



Pós-Graduação em Ciência da Computação

RODRIGO JOSÉ BATISTA COSTA SANTOS

**TÉCNICA DE CALIBRAÇÃO BASEADA NA ADAPTAÇÃO
DOS PARÂMETROS DE CÂMERA PARA AMBIENTES
COM DIFERENTES ÍNDICES DE REFRAÇÃO**



Universidade Federal de Pernambuco
posgraduacao@cin.ufpe.br
www.cin.ufpe.br/~posgraduacao

Recife
2018

Rodrigo José Batista Costa Santos

**Técnica de calibração baseada na adaptação dos parâmetros de câmera
para ambientes com diferentes índices de refração**

Este trabalho foi apresentado à Pós-Graduação em Ciência da Computação do Centro de Informática da Universidade Federal de Pernambuco como requisito parcial para obtenção do grau de Mestre Profissional em Ciência da Computação.

Orientadora: Prof^a Dra. Judith Kelner

Recife
2018

Catálogo na fonte
Bibliotecária Monick Raquel Silvestre da S. Portes, CRB4-1217

S237t Santos, Rodrigo José Batista Costa
Técnica de calibração baseada na adaptação dos parâmetros de câmera para ambientes com diferentes índices de refração / Rodrigo José Batista Costa Santos. – 2018.
117 f.: il., fig., tab.

Orientadora: Judith Kelner.
Dissertação (Mestrado) – Universidade Federal de Pernambuco. CIn, Ciência da Computação, Recife, 2018.
Inclui referências e apêndices.

1. Inteligência artificial. 2. Visão computacional. I. Kelner, Judith (orientadora). II. Título.

006.3

CDD (23. ed.)

UFPE- MEI 2019-012

Rodrigo José Batista Costa Santos

**Técnica de calibração baseada na adaptação dos parâmetros de câmera
para ambientes com diferentes índices de refração**

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Ciência da Computação da Universidade Federal de Pernambuco, como requisito parcial para a obtenção do título de Mestre em Ciência da Computação

Aprovado em: 23/02/2018.

BANCA EXAMINADORA

Prof. Dr. Hansenclever de França Bassani
Centro de Informática/UFPE

Prof. Dr. Carmelo José Albanez Bastos Filho
Escola Politécnica de Pernambuco

Profa. Dra. Judith Kelner
Centro de Informática / UFPE
(Orientadora)

*Aos meus pais, irmã, minha companheira Naiara
e a toda minha família e amigos que, com muito carinho e apoio,
não mediram esforços para que eu chegasse até esse etapa de minha vida.*

AGRADECIMENTOS

Agradeço primeiramente a Deus, por ter me dado todas as oportunidades que sempre tive e por colocar no meu caminho pessoas importantes para o meu crescimento pessoal e profissional. Agradeço aos meus pais, por sempre acreditarem em mim, me apoiando em todos os meus sonhos, aconselhando e apresentando ótimos exemplos. Assim como os meus demais familiares sempre presentes ao longo de minha jornada. Agradeço também a minha irmã, pela sua paciência em todos esses anos.

Um agradecimento mais do que especial a minha companheira Naiara pelo seu cuidado, compreensão, suporte, dedicação, carinho e paciência que demonstrou durante esse período de mestrado. Não posso deixar de agradecer principalmente a todos os auxílios, revisões e motivações que você concedeu nessa escrita, sem você não conseguiria passar por esse processo deste modo. Obrigado por tudo! Aos meus sogros, cunhado e cunhada por me fazerem sentir muito bem acolhido em sua família, proporcionando-me momentos de descontração e felicidade.

Agradeço aos amigos que fiz durante o período universitário, na UNICAP e na UFPE, pois eles tornaram o curso mais leve e o tempo mais curto, além de contribuírem diretamente no meu conhecimento adquirido ao longo desses períodos em discussões, estudos e trabalhos realizados em grupo. Também não poderia esquecer os amigos que fizeram parte da minha caminhada até chegar a faculdade, eles podem não saber, mas foram importantes para meu crescimento.

Não poderia deixar de agradecer a minha orientadora, Professora Dra. Judith, por ter me dado a oportunidade de ingressar no campo da pesquisa, além da sua confiança e suporte no meu trabalho. Também não posso deixar de agradecer aos membros do GRVM que colaboraram, de alguma forma, com o meu desenvolvimento profissional nesses últimos seis anos. Por fim, meus sinceros agradecimentos as entidades FACEPE e Petrobras que contribuíram financeiramente e estruturalmente para o desenvolvimento desta pesquisa.

RESUMO

Em visão computacional, aplicações que realizam extração métrica através de imagens obtidas por câmeras possuem como etapa crucial o procedimento de calibração, que consiste em determinar os parâmetros que descrevem o modelo matemático capaz de representar o processo da formação de imagem. Esse procedimento é uma das primeiras etapas a serem realizadas na maioria dessas aplicações e qualquer imprecisão ou erro cometido durante a sua realização é propagado para as demais etapas, afetando o resultado final obtido pelo sistema. Em aplicações subaquáticas, o procedimento de calibração de câmera tem ainda maior importância, devido aos efeitos de refração, existentes entre os diferentes meios, que devem ser incorporados no modelo. Na literatura, os trabalhos que calculam os parâmetros de câmera para o ambiente subaquático podem ser divididos em duas abordagens: técnicas que consideram a formação de imagem sendo realizada através de um único ponto de vista (SVP), utilizando-se a calibração no ambiente subaquático ou adaptando dos parâmetros calculados antes da submersão; e as técnicas que conceituam uma câmera subaquática sendo composta por múltiplos pontos de vista (nSVP). Essas abordagens diferem-se pela relação no custo-benefício entre a acurácia e complexidade da estimativa dos parâmetros que descrevem a câmera. Enquanto o modelo SVP diminui a complexidade dos algoritmos que o utilizam, o nSVP aumenta a quantidade de parâmetros, explicitando o efeito de refração, com o objetivo de incrementar a acurácia da descrição. Esta dissertação apresenta uma nova técnica para calcular os parâmetros de uma câmera subaquática através da adaptação do resultado obtido por uma calibração realizada antes de introduzi-la na água. Diferentemente dos trabalhos encontrados na literatura, a técnica proposta realiza a adaptação da calibração explicitando a aproximação para o efeito de distorção introduzido pela refração existente na formação de imagens por câmeras subaquáticas. Para tal, foi projetado um modelo genérico que incorpora a distorção da refração ocasionada pela mudança de meio em que a câmera foi calibrada para onde será utilizada. Experimentos realizados em cenários sintéticos e reais permitiram uma análise da correção e robustez da adaptação proposta, apresentando erros de reconstrução menores do que a técnica de adaptação da literatura, principalmente, quando o objeto alvo encontrou-se nas bordas das imagens.

Palavras-chave: Visão computacional subaquática. Modelo de câmera. Calibração de câmera. Adaptação de calibração.

ABSTRACT

Computer vision applications that perform metrical extraction through images obtained by cameras usually have, as a critical step, the calibration procedure. It consists in determining the parameters that describe the mathematical model capable of representing the process of image formation of that camera. This procedure is one of the first steps to be performed in most applications and any imprecision or error is propagated to the other steps, affecting the final result obtained by the system. In underwater applications, the calibration procedure of the camera is even more important, due to the effects of refraction that occur between the different mediums, which should be taken into account in the model. The camera calibration literature for the underwater environments can be divided into two approaches: techniques that consider the imaging being performed from a single viewpoint (SVP), by either calibrating in the underwater environment or adapting the parameters calculated before submersion; and the techniques that consider an underwater camera as being composed of multiple viewpoints (nSVP - non single viewpoint). These approaches differ by the trade-off between accuracy and the complexity of estimating the parameters that describe a camera. While the SVP model reduces the complexity of the algorithms that use it, the nSVP increases the quantity of parameters, better explaining the effects of refraction, in order to increase the accuracy of the description. This dissertation presents a new technique to calibrate an underwater camera by adapting the result obtained by the calibration performed before the submersion. In contrast to the underwater calibration literature, the proposed technique adapts the parameters through modeling the effect of distortion caused by refraction. To achieve this, we describe a generic model which takes refraction effects into account to adapt the camera parameters. Experimental results collected from synthetic and real scenarios allowed us to analyze the correctness and robustness of the proposed technique in comparison to the adaptation method most known in the literature. It achieved better results, in particular when the target object was near to the border of the image.

Keywords: Underwater computer vision. Underwater camera model. Camera calibration. Calibration adaptation.

LISTA DE ILUSTRAÇÕES

| | |
|--|----|
| Figura 1 – Geometria dos raios de uma câmera submersa em água com habitação plana. | 17 |
| Figura 2 – Imagens obtidas por câmeras em aplicações submarinas. | 17 |
| Figura 3 – Diagrama simplificado do campo de visão da câmera em uma caixa de estanque com uma interface plana (a) e interface domo (b). | 18 |
| Figura 4 – Representações de ponto e reta no plano projetivo. | 24 |
| Figura 5 – Modelo de formação de imagens pela câmera. | 25 |
| Figura 6 – Representação do modelo de câmera <i>pinhole</i> | 27 |
| Figura 7 – Representação dos efeitos de distorção em uma imagem. | 28 |
| Figura 8 – Sistema de coordenadas de imagem $[u, v]^T$ e câmera $[x_c, y_c]^T$ | 29 |
| Figura 9 – Exemplo de padrões de calibração 1D, 2D e 3D. | 32 |
| Figura 10 – Ângulos entre o raio de luz e uma interface com mudança de meio. | 36 |
| Figura 11 – Variação do campo de visão obtida por uma câmera no ar e na água. | 37 |
| Figura 12 – Padrões utilizados em calibrações para ambientes subaquáticos. | 40 |
| Figura 13 – Caminhos dos raios de luz incidentes no centro de uma câmera fixada em caixa de estanque com interface de vidro planar. | 42 |
| Figura 14 – Ilustração do método proposto em Agrawal et al. (2012). | 44 |
| Figura 15 – Exemplo do diagrama de execução para uma aplicação subaquática que utiliza o algoritmo CSAT. | 47 |
| Figura 16 – Fluxograma da técnica CSAT proposta para adaptar os parâmetros de câmera para ambientes com diferentes índices de refração. | 48 |
| Figura 17 – Cenário exemplificando uma câmera subaquática visualizando um cubo submerso em um meio diferente ao que está inserida | 49 |
| Figura 18 – Diferença entre a imagem obtida por uma câmera, sem distorção de lente, no meio 1 com $\eta_1 = 1,0$ (a) e no meio 2 com $\eta_2 = 1,333$ (b). | 50 |
| Figura 19 – Representação gráfica das quatro primeiras etapas do algoritmo CSAT. | 56 |
| Figura 20 – Mapa de calor dos valores que representam o fator de distorção, ocasionado pela mudança de meio, em cada ponto amostrado na imagem. | 57 |
| Figura 21 – Fluxograma da comparação entre as técnicas de calibração para o ambiente subaquático. | 61 |
| Figura 22 – Padrões utilizados para calibração (a) e validação nos experimentos (b). | 63 |
| Figura 23 – Cenário elaborado na ferramenta 3DS Max 2018 para a validação das técnicas de calibração. | 64 |
| Figura 24 – Amostragem do movimento (a) e orientação (b) do padrão em um dos cenários de calibração seca. | 66 |

| | |
|---|----|
| Figura 25 – Representação do 3DS Max sobre a distância entre a interface e a câ- mera, variando de 5 mm até 140 mm. | 67 |
| Figura 26 – Representação do 3DS Max sobre o posicionamento (a) e orientação (b) do padrão no cenário 01 com distância para a câmera fixa em 500 mm. | 68 |
| Figura 27 – Exemplos das imagens do conjunto de validação do cenário 01 com objeto fixo a 500 mm. | 69 |
| Figura 28 – Esquema de exemplificação da inserção de distorção causada pela mu- dança de meio. | 70 |
| Figura 29 – Gráfico de erro médio da reconstrução dos pontos do padrão obtido no Cenário 01 com objeto fixo sem 500 mm de distância da câmera. | 71 |
| Figura 30 – Exemplos das imagens do conjunto de validação do cenário 01 com objeto fixo a 1000 mm. | 72 |
| Figura 31 – Gráfico de erro médio da reconstrução dos pontos do padrão obtido no Cenário 01 com objeto fixo em 1000 mm de distância da câmera. | 73 |
| Figura 32 – Representação do 3DS Max sobre a translação do objeto de validação no cenário 02 nas 22 imagens produzidas. | 75 |
| Figura 33 – Exemplos das imagens do conjunto de validação do cenário 02 com objeto deslocado para a borda. | 76 |
| Figura 34 – Gráfico de erro médio da reconstrução dos pontos do padrão obtido no Cenário 02 com objeto deslocado para a borda. | 77 |
| Figura 35 – Exemplos das imagens do conjunto de validação do cenário 02 com objeto fixo no centro da imagem. | 78 |
| Figura 36 – Gráfico de erro médio da reconstrução dos pontos do padrão obtido no Cenário 02 com objeto fixo no centro da imagem. | 79 |
| Figura 37 – Gráfico das amplitudes dos erros médios e máximos obtidos pelas téc- nicas no Cenário 02 com objeto fixo no centro da imagem. | 79 |
| Figura 38 – Representação do 3DS Max sobre a translação do objeto de validação no cenário 03 nas 18 imagens produzidas. | 81 |
| Figura 39 – Exemplos das imagens do conjunto de validação do cenário 03 com objeto deslocado para borda. | 82 |
| Figura 40 – Gráfico de erro médio da reconstrução dos pontos do padrão obtido no Cenário 03 com objeto deslocado para borda. | 83 |
| Figura 41 – Exemplos das imagens do conjunto de validação do cenário 03 com objeto fixo. | 84 |
| Figura 42 – Gráfico de erro médio da reconstrução dos pontos do padrão obtido no Cenário 03 com objeto de fixo. | 85 |
| Figura 43 – Câmera subaquática, com caixa de estanque de interface planar, utili- zada nos experimentos reais desta dissertação. | 86 |

| | |
|---|-----|
| Figura 44 – Subconjunto das imagens utilizadas nas calibrações seca e molhada para a validação no tanque de água. | 87 |
| Figura 45 – Imagens dos vídeos utilizados para a validação no tanque de água. . . . | 88 |
| Figura 46 – Gráfico de erro médio da reconstrução no tanque de água com o primeiro conjunto de calibrações | 89 |
| Figura 47 – Gráfico de erro médio da reconstrução no tanque de água com o segundo conjunto de calibrações | 89 |
| Figura 48 – Geometria epipolar | 91 |
| Figura 49 – Subconjunto das imagens do Vídeo 01 analisado durante o experimento no ambiente submarino. | 92 |
| Figura 50 – Gráfico do erro epipolar obtido no Vídeo 01 do experimento no ambiente submarino, utilizando as técnicas CSMF, CSAT e CSAT-M. . . . | 93 |
| Figura 51 – Subconjunto das imagens do Vídeo 02 analisado durante o experimento no ambiente submarino. | 94 |
| Figura 52 – Gráfico do erro epipolar obtido no Vídeo 02 do experimento no ambiente submarino, utilizando as técnicas CSMF, CSAT e CSAT-M. . . . | 94 |
| Figura 53 – Subconjunto das imagens do Vídeo 03 analisado durante o experimento no ambiente submarino. | 95 |
| Figura 54 – Gráfico do erro epipolar obtido no Vídeo 03 do experimento no ambiente submarino, utilizando as técnicas CSMF, CSAT e CSAT-M. . . . | 96 |
| Figura 55 – Subconjunto das imagens do Vídeo 04 analisado durante o experimento no ambiente submarino. | 97 |
| Figura 56 – Gráfico do erro epipolar obtido no Vídeo 04 do experimento no ambiente submarino, utilizando as técnicas CSMF, CSAT e CSAT-M. . . . | 97 |
| Figura 57 – Gráfico comparativo entre as técnicas de adaptação usando erro epipolar médio, considerando desvio padrão. | 98 |
| Figura 58 – Mapa de calor da posição dos pontos detectados ao longo de cada vídeo analisado para o ambiente submarino. | 100 |

LISTA DE TABELAS

| | |
|---|-----|
| Tabela 1 – Resultado do cenário sintético 01 com objeto de validação fixo à 500 mm da câmera | 111 |
| Tabela 2 – Resultado do cenário sintético 01 com objeto de validação fixo à 1000 mm da câmera | 112 |
| Tabela 3 – Resultado do cenário sintético 02 com objeto de validação na borda da imagem | 113 |
| Tabela 4 – Resultado do cenário sintético 02 com objeto de validação no centro da imagem | 113 |
| Tabela 5 – Resultado do cenário sintético 03 com objeto de validação na borda da imagem | 114 |
| Tabela 6 – Resultado do cenário sintético 03 com objeto de validação com posição fixa | 114 |
| Tabela 7 – Resultado do experimento no tanque de água com o primeiro conjunto de calibração. | 116 |
| Tabela 8 – Resultado do experimento no tanque de água com o segundo conjunto de calibração. | 116 |
| Tabela 9 – Resultados obtidos nos experimentos realizados no ambiente submarino. | 117 |

LISTA DE ABREVIATURAS E SIGLAS

| | |
|--------|--|
| ANN | <i>Approximate Nearest Neighbor</i> |
| AUVs | <i>Autonomous Underwater Vehicles</i> |
| BIAS | <i>Basic Image Algorithms</i> |
| CM | Calibração Molhada |
| CMA-ES | <i>Covariance Matrix Adaptation Evolution-Strategy</i> |
| CSMF | Calibração Seca com Multiplicação do Foco |
| DLT | <i>Direct Linear Transformation</i> |
| FOV | <i>Field of View</i> |
| IR | Índice de refração |
| KLT | Kanade Lucas Tomasi |
| MATLAB | <i>Matrix Laboratory</i> |
| nSVP | <i>Non Single Viewpoint</i> |
| OpenCV | <i>Open Source Computer Vision</i> |
| RAC | <i>Radial Alignment Constraint</i> |
| RANSAC | <i>Random Sample Consensus</i> |
| RMS | <i>Root Mean Square</i> |
| SIFT | <i>Scale-Invariant Feature Transform</i> |
| SLAM | <i>Simultaneous Localisation and Mapping</i> |
| SVP | <i>Single Viewpoint</i> |
| VMS | <i>Vision Measurement System</i> |

