padrão fez uma primeira contribuição na etapa de captura, rastreamento e identificação grupal e individual dos marcadores que o compõem, e nesta etapa de otimização dos parâmetros, o padrão fornecerá mais uma contribuição: a modelagem da medida de colinearidade que seus marcadores devem manter como mais duas funções de controle que poderão ser inseridas para compor a função de erro global. Essa característica de colinearidade pode ser medida tanto no espaço 2D (na imagem) como no 3D (reconstruído no espaço de rastreamento). As duas funções objetivo para medir a colinearidade serão descritas a seguir.

#### **4.1.2.3.2. Colinearidade 2D e 3D**

Para medir a colinearidade 3D e 2D dos marcadores de nosso padrão, nós modelamos este cálculo como um problema de aproximação de linha. No caso da colinearidade 3D calculamos a melhor linha 3D que passe pelos marcadores reconstruídos em 3D que compõem nosso padrão e depois medimos a distância desses pontos à linha. No caso da colinearidade 2D, o processo é o mesmo descrito acima, só mudando os dados de entrada, que, nesse caso, são representados pelas coordenadas 2D dos marcadores que aparecem sobre a imagem.

Nossa estimativa é que a função de colinearidade 3D contribua para guiar a otimização dos parâmetros intrínsecos e extrínsecos, pois esses parâmetros são usados para a reconstrução 3D dos marcadores. Já a colinearidade 2D contribuirá para a otimização dos coeficientes de distorção das lentes, porque estes são os

responsáveis pela correção da posição das coordenadas 2D dos marcadores na imagem.

Para a implementação da função de colinearidade 3D e 2D, foi utilizado o método denominado “*Orthogonal distance regression*” [3], que permite achar a melhor solução para o problema de minimização de  $n$  distâncias ortogonais a uma linha definida por um conjunto de pontos.

Em nosso caso, os quatro marcadores que formam nosso padrão invariante definiram duas aproximações de linhas: uma linha  $L_{3D}$ , calculada usando os marcadores reconstruídos em 3D ( $P_{(X_w \ Y_w \ Z_w)}^{i=1...4}$ ), e uma segunda linha  $L_{2D}$ , que pode ser calculadas usando as coordenadas 2D na imagem dos mesmos marcadores ( $P_{(u \ v)}^{i=1...4}$ ).

As funções para colinearidade 3D e 2D são definidas como:

$$\mathcal{E}_{\text{Colinearidade 3D}} = \sum_{i=1...4}^n \text{dist} (L_{3D}, P_{(X_w \ Y_w \ Z_w)}^i) \quad (48)$$

$$\mathcal{E}_{\text{Colinearidade 2D}} = \sum_{i=1...4}^n \text{dist} (L_{2D}, P_{(u \ v)}^i) \quad (49)$$

Concluimos assim nossa função de erro global que será composta pelas seguintes funções de controle:

$$\mathcal{E}_{\text{Global}} = \left( \begin{array}{l} \mathcal{E}_{\text{Re proj. 3D-2D}} + \mathcal{E}_{\text{Re proj. 2D-2D}} + \mathcal{E}_{\text{dist. 3D}} + \\ \mathcal{E}_{\text{Colinearidade 3D}} + \mathcal{E}_{\text{Colinearidade 2D}} \end{array} \right) \quad (50)$$

Na implementação do método proposto neste trabalho, usamos os algoritmos de otimização Levenberg Marquardt [27][6] e Downhill simplex [24][27] para minimizar o erro estabelecido pela equação 50, na qual otimizaremos todos os parâmetros intrínsecos, extrínsecos e coeficientes de distorção das lentes. Para a implementação do algoritmo de otimização de Levenberg Marquardt, foi usada uma biblioteca em C++ chamada CMinpack[6].

## 4.2. Sumarizando o método

A seguir será mostrado um fluxograma que ilustra todos os processos que compõem o método de múltipla calibração de câmeras proposto neste trabalho. Também será apresentado um fluxograma mais detalhado para cada processo.

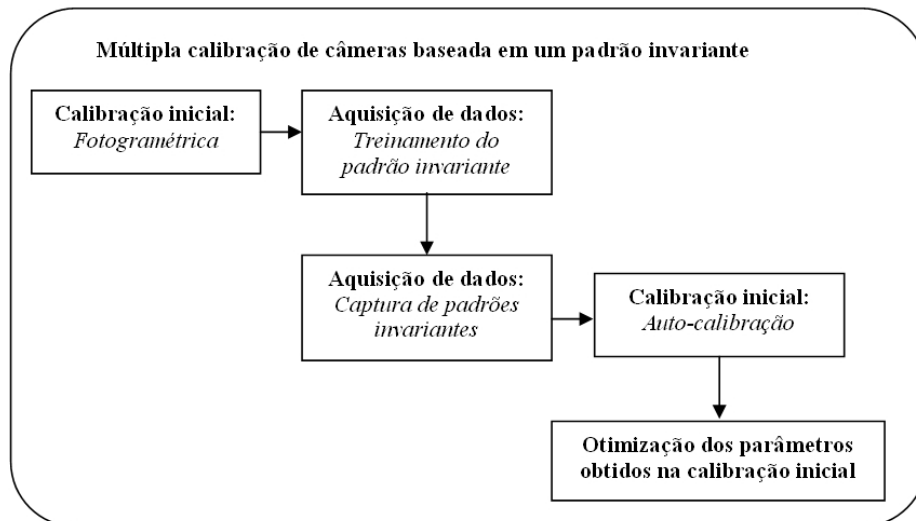


Figura 25: Fluxograma geral do método de múltipla calibração de câmeras proposto nesta tese.

O fluxograma apresentado na Figura 25 descreve, de uma forma global, os processos chave que compõem a base do método de múltipla calibração de câmeras proposto neste trabalho.

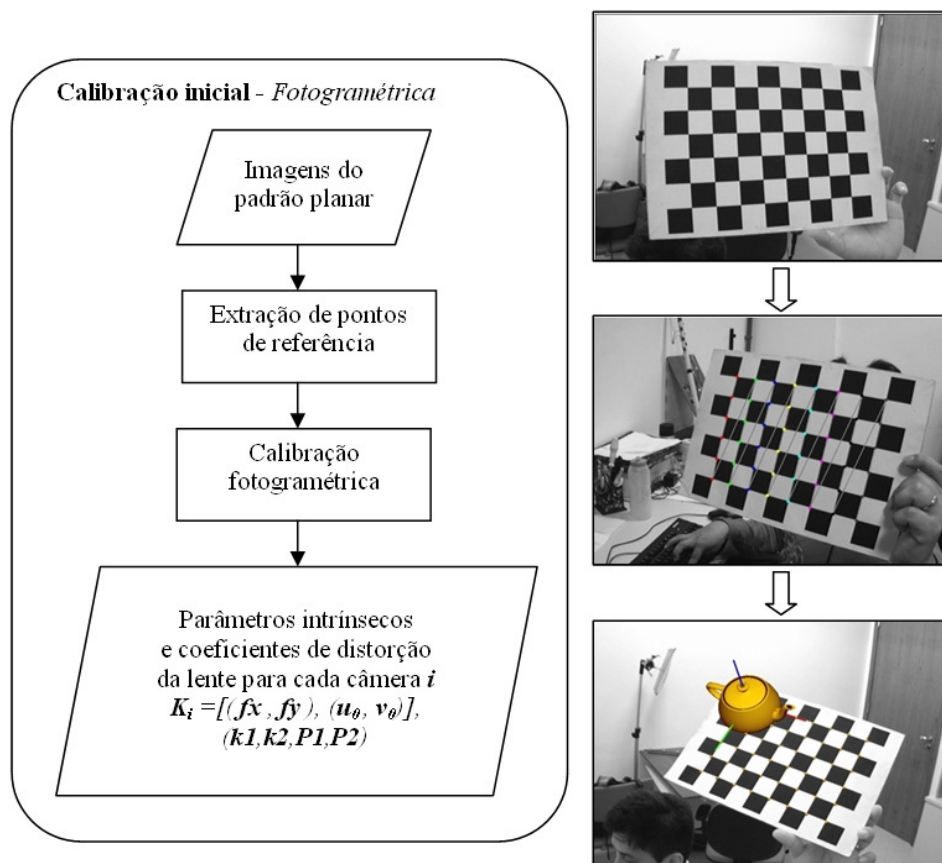


Figura 26: Fluxograma do processo de calibração fotogramétrica.

O primeiro passo na aplicação de nosso método é a calibração inicial, realizada em cada câmera do sistema com o método fotogramétrica (Figura 26).

O segundo passo consiste no treinamento dos padrões invariantes que serão usados no processo de múltipla calibração (Figura 27). Ressaltamos a importância de a calibração inicial fotogramétrica permitir calcular uma estimativa dos coeficientes de distorção da lente, para fazer a devida correção no momento da extração das posições 2D dos marcadores.

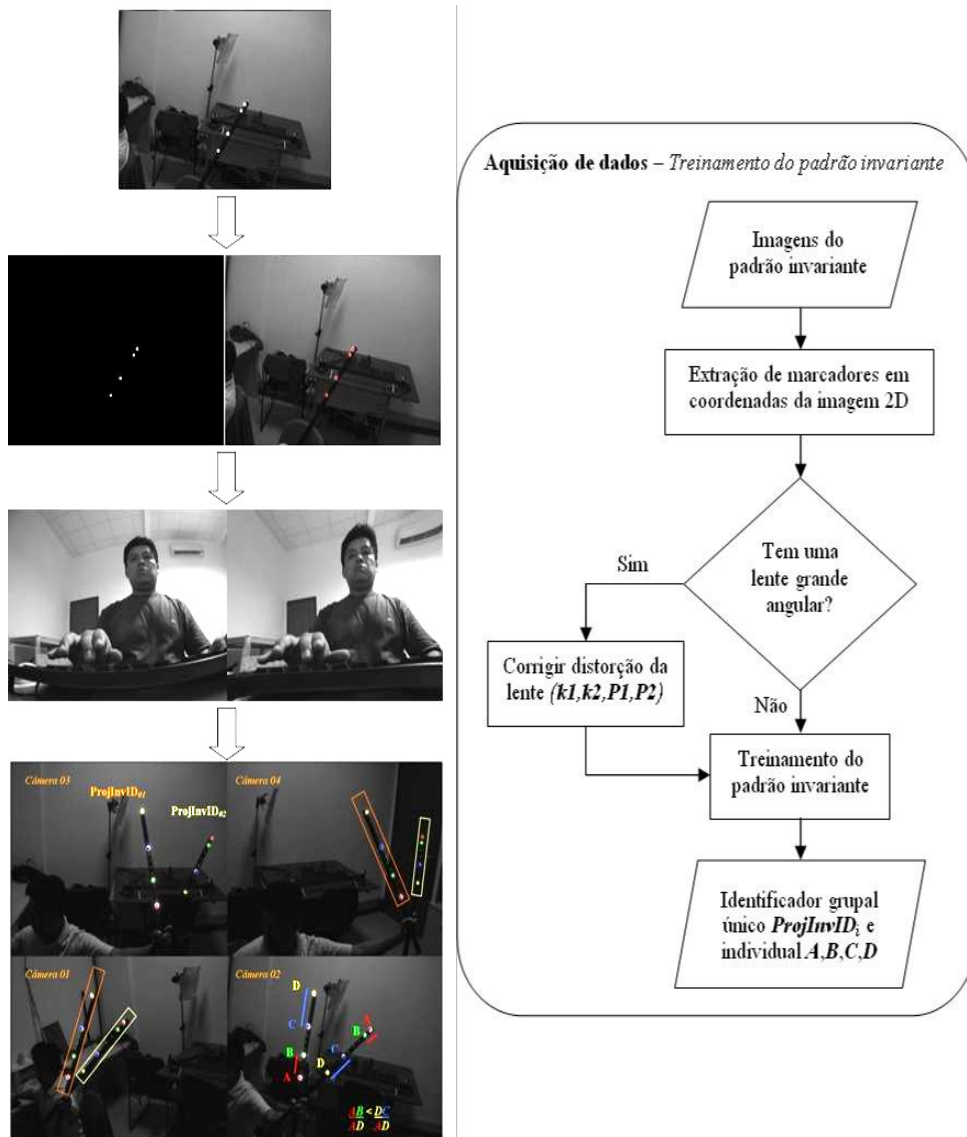


Figura 27: Fluxograma do processo de treinamento dos padrões invariantes.

O terceiro passo consiste na captura simultânea por todas as câmeras de vistas dos padrões invariantes previamente treinados (Figura 28). A partir dessas vistas, será gerada uma amostra de pontos correspondentes entre as imagens de cada par de câmeras. Essa amostra será usada na cálculo dos parâmetros extrínsecos para o qual o método de auto - calibração será aplicado.

