UNIVERSIDADE DE SÃO PAULO–USP
ESCOLA DE ENGENHARIA DE SÃO CARLOS
DEPARTAMENTO DE ENGENHARIA ELÉTRICA E DE COMPUTAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

Gustavo Teodoro Laureano

**Detecção Topológica de Padrões Xadrez
para Calibração de Câmeras**

São Carlos
2013

Gustavo Teodoro Laureano

**Detecção Topológica de Padrões Xadrez
para Calibração de Câmeras**

Tese de doutorado apresentada ao Programa de Engenharia Elétrica da Escola de Engenharia de São Carlos como parte dos requisitos para a obtenção do título de Doutor em Ciências.

Área de concentração: Processamento de Sinais e Instrumentação

Orientador: Prof^a. Dr^a. Maria Stela Veludo de Paiva

São Carlos

2013

Trata-se da versão corrigida da tese. A versão original se encontra disponível na EESC/USP que aloja o Programa de Pós-Graduação de Engenharia Elétrica.

AUTORIZO A REPRODUÇÃO TOTAL OU PARCIAL DESTE TRABALHO,
POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO, PARA FINS
DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

T314d   Teodoro Laureano, Gustavo
             Detecção Topológica de Padrões Xadrez para
             Calibração de Câmeras / Gustavo Teodoro Laureano;
             orientadora Maria Stela Veludo de Paiva. São Carlos,
             2013.

             Tese (Doutorado) - Programa de Pós-Graduação em
             Engenharia Elétrica e Área de Concentração em
             Processamento de Sinais e Instrumentação -- Escola de
             Engenharia de São Carlos da Universidade de São Paulo,
             2013.

             1. Padrões de Calibração. 2. Calibração de Câmeras.
             3. Reconhecimento de Padrões. 4. Detecção Topológica.
             I. Título.

FOLHA DE JULGAMENTO

Candidato: Bacharel **GUSTAVO TEODORO LAUREANO**.

Título da tese: "Detecção topológica de padrões xadrez para calibração de câmeras".

Data da defesa: 23/08/2013

Comissão Julgadora:

Resultado:

Profa. Dra. **Maria Stela Veludo de Paiva (Orientadora)**
(Escola de Engenharia de São Carlos/EESC)

APROVADO

Prof. Associado **Adilson Gonzaga**
(Universidade Federal de São Carlos/UFSCar)

APROVADO

Prof. Dr. **Fernando Santos Osório**
(Instituto de Matemática e Computação/ICMC)

APROVADO

Prof. Dr. **Clarimar José Coelho**
(Pontifícia Universidade Católica de Goiás/PUC-GO)

APROVADO

Prof. Dr. **Celso Aparecido de França**
(Universidade Federal de São Carlos/UFSCar)

APROVADO

Coordenador do Programa de Pós-Graduação em Engenharia Elétrica e Presidente da Comissão de Pós-Graduação:

Prof. Titular **Denis Vinicius Coury**

Este trabalho é dedicado à minha filha Luisa e minha esposa Regiane.

Agradecimentos

A Deus em primeiro lugar, o Grande Arquiteto do Universo, Senhor e Conhecedor de todos os mistérios que pairam em nossas mentes e ao nosso Senhor Jesus Cristo que é meu Caminho e minha Inspiração. Obrigado pelas oportunidades, por minhas conquistas, pelos desafios que tive que enfrentar e por seu amor infinito.

À minha orientadora Professora Doutora Maria Stela Veludo de Paiva, pelo apoio científico, pela a imensa ajuda na escrita deste texto, pelo carinho e cuidado que teve por mim todos esses anos.

À minha filha Luisa, a quem dedico este trabalho, por ser a minha maior fonte de felicidade, por ter me presenteado com sua chegada em minha vida e pelo seu amor que me deu forças para concluir esta pesquisa.

À minha amada esposa Regiane, que sempre me deu amor, pelo apoio e cuidado que tem por mim desde o primeiro momento que nos conhecemos. Obrigado por seu sacrifício e seu empenho em cuidar de nossa família, os quais foram o alicerce para essa minha conquista.

Aos meus pais, José Valoir Laureano e Laurice Teorodo Laureano, agradeço simplesmente por me amarem e pelo presente que são em minha vida. Obrigado Pai e Mãe por me dar o incentivo e a oportunidade de estudar.

Aos meus irmãos, Leonardo Teodoro Laureano e Vanessa Teodoro Laureano, por estarem sempre ao meu lado e pelo incentivo para eu prosseguir com meus estudos.

Ao Prof. Dr. Clarimar José Coelho, que conheci como orientador de Iniciação Científica e hoje o tenho como um dos meus melhores amigos. Obrigado por me iniciar na carreira acadêmica e abrir meus olhos para a pesquisa científica.

Agradeço aos meus grandes amigos Alexandre Coelho, Eduardo Sylvestre, Etienne Biasotto, Patrick Oliveira Santos, Marcelo Suetake, Helson Mineiro e Lucas Padilha pela amizade, companheirismo, apoio e os ótimos momentos que passamos juntos.

Aos funcionários do departamento de Engenharia Elétrica da USP/São Carlos, que sempre foram meus amigos, prestativos e extremamente competentes em suas atribuições. Obrigado pela conduta ética, educada e amigável.

À Escola de Engenharia de São Carlos/Universidade de São Paulo (EESC/USP) por me aceitar no programa de Doutorado e por me fornecer sua excelente infraestrutura de pesquisa.

Aos professores Adilson Gonzaga, Ivan Nunes da Silva, Evandro Luís Linhari Rodrigues e Carlos Dias Maciel do Departamento de Engenharia Elétrica e de Computação

da EESC/USP; aos professores Odemir Bruno, Eduardo Simões, Denis Fernando Wolf e Fernando Santos Osório do Instituto de Ciências Matemáticas e de Computação (ICMC); e ao professor Rodrigo Guido Capobianco do Instituto de Física de São Carlos (IFSC), dos quais tive a honra de ser aluno e por ministrarem suas disciplinas com excelência.

A todos aqueles que de alguma forma estiveram presentes na minha jornada pela busca do conhecimento.

À CAPES pelo apoio financeiro que foi de grande valia para a conclusão desse trabalho.

“Non Nobis Domine Non Nobis, Sed Nomini Tuo Da Gloriam.”

Resumo

Laureano, G. T. **Detecção Topológica de Padrões Xadrez para Calibração de Câmeras**. 138 p. Tese de doutorado – Escola de Engenharia de São Carlos, Universidade de São Paulo, 2013.

A identificação de pontos em padrões de calibração é relatada como uma fase trabalhosa nos processos de calibração de câmeras. Essa etapa é sensível à presença de ruídos e geralmente requer a identificação de muitos pontos de controle. A detecção automática de padrões de calibração é fundamental para a automatização desse processo. Os trabalhos existentes são poucos e implementam soluções semi-automáticas ou, quando automáticas, não lidam com imagens distorcidas, com inclinações acentuadas ou exigem a detecção de todos os pontos apresentados pelo padrão de calibração. O presente trabalho vem contribuir com essa área propondo uma metodologia, denominada *Chessboard Topological Detection* (ChTD), completamente automática, aplicável a imagens com alta distorção, independente da detecção completa do padrão e que apresenta melhor aproveitamento dos pontos e do conjunto de imagens de calibração. A metodologia proposta é baseada em três etapas: a detecção estrutural dos pontos de interesse; triangulação e filtragem dos pontos encontrados com base em regras topológicas; e a propagação e ajuste de coordenadas via conectividade dos pontos vizinhos. A partir dessa metodologia foi implementada uma ferramenta que permite extrair pontos de calibração de um padrão xadrez de forma simples e completamente automática. Na avaliação do método ChTD foi realizada uma comparação com a função *findChessBoardCorners* da biblioteca OpenCV usando conjuntos de imagens reais e sintéticas. Os testes realizados com imagens reais possibilitaram comparações numéricas da quantidade de pontos e de padrões detectados, e os com imagens sintéticas permitiram avaliar o ChTD diante de inclinações do padrão e ruídos controlados. Pelos resultados obtidos foi possível verificar que o ChTD foi superior ao método do OpenCV, apresentando menor dependência à inclinação do padrão, melhor aproveitamento dos pontos e das imagens de calibração e realizando a detecção em imagens distorcidas. O ChTD é executado em um único passo, diferenciando-se do método do OpenCV que faz transformações sucessivas da imagem. A metodologia desenvolvida é modularizada, possibilitando o uso de outros algoritmos em suas fases intermediárias.

Palavras-chave: Padrões de Calibração. Calibração de Câmeras. Reconhecimento de Padrões. Detecção Topológica.

Abstract

Laureano, G. T. **Topological Detection of Chessboard Patterns for Camera Calibration.** 138 p. Ph.D. Thesis – São Carlos School of Engineering, University of São Paulo, 2013.

The detection of calibration points is reported as a time consuming task in camera calibration systems. This task is sensitive to noise and usually requires identification of a large set of control points. A methodology of automatic detection of calibration patterns is essential for the automation of this process. Existing works are few and usually implement semi-automatic solutions. Automatic methods do not deal with distorted images, and patterns with high tilt angles or require the detection of all the calibration points. The present work aims to contribute to this area proposing a methodology named *Chessboard Topological Detection* (ChTD), which is completely automatic, applicable to images with high distortion, making better use of all detected points, avoiding calibration frames losses, regardless of the complete detection of the pattern features. The proposed methodology is based on three steps: the structural detection of points of interest, filtering and triangulation of the points based on topological rules and the adjust and coordinate propagation via connectivity of neighboring points inside the mesh. Based on this methodology, was implemented a tool that allows to extract calibration points of a chessboard pattern in a simple and automatic way. For the evaluation of the method ChTD, a comparison was made with the function `findChessBoardCorners`, which belongs to OpenCV library, using sets of synthetic and real images. Tests performed with real images allowed numerical comparisons of the number of points and detected patterns. The tests with synthetic images allowed to evaluate the ChTD facing the controlled tilt angle of the chessboard pattern and controlled image noise. The obtained results showed that the ChTD method was superior to OpenCV, with less dependence on the tilt angle of the calibration pattern, detecting more points, avoiding calibration images losses and detecting the pattern in distorted images. The ChTD runs in a single step, differing from OpenCV method that makes successive transformations of the image. The developed methodology is modular, enabling the use of different algorithms in its intermediates and some principal stages.

Keywords: Calibration Patterns. Camera Calibration. Pattern Recognition. Topological Detection..

Lista de Figuras

Figura 1	Exemplo de padrão xadrez.	24
Figura 2	Geometria projetiva.	36
Figura 3	Transformações projetivas.	39
Figura 4	Modelo <i>pinhole</i> de câmera.	41
Figura 5	Modelo <i>pinhole</i> de câmera em relação ao mundo exterior.	43
Figura 6	Padrão de calibração usado pelo trabalho de Brown (1971).	50
Figura 7	Exemplos de padrões 3D.	53
Figura 8	Exemplos de padrões 2D e 1D.	54
Figura 9	Padrão de calibração <i>Chessboard</i>	55
Figura 10	Outros marcadores de cantos e intersecções.	55
Figura 11	Padrões que usam retas e pontos.	55
Figura 12	Exemplos de padrões circulares.	56
Figura 13	Exemplos de padrões com marcadores ARTag.	58
Figura 14	Exemplos de outros padrões autoidentificáveis.	58
Figura 15	Erros de reprojeção de um sistema calibrado.	60
Figura 16	Metodologia de detecção de padrões xadrez.	72
Figura 17	Aparência de um canto-x.	73
Figura 18	Vizinhança de um canto-x.	75
Figura 19	Comparação da resposta do detector de cantos-x.	77
Figura 20	Exemplo do resultado gerado do detector de cantos-x adaptativo. . . .	78
Figura 21	Relação de áreas em uma imagem integral.	79
Figura 22	Exemplo de triangulação Delaunay.	81
Figura 23	Diagrama de classes para a implementação da triangulação Delaunay. .	81
Figura 24	Efeito das regras topológicas 1 e 2.	83
Figura 25	Efeito da terceira regra topológica.	83
Figura 26	Efeito da quarta regra topológica.	84
Figura 27	Demonstração do resultado do filtro topológico.	85
Figura 28	Propagação de coordenadas na malha de triângulos.	86
Figura 29	Configurações da malha de triângulos na propagação de coordenadas. .	87
Figura 30	Defasagem dos <i>pixels</i> no cálculo da relevância do canto-x.	89
Figura 31	Medida de similaridade do canto-x.	90
Figura 32	Interface do sistema de calibração.	91

Figura 33	Exemplos de imagens do conjunto Exemplo-OCV.	99
Figura 34	Imagens do conjunto Exemplo-OCV com pontos extras.	100
Figura 35	Resultados para a imagem 2 do conjunto Exemplo-OCV.	100
Figura 36	Resultados para a imagem 3 do conjunto Exemplo-OCV.	101
Figura 37	Exemplos de imagens do conjunto Bouguet.	102
Figura 38	Resultados para o conjunto de imagens Bouguet.	103
Figura 39	Comparação de resultados para a Imagem 11.	104
Figura 40	Comparação de resultados para a Imagem 20.	104
Figura 41	Comparação de resultados para a Imagem 19.	104
Figura 42	Comparação de resultados para a Imagem 18.	105
Figura 43	Exemplo do resultado do método ChTD para a imagem 18 de Bouguet.	105
Figura 44	Resultados de ChTD com pontos extras para as imagens de Bouguet.	106
Figura 45	Exemplos de imagens do conjunto OmiSample.	107
Figura 46	Exemplo da detecção do padrão das imagens <i>KaidanOmni</i> usando o método ChTD.	109
Figura 47	Exemplo da detecção do padrão das imagens <i>Ladybug</i> usando o método ChTD.	109
Figura 48	Exemplo da detecção do padrão das imagens <i>MiniOmni</i> usando o método ChTD.	110
Figura 49	Exemplo do resultado final do método ChTD após a fase de ajuste e ordenação de coordenadas.	110
Figura 50	Exemplos das imagens sintéticas usadas neste trabalho.	110
Figura 51	Porcentagem média de pontos detectados com a biblioteca OpenCV.	112
Figura 52	Porcentagem média de pontos detectados com o método ChTD.	112
Figura 53	Porcentagem média de quadros perdidos com a biblioteca OpenCV.	113
Figura 54	Porcentagem média de quadros perdidos pelo método ChTD.	114
Figura 55	Porcentagem média de pontos detectados com a biblioteca OpenCV após a aplicação do filtro gaussiano com $\sigma = 1, 5$	114
Figura 56	Porcentagem média de pontos detectados com o método ChTD após a aplicação do filtro gaussiano com $\sigma = 1, 5$	115
Figura 57	Porcentagem média de quadros perdidos com a biblioteca OpenCV.	115
Figura 58	Porcentagem média de quadros perdidos pelo método ChTD.	115
Figura 59	Resultado para padrões com oclusão parcial (Exemplo 1).	117
Figura 60	Resultado para padrões com oclusão parcial (Exemplo 2).	117

Lista de Tabelas

Tabela 1	Exemplos de <i>softwares</i> de detecção de padrões de calibração.	66
Tabela 2	Descrição do conjunto de imagens OmniSample.	96
Tabela 3	Parâmetros da câmera virtual para o conjunto de imagens sintéticas. .	97
Tabela 4	Resultados para o conjunto Exemplo-OCV.	99
Tabela 5	Resultados de ChTD para o conjunto Bouguet.	102
Tabela 6	Resultados do OpenCV para o conjunto Bouguet.	102
Tabela 7	Resultados do método ChTD para as imagens OmniSample.	107
Tabela 8	Resultados do OpenCV para as imagens OmniSample.	108
Tabela 9	Parâmetros de rotação para as imagens sintéticas.	111

Lista de siglas

0D zero dimensional

1D unidimensional

2D bidimensional

3D tridimensional

ARTag *Augmented Reality Tag*

ChTD *Chessboard Topological Detection*

DLT *Direct Linear Transformation*

DVS *Desktop Vision Systems*

PTAM *Parallel Tracking and Mapping*

RA Realidade Aumentada

SVD *Singular Value Decomposition*

Sumário

1	Introdução	23
1.1	Motivação	25
1.2	Justificativa	28
1.3	Objetivos	29
1.4	Contribuições do trabalho	30
1.5	Estrutura do trabalho	31
2	Fundamentos	33
2.1	Considerações Iniciais	33
2.2	Notação utilizada	33
2.3	Geometria projetiva	34
2.4	Transformações projetivas	37
2.5	Calibração de câmeras	40
2.5.1	Modelo de câmera <i>pinhole</i>	40
2.5.2	Estimação da matriz P	43
2.5.3	Considerações finais	45
3	Padrões de Calibração	47
3.1	Considerações iniciais	47
3.2	Padrões de calibração	47
3.3	Tipos de padrões	49
3.3.1	Marcadores de cantos e interseções	53
3.3.2	Marcadores circulares	56
3.3.3	Marcadores autocodificados	56
3.4	Precisão dos padrões de calibração	59
3.5	Considerações finais	61
4	Trabalhos Relacionados	63
4.1	Considerações iniciais	63
4.2	Definição do problema	63
4.3	Revisão dos trabalhos da área	65
4.4	Considerações finais	69

5	Detecção Automática de Padrões de Calibração	71
5.1	Considerações iniciais	71
5.2	Metodologia utilizada pelo algoritmo ChTD	71
5.3	Detector de cantos-x	73
5.4	Otimização do algoritmo de detecção	78
5.5	Filtragem topológica	80
5.6	Propagação e ajuste de coordenadas	86
5.7	Refinamento do local das características	88
5.8	Implementação da interface gráfica para o <i>software</i> de detecção	89
5.9	Considerações finais	92
6	Resultados Experimentais	95
6.1	Considerações iniciais	95
6.2	Materiais e métodos de avaliação	95
6.3	Experimentos com imagens reais	98
6.3.1	Imagens Exemplo-OCV	98
6.3.2	Imagens de Bouguet	101
6.3.3	Imagens OmniSample	106
6.4	Experimentos com imagens sintéticas	109
6.5	Resultados considerando oclusões	116
6.6	Considerações finais	116
7	Conclusões	119
7.1	Principais contribuições alcançadas pelo método ChTD	122
7.2	Trabalhos futuros	124
7.3	Trabalhos publicados e premiações	124
7.4	Trabalhos submetidos	125
7.5	Trabalhos a serem publicados	125
	Referências Bibliográficas	127

1 Introdução

A detecção automática de padrões de calibração é um assunto pouco abordado pelos trabalhos de calibração de câmeras. Esse processo busca extrair pontos em uma imagem de modo a fornecer os dados para os algoritmos de calibração. Nos trabalhos de calibração de câmeras a extração de pontos é uma fase considerada complicada por exigir alta precisão na detecção dos dados de calibração e baixo índice de ruído (TRUCCO; VERRI, 1998). Desse modo, a extração dos pontos requer uma análise minuciosa da imagem, sendo dependente de intervenção humana para garantir o levantamento seguro dos dados de calibração.

De modo geral os algoritmos de calibração são sensíveis a ruídos, exigem o conhecimento refinado sobre a posição dos pontos de calibração e múltiplas visões dos mesmos, o que torna a extração dos pontos uma tarefa trabalhosa e susceptível a erros quando executada manualmente (HARTLEY; ZISSERMAN, 2004; QI; LI; ZHENZHONG, 2010).

Os desdobramentos de um processo de calibração têm início com a definição de uma referência no ambiente real e de pontos de controle com coordenadas tridimensionais conhecidas. A câmera a ser calibrada observa uma cena do ambiente real de maneira que os pontos de controle são projetados em seu plano de imagem. O passo seguinte consiste em extrair e identificar as respectivas projeções de cada ponto de controle em coordenadas de imagem que, junto às coordenadas tridimensionais, formam o conjunto de dados de calibração.

Tradicionalmente, as metodologias de calibração de câmeras empregam o uso de objetos denominados *Padrões de Calibração* para facilitar a extração dos pontos de interesse. Esses objetos possuem dimensões físicas conhecidas e apresentam um conjunto de *Pontos de Controle* ou *Pontos de Calibração* que são representados por *características* visuais, cujas posições podem ser deduzidas em função de uma referência no próprio padrão e pela sua aparência regular (MALLON; WHELAN, 2007). A Figura 1 ilustra o padrão xadrez que é um dos padrões mais difundidos entre os trabalhos de calibração de câmeras. Esse padrão é formado por quadrados com dimensões similares e de cores diferentes, onde os pontos de controle são representados por características como cantos ou pelos pontos onde quadriláteros de cores diferentes se tocam, denominados *cantos-x*.

Nas aplicações onde é necessário relacionar grandezas espaciais em imagens com alguma unidade de medida, o termo “Calibração Métrica de Câmeras” é mais apropriado (BROWN, 1971). Nesse caso, é essencial o conhecimento da posição absoluta de cada característica apresentada (HARTLEY; ZISSERMAN, 2004). A detecção do padrão deve contemplar a identificação dos pontos de calibração e o reconhecimento de sua posição em relação a uma referência conhecida.

Figura 1: Exemplo de padrão xadrez. Os pontos correspondentes aos cantos dos quadriláteros e os cantos-x, onde quadriláteros de cores diferentes se tocam, representam os pontos de calibração. Fonte: Adaptado de Bouguet (2004).

A detecção automática de características em imagens pode ser considerada um problema da área de reconhecimento de padrões, que tem como objetivo principal classificar informações baseado em um conhecimento *a priori* sobre a característica de interesse (BISHOP, 2006). Para a detecção de pontos em imagens de padrões xadrez as informações são visuais e as características são dadas pela distribuição das intensidades dos *pixels* que forma um canto ou um canto-x como ilustrado na Figura 1.

Os problemas relacionados à detecção de padrões de calibração recaem sobre questões como a falta de unicidade da característica de interesse e a descaracterização provocada pela presença de ruído e distorções nas imagens (SCHMID; MOHR; BAUCKHAGE, 2000). Quando os pontos de interesse são definidos por características semelhantes, a aparência dos padrões de calibração impõe um alto grau de ambiguidade na imagem (DAO; SUGIMOTO, 2010). As características mais simples, como as da Figura 1, podem ocorrer no restante da cena por questões inerentes à própria imagem dos objetos que a compõem ou pela presença de ruídos (KANG; HA; JEONG, 2008). As distorções provocadas pelas lentes transformam as imagens de forma não linear desconfigurando as propriedades geométricas da característica de interesse (RUFLI; SCARAMUZZA; SIEGWART, 2008). Adicionalmente, a presença de ruído, sombras, iluminação não uniforme, distorções de perspectiva, escala e oclusões na imagem também têm papel importante na lista de dificuldades a serem superadas pelos trabalhos que buscam soluções para essa área (GONZALEZ; WOODS, 2006). A soma desses problemas faz da construção de métodos de detecção automáticos de padrões de calibração uma tarefa desafiadora (DAO; SUGIMOTO, 2010).

Na última década, a verificação de atributos topológicos tem sido empregada em trabalhos de detecção e casamento de padrões em vários segmentos da área de Visão Computacional (DELINGETTE; MONTAGNAT, 2000; DELINGETTE; MONTAGNAT, 2001; DONG; ZHOU, 2008; CHEN; FREEDMAN; LAMPERT, 2011; POKORNY et al., 2012). A detecção topológica de padrões faz uso de campo da matemática conhecido como Matemática Topológica, mais especificamente da Topologia Geral, que é fortemente baseada em teoria

dos conjuntos e geometria (HAZEWINKEL, 2001). A Topologia Geral considera atributos estruturais e qualitativos do objeto de análise como o relacionamento espacial de suas características, propriedades espaciais, conectividade, continuidade, bordas e forma.

A extração de atributos topológicos em dados reais é relatada como um problema de difícil solução. Uma das principais motivações de pesquisas nessa área tem sido focada na extração de informações topológicas sem a construção de modelos probabilísticos dos dados (POKORNY et al., 2012). No entanto, a avaliação de atributos topológicos é menos sensível a ruídos e admite a presença de distorções, mostrando-se adequada para as aplicações de detecção automática de padrões de calibração que são fortemente influenciadas por esses problemas (SHU; BRUNTON; FIALA, 2009).

O presente trabalho utiliza uma metodologia de detecção topológica dos padrões de calibração. A ideia de tratar esse problema por uma abordagem topológica não é nova, no entanto, a metodologia desenvolvida nesta tese é diferente das existentes na literatura por considerar a detecção parcial e a presença de ruídos inclusive no interior do padrão de calibração. Desse modo, esta tese vem contribuir com essa área de pesquisa, propondo um método automático de detecção de padrões xadrez que pode ser aplicado nos sistemas de calibração de câmeras com o objetivo de facilitar os processos de calibração e permitindo que essa tarefa seja realizada de acordo com as exigências estabelecidas pelos métodos de calibração e sem participação do usuário.

1.1 Motivação

A visão certamente é um dos sentidos mais desenvolvidos e mais complexos que se conhece. De fato, os sistemas naturais de visão possuem um alto grau de complexidade bem como uma alta taxa de sucesso nas operações que esses sistemas se propõem a desempenhar, seja a detecção ou a interpretação de informações visuais. Tomando como exemplo o sistema visual humano, não se pode negar que muitas tarefas são realizadas de maneira inconsciente e aparentemente trivial considerando somente dados visuais. Esse sistema evoluiu naturalmente por um grande período de tempo para aperfeiçoar o processamento de grandes volumes de informações que vão desde a retina até o córtex visual do cérebro por meio de uma rede complexa de sinapses neuronais, tomando proporções que fogem da compreensão humana (HUBEL, 1995).

A calibração de câmeras é uma dos campos de estudo da área de Visão Computacional que permite que as aplicações extrapolem o sentido natural da visão (DAVIES, 2012). Com uma câmera calibrada é possível realizar medições no espaço com precisão conhecida e em relação a uma unidade de comprimento. Os sistemas de visão que empregam câmeras calibradas podem ainda se beneficiar do cálculo de velocidade e do volume de objetos, o que não pode ser facilmente desempenhado pelos sistemas naturais de visão (WANG, 2013; SAKAS, 2002).

A calibração de câmeras é um tema de pesquisa que geralmente busca determinar as características geométricas de uma câmera de modo a modelar seu processo de formação de imagens (HARTLEY; ZISSERMAN, 2004). Desde 1970, quando houve um crescimento exponencial dos estudos na área de Visão Computacional, a calibração de câmeras é um assunto amplamente investigado em diversos centros de pesquisa, seja no meio acadêmico ou no setor industrial. Esse interesse pode ser observado pelo grande número de publicações nessa área e também pelo fato dessa tecnologia servir de base para um vasto campo de aplicações, estando presente desde a indústria do entretenimento até aplicações em medicina, engenharia e na área militar (QI; LI; ZHENZHONG, 2010).

A calibração de dispositivos é um assunto presente em quase todas as áreas do conhecimento, mas é essencialmente caracterizada pelo problema de estabelecer o relacionamento entre diferentes tipos de medidas (SKOOG; HOLLER; CROUCH, 2006). O termo “Calibração de Câmeras” possui um significado amplo, compreendendo qualquer operação de calibração realizada em uma câmera, no entanto, é frequentemente associado a um segmento específico, a Calibração Geométrica de Câmeras. Nesse caso, busca-se determinar os aspectos construtivos e geométricos de uma câmera, tendo como objetivo mapear os relacionamentos entre as medidas espaciais tridimensionais de pontos no ambiente com as medidas bidimensionais, ou distâncias, entre suas respectivas projeções dentro do plano de imagem. O mesmo termo também pode se referir à Calibração Fotométrica de Câmeras, que investiga a resposta radiométrica do sensor de imagem, normalmente aplicada ao mapeamento e à normalização de cores, mas que foge do contexto desse trabalho (SZELISKI, 2010).

Quando uma câmera está calibrada geometricamente, o conjunto de valores que descrever seu processo de formação de imagens é conhecido em função de um modelo matemático de câmera (TRUCCO; VERRI, 1998). Esses valores são parâmetros do modelo usado, divididos em parâmetros *intrínsecos*, que correspondem às características internas à câmera, como distância focal, centro real da imagem, eixo principal, inclinação entre os eixos de imagem, dimensões dos *pixels* e as distorções provocadas pelas lentes; e parâmetros *extrínsecos*, que descrevem a posição e orientação da câmera em relação a uma referência no ambiente.

Tradicionalmente muito esforço é dedicado à elaboração de métodos de calibração mais robustos e precisos, que focam essencialmente na modelagem matemática desse problema e na exatidão de suas soluções (SALVI; ARMANGUÉ; BATLLE, 2002). Esses estudos possibilitaram um amadurecimento significativo dessa área, no entanto a automatização do processo de calibração ainda é um assunto pouco abordado. Algumas operações importantes como extração de pontos em imagens, detecção de padrões, ajuste de modelos, retificação de imagens e os erros originados e propagados por esses estágios não recebem a devida atenção das pesquisas da área, embora sejam de suma importância para a qualidade da calibração (HEIKKILÄ; SILVÉN, 1997).

A detecção automática de padrões de calibração é ainda um problema em aberto, havendo poucos trabalhos publicados nessa área de pesquisa. Os estudos existentes, em sua maioria, propõem métodos semiautomáticos ou, quando se dizem automáticos, não conseguem lidar com imagens distorcidas, com inclinações acentuadas do padrão de calibração ou exigem a detecção de todos os pontos apresentados pelo padrão (BOUGUET, 2004; ESCALERA; ARMINGOL, 2010; BRADSKI, 2000).

A ausência de metodologias automáticas de detecção de pontos de controle dificulta o desenvolvimento e a manutenção de sistemas onde é necessária a calibração frequente dos dispositivos de captura de imagens. Isso pode ser observado em muitos setores, como na manutenção de câmeras usadas no setor industrial em aplicações de inspeção visual (CHIN; HARLOW, 1982; SCHAULIN, 2003; HASHIMOTO, 2012); em sistemas de visão ativa (YOUNG; MATAS; KITTLER, 1999), onde as câmeras usualmente se movimentam no espaço alterando seus parâmetros extrínsecos; na construção e manutenção de sistemas de mapeamento urbano (SHIODE, 2000; AKBARZADEH et al., 2006), quando as câmeras ficam expostas e sujeitas a avarias mecânicas que podem alterar tanto seus parâmetros internos e externos.

Na robótica industrial, braços robóticos estimam continuamente sua posição na linha de montagem observando padrões específicos a fim de determinar a posição correta das peças (KARABEGOVIĆ; VOJIC; DOLECEK, 2006). Na robótica móvel o emprego de câmeras possui propósitos variados, fornecendo dados para os sistemas de auto-localização, mapeamento de ambientes e estimação de movimentos (odometria visual) (DESOUZA; KAK, 2002; WILLIAMS; REID, 2010; CHOI et al., 2011). No contexto da navegação autônoma, uma câmera calibrada provê informações tridimensionais úteis para a detecção e reconhecimento de obstáculos (OHNISHI; IMIYA, 2006; CONRAD; DESOUZA, 2010). Nas aplicações de realidade aumentada (KLEIN; MURRAY, 2007) um sistema de visão fornece os dados necessários para a inserção de objetos virtuais no ambiente, tornando a interação homem-máquina mais imersivas.

Na medicina, o uso de imagens tem revolucionado o diagnóstico de enfermidades e conferindo um melhor acompanhamento dos pacientes. Com um sistema de visão calibrado pode-se obter informações que extrapolam a informação visual dos tecidos, chegando a medidas de comprimentos, volumes e dados sobre as texturas dos tecidos observados (CHANG; CHEN; HO, 1991; SAKAS, 2002). Na área de mobilidade e auxílio a pessoas com deficiências as informações tridimensionais extraídas do ambiente permitem a locomoção mais segura, com o aviso de imperfeições do solo e riscos de colisões (MOLTON et al., 1998).

Na fabricação de câmeras calibradas, como nas empresas PointGrey¹ e Tyzx², as câmeras, apesar de passarem por processos idênticos de fabricação, possuem pequenas imperfeições que as tornam únicas perante seus parâmetros intrínsecos. Também vem se tornando

---

¹Sítio da empresa PointGrey disponível em: <http://www.ptgrey.com/>, acessado em junho de 2013

²Sítio da empresa Tyzx disponível em: <http://www.tyzx.com/> acessado em junho de 2013

comuns os sistemas *Desktop Vision Systems* (DVS), onde normalmente empregam-se câmeras de baixo custo (ZHANG, 2000; TAVAKOLI; POURREZA, 2009).

Recentemente, o emprego de múltiplas câmeras tem se tornando mais comum em outros setores, como os que utilizam diversas câmeras para a vigilância e monitoração de ambientes (WANG, 2013), em aplicações semelhantes à de Wilburn et al. (2004) que chegam a usar uma matriz de cinquenta e duas (52) câmeras de baixo custo para aumentar a taxa de aquisição de quadros. Aplicações importantes incluem a reconstrução tridimensional (3D) como os trabalhos de Seitz et al. (2006) e Chrysostomou, Kyriakoulis e Gasteratos (2010), onde a dificuldade de reconstrução é reduzida quando todas as câmeras estão calibradas. Em aplicações desse tipo, a tarefa de calibração de todo o sistema é bastante demorada e tediosa, consumindo muito tempo de trabalho e sendo susceptível a erros quando existe intervenção manual.

Os trabalhos citados anteriormente mostram que a calibração de câmeras está inserida em um contexto abrangente de pesquisa e aplicações relacionadas à área de Visão Computacional. A construção de algoritmos de detecção automática de padrões de calibração pode beneficiar todos os campos onde é necessária a calibração de câmeras e demonstra um grande potencial de aplicações em sistemas que usam câmeras como instrumento de medida.

1.2 Justificativa

A literatura apresenta poucos trabalhos dedicados à detecção automática de padrões de calibração, sendo que estes apresentam algum tipo de limitação, seja pela exigência de intervenção do usuário, problemas com a detecção de padrões inclinados, pouco robusto à presença de ruídos e restrições quanto aos tipos de lentes utilizadas.