SUMÁRIO

| | | |
|----------|--|-----------|
| 1 | INTRODUÇÃO | 15 |
| 1.1 | DEFINIÇÃO DO PROBLEMA | 18 |
| 1.2 | OBJETIVO | 19 |
| 1.3 | CONTRIBUIÇÕES | 20 |
| 1.4 | ESTRUTURA DO DOCUMENTO | 20 |
| 2 | FUNDAMENTOS | 22 |
| 2.1 | NOTAÇÃO | 22 |
| 2.2 | GEOMETRIA PROJETIVA | 22 |
| 2.2.1 | Coordenadas homogêneas e outras definições | 23 |
| 2.2.2 | Plano projetivo | 23 |
| 2.2.3 | Transformação projetiva | 24 |
| 2.3 | MODELO DE CÂMERA | 24 |
| 2.3.1 | Mudança de coordenadas de mundo para coordenadas de câmera | 25 |
| 2.3.2 | Mudança de coordenadas de câmera para coordenadas de imagem | 26 |
| 2.3.3 | Distorção de lentes | 27 |
| 2.3.4 | Mudança de coordenadas de imagem para coordenadas em <i>pixel</i> | 29 |
| 2.3.5 | Composição das transformações | 30 |
| 2.4 | CALIBRAÇÃO DE CÂMERA | 31 |
| 2.5 | CONSIDERAÇÕES FINAIS | 33 |
| 3 | TRABALHOS RELACIONADOS | 34 |
| 3.1 | CÂMERA COM PONTO DE VISTA ÚNICO (SVP) | 35 |
| 3.1.1 | Calibração seca com adaptação dos parâmetros | 36 |
| 3.1.2 | Calibração molhada | 38 |
| 3.2 | CÂMERA COM MÚLTIPLOS PONTOS DE VISTA (NON-SVP) | 41 |
| 3.3 | CONSIDERAÇÕES FINAIS | 45 |
| 4 | CSAT - TÉCNICA PROPOSTA | 46 |
| 4.1 | VISÃO GERAL | 46 |
| 4.2 | DISTORÇÃO OCASIONADA PELA MUDANÇA DE MEIO | 48 |
| 4.2.1 | Formulação do problema | 48 |
| 4.2.2 | Função de distorção obtida pela variação de meio | 50 |
| 4.3 | ALGORITMO | 53 |
| 4.3.1 | Etapa 1: Amostragem de pontos | 54 |
| 4.3.2 | Etapa 2: Retroprojeção | 54 |

| | | |
|--------------|--|------------|
| 4.3.3 | Etapa 3: Adição da distorção do meio | 55 |
| 4.3.4 | Etapa 4: Adição da distorção de lente | 58 |
| 4.3.5 | Etapa 5: Otimização dos parâmetros do modelo | 58 |
| 4.4 | CONSIDERAÇÕES FINAIS | 59 |
| 5 | EXPERIMENTOS E ANÁLISE DOS RESULTADOS | 60 |
| 5.1 | CONFIGURAÇÕES DOS EXPERIMENTOS | 60 |
| 5.1.1 | Implementação | 61 |
| 5.1.2 | Método de avaliação comparativa | 61 |
| 5.2 | EXPERIMENTOS SINTÉTICOS | 63 |
| 5.2.1 | Cenário 01 | 65 |
| 5.2.1.1 | Objeto de validação fixo à 500 mm da câmera | 65 |
| 5.2.1.2 | Objeto de validação fixo à 1000 mm da câmera | 70 |
| 5.2.2 | Cenário 02 | 74 |
| 5.2.2.1 | Objeto de validação na borda da imagem | 74 |
| 5.2.2.2 | Objeto de validação no centro da imagem | 76 |
| 5.2.3 | Cenário 03 | 78 |
| 5.2.3.1 | Objeto de validação na borda da imagem | 80 |
| 5.2.3.2 | Objeto de validação com posição fixa | 82 |
| 5.3 | EXPERIMENTOS EM AMBIENTES REAIS | 85 |
| 5.3.1 | Tanque de água | 86 |
| 5.3.2 | Ambiente submarino | 88 |
| 5.3.2.1 | Vídeo 01 | 91 |
| 5.3.2.2 | Vídeo 02 | 93 |
| 5.3.2.3 | Vídeo 03 | 95 |
| 5.3.2.4 | Vídeo 04 | 96 |
| 5.3.2.5 | Comparação dos resultados | 98 |
| 5.4 | CONSIDERAÇÕES FINAIS | 99 |
| 6 | CONCLUSÃO | 101 |
| 6.1 | LIMITAÇÕES | 102 |
| 6.2 | CONTRIBUIÇÕES | 102 |
| 6.3 | TRABALHOS FUTUROS | 103 |
| | REFERÊNCIAS | 104 |
| | APÊNDICE A – RESULTADOS DOS CENÁRIOS SINTÉTICOS | 110 |
| | APÊNDICE B – RESULTADOS DOS CENÁRIOS REAIS | 115 |

1 INTRODUÇÃO

Visão computacional é a área de ciência da computação em que uma ou mais imagens digitais são analisadas, processadas e combinadas para extrair informações úteis sobre um ambiente ou objeto presente nessas imagens. Com o avanço tecnológico, essa área tornou-se cada vez mais importante no meio científico, com muitas propostas de novas aplicações e melhorias para automatizar e auxiliar atividades realizadas diariamente. Algumas das aplicações no escopo dessa área são reconstrução 3D, robôs autônomos, inspeção óptica automatizada, além de rastreamento e reconhecimento de objetos.

Nas últimas décadas, muitos cientistas começaram a reutilizar as técnicas de visão computacional convencionais para uso em ambientes submersos, mais comumente para cenários aquáticos. Nesses cenários, aplicações podem auxiliar na análise do desempenho de atletas de natação através de algoritmos que fornecem informações sobre, por exemplo, angulações, velocidade e tempo obtido em determinados trechos do tanque de água. Esses algoritmos podem extrair informações no espaço bidimensional (SILVA et al., 2013; OSBOROUGH; DALY; PAYTON, 2015; JESUS et al., 2012) ou tridimensional (CESERACCIU et al., 2011; PUEL et al., 2012; MCCABE; PSYCHARAKIS; SANDERS, 2011).

Por apresentar grande diversidade biológica e geológica, novas possibilidades no campo de telecomunicações e muitas fontes energéticas, como gás e petróleo, o oceano é um dos ambientes subaquáticos mais pesquisados nas últimas décadas. Como o ambiente submarino é hostil para o ser humano, devido as suas características físicas, normalmente as câmeras utilizadas nas aplicações de visão computacional são fixadas em veículos submarinos não tripulados (UUV - *Unmanned underwater vehicles*). Esses robôs podem ser classificados em veículo submarino autônomo (AUV - *Autonomous Underwater Vehicle*) e veículo operado remotamente (ROV - *Remotely Operated Vehicle*), diferenciando-se pelo modo em que realiza seu movimento: automático ou definido por usuário, respectivamente.

Dentre as muitas aplicações utilizadas no ambiente submarino, as mais comuns são: Medição do comprimento de animais marinhos para estimar a distribuição de biomassa ou da população (PETRELL et al., 1997; HARVEY et al., 2003; ROSEN et al., 2013; SANTANA-GARCON; NEWMAN; HARVEY, 2014); Mapeamento costeiro (DOUCETTE; HARVEY; SHORTIS, 2002) e do fundo do oceano (ROMAN; INGLIS; RUTTER, 2010); Levantamentos arqueológicos de naufrágios (BINGHAM et al., 2010; BOSCH et al., 2017); Monitoramento e inspeção de estruturas submarinas (BALDWIN, 1984; FORESTI, 2001; NEGAHDARIPOUR; FIROOZFAM, 2006; O'BYRNE et al., 2014); Modelagem virtual de sítios arqueológicos (DRAP et al., 2007); Reconstrução 3D de objetos submersos (BIANCO et al., 2011; SANTOS et al., 2015).

Um fator comum para todas essas aplicações de imagens subaquáticas é o alto nível de acurácia projetado ou especificado como requisito durante o desenvolvimento de

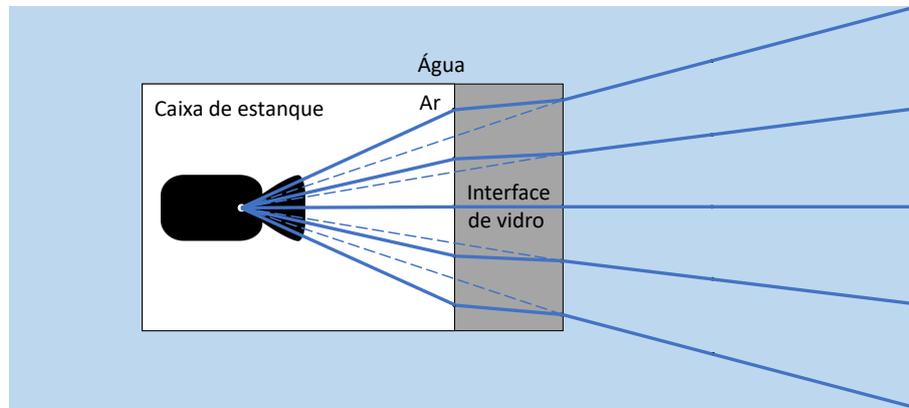
cada uma. Por exemplo, as aplicações que obtêm informações sobre as distribuições de biomassa ou população de animais marinhos dependem diretamente da acurácia das medidas de comprimento e qualquer imprecisão levará a erros significativos nas estimativas geradas. Esse comportamento também existe nas demais aplicações subaquáticas, pois, normalmente elas dependem de alguma informação métrica a ser calculada através das imagens formadas pelas câmeras, sendo obtida por meio de um algoritmo para reconstrução ou rastreamento 3D, por exemplo.

Para os algoritmos de visão computacional produzirem medições precisas e confiáveis, é necessário que as câmeras utilizadas pelas aplicações estejam devidamente calibradas. O processo de calibração envolve, normalmente, expor um objeto de características conhecidas, denominado padrão de calibração, para a câmera em diferentes pontos de vista. A calibração consiste em estimar os parâmetros que melhor relacionam os pontos característicos do padrão de calibração no mundo 3D para os pontos característicos do padrão detectados nas imagens obtidas pela câmera. Quanto mais acurada for a medição 3D e detecção na imagem 2D dos pontos característicos do padrão, melhor será a estimativa dos parâmetros de câmera.

Por determinar a relação entre o mundo 3D e as coordenadas de imagem, a calibração é uma etapa essencial quando utiliza um sistema de câmera para obter informações métricas sobre a cena através das imagens produzidas pelos sensores. Do ponto de vista prático, isso significa que os erros obtidos na estimativa dos parâmetros de câmera, realizado no procedimento de calibração, serão propagados para as próximas etapas, deteriorando a qualidade geral do sistema de visão computacional. Por isso, com a finalidade de eliminar a introdução de erros sistemáticos nas medições realizadas, pequenas perturbações existentes no processo de formação da imagem devem ser consideradas no modelo matemático que descreve o comportamento da câmera utilizada em um determinado ambiente ([SHORTIS, 2015](#)).

No ambiente subaquático, a calibração da câmera tem ainda maior importância, devido aos efeitos de refração, existentes entre os diferentes meios, que devem ser incorporados no modelo. Esse fato acontece pois, para submergir uma câmera, é necessário inseri-la em uma caixa de estanque para evitar a exposição do sensor ao líquido em que está sendo submerso. Essa caixa de estanque, normalmente, possui uma interface de vidro que divide os dois meios, o que a câmera habita e onde está sendo submersa, e devido a diferença na densidade encontrada em cada um dos meios, ocorre o efeito de refração. A consequência é a introdução de uma perturbação não-linear na formação da imagem, como apresentado na [Figura 1](#).

Figura 1 – Geometria dos raios de uma câmera submersa em água com habitação plana. Os segmentos pontilhados e contínuos representam os raios de luz sem e com o efeito de refração sendo incorporado em seu percurso, respectivamente.



Fonte: Elaborada pelo próprio autor.

Outros fatores que aumentam a complexidade do cenário subaquático são a incerteza adicional causada pela atenuação da luz no ambiente e os pequenos erros introduzidos devido a não uniformidade da interface da caixa de estanque (SHORTIS, 2015). Em consequência desses fatores, a condição de visibilidade, tanto em águas rasas (Figura 2a) como em profundas (Figura 2b), apresenta desafios para os algoritmos de detecção utilizados nos procedimentos de calibração.

Figura 2 – Imagens obtidas por câmeras em aplicações submarinas.

(a) 2,5m de profundidade



Fonte: Schechner e Karpel (2003).

(b) 2.000m de profundidade



Fonte: Petrobras.

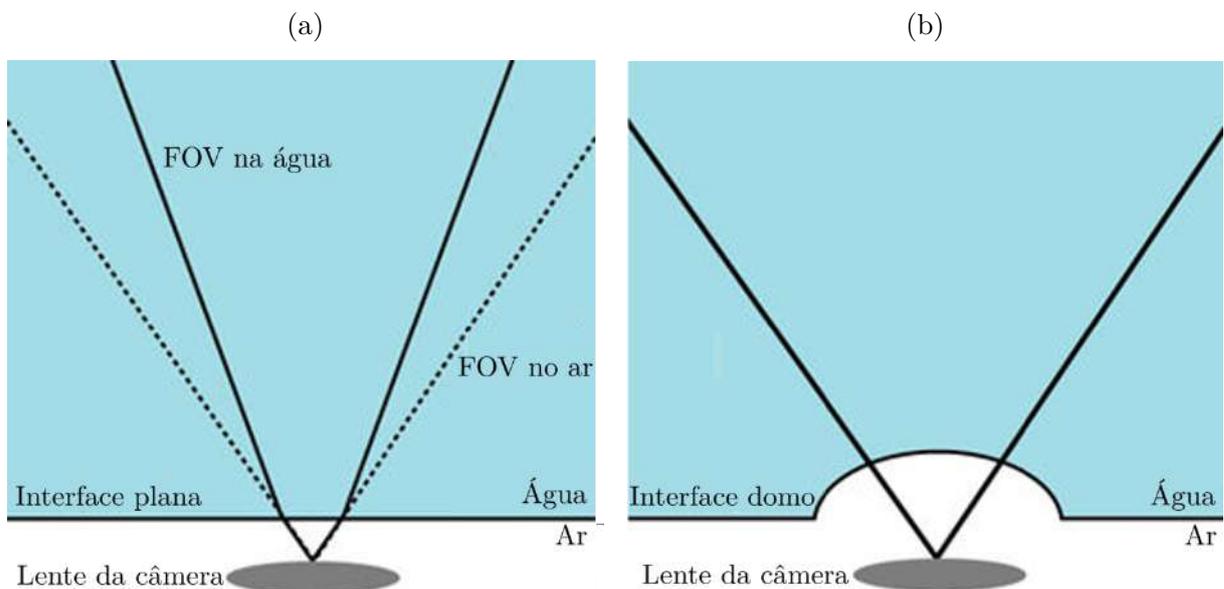
Dada a importância do procedimento de calibração e os desafios existentes nos ambientes subaquáticos, esta dissertação propõe uma técnica capaz de adaptar os parâmetros calculados, antes da câmera ser submersa, para descrever o processo de formação da imagem no ambiente em que a câmera será utilizada. Para alcançar esse propósito, foi desenvolvido um algoritmo genérico que altera os parâmetros de câmera para ambientes com diferentes densidades ópticas (também denominado de índice de refração), não havendo a necessidade de realizar o procedimento de calibração no ambiente de aplicação.

1.1 DEFINIÇÃO DO PROBLEMA

Para que uma câmera seja submersa, é necessário inseri-lá em uma caixa de estanque para evitar que o líquido externo entre em contato com os sensores. Por habitar nessa caixa, os raios de luz que percorrem do meio externo até o centro de projeção da câmera sofrem deslocamentos, quando comparado ao cenário onde há apenas um meio envolvido. Esse fato deve-se ao princípio físico da refração, que modifica a direção percorrida pelo raio de luz, uma vez que diferentes ambientes possuem índices de refração (IR) variados, causando assim uma mudança na velocidade da luz.

Atualmente, existem dois tipos de caixa de estanque: com interface plana ou em formato de domo. A diferença entre os dois tipos de interface, que normalmente possuem material de vidro, é apresentada na [Figura 3](#). A interface em domo foi desenvolvida para corrigir o efeito da refração existente na mudança de meio, sendo moldada como um arco de 180 graus ([JONSSON, 2013](#)). Esta interface óptica, quando acoplada a câmera, atua como uma lente divergente (negativa), evitando a perturbação inserida pela refração e tornando o campo de visão (FOV - *Field Of View*) da câmera igual, tanto dentro como fora do meio submerso. Porém, para a lente domo realizar o efeito desejado, o centro de projeção da câmera necessita estar perfeitamente alinhado com o centro do domo, o que na prática é difícil de alcançar ([BRÄUER-BURCHARDT et al., 2015](#)). Este fato, somado com a fácil fabricação de caixa de estanque com interface plana, tornando-a mais barata, faz com que grande parte das aplicações que utilizam câmeras submersas escolham interface plana e, por isso, esta será o foco dessa dissertação.

Figura 3 – Diagrama simplificado do campo de visão da câmera em uma caixa de estanque com uma interface plana (a) e interface domo (b).



Fonte: Adaptado de [Jonsson \(2013\)](#).

Como é apresentado no [Capítulo 3](#), a comunidade científica considera duas formas para modelar o efeito de refração em uma câmera submersa. A primeira abordagem considera a caixa de estanque e a câmera como um único sistema, sendo a perturbação causada pela refração absorvida nos parâmetros físicos e ópticos do modelo. Essa abordagem parte do pressuposto de que o eixo óptico da câmera é, aproximadamente, perpendicular ao plano da caixa de estanque, fazendo com que a refração existente seja modelada como um deslocamento radialmente simétrico em torno do ponto principal da imagem ([SHORTIS, 2015](#)). Em contrapartida, a segunda abordagem consiste em modelar o processo de refração explicitamente, aumentando a complexidade em razão da divisão dos conjuntos caixa de estanque e a câmera, que faz crescer a quantidade de parâmetros necessários.