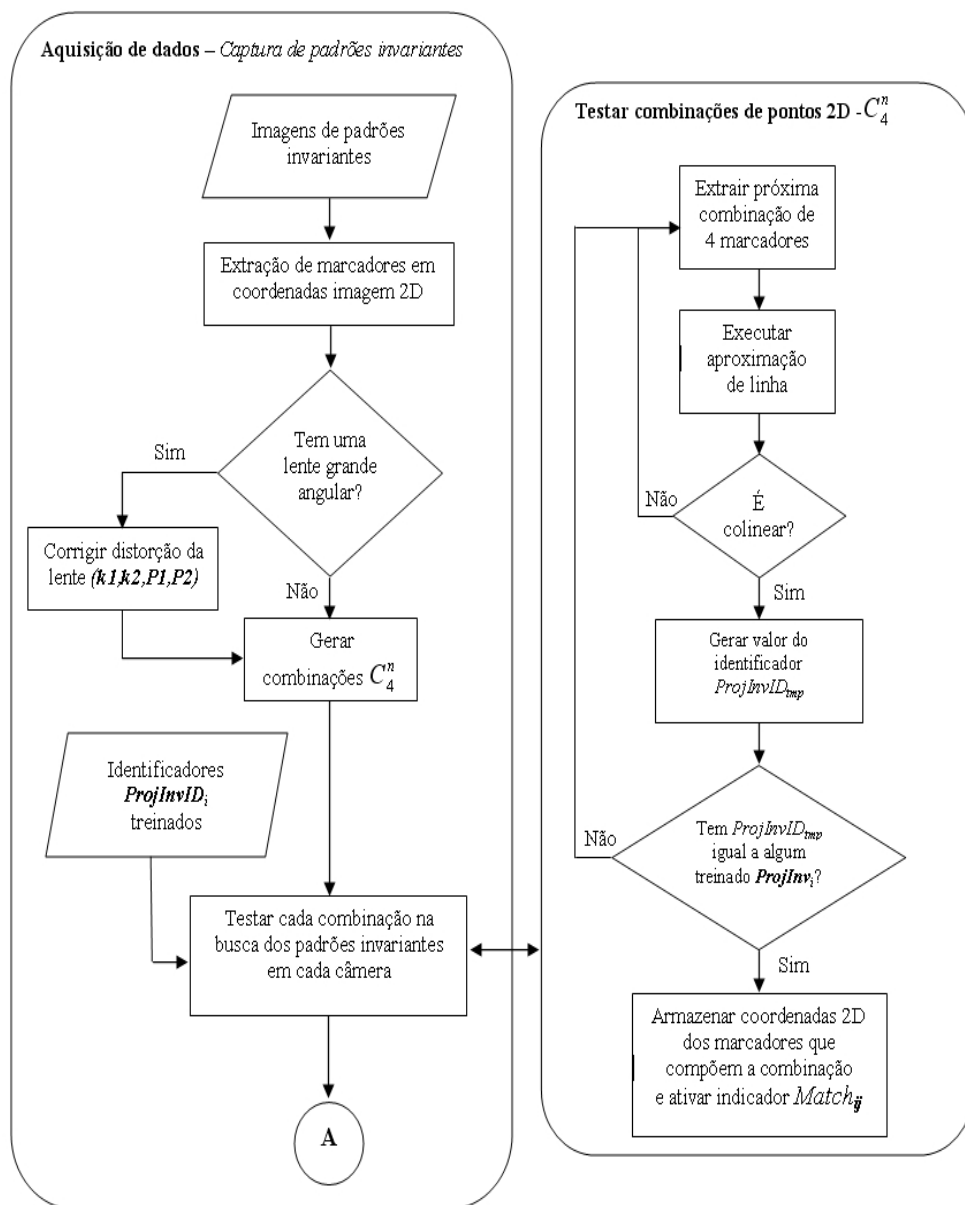


Figura 28: Fluxograma do processo de captura dos padrões invariantes previamente treinados.

Como resultado deste terceiro passo, nós teremos um conjunto de buffers  $CorrespondencePoints_{i,j}$  para cada par de câmeras  $(i,j)$  contendo pontos corretamente correspondidos entre elas (Figura 29). Esses pontos são definidos a partir das posições 2D dos marcadores que compõem nossos padrões invariantes e que foram extraídos a partir das vistas capturadas como amostra de calibração.

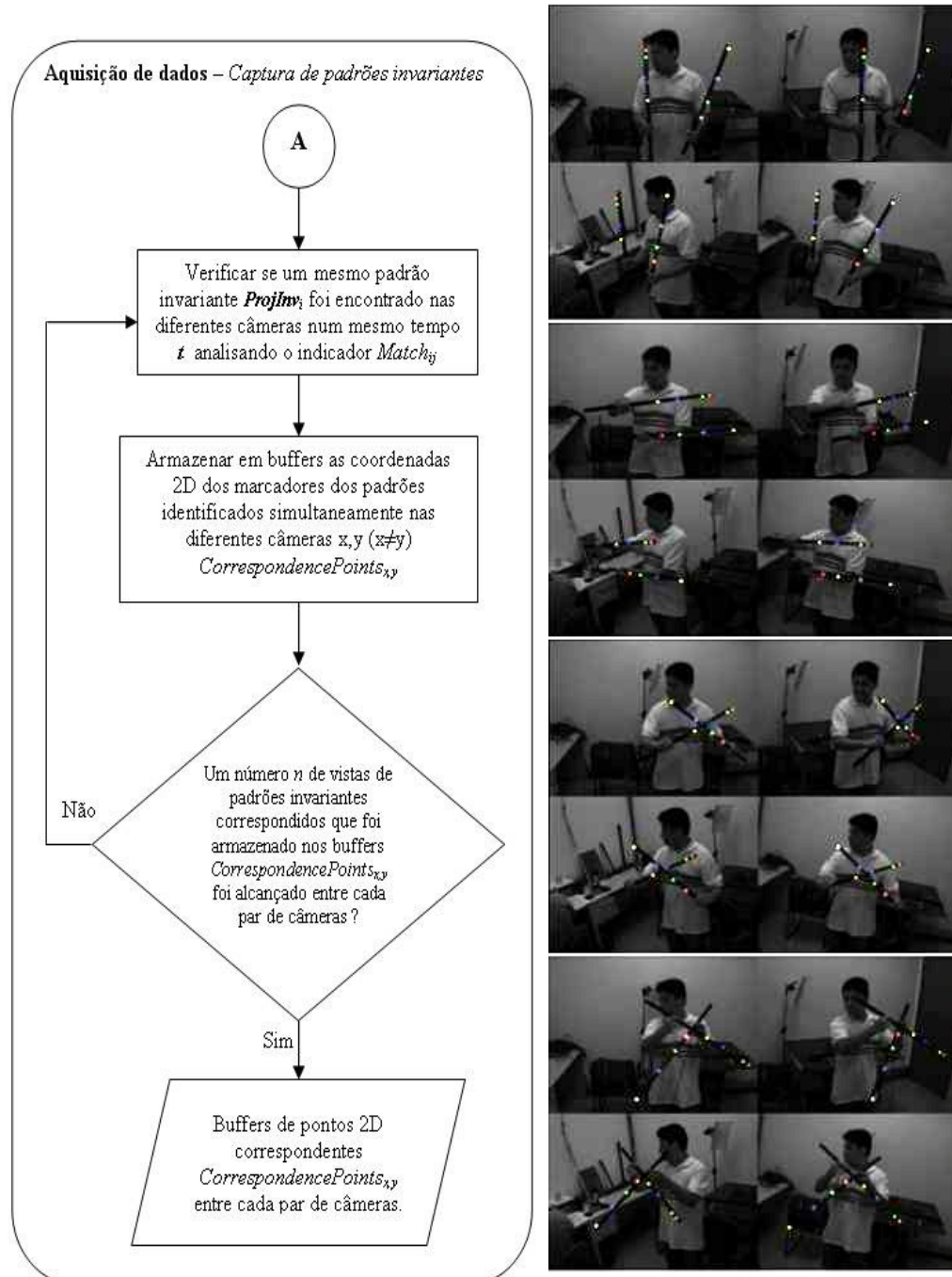


Figura 29: Fluxograma do processo de captura dos padrões invariantes previamente treinados (Continuação).

O quarto passo no processo de múltipla calibração consiste na aplicação do método de auto – calibração entre cada par de câmeras que formam nosso sistema (Figura 30). Como resultado recuperaremos os parâmetros extrínsecos que definem o posicionamento relativo entre cada par de câmeras. O último passo consiste na otimização de todos os parâmetros das câmeras previamente estimados pelos dois métodos de calibração inicial: fotogramétrico e auto–calibração, o resultado é a calibração final entre cada par de câmeras (Figura 31).

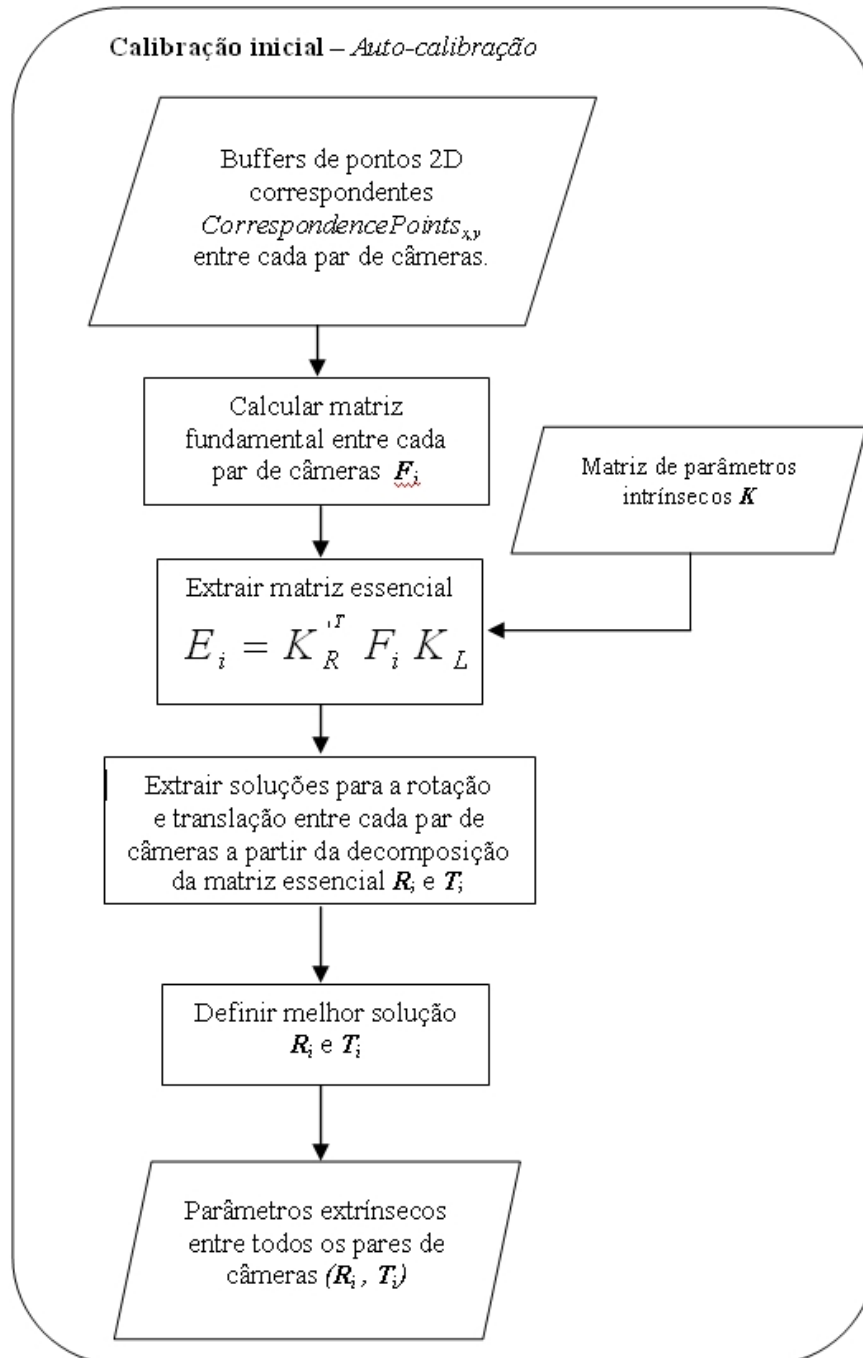


Figura 30: Fluxograma do processo de auto - calibração.