A identificação automática de pontos de calibração em imagens recai sobre um problema de reconhecimento de padrões visuais que representam os pontos de interesse. Essa tarefa é frequentemente vista como trivial pelos trabalhos da área já que os padrões são projetados para facilitar sua detectabilidade. Contudo, esse assunto é dificultado pelos problemas clássicos de processamento de imagens, tais como iluminação e ruído, e principalmente pela grande quantidade de ambiguidades induzida pela aparência do padrão de calibração e distorções provocadas pelas lentes das câmeras. Consequentemente, a determinação do conjunto de pontos de calibração ainda é uma tarefa que exige a intervenção do usuário. A automatização dessa tarefa precisa lidar com diversos fatores como a perda de informação, dados corrompidos e distorcidos (SOH, 1997).

Os métodos convencionais, como a função *findChessBoardCorners* implementada na biblioteca OpenCV (BRADSKI, 2000), apoiam-se sobre a necessidade de encontrar todos os pontos presentes no padrão de calibração para obter uma detecção segura. Outros trabalhos, como o *Toolbox* para Matlab desenvolvido por Bouguet (2004), exigem uma

participação do usuário em cada imagem analisada. As soluções disponíveis apresentam limitações quanto a presença de ruídos ou impõem restrições ao tipo de câmera a ser calibrada. Bouguet (2004), Chen, Benzeroual e Allison (2012), Bradski (2000) e Barreto et al. (2009) descrevem metodologias baseadas em detectores clássicos de cantos, como os de Harris e Stephens (1988) e Shi e Tomasi (1994), que são fortemente influenciados pela presença de ruídos (CHEN; ZHANG, 2005).

Escalera e Armingol (2010), Wang e Wu (2010), (YIMIN et al., 2010) e Chen, Benzeroual e Allison (2012) buscam características menos sensíveis a ruídos como retas e marcadores especiais, porém os métodos de detecção de retas não consideram distorções provocadas pelas lentes, apresentando falhas diante de câmeras convencionais (DUDA; HART, 1972). Os marcadores especiais funcionam como guias para a detecção dos pontos de calibração, desse modo a perda de marcadores inviabiliza a detecção de um grupo maior de características (FIALA; SHU, 2007).

Zheng et al. (2008), Shu, Brunton e Fiala (2009) e Dao e Sugimoto (2010) implementam detecções topológicas mas não tratam a presença de ruídos no interior do padrão e descartam muitos pontos indevidamente. De modo geral, os trabalhos relacionados à área de detecção automática de padrões de calibração não apresentam resultados comparativos entre as técnicas existentes, não realizam testes que permitem verificar a robustez de suas metodologias e se restringem à apresentação de resultados visuais.

A proposta desta pesquisa é desenvolver uma metodologia mais abrangente, buscando um melhor tratamento dos problemas relacionados ao ruído e admitindo a existência de falsos positivos. Adicionalmente, busca-se detectar o padrão de calibração mesmo em imagens distorcidas, melhorar o aproveitamento dos pontos e dos quadros de calibração e não restringir a extração de informações à detecção completa do padrão de calibração. A metodologia desenvolvida nesta pesquisa faz uso do padrão xadrez por apresentar melhores resultados na determinação dos parâmetros intrínsecos de câmera (MALLON; WHELAN, 2007) e por ser um dos objetos mais empregados nas aplicações de calibração de câmeras.

1.3 Objetivos

O objetivo principal desse trabalho é propor uma metodologia para a detecção topológica de padrões xadrez para calibração de câmeras, denominado *Chessboard Topological Detection* (ChTD), que inclui um detector de cantos-x estrutural e específico para a característica presente nos padrões de calibração xadrez; a filtragem topológica com base na regularidade desse padrão; a propagação e ajuste de coordenadas dos pontos detectados; e o aprimoramento da localização dos pontos finais.

Os objetivos específicos deste trabalho resumem-se na elaboração de uma metodologia de detecção de padrões xadrez que:

- seja robusta a ruídos, de modo a excluir falsos positivos;

- ❑ detecte o padrão sem a necessidade de detectar todos os pontos disponíveis;
- ❑ seja robusta à inclinação do padrão;
- ❑ seja independente do tipo de lente utilizada, permitindo a detecção do padrão mesmo em imagens altamente distorcidas;
- ❑ seja implementada em uma ferramenta de fácil uso pelo usuário final;

1.4 Contribuições do trabalho

O presente trabalho contribui com a área de detecção de padrões de calibração elaborando uma metodologia de detecção modularizada, baseada em atributos topológicos da característica de interesse, que consegue lidar melhor com o ruído presente na imagem, com um maior aproveitamento dos pontos disponíveis e com suporte à detecção de padrões de calibração em imagens distorcidas.

Cada fase da metodologia contribui nos seguintes aspectos:

- ❑ **Detector de cantos-x:** esse detector reduz a detecção de falsos positivos por considerar somente atributos estruturais da vizinhança de um canto-x ideal. Os atributos estruturais são extraídos de maneira adaptativa pela avaliação da vizinhança local de cada *pixel*. Considera somente uma amostra circular da vizinhança do *pixel* avaliado, permitindo a verificação rápida de sua estrutura. O mesmo detector é independente de um limiar global e seu resultado apresenta uma resposta classificatória que é dada em função do enquadramento do *pixel* na estrutura de um canto-x ideal;
- ❑ **Filtro topológico:** a análise topológica considera uma malha de triângulos construída sobre os cantos-x encontrados, permitindo uma verificação por meio de atributos topológicos e de conectividade. Essa abordagem torna o método mais robusto a ruídos, uma vez que os atributos representam regiões na imagem e são menos sensíveis a variações locais. As regras que definem o filtro topológico são verificadas em ordem crescente de rigorosidade, permitindo considerar mais pontos dentro da malha. A análise local de atributos topológicos verifica somente os aspectos relativos dos pontos de interesse, considerando disposição espacial de sua vizinhança de maneira independente de sua localização, o que permite considerar somente a regularidade do padrão, tornando a detecção menos dependente da distorção da imagem;
- ❑ **Propagação e ajuste de coordenadas:** permite que a posição relativa de cada ponto seja determinada, dando condições para a computação dos parâmetros intrínsecos da câmera. Caso o conjunto de pontos finais ofereça condições para a

determinação de suas posições absolutas no padrão, então o cálculo dos parâmetros extrínsecos pode ser realizado. Em ambos os casos, a ordenação dos pontos segue uma verificação topológica que explora a conectividade da malha filtrada. Essa abordagem torna a determinação das coordenadas dos pontos, que são referentes ao padrão de calibração, independente de oclusões ou falhas do detector de cantos-x, sendo realizada sempre entre triângulos que compartilhem uma aresta;

As três fases mencionadas definem a metodologia de detecção dos padrões de calibração implementada no método ChTD. A combinação dessas fases contribui para uma resposta mais estável desse método e menos dependente de distorções e da inclinação com que o padrão de calibração é apresentado à câmera. Além disso, proporciona um melhor aproveitamento dos pontos por imagem e do conjunto de imagens de calibração.

1.5 Estrutura do trabalho

A detecção automática de padrões de calibração é um assunto que está sempre relacionado à área de calibração de câmeras. No texto desta tese, frequentemente são feitas citações aos procedimentos de calibração e suas exigências para conseguir bons resultados. Esse é um comportamento comum entre as referências usadas neste trabalho. Por esse motivo, o Capítulo 2 traz uma revisão sobre os principais conceitos de calibração de câmeras. Nesse mesmo Capítulo foi usada uma nomenclatura normalizada, o que não ocorre na maioria dos trabalhos relacionados. O Capítulo 3 apresenta uma revisão sobre os padrões de calibração mais comuns na literatura e propõe duas metodologias de classificação desses aparatos. O Capítulo 4 apresenta o estado da arte do assunto relacionado a esta tese, discutindo as metodologias e resultados alcançados por este trabalho. O Capítulo 5 descreve a metodologia de detecção automática do método ChTD, que é a principal contribuição deste trabalho. O Capítulo 6 apresenta os testes e comparações realizadas para comprovar a eficiência do método ChTD. No Capítulo 7 as conclusões gerais sobre este trabalho são apresentadas.

2 Fundamentos

2.1 Considerações Iniciais

A calibração de câmeras é um tema quase sempre citado pela literatura de detecção automática de padrões de calibração. Do mesmo modo, o texto desta pesquisa faz uso frequente de referências à área de calibração de câmeras, tornando necessário um conhecimento básico sobre os princípios de geometria projetiva, transformações projetivas e os principais métodos de calibração de câmeras.

Este capítulo foi elaborado para situar o leitor sobre os conceitos envolvidos nos processos de calibração de câmeras e a importância da detecção segura dos dados de calibração. Inicialmente são introduzidos os conceitos da geometria projetiva, os quais formam a base que sustenta os processos matemáticos de calibração de câmeras. Mais adiante é apresentado um modelo de câmera que é melhor entendido relacionando propriedades da geometria projetiva tanto no espaço 3D e no espaço bidimensional (2D). São discutidas as transformações fundamentais do espaço projetivo: euclidiana, similaridade, *affine* e perspectiva, que modelam o sistema de formação de imagens de uma câmera. Em seguida é apresentado o algoritmo *Direct Linear Transformation* (DLT), que é a principal metodologia de estimação de parâmetros de câmera.

O texto das seguintes seções utiliza uma nomenclatura normalizada e resume a interpretação dos principais livros e artigos da área de Visão Computacional que estão ligados à geometria de câmeras e os métodos de calibração desses dispositivos. Partes do texto são baseadas nos trabalhos de Trucco e Verri (1998), Forsyth e Ponce (2002), Hartley e Zisserman (2004), Florczyk (2005) e Szeliski (2010), os quais apresentam uma discussão mais detalhada sobre esse assunto. O texto faz referências a conceitos de Álgebra Linear (STRANG, 1993), exigindo do leitor conhecimentos básicos sobre esse assunto.

2.2 Notação utilizada

A seguir, a discussão dos conceitos faz constantes referências a matrizes e elementos no espaço 2D e 3D, muitas vezes via representações diferentes. Sendo assim, algumas notações de escrita são definidas para facilitar a compreensão do texto.

A maioria das variáveis são representadas por letras pertencentes ao alfabeto latino com exceção de ângulos e planos que são expressos por letras gregas minúsculas e maiúsculas respectivamente. A representação de espaços vetoriais é feita por letras maiúsculas em negrito sem preenchimento, por exemplo $\mathbb{R}^2$ que é o espaço bidimensional euclidiano.

As constantes são expressas por letras minúsculas em itálico do tipo k . Vetores são letras em negrito podendo variar quanto ao espaço a que pertencem. Vetores em minúsculo são representações de elementos no espaço bidimensional e vetores em maiúsculo no espaço tridimensional. Por exemplo, $\mathbf{x}$ e $\mathbf{l}$ são elementos do espaço $\mathbb{R}^2$ mas não são compostos necessariamente por duas coordenadas. Da mesma forma, $\mathbf{X}$ é um elemento de $\mathbb{R}^3$. Para representar vetores em coordenadas homogêneas o símbolo $\sim$ é usado sobre a letra. Dessa forma, $\tilde{\mathbf{x}}$ é o vetor $\mathbf{x}$ expresso em coordenadas homogêneas em $\mathbb{R}^3$ e $\tilde{\mathbf{X}}$ é o vetor $\mathbf{X}$ expresso em coordenadas homogêneas em $\mathbb{R}^4$. A definição de coordenadas homogêneas é descrita na próxima seção. Matrizes são letras maiúsculas normais, sem negrito, por exemplo M . O símbolo $^\top$ representa a operação de transposição.

2.3 Geometria projetiva

Entender o processo de formação de imagens é essencial para a calibração de câmeras. Os parâmetros que modelam esse processo são os valores esperados pela calibração. Nesse contexto, uma das questões levantadas é como a imagem de um objeto é projetada em uma câmera, em seu plano de imagem, ou seja, qual a transformação que mapeia pontos pertencentes a um espaço 3D em um espaço 2D.

A geometria Euclidiana descreve de maneira elegante os relacionamentos entre elementos do espaço tridimensional partindo de conceitos intuitivos até a elaboração de teoremas sofisticados como a definição de distâncias, mas apresenta algumas contradições quando esses elementos são observados a partir de um ponto de vista, ou seja, por uma *perspectiva*. Nesse cenário, alguns dos postulados da geometria euclidiana não são aplicáveis, tal como o quinto que afirma a existência de retas paralelas e que elas não se cruzam (HARTSHORNE, 2000).

Em uma geometria que considera a perspectiva de visualização, os objetos têm aparência menor à medida que sua distância aumenta do ponto de observação e suas dimensões ao longo da linha de visão são relativamente menores àquelas que são perpendiculares. Essas propriedades fazem com que pontos situados em retas paralelas, à medida que se afastam do observador, pareçam convergir para um mesmo ponto que define a interseção entre essas retas. Desse modo, a propriedade de paralelismo é perdida na presença da perspectiva ao mesmo tempo que valores de distância também são degradados. Nesse caso persistem os postulados da definição de retas e o de colinearidade (CASSE, 2006) da geometria euclidiana.

A Geometria Projetiva é uma das geometrias que considera os efeitos provocados pela perspectiva. Ela compreende o estudo das propriedades geométricas que são invariantes a transformações projetivas (EGGAR, 1998). Enquanto que na geometria de Euclides a intersecção entre retas paralelas existe quando se considera a existência de pontos no infinito, na geometria projetiva os pontos no infinito são representados por pontos finitos.

Os elementos de uma geometria são essencialmente primitivas como pontos, retas e cônicas (HARTSHORNE, 2000). Eles são representados por vetores que admitem operações lineares de adição e escala, pertencendo a um espaço vetorial (HARTLEY; ZISSERMAN, 2004; CASSE, 2006).

Em uma geometria plana um ponto pode ser definido por duas coordenadas $(x, y) \in \mathbb{R}^2$. $\mathbb{R}^2$ é um espaço vetorial com duas dimensões e a combinação das coordenadas (x, y) é interpretadas como um único vetor dentro desse espaço. Desse modo, um ponto é definido por um vetor bidimensional do tipo $\mathbf{x} = [x \ y]^\top$.

De maneira semelhante, uma reta em um plano pode ser representada por uma equação da forma $ax + by + c = 0$, onde escolhas diferentes para as variáveis a , b e c geram retas diferentes. Assim, esse conjunto de parâmetros permite representar uma reta na forma vetorial $\mathbf{l} = [a \ b \ c]^\top \in \mathbb{R}^3$ que, apesar de ser um vetor tridimensional, descreve um elemento em $\mathbb{R}^2$, e representa um conjunto de pontos desse espaço. A relação entre retas e vetores é independente de qualquer fator de escala k , desde que $k \neq 0$, a reta $k\mathbf{l} = [ka \ kb \ kc]^\top = (ka)x + (kb)y + kc = 0$ é equivalente à reta $\mathbf{l}$ apesar de serem vetores com valores diferentes. Essa propriedade encaixa-se na definição de vetores equivalentes onde um vetor particular $\mathbf{l}$ na verdade representa uma classe de vetores em $\mathbb{R}^3$ definindo um conjunto de vetores homogêneos. O conjunto de vetores de classes equivalentes em $\mathbb{R}^3 - [0 \ 0 \ 0]^\top$ formam o espaço projetivo $\mathbb{P}^2$.

Uma reta $\mathbf{l} = [a \ b \ c]^\top$ passa por um ponto $\mathbf{x} = [x \ y]^\top$ se ambos satisfazem a equação $ax + by + c = 0$. Essa expressão equivale ao produto interno dos vetores $[x \ y \ 1]^\top \mathbf{l} = \tilde{\mathbf{x}}^\top \mathbf{l} = 0$, onde $\tilde{\mathbf{x}}$ representa o ponto $\mathbf{x}$ como um vetor em coordenadas homogêneas dentro de $\mathbb{R}^3$ adicionando 1 em sua última coordenada. Caso $\tilde{\mathbf{x}}$ seja um vetor $[kx \ ky \ k]^\top$, onde k é uma constante diferente de 0, a intersecção com a reta $\mathbf{l}$ se dá somente se $[x \ y \ 1]^\top \mathbf{l} = 0$. É natural então representar $\mathbf{x}$ por vetores homogêneos da forma $\tilde{\mathbf{x}} = [x_1 \ x_2 \ x_3]^\top$, sendo $\mathbf{x} = \begin{bmatrix} x_1 & x_2 \\ x_3 & x_3 \end{bmatrix}^\top$. Dessa maneira, $\mathbf{x}$ representa uma classe de vetores equivalentes em $\mathbb{R}^3$ que são elementos de $\mathbb{P}^2$.

Uma consequência direta da representação de retas e pontos em coordenadas homogêneas é a possibilidade de usar a álgebra linear para calcular seus relacionamentos. Por exemplo, a intersecção de duas retas $\mathbf{l}$ e $\mathbf{l}'$ é um ponto $\tilde{\mathbf{x}} = \mathbf{l} \times \mathbf{l}'$, onde o símbolo $\times$ é o produto cruzado desses vetores. De maneira semelhante, a reta que une dois pontos é dada por $\mathbf{l} = \mathbf{x} \times \tilde{\mathbf{x}}'$.

A operação de produto cruzado é equivalente ao cálculo do espaço nulo de uma matriz 3×3 . Supondo que $\mathbf{x}$ caia sobre as linhas $\mathbf{l}$ e $\mathbf{l}'$, pode-se dizer então que $\tilde{\mathbf{x}}$ é uma combinação linear desses dois vetores, de modo que uma matriz formada por eles é singular e, portanto, seu determinante é zero. Dessa forma pode-se afirmar que $\mathbf{x}$ é a intersecção das duas retas se e somente se

$$\det(\mathbf{M}) = \begin{vmatrix} \tilde{\mathbf{x}} & \mathbf{l} & \mathbf{l}' \end{vmatrix} = 0. \quad (1)$$

O mesmo raciocínio pode ser aplicado para o cálculo da reta que une dois pontos.

Outra vantagem de se usar coordenadas homogêneas é que pontos e retas no infinito passam a ter uma representação finita em $\mathbb{R}^3$. Supondo a intersecção de duas retas paralelas $\mathbf{l} = ax + by + c$ e $\mathbf{l}' = ax + by + c'$. Elas se encontram no ponto $\mathbf{l} \times \mathbf{l}' = (c' - c)[b \ -a \ 0]^\top$, que não tem sentido algum em $\mathbb{R}^2$ já que $\mathbf{x}$ seria representado por $[b/0 \ -a/0]^\top$, a não ser que seja considerado um ponto no infinito¹. Em $\mathbb{P}^2$ os pontos no infinito possuem a forma $[x_1 \ x_2 \ 0]^\top$ e são definidos apenas pela proporção $x_1:x_2$. O conjunto de pontos cuja terceira coordenada é igual a zero, ($x_3 = 0$), são conhecidos como pontos ideais, representados por $\tilde{\mathbf{x}}_\infty$ e correspondem a pontos no plano x_1 - x_2 em $\mathbb{R}^3$.

Dois pontos no infinito também definem uma reta no infinito. Sendo $\tilde{\mathbf{x}} = [x_1 \ x_2 \ 0]^\top$ e $\tilde{\mathbf{x}}' = [x'_1 \ x'_2 \ 0]^\top$, a reta que passa por esses dois pontos é dada por $\mathbf{l} = \tilde{\mathbf{x}} \times \tilde{\mathbf{x}}'$ e tem a forma geral $\mathbf{l}_\infty = k[0 \ 0 \ 1]^\top$. Consequentemente qualquer reta não paralela a $\mathbf{l}_\infty$ intercepta em algum ponto ideal $\tilde{\mathbf{x}}_\infty$.

A Figura 2 ilustra como essas primitivas são interpretadas em $\mathbb{R}^3$. Nessa figura, Π representa o plano de projeções. Cada ponto $\tilde{\mathbf{x}}$ em $\mathbb{R}^3$ projeta sua imagem $\mathbf{x}$ em Π . Dessa forma, $\mathbf{x}$ equivale a uma reta que atravessa a origem do sistema de coordenadas. A reta $\mathbf{l}$ que passa por dois pontos em Π define um plano que também passa pela origem desse sistema. O plano x_1 - x_2 representa $\mathbf{l}_\infty$ e os pontos no infinito correspondem a vetores nesse plano. O espaço projetivo $\mathbb{P}^2$ é dado pela combinação do plano de projeções Π e o plano x_1 - x_2 .

Figura 2: Geometria projetiva. Interpretação de $\mathbb{P}^2$ no espaço $\mathbb{R}^3$. Um ponto no espaço euclidiano projeta sua imagem no plano Π no ponto $\mathbf{x}$. Uma reta no plano Π define um plano que passa pela origem do sistema de coordenadas. O plano x_1 - x_2 corresponde a uma reta no infinito $\mathbf{l}_\infty$ e cada ponto nesse plano é um ponto no infinito $\tilde{\mathbf{x}}_\infty$. Fonte: Adaptada da Figura 2.1 do trabalho de Hartley e Zisserman (2004).

¹A constante $(c' - c)$ é desconsiderada por representar somente um fator de escala.

2.4 Transformações projetivas

Em um espaço projetivo tanto retas quanto pontos possuem a mesma representação (EGGAR, 1998), ou seja, em $\mathbb{P}^2$ ambos são vetores com 3 dimensões. Assim as operações entre pontos e retas podem ser tratadas da mesma maneira. Nesta seção, são apresentadas algumas das transformações básicas entre esses vetores e qual seu significado dentro do espaço projetivo.

Uma transformação projetiva é uma transformação linear que mapeia $\mathbb{P}^n \mapsto \mathbb{P}^n$ dentro do espaço $\mathbb{R}^{n+1}$. Ela possui a forma

$$\tilde{\mathbf{x}}' = \mathbf{H}\tilde{\mathbf{x}}, \quad (2)$$

onde $\mathbf{H}$ é uma matriz de dimensão $(n+1) \times (n+1)$ e representa um conjunto de transformações não singulares que mapeiam um plano de projeção em outro. Essa transformação preserva características projetivas em ambos os planos tais como a colinearidade e a proporção cruzada de distâncias (SONKA; HLAVAC; BOYLE, 2007). A restrição de não singularidade imposta a $\mathbf{H}$ significa que pontos distintos $\mathbf{x}$ são mapeados em pontos $\tilde{\mathbf{x}}'$ também distintos em outro plano. Outra característica dessa operação é que, uma vez que $\mathbf{H}$ representa uma transformação de vetores em coordenadas homogêneas, a transformação é invariante a qualquer fator de escala $k, \forall k \neq 0$, portanto somente as proporções entre seus elementos são importantes.

A transformação entre retas pode ser derivada por meio de seu relacionamento com pontos. Sabe-se que $\mathbf{l}^\top \tilde{\mathbf{x}} = 0$ é a condição exigida para a intersecção do ponto $\tilde{\mathbf{x}}$ com a reta $\mathbf{l}$, assim, com o objetivo de encontrar a transformação entre $\mathbf{l}'$ e $\mathbf{l}$, a condição de intersecção pode ser reescrita da forma $\mathbf{l}^\top \mathbf{H}^{-1} \mathbf{H} \tilde{\mathbf{x}} = 0$. Essa nova equação permite verificar que os pontos $\tilde{\mathbf{x}}' = \mathbf{H}\tilde{\mathbf{x}}$ caem sobre a linha $\mathbf{l}'^\top = \mathbf{l}^\top \mathbf{H}^{-1}$. Portanto,

$$\mathbf{l}' = \mathbf{H}^{-\top} \mathbf{l} \quad (3)$$

é a transformação que mapeia retas entre dois espaços projetivos, sendo $\mathbf{H}^{-\top}$ a notação equivalente a $(\mathbf{H}^{-1})^\top$.

Dentro do conjunto de operações que $\mathbf{H}$ pode representar, existem quatro subgrupos de transformações que são diferenciados em termos de quais propriedades permanecem invariantes após sua aplicação. São os grupos das transformações euclidianas, similares, *affine* e perspectivas. A seguir cada uma é apresentada.

1. **Transformação euclidiana:** também conhecida como transformação isométrica, corresponde a uma rotação seguida de uma translação. Possui a forma:

$$\tilde{\mathbf{x}}' = \mathbf{H}_E \tilde{\mathbf{x}} = \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}^\top & 1 \end{bmatrix} \tilde{\mathbf{x}} = \begin{bmatrix} \cos \theta & -\sin \theta & t_x \\ \sin \theta & \cos \theta & t_y \\ 0 & 0 & 1 \end{bmatrix} \tilde{\mathbf{x}}$$

Esse tipo de operação depende apenas de três variáveis: t_x e t_y que são responsáveis pela translação na primeira e na segunda coordenada respectivamente, e θ que define o ângulo de rotação. Nessa transformação as distâncias, os ângulos e as áreas permanecem inalterados já que R é uma matriz ortonormal e $\mathbf{t}$ é um deslocamento.

2. **Transformação de similaridade:** considera um fator de escala s na matriz de rotação. É similar à transformação euclidiana, mas com a liberdade de dilatação ou contração dos elementos no espaço, dado pelo escalar s . Nessa transformação a forma é preservada, os ângulos entre linhas não são alterados pela matriz de rotação, pela translação e nem pelo fator de escala. No entanto, as áreas não permanecem as mesmas para $s \neq 1$. Na forma matricial tem-se:

$$\tilde{\mathbf{x}}' = H_S \tilde{\mathbf{x}} = \begin{bmatrix} sR & \mathbf{t} \\ \mathbf{0}^\top & 1 \end{bmatrix} \tilde{\mathbf{x}} = \begin{bmatrix} s \cos \theta & -s \sin \theta & t_x \\ s \sin \theta & s \cos \theta & t_y \\ 0 & 0 & 1 \end{bmatrix} \tilde{\mathbf{x}}$$

3. **Transformação *Affine*:** Corresponde a uma transformação linear seguida de uma translação. Diferente das transformações euclidianas e de similaridade, uma transformação *Affine* não fica restrita somente à operação de rotação. A matriz R , agora chamada de A , admite qualquer conjunto de valores desde que $|A| \neq 0$.

$$\tilde{\mathbf{x}}' = H_A \tilde{\mathbf{x}} = \begin{bmatrix} A & \mathbf{t} \\ \mathbf{0}^\top & 1 \end{bmatrix} \tilde{\mathbf{x}} = \begin{bmatrix} a_{11} & a_{12} & t_x \\ a_{21} & a_{22} & t_y \\ 0 & 0 & 1 \end{bmatrix} \tilde{\mathbf{x}}$$

São invariantes a esse tipo de transformação a colinearidade, o paralelismo das linhas, as proporções entre os tamanhos de segmentos paralelos e a proporção entre áreas.

4. **Transformação perspectiva:** É a transformação mais importante para este trabalho. Ela modela uma transformação linear com a liberdade de incluir uma perspectiva de visualização, ou seja, permite que as projeções dos mesmos elementos variem em função da distância do plano de projeções. Essa transformação tem a forma:

$$\tilde{\mathbf{x}}' = H_P \tilde{\mathbf{x}} = \begin{bmatrix} A & \mathbf{t} \\ \mathbf{v}^\top & v \end{bmatrix} \tilde{\mathbf{x}} \quad (4)$$

A propriedade de perspectiva é alcançada pela substituição do vetor $\mathbf{0}^\top$ por um vetor $\mathbf{v}^\top$. Outro aspecto importante é que nem sempre é possível relacionar um fator de escala a essa operação, uma vez que v pode ser igual a 0. No entanto, essa característica não será explorada neste trabalho. v é considerado um número diferente de zero, sendo na maioria das vezes igual a 1.

Uma transformação projetiva pode ser decomposta por uma sequência de transformações mais simples. Na Equação (5), as matrizes H_S , H_A e H_P contribuem separadamente com suas características e invariâncias. H_P por exemplo carrega somente o vetor $\mathbf{v}^\top$ que é responsável pela perspectiva. A submatriz K é uma combinação linear e R uma matriz de rotação.

$$H = H_S H_A H_P = \begin{bmatrix} sR & \mathbf{t} \\ \mathbf{0}^\top & 1 \end{bmatrix} \begin{bmatrix} K & 0 \\ \mathbf{0}^\top & 1 \end{bmatrix} \begin{bmatrix} I & 0 \\ \mathbf{v}^\top & v \end{bmatrix} = \begin{bmatrix} A & \mathbf{t} \\ \mathbf{v}^\top & v \end{bmatrix} \quad (5)$$

De acordo com essa decomposição, a matriz A pode ser encarada como uma combinação linear da essência de cada tipo de transformação, $A = sRK + \mathbf{t}\mathbf{v}^\top$.

A Figura 3 apresenta o efeito de cada tipo de transformação citada neste texto. Ela ilustra o resultado das transformações discutidas sobre a figura de um quadrado. A translação está implícita em cada figura transformada já que estão deslocadas em relação à figura original. Como mencionado anteriormente, H_E é uma transformação euclidiana que admite também a rotação da figura. H_S adiciona a liberdade de escala, H_A é uma transformação linear não restrita a rotações e H_P permite a inserção de perspectiva na projeção.

Figura 3: Transformações projetivas.

Geometria projetiva em 3D

A maioria das propriedades da geometria projetiva $\mathbb{P}^3$ pode ser encarada como generalizações da geometria projetiva em um plano. Em $\mathbb{P}^3$ um ponto é representado por um vetor com quatro coordenadas homogêneas $\tilde{\mathbf{X}} = [X_1 \ X_2 \ X_3 \ X_4]^\top$, assim pode-se escrever $\mathbf{X} = \begin{bmatrix} \frac{X_1}{X_4} & \frac{X_2}{X_4} & \frac{X_3}{X_4} \end{bmatrix}^\top$. De forma semelhante ao espaço $\mathbb{P}^2$, quando $X_4 = 0$, o vetor $\tilde{\mathbf{X}}$ é a representação em coordenadas homogêneas de pontos no infinito em $\mathbb{R}^3$.

2.5 Calibração de câmeras

A calibração de câmeras é uma etapa vital para aplicações de visão computacional que precisam de informações métricas sobre o ambiente (HEMAYED, 2003). Ela pode ser resumida como o processo que busca determinar os parâmetros intrínsecos à geometria de formação de imagens, tais como a distância focal, ponto central de projeção, tamanho real de cada *pixel*, distorções induzidas pelo sistema de lentes e a inclinação entre os eixos de projeções. Muitas vezes, a calibração também considera parâmetros extrínsecos ao sistema como a posição (translação) e rotação da câmera em relação a um determinado sistema de coordenadas no mundo real (TRUCCO; VERRI, 1998). A estimação desses parâmetros é realizada relacionando pontos conhecidos no ambiente com suas respectivas projeções dentro do plano de imagem. Essa informação fornece os dados necessários para ajustar um modelo matemático simplificado que é apresentado a seguir.

Ter um sistema calibrado é equivalente a conhecer os valores numéricos dos parâmetros citados. Esses valores normalmente são dispostos como uma matriz de projeções P que modela a transformação projetiva de uma câmera específica (HARTLEY; ZISSERMAN, 2004).

2.5.1 Modelo de câmera *pinhole*

As câmeras, de uma maneira geral, constituem um sistema complexo que envolve um conjunto de lentes, espelhos e prismas (FORSYTH; PONCE, 2002). No entanto, para entender melhor o processo de formação de imagens, uma câmera normalmente é representada por meio de um modelo mais simples chamado de modelo *pinhole*. Esse modelo descreve uma câmera com um centro de projeções C , que é considerado a origem do sistema de coordenadas euclidiano, um ponto focal e um plano de projeções, que equivale ao plano de imagem da câmera. Essa representação é bem semelhante à Figura 2, diferenciando-se apenas quanto ao posicionamento do plano de projeções. Nesse caso, o plano de projeções não intercepta necessariamente o ponto $[0 \ 0 \ 1]^T$, mas sim o ponto focal $[0 \ 0 \ f]^T$, onde f é a distância focal da câmera e constitui uma de suas características construtivas. A Figura 4 ilustra esse modelo.

Pela Figura 4 é fácil perceber que um ponto no mundo real $\mathbf{X}_c = [X_c \ Y_c \ Z_c]^T$ projeta sua imagem no plano de imagem da câmera no ponto $\mathbf{x} = [x_c \ y_c]^T$, sendo definido pela intersecção da reta $\overline{C\mathbf{X}_c}$ com esse plano. O mapeamento é idêntico ao sistema projetivo $\mathbb{P}^2$, de forma que $\mathbf{X}_c$ é a representação de $\mathbf{x}$ em coordenadas homogêneas. Usando as relações entre triângulos equivalentes percebe-se que o ponto $\mathbf{X}_c$ é mapeado no ponto $\left[f \frac{X}{Z} \ f \frac{Y}{Z} \ f\right]^T$ sobre o plano de imagem, e que $\mathbf{x} = \left[f \frac{X}{Z} \ f \frac{Y}{Z}\right]^T$.

Se todos os pontos forem representados em coordenadas homogêneas, essa transfor-

Figura 4: Modelo *pinhole* de câmera.

mação pode ser expressa pela seguinte expressão matricial:

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} \mapsto \begin{bmatrix} fX_c \\ fY_c \\ Z \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} = \tilde{\mathbf{x}} = \mathbf{P}\tilde{\mathbf{X}}_c \quad (6)$$

A matriz $\mathbf{P}$ dessa expressão é chamada de matriz de projeção de câmera. Ela representa uma forma simples de modelagem, já que considera somente a distância focal como parâmetro. Por outro lado, essa mesma expressão pode comportar outras variáveis relacionadas aos aspectos construtivos desse dispositivo.