Resumidamente, o modelo da primeira abordagem representa a câmera como um único centro de projeção (SVP - *Single Viewpoint*), onde todos os raios de luz interceptam este centro, assim como os modelos utilizados para representar câmeras convencionais no ambiente terrestre. Por outro lado, a segunda abordagem considera que não há apenas um único centro de projeção (nSVP - *Non-Single Viewpoint*), ou seja, dependendo da localização do raio de luz, diferentes centros devem ser obtidos na formação de uma única imagem. Uma vez que o modelo considerado influencia diretamente na performance dos algoritmos métricos de visão computacional, este trabalho utilizará a primeira abordagem, com câmera SVP, para priorizar a redução da complexidade do modelo a ser utilizado no procedimento de calibração.

Por fim, com a finalidade de evitar que as dificuldades do meio submerso interfiram no processo de calibração, o problema abordado por este trabalho será o de determinar um novo conjunto de parâmetros que descrevam o comportamento da câmera no ambiente submerso a partir do resultado de uma calibração já existente. Em termos gerais, pode-se sintetizar de tal forma: dado um conjunto de parâmetros de câmera obtidos pela calibração em um ambiente com IR η_1 , instaura-se novos valores, através de uma adaptação, que representem as características física e óptica da câmera no ambiente onde será utilizada, possuindo IR η_2 .

1.2 OBJETIVO

O objetivo principal deste trabalho é propor uma técnica capaz de adaptar os parâmetros de câmera obtidos em um determinado meio para serem utilizados por aplicações empregadas em qualquer outro ambiente, sem a necessidade de recalibrar a câmera.

Em contraste com a literatura existente para calibração de câmera, a técnica proposta deverá evitar o processo de recalibração no ambiente de aplicação. Além disso, o aspecto da perturbação ocasionada na formação da imagem devido a refração também precisará ser levado em consideração, a fim de aumentar a acurácia dos parâmetros estimados. Para tal, os objetivos específicos dessa dissertação são:

1. Projetar um modelo de adaptação que incorpore a distorção ocasionada pela mudança de meio da calibração para o ambiente de aplicação;
2. Desenvolver um algoritmo que utilize o modelo de adaptação projetado;
3. Validar a técnica proposta em ambientes subaquáticos adaptando os parâmetros obtidos por uma calibração no meio terrestre.
4. Comparar a técnica proposta com as existentes no estado da arte para ambientes subaquáticos em diferentes estudos de caso, sintético e real.

1.3 CONTRIBUIÇÕES

A primeira contribuição desse trabalho é a descrição matemática do modelo de adaptação, que incorpora a perturbação ocasionada na formação da imagem, devido a mudança de meio no qual a calibração foi realizada para onde será destinada. Além disso, utilizando o modelo de adaptação, um algoritmo é apresentado em detalhes, sendo capaz de aproximar os parâmetros obtidos em um meio qualquer para outro, bastando conhecer apenas os índices de refração existentes em cada ambiente. Com o intuito investigativo, uma versão modificada do algoritmo proposto também foi desenvolvida e descrita. Por fim, realizou-se um estudo comparativo das técnicas propostas com duas existentes na literatura, que empregam o modelo SVP, em cenários sintéticos e reais, onde foram utilizados os ambientes submarino em alta profundidade e tanque de água, comumente adotados em aplicações de visão computacional subaquática. Todas as técnicas utilizadas foram implementadas pelo autor deste trabalho.

1.4 ESTRUTURA DO DOCUMENTO

O desenvolver dessa dissertação organiza-se da seguinte forma:

Capítulo 2 - Descreve fundamentos teóricos básicos necessários para contextualizar esta pesquisa. Introduzindo geometria projetiva, conceito necessário para o entendimento dos parâmetros intrínsecos da câmera e para a técnica de adaptação proposta. Segue apresentando o modelo de câmera utilizado nesse trabalho e algumas definições sobre a calibração.

Capítulo 3 - Apresenta trabalhos presentes na literatura que utilizam ou propõem novas técnicas de calibração de câmera utilizadas em ambientes subaquáticos. Alguns desses trabalhos serviram de base para o desenvolvimento da técnica de adaptação proposta e detalhada neste documento.

Capítulo 4 - Detalha o desenvolvimento da técnica proposta para adaptar a calibração realizada em um meio para outro qualquer, onde a câmera será usada. Para tal, é

exposto em detalhes os conceitos e as formulações utilizados na criação do modelo de adaptação, descrevendo também o algoritmo desenvolvido para calcular os novos parâmetros que descrevem a câmera no meio em que a aplicação será executada.

Capítulo 5 - Apresenta diferentes cenários sintéticos para verificar a influência de alguns fatores, inerentes a câmera e sua caixa de estanque, sobre a técnica proposta e as já existentes na literatura. Objetivando corroborar com as conclusões obtidas nos cenários sintéticos, experimentos são realizados em um cenário real correspondente a um tanque de água. Por fim, o resultado comparativo a cerca da técnica de adaptação proposta é discutido para uma aplicação de reconstrução 3D de dutos submarinos, utilizados na extração de gás e petróleo, através de um par de câmeras submersas no oceano.

Capítulo 6 - Argumenta sobre as conclusões obtidas durante o desenvolvimento e experimentos descritos nesse trabalho. Apresentando, também, direcionamentos para a continuação desta pesquisa, com novas investigações e melhorias a serem realizadas e validadas.

2 FUNDAMENTOS

Este capítulo apresenta alguns conceitos básicos e definições necessários para o entendimento do processo de calibração de câmeras. Inicialmente é descrita a notação de matrizes, vetores e pontos adotados no texto. Em seguida, na [Seção 2.2](#), são introduzidos conceitos de geometria projetiva, que formam a base para os modelos que representam as câmeras. A [Seção 2.3](#) apresenta o modelo matemático utilizado neste trabalho para descrever o comportamento físico e óptico da câmera usando um conjunto de parâmetros capaz de relacionar o ambiente tridimensional, no sistema de coordenadas de mundo, e bidimensional, no sistema de coordenada de imagem. Por fim, são apresentados alguns tipos de padrão de calibração, além das definições dos tipos de calibração de câmera e estéreo, utilizados nesta dissertação.

2.1 NOTAÇÃO

A notação utilizada ao longo dessa dissertação representará vetores como matrizes colunas, com letras do alfabeto latim e grego em negrito, como \mathbf{x} e Σ . Os elementos de um vetor são detalhados através de colchetes, ou seja, $\mathbf{x} = [x_1, x_2, x_3]^T$. Matrizes que possuem mais de uma coluna serão representadas utilizando letras com fonte sem serifa como \mathbf{R} e \mathbf{Q} . Enquanto valores escalares são denotados com letras minúsculas sem negrito, como a e σ .

Funções serão definidas com argumentos, entre parênteses, e retorno seguindo a notação descrita anteriormente. Por exemplo, $f(\mathbf{a})$ é uma função que retorna um escalar tendo um vetor como parâmetro e $F(\mathbf{v}, a)$ retorna uma matriz dado um vetor e um escalar.

2.2 GEOMETRIA PROJETIVA

Na área de computação gráfica e visão computacional a base matemática em destaque é a geometria projetiva, sendo utilizada para gerar imagens sintéticas, modelar a formação de imagem ou reconstruir objetos 3D. Essa geometria fornece o formalismo matemático capaz de descrever o modelo das câmeras e as suas transformações associadas, permitindo assim compreender as projeções 2D de objetos 3D realizadas durante a formação das imagens capturadas pelas câmeras. Por esse motivo, esta seção introduzirá os conceitos de geometria projetiva que serão utilizados para a descrição do modelo de câmera utilizado neste trabalho. Mais detalhes podem ser encontrados em [Hartley e Zisserman \(2004\)](#) e [??](#)).

2.2.1 Coordenadas homogêneas e outras definições

Um ponto no espaço projetivo \mathbb{P}^n é representado por um vetor com $(n + 1)$ coordenadas, ou seja, $\mathbf{x} = [x_1, \dots, x_{n+1}]^\top$, onde ao menos uma coordenada x_i deve ser diferente de zero. Dois pontos, \mathbf{x} e \mathbf{y} , são considerados iguais se houver um escalar λ , tal que $\mathbf{x} = \lambda\mathbf{y}$. Como a escala não é importante na geometria projetiva, os vetores descritos acima são chamados de coordenadas homogêneas de um ponto. A geometria projetiva representa pontos no infinito através da última coordenada, igualando-a a zero, ou seja, um ponto no infinito no espaço \mathbb{P}^n possui $x_{n+1} = 0$.

Considerando um ponto no espaço euclidiano \mathbb{R}^3 sendo representado por $[x, y, z]^\top$, tem-se que o seu correspondente no espaço projetivo \mathbb{P}^3 será representado por qualquer ponto $[x_1, x_2, x_3, x_4]^\top$ em que $x_4 \neq 0$. O mapeamento inverso também pode ser realizado através da divisão das coordenadas por x_4 , obtendo $[x_1/x_4, x_2/x_4, x_3/x_4]^\top$. Convencionalmente as coordenadas $[x, y, z]^\top$ e $[x_1, x_2, x_3, x_4]^\top$ são chamadas de coordenadas não homogêneas e homogêneas, respectivamente.

O mapeamento do ponto em \mathbb{R}^n para \mathbb{P}^n pode ser generalizado utilizando a relação

$$\underbrace{[x_1, x_2, \dots, x_n]}_{\text{Espaço euclidiano}} \longrightarrow \underbrace{[\lambda x_1, \lambda x_2, \dots, \lambda x_n, \lambda]}_{\text{Espaço projetivo}}, \quad (2.1)$$

onde $\lambda \neq 0$ é o parâmetro livre de escala.

2.2.2 Plano projetivo

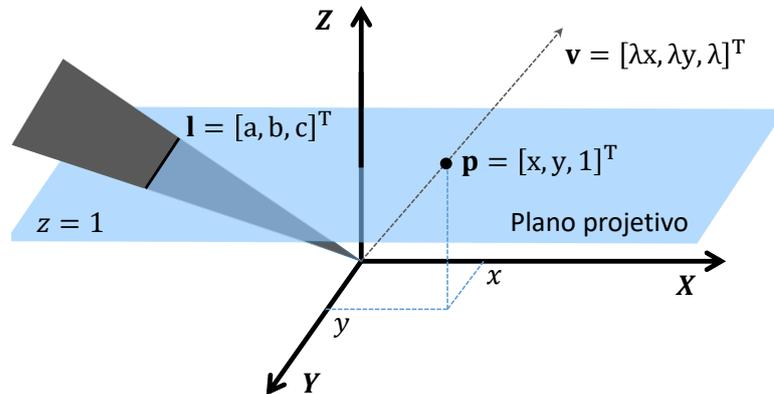
O plano projetivo \mathbb{P}^2 é o conjunto das retas no espaço euclidiano \mathbb{R}^3 que passa pela origem para representar pontos e retas, como apresentado na [Figura 4](#). A representação de um ponto $\mathbf{x} = [\lambda x, \lambda y, \lambda]^\top$ em \mathbb{P}^2 é dada pela interseção entre um raio que passa pela origem de \mathbb{R}^3 e o plano em $z = 1$. Enquanto uma reta $\mathbf{l} = [a, b, c]^\top$ é representada por um plano que passa na origem do \mathbb{R}^3 interceptando o plano em $z = 1$.

O ponto \mathbf{x} está localizado na reta \mathbf{l} se

$$\mathbf{l}^\top \mathbf{x} = 0. \quad (2.2)$$

Esta equação significa tanto que um ponto \mathbf{x} é representado por um conjunto de retas, quanto que uma reta \mathbf{l} é representada por um conjunto de pontos. Esses dois significados mostram que não há diferença entre pontos e retas em \mathbb{P}^2 , sendo explicado pelo princípio da dualidade ([HARTLEY; ZISSERMAN, 2004](#)). Na geometria projetiva, qualquer teorema que seja verdade para o plano projetivo pode ser modificado substituindo pontos por retas, ou o contrário, e o resultado também será verdadeiro. A partir da [Equação 2.2](#), pode-se determinar que a reta no infinito é representada por todos os elementos, com exceção da última coordenada, iguais a 0, ou seja, em \mathbb{P}^2 a reta no infinito é representada por $\mathbf{l} = [0, 0, c]^\top$

Figura 4 – Representações de ponto e reta no plano projetivo. O ponto \mathbf{p} em \mathbb{P}^2 é representado pela interseção entre o raio que passa pela origem em \mathbb{R}^3 com o plano em $z = 1$, enquanto a reta \mathbf{l} em \mathbb{P}^2 por um plano que passa pela origem do \mathbb{R}^3 e intercepta o plano em $z = 1$.



Fonte: Elaborada pelo próprio autor.

2.2.3 Transformação projetiva

A transformação entre dois espaços projetivos \mathbb{P}^n é representado por uma matriz de tamanho $(n + 1) \times (n + 1)$, dada por:

$$\mathbb{T} = \begin{bmatrix} A & \mathbf{t} \\ \mathbf{p}^T & s \end{bmatrix}, \quad (2.3)$$

onde, A é uma matriz $n \times n$, e \mathbf{t} um vetor de tamanho n , que combinados correspondem a uma transformação afim. O vetor \mathbf{p} , de tamanho n , representa a projeção em perspectiva aplicada em relação a um ponto de observação. O s é um valor que representa a escala da matriz e, geralmente, por se tratar de uma matriz independente de escala, possui valor 1.

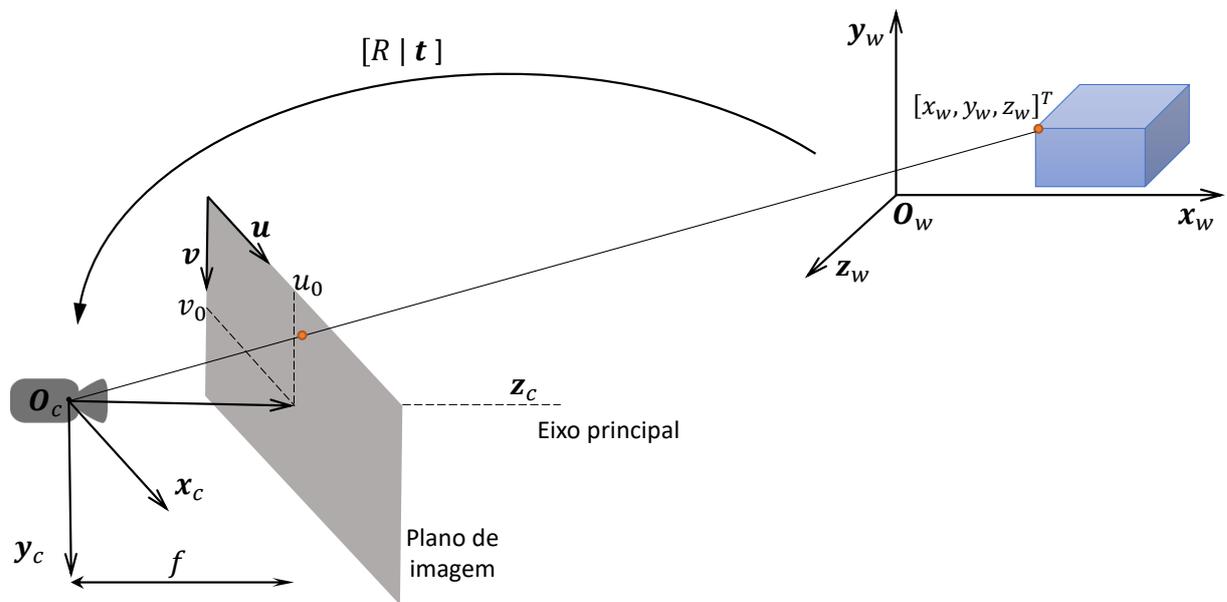
A transformação entre um espaço projetivo \mathbb{P}^{n+1} para outro no \mathbb{P}^n é representada por uma matriz de dimensões $n \times (n + 1)$. Este tipo de transformação, denominada transformação projetiva, é utilizada para projetar um ponto 3D no plano de imagem, como será definido na próxima seção.

2.3 MODELO DE CÂMERA

O modelo de câmera descreve matematicamente a relação entre o mundo real 3D e sua projeção no plano de imagem, como apresentado na [Figura 5](#). Essa modelagem é fundamental para a realização da calibração, possuindo grande influência na precisão da estimativa dos parâmetros que descrevem as características da câmera. Diversos modelos

foram propostos, os mais simples relacionam os pontos 3D da cena com os 2D projetados através de uma transformação linear, como é realizado por ??). Este tipo de modelo possui uma matemática mais fácil de ser aplicada e é computacionalmente mais eficiente, porém carece de uma boa precisão.

Figura 5 – Modelo de formação de imagens pela câmera. O ponto $[x_w, y_w, z_w]^T$, em coordenada de mundo, é transformado até o sistema de coordenadas de tela, em *pixel* com direções u e v . Esta transformação é realizada com as informações dos parâmetros que descrevem a câmera.



Fonte: Elaborada pelo próprio autor.

O modelo de câmera adotado é dividido em quatro etapas. A primeira consiste em uma transformação euclidiana 3D responsável por modificar o sistema de coordenadas do objeto de mundo para o de câmera. A segunda etapa aplica uma transformação projetiva para projetar os pontos para o plano de imagem. A terceira é responsável por modelar os efeitos causados pelas lentes da câmera através de uma transformação não linear. Por fim, uma transformação euclidiana 2D é aplicada para converter as coordenadas de imagem para o computador, em unidade de *pixels*. A seguir essas etapas serão detalhadas utilizando o modelo adotado nesta dissertação.

2.3.1 Mudança de coordenadas de mundo para coordenadas de câmera

O objetivo dessa etapa é transformar um ponto que está no sistema de coordenadas de mundo $\mathbf{P}_w = [x_w, y_w, z_w]^T$, centralizada em \mathbf{O}_w , em um ponto no sistema de coordenadas da câmera $\mathbf{P}_c = [x_c, y_c, z_c]^T$, centralizada em \mathbf{O}_c , como apresentado na Figura 5. Esta

transformação é modelada através de

$$\begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} = \mathbf{R} \begin{bmatrix} y_w \\ y_w \\ z_w \end{bmatrix} + \mathbf{t}, \quad (2.4)$$

onde \mathbf{R} é uma matriz de rotação, com 3×3 de dimensão, representada por elementos r_{ij} na linha i e coluna j , e \mathbf{t} é o vetor de translação, composto por $[t_x, t_y, t_z]^\top$. Sendo \mathbf{R} e \mathbf{t} responsáveis por modificar a orientação dos eixos e posição do centro do sistema, respectivamente. Esta transformação, em visão computacional, é denominada de pose do objeto em relação à câmera.