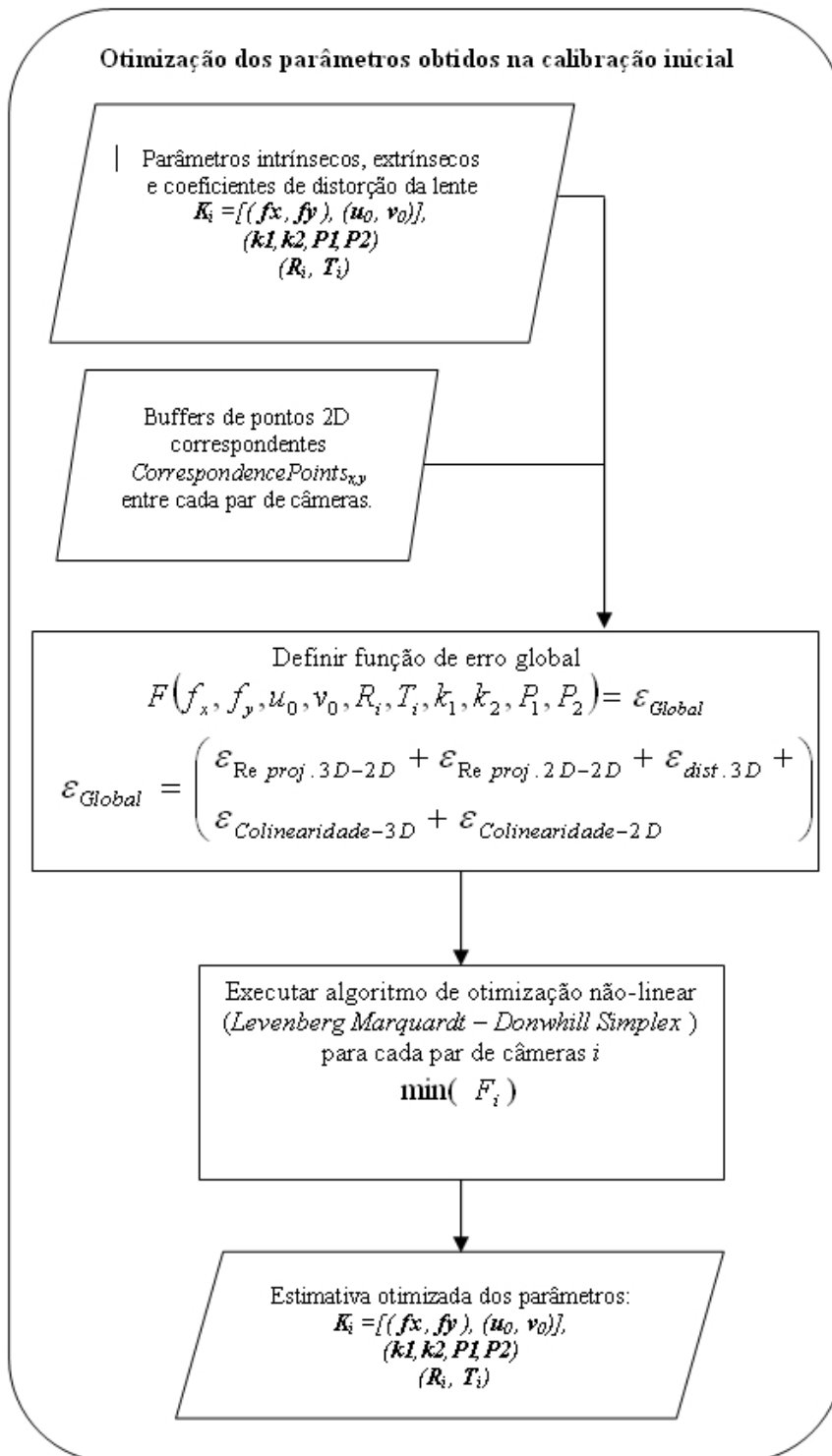


Figura 31: Fluxograma do processo de otimização final dos parâmetros das câmeras estimados na calibração inicial.

Com a exposição dos fluxogramas acima, descrevemos de uma forma resumida todo o processo de múltipla calibração de câmeras definido pelo método proposto neste trabalho. No capítulo seguinte, mostraremos alguns resultados da precisão alcançada na aplicação do nosso método na tarefa de calibração de um sistema de quatro câmeras montado por nós para testes.



## 5 Resultados

Neste capítulo apresentaremos os resultados de testes realizados usando o método de múltipla calibração de câmeras proposto neste trabalho, aplicado na calibração de dois sistemas formados por quatro câmeras. O objetivo desses testes é mostrar as contribuições que nosso método pode oferecer em comparação a outros trabalhos, como os citados no capítulo 3, aproveitando especialmente as características implícitas do novo padrão invariante proposto como componente importante do nosso método. Nossas contribuições tentam provar que:

- A característica de colinearidade que nossos padrões invariantes possuem contribui na obtenção de melhores resultados quando usada especificamente para aprimorar os resultados da otimização dos parâmetros das câmeras. Para isso, essa característica é modelada como mais duas funções de controle que serão inseridas na definição da função de erro global, componente chave do algoritmo de otimização. Nós compararemos os resultados obtidos incluindo ou não essas funções derivadas da colinearidade do padrão.
- A precisão alcançada na tarefa de recuperação de medidas 3D, após realizar a calibração das câmeras com nosso método, é melhor em comparação aos resultados obtidos com outros métodos de múltipla calibração de câmeras, como os apresentados academicamente em [4][28][32] e [37] e suas versões comerciais [1] e [13], que serão citados como referência para comparação com nosso método.

Para fazer a descrição dos nossos testes e dos resultados alcançados dividiremos este capítulo nas seguintes seções:

- Equipamento e ambiente de teste.
- Metodologia dos testes.
- Descrição dos resultados.

### 5.1. Equipamento e ambiente de teste

Para testar nosso método de múltipla calibração num sistema real de múltiplas câmeras montamos dois sistemas compostos por quatro câmeras cada um. A diferença entre os dois sistemas é o tipo de câmera que os compõem. Tanto as câmeras como seus acessórios adicionais são detalhados a seguir:

- 2 câmeras web, da marca Logitech, modelo QuickCam Pro 4000, com conexão USB 2.0, com taxa de captura de 30 fps. (Figura 32a).
- 4 câmeras monocromáticas, da marca Unibrain, modelo Fire-i, com conexão firewire, com taxa de captura de 30 fps. (Figura 32b e 32c)
- 3 cabos de 4.5m e 1 cabo de 1m para conexão firewire.
- O tipo de lente das câmeras tem distância focal de 4.3mm e ângulo de visão de 42° Horizontal, e 32° Vertical.

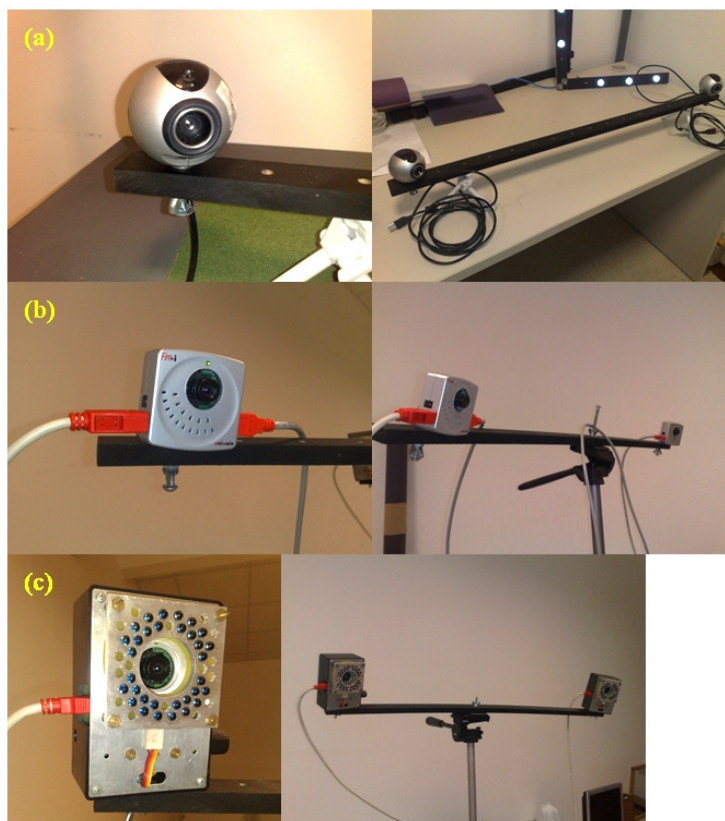


Figura 32: (a) Câmeras web Logitech, (b) e (c) Câmeras Fire-I

- Dois tripés sobre os quais foram construídas e posicionadas três varas de madeira para suporte das câmeras, o primeiro e segundo suporte (Figura 32b e 32c) permite colocar as câmeras a distância de 80, e o terceiro a 85 cm.



Figura 33: Três sistemas estéreo de câmeras formados pelas seis câmeras usadas para os testes.

Com o hardware descrito acima, foi possível montar dois sistemas de quatro câmeras, baseados em três sistemas estéreo de câmeras (Figura 33). Os sistemas de quatro câmeras são descritos a seguir:

- **Sistema X:** composto pelas quatro câmeras firewire (Figura 34a ).
- **Sistema Y:** composto por duas câmeras firewire e as duas câmeras web (Figura 34b ).

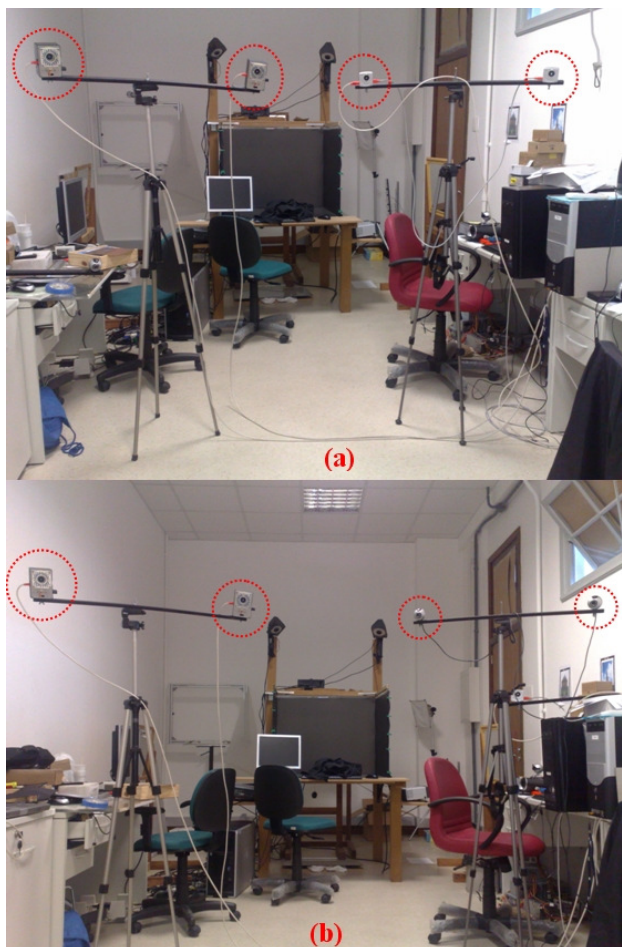


Figura 34: (a) Sistema X e (b) sistema Y de quatro câmeras.

Além das câmeras, foi necessário montar e construir nossos padrões invariantes formados por duas varas de madeira onde foram colocadas 4 lâmpadas pequenas alinhadas colinearmente (Figura 35). Também foi necessário criar um padrão planar, já mostrado na Figura 21, para a calibração fotogramétrica inicial de cada câmera.

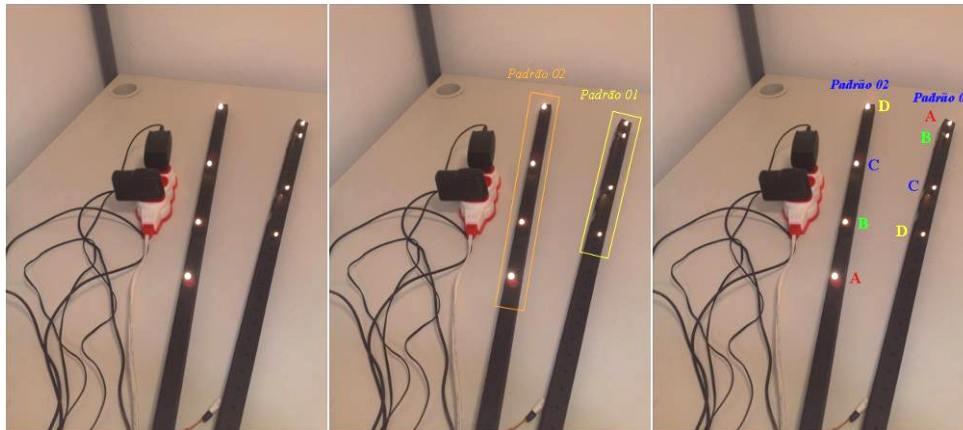


Figura 35: Dois padrões invariantes usados nos testes.

As distâncias a que foram colocadas as lâmpadas em cada padrão são:

- **Padrão 01:** AB (39.01 mm) , BC (140.196 mm), CD (99.156 mm).
- **Padrão 02:** AB (99.65 mm) , BC (139.505 mm), CD (180.565 mm).

Para executar os testes foi usado um computador composto por um processador Intel Core 2 Duo de 2.3Ghz, com 2GB de memória e foi necessária uma placa PCI de extensão de portas firewire para prover mais duas entradas firewire e assim poder conectar nossas quatro câmeras num só computador. Portanto, as câmeras foram conectadas ao computador por três saídas firewire, uma saída própria da placa mãe e duas fornecidas pela placa PCI de extensão utilizada. As duas saídas firewire da placa PCI foram conectadas individualmente a uma câmera Fire-i (Figura 32c), já a porta firewire da placa mãe serviu para conectar duas câmeras firewire em série (Figura 36). Isso foi possível porque esse modelo de câmera tem duas portas firewire que permitem sua conexão em série.

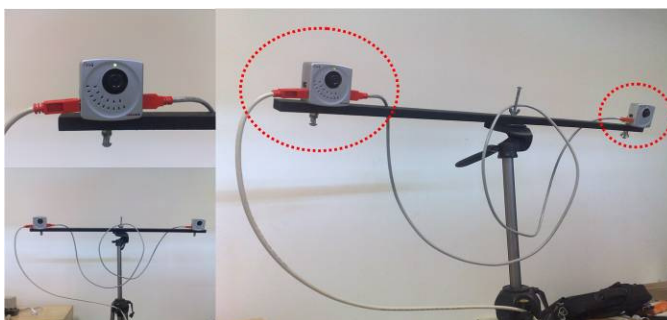


Figura 36: Duas câmeras firewire conectadas em série.