Na Equação (6), o modelo assume que o ponto principal $\mathbf{p}$ da Figura 4 coincide com o centro do plano de imagem, o que nem sempre é verdade devido às imperfeições no processo de montagem do sistema. Por esse motivo, deve ser considerado um deslocamento entre o centro real do plano de projeções e o ponto principal. Se p_x e p_y são os deslocamentos nos sentidos dos eixos x_c e y_c respectivamente, a transformação pode ser reescrita de forma que $[X_c \ Y_c \ Z_c]^\top \mapsto [f\frac{X}{Z} + p_x \ f\frac{Y}{Z} + p_y]^\top$. Em termos matriciais:

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} \mapsto \begin{bmatrix} fX_c + Z_cp_x \\ fY_c + Z_cp_y \\ Z \end{bmatrix} = \begin{bmatrix} f & 0 & p_x & 0 \\ 0 & f & p_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} \quad (7)$$

Outra questão a ser levantada é que nem sempre os eixos x_c e y_c formam um ângulo reto perfeito. Então é prudente considerar uma inclinação s entre esses eixos. Além disso, há a possibilidade dos *pixels* não serem perfeitamente quadrados mas sim retangulares, o que leva a fatores de escala diferentes em ambas direções. Uma forma de modelar tal eventualidade é considerar que f pode ser diferente em cada direção, sendo dependente do

número de *pixels* por unidade de distância m_x e m_y nas direções x_c e y_c respectivamente. Considerando todas essas variáveis na construção de um modelo mais genérico, a matriz P pode ser construída da seguinte maneira:

$$P = \begin{bmatrix} \alpha_x & s & x_0 & 0 \\ 0 & \alpha_y & y_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad (8)$$

onde $\alpha_x = fm_x$, $\alpha_y = fm_y$ e $[x_0 \ y_0]^T$ o deslocamento em função da proporção de *pixels* por unidade de distância, onde seus elementos são definidos por $x_0 = p_x m_x$ e $y_0 = p_y m_y$. P constitui a representação compacta de uma transformação homogênea de dimensão 3×4 que relaciona pontos no mundo real com suas projeções no plano de imagem em função de características intrínsecas à câmera. Quando a matriz P é decomposta na forma $P = K[I \mid \mathbf{0}]$, K é conhecida como a matriz de parâmetros intrínsecos ou matriz de calibração da câmera. Essa decomposição leva ao mapeamento dos pontos para a forma

$$\tilde{\mathbf{X}}_c \mapsto \tilde{\mathbf{x}} = K[I \mid \mathbf{0}]\tilde{\mathbf{X}}_c, \quad (9)$$

onde I é a matriz identidade.

Na Equação (9), os pontos no espaço, representados por $\tilde{\mathbf{X}}_c$, são pontos no sistema de coordenadas de câmera, ou seja, em relação ao centro $\mathbf{C}$. No entanto, na maioria das vezes, esses pontos estão disponíveis em função de outro sistema de coordenadas, chamado de sistemas de coordenadas de mundo. Nesse caso, existe uma relação entre esses dois sistemas que é descrita por uma rotação e uma translação. A Figura 5 ilustra essa situação. Os pontos no espaço possuem duas representações equivalentes, uma em função de coordenadas da câmera $\mathbf{X}_c$ e outra em função do sistema de coordenadas do mundo $\mathbf{X}_w$. $\mathbf{C}$ e $\mathbf{O}$ representam os centros de cada sistema, $\mathbf{t}$ é um vetor 3×1 de deslocamento e R uma matriz de rotação 3×3 formada por

$$R = R_z R_y R_x = \begin{bmatrix} \cos \phi & -\sin \phi & 0 \\ \sin \phi & \cos \phi & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos \gamma & 0 & \sin \gamma \\ 0 & 1 & 0 \\ -\sin \gamma & 0 & \cos \gamma \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{bmatrix}$$

sendo que ϕ , γ e θ são os ângulos de rotação dos eixos Z_w , Y_w e X_w respectivamente.

Em coordenadas homogêneas, $\tilde{\mathbf{X}}_c$ e $\tilde{\mathbf{X}}_w$ representam o mesmo ponto no plano de imagem, então $\tilde{\mathbf{X}}_c = R(\tilde{\mathbf{X}}_w - \tilde{\mathbf{C}}_w)$, onde $\tilde{\mathbf{C}}_w$ é o centro da câmera em coordenadas de mundo. Essa operação permite reescrever a Equação (9) considerando pontos em relação ao sistema de coordenadas de mundo:

$$\tilde{\mathbf{X}}_w \mapsto \tilde{\mathbf{x}} = KR[I \mid -\tilde{\mathbf{C}}_w]\tilde{\mathbf{X}}_w = K[R \mid \mathbf{t}]\tilde{\mathbf{X}}_w. \quad (10)$$

A matriz $[R \mid \mathbf{t}]$ constitui uma matriz 3×4 e representa os parâmetros extrínsecos do sistema de visualização, sendo $\mathbf{t} = -R\tilde{\mathbf{C}}_w$. A matriz $P = K[R \mid \mathbf{t}]$ modela o sistema

Figura 5: Modelo *pinhole* de câmera em relação ao mundo exterior.

projetivo da câmera levando em consideração parâmetros intrínsecos e extrínsecos, e nada mais é que um mapeamento linear entre pontos no mundo 3D e pontos bidimensionais no plano de imagem. Assim, a expressão de mapeamento entre esses pontos é definida por:

$$\tilde{\mathbf{x}} = \mathbf{P}\tilde{\mathbf{X}}_w \quad (11)$$

2.5.2 Estimação da matriz $\mathbf{P}$

O problema de estimação dessa matriz consiste em, dado um conjunto de correspondências $\mathbf{X}_i \leftrightarrow \mathbf{x}_i$, qual a matriz $\mathbf{P}$ que faz o mapeamento entre esses pontos, de forma que $\tilde{\mathbf{x}}_i = \mathbf{P}\tilde{\mathbf{X}}_i$, para todo i dentro do conjunto de correspondências? Pela Equação (11), observa-se que cada coordenada de $\tilde{\mathbf{x}}_i$ é uma combinação linear do ponto $\tilde{\mathbf{X}}_i$ definida por cada linha de $\mathbf{P}$. Assim

$$\lambda \tilde{\mathbf{x}}_i = \mathbf{P}\tilde{\mathbf{X}}_i = \begin{bmatrix} \mathbf{p}_1^\top \tilde{\mathbf{X}}_i \\ \mathbf{p}_2^\top \tilde{\mathbf{X}}_i \\ \mathbf{p}_3^\top \tilde{\mathbf{X}}_i \end{bmatrix}, \quad (12)$$

onde $\mathbf{p}_i^\top$ é a i -ésima linha de $\mathbf{P}$ e λ um fator de escala. Escrevendo $\tilde{\mathbf{x}}_i = [x_i \ y_i \ z_i]^\top$, o produto cruzado com $\mathbf{P}\tilde{\mathbf{X}}_i$ é dado por

$$\tilde{\mathbf{x}} \times \mathbf{P}\tilde{\mathbf{X}} = \begin{bmatrix} y_i \mathbf{p}_3^\top \tilde{\mathbf{X}}_i - z_i \mathbf{p}_2^\top \tilde{\mathbf{X}}_i \\ z_i \mathbf{p}_1^\top \tilde{\mathbf{X}}_i - x_i \mathbf{p}_3^\top \tilde{\mathbf{X}}_i \\ x_i \mathbf{p}_2^\top \tilde{\mathbf{X}}_i - y_i \mathbf{p}_1^\top \tilde{\mathbf{X}}_i \end{bmatrix} = \mathbf{0} \quad (13)$$

A Equação (13) pode ser remodelada como um sistema linear do tipo

$$\mathbf{A}_i \mathbf{p} = \begin{bmatrix} \mathbf{0}^\top & -z_i \tilde{\mathbf{X}}_i^\top & y_i \tilde{\mathbf{X}}_i^\top \\ z_i \tilde{\mathbf{X}}_i^\top & \mathbf{0}^\top & -x_i \tilde{\mathbf{X}}_i^\top \\ -y_i \tilde{\mathbf{X}}_i^\top & x_i \tilde{\mathbf{X}}_i^\top & \mathbf{0}^\top \end{bmatrix} \begin{bmatrix} \mathbf{p}_1 \\ \mathbf{p}_2 \\ \mathbf{p}_3 \end{bmatrix} = \mathbf{0} \quad (14)$$

onde $\mathbf{A}_i$ é uma matriz 3×12 construída pela correspondência i e $\mathbf{p}$ o vetor formado pelas linhas da matriz $\mathbf{P}$, totalizando doze incógnitas. No entanto, apesar de $\mathbf{A}_i$ possuir três equações, somente duas são linearmente independentes, uma vez que a terceira linha pode ser formada pela combinação linear das duas primeiras. Isso significa que um par de pontos correspondentes contribui com duas equações e pelo menos seis correspondências são necessárias para que o sistema tenha solução.

Uma solução exata para a Equação (14) exige um conhecimento preciso sobre as coordenadas de cada ponto, o que nem sempre é possível, já que as projeções assumem coordenadas no plano de imagem estando sujeitas a ruído e à truncagem no processo de digitalização. Se mais que seis correspondências estão disponíveis, esse sistema linear é classificado como super determinado e pode não ter solução exata. No entanto, é possível calcular uma aproximação de $\mathbf{p}$ pela minimização de $\|\mathbf{A}\mathbf{p}\|$. Uma solução direta para a Equação (14) é admitir que $\mathbf{p} = \mathbf{0}$, mas isso não seria uma solução interessante para o problema. Por esse motivo, a normalização de $\|\mathbf{A}\mathbf{p}\|$ é realizada sobre a restrição de $\|\mathbf{p}\| = 1$.

Esse método de estimação da matriz de projeção é chamado de Transformação Linear Direta, ou DLT como é mais conhecido (AZIZ; KARARA, 1971). A solução é alcançada empregando métodos numéricos tais como mínimos quadrados ou Decomposição em Valores Singulares, em inglês *Singular Value Decomposition* (SVD). Nesse último, o vetor solução $\mathbf{p}$ é o autovetor direito relacionado ao menor autovalor de $\mathbf{A}^\top \mathbf{A}$ constituindo o método mais utilizado pelos algoritmos de calibração (HARTLEY; ZISSERMAN, 2004).

O principal problema relacionado à calibração de câmeras é a imprecisão dos dados. Hartley e Zisserman (2004) apresentam uma versão mais detalhada no algoritmo DLT e ressaltam que, devido ao ruído e falta de precisão infinita nas coordenadas dos pontos no plano de imagem, o processo de estimação normalmente diverge da solução exata. Diante disso, como etapa de pré-processamento, eles introduzem uma normalização desses dados. Uma versão mais eficiente do algoritmo DLT inclui a normalização dos dados, a solução do sistema linear via SVD, uma etapa de ajuste fino conseguida por otimização não-linear e por último a desnormalização dos dados. Motivados pela mesma dificuldade, (TSAI, 1986) e (ZHANG, 2004b) também propõem métodos alternativos para esse processo.

Uma vez que a matriz $\mathbf{P}$ está disponível, os parâmetros de câmera podem ser determinados. A Equação (10) permite escrever $\mathbf{P}$ como o resultado da multiplicação entre as matrizes $\mathbf{K}$ e $[\mathbf{R} \mid \mathbf{t}] = [\mathbf{K}\mathbf{R} \mid \mathbf{K}\mathbf{t}]$. Sabe-se, pelo modelo *pinhole*, que a matriz de parâmetros intrínsecos $\mathbf{K}$ é triangular superior e que $\mathbf{Q}$ constitui uma matriz ortonormal que

descreve a rotação do sistema. Assim, o acesso aos parâmetros intrínsecos e extrínsecos pode ser obtido pela fatoração matricial RQ (AGULLO et al., 2010).

Uma versão da fatoração RQ, bem como a decomposição SVD, pode ser encontrada na biblioteca LAPACK, disponibilizada pela Netlib², que reúne um vasto número de operações de álgebra linear.

2.5.3 Considerações finais

Esse capítulo apresentou as principais ideias sobre a geometria projetiva e as transformações que são utilizadas nos modelos de câmera. Os conceitos discutidos funcionam como ferramentas cuja representação matemática é mais apropriada para o cálculo computacional. Também foi apresentado o modelo de câmera *pinhole* que é utilizado pela maioria dos trabalhos dedicados ao processo de calibração. Foi descrito o método linear de estimação da matriz P , que representa uma solução fechada para a calibração de câmeras. É importante ressaltar que parâmetros não-lineares como distorções provocadas pelas lentes não são considerados no algoritmo DLT devido à sua natureza puramente linear. No entanto, esse método muitas vezes serve como estimativa inicial para algoritmos iterativos de otimização não-linear que fornecem condições para o cálculo desses outros parâmetros. Em Salvi, Armangué e Batlle (2002) e Qi, Li e Zhenzhong (2010), é possível ter uma visão geral dos principais métodos de calibração que usam técnicas não-lineares para a otimização dos parâmetros de câmera.

---

²<http://www.netlib.org/> acessado em novembro de 2012

3 Padrões de Calibração

3.1 Considerações iniciais

Os padrões de calibração são objetos utilizados para facilitar a determinação de pontos de interesse e possuem papel importante nos processos tradicionais de calibração de câmeras. A grande diversidade de estudos na área de calibração de câmeras abre margens para o uso de diversos tipos de padrões, no entanto, existem poucos trabalhos científicos que abordam a classificação e avaliação desses aparatos. Este Capítulo apresenta uma visão geral sobre os diversos tipos de padrões de calibração e também uma abordagem classificatória dos mesmos. A Seção 3.4 expõe uma discussão sobre a construção desses padrões considerando sua precisão e as características que os compõem.

3.2 Padrões de calibração

Os algoritmos tradicionais de calibração de câmeras procuram estimar parâmetros de um modelo matemático, como o modelo *pinhole* da Seção 2.5.1, que representa o sistema projetivo de uma câmera real. Para isso, é comum considerar pontos reais e suas respectivas projeções no plano de imagem como os dados de entrada para a solução desse modelo. Os pontos reais, denominados *pontos de controle*, formam o conjunto de pontos cujas coordenadas são conhecidas com precisão em relação a uma referência presente no próprio ambiente.

O equacionamento do problema de calibração de câmeras (Seção 2.5.2) admite solução quando uma quantidade mínima de pontos é considerada. Para o algoritmo DLT por exemplo, são necessárias pelo menos oito correspondências para que o sistema tenha solução. No entanto, um processo de calibração que usa a quantidade mínima de pontos só apresenta bons resultados quando as posições dos pontos reais e suas respectivas projeções são conhecidas com precisão infinita, o que é dificilmente alcançado em situações práticas (HARTLEY; ZISSERMAN, 2004)

As fases do processo de calibração de câmeras são altamente dependentes da precisão com que os pontos de controle são identificados (TRUCCO; VERRI, 1998). Para que o processo seja restrito somente a um número mínimo de pontos é necessária uma métrica minuciosa do ambiente e também do plano de imagem. Em aplicações reais é muito difícil determinar a posição exata de cada ponto em ambos sistemas de coordenadas. Desse modo, o tamanho do conjunto de dados é aumentado na tentativa de minimizar a incerteza no cálculo dos parâmetros do modelo. Em um processo que requer intervenção manual

difficilmente atinge-se a acurácia necessária para o uso de poucos pontos e a identificação de um grande volume de dados é um exercício dispendioso para o operador do sistema. Por esses motivos, os trabalhos sobre calibração de câmeras frequentemente consideram o uso de *padrões de calibração*, ou *gabaritos*, como o principal meio de identificar seus pontos de controle.

Os padrões têm a função de tornar mais prática a identificação das coordenadas de cada ponto. Eles são aparatos visuais que usam cores de alto contraste para destacar *características* que evidenciam a presença de um ponto de interesse no mundo real (MALLON; WHELAN, 2007). As características são normalmente de natureza bidimensional, definidas pelo perfil da superfície de intensidades de *pixels* vizinhos dentro da imagem e são dispostas de uma maneira regular, de modo que cada uma define uma posição conhecida em relação à origem do padrão. Na maioria das vezes, a origem do padrão coincide com a origem do sistema de coordenadas do mundo real o que possibilita recuperar a posição 3D de cada ponto.

Em um padrão ideal, as características devem ser exclusivas à imagem do padrão de calibração para evitar ambiguidades com o restante da cena. Cada característica deve representar uma única posição no mundo exterior a fim de recuperar suas coordenadas 3D de maneira absoluta. Essa situação ideal facilita ao máximo a identificação dos pontos para os métodos manuais.

A principal vantagem do uso de padrões ideais é que a recuperação das coordenadas pode ser feita ao mesmo tempo que as características são extraídas da imagem. No caso particular em que cada característica é exclusiva ao padrão e auto identificável, cada uma delas é única na imagem e sua posição independe de qualquer outra referência. Por outro lado, construir padrões com essa complexidade exige o projeto individual de cada característica e também a elaboração de um algoritmo de extração específico para cada particularidade ou um sistema de reconhecimento mais complexo. Isso torna o trabalho difícil principalmente quando muitos pontos de controle são necessários.

Geralmente, por conta da facilidade de construção, a maioria dos trabalhos empregam características muito mais simples como retas, vértices, intersecções ou uma combinação desses elementos. Para reduzir a ambiguidade na cena, esses atributos são associados a *marcadores de calibração*, que são figuras geométricas ou combinações de características que dificilmente ocorrem naturalmente no ambiente. Os marcadores são dispostos de maneira regular, permitindo a localização relativa de um ou mais pontos de controle. É comum encontrar características ou marcos que funcionam exclusivamente como marcadores de orientação do padrão.

Um dos problemas com o uso de características como cantos, retângulos e circunferências, é que elas nem sempre são restritas à imagem do padrão de calibração. Elas podem ocorrer facilmente em ambiente antrópicos, onde é comum a presença de superfícies com bordas retangulares ou circulares e cantos por questões inerentes à imagem dos objetos

da cena ou pelo ruído presente na imagem. O uso de marcadores de calibração tende a diminuir esse problema, no entanto, a metodologia de extração desses marcadores baseia-se na detecção das mesmas características, comprometendo a localização dos mesmos.

O aumento do número de marcadores aumenta o número de características semelhantes na cena, o que torna o processo de localização um problema ambíguo e difícil de ser tratado. Em uma abordagem de processamento local, somente a posição relativa dos pontos de controle pode ser determinada, sendo necessária a detecção completa do padrão ou de outros marcadores específicos para determinar a posição real de cada ponto.

3.3 Tipos de padrões

De uma maneira geral, o objetivo de um padrão de calibração é representar pontos no mundo real. Não existe uma padronização de como esses dispositivos devam ser construídos a não ser pela restrição de que eles indiquem pontos no mundo 3D. Essa liberdade de construção possibilita que cada padrão seja projetado com o objetivo de fornecer melhores condições para os métodos de detecção do conjunto de pontos de controle. Por outro lado, a mesma liberdade implica também em uma grande diversidade de padrões, diminuindo a generalização do processo de calibração e dificultando a classificação desses objetos. No entanto, pode-se fazer uma diferenciação quanto à dimensão do objeto e quanto às características e marcadores usados para representar os pontos de controle.

É difícil determinar qual o primeiro tipo de padrão de calibração utilizado uma vez que a calibração de câmeras insere-se primeiramente no âmbito da fotogrametria (SLAMA, 1980). Essa é uma área mais antiga, onde seus métodos são caracterizados por operações puramente manuais dos processos de calibração de câmeras. Dentro do contexto de Visão Computacional, Brown (1971), um dos trabalhos pioneiros na área de calibração de câmeras e essenciais para sistemas de alta precisão, já utilizava um gabarito visual. Nesse trabalho Brown apresentou uma maneira de modelar principalmente as distorções provocadas pelas lentes de uma câmera. O método proposto baseava-se no fato de que linhas perfeitamente retas no mundo real deveriam ser projetadas também como linhas retas no plano de imagem a menos que sofram distorções. Para a construção do aparato de calibração, Brown posicionou uma chapa fotográfica atrás de sete linhas de prumo, daí o nome *plumb-line* para esse método de calibração, que era sensibilizada por uma fonte de luz. As linhas eram muito finas, espaçadas em 20mm de distância e seus pesos eram imersos em óleo para obter melhor estabilidade. Nesse trabalho em questão, cada chapa era rotacionada em 90° para uma segunda exposição.

A Figura 6 apresenta o padrão de calibração utilizado por Brown (1971). Ele se insere na classe de padrões bidimensionais cujas características de interesse são exclusivamente as retas que, neste caso, são produzidas pela sombra das linhas de prumo. Uma das vantagens dos métodos *plumb-line* é que somente a curvatura das linhas é suficiente para

Figura 6: Padrão de calibração usado pelo trabalho de Brown (1971). As linhas são perpendiculares e espaçadas em 20mm. Esse padrão resgata um dos primeiros métodos de calibração de câmeras baseado em retas denominado de calibração *plumb-line*, quando linhas de prumo eram usadas para gerar retas paralelas.

o processo de calibração sendo dispensável a localização de cada uma. Isso permite a utilização de marcadores situados em um único plano (ROSTEN; LOVELAND, 2011, seção 1.2.2, pg. 78).

Os padrões de calibração exercem forte influência sobre os sistemas de calibração e estão extremamente correlacionados aos métodos matemáticos empregados por esses sistemas (ALBARELLI; RODOLÀ; TORSELLO, 2010). Por esse motivo, é difícil dissociar a metodologia de estimação de parâmetros do tipo de dados que são utilizados no processo de calibração. Pode-se diferenciar os padrões como aqueles que fornecem coordenadas unidimensionais, bidimensionais ou tridimensionais, que também é um modo de classificação dos métodos de calibração. A variabilidade dos métodos se dá pelas diferentes formas de apresentar tais informações. Zhang (2004a) classifica os métodos de calibração de acordo com a dimensão dos conjunto de dados. Eles são categorizados como:

Calibração usando objetos 3D: Os métodos de calibração são baseados em técnicas convencionais, mais próximas à fotogrametria, que usam objetos cuja geometria no espaço tridimensional é conhecida rigorosamente. Os padrões usualmente consistem de 2 ou 3 planos perpendiculares entre si. Esse tipo de aparato requer uma métrica precisa de suas partes mecânicas e uma configuração mais elaborada. Uma alternativa ao uso de vários planos se dá quando somente um único plano é usado para fornecer coordenadas 3D. Nesse caso, a posição do plano deve ser conhecida de maneira precisa (TSAI, 1986). Nessa categoria, a estimação dos parâmetros de câmera pode ser realizada usando uma única imagem do gabarito e normalmente apresenta melhores resultados (ZHANG et al., 1997; SALVI; ARMANGUÉ; BATLLE, 2002; QI; LI; ZHENZHONG, 2010; SAMPER et al., 2013).

Calibração usando objetos 2D: Nessa classe os métodos requerem imagens de um único padrão plano em diferentes orientações e posições, sendo desnecessário o co-

nhecimento desse movimento. Os resultados são menos precisos em comparação às técnicas 3D, porém aceitáveis pela maioria dos sistemas DVS. A facilidade de construção dos padrões dessa categoria faz deles os aparatos mais usados pelas aplicações de calibração (CHATTERJEE; ROYCHOWDHURY, 2000; ZHANG, 2000; SAMPER et al., 2013).

Calibração usando objetos 1D: Os aparatos são formados por pontos alinhados sobre uma linha reta. O processo de calibração se dá de maneira semelhante aos que empregam padrões 2D pois também requer a apresentação do padrão em orientações e posições diferentes para a câmera. Em contrapartida essa técnica normalmente exige um número maior de imagens do padrão (ZHANG, 2004b; LOAIZA; RAPOSO; GATTASS, 2011).

Calibração sem o uso de objetos de referência: Os métodos dessa categoria, denominada autocalibração, consideram uma abordagem zero dimensional (0D) pois não empregam nenhum objeto de calibração. Os trabalhos dessa linha de pesquisa (HEMAYED, 2003) exigem que a câmera seja movimentada em uma cena estática e, aproveitando-se da rigidez do ambiente, tanto parâmetros intrínsecos como extrínsecos são estimados. No entanto, a acurácia é prejudicada pela grande quantidade de parâmetros desconhecidos e também por configurar um problema matemático mais complexo (ZHANG, 2004a). Uma restrição comum é que o ambiente deve fornecer imagens com grande quantidade de pontos de interesse tornando muito difícil a calibração em situações com predominância de imagens com cores homogêneas. Nenhuma informação sobre as posições dos pontos de controle é necessária desde que as correspondências entre eles em pelo menos duas imagens estejam disponíveis. Na maioria das publicações, a calibração é obtida somente sob um fator de escala e a capacidade métrica dos resultados precisa ser posteriormente ajustada (HARTLEY; ZISSERMAN, 2004).

Muito esforço científico é direcionado principalmente à qualidade e estabilidade dos resultados da calibração, para isso os trabalhos também recorrem a diferentes configurações mecânicas dos padrões (QI; LI; ZHENZHONG, 2010). Esse campo da ciência é bem explorado, principalmente quanto aos métodos numéricos empregados na solução dos modelos de câmera.

Pode-se verificar que os métodos que usam padrões 3D costumam apresentar melhores resultados. O conjunto de dados com informações 3D permite considerar um número maior de restrições inerentes ao sistema projetivo da câmera e os parâmetros são melhores ajustados. Essa é uma afirmação comum entre os pesquisadores e pode ser observada nos trabalhos de revisão e avaliação de métodos de calibração (SALVI; ARMANGUÉ; BATLLE, 2002; SAMPER et al., 2013). No entanto, esses padrões apresentam a desvantagem de serem

de difícil construção e normalmente exigem reparos e reavaliações de suas medidas todas as vezes que são utilizados.

Os algoritmos que utilizam padrões 2D são mais populares pela sua flexibilidade e por seus padrões serem de fácil construção. Com uma impressora disponível pode-se utilizar *softwares* de desenho digital para criar a imagem do gabarito e depois imprimi-la em uma folha de papel – recomenda-se colar a folha em uma peça plana e rígida para garantir que todas as características sejam coplanares. Em outro caso, pode-se criar a mesma imagem diretamente em uma plataforma de desenho usando ferramentas simples como régua e esquadro. Por outro lado, um número maior de variáveis do modelo são desconhecidas e mais imagens precisam ser processadas (SALVI; ARMANGUÉ; BATLLE, 2002; QI; LI; ZHENZHONG, 2010; STROBL; HIRZINGER, 2011).

Os padrões 1D se encaixam em uma vertente de pesquisa mais recente que é impulsionada principalmente por aplicações onde a calibração de múltiplas câmeras é necessária. Nesse tipo de problema é pouco provável que os padrões 3D e 2D sejam visualizados por todos os equipamentos de imagem simultaneamente, mas isso não é um problema para padrões 1D devido a sua natureza estrutural (ZHANG, 2004b).

A Figura 7(a) mostra um exemplo de padrão 3D. A imagem tem como fonte o trabalho de Zhang (2004a) mas é frequentemente relatado como o padrão usado por Tsai (1986). Nesse exemplo, dois planos de calibração são impressos, colados em uma placa rígida e plana sendo posteriormente posicionados de maneira perpendicular entre si. Na Figura 7(b) uma placa de madeira serve de apoio para marcadores que saltam fora do plano a fim de fornecer dados em diferentes profundidades. A construção desse padrão normalmente requer algumas horas de trabalho e a identificação das características é altamente dependente de intervenções manuais (CHIABRANDO et al., 2009). Forbes, Voigt e Bodika (2002) apresentam um sistema de calibração que se julga ser de baixo custo, completamente automático e preciso, contudo os autores usam padrões mais complexos, formados por várias faces e características autocodificadas. A Figura 7(c) ilustra alguns dos padrões no formato de cubo e icosaedro. Nesse exemplo, os marcadores são autocodificados (ver Seção 3.3.3).

A Figura 8(a) mostra a versão bidimensional do aparato de Tsai (1986). Esse padrão tornou-se famoso pela facilidade de construção e após a solução apresentada por Zhang (2000) que permite que o padrão seja movimentado livremente no ambiente. Já a Figura 8(b) ilustra um objeto pouco convencional que usa marcadores autocodificados (XIAO et al., 2010). A Figura 8(c) mostra o padrão de calibração que compõe a biblioteca OpenCV (BRADSKI, 2000). Esse padrão é caracterizado pela presença de círculos assimétricos, sendo relatado por apresentar melhores resultados de calibração para a reconstrução 3D quando comparado a padrões com marcos retangulares (MALLON; WHELAN, 2007). Ainda no mesmo conjunto de exemplos, as Figuras 8(d) e 8(e) apresentam padrões unidimensionais usados em Zhang (2004b) e Loaiza, Raposo e Gattass (2011) respectivamente.

Figura 7: Exemplos de padrões 3D. a) Padrão formado por dois planos perpendiculares (ZHANG, 2004a). b) Padrão usado um único plano cujos marcadores possuem profundidades diferentes (CHIABRANDO et al., 2009). c) Padrões que usam estruturas mais complexas. Nesse exemplo os marcadores usados são autocodificados (FORBES; VOIGT; BODIKA, 2002).

Outra maneira de classificar os objetos de calibração é atentando-se aos diferentes tipos de características e marcadores. Nos trabalhos mais comuns, o perfil visual do marco de calibração fornece tais características, sendo predominantemente os cantos as principais entidades de interesse. Em outros casos as propriedades inerentes ao próprio marco são relacionadas aos pontos de calibração, podendo ser desde o centro da figura geométrica ou a linearidade de suas bordas. As seções 3.3.1, 3.3.2 e 3.3.3 apresentam uma discussão sobre esses tipos de padrões.

3.3.1 Marcadores de cantos e interseções

Cantos e intersecções certamente são as características mais utilizadas na construção de padrões de calibração. Um canto, ou em inglês *corner*, pode ser definido como uma região da imagem cujo perfil de intensidades favorece a presença de bordas, obviamente na separação de regiões com cores de alto contraste, que formam ângulos. Outra definição mais simples admite que um canto é formado basicamente quando segmentos ou bordas com direções diferentes se encontram ou quando uma única borda apresenta uma curvatura abrupta (GONZALEZ; WOODS, 2006; NIXON; AGUADO, 2008). Já a intersecção, dentro do contexto de marcadores de calibração, é entendida como uma região onde a superfície de intensidades é caracterizada como um canto, mas que tende a ser mais simétrica. Neste trabalho essa característica é referenciada como *canto-x* (ver Seção 5.3).

Um dos registros mais antigos pode ser encontrado em um dos trabalhos mais influentes na área de calibração de câmeras, Tsai (1986). O autor empregou um padrão composto de quadrados pretos sobre um fundo branco igualmente espaçados onde os características correspondiam aos vértices, ou *cantos*, de cada marcador. Zhang (2000), outro trabalho não menos importante, também emprega o mesmo marcador. O uso desse padrão foi popularizado também devido aos trabalhos de Harris e Stephens (1988) e Shi e Tomasi

Figura 8: Exemplos de padrões 2D e 1D. Padrões bidimensionais (a,b,c) e unidimensionais (d,c). a) Aparato formado por marcadores retangulares impresso em papel (ZHANG, 2000); b) Padrão em forma de cruz com marcadores autocodificados (XIAO et al., 2010); c) Um dos padrões usados pela biblioteca OpenCV (BRADSKI, 2000); d) Padrão linear formado por esferas brancas dispostas em uma linha reta rígida (ZHANG, 2004b); e) As características estão dispostas linearmente e são representadas por pontos luminosos (LOAIZA; RAPOSO; GATTASS, 2011).

(1994) sobre detecção de cantos, que permitia automatizar a detecção de características e aprimorar a localização dos pontos de controle no plano de imagem. Esses detectores influenciaram fortemente uma boa parte dos estudos sobre calibração de câmeras que, a partir do final da década de 80, passaram a adotar cantos como o principal modo de representar seus pontos de calibração dentro da cena.

Os padrões frequentemente adotam uma estrutura formada por desenhos retangulares isolados como o da Figura 8(a). Outra variação se dá quando se tem quadrados conectados por seus vértices e intercalados por cores diferentes, simulando um tabuleiro de xadrez. A Figura 9 apresenta um exemplo do padrão *Chessboard*.

Alguns trabalhos apresentam marcadores diferentes para a mesma característica, como por exemplo Krüger e Wöhler (2011) que utilizam quatro padrões diferentes para avaliar seu método de refinamento da localização dos pontos de controle. Os autores usam 100 imagens de cada padrão e os pontos de controle têm sua localização inicial determinada manualmente (Figura 10). Já Habib e Morgan (2005) e Pullivelli (2005) usam intersecções de retas e pontos como características de interesse seguindo o mesmo princípio de faci-

Figura 9: Padrão de calibração *Chessboard*. As características correspondem às intersecções entre os quadrados.

dade de construção (Figuras 11(a) e 11(b) respectivamente). As retas contribuem para o cálculo da distorção das lentes e os pontos fornecem restrições métricas e geométricas no processo de calibração desenvolvido.

Figura 10: Outros marcadores de cantos. Nesse exemplo os marcadores fornecem características como cantos definidos por ângulos retos e cantos-x para representar a posição dos pontos de controle. Fonte: Krüger e Wöhler (2011).

Figura 11: Padrões que usam retas e pontos. As linhas retas contribuem para o cálculo dos parâmetros do modelo de distorções da lente e os pontos, representados pelas cruzes na imagem, fornecem informações para a estimação dos demais parâmetros intrínsecos à câmera. Fontes: Habib e Morgan (2005) e Pullivelli (2005).

3.3.2 Marcadores circulares

Os padrões dessa categoria são caracterizados, como o próprio nome já diz, pela presença de formas circulares. Nesse caso, as características de interesse não são diretamente relacionadas ao perfil de cores do marcador mas sim por suas propriedades, tais como o centro da circunferência, sua curvatura e retas tangenciais (MATEOS, 2000, seção 1.4).

Em detrimento da acurácia de detecção dos pontos de controle no padrão de Tsai e os de linhas retas, Heikkilä (2000) apresenta um estudo que comprova a superioridade do uso de marcadores circulares sobre retângulos isolados e usando como ponto de interesse o centro de cada forma cônica. Um outro trabalho do mesmo autor, (HEIKKILÄ, 1998), reporta que essa representação fornece meios mais apurados de detecção chegando a uma precisão de 1/100 contra 1/10 para os sistemas que usam marcadores retangulares (BOUGUET, 2004; KANNALA; BRANDT, 2006; XIAO; FISHER, 2010).