2.3.2 Mudança de coordenadas de câmera para coordenadas de imagem

Esta etapa é responsável pela projeção do ponto \mathbf{P}_c , do sistema de coordenadas 3D, para o ponto $\mathbf{p}_u = [x_u, y_u]^\top$ no plano de imagem 2D. A representação de câmera *pinhole* é amplamente utilizada em trabalhos de calibração para modelar a transformação projetiva. Este fato ocorre pois a câmera *pinhole* pode modelar precisamente a geometria e a óptica da maioria dos tipos de câmera utilizadas atualmente, como as Vidicon, CCD (*Charge-Coupled Device*) e CID (*Charge Injection Device*) (??).

A [Figura 6](#) apresenta como o modelo de câmera *pinhole* realiza a projeção de pontos. O mapeamento do ponto $[x_c, y_c, z_c]^\top$ para $[x_u, y_u]^\top$ é realizado interceptando o ponto central de projeção, também chamado de centro de câmera (\mathbf{O}_c). A linha perpendicular ao plano de imagem que passa por \mathbf{O}_c é chamada de eixo ou raio principal. Este eixo principal intercepta o plano de imagem no ponto $[u_0, v_0]^\top$, denominado de ponto principal, que está situado a uma distância f (denominada distância focal) de \mathbf{O}_c ([HARTLEY; ZISSERMAN, 2004](#)). Utilizando semelhança de triângulos na [Figura 6](#), obtemos

$$\frac{x_u}{f} = \frac{x_c}{z_c}, \quad (2.5)$$

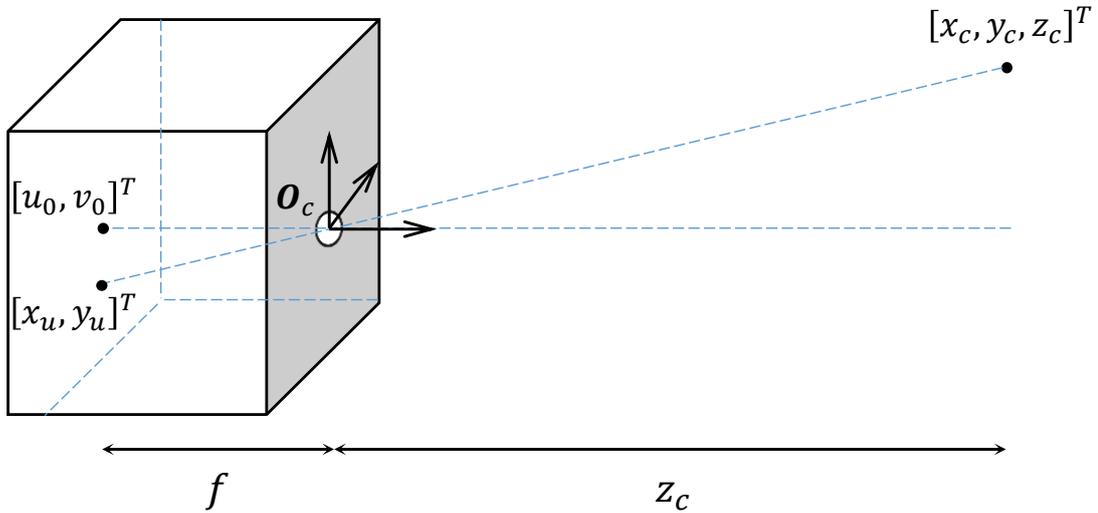
$$\frac{y_u}{f} = \frac{y_c}{z_c}. \quad (2.6)$$

A partir dessas relações é possível extrair uma transformação projetiva capaz de projetar o ponto do \mathbb{P}^3 para o \mathbb{P}^2 . Esta transformação é expressa na forma matricial por

$$\begin{bmatrix} \lambda x_u \\ \lambda y_u \\ \lambda \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_c \\ y_c \\ z_c \\ 1 \end{bmatrix}, \quad (2.7)$$

em que λ é o fator de escala homogênea.

Figura 6 – Representação do modelo de câmera *pinhole*. \mathbf{O}_c é o centro de câmera e $[u_0, v_0]^T$ é o ponto principal da imagem. O ponto $[x_c, y_c, z_c]^T$ é projetado no plano de imagem, a uma distância f (foco) da câmera, formando o ponto $[x_u, y_u]^T$.



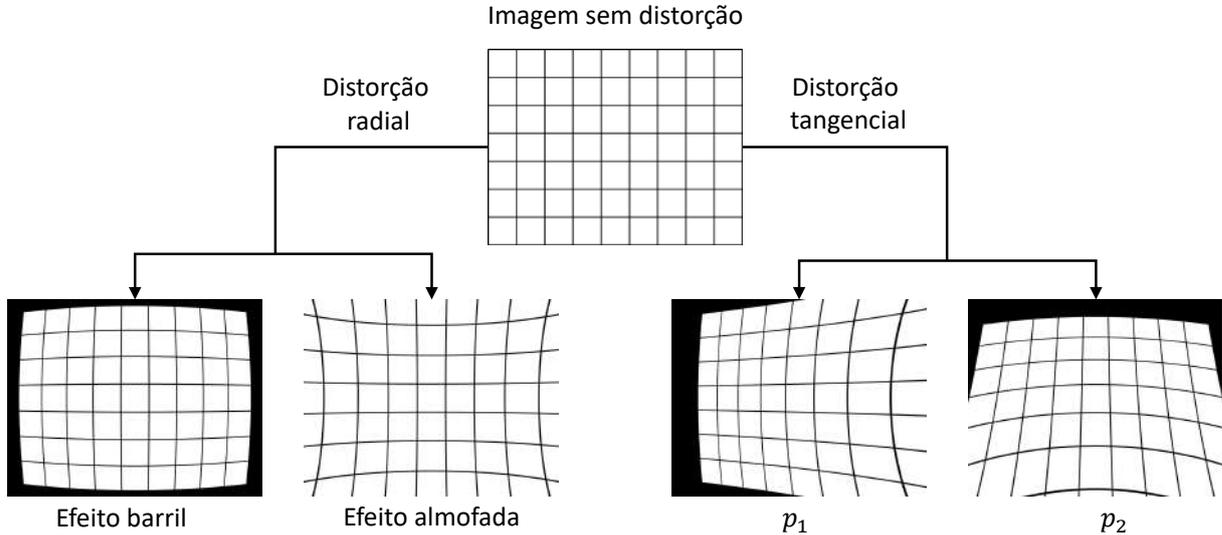
Fonte: Elaborada pelo próprio autor.

A transformação projetiva da [Equação 2.7](#) não é invertível, ou seja, cada ponto do sistema de coordenadas de câmera corresponde a um único ponto da imagem, enquanto um ponto no sistema de coordenadas de imagem corresponde a inúmeros pontos 3D de câmera. Sendo mais preciso, o ponto de imagem corresponde a uma linha infinita no espaço 3D.

2.3.3 Distorção de lentes

Diversas imperfeições são geradas na concepção e montagem das lentes de câmeras (??). Essas imperfeições não são modeladas pelo modelo linear de câmera *pinhole*, mas causam perturbações nos pontos no sistema de coordenadas de imagem, afetando a geometria do modelo. Os principais tipos de perturbações causadas são as distorções radiais e tangenciais, seus efeitos podem ser visualizados na [Figura 7](#).

Figura 7 – Representação dos efeitos de distorção em uma imagem. São apresentados os fenômenos do tipo barril e almofada, causados pela distorção radial, e os efeitos produzidos pelos parâmetros p_1 e p_2 da distorção tangencial, quando possuem valores positivos.



Fonte: Elaborada pelo próprio autor.

A distorção radial é causada principalmente pela falha no projeto da curvatura da lente e provoca deslocamento radial nos pontos, afastando-os ou aproximando-os, em relação ao ponto principal da imagem. Como é apresentado na Figura 7, a distorção radial gera dois tipos de fenômeno, denominados de barril e almofada. A perturbação do tipo barril é causado, nesse modelo, pela distorção radial negativa, realiza o deslocamento dos pontos aproximando-os do ponto principal da imagem, sendo geralmente associada às lentes grande angulares. Enquanto o efeito almofada é produzido, nesse modelo, pela distorção radial positiva, afastando os pontos em relação ao ponto principal da imagem, sendo geralmente associada às lentes telefoto para câmeras de longo alcance.

A distorção tangencial é causada pela descentralização e desalinhamento dos elementos físicos da lente (??). Os efeitos da distorção tangencial são menos perceptíveis na maioria das câmeras, além disso, adicioná-lo na modelagem aumenta a complexidade do modelo, por esses motivos diversos trabalhos não consideram esse tipo de perturbação da imagem. Entretanto, nesta dissertação este tipo de distorção está sendo incorporado ao modelo.

Propostas por Duane (1971), as expressões

$$\delta_r = \delta(\mathbf{p}_d, \mathbf{a}) = \begin{pmatrix} (x_d - u_0)d_x(a_1r^2 + a_2r^4 + a_3r^6 + \dots) \\ (y_d - v_0)d_y(a_1r^2 + a_2r^4 + a_3r^6 + \dots) \end{pmatrix} \quad (2.8)$$

e

$$\delta_t = \delta(\mathbf{p}_d, \mathbf{p}) = \begin{pmatrix} p_1[r^2 + 2(x_d - u_0)^2d_x^2] + 2p_2(x_d - u_0)d_x(y_d - v_0)d_y \\ p_2[r^2 + 2(y_d - v_0)^2d_y^2] + 2p_1(x_d - u_0)d_x(y_d - v_0)d_y \end{pmatrix} \quad (2.9)$$

são comumente utilizadas pela fotogrametria e também por trabalhos na área de visão computacional, para modelar as distorções radial (δr) e tangencial (δt). O $\mathbf{p}_d = [x_d, y_d]^\top$ é o ponto com distorção obtida na imagem capturada pela câmera, $[u_0, v_0]^\top$ corresponde ao ponto principal da imagem, $\{a_i\}$ representa os parâmetros de distorção radial e $\{p_i\}$ os de distorção tangencial, d_x e d_y são os tamanhos do *pixel* nas direções horizontal e vertical, respectivamente. Por fim,

$$r = \sqrt{(x_d - u_0)^2 d_x^2 + (y_d - v_0)^2 d_y^2} \quad (2.10)$$

é a distância do ponto da imagem até o ponto principal.

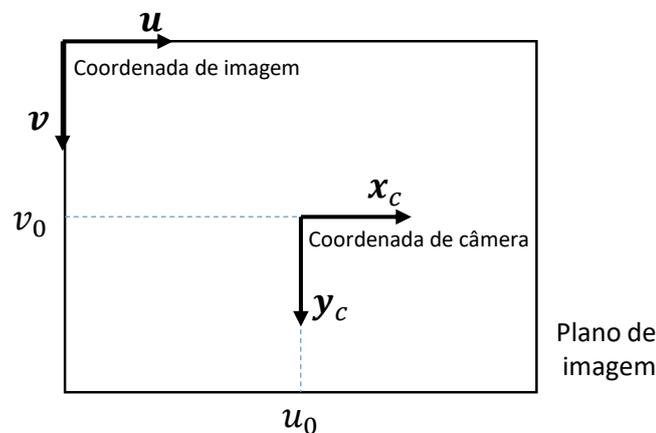
Sabendo que $\mathbf{p}'_u = [x'_u, y'_u]^\top$ é o ponto após a remoção dos efeitos de distorção, o problema de imperfeições das lentes removendo a distorção dos pontos obtido pela imagem da câmera foi modelado através de

$$\mathbf{p}'_u = \delta(\mathbf{p}_d, \mathbf{a}, \mathbf{p}) = \delta_r(\mathbf{p}_d, \mathbf{a}) + \delta_t(\mathbf{p}_d, \mathbf{p}) + \mathbf{p}_d \quad (2.11)$$

2.3.4 Mudança de coordenadas de imagem para coordenadas em *pixel*

Em visão computacional, normalmente, é adotado que o sistema de coordenadas de imagem no computador inicia-se no canto superior esquerdo da imagem e cresce na direção do lado direito no eixo horizontal e para baixo no eixo vertical. O sistema de coordenadas de imagem projetada pela câmera possui origem no chamado ponto principal $[u_0, v_0]^\top$ e segue a mesma direção do sistema de coordenadas de imagem, conforme ilustrado na [Figura 8](#). Além deste deslocamento, normalmente, os *pixels* da imagem não são perfeitamente quadrados possuindo valores diferentes nos eixos horizontal e vertical.

Figura 8 – Sistema de coordenadas de imagem $[u, v]^\top$ e câmera $[x_c, y_c]^\top$.



Fonte: Adaptado de [Hartley e Zisserman \(2004\)](#).

A transformação, na forma matricial, responsável por esta modificação de coordenadas é definida como

$$\begin{bmatrix} \lambda x_p \\ \lambda y_p \\ \lambda \end{bmatrix} = \begin{bmatrix} 1/d_x & 0 & u_0 \\ 0 & 1/d_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_{cam} \\ y_{cam} \\ 1 \end{bmatrix}, \quad (2.12)$$

onde d_x e d_y correspondem ao número de *pixels* por unidade de comprimento nas direções horizontal e vertical, respectivamente. $[u_0, v_0]^\top$ é o ponto principal da imagem que representa as coordenadas da projeção ortogonal do centro óptico no plano de imagem, modificando a origem do sistema.

2.3.5 Composição das transformações

Resumindo as etapas do modelo descrito para transformar o sistema de um ponto em coordenadas de mundo 3D para coordenadas em *pixels*, tem-se que

$$\begin{bmatrix} x_p \\ y_p \\ 1 \end{bmatrix} = \begin{bmatrix} 1/d_x & 0 & u_0 \\ 0 & 1/d_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix}, \quad (2.13)$$

sendo o ponto $[x_p, y_p]^\top$ a projeção no plano de imagem, no sistema de coordenadas em *pixels*, de $[x_w, y_w, z_w]^\top$ sem nenhum tipo de distorção. Para compará-lo com o ponto $[x_d, y_d]^\top$, capturado na imagem da câmera, é necessário remover a distorção existente neste ponto da imagem, de acordo com a equação [Equação 2.11](#), resultando em

$$e = d([x_p, y_p]^\top, [x'_u, y'_u]^\top), \quad (2.14)$$

sendo d a distância euclidiana entre dois pontos.

Para melhorar a acurácia da estimativa dos parâmetros de câmera durante o processo de otimização são utilizados termos individuais do erro, pois os métodos de otimização não lineares funcionam melhor quando possuem acesso completo aos componentes do erro (??). O erro associado aos parâmetros estimados com o modelo apresentado é dado por

$$e_x = u_0 + \left(\frac{f}{dx} \right) \frac{r_{11}x_w + r_{12}y_w + r_{13}z_w + t_x}{r_{31}x_w + r_{32}y_w + r_{33}z_w + t_z} - x_d + \frac{(\delta r_x + \delta t_x)}{d_x} \quad (2.15)$$

no eixo horizontal e

$$e_y = v_0 + \left(\frac{f}{dy} \right) \frac{r_{21}x_w + r_{22}y_w + r_{23}z_w + t_y}{r_{31}x_w + r_{32}y_w + r_{33}z_w + t_z} - y_d + \frac{(\delta r_y + \delta t_y)}{d_y} \quad (2.16)$$

no eixo vertical.

A distância focal (f), o tamanho do *pixel* (d_x e d_y), o ponto principal de imagem $[u_0, v_0]^T$ e as distorções de lente (a_1, a_2, a_3, p_1 e p_2) formam as características internas da câmera e são denominados de parâmetros intrínsecos. Enquanto os parâmetros extrínsecos da câmera correspondem à sua orientação (R) e posição (\mathbf{t}) em relação a um sistema de referência no mundo.

2.4 CALIBRAÇÃO DE CÂMERA

Calibração é o processo utilizado para definir ou corrigir determinados parâmetros utilizados como conjunto de entrada para um sistema. Em visão computacional, onde este processo envolve uma ou mais câmeras, as técnicas de calibração podem ser classificadas em quatro tipos, dependendo da inserção ou não de um objeto de medidas conhecidas para obter a correspondência entre os pontos do ambiente real e da imagem (??). A calibração que não insere objetos conhecidos no ambiente é denominada auto calibração e as que se baseiam em um objeto conhecido dividem-se de acordo com a dimensão dos objetos adicionados (1D, 2D e 3D).