Porém, esse tipo de conexão em série não garante que as câmeras estejam sincronizadas por hardware nem software, tampouco a arquitetura usada para ligar as quatro câmeras a um só computador garante algum tipo de sincronismo.

Todo o processo de captura, processamento das imagens e execução do algoritmo de múltipla calibração de câmeras proposto neste trabalho foi feito num só computador, ressaltando principalmente que os processos de captura de vídeo e processamento das imagens vindas das quatro câmeras do sistema rodaram numa resolução de imagem de  $640 \times 480$  pixels mantendo uma taxa de captura constante de  $\sim 30$ fps em todas as câmeras.

Na Figura 34, mostramos os dois sistemas de quatro câmeras utilizados, os quais enxergam uma área de rastreamento com as seguintes dimensões  $2 \text{ m} \times 2 \text{ m} \times 1.8 \text{ m}$  que é um volume padrão similar a outros testes realizados em trabalhos como [4][7][8][28][37] e [29].

## 5.2. Metodologia dos testes

O principal teste a ser realizado neste trabalho é aquele que consiste em calibrar cada sistema de quatro câmeras, e em seguida analisar qual é o grau de precisão que alcançamos após calibrados os sistemas usando o método de múltipla calibração de câmeras proposto nesta tese.

Com essa ideia de teste foi escolhida uma estratégia dividida em duas etapas:

- A primeira etapa consiste em calibrar cada sistema de câmeras usando uma amostra *A* de vistas capturadas com dois padrões invariantes criados para os testes. Com essa amostra inicial realizaremos o cálculo dos parâmetros das câmeras aplicando o método de múltipla calibração de câmeras proposto neste trabalho.
- A segunda etapa é na captura de uma segunda amostra *B*, muito mais numerosa e melhor distribuída que a amostra *A*, com a qual avaliaremos a precisão alcançada com os parâmetros da calibração realizada na primeira etapa.

Este tipo de estratégia de teste que usa duas amostras é adotado com o objetivo de medir de forma imparcial os resultados obtidos a partir da calibração feita com nosso método. Por esse motivo, os resultados para comparação só serão

conseguidos a partir de uma amostra diferente da utilizada na etapa de calibração das câmeras.

Uma vez feita a captura das nossas amostras para teste  $A$  e  $B$ , podemos continuar nossa metodologia de teste concentrando-nos agora em definir uma forma de provar que a característica de colinearidade que nosso método tenta explorar gera melhores resultados em comparação com o uso de outras características convencionais, comumente modeladas usando os diversos tipos de padrões de calibração descritos no capítulo 2.

Nós percebemos que cada método de múltipla calibração descrito no capítulo 3 possui uma relação direta entre as características implícitas do padrão de calibração utilizado e a definição da função de erro global usada pelo algoritmo de otimização. Como era de esperar, a função de erro global é composta por um conjunto de funções objetivo definidas exatamente com base nas características implícitas que o padrão de calibração usado oferece. Daí surge nossa ideia de comparação; nós modelaremos as mesmas funções de erro globais definidas pelos trabalhos descritos no capítulo 3 e faremos uma primeira calibração das câmeras usando apenas essa função de erro que chamaremos de *convencional*. Depois, aproveitando a característica de colinearidade de nossos padrões invariantes, faremos uma segunda calibração das câmeras modelando uma nova função de erro composta pela função de erro *convencional* e pelas funções de controle derivadas da característica de colinearidade do nosso padrão invariante. Ambas as calibrações serão feitas usando as mesmas informações fornecidas pela amostra de teste  $A$ .

Com essas duas calibrações realizadas, o próximo passo será testar a precisão dos parâmetros das câmeras que cada calibração gera quando testados sobre uma segunda amostra  $B$ , de vistas dos nossos padrões. Compararemos o erro médio obtido ao avaliar cada uma das funções de controle comumente usadas nas funções de erro *convencionais*, e mostraremos que a inclusão da característica de colinearidade permitiu obter parâmetros de calibração que geram melhores resultados que os alcançados usando só as funções objetivo *convencionais*.

Essa metodologia de teste permite criar um ambiente de teste geral onde poderemos simular, especialmente na etapa de otimização, os mesmos processos de calibração de câmeras realizados em outros métodos propostos no capítulo 3.

A seguir descreveremos os oito tipos de funções de erro que foram modeladas para a comparação:

- **Função de erro 1:**

$$\mathcal{E}_{Global} = ( \mathcal{E}_{Reproj. 2D-2D} + \mathcal{E}_{dist. 3D} )$$

Esta função tenta simular o ambiente de calibração definido pelos métodos de Borghese et al. [4], Kurillo et al. [14] e Uematsu et al. [37], onde todos os parâmetros das câmeras são otimizados usando as duas funções objetivo de reprojeção 2D-2D e da distância entre os marcadores esféricos montados nos extremos de uma vara usada no padrão unidimensional de calibração.

- **Função de erro 2:**

$$\mathcal{E}_{Global} = ( \mathcal{E}_{Reproj. 2D-2D} + \mathcal{E}_{dist. 3D} + \mathcal{E}_{Colinearidade 3D} + \mathcal{E}_{Colinearidade 2D} )$$

É a nossa versão da função 1 inserindo as funções de controle derivadas da colinearidade do nosso padrão.

- **Função de erro 3:**

$$\mathcal{E}_{Global} = ( \mathcal{E}_{Reproj. 3D-2D} + \mathcal{E}_{dist. 3D} )$$

Esta função tenta simular o ambiente de calibração definido pelo método descrito por Bouguet [5], onde todos os parâmetros das câmeras são otimizados tendo como referência as funções de reprojeção 3D-2D e as distâncias entre os marcadores definidos pelo padrão.

- **Função de erro 4:**

$$\mathcal{E}_{Global} = ( \mathcal{E}_{Reproj. 3D-2D} + \mathcal{E}_{dist. 3D} + \mathcal{E}_{Colinearidade 3D} + \mathcal{E}_{Colinearidade 2D} )$$

É nossa versão da função 3 na qual são inseridas as funções de controle derivadas da colinearidade do nosso padrão.

- **Função de erro 5:**

$$\mathcal{E}_{Global} = ( \mathcal{E}_{Reproj. 3D-2D} + \mathcal{E}_{Reproj. 2D-2D} + \mathcal{E}_{dist. 3D} )$$

Esta função tenta simular um ambiente de calibração onde todas as funções de controle como: reprojeção 3D-2D, reprojeção 2D-2D e

distância entre marcadores podem ser utilizadas, definindo uma função de erro global convencional robusta.

- **Função de erro 6:**

$$\mathcal{E}_{Global} = ( \mathcal{E}_{Reproj. 3D-2D} + \mathcal{E}_{Reproj. 2D-2D} + \mathcal{E}_{dist. 3D} + \mathcal{E}_{Colinearidade 3D} + \mathcal{E}_{Colinearidade 2D} )$$

É a nossa versão da função 5, adicionando as funções de controle derivadas da colinearidade do nosso padrão.

- **Função de erro 7:**

$$\mathcal{E}_{Global} = ( \mathcal{E}_{Reproj. 3D-2D} )$$

Esta função tenta simular o ambiente de calibração definido pelos métodos descritos por Svoboda et al. [32] e Pintaric et al. [28], onde só os parâmetros extrínsecos das câmeras são otimizados tendo como função de erro a função de reprojeção 3D-2D. Uma única função de controle é modelada como função de erro global devido às poucas informações fornecidas pelo tipo de padrão de calibração adimensional usada nesse método.

- **Função de erro 8:**

$$\mathcal{E}_{Global} = ( \mathcal{E}_{Reproj. 3D-2D} + \mathcal{E}_{Colinearidade 3D} + \mathcal{E}_{Colinearidade 2D} )$$

É nossa versão da função 7, inserindo as funções de controle derivadas da colinearidade do nosso padrão.

Nas funções 1 – 6, serão comparados os erros obtidos em relação às funções de controle: reprojeção 3D-2D, reprojeção 2D-2D e distância entre marcadores, já entre as funções 7 – 8 só podem ser comparados os erros da função de reprojeção 3D-2D.

O número de vistas capturadas de nossos padrões invariantes para a amostra *A*, é de 250 vistas (  $250 \times 4 = 1000$  pontos de referência) entre os três pares de câmeras que podem ser definidos em cada sistema de quatro câmeras. Para a amostra *B* usada para a comparação final dos resultados, foram capturadas 1000 vistas (  $1000 \times 4 = 4000$  pontos de referência ) tanto para o sistema **X** quanto para o sistema **Y**, compostos por quatro câmeras cada um, montados para nossos testes.

Na seção seguinte, serão apresentadas algumas tabelas mostrando os resultados obtidos com os valores dos parâmetros das câmeras após serem



calibrados com a amostra *A*, assim como o erro alcançado quando a amostra *B* foi analisada com esses parâmetros.

### 5.3. Descrição dos resultados

O primeiro resultado que desejamos expor são os valores dos identificadores dos padrões invariantes ( $ProjInvID_i$ ) após treinados em cada uma das seis câmeras usadas em nossos testes. A tabela 1 mostra esses valores.

	Padrão 01	Padrão 02
<b>Câmera 1 - Firewire</b>	[2.0257, 2.0275]	[2.2137, 2.2204]
<b>Câmera 2- Firewire</b>	[ 2.0258, 2.0271]	[ 2.2131, 2.2197]
<b>Câmera 3- Firewire</b>	[2.0258, 2.027]	[2.2125, 2.2192]
<b>Câmera 4- Firewire</b>	[2.0265, 2.0272]	[2.2123, 2.2189]
<b>Câmera 1 - USB</b>	[2.0259, 2.0278]	[2.2159, 2.2222]
<b>Câmera 2 - USB</b>	[ 2.026, 2.0279]	[2.2153, 2.2217]

Tabela 1: Valores dos identificadores para cada padrão invariante usado nos testes.

Como podemos observar os valores de nossos identificadores  $ProjInvID_i$  são definidos por intervalos nos quais a diferença entre o valor mínimo e máximo é ínfima. Porém, mesmo assim é importante que essa diferença seja levada em conta para conseguir uma correta identificação dos nossos padrões.

Conforme exposto na seção anterior, para testar se as funções de colinearidade definidas a partir do formato dos nossos padrões invariantes trazem algum tipo de contribuição ao cálculo dos parâmetros das câmeras, propomos testar nosso método de calibração usando oito possíveis funções de erro global modeladas com base em funções de controle *convencionais* nas quais foram inseridas as funções de controle derivadas da colinearidade de nossos padrões invariantes. Nós calibramos os dois sistemas de quatro câmeras **X** e **Y** com cada uma das oito funções de erro global previamente definidas usando uma amostra *A* de 250 vistas de nossos padrões colineares, que foi capturada independentemente em cada sistema de câmeras. Em seguida, com os valores dos parâmetros das câmeras obtidos com as calibrações realizadas em cada um dos sistemas, realizamos um teste para medir a precisão alcançada ao avaliar uma segunda amostra *B* de vistas dos nossos padrões invariantes. O teste de precisão consistiu em avaliar os erros médios obtidos nas funções de controle *convencionais* descritas na seção 2.1.3.1, que são:

- Erro médio da reprojeção 3D-2D avaliada para toda a amostra *B* de 1000 vistas dos nossos padrões invariantes. Cada vista contém 4

marcadores, isso representa a reconstrução 3D de  $(1000 \times 4 = )$  4000 marcadores que serão reprojatados em cada plano da imagem tanto da câmera da esquerda como da câmera da direita, o que representa 8000 pontos analisados. Este erro é expresso em pixels.

- Erro médio da medição da distância 3D entre os marcadores do padrão colinear reconstruídos em 3D. Como temos 4 marcadores no padrão então há um total de três distâncias a serem medidas por padrão, que na amostra *B* representará um total de  $(1000 \times 3 =)$  3000 medidas de distâncias entre marcadores analisadas, este erro médio é expresso em milímetros.
- Erro médio da reprojeção 2D-2D definido como a distância de cada linha epipolar ao ponto correspondente. Como uma linha epipolar pode ser gerada para cada marcador presente na amostra *B*, então temos 4000 marcadores que geram 4000 linhas epipolares, mas isso só no sentido câmera da esquerda para câmera da direita. Adicionando os casos no sentido inverso temos um total de 8000 distâncias avaliadas. Este erro é expresso em pixels.