A Figura 8(c) é um bom exemplo dessa classe de padrões, no entanto o uso de padrões simétricos é mais comum. A Figura 12 apresenta um exemplo de marcador circular convencional e uma variação do mesmo, denominado marcador em anel.

Figura 12: Exemplos de padrões circulares. a) Padrão 3D que usa círculos dispostos de maneira simétrica nas faces do objeto de calibração (HEIKKILÄ, 2000). b) Padrão 2D com uma variação de marcadores circulares. A imagem mais à direita mostra o ponto de interesse que é comum às circunferências do marcador (ZHOU; ZHANG, 2008).

3.3.3 Marcadores autocodificados

Os marcadores autocodificados, conhecidos também como autoidentificáveis, são marcas especiais que carregam consigo uma mensagem codificada. Essa abordagem é interessante por promover uma representação mais controlada dos pontos de controle no ambiente com o objetivo de reduzir a detecção de falsos positivos. Sua robustez baseia-se na regularidade do marcador e também em sua unicidade na cena, sendo muito difícil a ocorrência natural de marcas idênticas em uma mesma imagem. Todos os marcadores são

construídos sob um mesmo conjunto de regras e determinadas regiões são reservadas para a armazenagem da mensagem. Essas regras são relacionadas à geometria do marcador e a forma com que a mensagem é codificada (AHN; RAUH; KIM, 2001).

O uso desse tipo de padrão se tornou comum entre as aplicações DVS, principalmente aquelas relacionadas à Realidade Aumentada (RA) com o uso de marcadores *Augmented Reality Tag* (ARTag)¹. Esses marcadores são formados basicamente por um retângulo que delimita a região total do perfil e seu interior é composto por estruturas retangulares cuja distribuição espacial é usada para a codificação de mensagens (OWEN et al., 2002).

Com essas marcas é possível identificar a pose da câmera em relação ao marcador e assim inserir objetos virtuais tridimensionais na cena em tempo real (KATO; BILLINGHURST, 1999). Para o leitor que deseja trabalhar com marcadores ARTag e RA, o pacote *AR-Toolkit*² é um excelente ponto de partida. Outra aplicação comum se dá na robótica móvel para a autolocalização de robôs autônomos (FIALA, 2004). Padrões são associados a pontos específicos no mapa e, assim que visualizados, fornecem meios para que o robô determine sua posição.

Geralmente, as aplicações que exploram o uso de marcadores de realidade aumentada têm interesse somente nos parâmetros extrínsecos que modelam a orientação do ARTag em relação à câmera. Além disso, é comum haver múltiplos marcadores inseridos na cena em diversas orientações e posições. A orientação é calculada em função dos vértices do retângulo delimitador e a mensagem codificada em seu interior serve como atributo de distinção entre eles.

Dentro do contexto de calibração, o emprego de um único marcador ARTag é insuficiente para a recuperação dos parâmetros intrínsecos ao sistema. Isso porque cada marca fornece somente um conjunto limitado de pontos de controle que são restritos à região do retângulo delimitador da imagem. Os trabalhos que usam esses padrões costumam empregar vários marcadores ARTag e associar um para cada ponto de controle. A vantagem imediata dessa abordagem é que a construção da base de dados é realizada automaticamente, e praticamente livre de erros de correspondência. A posição 3D é automaticamente medida quando o marcador é decodificado. O trabalho de Fiala e Shu (2007) é um exemplo de aplicação que usa marcadores ARTag para representar o padrão de calibração, cuja metodologia se dá do mesmo modo como já foi explicado.

Outra característica desses sistemas é o tempo reduzido de calibração, visto que a correspondências dos pontos, etapa mais demorada de todo o processo, é automatizada. Em aplicações onde a calibração de câmeras é um processo rotineiro, como produtoras de conteúdo multimídia em geral, essa é uma ótima alternativa. Uma aplicação comercial que pode ser citada é a empresa LightCraft³ que fornece soluções para a indústria do

---

¹Detalhes sobre o marcador ARTag disponível em: www.artag.net.

²*ARToolkit* é uma biblioteca aberta que implementa mecanismos de detecção de marcadores auto-identificáveis e também de RA. Disponível em: www.hitl.washington.edu/artoolkit/.

³Sítio da empresa LightCraft disponível em: www.lightcrafttech.com/.

cinema e emissoras de televisão. Seu sistema permite que sejam inseridos objetos virtuais em filmagens ao vivo em campo, o que exige a constante calibração do sistema. Para isso seus produtos vem com um *kit* de calibração com marcadores autocodificados. As Figuras 8(b) e 7(c) se encaixam nessa classe de padrões. A Figura 13 apresenta o marcador usado por Fiala e Shu (2007) e da aplicação comercial citada. Ahn, Rauh e Kim (2001) apresentam uma revisão sobre alguns desses padrões, ilustrados na Figura 14.

Figura 13: Exemplos de padrões com marcadores ARTag. (a) Padrão usado em Fiala e Shu (2007). (b) Exemplo do padrão usado pela empresa LigthCraft.

Figura 14: Exemplos de outros padrões autoidentificáveis. A maioria é empregado no meio industrial. Fonte: Ahn, Rauh e Kim (2001).

Esses padrões são os que mais se aproximam dos padrões ideais devido a suas características de unicidade e autoidentificação, no entanto, existem algumas restrições para o uso dessas marcas. A detecção tem sucesso quando é possível decodificar a mensagem do marcador o que acaba exigindo boa nitidez da imagem. Essa restrição é sensível à posição do marcador, sendo difícil sua detecção quando os padrões encontram-se muito distantes da câmera. Esse efeito implica no uso de formas maiores que entra em conflito com a

necessidade de se ter muitos marcadores na imagem. Além disso, a detecção só pode ser feita se o marcador for visualizado por completo.

Entretanto, marcadores ARTag nem sempre são sinônimos de precisão. Seu algoritmo de detecção é baseado em imagens binárias que usualmente necessitam de limiares. Isso torna esse algoritmo muito sensível à iluminação. Em um ambiente de calibração, onde precisão faz toda a diferença, o uso de marcadores de RA inspira precaução.

Atualmente existe uma demanda por esses marcadores principalmente na indústria em aplicações de inspeção visual e automatização de processos (AHN; RAUH; KIM, 2001). Poucas aplicações são voltadas exclusivamente para a calibração de câmeras (FORBES; VOIGT; BODIKA, 2002; XIAO; FISHER, 2010; FORMAN et al., 2011).

3.4 Precisão dos padrões de calibração

Enquanto a dimensão do padrão tem forte ligação com o método e a estabilidade do algoritmo de calibração, a característica utilizada está mais correlacionada à precisão com que as projeções são aferidas no plano de imagem. Na verdade, a categoria de métodos é decisiva na qualidade dos resultados, porém nada adianta técnicas de estimação super avançadas se a localização das características é frequentemente perturbada por ruídos tendenciosos (ALBARELLI; RODOLÀ; TORSELLO, 2010).

Existe uma variedade de trabalhos que buscam aprimorar a localização dos pontos de controle. Como imagens são geradas a partir de uma amostragem espacial (GONZALEZ; WOODS, 2006, cap. 2), a localização das projeções é dada somente em valores discretos, o que tem efeito prejudicial na estimação dos parâmetros de câmera. As técnicas que tentam aproximar a posição correta das projeções no plano de imagem são chamadas de técnicas *subpixel*, pois conseguem extrair informações com uma resolução muito maior que a própria resolução do sensor de imagem (SHORTIS; CLARKE; SHORT, 1994).

A solução comumente empregada faz uso da maximização local da superfície gerada pela resposta dos próprios detectores de características, como os de Harris e Stephens e Shi e Tomasi. Algumas pesquisas já usam a característica linear dos marcadores e identificam a posição *subpixel* calculando a intersecção de retas que melhor se ajustam a suas bordas (YIMIN et al., 2010; ESCALERA; ARMINGOL, 2010; KRÜGER; WÖHLER, 2011). Outros trabalhos baseiam-se em técnicas mais simples, como o centro de massa da região próxima ao ponto de calibração (ZHAO et al., 2011; BENNETT; LASENBY, 2013).

Heikkilä (2000) propõe o uso de marcadores circulares considerando que marcadores retangulares tradicionais, como os usados no trabalho de Tsai (1986), são influenciados por um ruído tendencioso provocado principalmente pela assimetria inerente à característica de interesse. Nesse caso causado pela iluminação direcional dos cantos do marcador (ALBARELLI; RODOLÀ; TORSELLO, 2010). Outros artigos apresentam soluções com o mesmo tipo de marcador: (MATEOS, 2000; ZHOU; ZHANG, 2008; XIAO; FISHER, 2010; BAOZHANG;

YANPING, 2010).

Em aplicações que requerem ajustes mais finos, a escolha do conjunto: dimensão, marcador e algoritmo de otimização tem efeito direto na integridade dos resultados. Atualmente sabe-se que o modelo de câmera *pinhole* é um modelo razoável para o comportamento projetivo de uma câmera, no entanto, ainda não se sabe quais são todas as fontes de ruído envolvidas no processo de formação de imagens. Essa é uma área de pesquisa ainda em aberto.

Mallon e Whelan (2007) examinam o comportamento de detectores *subpixel* sobre marcadores circulares e em forma de “X”, provenientes do padrão *Chessboard*. Os autores consideram que, em câmeras reais, os padrões estão sujeitos a dois tipos de transformações. A primeira, a transformação projetiva que é definida por aspectos lineares e tem como consequência a posição 3D do ponto de controle. E a segunda, um conjunto de transformações não lineares oriundas das aberrações provocadas pelas lentes. Com isso, são considerados dois tipos de tendências, ou *bias*, no modelo: *bias* de perspectiva e *bias* de distorção. Como resultado da pesquisa tem-se que os marcadores circulares são invariantes à perspectiva enquanto que o padrão *Chessboard* lida melhor com as distorções provocadas pelas lentes.

A indicação de que os parâmetros de câmera são influenciados por ruídos tendenciosos tem implicação direta em como os resultados são avaliados. A maioria das técnicas assumem que o ruído presente no modelo possui uma distribuição gaussiana centrada na média, o que não acontece na prática.

Albarelli, Rodolà e Torsello (2010) apresentam uma imagem, Figura 15, que permite identificar o comportamento tendencioso do erro de reprojeção de uma câmera calibrada com um padrão xadrez.

Figura 15: Erros de reprojeção de um sistema calibrado. Do lado esquerdo o erro em relação aos pontos de controle no mundo 3D, do lado direito os erros no plano de imagem. Fonte: Albarelli, Rodolà e Torsello (2010).

Na Figura 15, os erros são ampliados por um fator de cem para facilitar sua visualização. No lado esquerdo da figura, os pontos em vermelho representam o ponto de

controle real no mundo 3D e os segmentos, em azul, a distância do ponto estimado. Esses resultados fortemente sugerem que as medições estão sujeitas a ruídos com distribuição tendenciosa e média diferente de zero. Do lado direito tem-se os erros de medição no plano de imagem. Nesse caso, o erro tem um comportamento mais isotrópico, o que indica a robustez do padrão xadrez em relação às distorções radiais.

Só recentemente alguns trabalhos tem atentado ao fato de que nem sempre os padrões são construídos com alto grau de precisão. Eles sofrem com imperfeições mecânicas, falta de acurácia das ferramentas de construção e também com a rugosidade da peça (STROBL; HIRZINGER, 2011). Os resultados sugerem que, além do ruído considerado em Mallon e Whelan (2007), a estimação também está sujeita a erros sistemáticos provocados pela falta de precisão dos equipamentos de impressão.

A avaliação do quanto as imprecisões construtivas do padrão de calibração, bem como os tipos de características visuais usadas, interferem no processo de calibração de câmeras ainda é um assunto que carece de estudos mais aprofundados.

3.5 Considerações finais

Esse Capítulo apresentou uma revisão sobre os padrões de calibração que aparecem como objetos acessórios dos trabalhos de calibração de câmeras. O texto contribui para elucidar a necessidade de se pesquisar tal assunto, evidenciando a relação desses objetos com os resultados finais de calibração. Duas metodologias de classificação foram propostas, uma baseada na dimensão do padrão e outra no tipo de marcador empregado. Adicionalmente, uma discussão sobre a modelagem do ruído e a precisão na construção dos padrões foi apresentada.

4 Trabalhos Relacionados

4.1 Considerações iniciais

Este Capítulo apresenta uma revisão de trabalhos que abordam problemas similares aos estudados nesta tese, envolvendo a detecção automática de padrões xadrez para a calibração de câmeras. Inicialmente o texto apresenta os principais problemas enfrentados pelos trabalhos que se propõem a detectar esses padrões e em seguida faz a apresentação e discussão do estado da arte desse assunto.

4.2 Definição do problema

Muitos dos processos de calibração de câmeras necessitam de uma base de dados que relaciona pontos de controle presentes no mundo real com suas respectivas projeções no plano de imagem. Essa exigência é comum nas aplicações onde as câmeras devem ser calibradas metricamente (BROWN, 1971). A extração dos pontos de controle muitas vezes é considerada trabalhosa quando requer a intervenção do usuário. Apesar da calibração ser um assunto bastante explorado, poucas metodologias de extração automática dos pontos internos de padrões de calibração estão disponíveis (QI; LI; ZHENZHONG, 2010).

A maioria dos métodos de calibração são construídos a partir de equações fechadas, o que os tornam sensíveis à presença de ruídos. Para obter resultados mais estáveis, esses métodos requerem dados consistentes e também o conhecimento mais preciso sobre a localização dos pontos que são considerados no processo de calibração (HEIKKILÄ, 2000). Por esses motivos, os trabalhos sobre calibração de câmeras necessitam da intervenção do usuário para a detecção segura dos pontos de calibração (ZHANG, 2000; HARTLEY; ZISSERMAN, 2004; TRUCCO; VERRI, 1998). Por outro lado, o uso de imagens digitais impõe a condição de que cada coordenada na imagem é representada de maneira discreta, sendo difícil identificar precisamente a localização desses pontos manualmente (KRÜGER; WÖHLER, 2011).

Alguns trabalhos sugerem o uso de muitos pontos de calibração, que representam diversas poses do padrão, de modo a obter uma melhor amostragem do sistema projetivo da câmera e, conseqüentemente, diminuir a influência da imprecisão relacionada à localização desses pontos (TSAI, 1986; TRUCCO; VERRI, 1998). Bouguet (2004), por exemplo, faz uso de 20 imagens com 156 pontos de controle para obter resultados aceitáveis na calibração de uma única câmera.

Diante dessas questões, é clara a necessidade de métodos de detecção automáticos, visto que a participação humana nesse processo faz da calibração uma etapa muito trabalhosa e susceptível a erros. Por outro lado, a implementação desses sistemas se depara com os problemas de reconhecimento de padrões, que precisam lidar com as diversas fontes de erro que interferem no processo de formação de imagens, com problemas de iluminação, distorções, ambiguidades e sinais ruidosos. Além disso, a necessidade de várias perspectivas do padrão leva a considerar desde imagens mais simples até cenas mais complexas. Nesse último caso, as características de calibração podem ser confundidas com características similares presentes no restante da cena, dificultando a detecção correta dos pontos que pertencem ao padrão. As combinações desses problemas dificultam a construção de métodos de detecção de padrões robustos.

Alguns trabalhos nem sempre consideram o uso de padrões no processo de calibração. Os métodos de auto-calibração por exemplo, inferem o valor dos parâmetros de câmera a partir de uma sequência de imagens em função da rigidez da cena observada (HEMAYED, 2003). Para esse tipo de procedimento, os pesquisadores usam uma combinação de restrições às características intrínsecas da câmera juntamente com restrições estabelecidas pelo movimento da mesma ou por características estáticas no ambiente. Nessa classe de métodos de calibração, muitas variáveis precisam ser estimadas, o que faz com que os métodos computacionais envolvidos sejam mais complexos. Além disso, os resultados normalmente são obtidos sob um fator de escala (ZHANG, 2004a), sendo difícil a calibração em relação a uma unidade métrica. Lichti e Kim (2011) fazem uma comparação de três das principais técnicas que usam essa abordagem de calibração.

Para os métodos tradicionais, qualquer conjunto de pontos conhecidos podem ser utilizados, desde que estejam devidamente correlacionados, assumam a faixa de precisão requerida e que não configurem situações degenerativas para os métodos numéricos empregados (HARTLEY; ZISSERMAN, 2004). Por isso a construção de padrões de calibração pode ser optativa.

Alguns trabalhos não chegam a demarcar pontos na cena, aproveitando-se do conhecimento de características naturais e utilizando técnicas pouco convencionais. Baltes (2001) por exemplo, usa texturas retangulares genéricas obtidas de cenas naturais para formar seu conjunto de dados. Nesse trabalho a estratégia do autor é usar cenas de arranha-céus cuja estrutura regular das janelas e fachadas formam uma textura retangular que é disposta como um *gride*. Com o conhecimento das dimensões dos prédios, os pontos de controle podem ser obtidos. O método de extração de características precisa identificar as janelas de maneira adequada para permitir a calibração do sistema.

Klaus et al. (2004) usam uma única imagem do céu em uma noite estrelada. Para atingir bons resultados, os autores usam 2000 estrelas como pontos de controle, considerando o fato de que suas posições angulares são conhecidas com boa acurácia. Com essa abordagem, a calibração é realizada com grande precisão mas erros podem ser induzidos

pelas aberrações das lentes e também pela refração atmosférica. Além disso, nem todas as câmeras conseguem registrar imagens nítidas do céu, dificultando o uso dessa técnica em câmeras convencionais.

Os dois trabalhos citados no parágrafo anterior, Baltes (2001) e Klaus et al. (2004), são exemplos de estratégias de calibração que são difíceis de se reproduzir e pouco convenientes para aplicações que funcionam dentro de ambientes fechados e que necessitam da recalibração constante de suas câmeras. Os métodos tradicionais de calibração optam pelo uso de padrões para a extração dos pontos de controle.

4.3 Revisão dos trabalhos da área

Uma das primeiras estratégias de identificação automática dos pontos de calibração pode ser encontrada no trabalho de Tsai (1986). Tsai propõe cinco passos para a detecção automática do padrão que é usado em seu trabalho (Figura 8(a)):

1. Captura da imagem em escala de cinza;
2. Binarização da imagem;
3. Ligar os cantos da imagem binária para extrair um conjunto aproximado de arestas de fronteira;
4. Varrer as bordas perpendicularmente para localizar bordas válidas usando interpolação;
5. Ajustar retas às bordas válidas e encontrar suas intersecções;

O procedimento descrito é vago e considera somente imagens com baixa distorção radial. No entanto, essa sequência de passos é adotada por alguns trabalhos de calibração (ZHANG, 2004b). Outros estudos também são inspirados por esse algoritmo, aplicando diretamente a Transformada Hough (DUDA; HART, 1972) na fase de ajuste de retas (WANG; WU, 2010; PULLIVELLI, 2005; HABIB; MORGAN, 2005).

Os trabalhos mais recentes usam técnicas mais elaboradas, com o objetivo de tornar o processo de calibração menos trabalhoso e mais automatizado, sendo propostos métodos de identificação completamente automáticos ou semiautomáticos. Algumas ferramentas são disponibilizadas gratuitamente, sendo as mais conhecidas o *toolbox* de calibração de câmeras para MatLab (BOUGUET, 2004) e as rotinas da biblioteca OpenCV (BRADSKI, 2000). Com base principalmente nessas duas bibliotecas, vários *softwares* se propõem a realizar a detecção automática dos padrões de calibração. A Tabela 1 apresenta uma lista de alguns dos *softwares* que são facilmente encontrados na literatura. Na primeira coluna está presente o nome do *software*, na segunda o nível de automação presente, na terceira a biblioteca utilizada, e na última a referência da aplicação. De uma maneira geral, pode-se

encontrar várias interfaces gráficas que se restringem à chamada de funções de bibliotecas já funcionais.

Tabela 1: Exemplos de *softwares* de detecção de padrões de calibração.

Nome da aplicação	Tipo de detecção	Algoritmo	Referência
<i>Camera Calibration Tools</i>	Semi-automática	OpenCV	(STOYANOV, 2011)
<i>JcamCalib</i>	Automática	OpenCV	(ORTEGA, 2005)
<i>GML C++</i>	Automática	OpenCV	(VELIZHEV, 2011)
<i>OCamCalib</i>	Automática	OpenCV*	(SCARAMUZZA, 2013)
<i>EasyCamCalib</i>	Automática	(BOUGUET, 2004)*	(BARRETO, 2009)
<i>ROS: Camera Calibration</i>	Automática	OpenCV	(BOWMAN, 2011)

* Algoritmos modificados.

O sistema disponibilizado pelo *toolbox* de Bouguet (2004) é usado para a calibração de câmeras com padrões xadrez. A aplicação é implementada usando o *software* MatLab e permite a detecção semiautomática dos pontos de calibração. O sistema implementado exige que o usuário forneça a posição de quatro pontos delimitadores do padrão em cada imagem do conjunto de calibração. O algoritmo calcula a transformação homográfica desses pontos em relação a um padrão ideal, sem distorções, estima a posição de cada ponto de controle e realiza uma busca local na imagem. O mesmo *toolbox* fornece um conjunto de 20 imagens em sua documentação, que é usado para a demonstração do sistema. Para esse conjunto, o usuário deve fornecer 80 pontos de imagem para que uma única câmera seja calibrada. Esse sistema de detecção é limitado a imagens com baixa distorção radial.

Chen, Benzeroual e Allison (2012) utilizam a mesma metodologia presente no *toolbox* para MatLab, mas automatizam a etapa de participação do usuário. Os autores usam quatro marcadores especiais que são posicionados nas extremidades do padrão xadrez para definir a região de busca dos pontos de calibração. Quando os marcadores extras são identificados, o processo de detecção se dá do mesmo modo que em Bouguet (2004). Essa abordagem de automatização só é possível quando os marcadores extras são reconhecidos com sucesso.

A biblioteca OpenCV implementa a detecção automática de padrões xadrez por meio da função *findChessBoardCorners*, que recebe como parâmetros a imagem do padrão de calibração e seu tamanho. O algoritmo apresenta uma estratégia de detecção baseada em sucessivas avaliações da imagem. Caso a imagem de entrada seja colorida, o algoritmo a transforma para a escala de cinza e segue os seguintes passos:

1. Binarização Adaptativa: nessa etapa espera-se que os quadriláteros do padrão xadrez sejam realçados pela binarização da imagem;
2. Erosão: a fase de erosão atua de modo a desconectar os marcadores do padrão xadrez, produzindo uma imagem formada por regiões isoladas;

3. Extração de contornos: nessa fase do algoritmo as regiões passam a ser representadas por seus contornos. Uma técnica de busca binária de contornos é empregada;
4. Ajuste de polígonos: o algoritmo ajusta polígonos a cada contorno extraído pela etapa passada. Se um contorno é representado por um polígono de quatro lados, então ele é classificado como um possível quadrilátero do padrão de calibração. Sucessivas etapas de erosão são aplicadas na tentativa de encontrar mais contornos válidos;
5. Conexão dos pontos: nessa fase o algoritmo tenta conectar os quadriláteros que foram ajustados às regiões isoladas em cada etapa de erosão. Os pontos são fundidos com base na heurística de que os cantos mais próximos representam um único ponto de calibração;
6. Por último, o algoritmo retorna o maior grupo de pontos conectados que foi encontrado. Se o tamanho desse grupo for igual ao tamanho do padrão passado como parâmetro, então a função considera que houve sucesso na detecção e realiza a ordenação dos pontos finais;

A metodologia implementada no OpenCV é dependente da identificação de todos os quadriláteros do padrão, o que nem sempre é possível devido aos problemas de iluminação e distorções que as imagens estão sujeitas, causando perda de quadros de calibração. A heurística que conecta quadriláteros vizinhos é baseada na proximidade dos cantos, o que pode ser facilmente violada se existirem regiões ruidosas próximas aos contornos avaliados. Além disso, se o padrão não é identificado nas primeiras iterações, o processamento fica lento por conta das sucessivas etapas de erosão da imagem.

Scaramuzza, Martinelli e Siegwart (2006) e Rufli, Scaramuzza e Siegwart (2008) disponibilizam outro *toolbox* de calibração de câmeras para Matlab que usa o OpenCV para a fase inicial de extração dos pontos de controle. Os autores alteram o algoritmo do OpenCV, modificando partes do pré-processamento e propondo uma nova heurística de ligação de pontos vizinhos. Os testes realizados comparam o algoritmo modificado com o original do OpenCV. Os resultados mostram que as modificações propostas melhoram a detecção de padrões xadrez em imagens distorcidas, geradas por câmeras omnidirecionais, com lentes olho de peixe, espelhos esféricos e hiperbólicos. Os testes realizados são restritos à avaliação quantitativa de pontos detectados por ambos os métodos e não são apresentados resultados com imagens convencionais. Os autores não disponibilizam as imagens utilizadas para as comparações com outros trabalhos. Como produto final, o método é disponibilizado como uma ferramenta de calibração de câmeras omnidirecionais, denominada OCamCalib¹. Barreto et al. (2009) e Hughes, Glavin e Jones (2011) também implementam metodologias de calibração específicas para imagem distorcidas.

---

¹Sítio do sistema OCamCalib disponível em: <https://sites.google.com/site/scarabotix/ocamcalib-toolbox>.

Escalera e Armingol (2010) identificam os pontos de controle como sendo as intersecções das retas definidas pelos retângulos presentes no padrão de calibração xadrez. Os autores desenvolveram uma sistema que emprega a Transformada de Hough, (DUDA; HART, 1972), sobre a resposta de um detector de cantos para identificar os cantos que são colineares na imagem. Como nem todas as características encontradas pertencem ao padrão de calibração, uma etapa de filtragem remove cantos ruidosos com uma segunda aplicação da Transformada Hough, agora sobre as intersecções das primeiras retas encontradas. Os pontos finais são definidos pelas intersecções das retas da segunda etapa. Os resultados apresentados se restringem à avaliação do método de calibração que também é proposto pelo trabalho. O método de detecção funciona bem para imagens com baixa distorção radial mas é prejudicado pela presença de retas extras ao padrão de calibração.

Fiala e Shu (2007) utilizam marcadores de realidade aumentada (*ARTag*) para representar os pontos de controle. A ideia principal desse trabalho é distinguir as características de interesse por meio de padrões regulares que dificilmente se repetem no restante da imagem. Os marcadores *ARTag* carregam códigos autoidentificadores que definem sua posição absoluta dentro do padrão de calibração. O sistema descrito é considerado robusto a alterações de iluminação por utilizar binarização adaptativa da imagem e não exige a identificação completa do padrão. Por outro lado, os marcadores usados exigem um detector mais complexo e devem ser únicos para cada ponto de calibração. Caso seja necessário aumentar o número de pontos de controle, novos marcadores *ARTag* devem ser projetados. Do mesmo modo que Escalera e Armingol (2010), os resultados presentes não avaliam o detector de padrões proposto pelo trabalho.

O sistema CAMcal de Shu, Brunton e Fiala (2009) usa o detector de cantos de Harris e Stephens (1988) para a extração dos pontos de interesse sobre a imagem de padrões xadrez que possuem marcadores circulares em seu interior. Os pontos encontrados são arranjados em uma malha de triângulos de modo a conectar cada ponto com seu vizinho mais próximo. Pares de triângulos são convertidos em quadriláteros comparando a cor média das regiões que cobrem na imagem. Se dois triângulos possuem cores semelhantes, ambos geram um quadrilátero em uma malha geométrica auxiliar. A posição de cada ponto é definida por meio da detecção de circunferências marcadoras de orientação. Os autores não explicam em detalhes como a conversão de triângulos é realizada e nem como as cores dos triângulos são comparadas. Os testes avaliam a resposta do algoritmo considerando câmeras diferentes sobre um banco de imagens criado pelos próprios autores do trabalho. Não são apresentadas comparações com outros métodos de detecção.

Zheng et al. (2008) também apresentam uma proposta baseada em triangulação dos pontos iniciais. Esse trabalho se diferencia por considerar atributos topológicos diretamente relacionados aos pontos detectados que puderam ser determinados por meio da convolução da imagem com padrões ou *templates* que maximizam a resposta dos pontos internos ao padrão xadrez. Do mesmo modo que os demais trabalhos, Zheng et al. se

limita à apresentação de resultados visuais.

Bevilacqua, Gherardi e Carozza (2008) detectam o padrão xadrez recalculando continuamente a transformação perspectiva da imagem. Um conjunto de pontos iniciais permite o cálculo dessa transformação, que é atualizada à medida que novos pontos são encontrados. A transformação perspectiva inversa permite reduzir as distorções geradas pela inclinação do padrão, facilitando a detecção dos pontos restantes. A ordenação do conjunto de pontos é realizada com base em uma estrutura de grafos, onde um algoritmo avalia relacionamentos angulares relativos entre os pontos detectados. Kang, Ha e Jeong (2008) empregam uma estratégia semelhante, mas aplicada à detecção de padrões com marcadores circulares.

O trabalho de Dao e Sugimoto (2010) propõe um reconhecimento de padrões xadrez não planos que é robusto a ruídos e à descontinuidade do padrão. A metodologia é construída em torno de um detector de cantos específico para as intersecções do padrão em questão, que avalia o número de alternâncias na vizinhança do *pixel* central. A filtragem e determinação das posições relativas de cada canto é alcançada com a verificação topológica das arestas que conectam todos os cantos encontrados. O trabalho apresenta somente resultados visuais do sistema implementado.

De uma maneira geral, os estudos referentes à detecção automática de padrões de calibração seguem uma ordem de processamento dividida em duas etapas principais: extração e filtragem de características. Muitos métodos recorrem a estratégias que são sensíveis à distorções da imagem ou abordam algum mecanismo topológico de filtragem dos pontos.

Outra ponto em comum aos métodos mencionados é o uso de detectores de cantos baseados no detector de Harris e Stephens (1988). No entanto sabe-se que esse detector não é adequado para o realce das características presentes em um padrão xadrez. Chen e Zhang (2005) e Krüger e Wöhler (2011) mostram que esse detector apresenta pouca precisão para o tipo de padrão utilizado, sendo inviável para o processo de calibração. Adicionalmente, esse tipo de detector consome muito tempo de processamento e faz uso de limiares globais para diferenciar os cantos de não-cantos. Por esse motivo é comum selecionar os n pontos mais representativos dentro da imagem. Na prática, dependendo da complexidade da cena, nada garante que os pontos de calibração tenham uma melhor resposta ao detector, o que força o parâmetro n assumir um número maior.

4.4 Considerações finais

A popularidade da biblioteca OpenCV permitiu o surgimento de muitos aplicativos dedicados à calibração automática de câmeras. Apesar de suas limitações, o algoritmo implementado pela função *findChessBoardCorners* ainda é um dos mais usados nos processos de calibração.

As limitações dos métodos automáticos de detecção de padrões xadrez tem influenciado

o surgimento de estudos que optam por métodos topológicos, visando uma metodologia mais abrangente aos diversos tipos de lentes e também menos sensível a ruídos. Dentre esses métodos, os trabalhos de Shu, Brunton e Fiala (2009) e Zheng et al. (2008) contribuíram para o desenvolvimento da pesquisa presente nesta tese. As abordagens topológicas adotadas são similares à implementada neste trabalho, mais precisamente a ideia de analisar as cores dos triângulos para a filtragem topológica e a conectividade de seus vizinhos. No entanto, a metodologia abordada no presente trabalho vai além dos resultados conseguidos por essas referências.

O detector estrutural desenvolvido em Dao e Sugimoto (2010) também contribuiu para o desenvolvimento desta pesquisa. A construção de detectores específicos permite uma filtragem logo no processo de extração de características, diminuindo a influência de falsos positivos no sistema de detecção.

Pelos trabalhos listados, percebe-se também que ainda há espaço para muitas pesquisas na área de detecção de padrões, principalmente voltadas para a criação de detectores de características mais robustos e análises topológicas mais eficientes.

5 Detecção Automática de Padrões de Calibração

5.1 Considerações iniciais

Este Capítulo apresenta o sistema de detecção de padrões xadrez para a calibração de câmeras desenvolvido nesta tese. A metodologia aqui desenvolvida faz uso de uma abordagem topológica para a identificação do padrão de calibração e por isso é denominada Detecção Topológica de Padrões Xadrez ou em inglês *Chessboard Topological Detection* (ChTD). O texto segue a ordem de processamento que é utilizada pelo sistema ChTD, apresentando resultados intermediários e justificando as estratégias de desenvolvimento. A metodologia aqui construída faz referências a trabalhos de outros autores como soluções para algumas etapas intermediárias. As referências são escolhidas com base na eficiência dos algoritmos que elas descrevem e por se encaixarem nas especificações do sistema de detecção desenvolvido nesta tese.

5.2 Metodologia utilizada pelo algoritmo ChTD

Esta parte do trabalho atenta-se à construção de um sistema de detecção de pontos para a calibração de câmeras dispostos sob um padrão xadrez. Dentre as fases do processo de calibração a determinação dos pontos de controle é uma etapa fundamental e muito trabalhosa, muitas vezes exigindo intervenção humana. A elaboração de sistemas robustos de detecção desses pontos visa suprir uma necessidade comum às aplicações de Visão Computacional onde a calibração métrica de câmeras é uma etapa indispensável.

Uma metodologia ideal deveria ser automática, rápida, livre de qualquer participação manual, robusta a ruídos, menos dependente de limiares e com boa precisão na localização dos pontos de controle. Além disso, a oclusão parcial do padrão ou a interferência de pontos ruidosos não deveria causar uma perda completa da imagem, uma vez que a calibração geralmente pode ser realizada mesmo com uma quantidade menor de pontos de controle que são fornecidos pelo padrão.

O presente trabalho busca como solução para a detecção automática de padrões xadrez a construção de um detector estrutural dos pontos de controle e uma verificação topológica da vizinhança desses pontos com base na regularidade do padrão utilizado. O detector estrutural tem o objetivo de reduzir a detecção de características que não pertençam ao padrão de calibração enquanto que a análise topológica torna o resultado final robusto à

presença de ruído.