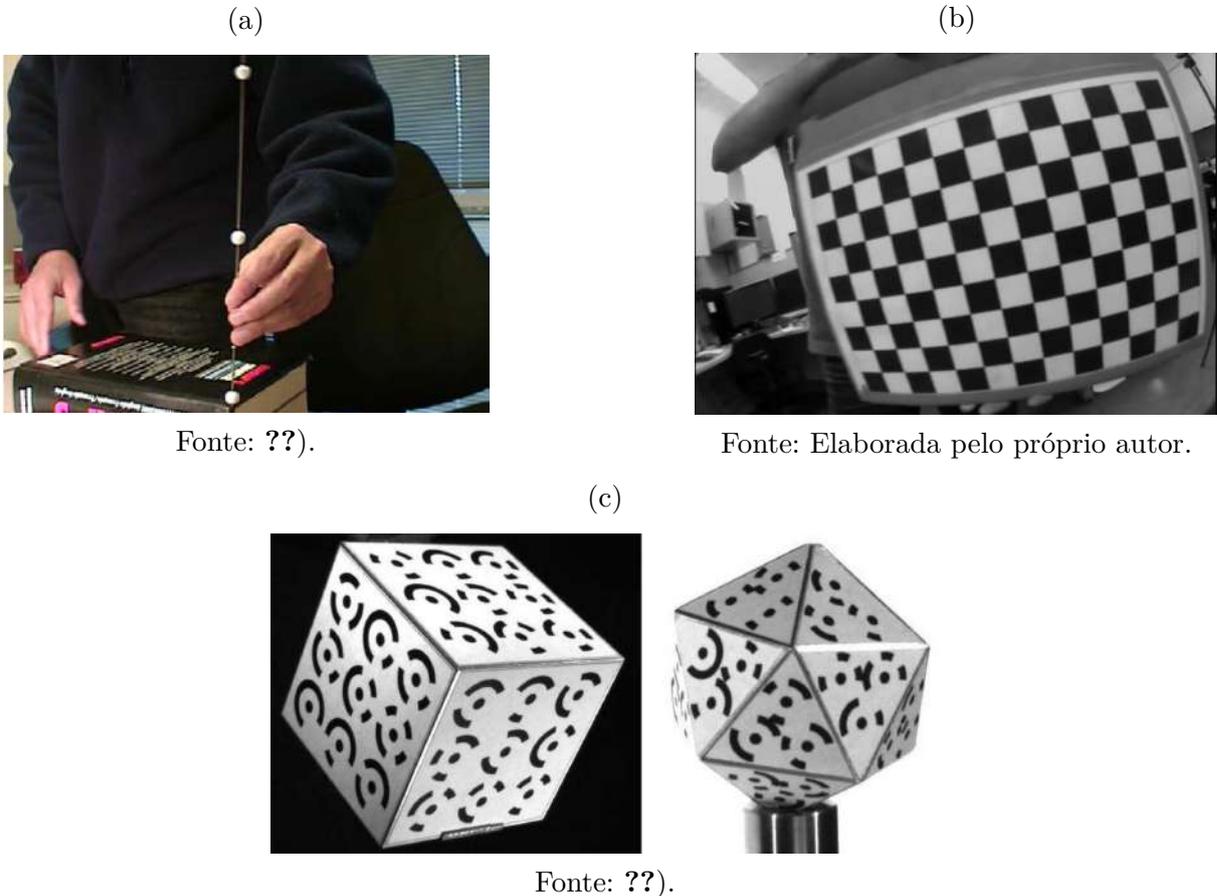
Auto calibração - Esse tipo de calibração requer apenas pontos fornecidos pela imagem da câmera, por isso pode ser considerada como uma calibração 0D. Necessita de pelo menos três imagens para calcular os parâmetros intrínsecos e extrínsecos da câmera. Por não utilizar qualquer objeto conhecido na cena, este tipo de calibração necessita estimar um maior número de parâmetros, o que torna o problema mais complexo matematicamente, resultando em uma calibração menos acurada que a baseada em objetos.

Calibração baseada em objeto 1D - O objeto de calibração é formado por pontos colineares dispostos em forma de uma linha, onde essas técnicas realizam movimentos do objeto em torno de um ponto fixo. Exemplo de um padrão deste tipo é apresentado na [Figura 9a](#).

Calibração baseada em objeto 2D - Esta técnica consiste em exibir um padrão planar para a câmera em diferentes orientações. Para a grande maioria das aplicações de visão computacional esta é uma calibração com acurácia satisfatória e que requer pouco esforço para ser realizada. Este padrão pode ser visualizado na [Figura 9b](#).

Calibração baseada em objeto 3D - O objeto de calibração, normalmente formado por dois ou três planos ortogonais ([Figura 9c](#)), é apresentado para a câmera em diferentes orientações. Este tipo de calibração exibe uma boa precisão na estimativa dos parâmetros, porém possui uma implementação mais complexa e necessita de maior esforço para elaborar o cenário de calibração.

Figura 9 – Exemplo de padrões de calibração 1D, 2D e 3D: (a) Padrão linear formado por três esferas brancas e uma linha rígida. (b) Padrão planar no formato de tabuleiro de xadrez usado na implementação inicial deste trabalho. (c) Padrão 3D de estrutura mais complexa com marcadores autocodificados.



Esses objetos de dimensões conhecidas utilizados nos procedimentos de calibração são comumente denominados de padrão de calibração. Nesta dissertação, as calibrações são realizadas utilizando um objeto 2D de dimensões conhecidas. Usualmente, na literatura de visão computacional, existem dois tipos de calibração: de câmera e estéreo.

A calibração de câmera é o processo realizado para determinar os parâmetros intrínsecos e extrínsecos da câmera, relativo a um sistema de coordenadas de mundo, a partir da relação analítica entre o espaço de mundo 3D e o espaço de imagem 2D (??). Esta é uma etapa fundamental para aplicações que realizam medições em um ambiente real através da imagem capturada pelo dispositivo. Algoritmos de visão computacional usados, por exemplo, para realizar reconstrução densa, inspeção visual e localização de objetos necessitam de informações precisas que tornem possível a correspondência entre o mundo real (tridimensional) e a imagem (bidimensional) (??). Essa correspondência é descrita por um modelo matemático que define as características intrínsecas e extrínsecas da câmera através de fórmulas, como descrito anteriormente na [Seção 2.3](#).

Sistemas que possuem mais de uma câmera, além de realizar o processo usual de

calibração, normalmente, precisam também conhecer a posição relativa entre estes dispositivos. A calibração estéreo determina os parâmetros que compõem a transformação responsável por modificar a orientação e o centro de câmera de um dispositivo para o outro, podendo ser calculada em conjunto com a calibração de câmera ou em separado, utilizando os parâmetros intrínsecos previamente determinados.

A transformação que relaciona duas câmeras é composta por um vetor translação e uma matriz rotação no ambiente 3D e cria uma relação entre as imagens obtidas pelas câmeras. Essa relação torna possível a recuperação de informações 3D a partir de imagens através de algoritmos de visão computacional para a reconstrução da cena. Esses algoritmos são fundamentais para o campo de robótica, assim como para pesquisas em diversas áreas através do mapeamento de regiões.

2.5 CONSIDERAÇÕES FINAIS

Este capítulo apresentou os tópicos mais relacionados com o desenvolvimento deste trabalho, sendo fundamentais para a compreensão da proposta desta dissertação de mestrado. Inicialmente, foi apresentado a notação utilizada ao longo de todo este trabalho para, em seguida, indicar os principais conceitos sobre a geometria projetiva e suas transformações, modelo *pinhole* e distorção de lente, além de definições sobre o processo de calibração de câmara. Todos esses conceitos explicitados neste capítulo são direta ou indiretamente utilizados para a formulação da técnica proposta, apresentada no [Capítulo 4](#). Maiores detalhes sobre os tópicos abordados devem ser encontrados nas referências citadas ao longo do deste capítulo.

3 TRABALHOS RELACIONADOS

Calibração de câmera é usualmente utilizada como o primeiro passo para diversos sistemas de visão computacional, sendo crucial quando necessário extrair informações métricas a partir de imagens obtidas por este sensor. Por isso, com a finalidade de evitar a introdução de erros sistemáticos nas medições desses sistemas, qualquer imperfeição causada durante o processo de formação da imagem deve ser modelada, ou seja, considerada na matemática que descreve relação entre o mundo 3D e o plano de imagem. Calibração de câmeras convencionais, realizadas no ambiente terrestre, é um tópico bem estabelecido na literatura, tendo como um dos pioneiros o trabalho de [Duane \(1971\)](#). Posteriormente, diferentes algoritmos foram propostos, encontrando-se entre os mais populares: [Tsai \(1987\)](#), [Heikkila e Silven \(1997\)](#), [Zhang \(2000\)](#) e [Bouguet \(2004\)](#). Esses trabalhos são utilizados em calibrações subaquáticas ou servem como base para o desenvolvimento de novas técnicas.

Assumindo que alguns parâmetros referentes à câmera são fornecidos pelo fabricante, Tsai desenvolveu uma técnica de calibração com duas etapas ([TSAI, 1987](#)). Dispondo de correspondências entre pontos 3D-2D, a primeira etapa do algoritmo utiliza um método linear para resolver o sistema de equações baseando-se na restrição de alinhamento radial (RAC - *Radial Alignment Constraint*). Devido aos graus de liberdade desse sistema de equações, é necessário obter para cada imagem pelo menos 7 correspondências. A segunda etapa realiza um refinamento iterativo minimizando os erros de reprojeção, assumindo que o efeito ocasionado pela distorção da lente é compensado apenas por um coeficiente de distorção radial. Esta técnica faz uso de uma ou mais imagens de um padrão 3D ou planar com medidas conhecidas.

Em [Heikkila e Silven \(1997\)](#) o procedimento de duas etapas foi estendido para quatro passos. O primeiro passo consiste em determinar uma estimativa inicial para os parâmetros da câmera (intrínsecos e extrínsecos, sem considerar os efeitos de distorção da lente) através do algoritmo de transformação linear direta (DLT - *Direct Linear Transformation*) ([ABDEL-AZIZ; KARARA, 1971](#)), considerando o modelo *pinhole* e que os pontos de calibração são não coplanares. No segundo passo, o método de otimização de Levenberg-Marquardt ([LEVENBERG, 1944; MARQUARDT, 1963](#)) é utilizado para refinar todos os parâmetros da câmera, incluindo os de distorção radial e tangencial da lente (dois coeficientes para cada), minimizando o erro de reprojeção. As demais etapas são realizadas para desempenhar uma correção da imagem, incluindo remoção da distorção, utilizando os parâmetros calculados nos passos iniciais. O método funciona com uma ou várias imagens de um padrão 2D ou 3D.

O algoritmo proposto em [Zhang \(2000\)](#) foi desenvolvido para ser uma técnica flexível que facilite o processo de calibração. Utilizando pelo menos duas imagens de um padrão planar em diferentes orientações, é realizado a correspondência entre os pontos 3D do

padrão, previamente conhecido, e os pontos 2D capturados pela imagem. Sem perder a generalidade, é usada a restrição de que o padrão está no plano $Z = 0$ para calcular a transformação projetiva, ou homografia, entre os pontos correspondidos. Com a homografia calculada, os parâmetros intrínsecos e extrínsecos da câmera são inicializados considerando que não há distorção. Na segunda etapa, todos os parâmetros (incluindo os dois coeficientes de distorção radial, considerados por Zhang) são refinados através da minimização algébrica não-linear do erro de reprojeção com o algoritmo de Levenberg-Marquardt.

Em Bouguet (2004) é descrito e disponibilizado uma implementação de um software para calibrar câmeras usando a ferramenta MATLAB (*Matrix Laboratory*). Seu algoritmo baseia-se no modelo de câmera proposto por Heikkila e Silven (1997), com dois coeficientes para distorção radial e dois para tangencial. A inicialização dos parâmetros da câmera é realizada utilizando a técnica apresentada por Zhang (2000). Para obter uma melhor amostragem do sistema projetivo da câmera e, conseqüentemente, diminuir a influência da imprecisão relacionada à localização desses pontos, Jean-Yves Bouguet sugere o uso de, pelo menos, 20 imagens com 156 pontos de controle para obter resultados aceitáveis na calibração de câmera. Outra implementação bastante utilizada para realizar calibração é disponibilizada pela biblioteca de visão computacional de código aberto (OpenCV - *Open Source Computer vision*), através da linguagem C++ (BRADSKI, 2000).

Este capítulo apresenta uma revisão de trabalhos que abordam problemas similares ao estudado nessa dissertação, envolvendo calibração de câmeras para ambientes subaquáticos, que, devido a introdução de efeitos causados pela refração nas interfaces planas (ar-vidro e vidro-água) da caixa de estanque, ganha uma importância ainda maior. Tais trabalhos foram classificados em duas classes: na primeira, estarão as técnicas em que o modelo considera uma câmera com ponto de vista único; na segunda, constam as técnicas com modelos mais genéricos, assumindo que uma câmera pode ter diferentes centros de projeção.

3.1 CÂMERA COM PONTO DE VISTA ÚNICO (SVP)

Esta classe de técnicas está composta por trabalhos onde o modelo de câmera garante que todos os raios passam por um único centro de projeção, considerando que a câmera possui um único ponto de vista. Para isso, partem do pressuposto que os efeitos causados pela mudança do índice de refração podem ser absorvidos por parâmetros físicos da câmera, como a distância focal e a distorção radial, modelada pelo modelo *pinhole* com correção de aberrações ocasionadas pela lente da câmera.

Essas técnicas são defendidas em Kang, Wu e Yang (2012), onde é realizado um estudo sistemático sobre a influência da refração na reconstrução 3D subaquática para avaliar a performance do modelo SVP em imagens sintéticas e reais. Tal avaliação revelou que esse modelo, com a correção da distorção radial e o ajuste da distância focal, consegue

aproximar bem o efeito causado pela refração, alcançando resultados com boa precisão em reconstruções 3D no ambiente subaquático.

Algoritmos que utilizam a abordagem SVP para realizar a calibração em ambientes subaquáticos podem ser divididos em duas categorias: A primeira contendo técnicas que realizam a calibração no meio terrestre (calibração seca) e adaptam alguns parâmetros intrínsecos à câmera para obter um resultado válido no meio subaquático. Enquanto a segunda categoria, explana os trabalhos que calibram a câmera no mesmo meio em que será utilizada (submersa na água), e que também pode ser denominada de calibração molhada. Nas próximas subseções serão detalhados trabalhos que seguem essas duas abordagens.

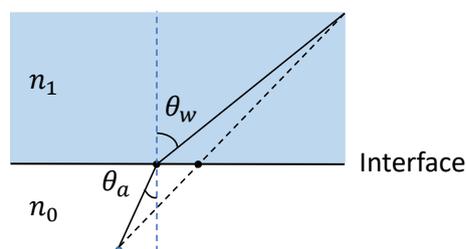
3.1.1 Calibração seca com adaptação dos parâmetros

Fryer e Fraser (1986), baseando-se no trabalho de Duane (1971), realiza um estudo para prever uma calibração válida na água utilizando uma calibração seca. Para isso, faz uso de um modelo de câmera *pinhole*, corrigindo a aberração das lentes com três coeficientes para distorção radial e dois para a tangencial. Considerando que a interface de vidro tem efeito desprezível no deslocamento do raio de luz, Fryer e Fraser adaptaram sua calibração seca compensando a refração através da multiplicação do parâmetro de foco por um fator de 1,34. Além disso, o efeito de distorção radial δr na água, quando se tem uma interface plana, deve ser modificada considerando

$$\delta r = \left(\frac{\cos \theta_w}{\cos \theta_a} - 1 \right) r, \quad (3.1)$$

onde r é a distorção radial no ar, θ_w é o ângulo entre a interface e o raio da água e θ_a é o ângulo entre a interface e o raio no ar, como exemplificado na Figura 10. É importante ressaltar que, em aplicações comuns para câmeras perspectivas, esses ângulos necessários para computar a distorção são geralmente desconhecidos.

Figura 10 – Ângulos entre o raio de luz e uma interface com mudança de meio.



Fonte: Elaborada pelo próprio autor.

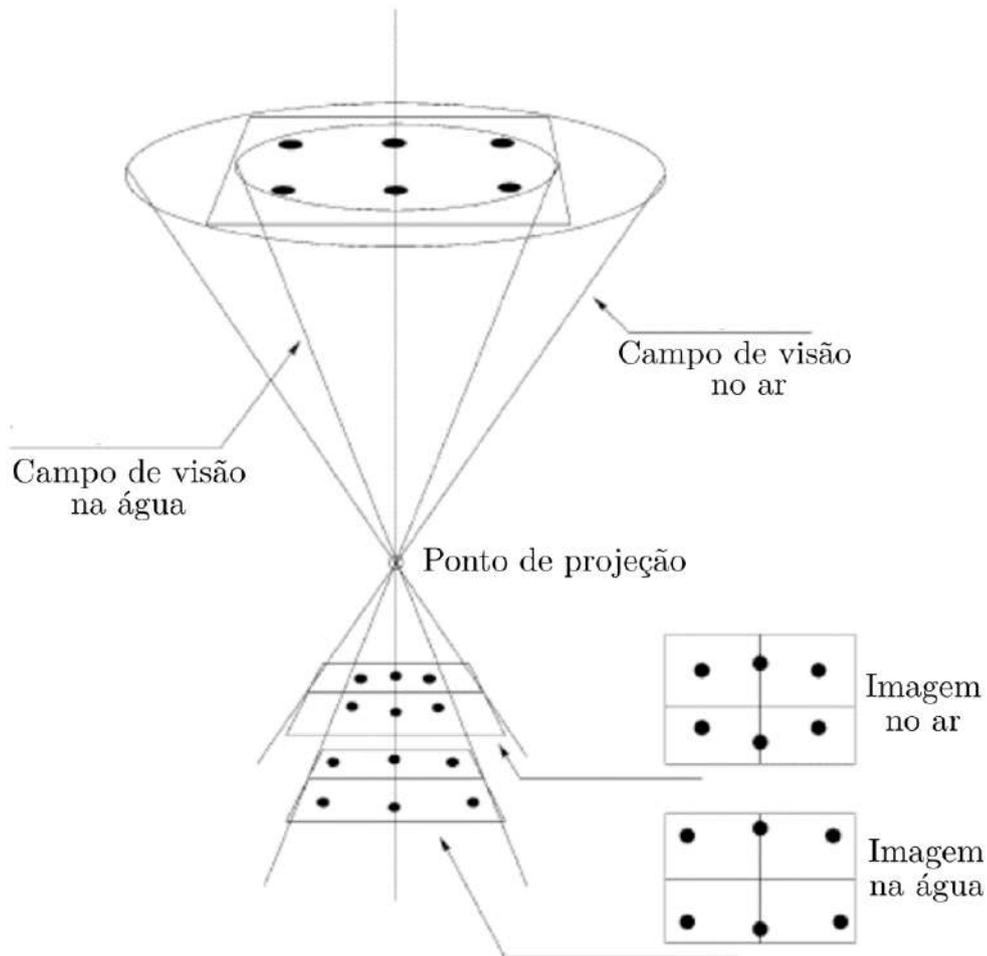
Em Lavest, Rives e Lapresté (2003) é publicado um estudo similar ao de Fryer e Frasers, baseado na teoria óptica e em experimentos. Nesse trabalho, é alegado que os parâmetros de calibração da câmera submersa em água podem ser inferidos a partir dos calculados no ar, desde que a superfície de entrada externa da câmera seja planar. Lavest

et al. demonstram que a distância focal pode ser adaptada de um meio para o outro através da relação

$$f_1 = \frac{n_1}{n_0} f_0, \quad (3.2)$$

onde f_1 e f_0 representam os valores de distância focal da câmera nos diferentes meios e n_1 e n_0 os seus índices de refração. No caso da adaptação da calibração realizada no ar para a submersa em água, basta multiplicar o foco calculado no ar por 1,333 (índice de refração na água). A variação do campo de vista, ocasionada pela mudança no foco, obtida pela câmera no ar e submerso na água é apresentada na [Figura 11](#).

Figura 11 – Variação do campo de visão obtida por uma câmera no ar e na água.



Fonte: Adaptado de [Lavest, Rives e Lapresté \(2003\)](#).