Nas tabelas 2 e 3 mostraremos o erro médio alcançado ao avaliar a amostra *B* de 1000 vistas de nossos padrões invariantes capturadas e analisadas em cada um de nossos sistemas de quatro câmeras *X* e *Y*:

Câmeras 1 e 2	Função 01	Função 02	Função 03	Função 04	Função 05	Função 06	Função 07	Função 08
Reprojeção 3D-2D (pixels)	0,403	0,391	0,444	0,422	0,402	0,391	0,369	0,371
Distância 3D (mm)	1,499	1,472	1,392	1,455	1,501	1,478		
Reprojeção 2D-2D (pixels)	0,806	0,781	0,889	0,844	0,805	0,782	0,738	0,7411
Câmeras 2 e 3								
Reprojeção 3D-2D (pixels)	0,426	0,422	0,439	0,4372	0,421	0,424	0,397	0,394
Distância 3D (mm)	1,691	1,664	1,719	1,713	1,677	1,666		
Reprojeção 2D-2D (pixels)	0,852	0,844	0,877	0,874	0,842	0,847	0,793	0,788
Câmeras 3 e 4								
Reprojeção 3D-2D (pixels)	0,133	0,132	0,182	0,177	0,128	0,127	0,14	0,135
Distância 3D (mm)	0,794	0,791	0,785	0,772	0,801	0,794		
Reprojeção 2D-2D (pixels)	0,267	0,264	0,364	0,355	0,256	0,254	0,28	0,27

Tabela 2: Resultados do erro médio medidos na amostra *B*, 1000 vistas, capturada para teste da calibração realizada do sistema *X* de quatro câmeras firewire.

Nas tabelas 2 e 3, os resultados da calibração realizada com as funções de erro *convencionais* (funções de número ímpar) são, na maioria das vezes, piores

em comparação com as obtidas pela calibração usando as funções modeladas com as funções de controle derivadas da colinearidade (funções de número par).

Câmeras 1 e 2	Função 01	Função 02	Função 03	Função 04	Função 05	Função 06	Função 07	Função 08
Reprojeção 3D-2D (pixels)	0,716	0,691	0,632	0,599	0,765	0,762	0,581	0,575
Distância 3D (mm)	1,914	1,694	1,262	1,174	2,164	2,162		
Reprojeção 2D-2D (pixels)	1,432	1,382	1,263	1,199	1,529	1,524	1,16	1,15
Câmeras 2 e 3								
Reprojeção 3D-2D (pixels)	1,5514	1,5506	1,556	1,559	1,557	1,556	1,552	1,551
Distância 3D (mm)	1,816	1,809	1,833	1,83	1,795	1,787		
Reprojeção 2D-2D (pixels)	3,099	3,098	3,111	3,114	3,112	3,109	3,098	3,097
Câmeras 3 e 4								
Reprojeção 3D-2D (pixels)	0,5195	0,5194	0,557	0,546	0,518	0,517	0,513	0,512
Distância 3D (mm)	1,095	1,093	1,115	1,068	1,116	1,092		
Reprojeção 2D-2D (pixels)	1,0389	1,0387	1,114	1,093	1,035	1,034	1,025	1,024

Tabela 3: Resultados do erro médio medidos na amostra  $B$ , 1000 vistas, capturada para teste da calibração realizada do sistema  $Y$  de quatro câmeras, duas firewire e duas usb.

Nas seguintes tabelas mostramos em resumo o ganho ou perda calculados percentualmente ao comparar os resultados mostrados nas tabelas 2 e 3.

Câmeras 1 e 2	Ganho % função 02 para 01	Ganho % função 04 para 03	Ganho % função 06 para 05	Ganho % função 08 para 07
Reprojeção 3D-2D (%)	2,98	4,95	2,74	-0,54
Distância 3D (%)	1,8	-4,53	1,53	
Reprojeção 2D-2D (%)	3,1	5,06	2,86	-0,42
Total acumulado (%)	7,88	5,48	7,13	-0,96
Câmeras 2 e 3				
Reprojeção 3D-2D (%)	0,92	0,41	-0,71	0,76
Distância 3D (%)	1,59	0,35	0,66	
Reprojeção 2D-2D (%)	0,94	0,34	-0,59	0,63
Total acumulado (%)	3,45	1,1	-0,64	1,39
Câmeras 3 e 4				
Reprojeção 3D-2D (%)	0,75	2,74	0,78	3,57
Distância 3D (%)	0,38	1,66	0,87	
Reprojeção 2D-2D (%)	1,12	2,47	0,78	3,57
Total acumulado (%)	2,25	6,87	2,43	7,14

Tabela 4: Ganho percentual ao analisar os resultados apresentados na tabela 1.

Nas tabelas 4 e 5, analisamos se a calibração realizada usando as funções de erro global que levaram em conta a colinearidade dos nossos padrões (funções pares) obtiveram menor erro médio em comparação com os resultados gerados pela calibração realizada com as funções de erro global *convencionais* (funções ímpares). Quando a função par teve um menor erro nas funções objetivo analisadas, o ganho percentual é positivo, no caso de perda, ou seja, quando a função par não gerou ganho, o valor percentual é negativo.

	Ganho % função 02 para 01	Ganho % função 04 para 03	Ganho % função 06 para 05	Ganho % função 08 para 07
Câmeras 1 e 2				
Reprojeção 3D-2D (%)	3,49	5,19	0,39	1,03
Distância 3D (%)	11,49	6,97	0,09	
Reprojeção 2D-2D (%)	3,49	5,07	0,33	0,86
Total acumulado (%)	18,47	17,23	0,81	1,89
Câmeras 2 e 3				
Reprojeção 3D-2D (%)	0,05	-0,19	0,07	0,06
Distância 3D (%)	0,39	0,16	0,45	
Reprojeção 2D-2D (%)	0,03	-0,09	0,09	0,03
Total acumulado (%)	0,47	-0,12	0,61	0,09
Câmeras 3 e 4				
Reprojeção 3D-2D (%)	0,02	1,97	0,19	0,02
Distância 3D (%)	0,18	4,22	2,15	
Reprojeção 2D-2D (%)	0,02	1,89	0,09	0,09
Total acumulado (%)	0,22	8,08	2,43	0,11

Tabela 5: Ganho percentual ao analisar os resultados apresentados na tabela 2.

Como pode ser observado nas tabelas 4 e 5, na maioria das vezes conseguimos obter um erro médio menor gerado pelas funções pares em relação às ímpares. Tivemos casos em que conseguimos alcançar um ganho de 18%. Porém, o ganho percentual médio ficou em 3,9 %, o que representa uma melhoria dos resultados obtidos quando usados os parâmetros das câmeras obtidos da calibração com funções pares, que internamente levaram em conta as funções objetivo derivadas da colinearidade do nosso padrão. É importante ressaltar também que nosso método não é infalível e, como também se observa em nossas tabelas, houve alguns casos em que nossas funções pares não tiveram ganho. A perda no pior caso não nunca foi maior a 0,96%.

Com os resultados mostrados nas tabelas acima podemos afirmar que realmente a característica de colinearidade de nossos padrões ajudou a melhorar calibração dos parâmetros de nossas câmeras, provando a contribuição que nosso método propõe com o uso dos novos padrões invariantes.

Outra observação importante a ser feita é ressaltar qual foi o erro médio alcançado em relação à recuperação das distâncias 3D entre os marcadores de nossos padrões invariantes. Esse tipo de medida de erro é muito utilizada quando são comparados diversos sistemas de rastreamento óptico, nos quais quanto menor for o erro maior será a precisão alcançada no processo de reconstrução 3D da posição dos nossos marcadores. A precisão neste tipo de medida é tão importante, que nos sistemas comerciais não é admitido um erro maior do que 5 mm para áreas de rastreamento semelhantes às usadas em nossos testes ( 2 m × 2 m × 2 m ). Como é de se esperar, uma reconstrução 3D com nível de precisão milimétrica,

está diretamente relacionada a uma calibração precisa dos parâmetros das nossas câmeras. Com base nos valores obtidos e mostrados para este tipo de erro em nossas tabelas 2 e 3, podemos afirmar que o maior erro obtido foi de 2,162 mm e o menor foi de 0,772mm. O erro médio geral analisando tanto os sistemas **X** e **Y** foi de  $1.43 \pm 0,7$ mm. Esse valor é muito menor que o máximo permitido de 5 mm.

A seguir mostraremos uma tabela de comparação da precisão do erro alcançado e descrito por outros métodos de calibração proposto em artigos acadêmicos, como os citados no capítulo 3 desta tese, assim como por especificações técnicas recolhidas de alguns produtos comerciais (Tabela 6).

<b>Método apresentado por</b>	<b>Valor do erro médio em (mm)</b>
<b>Acadêmicos</b>	
Borghese et al. [4] (2000)	2
Pintaric et al. [28] (2007)	5,3
Uematsu et al. [37] (2007)	2
<b>Obtido com o método proposto neste trabalho</b>	<b>1,5</b>
<b>Produtos comerciais</b>	
A.R.T. GmbH [1] (2009)	2
IOTracker [13] (2009)	5
VICON [39] (2009)	0,1

Tabela 6: Comparação valor do erro médio referente da medição da distância entre marcadores definidos em métodos acadêmicos e produtos comerciais.

A comparação acima pode ser um pouco superficial, devido aos diferentes equipamentos que podem ter sido utilizados para obter essas medidas. Porém, é importante ressaltar que em todos os trabalhos citados na tabela 6, os autores afirmam que usam hardware composto por câmeras de uso semi-industrial que permite usá-las junto a algum equipamento de sincronismo eletrônico que garante uma correta captura de imagens simultaneamente entre todas as câmeras. Segundo eles, isso ajuda a obter uma maior precisão na recuperação de medidas. Em nosso caso, é bom lembrar que nossos resultados foram obtidos usando câmeras sem sincronismo nenhum, e no caso das câmeras web utilizadas, elas estariam classificadas como câmeras de baixa qualidade, mas mesmo assim foi possível alcançar uma boa precisão milimétrica com as mesmas.

Finalmente, neste capítulo de resultados é importante ressaltar a importância e a contribuição do novo tipo de padrão de calibração usado como componente chave do método de múltipla calibração de câmeras descrito nesta tese. Este padrão de calibração trouxe robustez ao processo de aquisição de dados e nova

informação explorada pelo processo de otimização dos parâmetros das câmeras. A sua estrutura simples e compacta ofereceu novas vantagens em relação a outros padrões semelhantes, mas a técnica desenvolvida para sua correta e acertada identificação, validada em trabalhos acadêmicos como [15] e [33], faz dele uma opção mais robusta para a modelagem de um padrão de calibração. Algumas vantagens de nosso padrão invariante em relação aos padrões unidimensional e adimensional podem ser resumidas da seguinte forma:

- Maior número de marcadores por vista do padrão capturado, 4 pontos de referência em comparação com 2, no caso do padrão unidimensional, e 1, no caso do adimensional.
- Capacidade de usar mais de um padrão invariante simultaneamente para a captura de amostras de pontos referenciais.
- Maior robustez em relação à identificação grupal e individual dos padrões e os marcadores que o compõem, com base na técnica de reconhecimento desenvolvida para este tipo de padrão [15].

A principal desvantagem de nosso padrão invariante é derivada de problemas de oclusão, isso quer dizer que se algum marcador do padrão é ocluído, o resto dos marcadores não poderá ser identificado. Porém, esta desvantagem está presente também nos outros tipos de padrões: unidimensional e adimensional.

No capítulo seguinte, serão apresentadas nossas conclusões e propostas de trabalhos futuros.

## 6 Conclusões

Neste trabalho de tese, foram propostas a modelagem e a implementação de um novo método de múltipla calibração de câmeras que traz como novidade o uso de um novo tipo de padrão de calibração chamado de padrão invariante. As principais contribuições do método proposto estão diretamente relacionadas à robustez na identificação dos marcadores que compõem o novo padrão – especialmente aproveitado na etapa de aquisição de dados do nosso método – e na exploração da característica de colinearidade que o padrão invariante possui para a obtenção de melhores resultados no cálculo dos parâmetros das câmeras, realizado na última etapa de otimização do nosso método de múltipla calibração.

Nós testamos se a inserção da característica de colinearidade no processo de calibração das câmeras realmente geraria algum tipo de melhora, para isso, modelamos essa característica como duas funções objetivo que foram inseridas na função de erro global usada pelo algoritmo de otimização na terceira etapa do método proposto. Para a comparação do nosso método com outros, foi necessário também realizar a calibração das câmeras usando uma função de erro global composta por funções objetivo *convencionais*, as que são usadas comumente por outros métodos de múltipla calibração. Conforme apresentado nas tabelas 2, 3, 4 e 5, os resultados obtidos com a calibração feita com a inclusão da característica de colinearidade quase sempre geraram resultados com erro médio menor em relação às calibrações feitas com funções erro que não levaram em conta a colinearidade.

Os testes também mostraram ainda que foi possível alcançar um alto grau de precisão milimétrica na tarefa de recuperação das distâncias entre os marcadores do nosso padrão após serem reconstruídos em 3D. O erro médio alcançado foi de 1.5mm, que ficou bem abaixo do limite de 5mm proposto pelos sistemas comerciais de rastreamento óptico. Porém, o mais ressaltante em termos de resultados foi constatar que o erro médio se manteve nos dois sistemas de quatro câmeras **X** e **Y** montados para nossos testes, cuja principal diferença em relação aos sistemas comerciais é o tipo de hardware com o qual trabalhamos. Nossos sistemas de câmeras não têm nenhum tipo de sincronismo eletrônico entre elas, o

que é uma característica básica dos sistemas comerciais; ademais, a qualidade das nossas câmeras não é muito alta – ambos os tipos de câmeras usados para testar nossa calibração (câmeras web e câmeras firewire) estão abaixo da média das câmeras usadas pelos sistemas comerciais. No entanto, até mesmo as simples câmeras web de baixo custo forneceram precisão equivalente à das nossas câmeras firewire de melhor qualidade.