A metodologia aqui desenvolvida segue a seguinte sequência de processamento: Extração dos candidatos a pontos de calibração; Triangulação; Seleção dos pontos por meio de um filtro topológico; Propagação e ajuste de coordenadas; Refinamento da posição das características. A Figura 16 apresenta a sequência de passos e a visualização dos resultados esperados em cada etapa.

Figura 16: Metodologia de detecção de padrões xadrez.

São considerados como dados de entrada imagens em escala de cinza. Os pontos de interesse são entendidos como as posições na imagem cuja vizinhança se assemelha a imagens da intersecção de quatro regiões retangulares cujas bordas assumem o formato de “X”, denominados *cantos-x*. A Seção 5.3 apresenta a definição de um canto-x e também o método de detecção estrutural e específico para essa característica.

O conjunto de cantos-x define uma malha de triângulos que relaciona cada ponto com seus vizinhos mais próximos. Com base na estrutura quadriculada do padrão xadrez, um filtro que verifica atributos topológicos permite identificar quais características são válidas, evitando que pontos não pertencentes ao padrão sejam registrados no resultado final. Assume-se que cada quadrilátero do padrão seja formado por dois triângulos de mesma cor e suas propriedades de vizinhança são avaliadas. Os triângulos que não se encaixam nas regras definidas pelo filtro topológico da Seção 5.5 são classificados como inválidos. São descartados os pontos cujos triângulos não obedecem a regularidade do padrão.

A propagação e ajuste de coordenadas é a fase em que é realizado o casamento entre as projeções detectadas e os pontos no mundo real que são apresentados pelo padrão de calibração. A posição correta de cada ponto do padrão é determinada com base no conjunto de triângulos válidos. Inicialmente toma-se como origem do plano Cartesiano o ponto que mais se aproxima do centro de massa dos pontos detectados. A malha geométrica estabelece a posição relativa de cada ponto de calibração e, se o tamanho do

padrão é conhecido e um número mínimo de pontos nos sentidos x e y são identificados, então a posição absoluta de cada ponto é determinada. A Seção 5.6 descreve esse processo.

Para evitar o processamento desnecessário, cada vértice da malha filtrada define uma região onde o refinamento da localização dos pontos de calibração é executado. É atribuído um valor de relevância a cada ponto detectado e sua localização é calculada com maior resolução (precisão *sub-pixel*). Um método específico para cantos-x é empregado e a posição do ponto na imagem é atualizada com base no centro de massa da região vizinha a esse ponto. Essa etapa normalmente envolve limiarização da resposta dos detectores, no entanto ela é beneficiada pela definição de uma região de procura. A Seção 5.7 descreve como a posição de cada canto-x é aprimorada.

Como resultado final espera-se um par de matrizes que contém as coordenadas dos pontos classificados como cantos-x pertencentes ao padrão de calibração e suas respectivas posições no mundo real. Esses resultados formam a base de dados usadas nos métodos de calibração de câmeras.

5.3 Detector de cantos-x

O primeiro passo do algoritmo proposto por este trabalho é a detecção das características candidatas a pontos de calibração. De acordo com o padrão utilizado, os cantos de interesse possuem uma estrutura parecida com uma forma de “X” que intercala duas cores de alto contraste. Esse alternância de cores define quatro conjuntos de *pixels* que são categorizados em duas classes: claros e escuros. O termo *contraste* refere-se à diferença entre as intensidades dos *pixels* analisados.

A Figura 17 apresenta a estrutura de um canto-x em três imagens sintéticas seguidas de três exemplos de imagens reais. As imagens sintéticas permitem uma análise da característica de interesse sem as perturbações causadas pelo processo de formação das imagens reais.

Figura 17: Aparência de um canto-x. A imagem mostra três imagens sintéticas do canto-x e três exemplos de como o perfil de intensidades é amostrado por uma câmera real.

Da esquerda para a direita, a primeira imagem mostra um canto-x ideal onde a superfície de intensidades forma bordas bem definidas e perpendiculares, separando as classes de *pixels* claros e escuros de maneira simétrica com a mesma proporção de pontos em cada conjunto. Já na segunda imagem, a aparência do mesmo padrão foi modificada por

uma transformação *Affine* de modo que as bordas não preservam mais a característica perpendicular. Nesse exemplo, a quantidade de pontos com cores diferentes também é alterada, mas a proporção entre as áreas dos conjuntos da mesma classe permanece constante. A terceira imagem é produto de uma transformação perspectiva. Ela mostra que a razão entre a área dos conjuntos de uma mesma classe também pode ser afetada.

Em situações reais a imagem capturada também está sujeita à perturbações pela distorção projetiva e por condições de iluminação, amostragem espacial, quantificação da luminância de cada ponto, ruídos no equipamento de captura e distorções provocadas pelas lentes. É comum uma imagem ligeiramente borrada e com variações de cores dentro da mesma classe de cor. As três últimas imagens da Figura 17 mostram alguns exemplos de cantos-x reais com diferentes transformações perspectivas. Elas apresentam falta de definição nas bordas e flutuações dos valores de intensidades, normalmente distribuídas. Na última imagem a distorção é mais acentuada e chega a descaracterizar o ponto central. Os pontos mais próximos ao centro da figura tendem a ter cores mais semelhantes à média de toda a imagem. Esses fenômenos tornam complicada a elaboração de detectores robustos que conseguem lidar com todas as perturbações da imagem dessa característica.

Observa-se pelas imagens reais da Figura 17 que duas propriedades permanecem inalteradas: o número de alternâncias entre os quatro conjuntos se mantém o mesmo e o ponto central tem grande tendência à média de toda a região. Por esses motivos, neste trabalho, um canto-x é modelado como o *pixel* cuja vizinhança possui quatro alternâncias de alto contraste e sua intensidade é próxima à média de seus vizinhos. Sendo $\Psi = \{\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_n\}$ o conjunto de *pixels* angularmente ordenados em relação a $\mathbf{p}_c$ que define todos os *pixels* pertencentes à borda de uma circunferência de Bresenham (BRESENHAM, 1977) tendo $\mathbf{p}_c$ como origem, o número de alternâncias de alto contraste pode ser definido pela Equação (15):

$$N_{alt} = \sum_{i=1}^n \begin{cases} 1, & I(\mathbf{p}_i) > T_h \wedge I(\mathbf{p}_{i-1}) < T_l \\ 1, & I(\mathbf{p}_i) < T_l \wedge I(\mathbf{p}_{i-1}) > T_h \\ 0, & \text{caso contrário} \end{cases} \quad (15)$$

onde $\mathbf{p}_i \in \Psi$, $I(\mathbf{p}_i)$ expressa a intensidade do *pixel* $\mathbf{p}_i$. T_l e T_h são os limiares inferior e superior respectivamente. A Figura 18 ilustra como o conjunto vizinhança de um canto-x e suas as alternâncias de contraste são vistas neste trabalho.

Essa modelagem é semelhante à que Klein e Murray (2007) utilizam para identificar os pontos de controle de calibração em sua aplicação denominada *Parallel Tracking and Mapping* (PTAM). Os autores modelam as variáveis de forma que o tamanho da vizinhança é fixado em $n = 16$, $T_l = m - gate$ e $T_h = m + gate$, onde $m = \frac{1}{n} \sum_{i=1}^n I(\mathbf{p}_i)$ e a variável *gate* modela a sensibilidade da operação de modo que um valor pequeno permite identificar cantos-x em regiões com menor contraste. O *pixel* $\mathbf{p}_c$ é classificado como um canto-x se a expressão lógica na Equação (16) for aceita. O detector descrito por

Figura 18: Vizinhança de um canto-x. A circunferência é interpretada como uma sequência linear de pontos.

essa referência tem a vantagem de possuir custo computacional reduzido, uma vez que somente uma parcela da vizinhança do *pixel* de interesse é considerada. O algoritmo do detector PTAM está presente no código disponibilizado pelos autores¹. Outros detectores semelhantes podem ser encontrados em Rosten e Drummond (2006), Sun et al. (2008) e Zhao et al. (2011).

$$(N_{alt} = 4) \wedge (T_l \leq I(\mathbf{p}_c) \leq T_h) \quad (16)$$

Considerando uma imagem de entrada, o número de alternâncias impõe grande parte da restrição necessária para a classificação de um canto-x, uma vez que $I(\mathbf{p}_c)$ tende a estar no intervalo médio proposto. Por outro lado, a variável *gate* restringe a diferença de intensidades entre as classes de *pixels* claros e escuros em um intervalo fixo igual a $2 \times gate$, que é sensivelmente afetada em situações onde a iluminação não é controlada. A vizinhança é amostrada considerando somente os *pixels* que pertencem à borda da circunferência, descartando completamente as intensidades mais próximas ao centro da janela de observação. Além disso o modelo descrito supõe que o valor do ponto central seja rigorosamente próximo à média das intensidades dos pontos de Ψ . Na verdade, a intensidade de $\mathbf{p}_c$ depende fortemente dos vizinhos mais próximos e tende a perder correlação com a média do conjunto Ψ à medida que o raio da circunferência aumenta.

Antes de detalhar o detector de cantos-x usado nesta tese, algumas considerações sobre as perturbação na imagem devem ser observadas. Considerando as imagens reais da Figura 17, as variações de cores em regiões mais homogêneas podem ser justificadas pelos ruídos eletrônicos do equipamento de captura de imagem e também pelos truncamentos do processo de quantificação da intensidade de luz que o sensor de imagem recebe. Esse efeito está relacionado à resolução do conversor analógico/digital utilizado. Adicionalmente, a falta de precisão cromática da ferramenta que confecciona o padrão de calibração tam-

¹Sítio do código PTAM disponível em: <http://www.robots.ox.ac.uk/~gk/PTAM/>

bém tem participação significativa nesse problema. Outras perturbações como sombras, iluminação direcional e inclinação afetam diretamente como a luz do padrão de calibração é refletida para a câmera.

O processo de formação de imagens também pode explicar boa parte dos efeitos encontrados nas imagens reais. O borramento da figura é geralmente associado à falta de foco e também pela amostragem espacial feita pelo sensor. Quando o foco encontra-se mal ajustado, a refração das lentes faz com que a imagem atinja o sensor já com perda de definição. Cada célula desse sensor também responde de acordo com a quantidade média de radiação que incide sobre sua superfície. Desse modo, quando a projeção de uma borda encontra-se exatamente sobre uma célula, a intensidade mensurada é na verdade um valor médio das duas classes de cores que ela separa. A falta de definição sobre a imagem das bordas é geralmente mais expressiva já que a radiação calculada carrega consigo informações sobre duas regiões diferentes, ou seja, regiões de alto contraste. Maiores detalhes sobre o processo de formação de imagens podem ser encontrados em Forsyth e Ponce (2002), Gonzalez e Woods (2006) e Szeliski (2010).

Diante das observações do parágrafo anterior, entende-se que uma imagem digital é construída sob uma amostragem da radiação refletida pelo padrão real. Nesse caso, cada amostra carrega consigo informações de vizinhança que devem participar da classificação do ponto como canto-x. É evidente que outros fatores também perturbam o comportamento do sinal, como por exemplo a iluminação direcional, a inclinação perspectiva e a rugosidade do objeto que adicionam uma tendência na reflexão da luz e maneira com que ela é difusa pela superfície (BLINN, 1977; SCHLICK, 1994). No entanto, a técnica implementada nesta tese considera uma região muito pequena, estritamente local à característica onde esse comportamento é minimizado. Desse modo, considera-se que o padrão de calibração é próximo a uma superfície Lambertiana onde os efeitos causados por esses fenômenos não são significativos (OREN; NAYAR, 1994).

O detector de cantos-x desenvolvido neste trabalho considera que o número de alterações dado pela Equação (15) é suficientemente restritivo à característica de interesse, mas propõe um cálculo adaptativo dos limiares T_l e T_h de modo a considerar uma melhor representação dos *pixels* vizinhos.

Seja uma janela W centrada em $\mathbf{p}(x, y)$ e com largura w , a média μ_W e o desvio padrão σ_W de W é dado pelas Equações 17 e 18 respectivamente.

$$\mu_W(x, y) = \frac{1}{w^2} \sum_{i=-\frac{w}{2}}^{\frac{w}{2}} \sum_{j=-\frac{w}{2}}^{\frac{w}{2}} I(\mathbf{p}(x+i, y+j)) \quad (17)$$

$$\sigma_W(x, y) = \frac{1}{w^2 - 1} \sum_{i=-\frac{w}{2}}^{\frac{w}{2}} \sum_{j=-\frac{w}{2}}^{\frac{w}{2}} [I(\mathbf{p}(x+i, y+j)) - \mu_W(x, y)]^2 \quad (18)$$

A notação $\mathbf{p}(x, y)$ representa o *pixel* como o vetor $\mathbf{p} = [x \ y]^\top$. Os limiares T_l e T_h são calculados de acordo com a Equação (19).

$$\begin{aligned} T_l &= \mu_W - k\sigma_W \\ T_h &= \mu_W + k\sigma_W \end{aligned} \quad (19)$$

Seja r o raio da circunferência de Bresenham, a largura w é dimensionada como $w = 2 \times (r + 2)$ para considerar um *pixel* além da borda definida por Ψ . k é uma constante que assume valores entre $[0, 1]$ e ajusta a sensibilidade do detector.

Considerando cada *pixel* como uma média de seus vizinhos mais próximos, o conjunto Ψ realiza uma boa amostragem do número de alternâncias em volta do ponto central. μ_W agrega informação de todos os pontos que pertencem a W enquanto que σ_W descreve a variabilidade média das alternâncias. Assim, se o interior da janela de observação é mais homogêneo, a variância é menor e restringe o intervalo em que a intensidade do ponto central deve estar. Quando existe grande variação de cor o intervalo médio é maior e a quantidade de alternâncias é mais restritiva. Desse modo, o detector que usa os limiares da Equação (19) considera valores mais adequados à janela de análise. Se $I(\mathbf{p}(x, y)) \approx \mu_W$ e $N_{alt} = 4$, o ponto central pode ser classificado como um canto-x.

As Figuras 19(a) e 19(b) apresentam o resultado típico do detector PTAM e do desenvolvido neste trabalho respectivamente. Os *pixels* em vermelho são classificados como cantos-x. Em ambos os casos a detecção classifica pontos em torno do *pixel* de interesse definindo uma região de cantos-x válidos. Na primeira imagem, o algoritmo de PTAM apresenta um comportamento não uniforme em torno do *pixel* central e gera regiões desconectadas. Com a abordagem estatística sugerida, essa região é mais simétrica e centrada na posição do ponto de interesse. A Figura 20 mostra cantos-x em uma imagem simples e outra com o fundo mais complexo.

(a) Detector de PTAM.

(b) Detector descrito por este trabalho.

Figura 19: Comparação da resposta do detector de cantos-x. A primeira imagem mostra o resultado típico do detector usado em Klein e Murray (2007) com $gate = 20$ como sugerido pelos próprios autores. Na segunda imagem, o detector adaptativo desenvolvido neste trabalho foi executado com $k = 0,7$. A resposta apresenta um comportamento mais uniforme em torno do ponto de interesse.

Figura 20: Exemplo do resultado gerado do detector de cantos-x adaptativo.

O resultado final do detector desenvolvido nesta seção é um conjunto de respostas lógicas que representam as regiões mais prováveis em que os verdadeiros cantos-x estão localizados. Por conta da resposta mais simétrica do detector apresentado, uma boa estimativa da localização real do *pixel* de interesse pode ser obtida calculando-se o centro médio da região detectada. Para viabilizar esse procedimento é preciso extrair os componentes conectados que definem cada região. Esse problema é conhecido em Visão Computacional como rotulação de *pixels* conectados. O algoritmo usado neste trabalho pode ser encontrado em He et al. (2009). Os autores apresentam uma revisão dos principais algoritmos para a extração de regiões conectadas e propõem uma solução eficiente reduzindo a quantidade de memória e visitando cada ponto da imagem somente duas vezes. Essas características fazem da técnica de He et al. (2009) uma escolha apropriada para aplicações que exigem pouco tempo de processamento.

5.4 Otimização do algoritmo de detecção

Para manter a eficiência do algoritmo proposto, os valores μ_W e σ_W são calculados usando o conceito de *imagem integral*. Imagem integral é uma representação intermediária de uma imagem construída por meio de uma função acumulativa. Essa representação se tornou mais conhecida em Visão Computacional a partir do trabalho de Viola e Jones (2001) mas é um conceito antigo da teoria de probabilidades, frequentemente usada no cálculo de funções de probabilidades acumulativas multidimensionais (RENCER, 2002). Uma referência mais antiga também pode ser encontrada em Crow (1984) dentro do contexto de computação gráfica.

A imagem integral Int na posição (x, y) contém a soma dos valores de todos *pixels* acima e à esquerda de (x, y) . Formalmente, ela pode ser definida como:

$$Int(x, y) = \sum_{i \leq x} \sum_{j \leq y} I(i, j) \quad (20)$$

Essa formulação pode ser calculada em apenas um passo de acordo com uma relação de recorrência:

$$\begin{aligned} S(x, y) &= S(x, y - 1) + I(x, y) \\ Int(x, y) &= Int(x - 1, y) + S(x, y) \end{aligned} \quad (21)$$

onde $S(x, y)$ denota a soma acumulativa da linha y até a posição x . Essa estrutura permite recuperar a soma de valores de regiões retangulares de maneira rápida e eficiente a partir de apenas três operações de soma entre quatro posições da imagem. A Figura 21 mostra o relacionamento das áreas dentro de uma imagem integral. As letras A , B , C e D representam regiões na imagem original. Em Int , o ponto $\mathbf{p}_a$ tem valor igual à soma de todos os *pixels* de A , $\mathbf{p}_b = A + B$, $\mathbf{p}_c = A + C$ e $\mathbf{p}_d$ representa a área $A + B + C + D$.

Figura 21: Relação de áreas em uma imagem integral.

Formalmente, seja uma janela W definida pelos pontos $\mathbf{p}_a$, $\mathbf{p}_b$, $\mathbf{p}_c$ e $\mathbf{p}_d$, a soma de todos seus elementos é dada pela Equação (22).

$$A_W(x, y) = Int(\mathbf{p}_d) - Int(\mathbf{p}_b) - Int(\mathbf{p}_c) + Int(\mathbf{p}_a) \quad (22)$$

De maneira análoga, calcula-se A_W^2 como:

$$A_W^2(x, y) = Int^2(\mathbf{p}_d) - Int^2(\mathbf{p}_b) - Int^2(\mathbf{p}_c) + Int^2(\mathbf{p}_a) \quad (23)$$

onde Int^2 é a imagem integral do quadrado dos *pixels* da imagem I . μ_W e σ_W podem ser calculados como:

$$\mu_W = \frac{A_W}{w^2} \quad (24)$$

$$\sigma_W = \frac{A_W^2 - w^2 \mu_W^2}{w^2 - 1} \quad (25)$$

O custo computacional das Equações 24 e 25 é significativamente menor do que as Equações 17 e 18. Com esse artifício, a área e o desvio padrão de W podem ser obtidos em tempo constante independentemente do tamanho da janela usada.

5.5 Filtragem topológica

Um filtro topológico considera regras topológicas que definem propriedades conectivas e da organização espacial atributos que modelam a característica de interesse que devem ser conservadas. Os cantos-x detectados na imagem nem sempre pertencem ao padrão de calibração, sendo necessária uma identificação de quais pontos são válidos. Neste trabalho, essa verificação é feita considerando a regularidade da vizinhança do padrão utilizado. Se os pontos detectados recaem sobre os cantos-x do padrão de calibração, então seus pontos vizinhos tem propriedades definidas pelos marcadores dessas características.

A determinação dos vizinhos de cada ponto pode ser entendido como um problema de triangulação na criação de malhas geométricas (BERN; EPPSTEIN, 1992). Uma malha geométrica é uma coleção de vértices, arestas e faces definidas por polígonos básicos, tais como triângulos, interligados de acordo com sua vizinhança. Esse assunto compreende um vasto campo de estudos na área de Computação Gráfica, normalmente solucionado via triangulação de Delaunay (BERG; CHEONG; KREVELD, 2008).

A triangulação de Delaunay é um problema clássico de geometria computacional. Para o caso bidimensional, dado um conjunto de pontos em um plano, a triangulação de Delaunay é única, de forma que a circunferência que passa pelos pontos de cada triângulo criado não contém nenhum vértice (O'ROURKE, 1998, Cap. 5). Isso garante que os triângulos sejam formados sempre pelos vértices mais próximos. O algoritmo de triangulação implementado nesse trabalho foi desenvolvido por Guibas, Knuth e Sharir (1992) e baseia-se numa metodologia incremental que executa em tempo $O(n \log(n))$.

A Figura 22 apresenta um exemplo de triangulação de Delaunay sobre uma imagem do padrão xadrez, sendo que A , B , C e D representam quatro cantos-x detectados na imagem original. A única triangulação correta é formada pelos triângulos $\Delta(A, B, C)$ e $\Delta(A, C, D)$. Os triângulos $\Delta(A, B, D)$ e $\Delta(B, C, D)$ são ditos inválidos porque as circunferências que passa por eles englobam os pontos C e A respectivamente. Isso fere a restrição imposta por Delaunay fazendo com que a aresta $\overline{BD}$ também seja inválida.

A implementação da triangulação Delaunay deste trabalho faz uso de uma estrutura redundante de relacionamentos. A estrutura de dados foi projetada para fornecer acesso facilitado aos vizinhos a partir de qualquer uma das três entidades geométricas básicas seguintes:

- **Vértice:** é um ponto na imagem que pode estar conectado a qualquer quantidade de arestas.

Figura 22: Exemplo de triangulação Delaunay. Os triângulos $\Delta(A, B, D)$ e $\Delta(B, C, D)$ são considerados inválidos pois existem vértices no interior da circunferência formada por seus vértices.

- **Aresta:** é o segmento de reta que liga exatamente dois vértices. Cada aresta possui nenhum ou no máximo dois triângulos vizinhos.
- **Triângulo:** é a entidade com três vértices e três arestas com relacionamento consistente. Dois triângulos são vizinhos quando compartilham uma aresta.

O vértice é a única entidade que pode ser criada arbitrariamente. Os triângulos e arestas são construídos pelo algoritmo de triangulação e expressam os relacionamentos entre os vértices. A Figura 23 ilustra o diagrama de classes simplificado usado no desenvolvimento do *software* de triangulação. Os atributos x_i e y_i são as coordenadas de imagem de cada vértice, inicialmente definidos pelo detetor de cantos-x da Seção 5.3. Os atributos x_w e y_w são as coordenadas do padrão de calibração relativas às posições do mundo real. Esses atributos são usados pela fase de propagação e ajuste de coordenadas na Seção 5.6. A classe *MalhaDelaunay* agrega uma coleção de *Vértices*, *Arestas* e *Triângulos*. Essa estrutura de dados favorece a navegação mútua entre as diferentes classes.

Figura 23: Diagrama de classes para a implementação da triangulação Delaunay. O relacionamento redundante permite um rápido acesso aos demais componentes vizinhos.

A descrição completa do algoritmo de triangulação de Delaunay exigiria uma revisão abrangente de fundamentos de geometria computacional, o que foge do escopo deste

trabalho. No entanto, de maneira resumida, o algoritmo pode ser explicado pelos seguintes passos:

- Passo 1:** Primeiramente, os vértices são ordenados em função de suas coordenadas de imagem;
- Passo 2:** Quatro vértices extras, denominados de vértices âncoras são inseridos no conjunto de vértices para definir dois triângulos cujas áreas sobreponham toda a área da imagem;
- Passo 3:** Cada vértice detectado é inserido individualmente. O algoritmo percorre a lista de triângulos e extrai aquele que possui o novo vértice em seu interior;
- Passo 4:** O triângulo detectado é removido da malha e o vértice inserido forma três novos triângulos com os vértices do triângulo excluído que satisfazem a restrição imposta por Delaunay;
- Passo 5:** A verificação da circunferência de cada triângulo é propagada pela malha para todos os vizinhos dos triângulos modificados;
- Passo 6:** O processo retorna ao **Passo 3** e se repete até que todos os vértices sejam inseridos;
- Passo 7:** Os vértices âncoras são removidos junto com os triângulos que formam.

Como resultado tem-se uma malha que relaciona cada canto-x com seu vizinho mais próximo formando triângulos que cobrem as regiões homogêneas. Essas regiões correspondem aos retângulos do padrão xadrez. A ordenação realizada no **Passo 1** permite que a identificação do triângulo do **Passo 2** não precise varrer toda a lista de triângulos, uma vez que o próximo ponto a ser inserido sempre segue a sequência de ordenação de suas coordenadas.

Com a malha construída, os vértices passam por uma filtragem topológica para excluir aqueles que não satisfazem a regularidade do padrão. Os vértices internos ao padrão xadrez compartilham triângulos de cores diferentes de uma forma regular. O filtro topológico assume duas restrições principais: a validade do triângulo e a conectividade entre eles.

São definidos como válidos os triângulos que possuem todo seu interior preenchido por uma única cor. Cada quadrilátero do padrão é representado por dois triângulos de mesma cor. Os quadriláteros que não possuem nenhum triângulo vizinho de cor diferente não satisfazem a condição de conectividade do padrão. O filtro topológico segue uma sequência de verificação de 4 regras:

- Regra 1:** São válidos os triângulos que não possuem transição de cor em seu interior;
- Regra 2:** São inválidos os triângulos que não possuem pelo menos um vizinho de mesma cor;

Regra 3: Considerando o triângulo T e Φ o conjunto de triângulos vizinhos de mesma cor, se a soma entre o número de vizinhos de cor diferente de T e o número de vizinhos de cor diferente de todos os triângulos de Φ for menor que 2, T é fracamente conectado à malha xadrez e assim classificado como inválido.

Regra 4: Os Triângulos que não possuem exatamente 1 triângulo vizinho com mesma cor é considerado inválido.

Na **Regra 1**, o interior dos triângulos é analisado seguindo o algoritmo de preenchimento de polígonos *Scan-line*, outro método comum em Computação Gráfica (FOLEY et al., 1990). A **Regra 2** parte do princípio de que, se o triângulo não forma um quadrilátero com outro triângulo, então ele não pertence ao padrão de calibração. A conectividade é garantida pela **Regra 3** que exclui todos os quadriláteros que não formam vizinhança com quadriláteros de cor diferente, removendo aqueles que não estão fortemente conectados com o restante da malha. A **Regra 4** remove situações ambíguas no interior da malha. As Figuras 24, 25 e 26 ilustram os efeitos que as regras citadas causam sobre a malha.

Figura 24: Efeito das regras topológicas 1 e 2.

Na Figura 24 a **Regra 1** remove os triângulos sobre as regiões fora do padrão e também aqueles que são influenciados pela presença de falsos negativos no interior do padrão de calibração. A **Regra 2** detecta os triângulos que não podem participar dos testes seguintes.

Figura 25: Efeito da terceira regra topológica.

Os efeitos causados pela **Regra 3** na Figura 25 reforçam a conectividade de cada quadrilátero com a malha, removendo aqueles que possuem pelo menos duas bordas com quadriláteros de cor oposta. A Figura 26 mostra as situações em que a **Regra 4** desempenha papel significativo na remoção de ruídos internos. Os falsos positivos no interior do padrão geram ambiguidades topológicas que estão de acordo com a regularidade dos marcadores.

Figura 26: Efeito da quarta regra topológica.

Nos exemplos apresentados pela Figura 26, observa-se que um falso positivo cria triângulos bem conectados à malha e ficam imperceptíveis à condição de conectividade imposta pela **Regra 3**. Por esse motivo, a **Regra 4** desempenha um papel mais severo, restringindo a análise somente ao número de vizinhos de mesma cor de cada triângulo. Se um triângulo possui mais que 1 vizinho de cor semelhante, então um de seus vértices está localizado em uma região de cor homogênea que, consequentemente, deve ser classificado como ruído. A exclusão de triângulos internos ao padrão de calibração não afeta significativamente os vértices que o formam, uma vez que os triângulos removidos compartilham seus vértices com os demais triângulos válidos. Com uma verificação sucessiva da **Regra 4** somente os pontos ruidosos são desconectados.

O algoritmo de filtragem executa a **Regra 1** uma única vez, e as demais são aplicadas sequencialmente em laço de repetição para a malha. Se um triângulo não satisfaz a **Regra 2** então as próximas regras não precisam ser verificadas. O mesmo se aplica às outras regras. A sequência de execução é interrompida quando não há mais nenhum triângulo classificado como inválido. Em seguida, os vértices que não formam nenhum triângulo também são descartados. A Figura 27 apresenta o resultado da filtragem topológica em uma imagem real. A primeira imagem contém a malha de triângulos construída a partir de todos os cantos detectados, e a segunda o resultado após a aplicação do filtro apresentado nesta seção.

Figura 27: Demonstração do resultado do filtro topológico. Da esquerda para a direita, a malha de triângulos original e após a aplicação do filtro topológico.

A comparação entre as cores dos triângulos é realizada sobre uma versão binarizada da imagem original. Essa é uma etapa importante da validação dos pontos pois a classificação dos triângulos considera o resultado do algoritmo de binarização. Para minimizar os efeitos do uso de limiares, a binarização deve ser adaptativa. Bradley e Roth (2007) desenvolveram um algoritmo de binarização que lida bem com as variações de iluminação comuns nessa aplicação de detecção de padrões xadrez. Os autores implementaram uma extensão do algoritmo de Wellner (1993), aumentando o tamanho da janela de análise. A imagem binarizada é construída de acordo com a Equação (26).

$$B(x, y) = \begin{cases} 1, & I(x, y) \cdot w^2 \leq A_W \cdot (1 - t) \\ 0, & I(x, y) \cdot w^2 > A_W \cdot (1 - t) \end{cases} \quad (26)$$

onde B é a imagem binarizada e t um fator de ajuste. No presente trabalho a variável t assume o valor 0.15 como sugerido em Wellner (1993). Da mesma forma que o detector de cantos-x da Seção 5.3 o cálculo da Equação (26) é beneficiado pelo uso de imagem integral.

A fase de binarização pode ser influenciada por problemas provenientes do processo de aquisição de imagens e pode haver transições de cores no interior de alguns triângulos. Esse comportamento é observado principalmente nas bordas onde uma faixa de valores intermediários são considerados indevidamente como brancos ou pretos, ocasionando o surgimento de triângulos brancos com bordas pretas e também o contrário. Na prática, a verificação de transição de cor é feita em uma região mais interna do triângulo, ignorando as bordas com uma distância de dois *pixels*, onde é mais frequente essa imprecisão.

5.6 Propagação e ajuste de coordenadas

Após a aplicação do filtro topológico, espera-se que a malha de triângulos possua as mesmas características topológicas do padrão de calibração. Resta então associar cada vértice às coordenadas reais do padrão.

Para isso, um par de triângulos vizinhos de cores iguais são selecionados arbitrariamente, T_r e T_v , chamados de triângulo de referência e triângulo vizinho. Dos três vértices que compõem T_r , é definido como a origem v_o aquele que possui T_v como triângulo oposto. Aos vértices restantes v_x e v_y , são atribuídos os sentidos horizontal e vertical do plano cartesiano.

Figura 28: Propagação de coordenadas na malha de triângulos.

Os triângulos T_r e T_v da Figura 28 definem o início e o sentido da correspondência. A estrutura de dados que representa um vértice (Figura 23) possui dois sistemas de coordenadas, um em relação à imagem e outro em relação ao padrão de calibração. Um dos vértices opostos à borda compartilhada por T_r e T_v é escolhido como origem. Nesse exemplo, v_o assume as coordenadas $[0 \ 0]^T$ para as coordenadas do padrão. v_x é escolhido entre os outros dois pontos de T_r , como aquele que possui maior variação horizontal em coordenadas de imagem. Caso os pontos possuam a mesma variação, v_x e v_y são escolhidos arbitrariamente. As coordenadas são propagadas em função dos vértices definidos. As coordenadas de padrão dos vértices v_x e v_y são incrementadas em uma unidade correspondente ao seu sentido. Da mesma forma v_{op} , que é o vértice oposto a T_r , também pode ser determinado. Nesse momento, o sinal do incremento pode ser escolhido em função das coordenadas de imagem de cada vértice. Assumindo que os valores de coordenadas no sentido de x aumentem da esquerda para a direita e em y de baixo para cima, a regra de atribuição de sentidos e incremento é dada por:

- Se o valor $|v_x - v_o| > |v_y - v_o|$ em coordenadas de imagem, então v_x está em concordância com o eixo x do plano Cartesiano. Se não, os sentidos dos eixos devem ser trocados.

- Se o componente x do vetor $v_x - v_o$ for maior que 0, então o incremento utilizado é +1 e -1 caso contrário. O mesmo raciocínio é aplicado no sentido do eixo y .

A propagação das coordenadas consiste em estabelecer as posições relativas dos vértices vizinhos. Considerando agora a situação onde T_r é um triângulo cujos vértices já possuem correspondências definidas. v_o é o vértice oposto ao triângulo de mesma cor, v_x o vértice na direção x e v_y na direção y . T_v é um triângulo vizinho de cor diferente de T_r compartilhando a borda e que contém v_o . T_v forma outro quadrilátero de mesma cor com T_{op} , compartilhando a borda e_{op} e v_v é o vértice de T_v que deverá ser determinado. Em função da malha, quatro configurações de triangulação são possíveis.

Figura 29: Configurações da malha de triângulos na propagação de coordenadas.

A Figura 29, apesar de apresentar quatro imagens, ilustra duas situações distintas. A posição de v_v e v_{op} depende do relacionamento entre e_{op} e v_o . Nas Figuras 29(a) e 29(d), a borda e_{op} é oposta ao vértice v_o , enquanto que nas outras a borda e_{op} contém v_o . Além disso, como as posições são relativas, a borda e pode ser formada tanto por v_x quanto por v_y , o que duplica o número de situações apresentadas.