Com a finalização de seus experimentos, Lavest et al. concluem que o ponto principal da imagem permanece inalterado entre calibrações do ar e água. Além de afirmar que a distorção da água, $d_w(u')$, e do ar, $d_a(u)$, segue a relação

$$1,333(u + d_a(u)) = u' + d_w(u'), \quad (3.3)$$

onde u e u' representam o mesmo ponto do objeto 3D projetado na imagem, com distorção, obtida nos meios terrestre e aquático, respectivamente. Essa relação indica que um ponto sem distorção no ar multiplicado pelo fator de 1,333 é igual ao ponto sem distorção na água (com os parâmetros de distorção calculados no meio subaquático). Resumidamente, conclui-se que realizar uma calibração seca e adaptar o foco resulta em parâmetros similares a realização de uma calibração molhada.

Em [Ferreira, Costeira e Santos \(2005\)](#) é proposta uma aproximação da refração por linearização e retenção de termos de primeira ordem da série de Taylor, atingindo assim um centro de imagem virtual ao qual convertem todos os raios de luz, como exemplificado por [Lavest, Rives e Lapresté \(2003\)](#). Utilizando câmera estéreo, esta aproximação permite o uso de algoritmos padrões na literatura de visão computacional para a realização de correspondência estéreo. Os resultados obtidos mostraram que o erro de reconstrução, devido à refração, é pequeno, desde que as câmeras olhem perpendicularmente para a superfície da água.

3.1.2 Calibração molhada

Shortis e Harvey apresentam o desenvolvimento de um sistema estéreo de câmeras para medir informações sobre peixes ([SHORTIS; HARVEY, 1998](#)). O modelo de câmera utilizado não possui nenhum termo explícito para modelar o efeito ocasionado pela refração na interface entre a caixa de estanque e o meio aquático, uma vez que é afirmado pelos autores que esses efeitos causam um deslocamento radialmente simétrico em torno do ponto principal da imagem. Portanto, esse efeito ocasionado pela refração é absorvido nos parâmetros de distorção radial, que modelam esse mesmo tipo de deslocamento dos pontos na imagem. A [Figura 12a](#) apresenta o padrão de calibração utilizado neste sistema estéreo que resultou em um software de medições para ambientes subaquáticos (VMS - *Vision Measurement System* ([GEOMETRIC, 2001](#))).

Devido a importância da calibração de câmeras nas aplicações submarinas e aos desafios apresentados por este ambiente, [Bryant et al. \(2000\)](#) apresentaram um algoritmo para identificação de pontos do padrão de calibração robusto para ambientes turvos. Este algoritmo baseia-se na indexação de invariantes planares calculados a partir dos pontos 3D do padrão conhecido, facilitando a correspondência entre os pontos em um cenário onde a detecção é mais difícil de ser realizada, devido a quantidade de ruído e a baixa luminosidade. A calibração dos parâmetros intrínsecos da câmera é realizada através do algoritmo de Tsai. Esse algoritmo de calibração também é utilizado em [Li et al. \(2016\)](#) para calibrar uma câmera GoPro Hero 3 *black edition*. Após a calibração, Li et al. utilizaram a câmera para realizar a reconstrução 3D submersa do leito de um rio.

Uma técnica de auto-calibração para uma câmera montada em um veículo subaquático projetado para realizar a reconstrução 3D de cenas subaquáticas é apresentado em [Pessel, Opderbecke e Aldon \(2003\)](#). Pessel et al. utilizam uma trajetória de câmera predefinida

para rastrear características naturais encontradas na cena através da técnica de Kanade-Lucas-Tomasi (KLT) (TOMASI; KANADE, 1991). Para aumentar a quantidade e melhorar a qualidade das correspondências, foram rastreados um maior número de características para, posteriormente, utilizar um método robusto de identificação (RANSAC - *Random Sample Consensus*) de correspondências corretas. Neste trabalho, a distorção não é modelada e os parâmetros intrínsecos são calculados a partir do cálculo da matriz essencial e fundamental extraídas através da geometria epipolar entre as imagens.

Em Eustice, Pizarro e Singh (2008), a calibração é realizada dentro da água, utilizando o método planar de Zhang com o modelo de distorção radial e tangencial empregado por Heikkila e Silven, para compensar a efeitos causados pela lente e refração. Os parâmetros da câmera calculados são utilizados em AUVs para exploração e navegação do fundo do oceano através da técnica de localização e mapeamento simultâneo (SLAM - *Simultaneous Localisation and Mapping*). Outro trabalho que utiliza a técnica de Zhang é desenvolvido em Meline, Triboulet e Jouvencel (2010) para realizar uma reconstrução da arqueologia submarina. Ao contrário de Eustice et al., o modelo de câmera considerado é o mesmo proposto por Zhang, considerando distorção apenas como dois coeficientes radiais.

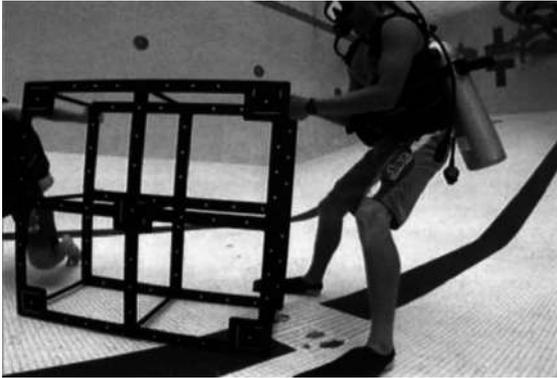
Em Rahman et al. (2013) é apresentada uma pesquisa sobre a diferença da calibração de câmera no ambiente terrestre e submerso em água. A calibração é realizada capturando imagens de um padrão circular 3D em formato de cubo, como pode ser visto na Figura 12b. A fim de validar o algoritmo proposto, são realizados comparações entre os resultados obtidos com os apresentados pela técnica de Heikkila e Silven. A comparação efetuada entre os algoritmos de calibração é realizada através do erro de reconstrução dos pontos 3D, onde o algoritmo proposto apresenta melhores resultados. Rahman et al. afirmam que, para reconstrução de objetos no ambiente submerso, é necessário realizar a calibração no mesmo meio, devido a diferença que os ambientes exercem sobre o modelo.

Com a finalidade de reduzir os esforços de uma calibração subaquática, Oleari et al. propuseram um método que não utiliza um padrão de calibração submerso (OLEARI et al., 2015). Os autores apresentam um método para calibrar um sistema estéreo de câmeras dividido em duas etapas. Na primeira etapa é realizada uma calibração padrão no ar. Na segunda etapa, os parâmetros da câmera são otimizados usando o conhecimento prévio sobre a forma de tubos cilíndricos submersos, como os dutos usados para extração de gás e petróleo. Essa otimização desempenha uma correção automática, principalmente da distorção radial, com base na retificação das superfícies cilíndricas. Oleari et al. testaram o algoritmo proposto com dutos de diferentes configurações em uma profundidade de 10m.

Em Silvatti et al. (2012) é proposto um estudo comparativo entre três técnicas de calibração para um sistema de câmera estéreo, aplicado para análise de performance de atletas que praticam natação. As técnicas de calibração são: DLT modificado, não modelando explicitamente as características física da câmera (HATZE, 1988); Calibração não linear através do movimento de bastão, com objeto triédrico fixo para determinar o

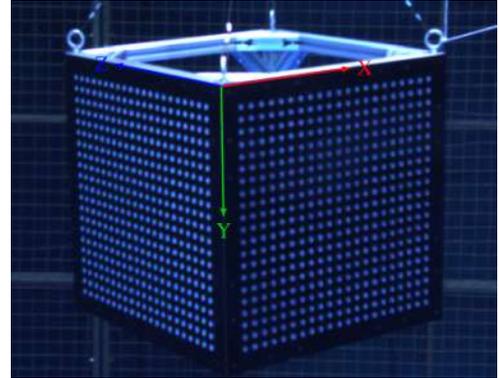
Figura 12 – Padrões utilizados em calibrações para ambientes subaquáticos: (a) Cubo aberto, fabricado com alumínio leve, de cor preta com 56 alvos circulares brancos usado por Shortis; (b) Padrão de calibração com dois planos de 21x21 pontos circulares utilizados por Rahman; (c) Padrão utilizado para a calibração com varinha; (d) Padrão planar formado por um tabuleiro de xadrez.

(a)



Fonte: Shortis (2015).

(b)



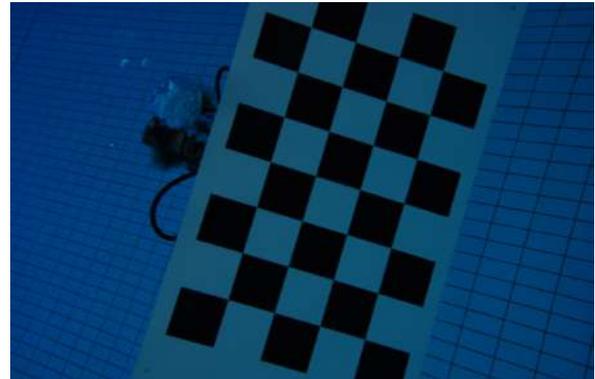
Fonte: Rahman et al. (2013).

(c)



Fonte: Silvatti et al. (2012).

(d)



Fonte: Jordt-Sedlazeck e Koch (2012).

sistema de coordenadas para a calibração (CERVERI; BORGHESE; PEDOTTI, 1998); Método proposto por Zhang (2000). Os objetos utilizados nos testes de calibração foram o bastão com a referência triédrica (Figura 12c), além de um padrão planar similar ao apresentado na Figura 12d. Silvatti et al. apresentaram vantagens e desvantagens de cada uma das técnicas executadas e concluíram que o método de Zhang provê melhores resultados, mas que a calibração com o movimento de bastão também pode ser utilizado. Posteriormente, Silvatti apresentou uma expansão deste estudo comparativo em sua dissertação (SILVATTI et al., 2013).

Alguns trabalhos de visão computacional para aplicações subaquáticas utilizam ferramentas de calibração de câmeras que, inicialmente, foram desenvolvidas para calibrações

terrestres. Brandou et al. utilizam a ferramenta desenvolvida por Jean-Yves Bouguet para estimar os parâmetros intrínsecos e a relação entre o par de câmeras, empregadas para executar a reconstrução 3D do ambiente através de uma trajetória predefinida (BRANDOU et al., 2007). Esta mesma técnica de calibração é utilizada por Han Hu, em sua dissertação (HU, 2015), para desenvolver um algoritmo de correspondência quase-denso de características, a fim de robustificar a reconstrução de ambientes subaquáticos. Assim como o trabalho de Brandou et al., o algoritmo de reconstrução de Hu utiliza um sistema de câmeras estéreo. Enquanto Onmek et al. propõem um algoritmo de reconstrução 3D subaquática, para apenas uma câmera, de artefatos arqueológicos que utiliza uma técnica de calibração baseada nos métodos de Bouguet e Zhang (ONMEK et al., 2017).

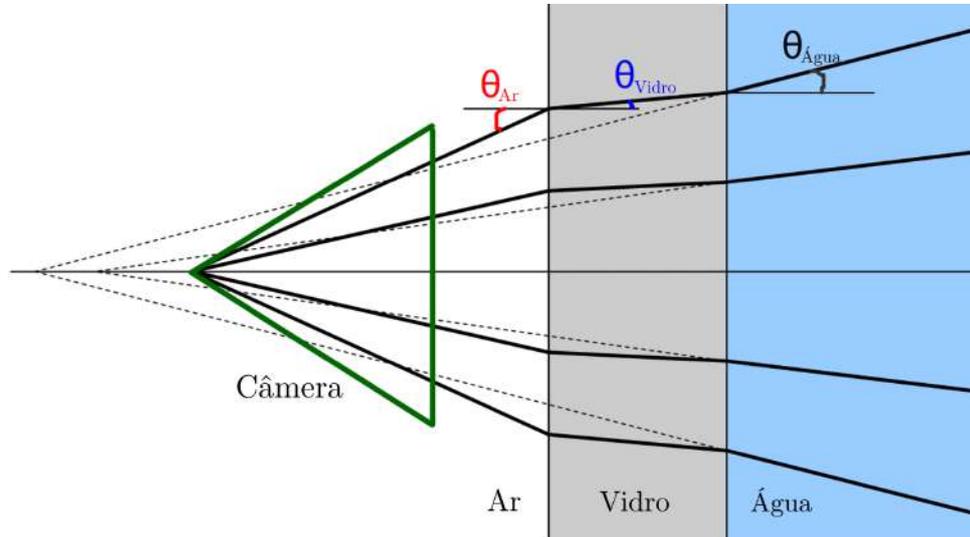
3.2 CÂMERA COM MÚLTIPLOS PONTOS DE VISTA (NON-SVP)

Essa abordagem considera que o modelo de uma câmera subaquática não pode dispor apenas de um centro de projeção. É defendido que, devido as refrações ocasionadas nas interfaces água-vidro e vidro-ar da caixa de estanque, uma câmera subaquática deve possuir diferentes centros de projeções para a formação de uma única imagem. A Figura 13 apresenta a ideia que motivou a criação desse modelo com múltiplos pontos de vista para câmeras subaquáticas. Trabalhos como Treibitz et al. (2012) e Sedlazeck e Koch (2012) criticam o uso de modelo SVP para o cenário subaquático, afirmando que sempre haverá alguns erros sistemáticos, devido aos efeitos refrativos, não incorporados ao modelo. Apesar de melhorar a corretude do modelo, essa abordagem aumenta a complexidade do processo geométrico de correção, demandando mais tempo das aplicações de visão computacional que as utilizam. Isso ocorre porque, efetivamente, essa abordagem baseia-se na simulação do trajeto que os raios de luz percorreriam no mundo real.

Um modelo matemático para imagens subaquáticas obtidas através de uma única interface planar é introduzido em Chari e Sturm (2009). Ao formular a refração dos raios na interface, Chari e Sturm derivaram uma matriz fundamental refrativa de 12×12 que correlaciona duas câmeras perspectivas que compartilham a mesma superfície refrativa. Também é derivada a relação entre um plano da cena 3D e sua imagem após a refração, através de homografia refrativa. Os autores descreveram como calcular, em circunstância especial, a pose relativa entre duas câmeras subaquáticas. Porém, nenhum resultado experimental foi apresentado. Como não foi encontrado qualquer continuação desse trabalho, seus resultados permanecem como considerações teóricas de valor conceitual.

Em contrapartida, Gedge et al. desenvolveram um método de calibração, baseado nos mesmos princípios físicos apresentado por Chari e Sturm, com ênfase na computação dos resultados, em vez do comportamento teórico deles (GEDGE; GONG; YANG, 2011). O método apresenta uma solução para múltiplas câmeras com interfaces de refração independentes, calculando os vetores normais das câmeras e a da distância referente à câmera principal do sistema estéreo em relação à interface. O algoritmo identifica pontos carac-

Figura 13 – Caminhos dos raios de luz incidentes no centro de uma câmera fixada em caixa de estanque com interface de vidro planar. As linhas tracejadas demonstram os raios provenientes da água sem considerar a refração entre as interfaces ar-vidro e vidro-água, exemplificando que o modelo de câmera não deve ter um único centro de projeção.



Fonte: Adaptado de Sedlazeck e Koch (2011).

terísticos do padrão de calibração correspondentes nas imagens geradas pelas câmeras e minimiza o erro epipolar, através da técnica de otimização não linear de Levenberg-Marquardt. Nesse trabalho, a interface de vidro é considerada muito fina e, por isso, é modelada apenas a refração ocasionada pela mudança do meio água-ar. Um estudo mais aprofundado deste trabalho é apresentado na dissertação de Gedge (2011).

Um método similar de calibração estéreo é proposto por Anne Sedlazeck e Reinhard Koch em (SEDLAZECK; KOCH, 2011). Diferente de Gedge et al., os autores modelam a interface da caixa de estanque com duas refrações (água-vidro e vidro-ar), retiram a necessidade de inserção de um padrão de calibração e modificam a função de erro a ser minimizada, adicionando o conceito de projeções em perspectiva de câmeras virtuais. Possuindo uma sistema de câmera estéreo previamente calibrado com um dos métodos já estabelecidos para calibração no ar, características são extraídas nas imagens subaquáticas utilizando o algoritmo de detecção e descrição de características invariantes a escalas (SIFT - *Scale-Invariant Feature Transform*) (LOWE, 1999). As correspondências dos pontos entre as imagens são realizadas através da técnica de aproximação de vizinho mais próximo (ANN - *Approximate Nearest Neighbor*) (ARYA et al., 1998), sendo robustificada com o uso do método de RANSAC para evitar combinações erradas de pontos. Esse algoritmo, que está disponibilizado na biblioteca de algoritmos básicos para imagem (BIAS - *Basic Image Algorithms*) (MIP, CAU Kiel, Germany, 2015), foi implementado em C/C++ e avaliado em ambiente sintético e real.