Com relação ao novo padrão de calibração definido como padrão invariante, podemos ressaltar que este padrão possui as mesmas vantagens oferecidas pelos padrões unidimensional e adimensional, comumente usados em diferentes métodos de calibração. Porém, o padrão invariante oferece maior robustez no processo de identificação grupal e individual dos marcadores que o compõem, porque existe uma técnica específica para o reconhecimento e a identificação deste tipo de padrão [15].

Finalmente, também queremos ressaltar a proposta do framework conceitual criado nesta tese para analisar os diferentes métodos de múltipla calibração citados nos trabalhos relacionados. Nós modelamos o framework com o propósito de definir melhor as etapas que compõem o processo de múltipla calibração de câmeras, e achamos que ele pode representar uma ferramenta muito útil na análise de algum outro método que possa ser proposto no futuro.

### **6.1. Trabalhos futuros**

O tema de tese apresentado neste trabalho teve como resultado a proposta de um novo método de múltipla calibração de câmeras, este tipo de método tem uma ampla gama de aplicações onde pode ser utilizado, especialmente aquelas onde a calibração de múltiplas câmeras representa o coração do funcionamento do sistema. De uma forma geral, como exemplos de aplicações onde nosso método poderia ser testado, citaremos as seguintes:

- Sistemas de rastreamento óptico.
- Sistema de captura de movimento.
- Modelagem e criação de interfaces de interação 3D muito usada em aplicações de realidade virtual e aumentada.
- Scanner 3D.



- Aplicações relacionadas à engenharia civil e de estruturas, onde se procura algum tipo de aplicação que permita a extração de medidas de objetos e estruturas a partir da análise de imagens capturadas por um conjunto de câmeras, evitando a medição manual das mesmas.
- Aplicações de treinamento cirúrgico, onde são usados sistemas estéreo de câmeras para rastrear os movimentos que um cirurgião faria durante uma cirurgia virtual onde podem ser simuladas cirurgias de alto risco, o que permite melhorar o desempenho e aumentar a confiança do cirurgião.

Como podemos observar, os exemplos acima citados podem ser entendidos como uma visão a longo prazo da variedade de aplicações onde podemos usar nosso método. Já a médio e curto prazo, nós propomos alguns projetos mais concretos como:

- Testar nosso método de calibração num sistema composto pelas seis câmeras usadas nos testes expostos nesta tese, o objetivo é implementar um sistema de rastreamento óptico com este conjunto de câmeras, que seja a base para a criação de um sistema de captura de movimento.
- Outra área onde testaremos nosso método é na implementação de uma aplicação dirigida à recuperação de medidas métricas semelhante a sistemas de scanner 3D. Este tipo de aplicação é implementado sobre um sistema de câmeras em estéreo. Esta ideia está apoiada nos resultados com precisão milimétrica alcançados no teste de recuperação de medidas realizado para testar nosso método.

Finalmente, seria bom também aprofundar e aprimorar o processo de extração e identificação dos marcadores do nosso padrão, a partir das imagens capturadas nas diferentes câmeras. Um exemplo deste aprimoramento foi apresentado no último ano pelo sistema comercial OptiTrack [26], onde se reportaram ganho de precisão derivado da aplicação desta nova técnica na extração da posição 2D dos marcadores na imagem.

## 7 Referências Bibliográficas

1. **ADVANCED REAL TIME TRACKING GmbH**, <http://www.ar-tracking.de/>, Julho, 2009.
2. BAZARAA M.S., SHERALI D. H., SHETTY C. M. **Nonlinear Programming Theory and Algorithms**. Second Edition. John Wiley & Sons, Inc. ISBN 0-471-55793-5.
3. BOGGS P.T., BYRD R.H., ROGERS J.E. AND SCHNABL R.B., **User's reference guide for ODRPACK**. Software for weighted orthogonal distance regression, version 2.01, National Institute of Standards and Technology, Gaithersburg, USA (1992).
4. BORGHESE A. and CERVERI P., **Calibrating a video camera pair with a rigid bar**. Pattern Recognition, 2000.
5. BOUGUET J. Y., **Camera Calibration Toolbox for Matlab**, [http://www.vision.caltech.edu/bouguetj/calib\\_doc/htmls/example5.html](http://www.vision.caltech.edu/bouguetj/calib_doc/htmls/example5.html), Julho, 2009.
6. DEVERNAY F., **C/C++ Minpack**, <http://devernay.free.fr/hacks/cminpack.html>, Julho 2009.
7. DORFMÜLLER, K., WIRTH, H., **Real-Time Hand and Head Tracking for Virtual Environments Using Infrared Beacons**, Proceedings of the International Workshop on Modelling and Motion Capture Techniques for Virtual Environments, pp. 113 – 127, London, UK, 1998.
8. DORFMÜLLER-ULHAAS, K., **Optical Tracking - From User Motion To 3D Interaction**, Tese de Doutorado, Vienna University of Technology, Institut 186 für Computergraphik und Algorithmen, Novembro, 2002.
9. FISCHLER, M. A., BOLLES, R. C. **Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image**

- Analysis and Automated Cartography.** Comm. of the ACM, Vol 24, pp 381-395, 1981.
10. HARTLEY R. and ZISSERMAN A. **Multiple View Geometry in Computer Vision.** Cambridge University Press, Cambridge, 2000.
  11. HEIKKILA J., SILVEN O., **A Four-Step Camera Calibration Procedure with Implicit Image Correction.** In Proc. of IEEE Computer Vision and Pattern Recognition, pp. 1106-1112, 1997.
  12. **Intel Open Source Computer Vision Library,** Intel Research, <http://www.intel.com/technology/computing/opencv/index.htm>, Julho 2009.
  13. **IOTRACKER ADVANCED OPTICAL MOTION,** <http://www.iotracker.com/>, June 2009.
  14. KURILLO G., Li Z., BAJCSY R., **Wide-Area External Multi-Camera Calibration Using Vision Graphs and Virtual Calibration Object,** In Proceedings of 2nd ACM/IEEE International Conference on Distributed Smart Cameras (ICDSC '08), Stanford, CA, 2008.
  15. LOAIZA M., RAPOSO B. A., GATTASS M., **A Novel Optical Tracking Algorithm for Point-Based Projective Invariant Marker Patterns,** 3rd International Symposium on Visual Computing – ISVC 2007, Lake Tahoe, Nevada/California, EUA. Advances in Visual Computing - Lecture Notes in Computer Science. Vol. 4841, pp. 160-169. Springer-Verlag, 2007.
  16. LOAIZA, M. E., RAPOSO, A. B., GATTASS, M. **A Novel Optical Tracking Algorithm for Point-Based Projective Invariant Marker Patterns.** In: C. J. P. de Lucena (ed.), Monografias em Ciência da Computação, n.31/07. Departamento de Informática, PUC-Rio, Rio de Janeiro, 2007 (ISSN 0103-9741)
  17. LOAIZA, M. E. F.. **Implementação de um Dispositivo de Rastreamento Óptico com 6 Graus de Liberdade para Interação com Aplicações de Realidade Virtual,** Dissertação de Mestrado, DI, PUC / RJ, Abril, 2005.

18. MA, Y., SOATTO, S., KOSECKÁ, J., SASTRY, S.S. **An Invitation to 3-D Vision: From Images to Geometric Models**, Springer-Verlag, 2005.
19. MARQUES C.C.C., **Um Sistema de Calibração de Câmera**, Dissertação de Mestrado, Programa de Pós-Graduação em Matemática, Universidade Federal de Alagoas, AL, Fevereiro, 2007.
20. MEER P., LENZ R. AND RAMAKRISHNA S. **Efficient Invariant Representations**. International Journal of Computer Vision 26, 137–152, 1998.
21. MEER P., LENZ R. AND RAMAKRISHNA S. **Correspondence of Coplanar Features through P2-Invariant Representations**. In Applications of Invariance in Computer Vision, Springer, Heidelberg, 473– 492, 1993.
22. MITCHELSON J., HILTON A., **Wand-based multiple camera studio calibration**. CVSSP Technical Report VSSP-TR-2/2003.
23. MURRAY, N., GOULERMAS, Y., FERNANDO, T. **Visual Tracking for a Virtual Environment**, HCI International, Vol. 1, pp. 1198-1202, 2003.
24. NELDER J. A. and MEAD R., **A simplex method for function minimization**. Computer Journal, vol. 7:308-313, 1965.
25. OLSEN B.D. and HOOVER A., **Calibrating a camera network using a domino grid**, Pattern Recognition Journal, Vol. 34, 1105-1117, 2001.
26. **Optical Motion Capture Solution - OptiTrack**, <http://www.naturalpoint.com/optitrack/>, Junho 2009.
27. PRESS W.H., TEUKOLSKY S.A., VETTERLING W.T., AND FLANNERY B.P. **Numerical Recipes in C**. The Art of Scientific Computing. Cambridge University Press, Cambridge, 2 edition, 1999.
28. PINTARIC T. and KAUFMANN H., **Affordable Infrared-Optical Pose Tracking for Virtual and Augmented Reality**, IEEE VR Workshop on Trends and Issues in Tracking for Virtual Environments, Charlotte, NC (USA), 2007.

29. RIBO, M., PINZ, A., FUHRMANN, A.L., **A New Optical Tracking System for Virtual and Augmented Reality Applications**, IEEE Instrumentation and Measurement Technology Conference, Budapest, Hungary, Maio, 2001.
30. ROTHWELL, C., FAUGERAS, O., CSURKA, G., **A Comparison of Projective Reconstruction Methods for Pairs of Views**, Computer Vision and Image Understanding, vol. 68, No 1, pp. 37-58, Outubro, 1997
31. SALVI, J., ARMANGUE, X., PAGES, J., **A Survey Addressing The Fundamental Matrix Estimation Problem**, IEEE International Conference on Image Processing, Thessaloniki, Greece, Vol. 2, No 8, pp. 209-212, Outubro, 2001.
32. SVOBODA T., MARTINEC D. and PAJDLA T., **A Convenient Multi-Camera Self-Calibration for Virtual Environments**, PRESENCE: Teleoperators and Virtual Environments, MIT Press, August 2005.
33. TEIXEIRA, L., LOAIZA, M., RAPOSO, A. B., GATTASS, M. **Augmented Reality Using Projective Invariant Patterns**. 4th International Symposium on Visual Computing – ISVC 2008 (Advances in Visual Computing - Lecture Notes in Computer Science, vol. 5358), Las Vegas, Nevada, EUA., p.520-529. Springer-Verlag, 2008.
34. TORR, P.H.S. , MURRAY, D.W., **The Development and Comparison of Robust Methods for Estimating the Fundamental Matrix**, International Journal Computer Vision , pp.1-33, July 1996.
35. TSAI, R.Y. **An Efficient and Accurate Camera Calibration Technique for 3D Machine Vision**, CVPR'86 Proceeding, IEEE Computer Society Conference on Computer Vision and Pattern Recognition, FL, Junho 22-26, pp. 364-373, 1986.
36. TSAI, R.Y., **Metrology Using Off-the-Shelf TV Cameras and Lenses**, IEEE Journal of Robotics and Automation, Vol. 3, No. 4, pp. 323-344, August 1987.

37. UEMATSU Y., TESHIMA T., SAITO H., HONGHUA C., **D-Calib: Calibration Software for Multiple Cameras System**, 14th International Conference on Image Analysis and Processing (ICIAP 2007), 285-290, 2007.
38. VELHO L., CARVALHO P., SOARES E., SA A., MONTENEGRO A., PEIXOTO A., and RIVERA L. A. **Fotografia 3D**. 25 Colóquio Brasileiro de Matemática - IMPA, Rio de Janeiro, 2005.
39. **Vicon Motion Systems**, <http://www.vicon.com/>. Julho 2009.
40. ZHANG Z., **A Flexible New Technique For Camera Calibration**, IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(11):1330-1334, 2000.

## 8 Apêndice A: Modelo de câmera

Um modelo de câmera é uma forma conceitual de definir e detalhar como funciona internamente um dispositivo de captura de imagem como, por exemplo, uma câmera fotográfica ou de vídeo. O objetivo do modelo é ajudar a compreender o processo que leva um ponto tridimensional do espaço físico enxergado por uma câmera para um ponto bidimensional posicionado dentro da imagem capturada desse mesmo espaço físico visto pela câmera.

Uma forma de modelar esse processo é usar um modelo básico de câmera conhecido como câmera Pinhole (Figura A1). O modelo de câmera Pinhole é definido como uma transformação em perspectiva composta por um conjunto de transformações entre sistemas distintos de coordenadas que permitem modelar a correspondência entre pontos do espaço e pontos da imagem. Primeiramente, definiremos esses sistemas de coordenadas tal como apresentados em [38] e depois descreveremos as transformações correspondentes entre esses sistemas.