Como T_v é vizinho de T_r , seus dois vértices já possuem correspondências, no entanto, seu terceiro vértice v_v ainda é indeterminado. Diante das quatro possibilidades da malha, são definidas as seguintes regras:

Caso v_o seja um vértice oposto a e_{op} :

$$\square \text{ Se } e \text{ contém } v_x, \text{ então } v_v = [v_y^{(x)} \quad 2v_o^{(y)} - v_y^{(y)}]^\top;$$

$$\square \text{ Se } e \text{ contém } v_y, \text{ então } v_v = [2v_o^{(x)} - v_x^{(x)} \quad v_x^{(y)}]^\top;$$

Caso v_o não seja um vértice oposto a e_{op} , então e_{op} contém v_o , desse modo:

$$\square \text{ Se } e \text{ contém } v_x, \text{ então } v_v = [v_x^{(x)} \quad 2v_o^{(y)} - v_y^{(y)}]^\top;$$

$$\square \text{ Se } e \text{ contém } v_y, \text{ então } v_v = [2v_o^{(x)} - v_x^{(x)} \quad v_y^{(y)}]^\top;$$

Entende-se por $v^{(\cdot)}$ a coordenada $(\cdot)$ do vértice v . Assim, como T_v possui todos os seus vértices definidos, então os vértices de T_{op} também são determinados.

Nesse algoritmo, toda vez que um triângulo é visitado, os vértices do triângulo corrente e do oposto são calculados. O algoritmo é realizado recursivamente para cada triângulo vizinho do par T_v e T_{op} que ainda não foi acessado. Esse algoritmo executa em tempo $O(\frac{n}{2})$, onde n é a quantidade de triângulos presente na malha.

As coordenadas xw e yw definidas na estrutura de dados usada no algoritmo de triangulação são definidas via propriedades de vizinhança e posteriormente ajustadas de forma que não apresentem índices negativos. Assim, considerando um padrão xadrez formado por $n_x \times n_y$ cantos-x tal que $n_x \neq n_y$, a determinação da posição absoluta de cada vértice pode ser encontrada sempre que o conjunto de pontos contém pelo menos um vértice com uma das coordenadas máximas em x e y do padrão. As ordenadas são invertidas caso a coordenada máxima em x ultrapasse o tamanho n_x quando $n_y > n_x$.

Uma vez que a malha segue as mesmas regras topológicas do padrão xadrez, cada vértice corresponde a um ponto de calibração e obstruções encontradas na malha de triângulos não impedem a propagação das coordenadas entre os vértices, se existir pelo menos um triângulo conectando-os ao restante da malha.

5.7 Refinamento do local das características

Uma vez que a qualidade da calibração está diretamente ligada à precisão com que a posição das características são encontradas (HARTLEY; ZISSERMAN, 2004), faz-se necessário uma técnica de refinamento da posição. Os algoritmos tradicionais, como o detector descrito em Harris e Stephens (1988) e Shi e Tomasi (1994), varrem toda a imagem e usam limiares para diferenciar cantos de não-cantos. A posição é aprimorada maximizando funções quadradas que se ajustam ao perfil das intensidades locais da vizinhança

de cada *pixel*. O limiar escolhido tem impacto direto sobre a qualidade de resposta desses detectores, por isso, normalmente são classificados como cantos válidos os N *pixels* com maior resposta ao operador. Krüger e Wöhler (2011) apresentam um estudo que comprova a deficiência desses detectores tradicionais quando a característica de interesse é um canto-x.

A formulação do detector de cantos-x adaptativo da Seção 5.3 não garante que somente um *pixel* seja classificado como canto. Na verdade, uma região próxima ao canto-x é encontrada. O cálculo do ponto médio é uma boa aproximação da própria região mas pode divergir da posição real do ponto de interesse, caso ela não seja perfeitamente simétrica. Para lidar com esse problema, o presente trabalho apresenta a Equação (27) que descreve uma medida de relevância para cada canto, contabilizando o grau de simetria do conjunto Φ dos cantos-x detectados.

$$S = \sum_{i=1}^{\frac{n}{4}} \left| I(\mathbf{p}_i) + I(\mathbf{p}_{i+\frac{n}{2}}) - I(\mathbf{p}_{i+\frac{n}{4}}) - I(\mathbf{p}_{i+\frac{3n}{4}}) \right| \quad (27)$$

As intensidades opostas na circunferência de análise, separadas em 180° , devem ser semelhantes enquanto que as separadas em 90° devem representar uma maior variação de cor. A magnitude da Equação (27) tende a ser grande quando o conjunto Φ é construído em torno de um canto-x. A Figura 30 exemplifica a localização de cada *pixel* usado pela equação em relação à circunferência de vizinhança.

Figura 30: Defasagem dos *pixels* no cálculo da relevância do canto-x. A figura apresenta os *pixels* opostos em 0° , 90° , 180° e 270° partindo de cima pelo sentido horário.

A Figura 31 ilustra o resultado do cálculo de simetria sobreposto à imagem original. A imagem original foi normalizada para realçar o resultado gerado pela Equação (27). A posição do ponto de interesse é atualizada calculando o centro de massa da região em torno do canto-x.

5.8 Implementação da interface gráfica para o *software* de detecção

A interface foi desenvolvida com a intenção de disponibilizar o método ChTD com um algoritmo de calibração, com o objetivo de tornar o processo de calibração de câmeras

Figura 31: Medida de similaridade do canto-x. A posição real do canto-x é aproximada pelo centro de massa da região detectada.

mais simples para o usuário. Para isso, tanto o detector descrito neste capítulo quanto um algoritmo de calibração foram implementados. Como método de calibração foi escolhido o de (ZHANG, 2000), por ser uma técnica que permite a calibração de câmeras via padrões planos e também por ser uma das técnicas mais consolidadas nessa área.

Utilizando a linguagem C/C++, foram implementados todos os algoritmos adicionais para o funcionamento do método ChTD, que inclui os seguintes:

- ❑ Algoritmos básicos para o processamento de imagens:
  - Conversão de imagens coloridas para a escala de cinza;
  - Convolução de sinais bidimensionais;
  - Algoritmos de desenho em imagens digitais;
- ❑ Criação de imagens integral do trabalho de Viola e Jones (2001) e os algoritmos de cálculo de média e desvio padrão otimizados;
- ❑ O algoritmo de triangulação Delaunay apresentado por Guibas, Knuth e Sharir (1992) usando a estrutura de dados da Figura 23;
  - Foram implementados algoritmos inerentes à geometria computacional embasados em Álgebra Linear.
- ❑ A binarização adaptativa de imagens digitais descrita no trabalho de Bradley e Roth (2007).
- ❑ Algoritmo *Scan-Line* para a verificação de variação de cores no interior dos triângulos;
- ❑ Algoritmo de rotulação de *pixels* conectados desenvolvido por He et al. (2009) e a extração de características de cada rótulo.
- ❑ Algoritmo de calibração de câmeras de Zhang (2000):
  - Operações básicas entre matrizes: Adição, Subtração e Multiplicação;

- Transposição de matrizes;
- Decomposição de matrizes: SVD², RQ e Cholesky;
- Solução de sistemas lineares via as decomposições matriciais listadas no item anterior;
- Algoritmo DLT;
- Algoritmo Levenberg-Marquardt para a fase de otimização não linear requerida pelo sistema de calibração proposto em Zhang (2000).

O sistema foi desenvolvido sobre a plataforma Ubuntu 12.10³, usando a biblioteca gtkmm-2.4⁴ para a construção da interface gráfica. No total, quatro módulos formam o sistema. A Figura 32 apresenta uma visão geral da interface criada e, a seguir, cada módulo é comentado.

Figura 32: Interface do sistema de calibração.

Descrição dos componentes do *software*:

- ❑ **Módulo 1:** Esse módulo é responsável pela interface com a câmera. O acesso ao dispositivo é feito via uma biblioteca nativa dos sistemas Linux de nome *video4linux*. Isso permite que qualquer câmera que seja reconhecida pelo sistema operacional seja listada e possa ser configurada por essa parte do programa. Por motivos práticos, esse módulo foi projetado e desenvolvido como um programa independente e com

²O algoritmo de decomposição SVD foi adaptado a partir da implementação disponível no livro Press et al. (1992).

³<http://www.ubuntu.com/>

⁴<http://www.gtkmm.org/>

interface gráfica própria, que pode ser adicionado a qualquer outro *software* que utilize a biblioteca *gtkmm*. Dessa maneira, qualquer outra aplicação que necessite de uma interface de câmera pode utilizá-lo, tornando sua implementação menos trabalhosa e mais simples. O botão *Start* inicia a captura de *frames* usando as configurações selecionadas.

- ❑ **Módulo 2:** O detector de padrões exige dois parâmetros de entrada que definem as dimensões do padrão em xadrez utilizado. O botão *Capture* inicia o detector de cantos. Um tempo de 2 segundos entre cada detecção permite que o padrão seja reposicionado para a próxima captura.
- ❑ **Módulo 3:** Nessa parte do programa é apresentada a lista de imagens capturadas pelo detector de padrões. É disponibilizado o botão *Calibrate* que dispara o algoritmo de calibração. O resultado é escrito em um arquivo e também na linha de comando.
- ❑ **Módulo 4:** Nesse módulo, o vídeo produzido pelo módulo 1 é apresentado. Quando o botão *Capture* é pressionado, o resultado do detector é sobreposto ao vídeo original. Quando o vídeo é interrompido, o resultado do detector pode ser apresentado em função da seleção da lista de imagens do módulo 3.

5.9 Considerações finais

O presente Capítulo descreveu o método ChTD, que é alvo principal do estudo desenvolvido neste trabalho. A metodologia desse método é aplicada na detecção automática de padrões xadrez com o objetivo de fornecer dados para aplicações de calibração de câmeras. O método é composto por cinco etapas de processamento, definidas pelo detector de cantos-x, construção da malha geométrica, filtragem topológica dos cantos detectados, pelo ajuste de coordenadas e pelo refinamento de posição das características.

O detector de cantos-x é equacionado com o objetivo de realizar melhores considerações sobre as variações de intensidades de *pixels* de sua janela de observação. Para isso, a detecção é realizada com base na média e no desvio padrão de cada região. O cálculo é otimizado com o uso de imagens integral.

A filtragem topológica faz uma análise exclusivamente local da vizinhança de cada canto-x com base em uma malha de triângulos gerada pela triangulação de Delaunay. Cada triângulo assume características da região que cobre na imagem permitindo uma verificação topológica das características da imagem. Essa fase pode gerar alguns falsos negativos, mas, sobre uma perspectiva de calibração onde pontos ruidosos interferem significativamente no resultado final, é mais prudente descartar aqueles pontos que não estão plenamente de acordo com a regularidade do padrão.

A fase de ajuste de coordenadas compreende também a propagação de coordenadas dentro da malha de triângulos resultante do filtro topológico. Essa é uma etapa pouco explicada pela maioria dos trabalhos relacionados ao assunto dessa tese, no entanto, este trabalho descreve um método de ordenação dos pontos detectados explorando as características conectivas de uma malha de triângulos.

Por último, o refino da localização de cada ponto é realizado com base na simetria que é esperada em um *pixel* classificado como canto-x. A posição final é obtida extraindo o centro de massa da região relacionada ao canto-x de interesse.

6 Resultados Experimentais

6.1 Considerações iniciais

Este Capítulo faz a exposição dos resultados experimentais da metodologia apresentada pelo Capítulo 5 e a comparação dos resultados obtidos com trabalhos relacionados. Os dados discutidos referem-se à detecção de padrões xadrez, alvo principal do trabalho desenvolvido nesta tese. Os testes são realizados sobre conjuntos de imagens reais e simuladas. No final são apresentados exemplos de detecções quando o padrão encontra-se parcialmente ocluído.

6.2 Materiais e métodos de avaliação

A avaliação de sistemas de detecção automática de padrões de calibração é uma questão que ainda merece maior discussão. Os trabalhos especificamente direcionados à detecção de padrões de calibração dificilmente apresentam métodos de avaliação dos seus algoritmos. Além disso não existe uma metodologia padronizada para a comparação da eficiência dos algoritmos e, de uma maneira geral, somente resultados visuais são discutidos.

Dentro do contexto de calibração de câmeras, espera-se como resultado de um detector de padrões de calibração um conjunto de dados que expresse o relacionamento coerente entre os pontos de controle do mundo real e suas projeções no plano de imagem da câmera que se deseja calibrar. Desse modo, é importante que o conjunto esteja livre de ruídos, agregue todos os pontos disponibilizados pelo padrão e que os pares de dados, que relacionam pontos e suas respectivas projeções, estejam devidamente determinados. A diversidade de pontos capturados em posições e orientações diferentes também é fator importante para a construção da base de calibração. Isso exige que o padrão seja exposto à câmera de formas diferentes de maneira a se obter uma melhor amostra do sistema projetivo que se deseja estimar. É importante então que o padrão seja detectado mesmo em situações onde ele encontra-se muito inclinado em relação à câmera.

Também é desejável que o processo de calibração seja realizado de maneira automática, que a maior quantidade de imagens do padrão seja aproveitada e que seja independente da qualidade das imagens que a câmera consegue capturar. Isso significa que o algoritmo de detecção deve ser robusto à presença de ruídos externos e internos à imagem do padrão de calibração.

A metodologia de análise usada neste Capítulo é construída buscando avaliar o comportamento do detector de padrões quanto a às expectativas listadas nos dois parágrafos

anteriores. Para isso, o sistema de detecção do Capítulo 5 é colocado à prova perante conjuntos de imagens reais e sintéticas. As imagens reais permitem uma comparação direta com outras técnicas enquanto que as imagens sintéticas fornecem meios controlados para verificar a eficiência das técnicas propostas.

São usados três bancos de imagens reais, constituídos por imagens de padrões xadrez em diversas orientações e posicionamentos. Esses bancos representam situações comuns à maioria dos sistemas de calibração, onde as imagens do padrão são primeiramente capturadas e depois processadas de maneira *off-line*. Os bancos de imagens são denominados:

- Exemplo-OCV: esse banco está presente no código fonte da biblioteca OpenCV para exemplificar a funcionalidade da função *findChessBoardCorners*. Ele é formado por 13 imagens numeradas de 1 à 14 – a imagem 10 não está disponível – com resolução 640×480 , cujo padrão de calibração apresenta 9×6 cantos-x.
- Bouguet: é disponibilizado com o *toolbox* de calibração de câmeras para Matlab em Bouguet (2004). Estão presentes 20 imagens de resolução 640×480 que contém um padrão xadrez com 13×12 cantos-x, explorando orientações mais acentuadas.
- OmniSample: é constituído por 6 conjuntos de imagens com alta distorção radial, capturadas por diferentes tipos de lentes. Esse conjunto é disponibilizado no sítio¹ do *software* de calibração de câmeras OCamCalib, que é detalhado nos trabalhos de Scaramuzza, Martinelli e Siegwart (2006) e Rufli, Scaramuzza e Siegwart (2008). A Tabela 2 detalha esse conjunto de imagens.

Tabela 2: Descrição do conjunto de imagens OmniSample.

Conjunto	Nome	Resolução	Cantos-x	n. de imagens
1	<i>Fisheye</i>	752×480	7×6	8
2	<i>KaidanOmni</i>	640×480	11×7	17
3	<i>Ladybug</i>	1024×768	8×5	13
4	<i>MiniOmni</i>	752×480	11×7	15
5	<i>Omni</i>	1280×960	7×6	14
6	<i>VMRImage</i>	1024×768	7×6	10

As imagens sintéticas usadas nos experimentos desta tese foram geradas pelo trabalho de Datta, Kim e Kanade (2009), com o objetivo de testar sua metodologia de calibração de câmeras. Esse conjunto é ideal para avaliar os algoritmos de detecção, uma vez que ele apresenta diversas situações de exposição do padrão aos métodos de calibração. O conjunto é constituído por um padrão xadrez ideal com 10×7 cantos-x que é projetado

¹Sítio do *software* OCamCalib disponível em: <https://sites.google.com/site/scarabotix/ocamcalib-toolbox>

em uma câmera virtual variando combinações de inclinação de $\pm 15^\circ$, $\pm 30^\circ$ e $\pm 45^\circ$ nos eixos X , Y e variações de inclinação de $\pm 5^\circ$ no eixo Z , totalizando 26 imagens para cada módulo de inclinação. Esse banco de imagens agrega também imagens de padrões com marcadores circulares e no formato de anel mas, evidentemente, somente o padrão xadrez é considerado nos experimentos aqui apresentados. A Tabela 3 lista os parâmetros intrínsecos adotados na construção da câmera virtual. Os parâmetros $k1$, $k2$ e $k3$ são referentes à distorção radial, x_0 e y_0 é a posição do ponto principal e f_x e f_y as distâncias focais (Capítulo 2.5).

Tabela 3: Parâmetros da câmera virtual para o conjunto de imagens sintéticas.

Parâmetros:	$k1$	$k2$	$k3$	$p1$	$p2$	x_0	y_0	f_x	f_y
Valores:	-0.7	-0.3	8	0.0004	0.00005	360.5	240.5	1150	1150

A maioria dos resultados apresentados neste Capítulo expõe uma comparação entre a técnica de detecção desta tese principalmente com a função *findChessboardCorners* da biblioteca OpenCV. Essa função é referenciada no texto pelo mesmo nome da biblioteca a qual pertence. As imagens sintéticas permitem uma avaliação mais consistente dos algoritmos de detecção diante da adição controlada de ruídos e posicionamento do padrão na imagem.

Dos trabalhos listados no Capítulo 4, alguns não são avaliados pois requerem participação do usuário ou da detecção de marcadores especiais (BOUGUET, 2004; ESCALERA; ARMINGOL, 2010; SHU; BRUNTON; FIALA, 2009) e outros, que propõem métodos classificados como automáticos Fiala e Shu (2007), apresentam soluções que necessitam de padrões específicos, o que foge do problema de detecção automática que é questionado nesta tese. As técnicas automáticas de detecção de padrões xadrez como as descritas em Sun et al. (2008) e Wang e Wu (2010) apresentam resultados visuais e nenhuma avaliação quantitativa é discutida. Esses trabalhos não disponibilizam o código de detecção e/ou a metodologia usada requer muitas etapas intermediárias que não são descritas exatamente como são implementadas. Esses fatores dificultam a avaliação e/ou a reprodução fiel dos algoritmos em questão.

A técnica comparada se restringe ao algoritmo presente na biblioteca OpenCV, dado pela função *findChessBoardCorners*, por estar mais de acordo com as especificações de detectores completamente automáticos dos padrões de calibração que esta tese investiga. Os algoritmos foram implementados em linguagem C/C++ utilizando como plataforma de desenvolvimento o *software* Eclipse CDT 8.2 Kepler² e um computador com processador AMD Phenom II X6 1100T de 6 processadores e com 8GB de memória RAM. Os *softwares* executam sobre o sistema operacional Linux Ubuntu³ versão 12.10 de 64 bits.

²*Software* Eclipse disponível em: <http://www.eclipse.org/cdt/>

³Sistema Operacional Ubuntu disponível em: <http://www.ubuntu.com/>.

As imagens apresentadas são geradas exclusivamente pelos códigos desenvolvidos neste trabalho. Da biblioteca OpenCV somente a função *findChessBoardCorners* é utilizada. Os gráficos são construídos usando o *software* Matlab R2012a⁴.

6.3 Experimentos com imagens reais

Os experimentos com imagens reais avaliam a resposta dos detectores analisados em função dos três bancos de imagens descritos na seção anterior. Os resultados são gerados aplicando o algoritmo em cada imagem, contabilizando a quantidade de pontos encontrados pelos algoritmos de detecção. São considerados válidos somente os cantos-x definidos pelo padrão de calibração. Devido ao equacionamento do detector de cantos-x apresentado neste trabalho, todas as imagens avaliadas são suavizadas por um filtro gaussiano com $\sigma = 1$. Esse valor foi escolhido por ser equivalente a uma leve perda de foco da câmera, proporcionando a remoção de ruídos, de modo a não interferir nos resultados gerados pela função do OpenCV.

Os testes que usam os bancos de imagens Exemplo-OCV e Bouguet são executados fixando o parâmetro k da Equação 19 em $k = 0,5$ e o raio da circunferência de vizinhança em $r = 4$. Os testes realizados com o banco OmniSample têm o raio fixado em $r = 5$ para as imagens com resolução igual a 1024×768 , por serem imagens com resolução maior. O valor de k foi escolhido como $0,5$ por ser uma escolha intermediária perante seu significado dentro da equação do detector de cantos. Assim, considera-se uma alternância de regiões da imagem quando a intensidade do *pixel* ultrapassa ou fica inferior a 50% do desvio médio das cores dos *pixels* contidos na janela de observação.

Essa padronização permite uma avaliação mais consistente e menos tendenciosa dos resultados gerados pelo sistema ChTD quando comparados com os produzidos pelo algoritmo da biblioteca OpenCV. Na prática, os resultados podem ser melhorados com pequenas modificações, específicas para cada imagem, como por exemplo: aumentando o raio de vizinhança para imagens com resolução maior e aumentando o valor do σ do filtro gaussiano. De uma maneira geral, o sistema ChTD apresenta melhores resultados para imagens mais borradas com σ maior que 1.

6.3.1 Imagens Exemplo-OCV

Primeiramente são avaliados os resultados para o banco Exemplo-OCV. Nessas imagens os padrões estão completamente visíveis e assumem poucas orientações com inclinações acentuadas. A Figura 33 mostra algumas das imagens que compõem esse conjunto. Nesse banco a detecção do padrão é realizada com sucesso pela função do OpenCV, detec-

---

⁴Sítio do *software* Matlab disponível em: <http://www.mathworks.com/>.

tando o padrão de calibração em todas as imagens. Por esse motivo, somente os resultados do método ChTD são apresentados.

Figura 33: Exemplos de imagens do conjunto Exemplo-OCV.

A Tabela 4 apresenta a quantidade de pontos detectados pelo sistema ChTD em cada imagem do banco Exemplo-OCV. A linha “Cantos-x” mostra a quantidade total de cantos-x encontrados na cena pelo método ChTD e a linha “Pontos finais” lista a quantidade final de pontos após a aplicação do filtro topológico. A linha “Pontos extras” contabiliza os falsos positivos, que representam a quantidade de pontos encontrados dentro do padrão mas que ainda obedecem sua regularidade.

São esperados um total de 54 cantos-x por imagem, no entanto, o padrão usado pelo conjunto Exemplo-OCV apresenta bordas muito estreitas, o que permite que os retângulos mais externos ao padrão sejam fundidos com a borda do próprio padrão ou com o restante da cena. Isso faz com que o sistema de detecção de cantos-x do método ChTD identifique o cantos desses retângulos também como pontos de calibração. Esse efeito pode ser observado na Figura 34, a qual apresenta os pontos detectados nas imagens 1, 7, 9 e 13 desse banco que tiveram pontos extras contabilizados.

Tabela 4: Resultados para o conjunto Exemplo-OCV.

Imagem:	1	2	3	4	5	6	7	8	9	11	12	13	14
Cantos-x:	197	155	129	145	150	209	172	131	159	170	132	175	156
Pontos finais:	57	54	54	54	54	54	57	54	57	54	54	59	54
Pontos extras:	3	0	0	0	0	0	3	0	3	0	0	5	0

Os dados da Tabela 4 mostram que o método ChTD pode ser equiparado com a detecção desempenhada pela função *findChessBoardCorners*. O sistema aqui construído consegue encontrar todos os pontos do padrão em todas as imagens, sendo penalizado somente pelo aspecto construtivo do mesmo que favorece a identificação dos cantos mais externos. As Figuras 35 e 36 exemplificam a sequência de execução para a segunda e terceira imagem do conjunto onde o método ChTD obteve sucesso.

Figura 34: Imagens do conjunto Exemplo-OCV com pontos extras. Os pontos extras são pontos mais externos pertencentes ao padrão de calibração que foram erroneamente classificados como cantos-x mas que ainda pertencem ao plano de calibração.

Figura 35: Resultados para a imagem 2 do conjunto Exemplo-OCV. Em (a) a imagem original, (b) está presente a malha formada por todos os cantos-x encontrados, (c) a malha de triângulos após a aplicação do filtro topológico e (d) os pontos após o ajuste de coordenadas.

Figura 36: Resultados para a imagem 3 do conjunto Exemplo-OCV. Em (a) a imagem original, (b) está presente a malha formada por todos os cantos-x encontrados, (c) a malha de triângulos após a aplicação do filtro topológico e (d) os pontos após o ajuste de coordenadas.

6.3.2 Imagens de Bouguet

A avaliação do conjunto de imagens de Bouguet se dá do mesmo modo que na Seção 6.3.1. Neste caso também são apresentados os resultados usando a biblioteca OpenCV. Esse conjunto de imagens é composto por 20 imagens de um padrão com 13×12 cantos-x que totalizam 156 características. As imagens exploram uma boa variabilidade de inclinações do padrão, acentuadas o suficiente para dificultar a detecção do mesmo. A Figura 37 mostra alguns exemplos das imagens contidas nesse banco.

Esse banco de imagens é considerado problemático por apresentar baixo contraste e também por descaracterizar boa parte dos cantos-x nas extremidades do padrão de calibração diante das inclinações acentuadas. Por outro lado, esse banco representa um desafio maior devido a esses problemas mencionados.

A Tabela 5 descreve os resultados alcançados pelo algoritmo ChTD e a Tabela 6 os da biblioteca OpenCV. Para facilitar a comparação, a Figura 38 ilustra as quantidades listadas por ambas as tabelas. De acordo com os resultados, observa-se que o algoritmo ChTD consegue encontrar a grande maioria de pontos de calibração e demonstra ser mais estável que os resultados produzidos pela biblioteca OpenCV.

Uma questão importante deve ser colocada antes da avaliação do sistema de detecção implementado pela biblioteca OpenCV. A função *findChessBoardCorners* tem como retorno um valor lógico, sendo verdadeiro quando todo o conjunto de pontos do padrão é detectado e também quando o algoritmo consegue interligar e ordenar os pontos encon-

Figura 37: Exemplos de imagens do conjunto Bouguet.

Tabela 5: Resultados de ChTD para o conjunto Bouguet.

Imagem:	1	2	3	4	5	6	7	8	9	10
Cantos-x:	196	177	179	179	182	167	174	178	179	191
Pontos finais:	156	156	156	156	156	156	156	156	159	159
Pontos extras:	0	0	0	0	0	0	0	0	3	3

Imagem:	11	12	13	14	15	16	17	18	19	20
Cantos-x:	183	177	177	181	165	171	167	162	177	167
Pontos finais:	156	156	163	162	156	156	156	140	156	155
Pontos extras:	0	0	7	6	0	0	0	0	0	0

Tabela 6: Resultados do OpenCV para o conjunto Bouguet.

Imagem:	1	2	3	4	5	6	7	8	9	10
Pontos detectados:	0	0	0	49	154	0	153	156	156	155

Imagem:	11	12	13	14	15	16	17	18	19	20
Pontos detectados:	156	155	0	154	154	154	156	141	136	156

Figura 38: Resultados para o conjunto de imagens Bouguet.

trados. Para o banco de Bouguet, a função do OpenCV produz uma resposta verdadeira somente para as imagens 9, 11, 17 e 20, causando uma perda de 80% dos quadros disponíveis.

A Tabela 6 foi construída com um resultado intermediário do algoritmo, interno à função referenciada, de maneira a enfatizar o sistema de detecção de cantos dessa biblioteca. A totalidade dos pontos são referentes à quantidade de pontos detectados por grupos diferentes e não significa que eles estão plenamente conectados. Pela função *findChessBoardCorners*, quando o padrão não é detectado, os cantos que foram identificados podem ser acessados, porém eles não formam um único conjunto conexo. Isso impede a determinação das coordenadas de cada ponto e, portanto, eles não são ordenados. Essa função recebe como parâmetro o tamanho do padrão de calibração, e a ordenação é realizada somente com a garantia de que foi encontrado um grupo de pontos interligados com a quantidade exata passada como parâmetro.

Pelo gráfico da Figura 38 pode-se notar que o algoritmo da biblioteca OpenCV não consegue detectar o padrão de calibração, mesmo quando todos os cantos-x são encontrados. O que comprova a deficiência do mecanismo que determina a vizinhança entre pontos próximos. O detector ChTD mostra-se superior ao OpenCV, detectando e ordenando os pontos de interesse em todos os quadros de imagem. O traço “Todos os cantos-x” permite visualizar o montante de pontos que são filtrados pelo algoritmo topológico. As Figuras 39, 40, 41 e 42 mostram a comparação do resultado intermediário da biblioteca OpenCV perante o resultado do detector desenvolvido neste trabalho para as imagens 11, 20, 19 e 18 da base de Bouguet respectivamente.

As Figuras 39 e 40 mostram dois exemplos onde os dois métodos conseguiram detectar o padrão com sucesso. O método ChTD sobressai sobre o resultado do OpenCV na Figura 41, detectando o padrão completamente enquanto o OpenCV apresenta resultados inferiores pois não consegue identificar todos os pontos e também apresenta falsos positivos.

Na Figura 42 percebe-se que a função da biblioteca avaliada não lida bem com uma

(a) Resultado OpenCV

(b) Resultado ChTD

Figura 39: Comparação de resultados para a Imagem 11.

(a) Resultado OpenCV

(b) Resultado ChTD

Figura 40: Comparação de resultados para a Imagem 20.

(a) Resultado OpenCV

(b) Resultado ChTD

Figura 41: Comparação de resultados para a Imagem 19.

Figura 42: Comparação de resultados para a Imagem 18.

inclinação mais acentuada do padrão, principalmente na região onde não foi possível identificar os cantos existentes. Esse efeito também pode ser observado na malha final do método ChTD mas não interfere na detecção do padrão. Com o ChTD, a ausência da malha sobre essa região mais prejudicada pela inclinação não significa que os cantos-x não foram corretamente identificados. Nesse caso específico, os triângulos construídos a partir da triangulação Delaunay não cobrem áreas homogêneas da imagem, o que leva à sua exclusão de acordo com as regras do filtro topológico. A Figura 43 apresenta a malha inicial construída sobre a Imagem 18 de Bouguet e a ordenação final alcançada para o ChTD.

Figura 43: Exemplo do resultado do método ChTD para a imagem 18 de Bouguet. O filtro topológico remove pontos que foram corretamente detectados pelo detector de cantos-x em função da triangulação de Delaunay. Essa figura apresenta o caso da Imagem 18 de Bouguet, onde os triângulos não cobrem regiões homogêneas da imagem.

O algoritmo ChTD também é afetado pela borda externa do padrão de calibração usado por este conjunto de imagens. Nesse caso específico, a borda de separação entre o padrão e o restante da imagem é inexistente. Da mesma forma que nas imagens do banco Exemplo-OCV, os quadriláteros mais externos formam condições para o surgimento de cantos-x com o restante da imagem. Também nesse caso, os pontos extras ainda seguem a regularidade do padrão. Considerando que a presença de pontos extras inviabiliza a

detecção do padrão, a percentagem de quadros perdidos pelo método ChTD é de 20% para o banco de Bouguet contra 80% do algoritmo presente no OpenCV. A Figura 44 mostra as imagens onde ocorrem os pontos extras.

Figura 44: Resultados de ChTD com pontos extras para as imagens de Bouguet. Da esquerda para a direita e de cima para baixo as imagens: 9, 10, 13 e 14 do banco de Bouguet.

6.3.3 Imagens OmniSample

O banco de imagens OmniSample compreende uma classe diferente de problemas para a detecção de padrões de calibração. Ele é constituído de imagens com alta distorção radial comumente geradas por câmeras com lentes olho de peixe, espelhos hiperbólicos e esféricos. Nesse tipo de equipamento as imagens capturadas são muito distorcidas e dificultam o processamento de imagens ou aplicações de Visão Computacional de uma maneira geral. É comum uma fase de retificação de imagens, onde é realizada a remoção da distorção, antes de qualquer processamento digital. A retificação só é possível quando o sistema está pelo menos parcialmente calibrado, o que torna fundamental o papel desempenhado pelos detectores automáticos de padrões de calibração. A Figura 45 mostra alguns exemplos do conjunto OmniSample.

As Tabelas 7 e 8 resumizam os resultados gerados pelo método ChTD e OpenCV sobre esse banco de imagens respectivamente. As colunas “Nº de cantos-x” listam a quantidade esperada de pontos para cada imagem. As colunas “Min”, “Max” apresentam a quantidade mínima e máxima de pontos detectados em todo o conjunto respectivamente. As

Figura 45: Exemplos de imagens do conjunto OmiSample.

colunas “ μ ” representam a quantidade média de pontos encontrados por imagem do conjunto analisado. O cálculo é realizado somando todos os pontos encontrados no conjunto e dividindo esse valor pelo número de imagens de cada conjunto. As colunas “ σ ” mostram o desvio padrão de pontos detectados no conjunto tomando por base cada imagem. Os valores das colunas “Acerto médio” são calculados dividindo os valores das colunas “ μ ” pela coluna “Nº de cantos-x” seguido de uma multiplicação por 100 para representarem porcentagens. As colunas “Detecções” das Tabelas 7 e 8 expõem a porcentagem de padrões detectados por cada método no conjunto. O tamanho de cada padrão e a quantidade de imagens de cada conjunto estão listados na Tabela 2 que descreve esse banco de imagens.

Tabela 7: Resultados do método ChTD para as imagens OmiSample.

Conjunto	Nº de cantos-x	Min	Max	μ	σ	Acerto médio (%)	Detecção (%)
<i>Fisheye</i>	42	18	41	31,1	6,8	74,1	100
<i>KaidanOmni</i>	77	30	67	52,5	13,3	68,1	100
<i>Ladybug</i>	40	36	40	39,7	1,1	99,2	100
<i>MiniOmni</i>	77	61	77	69,8	5,2	84,4	100
<i>Omni</i>	42	36	42	40,4	1,9	96,3	100
<i>VRMImage</i>	42	38	42	39,8	1,6	94,8	100

Pelas últimas colunas das Tabelas 7 e 8, pode ser observado que o método ChTD consegue aproveitar todos os quadros de imagem enquanto que o OpenCV apresenta re-

Tabela 8: Resultados do OpenCV para as imagens OmniSample.