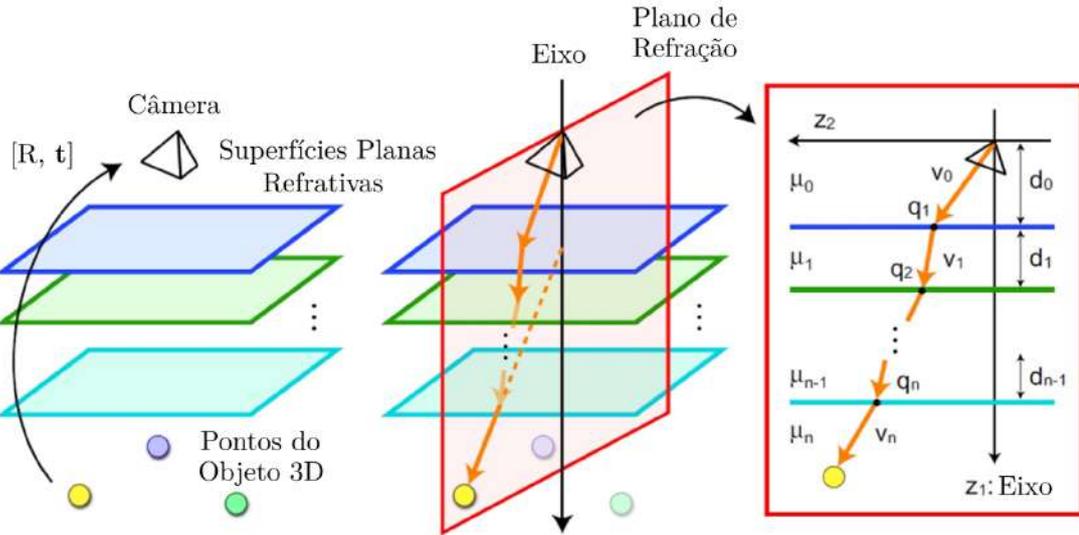
Gili Telem e Sagi Filin propuseram um modelo que determina um ponto no eixo óptico

da câmera para cada ponto da imagem, servindo como um novo centro de projeção virtual que relaciona as coordenadas 3D com os pontos 2D utilizando uma projeção perspectiva (TELEM; FILIN, 2010). Os parâmetros do modelo são determinados através de uma calibração dividida em duas etapas: a primeira, realiza uma calibração “seca” para estimar os parâmetros intrínsecos da câmera (distância focal; ponto principal da imagem; coeficientes de distorção de lente), enquanto a segunda etapa desempenha uma calibração “molhada” para estimar os outros parâmetros referentes a refração nas interfaces ar-vidro-água (a espessura do vidro, distância do centro de projeção da câmera para a interface, índice de refração relativo entre os meios, a inclinação e direção do eixo óptico da câmera em relação a normal na interface). Uma abordagem modificada negligencia a direção do eixo óptico e a espessura da interface, pois esses fatores podem ser facilmente absorvidos pelos parâmetros físicos da câmera. O modelo foi avaliado em ambiente sintético e real, através de testes realizados em piscina e mar aberto.

Agrawal et al. propõem um método de calibração genérico para câmeras subaquáticas baseando-se no modelo refrativo físico, pressupondo que uma câmera visualize o objeto através de múltiplas interfaces planas refrativas paralelas (AGRAWAL et al., 2012). Os autores salientam que o modelo de refração plana corresponde a uma câmera axial, ou seja, todos os raios de imagem passam por um eixo perpendicular às interfaces refrativas, denominado “eixo de refração”. Além disso, o caminho de um dado raio de luz está localizado em um “plano de refração”. O conjunto de todos esses planos se interceptam no eixo de refração, onde também está localizado o centro de câmera. Os autores usam essas observações para desenvolver o método de calibração, calculando os parâmetros de refração do modelo refrativo plano para múltiplas camadas através da resolução de equações lineares. O sistema de equação é obtido a partir das correspondências de pontos 2D-3D de um padrão planar, como apresentando na Figura 12d, em uma única imagem. Uma visão geral da ideia proposta por Agrawal et al. é apresentada na Figura 14.

Anne Jordt-Sedlazeck e Reinhard Koch apresentam um método de calibração baseado em parâmetros geométricos e radiométricos que estende o método de calibração proposto por Agrawal et al. (JORDT-SEDLAZECK; KOCH, 2012). Os autores descrevem o algoritmo dividindo-o nas etapas de inicialização, otimização geométrica e radiométrica. Na primeira etapa, inicialmente são determinados linearmente as poses da câmera em relação ao padrão de calibração, as distâncias das interfaces e suas normais através do método descrito por Agrawal et al. (2012). Em seguida, uma otimização não linear, no espaço 3D, das poses é realizada baseando-se nos pontos característicos do padrão de calibração, exemplificado em Figura 12d. A segunda etapa realiza uma otimização em todas as câmeras usando uma abordagem de análise por síntese dentro de uma rotina de otimização utilizando o algoritmo evolucionário CMA-ES (*Covariance Matrix Adaptation Evolution-Strategy*) (HANSEN; OSTERMEIER, 2001). Por último, os componentes radiométricos são computados, permitindo que esses parâmetros corrijam e aprimorem as cores das imagens

Figura 14 – Ilustração do método proposto em Agrawal et al. (2012).



Fonte: Adaptado de Agrawal et al. (2012).

subaquáticas. No entanto, eles não compararam seus resultados com o método linear de Agrawal et al. e seus resultados não possuem gabarito (*ground-truth*).

Em Dolereit e Kuijper (2014) é apresentado um estudo inicial para inferir os parâmetros de câmera subaquática a partir dos obtidos em uma calibração terrestre. Os autores consideram que a câmera subaquática pode ser representada por múltiplas câmeras virtuais que seguem o modelo *pinhole*, calculando novos centros de projeções ao longo do eixo óptico da câmera. Para realizar os cálculos, é considerado que as informações sobre os índices de refrações do vidro e da água, além da rotação e distância do centro da câmera para a interface e espessura do vidro, são conhecidos. O problema é que esses valores dificilmente são fornecidos com correteude pelos fabricantes e, por isso, apenas resultados sintéticos são apresentados nesse trabalho.

Em um trabalho mais recente, Dolereit et al. apresentam uma calibração para um sistema estéreo que compartilham uma única interface plana (DOLEREIT; LUKAS; KUIJPER, 2015). Os autores renunciam a proposta de inferir os parâmetros da câmera subaquática a partir de uma calibração seca, utilizando uma abordagem similar à proposta por Agrawal et al. (2012) para determinar as localizações virtuais dos pontos de objeto e calcular o eixo de refração. Já Yau, Gong e Yang (2013) estendem o trabalho apresentado em Agrawal et al. (2012) para considerar o efeito de dispersão da luz, desenvolvendo restrições nos parâmetros de calibração com a finalidade de melhorar a precisão dos cálculos.

Em Traffelet et al. (2016) é apresentada uma calibração onde o modelo utilizado para representar a refração existente na caixa de estanque assume que a espessura da interface e o índice de refração do ar são conhecidos, possuindo como parâmetros os índices de refração do vidro e água, a distância câmera-interface e os ângulos de inclinação do raio na interface de vidro. O procedimento de calibração é, inicialmente, realizado no ar, utilizando

um padrão de calibração convencional para calcular os parâmetros intrínsecos da câmera. Posteriormente, uma sequência de vídeo é capturada com a câmera submersa observando o objeto alvo em movimento. Utilizando essas informações, os parâmetros que modelam a refração da caixa de estanque são calculados através de uma minimização dos erros de reprojeção das observações do objeto de calibração no espaço de imagem.

Além disso, baseando-se na contribuição realizada por Yau et al., modelando o efeito de dispersão da luz, Xida Chen e Yee-Hong Yang apresentam em [Chen e Yang \(2017\)](#) uma solução fechada para o cálculo da distância entre a interface e o centro de projeção da câmera, uma vez que a normal da interface refrativa seja conhecida. Esse método proposto tem como principal vantagem o fato de não requerer a inserção de nenhum padrão de calibração, sendo capaz de encontrar uma solução única dos parâmetros de câmera.

3.3 CONSIDERAÇÕES FINAIS

Neste capítulo, foram discutidos os principais trabalhos encontrados na literatura que possuem relação direta com esta dissertação de mestrado. Explanou-se diferentes abordagens para calibração em ambientes subaquáticos, como a técnica da câmera com ponto de vista único (SVP), utilizando-se a calibração molhada ou seca com adaptação dos parâmetros, e a técnica com múltiplos pontos de vista (nSVP), sugerindo que o modelo de uma câmera subaquática não pode dispor de apenas um centro de projeção. Enquanto os trabalhos que utilizam a primeira abordagem valorizam a simplificação do modelo de câmera utilizado, realizando aproximações para obter uma calibração válida para o meio subaquático, os da segunda abordagem aumentam a complexidade para incrementar a acurácia dos cálculos para determinar os parâmetros da câmera.

Esta dissertação abordará o problema de calibração subaquática priorizando a redução da complexidade do modelo, através de uma técnica de adaptação que utiliza apenas as informações obtidas por uma calibração seca realizada previamente. Priorizando, também, o incremento da acurácia dos parâmetros adaptados em relação a técnica existente na literatura. Com o intuito de proporcionar os conceitos essenciais para a compreensão da proposta desta dissertação, o capítulo seguinte aponta a notação utilizada ao longo deste trabalho e os fundamentos básicos de geometria projetiva, modelo e calibração de câmera.

4 CSAT - TÉCNICA PROPOSTA

Este capítulo apresenta, em detalhes, a técnica proposta de adaptação dos parâmetros de câmera, calculados em um determinado meio, para serem utilizados em diferentes ambientes. Apesar da técnica ser genérica, este trabalho de mestrado ressalta o seu uso para uma aplicação comumente utilizada em visão computacional, onde uma câmera calibrada em ambiente terrestre é operada em cenários subaquáticos. Essa adaptação tem como principal objetivo determinar novos parâmetros que descrevam a formação da imagem no ambiente subaquático, a partir dos parâmetros obtidos na calibração seca, sem necessitar submergir a câmera. Inicialmente, na [Seção 4.1](#), é apresentada uma visão geral sobre a técnica para, em seguida, na [Seção 4.2](#), expor a formulação geral sobre a perturbação ocasionada no ponto devido a mudança no índice de refração do meio externo e interno à câmera. Com base nessa formulação, o algoritmo CSAT é descrito detalhadamente na [Seção 4.3](#), concluindo com as considerações finais do capítulo na [Seção 4.4](#).

4.1 VISÃO GERAL

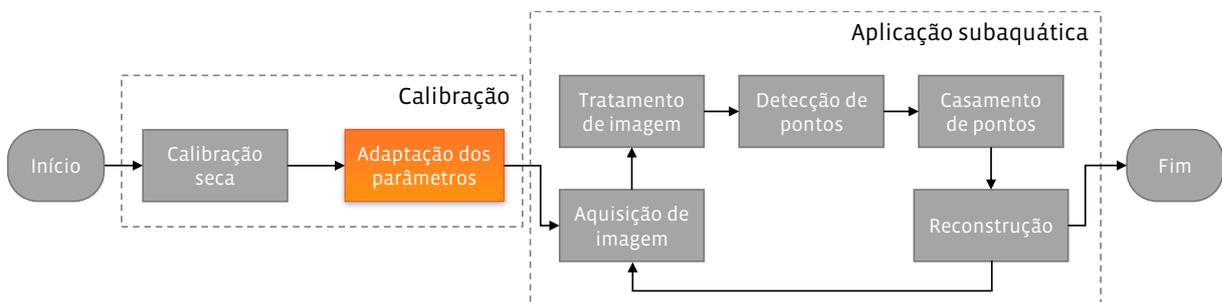
Como foi apresentado no [Capítulo 3](#), na literatura existe uma grande diversidade de trabalhos propostos com a finalidade de calibrar câmeras para ambientes subaquáticos. Essa ampla quantidade de técnicas existe devido a uma relação de custo-benefício (*trade-off*) entre a acurácia e complexidade da estimativa dos parâmetros que descrevem a câmera. Enquanto alguns trabalhos valorizam a diminuição da complexidade, realizando aproximações para obter uma calibração válida para o meio subaquático, outros prezam pelo aumento, com a intuito de melhorar o cálculo dos parâmetros da câmera. Esse acréscimo na complexidade engloba tanto a realização da calibração submersa no ambiente de aplicação, uma vez que é essencial submergir todos os aparatos necessários para o processo, quanto o enriquecimento do modelo de descrição da câmera, visto que esse fato promove modificações nos algoritmos tradicionais de visão computacional, além de, normalmente, necessitar de pelo menos uma calibração seca e outra molhada para realizar a estimativa dos parâmetros.

Esta dissertação aborda o problema de calibração para câmeras a serem utilizadas em ambiente subaquático, priorizando a diminuição da complexidade necessária para a técnica proposta, conseqüentemente reduzindo as das aplicações. Para tal, foi desenvolvido um método de adaptação dos parâmetros obtidos através da realização de uma calibração seca, onde não se faz necessário capturar imagens subaquáticas. Essa abordagem é semelhante à proposta em [Lavest, Rives e Lapresté \(2003\)](#), porém, sem aumentar a dificuldade do procedimento, prioriza-se também o incremento da acurácia dos parâmetros adaptados. Para esse fim, desenvolveu-se um modelo que considera as perturbações ocasionadas, em

decorrência da variação do meio em que a câmera está inserida, durante o processo de formação das imagens.

Com a finalidade de evidenciar como a técnica proposta é utilizada, a [Figura 15](#) apresenta um diagrama que ilustra o passo a passo de uma aplicação de visão computacional que realiza a reconstrução 3D, através de um sistema estéreo de câmeras, no ambiente subaquático. Essa aplicação descrita é a mesma utilizada nos experimentos de validações no ambiente submarino, além de ser bastante similar ao usado nos cenários sintéticos e tanque de água, apresentados no [Capítulo 5](#). A caixa destacada em laranja representa o estágio da técnica de adaptação proposta. Os dois estágios iniciais do sistema são responsáveis pela determinação dos parâmetros que descrevem o comportamento da câmera no ambiente subaquático, sendo utilizados como entrada para os subsequentes. Os demais estágios representam a aplicação de reconstrução subaquática e são executados de forma cíclica.

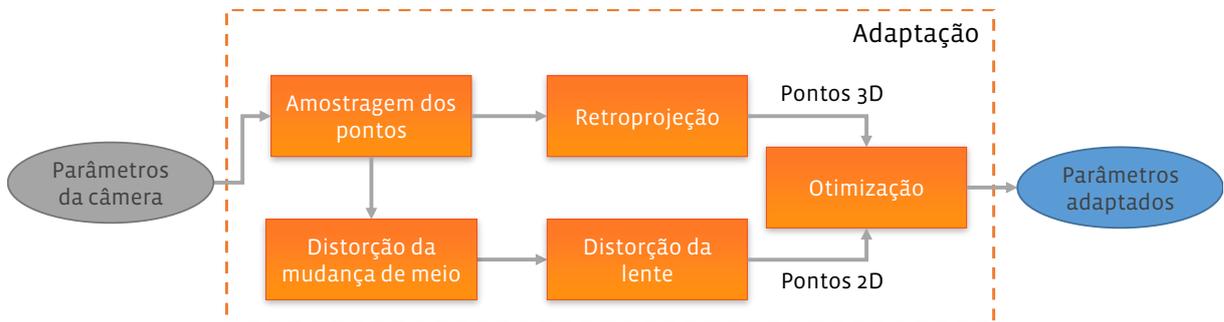
Figura 15 – Exemplo do diagrama de execução para uma aplicação subaquática que utiliza o algoritmo CSAT. A técnica proposta é empregada após uma calibração seca, antes de iniciar a aplicação subaquática. Esse fluxo é o mesmo utilizado nos experimentos realizados no ambiente submarino apresentado no [Capítulo 5](#).



Fonte: Elaborada pelo próprio autor.

Uma visão geral da técnica de adaptação dos parâmetros que descrevem a câmera para um meio diferente é apresentada na [Figura 16](#). Basicamente, o procedimento proposto consiste em, dado o índice de refração η_1 e os parâmetros calculados pela calibração no meio 1, além do índice de refração η_2 do meio em que a câmera será inserida, determinar um conjunto de pontos 3D e seus correspondentes 2D, distorcido pelos efeitos da mudança de meio e da lente, para, utilizando um algoritmo de otimização, calcular as variáveis do modelo que descrevem matematicamente a câmera no meio 2. Mais detalhes sobre o algoritmo proposto serão fornecidos na [Seção 4.3](#), porém, primeiramente será aprofundado como foi determinada a distorção causada na formação da imagem devido a mudança de meio no caminho de luz do ambiente onde a câmera está submersa para o seu centro de projeção.

Figura 16 – Fluxograma da técnica CSAT proposta para adaptar os parâmetros de câmera para ambientes com diferentes índices de refração.



Fonte: Elaborada pelo próprio autor.

4.2 DISTORÇÃO OCACIONADA PELA MUDANÇA DE MEIO

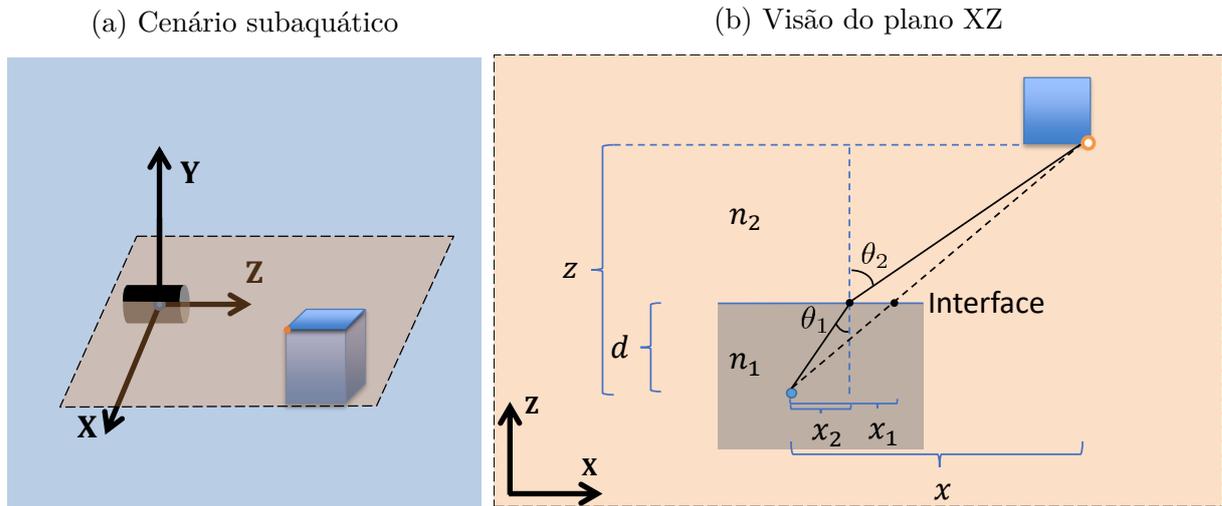
A técnica proposta nessa dissertação, denominada CSAT, para adaptar os parâmetros de uma calibração obtidos no meio 1 para o meio 2, precisa considerar a perturbação causada na formação da imagem devido a refração na interface entre os diferentes meios. Esta seção descreverá a formulação do problema, assim como os cálculos realizados, para determinar a distorção inserida pela refração na formação da imagem.