### 8.1.1. Sistemas de coordenadas

Os sistemas de coordenadas considerados no processo de modelagem de uma câmera são descritos a seguir:

1. **Sistema de coordenadas do mundo (SCM)** é um sistema tridimensional estabelecido para modelar o nosso cenário real ou cenário físico enxergado pela câmera. É definido num sistema de unidade de medida métrico e é mostrado na Figura A1 como o sistema de coordenadas definido pelos eixos ( $X_w, Y_w, Z_w$ ).
2. **Sistema de coordenadas da câmera (SCC)** é um sistema tridimensional cuja origem fica definida pelo centro óptico da câmera. Os eixos X e Y são definidos de modo a serem paralelos às bordas da imagem formada sobre o plano da imagem. A distância entre o centro óptico da câmera e o plano da imagem é chamada

distância focal. Esse sistema de coordenadas está definido pelos eixos  $(X_c, Y_c, Z_c)$  na Figura A1.

3. **Sistema de coordenadas da imagem (SCI)** é um sistema bidimensional definido sobre o plano da imagem ou plano de projeção, cuja origem fica determinada pela projeção da origem do sistema da câmera sobre o plano de imagem. Esse sistema de coordenadas está definido pelos eixos  $(u', v')$  na Figura A1.
4. **Sistema de coordenadas em pixel (SCP)** é um sistema bidimensional com coordenadas são expressadas em pixels. Esse sistema permite definir a posição de um ponto na imagem com relação à grade de pixels do sensor óptico da câmera. A origem desse sistema costuma ser definida no canto superior ou inferior esquerdo da imagem. Esse sistema de coordenadas está definido pelos eixos  $(u, v)$  na Figura A1.

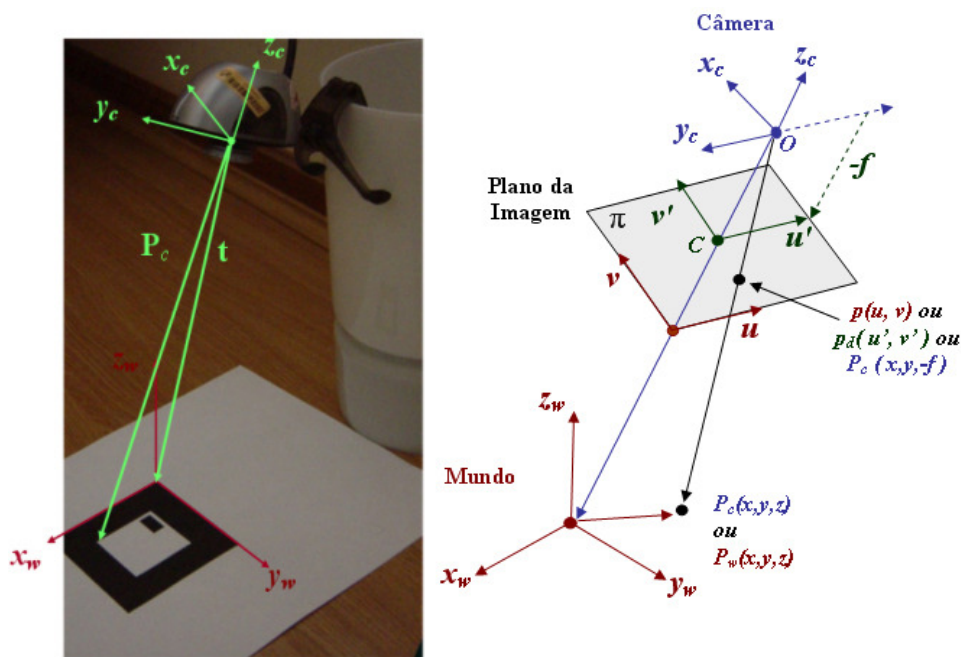


Figura A1: Modelo de captura da câmera Pinhole [17].

### 8.1.2. Transformações

As transformações do processo de imageamento de um ponto no mundo através dos diferentes sistemas de coordenadas expostas na seção anterior são descritas a seguir:



- A primeira transformação permite expressar as coordenadas de um ponto no mundo (SCM) no sistema de coordenadas da câmera (SCC). Essa mudança de coordenadas é definida por uma transformação que inclui uma matriz de rotação e um vetor de translação.
- A segunda transformação permite expressar a projeção em perspectiva de um ponto definido em coordenadas da câmera (SCC) para uma posição 2D sobre o plano da imagem (SCI).
- A terceira transformação permite expressar um ponto definido no plano da imagem (SCI) em uma posição na superfície de células fotosensíveis que o sensor óptico da câmera possui (SCP).
- Uma última transformação que pode ser considerada no processo de imageamento é a correção da distorção da lente. No modelo de câmera Pinhole, essa distorção não é considerada porque o componente responsável, ou seja a lente, só está presente nas câmeras reais. O objetivo da lente é controlar a incidência dos raios de luz sobre o sensor óptico da câmera, mas nesse processo a lente gera um tipo de distorção que muda a posição de um ponto que foi projetado sobre o plano da imagem, afetando especialmente a transformação entre o SCC e o SCI.

Para moldar cada uma das transformações intermediárias definidas entre o sistema de coordenadas do mundo, da câmera, do plano da imagem e do sensor óptico, é necessário definir alguns parâmetros que nos ajudem a modelar teoricamente cada uma dessas transformações. Explicaremos a seguir estes três tipos de parâmetros:

1. Parâmetros extrínsecos.
2. Parâmetros intrínsecos.
3. Coeficientes de distorção da lente.

### 8.1.3. Parâmetros extrínsecos

Os parâmetros extrínsecos permitem modelar a transformação rígida de uma posição 3D expressa no SCM para uma outra posição 3D correspondente expressa no SCC.

Esse tipo de parâmetros são definido por dois componentes:

- Uma matriz de rotação, que permite rodar o sistema de eixos base do SCM para deixá-los alinhados com os eixos base do SCC.
- Um vetor de translação, que permite trasladar o centro do SCM para a mesma posição do centro do SCC.

Esses parâmetros e sua aplicação no processo de transformação entre coordenadas do sistema do mundo e da câmera são modelados matematicamente na equação 1a, e exemplificados visualmente na Figura A2.

$$P_c \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} = \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix} P_w \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \quad (1a)$$

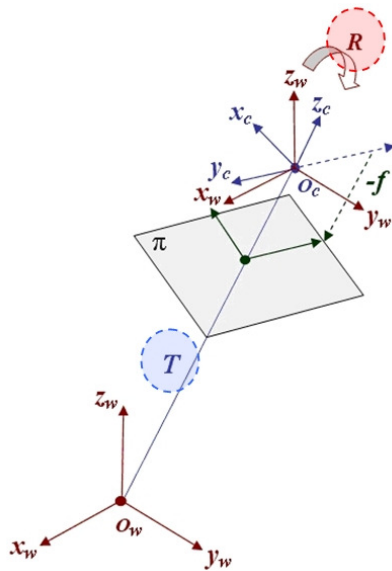


Figura A2: Parâmetros extrínsecos na modelagem da transformação entre os SCM para o SCC [17].

#### 8.1.4. Parâmetros intrínsecos

Este tipo de parâmetro é o utilizado para representar as características internas da câmera, como distância focal  $f$  e centro da imagem  $O$  (Figura A3). Os parâmetros intrínsecos modelam a projeção de pontos 3D expressos no SCC para pontos correspondentes 2D posicionados sobre o plano da imagem definido no SCI (Figura A3).

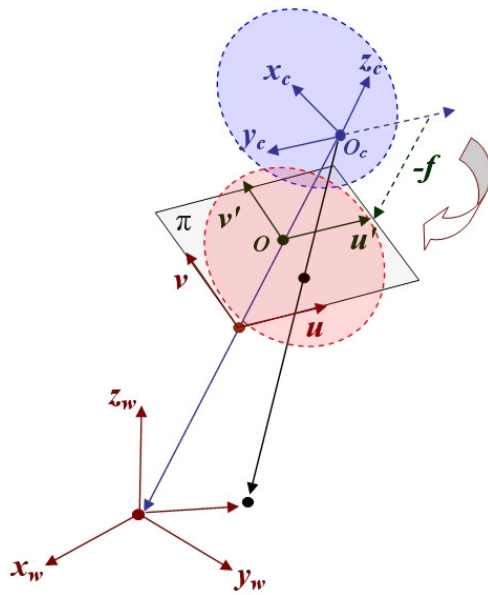


Figura A3: Parâmetros intrínsecos na modelagem da transformação entre o SCC para o SCI [17].

Um ponto  $p_c$  expressado no SCC gera um raio que sai da origem do SCC e chega a esse ponto, e no trajeto o raio intercepta o plano da imagem. O ponto de interseção  $p'$  representa a projeção do ponto 3D no SCC sobre o plano de imagem 2D expresso no SCI (Figura A4).

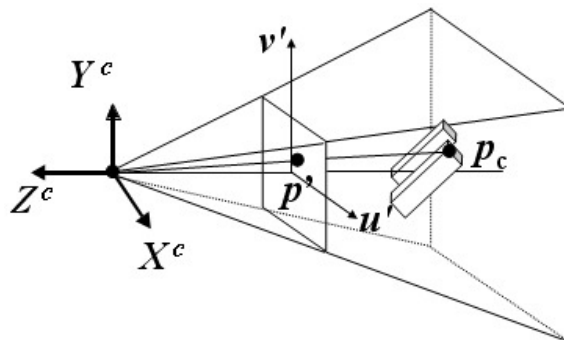


Figura A4: Projeção de um ponto  $P_c$  no SCC para um ponto  $P'$  no SCI [17].

A projeção de um ponto no SCC para o SCI é modelado matematicamente nas equações 2a e 3a.

$$p' \begin{bmatrix} u' \\ v' \\ 1 \end{bmatrix} \cong \begin{bmatrix} -f & 0 & 0 & 0 \\ 0 & -f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} P_c \begin{pmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{pmatrix} \quad (2a)$$

$$p' \begin{pmatrix} u' \\ v' \end{pmatrix} = -f \begin{pmatrix} X_c / Z_c \\ Y_c / Z_c \end{pmatrix} \quad (3a)$$

A transformação projetiva descrita na equação 2a é não inversível, e demonstra que cada ponto do espaço corresponde a um único ponto na imagem, mas um ponto na imagem pode corresponder a uma infinidade de pontos do espaço. Um ponto  $(u', v')$  na imagem é a projeção de todos os pontos espaço no SCC da forma  $\lambda(u', v', -f)$ .

Após ter projetado um ponto sobre o plano da imagem, podemos também descobrir a posição do mesmo na grade de pixels do sensor óptico da câmera. Isso é feito fazendo uma transformação entre o SCI e o SCP. Quando um raio luminoso atinge o plano de formação da imagem, ele é registrado pelo sensor óptico na câmera. Esse sensor é composto por uma matriz retangular de células fotossensíveis, mas algumas imperfeições na grade de células podem ocorrer:

- O espaçamento entre linhas pode ser diferente do espaçamento entre colunas.
- As colunas podem não ser dispostas perpendicularmente às linhas (devido a imperfeições geométricas de construção ou devido ao processo de aquisição dos valores dos sensores).

Finalmente, nós precisamos transladar a origem do SCI usualmente posicionada no centro da imagem, para a origem do SCP, que costuma se localizar num dos cantos superior ou inferior esquerdo da imagem. A transformação do SCI para SCP pode ser definida pela equação 4a.

$$p \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} s_x & \tau & u_c \\ 0 & s_y & v_c \\ 0 & 0 & 1 \end{bmatrix} P' \begin{bmatrix} u' \\ v' \\ 1 \end{bmatrix} \quad (4a)$$

Os coeficientes mostrados na matriz acima junto com a distância focal definem os parâmetros intrínsecos da câmera. Cada um dos coeficientes da matriz acima será descrito a seguir:

- $s_x$  e  $s_y$  representam o número de pixels por unidade de comprimento nas direções horizontal e vertical respectivamente. Na maior parte das câmeras, idealmente  $s_x$  e  $s_y$  são iguais, ou seja, os pixels são quadrados.
- $u_c$  e  $v_c$  fornecem a posição, em pixels, da projeção ortogonal da origem do SCC sobre o plano da imagem. Essa posição na maioria

das câmeras é definida pelo valor referente à metade das dimensões da imagem (largura/2, altura/2).

- $\tau$  é a tangente do ângulo que as colunas de pixels formam com a perpendicular às linhas. Na maior parte das câmeras, idealmente as colunas são perpendiculares às linhas, ou seja,  $\tau = 0$ .

A figura A5 exemplifica essa transformação entre o SCI e o SCP, e as equações 5a e 6a definem a transformação que modela a projeção de um ponto no SCC para um ponto no SCP, nas quais são utilizados todos os parâmetros intrínsecos antes descritos.