Conjunto	Nº de cantos-x	Min	Max	μ	σ	Acerto médio (%)	Deteccção (%)
<i>Fisheye</i>	42	0	42	5,3	14,8	12,5	25
<i>KaidanOmni</i>	77	0	0	0,0	0	0	0
<i>Ladybug</i>	40	40	40	40,0	0	100	100
<i>MiniOmni</i>	77	0	77	56,9	26,1	73,9	6,66
<i>Omni</i>	42	0	42	38,0	11,0	90,5	42,9
<i>VMRImage</i>	42	42	42	42,0	0	100	100

sultados inferiores para essa classe de imagens distorcidas, apresentando falhas na maioria das imagens consideradas.

Nas imagens *Ladybug* e *VMRImage*, o algoritmo da biblioteca OpenCV se mostrou superior ao implementado nesta tese. No método ChTD, a detecção do padrão depende do detector de cantos-x que apresentou problemas com o tipo de iluminação especular desses dois conjuntos. Algumas partes do padrão estão fortemente iluminadas, o que leva à descaracterização da aparência dos cantos-x presentes no padrão. Os melhores resultados do OpenCV nos testes de *Ladybug* e *VMRImage* podem ser justificados pela abordagem de detecção de seu algoritmo que é baseada na detecção de contornos. Além disso, a estratégia empregada pelo OpenCV realiza testes com sucessivas transformações morfológicas tentando identificar a maior quantidade de pontos. Isso faz do algoritmo implementado mais robusto a ruídos e às condições presentes nesses dois conjuntos de imagens, mas também requer mais avaliações da imagem. O método ChTD tem um resultado inferior mas consegue identificar quase a totalidade dos pontos disponíveis pelo padrão de calibração analisando a imagem em um único passo. É importante ressaltar que a falta de alguns pontos não impede que a detecção final do padrão xadrez tenha sucesso pelo método ChTD, o que não acontece no algoritmo do OpenCV.

O conjunto *KaidanOmni* representa situações onde é mais difícil a detecção do padrão xadrez. O padrão em questão é maior, formado por 77 cantos-x, e as imagens desse conjunto possuem menor resolução do que as imagens dos demais conjuntos desse banco. Por serem imagens com alta distorção radial, a região que contém o padrão é significativamente menor, reduzindo a definição das características de interesse. Nesses exemplos, o OpenCV não consegue identificar nenhum ponto de calibração enquanto que o método ChTD tem um acerto médio dos pontos de 68,1%, aproveitando todos os quadros de imagem (Tabelas 7 e 8).

No conjunto de imagens OmniSample, os padrões utilizados possuem uma borda de separação maior entre o padrão de calibração e o restante da cena, o que evita a detecção dos cantos nos quadriláteros mais externos. Desse modo, nenhum ponto extra foi identificado nas imagens desse conjunto. As Figuras 46, 47 e 48 apresentam exemplos da detecção do padrão sobre algumas das imagens mais distorcidas desse banco. A Figura

49 mostra um exemplo da ordenação gerada pelo método ChTD para as imagens *Omni* e *Fisheye*.

Figura 46: Exemplo da detecção do padrão das imagens *KaidanOmni* usando o método ChTD. Do lado esquerdo a malha de triângulos inicial e do lado direito o detalhe da detecção do padrão sobre a imagem binarizada.

Figura 47: Exemplo da detecção do padrão das imagens *Ladybug* usando o método ChTD. Do lado esquerdo a malha de triângulos inicial e do lado direito a detecção do padrão sobre a imagem binarizada.

6.4 Experimentos com imagens sintéticas

Nesta seção de experimentos, os métodos ChTD e OpenCV são avaliados em condições controladas de ruído na imagem e de inclinação do padrão de calibração. Para os testes são usadas as imagens disponibilizadas por Datta, Kim e Kanade (2009) descritas na Seção 6.2. Para uma melhor avaliação em relação à inclinação do padrão, as imagens foram agrupadas em 3 grupos conforme a inclinação presente nos eixos X e Y : $\text{Imagens} \pm 15$, $\text{Imagens} \pm 30$ e $\text{Imagens} \pm 45$. A Figura 50 apresenta alguns exemplos das imagens disponíveis.

Os testes são realizados variando a variável σ que controla o nível do ruído gaussiano adicionado à imagem. Para cada σ são realizadas 10 execuções em cada imagem e é

Figura 48: Exemplo da detecção do padrão das imagens *MiniOmni* usando o método ChTD. Do lado esquerdo a malha de triângulos inicial e do lado direito o detalhe da detecção do padrão sobre a imagem binarizada.

(a) Ordenação dos pontos na imagem *Omni*. (b) Ordenação dos pontos na imagem *Fisheye*.

Figura 49: Exemplo do resultado final do método ChTD após a fase de ajuste e ordenação de coordenadas. Em (a) um exemplo para as imagens *Omni* e em (b) para as imagens *Fisheye*.

(a) Imagem com rotação $X = 15^\circ, Y = 15^\circ$ e $Z = -5^\circ$ (b) Imagem com rotação $X = 30^\circ, Y = 30^\circ$ e $Z = 5^\circ$ (c) Imagem com rotação $X = 45^\circ, Y = 45^\circ$ e $Z = -5^\circ$

Figura 50: Exemplos das imagens sintéticas usadas neste trabalho.

	Imagens $\pm 15^\circ$			Imagens $\pm 30^\circ$			Imagens $\pm 45^\circ$		
	<i>X</i>	<i>Y</i>	<i>Z</i>	<i>X</i>	<i>Y</i>	<i>Z</i>	<i>X</i>	<i>Y</i>	<i>Z</i>
Imagem 1	0	15	0	0	30	0	0	45	0
Imagem 2	0	15	5	0	30	5	0	45	5
Imagem 3	-15	0	5	-30	0	5	-45	0	5
Imagem 4	-15	15	0	-30	30	0	-45	45	0
Imagem 5	-15	-15	5	-30	-30	5	-45	-45	5
Imagem 6	15	-15	5	30	-30	5	45	-45	5
Imagem 7	0	0	-5	0	0	-5	0	0	-5
Imagem 8	0	-15	-5	0	-30	-5	0	-45	-5
Imagem 9	-15	15	5	-30	30	5	-45	45	5
Imagem 10	15	0	-5	30	0	-5	45	0	-5
Imagem 11	15	-15	0	30	-30	0	45	-45	0
Imagem 12	-15	15	-5	-30	30	-5	-45	45	-5
Imagem 13	15	15	-5	30	30	-5	45	45	-5
Imagem 14	0	0	5	0	0	5	0	0	5
Imagem 15	0	-15	5	0	-30	5	0	-45	5
Imagem 16	15	0	0	30	0	0	45	0	0
Imagem 17	15	0	5	30	0	5	45	0	5
Imagem 18	15	15	0	30	30	0	45	45	0
Imagem 19	-15	15	5	-30	30	5	-45	45	5
Imagem 20	0	-15	0	0	-30	0	0	-45	0
Imagem 21	0	-15	0	0	-30	0	0	-45	0
Imagem 22	0	15	-5	0	30	-5	0	45	-5
Imagem 23	-15	0	-5	-30	0	-5	-45	0	-5
Imagem 24	-15	15	0	-30	30	0	-45	45	0
Imagem 25	-15	-15	-5	-30	-30	-5	-45	-45	-5
Imagem 26	15	-15	-5	30	-30	-5	45	-45	-5

Tabela 9: Parâmetros de rotação para as imagens sintéticas. As colunas *X*, *Y* e *Z* apontam o ângulo usado para a rotação de cada eixo.

contabilizada a quantidade média de pontos detectados no padrão de calibração. O fator de ajuste do detector de cantos do método ChTD é $k = 4$ e as imagens são suavizadas por um filtro gaussiano com $blur = 1$. Nestes testes, a nomenclatura $blur$ é o desvio padrão do filtro gaussiano, sendo usada para evitar confusões com o σ do ruído gaussiano. A Figura 51 apresenta o gráfico que relaciona os resultados para o número médio de pontos encontrados em cada classe de imagem sintética utilizada para a biblioteca OpenCV. O eixo x dos gráficos representa a intensidade de ruído gaussiano e o eixo y lista as imagens de acordo com a sequência apresentada pela Tabela 9. As cores mais próximas à cor vermelha significam alta taxa de identificação enquanto que as próximas à cor azul uma baixa taxa de identificação dos pontos do padrão. A quantidade máxima de cantos encontrados foi de 70, o que é esperado pelo padrão de calibração utilizado.

Observa-se pela Figura 51 que o algoritmo presente na biblioteca OpenCV possui

Figura 51: Porcentagem média de pontos detectados com a biblioteca OpenCV.

boa resposta para imagens onde o padrão encontra-se pouco inclinado, sendo robusto à presença de ruídos nas imagens do grupo de $\pm 15^\circ$. Nos grupos de imagens com inclinações mais acentuadas, $\pm 30^\circ$ e $\pm 45^\circ$, o OpenCV possui dificuldades em detectar os pontos de interesse, principalmente naquelas onde as inclinações acontecem ao mesmo tempo nos dois eixos X e Y . Pode-se dizer que existe uma dependência direta à forma com que o padrão é apresentado para a câmera. Dentro do contexto de calibração de câmeras, onde é importante uma boa amostragem do sistema projetivo da câmera, isso impede que a estimação dos parâmetros de câmera se aproximem dos valores reais. A Figura 52 mostra o mesmo gráfico para o método ChTD.

Figura 52: Porcentagem média de pontos detectados com o método ChTD.

Pode-se verificar que o método ChTD é mais estável em relação à inclinação do padrão a partir da Figura 52. A quantidade média de pontos detectados é dependente de um valor de $40 < \sigma < 50$. Isso aponta uma sensibilidade ao ruído presente na imagem. Para valores de σ muito altos a vizinhança do *pixel* avaliada pelo detector de cantos-x do ChTD

é diretamente afetada pela presença do ruído, que leva a uma quantificação do número de alternâncias maior que a esperada. Além disso o ruído também influencia o *pixel* central e a média da janela de observação afetando diretamente a resposta do detector utilizado. É de se esperar então uma resposta correlacionada ao nível de ruído da imagem. Entretanto, a ausência de pontos detectados no resultado final não significa necessariamente que os todos os pontos não foram detectados. Na última etapa do algoritmo, alguns pontos podem ser desconsiderados em função do filtro topológico aplicado. Se a resposta da binarização da imagem muda em função do ruído então é de se esperar que as regiões de imagem que pertencem aos quadriláteros do padrão não sejam mais homogêneas e possuam transições de cores em seu interior.

Outro fator importante a ser avaliado é que, pela formulação do método ChTD, a ausência de alguns cantos-x não significa que o padrão de calibração não foi identificado. Uma métrica que permite avaliar essa questão é a taxa de quadros perdidos referente a cada método considerado. As Figuras 53 e 54 apresentam essa taxa para o método do OpenCV e o método ChTD. As imagens permitem comparar a quantidade de quadros perdidos em ambos os algoritmos analisados. Nessas figuras, as cores mais avermelhadas relatam uma maior taxa de perdas de quadros e as mais azuladas um sucesso maior na detecção do padrão de calibração.

Figura 53: Porcentagem média de quadros perdidos a biblioteca OpenCV.

A análise da taxa média de quadros perdidos permite concluir que o método ChTD tem um comportamento superior ao método presente na biblioteca OpenCV. Nessa biblioteca, a resposta de detecção é em função da identificação completa do conjunto de característica do padrão de calibração, restringindo a resposta final do sistema de detecção. Já no método desta tese, mesmo quando o detector de cantos-x falha, o algoritmo de detecção do padrão ainda consegue extrair informações da imagem referentes aos pontos de calibração. Mesmo para uma quantidade reduzida de pontos, o ChTD ainda constrói a base de dados de calibração.

Figura 54: Porcentagem média de quadros perdidos pelo método ChTD.

Para verificar a influência do ruído na resposta dos métodos ChTD e OpenCV, as Figuras 55, 56, 57 e 58 apresentam gráficos similares dos resultados discutidos anteriormente para as imagens de entrada suavizadas com um filtro gaussiano construído com $blur = 1, 5$. Nessas figuras observa-se que ambos métodos apresentam melhoras na detecção dos pontos de calibração.

Figura 55: Porcentagem média de pontos detectados com a biblioteca OpenCV após a aplicação do filtro gaussiano com $blur = 1, 5$.

Nas Figuras 55 e 56, pode-se notar que o pré-processamento utilizado na imagem favorece o resultado do método desenvolvido neste trabalho. O filtro gaussiano aplicado reduz a quantidade de ruído na imagem e o ChTD permanece estável em relação à inclinação do plano de calibração ao mesmo tempo que melhora sua taxa média de pontos detectados.

Do mesmo modo, as Figuras 57 e 58 permitem comparar a taxa de quadros perdidos entre as duas técnicas avaliadas. O algoritmo do OpenCV, na média, não perde nenhum quadro de imagem para o grupo de $\pm 15^\circ$, o que é um ótimo resultado. Isso pode ser

Figura 56: Porcentagem média de pontos detectados com o método ChTD após a aplicação do filtro gaussiano com $blur = 1, 5$.

Figura 57: Porcentagem média de quadros perdidos a biblioteca OpenCV após a aplicação do filtro gaussiano com $blur = 1, 5$.

Figura 58: Porcentagem média de quadros perdidos pelo método ChTD após a aplicação do filtro gaussiano com $blur = 1, 5$.

observado pela barra de cores que apresenta valor máximo de 90% de perdas. Por outro lado, ele permanece muito dependente da inclinação do padrão de calibração. Já o método ChTD tem seus resultados melhorados significativamente. Isso mostra uma dependência entre a resposta produzida por este método e o grau de suavidade que a imagem de entrada apresenta, o que é de se esperar pela modelagem do detector de canto-x que ele implementa.

6.5 Resultados considerando oclusões

Nesta seção de resultados, o método ChTD é colocado diante de imagens de padrões de calibração que possuem parte de seus cantos-x oclusos por objetos que compõem a cena. Aqui é exemplificado a eficiência do método em questão com relação ao seu filtro topológico e também a fase de ajuste de coordenadas. Para isso, foram usadas algumas imagens forçando oclusões do padrão de calibração. Os resultados apresentados nesta seção se restringem a apresentação visual do conjunto de pontos ordenados pelo método ChTD.

Nas Figuras 59 e 60 estão ilustrados os resultados do algoritmo ChTD mediante a oclusão parcial do padrão. Na sequência é apresentada a triangulação de todos os cantos-x, o resultado após o filtro topológico, coordenadas ordenadas e o plano reprojetoado por meio da matriz projetiva que foi calculada pelos pontos detectados. Esses exemplos representam situações onde a perda de cantos-x é dada pela falta de informação presente na imagem em vez das falhas provocadas pelo detector de cantos-x. Pode-se verificar que os cantos-x ocultos não interferem na propagação das coordenadas, sendo possível ordenar os pontos encontrados e utilizar o máximo de características de calibração para a estimação dos parâmetros de câmera. Os pontos são ordenados pela etapa de propagação sempre entre os triângulos que compartilham pelo menos uma aresta.

6.6 Considerações finais

Este Capítulo apresentou a comparação do método ChTD com o algoritmo implementado pela função *findChessBoardCorners* da biblioteca OpenCV que é um dos mais usados pelos trabalhos relacionados à área de calibração de câmeras.

A comparação foi realizada com o uso de bancos padronizados de imagens reais e sintéticas. As imagens reais permitiram uma avaliação numérica do método desenvolvido nesta tese, enquanto que os testes com imagens sintéticas permitiram a verificação do comportamento da metodologia de detecção do ChTD diante de perturbações controladas, como a inclinação do padrão e o nível de ruído presente nas imagens.

O método ChTD conseguiu ser superior em relação ao OpenCV pela quantidade de acerto médio atingido em cada grupo de imagens avaliado e por ser mais invariante ao

Figura 59: Resultado para padrões com oclusão parcial (Exemplo 1).

Figura 60: Resultado para padrões com oclusão parcial (Exemplo 2).

posicionamento do padrão perante a câmera. Outra questão importante a ser observada é que o método ChTD consegue detectar o padrão de calibração mesmo quando todos seus pontos de controle não sejam detectados. A ausência de alguns pontos no conjunto final de cantos encontrados pode ser ocasionada tanto pela presença de ruído no interior do padrão de calibração quanto pela oclusão parcial do mesmo. O método ChTD é mais robusto a essas duas situações, conseguindo extrair informações da imagem sempre que seja possível encontrar pontos que sigam a regularidade do padrão xadrez.

A detecção topológica dos pontos de calibração do método ChTD processa de maneira local ao ponto de interesse e consegue lidar com a alta distorção presente nas imagens, permitindo extrair informações de todas os quadros avaliados. O algoritmo é executado uma única vez, enquanto que o OpenCV faz uma busca interna pelo melhor resultado, alterando parâmetros do seu algoritmo implementado.

7 Conclusões

Este trabalho apresentou uma proposta de metodologia aplicada à detecção automática de padrões xadrez para a calibração de câmeras, denominada ChTD. A metodologia empregada no método ChTD tem por base três etapas principais: um detector de cantos-x estrutural e específico para a característica presente nos padrões de calibração xadrez; a filtragem topológica com base na regularidade desse padrão; e a propagação e ajuste de coordenadas dos pontos detectados.

O detector de cantos implementado no ChTD foi baseado no trabalho de (KLEIN; MURRAY, 2007). No entanto, o algoritmo foi modificado para realizar uma detecção adaptativa em função de uma janela de vizinhança do ponto de interesse. A filtragem topológica é similar à proposta de Shu, Brunton e Fiala (2009), diferindo quanto aos atributos topológicos analisados e o modo com que a filtragem é desempenhada. Nessa etapa é construída uma malha de triângulos que, diferentemente da proposta de Shu, Brunton e Fiala (2009), evita que os triângulos sejam removidos indevidamente, uma vez que eles podem representar conexões entre pontos válidos. A propagação de coordenadas foi um algoritmo desenvolvido para ordenar os pontos detectados de maneira coerente com a regularidade do padrão. Essa etapa é robusta a ruídos, uma vez que considera a conectividade da malha de triângulos já filtrada.

A partir dessa metodologia foi implementada uma ferramenta que permite extrair pontos de calibração de imagens de um padrão xadrez de forma simples e completamente automática.

A estratégia de avaliação do método ChTD teve por base uma comparação com a função *findChessBoardCorners* da biblioteca OpenCV (BRADSKI, 2000), usando imagens reais e sintéticas. A primeira, com o objetivo de realizar comparações numéricas da quantidade de pontos e de padrões detectados com base em conjuntos de imagens padronizados, e a segunda, permitindo uma avaliação mais generalizada do método ChTD.

Os testes apresentados foram executados sobre três bancos de imagens reais que são disponibilizados por trabalhos influentes no campo de detecção automática de padrões de calibração. O primeiro banco pertence à biblioteca OpenCV, sendo formado por imagens de fácil detecção e foram usadas para a validação do método ChTD perante essa mesma biblioteca. O segundo banco pertence ao *Toolbox* para Matlab de calibração de câmeras desenvolvido em Bouguet (2004). Esse banco contém imagens de um padrão xadrez com 156 pontos de calibração, com variações de inclinação em relação à câmera. Os teste sobre as imagens de Bouguet mostraram que o método ChTD produz melhores resultados que o método presente no OpenCV, sendo superior na quantidade de pontos detectados, na

quantidade de padrões detectados e também na estabilidade da resposta de detecção do algoritmo. Essas afirmações podem ser comprovadas pelos dados listados nas Tabelas 6 e 5.

Nesses dois primeiros bancos de imagens, o método ChTD identificou pontos extras do padrão de calibração devido às características construtivas do mesmo, que favorecia a identificação de seus quadriláteros mais externos. Apesar disso, o número de quadros em que o padrão foi detectado com sucesso foi superior ao alcançado pelo OpenCV, reduzindo a taxa de perda de quadros obtidas pelo OpenCV de 80% para somente 20% sobre o banco de Bouguet.

O terceiro banco é formado por 6 conjuntos de imagens com alta distorção, explorando diferentes tipos de lentes (SCARAMUZZA; MARTINELLI; SIEGWART, 2006; RUFLI; SCARAMUZZA; SIEGWART, 2008). Dentre os 6 conjuntos avaliados, no que gerou pior resultado, o método ChTD apresentou 68,1% de acerto dos pontos disponíveis quando o outro algoritmo avaliado não consegue identificar nenhum ponto do padrão. Nesse banco o OpenCV apresentou resultados melhores para dois conjuntos de imagens, identificando com sucesso a quantidade total de pontos de calibração e detectando todos os padrões apresentados, obtendo uma taxa de acerto de 100% para esses dois conjuntos de imagens. Por outro lado, o resultado gerado pelo método ChTD ficou bem próximo ao algoritmo comparado, apresentando resultados inferiores apenas na taxa de acerto médio de pontos que foi de 99,2% para as imagens do conjunto 3 e de 94,5% para o conjunto 6. Para os outros conjuntos de imagens, o método ChTD apresentou resultados significativamente superiores como apresentado nas Tabelas 7 e 8. Os resultados obtidos sobre o terceiro banco de imagens podem ratificar a mesma conclusão tirada dos dois primeiros conjuntos de testes, com a vantagem de nenhum canto extra ser contabilizado.

Também foi testado um conjunto de imagens sintéticas formado por imagens onde o padrão de calibração encontra-se com orientações bem definidas (DATTA; KIM; KANADE, 2009), permitindo uma verificação controlada do efeito do ruído sobre os métodos avaliados. Por meio dos resultados produzidos por esse conjunto de imagens, foi possível comprovar uma menor dependência do método ChTD em relação à inclinação do padrão de calibração enquanto que o método implementado no OpenCV é fortemente afetado por esse problema. Esse comportamento pode ser visualizado nas Figuras 51 e 52. Usando esse mesmo banco, observa-se que a estratégia de detecção do padrão de calibração que é implementada pelo ChTD consegue aproveitar a maioria dos quadros de calibração, mostrando-se superior à resposta final do OpenCV. As Figuras 53 e 54 permitem essa última verificação.

O detector de cantos-x presente no método ChTD foi projetado para trabalhar com imagens suavizadas, o que o tornou um pouco mais sensível ao ruído do que o algoritmo desenvolvido no OpenCV para imagens com pouca inclinação. Essa sensibilidade pode ser observada nas Figuras 57 e 58, onde se percebe uma faixa de atuação em níveis de

ruídos mais bem definidos. Entretanto, os resultados obtidos pela metodologia desenvolvida neste trabalho é mais estável às variações da apresentação do padrão de calibração em relação ao OpenCV, tendo seu desempenho melhorado com o uso de imagens mais suavizadas.

Com base nos testes realizados, pôde-se comprovar que o método ChTD melhora a taxa de aproveitamento de quadros de imagens de calibração em comparação à principal técnica de detecção automática de padrões xadrez, disponível na biblioteca OpenCV (BRADSKI, 2000) por meio da função *findChessBoardCorners*, uma vez que a detecção não se restringe à quantidade total de pontos disponíveis. Além disso, apresenta a vantagem de detectar pontos de calibração de forma consistente, tendo por base a organização topológica de seus vizinhos se comparados ao OpenCV. A função *findChessBoardCorners* é referenciada no texto pelo mesmo nome da biblioteca a qual pertence.

O detector de cantos-x implementado no ChTD possui pontos positivos sobre outros detectores de cantos da literatura: o detector é específico para a característica de interesse; os cálculos empregados pelo detector são facilmente otimizados com o uso de imagens integral; a metodologia de detecção realiza uma detecção adaptativa para cada ponto em função de sua vizinhança, sendo menos dependente de um valor global de limiarização; outra vantagem é que, ao mesmo tempo que o detector de cantos-x avalia a janela de observação, uma resposta classificatória é obtida imediatamente após a análise da circunferência vizinha.

Outro aspecto importante do método ChTD é que a detecção é realizada em um único passo, ou seja, o algoritmo é executado somente uma única vez para toda a imagem. Em contrapartida, o processamento desempenhado pelo método presente na biblioteca OpenCV faz uso de sucessivas transformações da imagem original retornando o melhor resultado obtido. Apesar do sistema de detecção do OpenCV ter sido projetado para a calibração de câmeras em tempo real, a resposta da função *findChessBoardCorners* é demorada quando o padrão não é detectado nas primeiras iterações.

Adicionalmente, a metodologia usada pelo ChTD possui uma formulação modularizada, sendo possível utilizar outros algoritmos em suas fases principais. O detector de cantos-x fornece um conjunto de dados que é avaliado pelo filtro topológico do ChTD. Esse conjunto de pontos pode ser gerado por outro algoritmo de detecção de características, independentemente do restante da aplicação. Entretanto, a qualidade desse conjunto pode ter forte influência no resultado do filtro topológico. É importante que os cantos-x estejam próximos de sua posição real na imagem para que os triângulos da malha geométrica cubram corretamente as regiões mais homogêneas. Por isso, o uso do detector de cantos-x desenvolvido nesta tese é mais recomendado.

O filtro topológico do método ChTD consegue considerar todos os triângulos que possivelmente formam quadriláteros válidos sobre o padrão de calibração. As regras de filtragem são verificadas seguindo uma ordem de menor para maior grau de restrição,

removendo somente os triângulos cujas propriedades topológicas são ambíguas. Essas características do filtro topológico possibilitam considerar um maior número de triângulos, descartando somente aqueles que se encontram fracamente conectados à malha.

A etapa de propagação e ajuste de coordenadas também pode ser realizada por outros métodos de ordenação. No entanto, o método de ordenação proposto neste trabalho é mais robusto a ruídos e executa em tempo $O(\frac{n}{2})$, sendo n o número de triângulos na malha filtrada.

Para as fases intermediárias, os algoritmos utilizados foram escolhidos com base nas discussões e resultados que seus autores apresentaram, mas isso não impede que outros algoritmos sejam avaliados.

De uma maneira geral, o método desenvolvido nesta tese apresentou melhores resultados para as mesmas imagens processadas pelo OpenCV, com as principais vantagens de detectar uma maior quantidade de cantos, de ser mais independente à inclinação do padrão de calibração, por detectar o padrão mesmo quando todos os cantos não puderam ser identificados e por realizar o procedimento em um único passo. Além disso, o ChTD é altamente recomendado para aplicações que requeiram dados sobre padrões xadrez com inclinações acentuadas em relação câmera. Isso porque o método ChTD é independente do resultado de qualquer detector de pontos de interesse. Essa é uma característica que deve ser considerada pois, mediante as inúmeras perturbações que as imagens estão sujeitas, torna-se difícil a detecção de todos os pontos de calibração.

7.1 Principais contribuições alcançadas pelo método ChTD

Pode-se destacar as seguintes contribuições obtidas com a implementação da metodologia do ChTD:

- ❑ A elaboração do detector de cantos-x. Esse detector é específico para as características presentes nos padrões xadrez, o que permitiu reduzir a quantidade de falsos positivos iniciais. A detecção é realizada de forma adaptativa, com base em limiares que são calculados em função da janela de vizinhança do ponto de interesse. Esse detector apresenta algumas vantagens por permitir cálculos otimizados e também por verificar uma pequena parcela da região vizinha para a contagem do número de alternâncias em torno do *pixel* central;
- ❑ A construção de um filtro topológico com base na regularidade do padrão xadrez. A avaliação de características topológicas aumenta a robustez do algoritmo à presença de ruídos, sejam eles internos ou externos ao padrão, permitindo desconsiderar os pontos de controle que não possam ser seguramente classificados como características

apresentadas pelo padrão xadrez. Essa estratégia é importante dentro do contexto de calibração de câmeras onde é indesejável a presença de *outliers*, ou seja, pontos que se diferenciam muito do restante dos dados encontrados;

- ❑ A propagação e ajuste de coordenadas seguindo as características topológicas de triângulos conectados. Essa abordagem permite a determinação das coordenadas do padrão em função da conectividade dos triângulos que compõem a malha filtrada considerando suas características topológicas. O tempo de execução só depende da quantidade de triângulos encontrados na malha final;
- ❑ A estratégia de detecção de padrões é menos dependente da quantidade de pontos encontrados. A malha geométrica definida pela vizinhança de cada ponto permite recuperar a posição relativa de cada característica e extrair os dados de calibração mesmo quando parte da informação não está disponível ou corrompida pela presença de ruído.
- ❑ A detecção do padrão em imagens com alta distorção radial. A filtragem topológica independe de características globais, sendo possível uma avaliação local de cada triângulo presente na malha. Os parâmetros considerados pelas regras definidas pelo filtro topológico são baseados na regularidade do padrão, que permanece inalterada mesmo em imagens muito distorcidas.
- ❑ A implementação de uma interface gráfica que foi idealizada para facilitar a calibração de câmeras usando os dados gerados pelo método ChTD. Nela, também está implementado o método de calibração de câmeras de Zhang (2000), um dos mais empregados pelos estudos da área. Isso torna o processo de calibração de câmeras mais conveniente para aplicações que requeiram uma rápida determinação dos parâmetros de câmera;
- ❑ Por último, as propostas de classificação dos padrões de calibração que são frequentemente empregados pela pesquisa científica ou pelos *softwares* comerciais. Foram sugeridas duas abordagens de categorização, baseada na dimensão dos padrões e também em função dos marcadores e características que o compõem.

O resultado final do ChTD também pode ser aplicado facilmente à calibração de sistemas multioculares, uma vez que a detecção do padrão se dá de forma mais estável, sempre que é possível identificar a posição absoluta de cada ponto encontrado em função do sistema de coordenadas do mundo real.

7.2 Trabalhos futuros

Uma primeira proposta para trabalhos futuros é a investigação da dependência do raio de vizinhança analisado em função do borramento da imagem. Essa pesquisa daria condições de dimensionar melhor tais parâmetros e também uma melhor compreensão das limitações do método ChTD.

Outra forma de melhorar a detecção do padrão seria: dado que o padrão esteja completamente visível na imagem e se nem todos os pontos foram identificados corretamente, ou seja, se a quantidade de cantos-x na malha filtrada for menor que o número de cantos-x apresentados pelo padrão, os outros pontos podem ter sua localização estimada pela re-projeção de suas coordenadas no plano de imagem. Essas projeções definiriam uma região de procura mais plausível, onde o ponto inicialmente tido como falso negativo deva estar. Essa seria uma estratégia com grandes chances de aumentar o acerto médio do método ChTD.

Uma outra abordagem pode surgir de um processamento que emprega o uso pirâmides de imagens, investigando uma abordagem multiresolução para a detecção inicial do padrão de calibração.

Também podem ser sugeridas algumas modificações do método ChTD:

- ❑ A análise da transição de cores dentro dos triângulos da malha, realizada pela **Regra 1** do filtro topológico, pode ser verificada em função de um fator de ajuste que modela uma porcentagem aceitável de área que possui cores diferentes. Isso pode evitar que o processo de binarização de imagens tenha papel importante na exclusão de triângulos da malha;
- ❑ Podem ser anexadas às regras topológicas verificações sobre a geometria e forma dos triângulos que são classificados como triângulos válidos. Algumas características como área do triângulo, e descritores de formas poderiam ser avaliados considerando que esses atributos apresentem uma variação suave em relação aos atributos dos triângulos vizinhos;

7.3 Trabalhos publicados e premiações

- ❑ LAUREANO, G. T.; PAIVA, M. S. V. de; SOARES, A. S.. Topological detection of chessboard pattern for camera calibration. In: Proceedings of The 17th International Conference on Image Processing, Computer Vision & Pattern Recognition, 2013, Las Vegas.
- ❑ LAUREANO, G. T.; PAIVA, M. S. V. de; COELHO, C. J.. Detecção Topológica de Padrões para Calibração de Câmeras. In: Anais do VIII Workshop de Visão Computacional - VIII WVC, 2012, Goiânia.

- Prêmio de Terceiro melhor artigo publicado no VIII Workshop de Visão Computacional 2012 - VIII WVC2012, Instituto de Informática da Universidade Federal de Goiás - INF/UFG.

7.4 Trabalhos submetidos

- LAUREANO, G. T.; PAIVA, M. S. V. de; SOARES, A. S.. Topological detection of chessboard pattern for camera calibration. In: Proceedings of The International Conference on Computer Vision, ICCV2013, Sydney.

7.5 Trabalhos a serem publicados

- LAUREANO, G. T.; PAIVA, M. S. V. de. A topological approach for chessboard detection in images. 2014 IEEE Conference on Computer Vision and Pattern Recognition, 2014, Ohio.
- LAUREANO, G. T.; PAIVA, M. S. V. de. ChTD: A topological approach for chessboard detection in images. Machine Vision and Applications.

Referências Bibliográficas

AGULLO, E. et al. QR Factorization of Tall and Skinny Matrices in a Grid Computing Environment. In: **24th IEEE International Parallel and Distributed Processing Symposium (IPDPS 2010)**. Atlanta, États-Unis: [s.n.], 2010. Disponível em: <<http://hal.inria.fr/inria-00548900>>.

AHN, S. J.; RAUH, W.; KIM, S. I. Circular coded target for automation of optical 3D-measurement and camera calibration. **International Journal of Pattern Recognition and Artificial Intelligence**, 2001. World Scientific Publishing Company, v. 15, n. 6, p. 905–919, 2001. Disponível em: <<http://www.worldscientific.com/doi/pdf/10.1142/S0218001401001222>>.

AKBARZADEH, A. et al. Towards Urban 3D Reconstruction from Video. In: **Third International Symposium on 3D Data Processing, Visualization, and Transmission (3DPVT'06)**. IEEE, 2006. p. 1–8. ISBN 0-7695-2825-2. Disponível em: <<http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=4155703>>.

ALBARELLI, A.; RODOLÀ, E.; TORSELLO, A. Robust Camera Calibration using Inaccurate Targets. **Procedings of the British Machine Vision Conference 2010**, 2010. British Machine Vision Association, p. 16.1–16.10, 2010. Disponível em: <<http://www.bmva.org/bmvc/2010/conference/paper16/index.html>>.

AZIZ, A. Y. I.; KARARA, H. M. Direct linear transformation into object space coordinates in close-range photogrammetry. In: **Proc. of the Symposium on Close-Range Photogrammetry**. Urbana, Illinois: [s.n.], 1971. p. 1–18.

BALTES, J. Camera calibration using rectangular textures. In: **Proceedings of the International Workshop on Robot Vision**. London, UK, UK: Springer-Verlag, 2001. (RobVis '01), p. 245–251. ISBN 3-540-41694-3. Disponível em: <<http://dl.acm.org/citation.cfm?id=646995.711269>>.

BAOZHANG, L.; YANPING, C. Study on edge subpixel location of ellipse in computer vision measurement. In: **2010 3rd International Congress on Image and Signal Processing**. Ieee, 2010. p. 1684–1688. ISBN 978-1-4244-6513-2. Disponível em: <<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5457474>>.

BARRETO, J. et al. Automatic camera calibration applied to medical endoscopy. In: **Proceedings of the 20th British Machine Vision Conference, London, UK**. [s.n.], 2009. Disponível em: <<http://perception.inrialpes.fr/Publications/2009/BRSF09>>.

BARRETO, J. P. **EasyCamCalib Software**. 2009. Disponível em: <<http://arthronav-isr.uc.pt/easycamcalib/>>.

BENNETT, S.; LASENBY, J. ChESS-Quick and Robust Detection of Chess-board Features. **CoRR**, 2013. abs/1301.5, p. 1–13, 2013. Disponível em: <<http://www-com-serv.eng.cam.ac.uk/~sb476/publications/ChESS.pdf> <http://arxiv.org/abs/1301.5491>>.

BERG, M. D.; CHEONG, O.; KREVELD, M. V. **Computational Geometry: Algorithms and Applications**. Third. [S.l.]: Springer-Verlag ISBN: 978-3-540-77973-5, 2008.

BERN, M.; EPPSTEIN, D. Mesh generation and optimal triangulation. **Computing in Euclidean geometry**, 1992. World Scientific, v. 1, n. 1, p. 23–90, 1992. Disponível em: <<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.41-.3221\rep=rep1\type=pdf>>.

BEVILACQUA, A.; GHERARDI, A.; CAROZZA, L. Automatic Perspective Camera Calibration Based on an Incomplete Set of Chessboard Markers. **2008 Sixth Indian Conference on Computer Vision, Graphics & Image Processing**, 2008. Ieee, p. 126–133, dez. 2008. Disponível em: <<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4756061>>.

BISHOP, C. M. **Pattern Recognition and Machine Learning (Information Science and Statistics)**. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 2006. ISBN 0387310738.

BLINN, J. F. Models of light reflection for computer synthesized pictures. **SIGGRAPH Comput. Graph.**, 1977. ACM, New York, NY, USA, v. 11, n. 2, p. 192–198, jul. 1977. ISSN 0097-8930. Disponível em: <<http://doi.acm.org/10.1145/965141.563893>>.

BOUGUET, J. Camera calibration toolbox for matlab. 2004. 2004. Disponível em: <http://www.vision.caltech.edu/bouguetj/calib_doc/>.

BOWMAN, J. **ROS: Camera Calibration**. 2011. Disponível em: <http://www.ros.org/wiki/camera_calibration/Tutorials/MonocularCalibration>.

BRADLEY, D.; ROTH, G. Adaptive thresholding using the integral image. **Journal of Graphics Tools**, 2007. p. 13–21, 2007.

BRADSKI, G. The OpenCV Library. **Dr. Dobb's Journal of Software Tools**, 2000. 2000. Disponível em: <<http://opencv.willowgarage.com/>>.

BRESENHAM, J. A linear algorithm for incremental digital display of circular arcs. **Commun. ACM**, 1977. ACM, New York, NY, USA, v. 20, n. 2, p. 100–106, fev. 1977. ISSN 0001-0782. Disponível em: <<http://doi.acm.org/10.1145/359423.359432>>.

BROWN, D. C. Close-range camera calibration. **PHOTOGRAMMETRIC ENGINEERING**, 1971. v. 37, n. 8, p. 855–866, 1971.

CASSE, R. **Projective Geometry An Introduction**. [S.l.]: Oxford, 2006. ISBN 0–19–929885–8.

CHANG, L.-W.; CHEN, H.-W.; HO, J.-R. Reconstruction of 3D medical images: A nonlinear interpolation technique for reconstruction of 3D medical images. **CVGIP: Graphical Models and Image Processing**, 1991. v. 53, n. 4, p. 382–391, jul. 1991. ISSN 10499652. Disponível em: <<http://www.sciencedirect.com/science/article/pii/104996529190041H> <http://linkinghub.elsevier.com/retrieve/pii/104996529190041H>>.

CHATTERJEE, C.; ROYCHOWDHURY, V. P. Algorithms for coplanar camera calibration. **Machine Vision and Applications**, 2000. v. 12, n. 2, p. 84–97, ago. 2000. ISSN 0932-8092. Disponível em: <<http://link.springer.com/10.1007/s001380050127>>.

CHEN, C.; FREEDMAN, D.; LAMPERT, C. H. Enforcing topological constraints in random field image segmentation. In: **CVPR**. IEEE, 2011. p. 2089–2096. Disponível em: <<http://dblp.uni-trier.de/db/conf/cvpr/cvpr2011.htmlChenFL11>>.

CHEN, D.; ZHANG, G. A new sub-pixel detector for x-corners in camera calibration targets. In: **13th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision**. Citeseer, 2005. Disponível em: <<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.87-5833&rep=rep1&type=pdf>>.

CHEN, J.; BENZEROUAL, K.; ALLISON, R. S. Calibration for high-definition camera rigs with marker chessboard. **2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops**, 2012. Ieee, p. 29–36, jun. 2012. Disponível em: <<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6238905>>.

CHIABRANDO, F. et al. Sensors for 3D Imaging: Metric Evaluation and Calibration of a CCD/CMOS Time-of-Flight Camera. **Sensors (Basel, Switzerland)**, 2009. v. 9, n. 12, p. 10080–96, jan. 2009. ISSN 1424-8220.

CHIN, R. T.; HARLOW, C. A. Automated Visual Inspection: A Survey. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, 1982. PAMI-4, n. 6, p. 557–573, nov. 1982. ISSN 0162-8828.

CHOI, S. et al. What does ground tell us? monocular visual odometry under planar motion constraint. In: **Control, Automation and Systems (ICCAS), 2011 11th International Conference on**. [S.l.: s.n.], 2011. p. 1480–1485. ISSN 2093-7121.

CHRYSOSTOMOU, D.; KYRIAKOULIS, N.; GASTERATOS, A. Multi-camera 3d scene reconstruction from vanishing points. In: **Imaging Systems and Techniques (IST), 2010 IEEE International Conference on**. [S.l.: s.n.], 2010. p. 343–348.

CONRAD, D.; DESOUSA, G. Homography-based ground plane detection for mobile robot navigation using a modified em algorithm. In: **Robotics and Automation (ICRA), 2010 IEEE International Conference on**. [S.l.: s.n.], 2010. p. 910–915. ISSN 1050-4729.

CROW, F. C. Summed-area tables for texture mapping. **SIGGRAPH Comput. Graph.**, 1984. ACM, New York, NY, USA, v. 18, n. 3, p. 207–212, jan. 1984. ISSN 0097-8930. Disponível em: <<http://doi.acm.org/10.1145/964965.808600>>.

DAO, V. N.; SUGIMOTO, M. A Robust Recognition Technique for Dense Checkerboard Patterns. **2010 20th International Conference on Pattern Recognition**, 2010. Ieee, p. 3081–3084, ago. 2010. Disponível em: <<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5597291>>.

DATTA, A.; KIM, J.-S.; KANADE, T. Accurate camera calibration using iterative refinement of control points. In: **2009 IEEE 12th International Conference on Computer Vision Workshops, ICCV Workshops**. IEEE, 2009. p. 1201–1208. ISBN 978-1-4244-4442-7. Disponível em: <http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5457474>.

DAVIES, E. R. **Computer and machine vision theory, algorithms, practicalities**. 4th. ed. Waltham, Mass.: Academic Press, 2012. ISBN 0123869080, 9780123869081. Disponível em: <http://www.amazon.com/Computer-Machine-Vision-Fourth-Edition-/dp/0123869080/ref=dp_ob_title_bk>.

DELINGETTE, H.; MONTAGNAT, J. Shape and topology constraints on parametric active contours. **Computer Vision and Image Understanding**, 2000. v. 83, p. 140–171, 2000.

_____. Shape and topology constraints on parametric active contours. **Computer Vision and Image Understanding**, 2001. v. 83, n. 2, p. 140 – 171, 2001. ISSN 1077-3142. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1077314201909208>>.

DESOUZA, G.; KAK, A. Vision for mobile robot navigation: a survey. **Pattern Analysis and Machine Intelligence, IEEE Transactions on**, 2002. v. 24, n. 2, p. 237–267, 2002. ISSN 0162-8828.

DONG, J.; ZHOU, M. Medical image segmentation based on topology correlation of regions. In: **BMEI (2)**. IEEE Computer Society, 2008. p. 142–146. Disponível em: <<http://dblp.uni-trier.de/db/conf/bmei/bmei2008-2.htmlDongZ08>>.

DUDA, R. O.; HART, P. E. Use of the Hough transformation to detect lines and curves in pictures. **Communications of the ACM**, 1972. v. 15, n. April 1971, p. 11–15, 1972. Disponível em: <<http://dl.acm.org/citation.cfm?id=361242>>.

EGGAR, M. H. Pinhole cameras, perspective, and projective geometry. **The American Mathematical Monthly**, 1998. Mathematical Association of America, v. 105, n. 7, p. pp. 618–630, 1998. ISSN 00029890. Disponível em: <<http://www.jstor.org/stable/2589244>>.

ESCALERA, A. la; ARMINGOL, J. M. Detección Automática de un Patrón para la Calibración de Cámaras. **Revista Iberoamericana de Automática e Informática Industrial (RIAI)**, 2010. v. 7, n. 4, p. 83–94, 2010. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1697791210700637> [http://dx.doi.org/10.1016/S1697-7912\(10\)70063-7](http://dx.doi.org/10.1016/S1697-7912(10)70063-7)>.

FIALA, M. Vision guided control of multiple robots. In: **First Canadian Conference on Computer and Robot Vision, 2004. Proceedings**. Ieee, 2004. p. 241–246. ISBN 0-7695-2127-4. Disponível em: <<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1301450>>.

FIALA, M.; SHU, C. Self-identifying patterns for plane-based camera calibration. **Machine Vision and Applications**, 2007. v. 19, n. 4, p. 209–216, ago. 2007. ISSN 0932-8092. Disponível em: <<http://link.springer.com/10.1007/s00138-007-0093-z>>.

FLORCZYK, S. **Robot Vision: Video-Based Indoor Exploration with Autonomous and Mobile Robots**. [S.l.]: Wiley-VCH, 2005. ISBN 978-3-527-40544-2.

FOLEY, J. D. et al. **Computer graphics: principles and practice (2nd ed.)**. Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc., 1990. ISBN 0-201-12110-7.

FORBES, K.; VOIGT, A.; BODIKA, N. An inexpensive, automatic and accurate camera calibration method. In **Proceedings of the Thirteenth Annual South African Workshop on Pattern Recognition (PRASA)**, 2002. 2002. Disponível em: <<http://www.prasa.org/proceedings/2002/prasa02-01.pdf>>.

FORMAN, C. et al. Self-encoded marker for optical prospective head motion correction in MRI. **Medical image analysis**, 2011. Elsevier B.V., v. 15, n. 5, p. 708–19, out. 2011. ISSN 1361-8423. Disponível em: <<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3164440&tool=pmcentrez&rendertype=abstract>>.

FORSYTH, D. A.; PONCE, J. **Computer Vision: A Modern Approach**. [S.l.]: Prentice Hall, 2002. 599–619 p. ISSN 18727786. ISBN 0130851981.

GONZALEZ, R. C.; WOODS, R. E. **Digital Image Processing (3rd Edition)**. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 2006. ISBN 013168728X.

GUIBAS, L.; KNUTH, D.; SHARIR, M. Randomized incremental construction of delaunay and voronoi diagrams. **Algorithmica**, 1992. Springer New York, v. 7, p. 381–413, 1992. ISSN 0178-4617. 10.1007/BF01758770. Disponível em: <<http://dx.doi.org/10.1007/BF01758770>>.

HABIB, A.; MORGAN, M. Stability analysis and geometric calibration of off-the-shelf digital cameras. **Photogrammetric Engineering & Remote Sensing**, 2005. v. 71, n. 6, p. 733–741, 2005. Disponível em: <http://www.asprs.org/a/publications/pers-2005journal/june/2005_june_733-741.pdf>.

HARRIS, C.; STEPHENS, M. A Combined Corner and Edge Detector. In: **Proceedings of the Alvey Vision Conference 1988**. Alvey Vision Club, 1988. p. 147–151. Disponível em: <<http://www.bmva.org/bmvc/1988/avc-88-023.html>>.

HARTLEY, R. I.; ZISSERMAN, A. **Multiple View Geometry in Computer Vision**. Second. [S.l.]: Cambridge University Press, ISBN: 0521540518, 2004.

HARTSHORNE, R. **Geometry: Euclid and Beyond**. [S.l.]: Springer, 2000. ISBN 0-387-98650-2.

HASHIMOTO, K. A flexible visual inspection system combining pose estimation and visual servo approaches. In: **2012 IEEE International Conference on Robotics and Automation**. [S.l.]: IEEE, 2012. p. 1304–1309. ISBN 978-1-4673-1405-3.

HAZEWINKEL, M. **Topology, general**. **Encyclopedia of Mathematics**. [S.l.]: Springer, 2001. ISBN 978-1-55608-010-4.

HE, L. et al. Fast connected-component labeling. **Pattern Recognition**, 2009. v. 42, n. 9, p. 1977 – 1987, 2009. ISSN 0031-3203. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0031320308004573>>.

HEIKKILÄ, J. Moment and curvature preserving technique for accurate ellipse boundary detection. In: **Proceedings of 14th International Conference on Pattern Recognition, 1998**. Brisbane, Qld.: [s.n.], 1998. v. 1, p. 734–737. ISBN 0-8186-8512-3. ISSN 1051-4651. Disponível em: <http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=711250>.

_____. Geometric camera calibration using circular control points. **Pattern Analysis and Machine Intelligence, IEEE Transactions on**, 2000. v. 22, n. 10, p. 1066–1077, 2000. ISSN 0162-8828.

HEIKKILÄ, J.; SILVÉN, O. A four-step camera calibration procedure with implicit image correction. In: **CVPR. IEEE Computer Society**, 1997. p. 1106–1112. Disponível em: <<http://dblp.uni-trier.de/db/conf/cvpr/cvpr1997.html#HeikkilaS97>>.

HEMAYED, E. A survey of camera self-calibration. **Proceedings of the IEEE Conference on Advanced Video and Signal Based Surveillance, 2003.**, 2003. IEEE Comput. Soc, p. 351–357, july 2003. Disponível em: <<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1217942>>.

HUBEL, D. H. **Eye, Brain, and Vision (Scientific American Library, No 22)**. 2nd. ed. W. H. Freeman, 1995. Paperback. ISBN 0716760096. Disponível em: <<http://www.amazon.com/exec/obidos/redirect?tag=citeulike07-20&path=ASIN/0716760096>>.

HUGHES, C.; GLAVIN, M.; JONES, E. Simple fish-eye calibration method with accuracy evaluation. **Electronic Letters on Computer Vision and Image Analysis**, 2011. v. 10, n. 1, p. 54–62, 2011. Disponível em: <<http://elcvia.cvc.uab.es/article/viewFile/396/245>>.

KANG, D.-J.; HA, J.-E.; JEONG, M.-H. Detection of calibration patterns for camera calibration with irregular lighting and complicated backgrounds. **International Journal of Control, Automation, and Systems**, 2008. v. 6, n. 5, p. 746–754, 2008.

KANNALA, J.; BRANDT, S. S. A generic camera model and calibration method for conventional, wide-angle, and fish-eye lenses. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, 2006. IEEE, v. 28, n. 8, p. 1335–1340, 2006. Disponível em: <http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1642666>.

KARABEGOVIĆ, I.; VOJIC, S.; DOLECEK, V. 3d vision in industrial robot working process. In: **Power Electronics and Motion Control Conference, 2006. EPE-PEMC 2006. 12th International**. [S.l.: s.n.], 2006. p. 1223–1226.

KATO, H.; BILLINGHURST, M. Marker tracking and hmd calibration for a video-based augmented reality conferencing system. In: **Proceedings of the 2nd IEEE and ACM International Workshop on Augmented Reality**. Washington, DC, USA: IEEE Computer Society, 1999. (IWAR '99), p. 85–. ISBN 0-7695-0359-4. Disponível em: <<http://dl.acm.org/citation.cfm?id=857202.858134>>.

KLAUS, A. et al. Camera calibration from a single night sky image. In: **Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on**. [S.l.: s.n.], 2004. v. 1, p. I–151–I–157 Vol.1. ISSN 1063-6919.

KLEIN, G.; MURRAY, D. Parallel tracking and mapping for small AR workspaces. In: **Proc. Sixth IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR'07)**. Nara, Japan: [s.n.], 2007.

KRÜGER, L.; WÖHLER, C. Accurate chequerboard corner localisation for camera calibration. **Pattern Recognition Letters**, 2011. v. 32, n. 10, p. 1428–1435, jul.

2011. ISSN 01678655. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0167865511001024>>.

LICHTI, D. D.; KIM, C. A Comparison of Three Geometric Self-Calibration Methods for Range Cameras. **Remote Sensing**, 2011. v. 3, n. 12, p. 1014–1028, maio 2011. ISSN 2072-4292. Disponível em: <<http://www.mdpi.com/2072-4292/3/5/1014/>>.

LOAIZA, M. E.; RAPOSO, A. B.; GATTASS, M. Multi-camera calibration based on an invariant pattern. **Computers and Graphics**, 2011. v. 35, n. 2, p. 198 – 207, 2011. ISSN 0097-8493. Virtual Reality in Brazil, Visual Computing in Biology and Medicine, Semantic 3D media and content, Cultural Heritage. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0097849310002050>>.

MALLON, J.; WHELAN, P. F. Which pattern? Biasing aspects of planar calibration patterns and detection methods. **Pattern Recognition Letters**, 2007. v. 28, n. 8, p. 921–930, jun. 2007. ISSN 01678655. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0167865506003114>>.

MATEOS, G. G. A camera calibration technique using targets of circular features. In: **Proceedings of V IberoAmerican Symposium on Pattern Recognition SIARP**. Lisbon, Portugal: [s.n.], 2000.

MOLTON, N. et al. A stereo vision-based aid for the visually impaired. **Image and Vision Computing**, 1998. v. 16, n. 4, p. 251 – 263, 1998. ISSN 0262-8856. Vision-Based Aids for the Disabled. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0262885697000875>>.

NIXON, M. S.; AGUADO, A. S. **Feature Extraction and Image Processing**. Second edition. [S.l.]: Academic Press ISBN: 978-0-12-372538-7, 2008.

OHNISHI, N.; IMIYA, A. Dominant plane detection from optical flow for robot navigation. **Pattern Recognition Letters**, 2006. v. 27, n. 9, p. 1009 – 1021, 2006. ISSN 0167-8655. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0167865505003703>>.

OREN, M.; NAYAR, S. K. Generalization of Lambert's reflectance model. **Proceedings of the 21st annual conference on Computer graphics and interactive techniques - SIGGRAPH '94**, 1994. ACM Press, New York, New York, USA, p. 239–246, 1994. Disponível em: <<http://portal.acm.org/citation.cfm?doid=192161.192213>>.

O'ROURKE, J. **Computational Geometry in C**. 2nd. ed. New York, NY, USA: Cambridge University Press, 1998. ISBN 0521640105.

ORTEGA, H. **JCamCalib: A Camera Calibration Utility**. 2005. Disponível em: <<http://sourceforge.net/projects/jcamcalib/>>.

OWEN, C. B. C. et al. What is the best fiducial? In: **The First IEEE International Augmented Reality Toolkit Workshop**. Darmstadt, Germany: IEEE, 2002. p. 98—105. ISBN 0-7803-7680-3. Disponível em: <http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1107021 <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1107021>>.

POKORNY, F. T. et al. Topological constraints and kernel-based density estimation. In: . [S.l.]: Workshop on Algebraic Topology and Machine Learning, Advances in Neural Information Processing Systems (NIPS), 2012. p. 1–5.

PRESS, W. H. et al. **Numerical recipes in C (2nd ed.): the art of scientific computing**. New York, NY, USA: Cambridge University Press, 1992. ISBN 0-521-43108-5.

PULLIVELLI, A. **Low-Cost Digital Cameras: Calibration, Stability Analysis, and Applications**. Tese (Master of Science) — University of Calgary, 2005. Disponível em: <<http://www.geomatics.ucalgary.ca/links/GradTheses.html>>.

QI, W.; LI, F.; ZHENZHONG, L. Review on camera calibration. In: **2010 Chinese Control and Decision Conference**. [S.l.]: IEEE, 2010. p. 3354–3358. ISBN 978-1-4244-5181-4.

RENCER, A. C. **Methods of Multivariate Analysis**. Second edition. [S.l.]: A JOHN WILEY & SONS, INC. PUBLICATION, 2002. Paperback. ISBN 9780470178966.

ROSTEN, E.; DRUMMOND, T. Machine learning for high-speed corner detection. In: **Proceedings of the 9th European conference on Computer Vision - Volume Part I**. Berlin, Heidelberg: Springer-Verlag, 2006. (ECCV'06), p. 430–443. ISBN 3-540-33832-2, 978-3-540-33832-1. Disponível em: <http://dx.doi.org/10.1007/11744023_34>.

ROSTEN, E.; LOVELAND, R. Camera distortion self-calibration using the plumb-line constraint and minimal hough entropy. **Machine Vision and Applications**, 2011. Springer-Verlag, v. 22, n. 1, p. 77–85, 2011. ISSN 0932-8092. Disponível em: <<http://dx.doi.org/10.1007/s00138-009-0196-9>>.

RUFLI, M.; SCARAMUZZA, D.; SIEGWART, R. Automatic detection of checkerboards on blurred and distorted images. **2008 IEEE/RSJ International Conference on Intelligent Robots and Systems**, 2008. Ieee, p. 3121–3126, set. 2008. Disponível em: <<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4650703>>.

SAKAS, G. Trends in medical imaging: from 2D to 3D. **Computers & Graphics**, 2002. Elsevier, v. 26, n. 4, p. 577–587, ago. 2002. ISSN 00978493. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0097849302001036>>.

SALVI, J.; ARMANGUÉ, X.; BATLLE, J. A comparative review of camera calibrating methods with accuracy evaluation. **Pattern recognition**, 2002. v. 35, n. 7, p. 1617–1635, 2002. ISSN 0031-3203. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0031320301001261>>.

SAMPER, D. et al. Analysis of tsai calibration method using two- and three-dimensional calibration objects. **Machine Vision and Applications**, 2013. Springer-Verlag, v. 24, n. 1, p. 117–131, 2013. ISSN 0932-8092. Disponível em: <<http://dx.doi.org/10.1007/s00138-011-0398-9>>.

SCARAMUZZA, D. **OCamCalib: Omnidirectional Camera Calibration Toolbox for Matlab**. 2013. Disponível em: <<https://sites.google.com/site/scarobotix/ocamcalib-toolbox>>.

- SCARAMUZZA, D.; MARTINELLI, A.; SIEGWART, R. A Toolbox for Easily Calibrating Omnidirectional Cameras. **2006 IEEE/RSJ International Conference on Intelligent Robots and Systems**, 2006. Ieee, p. 5695–5701, out. 2006. Disponível em: <<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4059340>>.
- SCHAULIN, M. 3D measuring methods for visual inspection of advanced packages. In: **26th International Spring Seminar on Electronics Technology: Integrated Management of Electronic Materials Production, 2003**. IEEE, 2003. p. 427–432. ISBN 0-7803-8002-9. Disponível em:
- SCHLICK, C. A Survey of Shading and Reflectance Models. **Computer Graphics Forum**, 1994. v. 13, n. 2, p. 121–131, maio 1994. ISSN 0167-7055. Disponível em: <<http://doi.wiley.com/10.1111/1467-8659.1320121>>.
- SCHMID, C.; MOHR, R.; BAUCKHAGE, C. Evaluation of interest point detectors. **International Journal of computer vision**, 2000. v. 37, n. 2, p. 151–172, 2000. Disponível em: <<http://link.springer.com/article/10.1023/A:1008199403446>>.
- SEITZ, S. et al. A comparison and evaluation of multi-view stereo reconstruction algorithms. In: **Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on**. [S.l.: s.n.], 2006. v. 1, p. 519–528. ISSN 1063-6919.
- SHI, J.; TOMASI, C. Good features to track. In: **Proceedings of IEEE Conference on Computer Vision and Pattern Recognition CVPR-94**. IEEE Comput. Soc. Press, 1994. p. 593–600. ISBN 0-8186-5825-8. Disponível em: <http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=323794 <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=323794>>.
- SHIODE, N. 3d urban models: Recent developments in the digital modelling of urban environments in three-dimensions. **GeoJournal**, 2000. Springer Netherlands, v. 52, p. 263–269, 2000. ISSN 0343-2521. 10.1023/A:1014276309416. Disponível em: <<http://dx.doi.org/10.1023/A:1014276309416>>.
- SHORTIS, M. R.; CLARKE, T. A.; SHORT, T. A comparison of some techniques for the subpixel location of discrete target images. In: **Procedings of SPIE 2350, Videometrics III, 239**. [S.l.: s.n.], 1994. v. 2350.
- SHU, C.; BRUNTON, A.; FIALA, M. A topological approach to finding grids in calibration patterns. **Machine Vision and Applications**, 2009. v. 21, n. 6, p. 949–957, out. 2009. ISSN 0932-8092. Disponível em: <<http://link.springer.com/10.1007/s00138-009-0202-2>>.
- SKOOG, D. A.; HOLLER, J. F.; CROUCH, S. R. **Principles of Instrumental Analysis**. 6. ed. Brooks Cole, 2006. Hardcover. ISBN 0495012017. Disponível em: <<http://www.amazon.com/exec/obidos/redirect?tag=citeulike07-20&path=ASIN-/0495012017>>.
- SLAMA, C. C. **Manual of Photogrammetry**. 4. ed. [S.l.]: American Society of Photogrammetry, 1980.
- SOH, L. Robust recognition of calibration charts. **6th International Conference on Image Processing and its Applications**, 1997. Iee, v. 1997, p. 487–491,

1997. Disponível em: <http://digital-library.theiet.org/content/conferences/10.1049/cp/_19970941>.

SONKA, M.; HLAVAC, V.; BOYLE, R. **Image Processing, Analysis, and Machine Vision**. [S.l.]: Thomson-Engineering, 2007. ISBN 049508252X.

STOYANOV, D. **Camera Calibration Tools**. 2011. Disponível em: <<http://www0.cs.ucl.ac.uk/staff/Dan.Stoyanov/calib/>>.

STRANG, G. **Introduction to Linear Algebra**. [S.l.]: Wellesley-Cambridge Press, 1993.

STROBL, K. H.; HIRZINGER, G. More accurate pinhole camera calibration with imperfect planar target. In: **2011 IEEE International Conference on Computer Vision Workshops (ICCV Workshops)**. IEEE, 2011. p. 1068–1075. ISBN 978-1-4673-0063-6. Disponível em: <<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6130369>>.

SUN, W. et al. Robust checkerboard recognition for efficient nonplanar geometry registration in projector-camera systems. In: **Proceedings of the 5th ACM/IEEE International Workshop on Projector camera systems**. New York, NY, USA: ACM, 2008. (PROCAMS '08), p. 2:1–2:7. ISBN 978-1-60558-272-6. Disponível em: <<http://doi.acm.org/10.1145/1394622.1394625>>.

SZELISKI, R. **Computer Vision: Algorithms and Applications**. 1st. ed. New York, NY, USA: Springer-Verlag New York, Inc., 2010. ISBN 1848829345, 9781848829343.

TAVAKOLI, H.; POURREZA, H.-R. An automated camera calibration framework for desktop vision systems. In: **Advances in Computational Tools for Engineering Applications, 2009. ACTEA '09. International Conference on**. [S.l.: s.n.], 2009. p. 96–100.

TRUCCO, E.; VERRI, A. Introductory Techniques for 3D Computer Vision. In: _____. Upper Saddle, NJ: Prentice-Hall, Inc., 1998. cap. Dealing with Imaging Noise, p. 56–62.

TSAI, R. Y. An efficient and accurate camera calibration technique for 3d machine vision. In: **Proceedings of IEEE Conference on Computer Vision and Pattern Recognition**. [S.l.: s.n.], 1986. p. 364 – 374.

VELIZHEV, A. **GML C++ Camera Calibration Toolbox**. 2011. Disponível em: <<http://graphics.cs.msu.ru/en/node/909>>.

VIOLA, P.; JONES, M. Rapid object detection using a boosted cascade of simple features. In: **Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on**. [S.l.: s.n.], 2001. v. 1, p. I–511–I–518 vol.1. ISSN 1063-6919.

WANG, X. Intelligent multi-camera video surveillance: A review. **Pattern Recognition Letters**, 2013. v. 34, n. 1, p. 3–19, jan. 2013. ISSN 01678655. Disponível em: <<http://dx.doi.org/10.1016/j.patrec.2012.07.005>>.

- WANG, Z.; WU, Y. Recognition of corners of planar pattern image. (**WCICA**) **2010 8th World Congress on Intelligent Control and Automation**, 2010. n. 2008466, p. 6342–6346, 2010. Disponível em: <http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5554349>.
- WELLNER, P. D. **Adaptive Thresholding for the DigitalDesk**. [S.l.], 1993.
- WILBURN, B. et al. High-speed videography using a dense camera array. In: **Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004**. IEEE, 2004. v. 2, p. 294–301. ISBN 0-7695-2158-4. Disponível em: <<http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=1315176>>.
- WILLIAMS, B.; REID, I. On combining visual slam and visual odometry. In: **Robotics and Automation (ICRA), 2010 IEEE International Conference on**. [S.l.: s.n.], 2010. p. 3494–3500. ISSN 1050-4729.
- XIAO, Y.; FISHER, R. Accurate Feature Extraction and Control Point Correction for Camera Calibration with a Mono-Plane Target. **Procedings of International Conference on 3D Data Processing, Visualization and Transmission 2010**, 2010. 2010. Disponível em: <http://homepages.inf.ed.ac.uk/rbf/CHIROPING/CAMCAL/camcal_files/3DPVT10_final_v3.pdf>.
- XIAO, Z. et al. A cross-target-based accurate calibration method of binocular stereo systems with large-scale field-of-view. **Measurement**, 2010. Elsevier Ltd, v. 43, n. 6, p. 747–754, jul. 2010. ISSN 02632241. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0263224110000308>>.
- YIMIN, L. et al. A novel approach to sub-pixel corner detection of the grid in camera calibration. In: **Computer Application and System Modeling (ICCSM), 2010 International Conference on**. [S.l.: s.n.], 2010. v. 5, p. V5–18–V5–21.
- YOUNG, R.; MATAS, J.; KITTLER, J. On camera calibration for scene model acquisition and maintenance using an active vision system. In: **Computer Vision Systems**. Springer Berlin Heidelberg, 1999, (Lecture Notes in Computer Science, v. 1542). p. 509–521. ISBN 978-3-540-65459-9. Disponível em: <http://dx.doi.org/10.1007/3-540-49256-9_30>.
- ZHANG, X. et al. Camera calibration based on 3D-point-grid. **Image Analysis and Processing**, 1997. Springer Berlin Heidelberg, v. 1310, p. 636–643, 1997. Disponível em: <http://link.springer.com/chapter/10.1007/3-540-63507-6_255> http://dx.doi.org/10.1007/3-540-63507-6_255.
- ZHANG, Z. A flexible new technique for camera calibration. **Pattern Analysis and Machine Intelligence, IEEE Transactions on**, 2000. IEEE, v. 22, n. 11, p. 1330–1334, 2000. ISSN 01628828. Disponível em: <<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=888718> http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=888718>.
- _____. Camera calibration - Emerging Topics in Computer Vision. In: . [S.l.]: Prentice Hall Professional Technical Reference, 2004. p. 4–43.

_____. Camera calibration with one-dimensional objects. **Pattern Analysis and Machine Intelligence, IEEE ...**, 2004. v. 26, n. 7, p. 892–899, 2004. Disponível em: <http://ieeexplore.ieee.org/xpls/abs/_all.jsp?arnumber=1300559>.

ZHAO, F. et al. An automated x-corner detection algorithm (axda). **Journal of Software**, 2011. v. 6, n. 5, p. 791–797, 2011. Disponível em: <<https://academypublisher.com/~academz3/ojs/index.php/jsw/article/view/0605791797>>.

ZHENG, Z. et al. Feature Detection and Correspondence for Camera Calibration. **International Journal of Information Acquisition**, 2008. v. 05, n. 01, p. 41–50, mar. 2008. ISSN 0219-8789. Disponível em: <<http://www.worldscientific.com/doi/abs/10.1142/S021987890800148X>>.

ZHOU, F.; ZHANG, W. Accurate Detection and Automatic Match Method for Feature Points Using Concentric Circles Array. In: FANG, J.; WANG, Z. (Ed.). [s.n.], 2008. v. 7129, n. 37, p. 71290C–71290C–6. Disponível em: <<http://spiedigitallibrary.org/proceeding.aspx?doi=10.1117/12.807409>>.