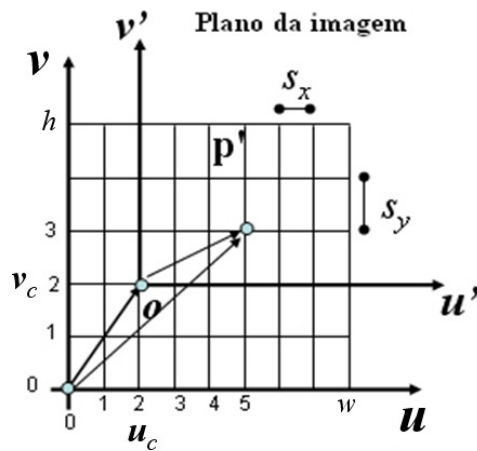


Figura A5: Transformação do SCI para o SCP [17].

$$p \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \cong \begin{bmatrix} s_x & \tau & u_c \\ 0 & s_y & v_c \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -f & 0 & 0 & 0 \\ 0 & -f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} P_c \begin{pmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{pmatrix} \quad (5a)$$

$$p \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \cong \begin{bmatrix} f s_x & f \tau & u_c \\ 0 & f s_y & v_c \\ 0 & 0 & 1 \end{bmatrix} P_c \begin{pmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{pmatrix} \quad (6a)$$

Finalmente, na equação 7a é definida a transformação que leva um ponto no SCM para sua projeção no SCP, o que é conseguido multiplicando as matrizes de transformação definidas pelos parâmetros intrínsecos e extrínsecos da câmera.

$$P \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \cong \begin{bmatrix} f s_x & f \tau & u_c \\ 0 & f s_y & v_c \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} P_w \begin{pmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{pmatrix} \quad (7a)$$

Como podemos observar, os parâmetros  $f$ ,  $s_x$  e  $s_y$  aparecem na transformação da câmera como os produtos  $f s_x$  e  $f s_y$ . Muitas vezes é difícil calcular os valores individuais de cada parâmetro, a menos que se tenha informação técnica dos valores  $s_x$  e  $s_y$  por parte do fabricante do sensor óptico da câmera. Por esse motivo é comum a matriz de parâmetros intrínsecos ser descrita como na equação 8a.

$$K = \begin{bmatrix} f s_x & f \tau & u_c \\ 0 & f s_y & v_c \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} f_x & s & u_c \\ 0 & f_y & v_c \\ 0 & 0 & 1 \end{bmatrix} \quad (8a)$$

Por fim, a multiplicação das matrizes que representam nossos parâmetros intrínsecos e extrínsecos define um novo tipo de matriz única chamada de matriz de projeção da câmera  $P$ , definida na equação 9a.

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \cong \begin{bmatrix} f_x & s & u_c \\ 0 & f_y & v_c \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}$$

$$P = \begin{bmatrix} f_x & s & u_c \\ 0 & f_y & v_c \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} \Rightarrow \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \cong P \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \quad (9a)$$

### 8.1.5. Coeficientes de distorção das lentes

Até agora, a descrição dos parâmetros usados para modelar a câmera real seguiu a definição dada pelo modelo de câmera ideal Pinhole, mas na realidade as câmeras fotográficas ou de captura de vídeo possuem mais um componente que deve ser considerado em nossa modelagem: a lente.

A lente é um equipamento usado no processo de captura óptica para controlar a incidência dos raios de luz sobre a área do sensor óptico. Controlando o grau de incidência, pode-se aumentar ou diminuir o campo de visão da câmera, mas nesse processo um tipo de distorção é inserido na imagem final vista pela câmera. Estes dois tipos de distorção visual produzidas sobre a imagem por causa das lentes tipificam bem este problema (Figura A6):

- A distorção de abaulamento ( *Barrel* ) está associada com as lentes grandes angulares (ou de zoom mínimo). Ela faz com que as imagens pareçam esféricas (curvadas para fora).

- Em oposição, a distorção de afunilamento ( *Pincushion* ) está associada com as lentes telescópicas (ou de zoom máximo) e fotos embaixo da água. As imagens parecem espremidas em torno do centro (distorcidas para dentro). O afunilamento é, muitas das vezes, menos notório que o abaulamento.

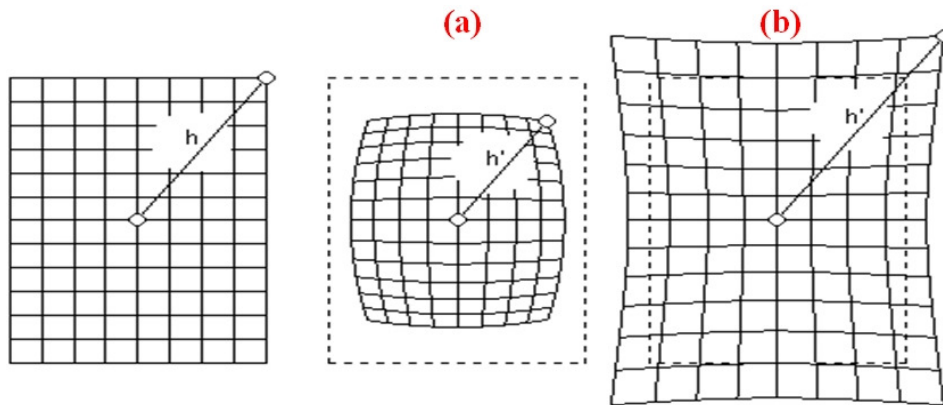


Figura A6: Distorção causada por distintos tipos de lente - (a) Barrel, (b) Pincushion.

(Fonte: <http://toothwalker.org/optics/distortion.html>)

É importante corrigir esse tipo de distorção especialmente na implementação de aplicações nas quais algum tipo de recuperação de medida será feita a partir das imagens capturadas.

Na fotogrametria, tenta-se modelar esse tipo de distorção com base em três tipos de componentes: distorção radial ( *radial* ), distorção de descentrada ( *decentering* ) e distorção de prisma ( *thin prism* ). Nós consideramos nesta descrição só as duas primeiras.

Tanto a distorção radial como a descentrada são aproximadas por meio de funções polinomiais, que usam um conjunto de coeficientes definidos como (  $k1, k2$  ) e (  $p1, p2$  ) para modelar o desvio produzido pela distorção radial e descentrada respectivamente. As distorções afetam especificamente a transformação entre o SCC e o SCI, e é justamente sobre a posição 2D do ponto (  $u', v'$  ) projetado no plano da imagem que a distorção da lente é aplicada. Definiremos (  $u'_d, v'_d$  ) como sendo igual ao ponto (  $u', v'$  ), ou seja, o ponto distorcido que é capturado sobre o plano da imagem numa câmera real, e o ponto (  $u'_u, v'_u$  ) como sendo o mesmo ponto, mas sem a distorção da lente.

As equações 10a, 11a, 12a descrevem essa correção da distorção que correlacionando um ponto distorcido (  $u'_d, v'_d$  ) para um sem distorção definido por (  $u'_u, v'_u$  ).

$$u'_u = u'_d * (1 + k1*r^2 + k2*r^4) + 2*p1*u'_d * v'_d + p2*(r^2 + 2*u'_d^2) \quad (10a)$$

$$v'_u = v'_d * (1 + k1*r^2 + k2*r^4) + p1*(r^2 + 2*v'_d^2) + 2*p2*u'_d * v'_d \quad (11a)$$

$$r^2 = u'_d^2 + v'_d^2 \quad (12a)$$

A modelagem apresentada mostrou os diferentes sistemas de coordenadas, transformações e parâmetros a serem considerados no processo de imageamento de ponto no mundo para um ponto no plano da imagem na câmera. Essas informações serão de muita utilidade na descrição das seções seguintes, que definirão o processo de calibração de múltiplas câmeras simultaneamente.



## 9 Apêndice B: Projetivas Invariantes

As características projetivas invariantes têm como base o estudo das propriedades da razão cruzada (*cross ratio*), a qual é definida pela razão cruzada entre as distâncias definidas por um conjunto de 4 pontos colineares. Foi baseado nesse estudo que alguns trabalhos na área de reconhecimento de padrões foram propostos. Esses trabalhos exploraram essas características como uma ferramenta para criar e reconhecer individualmente padrões de pontos definidos num formato específico. Alguns dos trabalhos mais importantes nessa linha foram desenvolvidos por Meer et al.[20][21] e como resultado foi a proposta de uma nova técnica conhecida como “ $P^2$ -invariant”. Essa técnica foi testada no reconhecimento de padrões extraídos a partir de imagens aéreas de construções e campos agrícolas. As imagens aéreas têm pouca distorção por causa da distância na qual foram tomadas, o que permitiu obter bons resultados nos testes realizados. Porém, a pouca amostragem e má qualidade das imagens usadas para o cálculo do identificador “ $P^2$ -invariant” criou uma alta instabilidade no processo de reconhecimento dos padrões quando a variação entre as imagens usadas para cálculo do valor do identificador “ $P^2$ -invariant” e as de comparação era muito significativa.

Já no trabalho proposto por Loaiza et al.[15] essa mesma técnica foi testada em outro ambiente, onde se tinha total controle sobre: as imagens usadas para calcular o valor do identificador “ $P^2$ -Invariant”, a definição e criação dos padrões a serem rastreados e os valores da distorção das lentes que são mais ressaltantes por conta da pouca distância de captura entre a câmera e os marcadores que compunham os padrões. Como resultado desse trabalho foi proposto um novo tipo de algoritmo baseado no valor “ $P^2$ -Invariant” que está composto por duas etapas: treinamento e reconhecimento. O trabalho desenvolvido em [15] provou que é possível criar um identificador único para padrões colineares de 4 pontos e padrões coplanares de 5 pontos baseado no valor do identificador “ $P^2$ -Invariant”, tendo como procedimento chave uma robusta e bem distribuída amostragem de imagens dos padrões a serem usados para cálculo do valor do identificador “ $P^2$ -

Invariant”. Esse treinamento mostrou que por causa de ruídos na extração da posição 2D dos marcadores dos padrões na imagem, o cálculo do valor do identificador “P<sup>2</sup>-invariant” não era estável e único, pelo contrario, ele precisa de um intervalo de variação que permita uma maior flexibilidade no processo de identificação dos padrões. Foi assim que o novo identificador proposto pelo trabalho de Loaiza et. al.[15] ficou definido como um intervalo entre um valor mínimo e máximo dos valores do identificador “P<sup>2</sup>-invariant” obtido para cada imagem do padrão capturada na amostra de treinamento. Este intervalo mostrou-se bem robusto, especialmente quando foi testado o reconhecimento de mais de um padrão, previamente treinados, que compartilhavam um mesmo formato: colinear ou coplanar. O reconhecimento individual de vários padrões com formatos semelhantes usando o novo identificador definido em [15] foi realizado de forma satisfatória, mesmo quando tinha-se dentro de uma mesma imagem mais de um destes padrões simultaneamente [15][16][33].

### 9.1. Cálculo do P<sup>2</sup>-Invariant para 4 pontos colineares

A seguir será descrito como é feito o cálculo do valor do identificador “P<sup>2</sup>-invariant”, base do algoritmo de reconhecimento do padrão de calibração usado neste trabalho de tese.

O valor fundamental para o cálculo da projetiva invariante “P<sup>2</sup>-invariant” é a razão cruzada (*cross ratio*) “ $\lambda$ ” definida por 4 pontos colineares definidos como  $A_i$  para  $i = 1, \dots, 4$ .

$$\lambda = \frac{(A_1A_3)(A_2A_4)}{(A_3A_2)(A_4A_1)} \quad (1b)$$

Onde  $A_iA_j$  é o segmento orientado definido pelos pontos  $A_i$  e  $A_j$ . Se nós consideramos que o etiquetamento deste 4 pontos pode variar, nós teríamos  $4! = 24$  diferentes combinações de 4 pontos colineares a serem avaliadas. Estas 24 combinações, segundo o trabalho de Meer et. al.[20], produzem só 6 valores diferentes de razão cruzada, os quais são:

$$\begin{aligned} \lambda_1 &= \lambda & \lambda_2 &= \frac{1}{\lambda} & \lambda_3 &= \frac{\lambda-1}{\lambda} \\ \lambda_4 &= \frac{\lambda}{\lambda-1} & \lambda_5 &= \frac{1}{1-\lambda} & \lambda_6 &= 1-\lambda \end{aligned} \quad (2b)$$

As soluções de qualquer polinômio simétrico construído com alguma das 6 razões cruzadas obtidas das permutações dos nossos 4 pontos colineares é também invariante à permutação dos índices de nossos pontos [20]. Esses polinômios são também invariantes à projeção por causa da razão cruzada com base na qual são calculados. Meer et. al.[20] analisaram também como era o comportamento das soluções obtidas pelos polinômios montados com as expressões definidas nas equações 2b, onde por exemplo para um polinômio simples baseado no  $\sum_{i=1}^6 \lambda_i$  a solução trivial é 3. Já no caso de outros polinômios não triviais usando as mesmas equações 2b, foi feita uma pesquisa sistemática onde foram montados diferentes polinômios de segundo grau baseados no produto das razões como  $\lambda_i \lambda_j$ , e concluíram que essas variações podem ser expressas como um combinação linear dos seguintes polinômios:

$$J_1[\lambda] = \frac{\lambda^6 - 3\lambda^5 + 3\lambda^4 - \lambda^3 + 3\lambda^2 - 3\lambda + 1}{\lambda^2(\lambda-1)^2} \quad (3b)$$

$$J_2[\lambda] = \frac{2\lambda^6 - 6\lambda^5 + 9\lambda^4 - 8\lambda^3 + 9\lambda^2 - 6\lambda + 2}{\lambda^2(\lambda-1)^2}$$

$$J_3[\lambda] = 3$$

$$J_4[\lambda] = 3$$

Como resultado, o valor utilizado para definir o valor do identificador “P<sup>2</sup>-invariant” é a razão definida por:

$$J[\lambda] = \frac{J_2[\lambda]}{J_1[\lambda]} \quad (4b)$$

O valor obtido em diferentes testes realizados em [20] ficou sempre dentro do intervalo de 2.0 a 2.8.