4.2.1 Formulação do problema

A fim de um melhor entendimento sobre o problema, foi examinado um cenário tridimensional, onde uma câmera, com centro de projeção sendo representado por um ponto azul, inserida em uma caixa de estanque, simbolizada por um cilindro preto, visualiza um cubo azul, como apresentado na [Figura 17a](#). Para efeito de simplificação, mas sem perder a generalidade do problema, devido a simetria em relação ao eixo óptico da câmera (Z), considerou-se apenas o espaço bidimensional, correspondente ao plano XZ . Na visão do plano XZ , exposta na [Figura 17b](#), o cubo passa a ser representado como um quadrado, assim como a caixa de estanque, que possui dimensões maiores para auxiliar a explanação do que ocorre em seu interior. Essa visão 2D evidencia uma versão simplificada do esquemático que representa o problema de distorção ocasionado pela mudança de meio.

Analisando a [Figura 17b](#), é possível verificar a diferença entre os raios de luz representados pela linha contínua e tracejada. O caminho contínuo representa o percurso obtido do ponto laranja até o centro de projeção da câmera, refratando na interface, enquanto o tracejado consiste na rota trilhada pela luz do ponto até o centro de projeção da câmera, considerando que todos os elementos estão situados no mesmo meio, igual ao do interior da caixa de estanque. O caminho tracejado é obtido pela calibração ocorrida no meio 1, sendo a representação do caminho de luz que produz a formação do ponto na imagem sem considerar o efeito de distorção aplicado pela lente, ou seja, demonstra a retroprojeção do ponto de imagem após a remoção, descrita na [Equação 2.11](#), da distorção causada pela

Figura 17 – Cenário exemplificando uma câmera subaquática visualizando um cubo submerso em um meio diferente ao que está inserida. Sendo apresentada uma visão 3D do cenário (a) e do plano XZ (b), simplificando e expondo o caminho percorrido pelo raio de luz, do objeto ao sensor, quando considerado a mudança de meio, admitindo os índices de refração η_1 e η_2 , (linha preta contínua) e considerando que objeto e sensor estivessem no mesmo ambiente, com índice de refração η_1 , (linha preta tracejada).



Fonte: Elaborada pelo próprio autor.

lente.

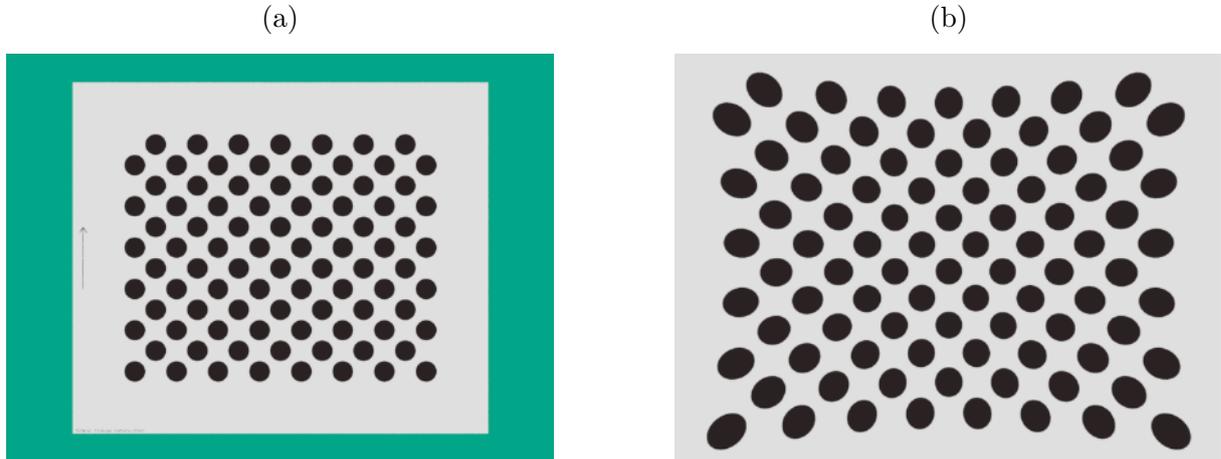
Resumindo, a definição do problema para a versão simplificada deve determinar o ponto x_2 , distorcido pela mudança de meio, tal que

$$x_2 = \lambda(x_1), \quad (4.1)$$

onde x_1 é o ponto após a remoção da distorção de lente, considerando todo o meio com índice de refração η_1 , e λ a função que deve ser determinada para aplicar distorção no ponto quando a caixa de estanque é submersa no meio com índice de refração η_2 .

O efeito de perturbação existente devido a mudança de meio é apresentado na [Figura 18](#), onde um objeto fixo distanciado por 250 mm da câmera é visualizado nos meios com índices de refração de 1,0 e 1,333 para representar o ar e a água, respectivamente. A próxima seção demonstrará os cálculos realizados para obter a distorção ocasionada pela mudança de meio no processo de formação de imagem. Para tal, os índices de refração dos meios envolvidos no processo de adaptação, η_1 e η_2 , devem ser conhecidos e consideraremos que a distância, d , do centro de projeção da câmera para a interface é muito menor que a distância, z , do centro de projeção para o objeto de interesse.

Figura 18 – Diferença entre a imagem obtida por uma câmera, sem distorção de lente, no meio 1 com $\eta_1 = 1,0$ (a) e no meio 2 com $\eta_2 = 1,333$ (b). Imagens geradas através da ferramenta de modelagem 3DS Max para comparar a visualização do objeto fixo a uma distância de 250 mm do centro de projeção e com interface da caixa de estanque a 30 mm.



Fonte: Elaborada pelo próprio autor.

4.2.2 Função de distorção obtida pela variação de meio

Começaremos definindo os pontos u_1 e u_2 como sendo os pontos provenientes das interseções dos raios de luz, projetados pela câmera no plano de imagem, considerando o objeto presente nos meios 1 e 2 com a interface da caixa de estanque, respectivamente. Esses pontos são definidos como

$$u_1 = \frac{f x_1}{d} + c_x \quad \text{e} \quad u_2 = \frac{f x_2}{d} + c_x, \quad (4.2)$$

onde f e d são as distâncias focal da câmera e do centro de projeção para a interface, respectivamente, e c_x é o deslocamento horizontal do ponto principal da imagem. A partir da [Equação 4.2](#), podemos determinar a posição do ponto na interface, através da retro-projeção dos pontos de imagens, como

$$\begin{aligned} x_1 &= \frac{d u_1}{f} - \frac{d c_x}{f} \\ &= \frac{d}{f} (u_1 - c_x), \end{aligned} \quad (4.3)$$

para o ponto formado sem considerar o meio em que o objeto está submerso, e

$$\begin{aligned} x_2 &= \frac{d u_2}{f} - \frac{d c_x}{f} \\ &= \frac{d}{f} (u_2 - c_x), \end{aligned} \quad (4.4)$$

para o ponto produzido com a distorção ocasionada pela mudança do meio na propagação da luz do objeto para a câmera.

Para a resolução do problema, iniciaremos pela lei de refração, ou também conhecida como Lei de Snell, que correlaciona os segmentos de retas dos meios envolvidos, descrevendo a relação entre os ângulos de incidência, θ_2 , e de refração, θ_1 , quando refere-se ao caminho de luz entre dois meios isotrópicos diferentes. Essa lei é expressa matematicamente por

$$\eta_1 \operatorname{sen}(\theta_1) = \eta_2 \operatorname{sen}(\theta_2). \quad (4.5)$$

Para utilizar a [Equação 4.5](#), necessitamos determinar os ângulos envolvidos na operação. Através da relação trigonométrica, e considerando a [Equação 4.4](#), obtém-se o ângulo de refração como

$$\begin{aligned} \theta_1 &= \arctan\left(\frac{x_2}{d}\right) \\ &= \arctan\left(\frac{u_2 - c_x}{f}\right). \end{aligned} \quad (4.6)$$

Enquanto o ângulo de incidência é definido por

$$\theta_2 = \arctan\left(\frac{x - x_2}{z - d}\right), \quad (4.7)$$

sendo x e z as distâncias horizontal e vertical, respectivamente, do centro de projeção para o ponto do objeto. Através da semelhança de triângulo, e utilizando a [Equação 4.3](#), podemos definir a distância horizontal como

$$\begin{aligned} x &= \frac{z x_1}{d} \\ &= \frac{z(u_1 - c_x)}{f}. \end{aligned} \quad (4.8)$$

Aplicando as equações [4.4](#) e [4.8](#) na definição do ângulo de incidência, expandimos a [Equação 4.7](#) para

$$\theta_2 = \arctan\left(\frac{\frac{z(u_1 - c_x)}{f} - \frac{d(u_2 - c_x)}{f}}{z - d}\right). \quad (4.9)$$

Uma vez que a distância do centro de projeção da câmera para a interface da caixa de estanque, em geral, é muito menor do que a distância para os objetos alvo das aplicações ($z \gg d$), consideramos

$$k = \frac{z}{d}, \quad (4.10)$$

onde k é um escalar que, devido ao pressuposto definido anteriormente, tende a infinito, isto é, $k \rightarrow \infty$. Substituindo a relação da [Equação 4.10](#) na [Equação 4.9](#) e considerando que

$$\bar{x}_1 = \frac{u_1 - c_x}{f} \quad \text{e} \quad \bar{x}_2 = \frac{u_2 - c_x}{f}, \quad (4.11)$$

são a representação do ponto normalizado nos meios 1 e 2, respectivamente, obtemos

$$\begin{aligned} \theta_2 &= \lim_{k \rightarrow \infty} \arctan \left(\frac{k d \bar{x}_1 - d \bar{x}_2}{k d - d} \right) \\ &= \arctan \left(\lim_{k \rightarrow \infty} \frac{k \bar{x}_1 - \bar{x}_2}{k - 1} \right). \end{aligned}$$

Como a fração do limite apresenta uma indeterminação, aplica-se o Teorema de L'Hôpital, também denominada de Regra de Cauchy, derivando o numerador e denominador da fração, para posteriormente, definir o limite da expressão. Dessa forma, finalmente concluímos que

$$\begin{aligned} \theta_2 &= \arctan \left(\lim_{k \rightarrow \infty} \frac{\bar{x}_1}{1} \right) \\ &= \arctan(\bar{x}_1). \end{aligned} \quad (4.12)$$

A partir da [Equação 4.11](#) e substituindo as definições para os ângulos de incidência e de refração apresentadas nas equações [4.6](#) e [4.12](#), respectivamente, na Lei de Snell, expandimos a [Equação 4.5](#) para

$$\eta_1 \operatorname{sen}(\operatorname{arctg}(\bar{x}_2)) = \eta_2 \operatorname{sen}(\operatorname{arctg}(\bar{x}_1)),$$

onde, após aplicar a relação trigonométrica que define

$$\operatorname{sen}(\operatorname{arctg}(\theta)) = \frac{\theta}{\sqrt{1 + \theta^2}},$$

desenvolve-se

$$\begin{aligned} \eta_1 \frac{\bar{x}_2}{\sqrt{1 + \bar{x}_2^2}} &= \eta_2 \frac{\bar{x}_1}{\sqrt{1 + \bar{x}_1^2}} \\ \frac{\bar{x}_2^2}{1 + \bar{x}_2^2} &= \left(\frac{\eta_2}{\eta_1} \right)^2 \frac{\bar{x}_1^2}{1 + \bar{x}_1^2} \\ \bar{x}_2^2 + \bar{x}_2^2 \bar{x}_1^2 &= \left(\frac{\eta_2}{\eta_1} \right)^2 \bar{x}_1^2 + \left(\frac{\eta_2}{\eta_1} \right)^2 \bar{x}_1^2 \bar{x}_2^2 \\ \bar{x}_2^2 (\bar{x}_1^2 - \left(\frac{\eta_2}{\eta_1} \right)^2 \bar{x}_1^2 + 1) &= \left(\frac{\eta_2}{\eta_1} \right)^2 \bar{x}_1^2. \end{aligned}$$

Para encontrar o ponto normalizado que pertence ao caminho que percorre os dois meios, \bar{x}_2 , em função do ponto normalizado presente na trajetória que considera apenas o meio em que a câmera está inserida, \bar{x}_1 , define-se

$$\bar{x}_2 = \left(\frac{\eta_2}{\eta_1} \left(\frac{1}{\sqrt{\bar{x}_1^2 \left(1 - \left(\frac{\eta_2}{\eta_1} \right)^2 \right) + 1}} \right) \right) \bar{x}_1. \quad (4.13)$$

Substituindo-se a [Equação 4.11](#) em [Equação 4.13](#), transformando os pontos para coordenadas de imagem, obtemos que

$$u_2 = \left(\frac{\eta_2}{\eta_1} \left(\frac{1}{\sqrt{\left(\frac{u_1 - c_x}{f} \right)^2 \left(1 - \left(\frac{\eta_2}{\eta_1} \right)^2 \right) + 1}} \right) \right) (u_1 - c_x) + c_x \quad (4.14)$$

é a expressão que aplica distorção ocasionada pela refração em um ponto u_1 . Devido a simetria em torno do eixo óptico da câmera, podemos generalizar a [Equação 4.14](#), que considera apenas a dimensão horizontal da imagem, para a versão 2D, admitindo que \mathbf{c}_0 descreve o ponto principal da imagem, $[c_x, c_y]^\top$, temos que

$$\mathbf{u}_2 = \left(\frac{\eta_2}{\eta_1} \left(\frac{1}{\sqrt{\left(\frac{\|\mathbf{u}_1 - \mathbf{c}_0\|}{f} \right)^2 \left(1 - \left(\frac{\eta_2}{\eta_1} \right)^2 \right) + 1}} \right) \right) (\mathbf{u}_1 - \mathbf{c}_0) + \mathbf{c}_0. \quad (4.15)$$

4.3 ALGORITMO

Após descrever a modelagem da distorção ocorrida devido a variação no índice de refração do meio, essa seção descreve o algoritmo de calibração proposto para adaptar os parâmetros de câmera, calculados no meio 1, a fim de descreverem o comportamento da câmera em ambientes, meio 2, com diferentes índices de refração. A ideia geral do algoritmo consiste em: dada a calibração realizada no meio 1, amostra-se pontos 2D, ao longo de toda a imagem gerada pela câmera, e os seus pontos 3D correspondentes. Sabendo que a localização dos pontos 3D independe do meio em que estão presentes, mas que os pontos 2D variam devido a função de distorção do meio, modelada anteriormente, determina-se os pontos 2D com distorção e realiza-se um novo procedimento de calibração utilizando alguns dos algoritmos tradicionais na literatura.

O pseudocódigo da implementação do algoritmo de adaptação proposto é apresentado no [Algoritmo 1](#). Como requisito, são necessários os parâmetros da câmera no meio 1 (\mathbf{K}_1 , \mathbf{a}_1 , \mathbf{p}_1), informações sobre as dimensões das imagens geradas (largura w e altura h), além da relação entre os índices de refração (η) do meio 1 e do 2. Resumidamente, a técnica consiste em cinco etapas:

Algoritmo 1: Pseudocódigo do algoritmo de adaptação proposto.

Entrada: \mathbf{K}_1 , \mathbf{a}_1 , \mathbf{p}_1 , w , h , n

Saída: \mathbf{K}_2 , \mathbf{a}_2 , \mathbf{p}_2

```

1 inicio
2    $\{\mathbf{p}_i\} \leftarrow \text{AMOSTRAGEM}(w, h, s)$ 
3   para cada  $\mathbf{p}_i \in \{\mathbf{p}_i\}$  fazer
4      $\mathbf{X}_i \leftarrow \text{RETROPROJEÇÃO}(\mathbf{p}_i, \mathbf{K}_1)$ 
5      $\mathbf{u}_i \leftarrow \text{DISTORÇÃOMEIO}(\mathbf{p}_i, \mathbf{K}_1, n)$ 
6      $\mathbf{u}_i \leftarrow \text{DISTORÇÃOLENTE}(\mathbf{u}_i, \mathbf{a}_1, \mathbf{p}_a)$ 
7   fin
8    $\mathbf{K}_2, \mathbf{a}_2, \mathbf{p}_2 \leftarrow \text{OTIMIZAÇÃO}(\{\mathbf{X}_i\}, \{\mathbf{u}_i\}, \mathbf{K}_1, \mathbf{a}_1, \mathbf{p}_1)$ 
9 fin

```

4.3.1 Etapa 1: Amostragem de pontos

Nessa etapa, pontos são uniformemente amostrados ao longo de todo o espaço da imagem gerada pela câmera, com dimensões conhecidas por largura w e altura h . Para tal, é utilizado um coeficiente de amostragem, s , que determina a distância existente entre cada ponto produzido. Portanto, a técnica emprega um total de

$$a = \frac{w h}{s} \quad (4.16)$$

amostras geradas no conjunto $\{\mathbf{p}_i\}$ e que serão utilizadas nas demais etapas do algoritmo. A [Figura 19a](#) apresenta, em uma visão bidimensional do cenário, a distribuição proporcional das amostragens ao longo do plano de projeção da câmera.

4.3.2 Etapa 2: Retroprojeção

Cada ponto amostrado \mathbf{p}_i , onde $i = 1 \dots a$, na etapa anterior é retroprojetado para o espaço tridimensional, obtendo-se o vetor que representa o caminho percorrido pelo raio de luz, através de

$$\mathbf{L}_i = \mathbf{K}_1^{-1} \mathbf{p}_i, \quad (4.17)$$

originado no centro de projeção da câmera e que intercepta o ponto amostrado no plano de projeção. Essa câmera é representada pelo modelo *pinhole*, sem considerar os efei-

tos de distorção, contendo os parâmetros K_1 determinados pela calibração seca efetuada previamente.

A representação gráfica dessa etapa é indicada na [Figura 19b](#), onde cada reta laranja demonstra o caminho de luz percorrido no espaço 3D para produzir o ponto amostrado no plano de imagem. Devido a dualidade ponto-reta existente na geometria projetiva, conforme descrito em [Seção 2.2](#), podemos considerar que o vetor L_i gerado através de retroprojeção da amostra no plano de imagem representa o ponto 3D X_i .

4.3.3 Etapa 3: Adição da distorção do meio

Empregando os parâmetros intrínsecos do modelo *pinhole* da câmera (f e \mathbf{c}_0) previamente calculados através de uma calibração no meio 1, a terceira etapa tem como objetivo incorporar aos pontos amostrados \mathbf{p}_i a perturbação ocorrida devido a refração, existente na interface que divide os diferentes meios envolvidos, na formação da imagem. Para tal, como discutido na [Seção 4.2](#), todos os pontos amostrados são deslocados utilizando a [Equação 4.15](#), aplicando um fator de distorção para cada uma das amostras, como apresentado na [Figura 19c](#).

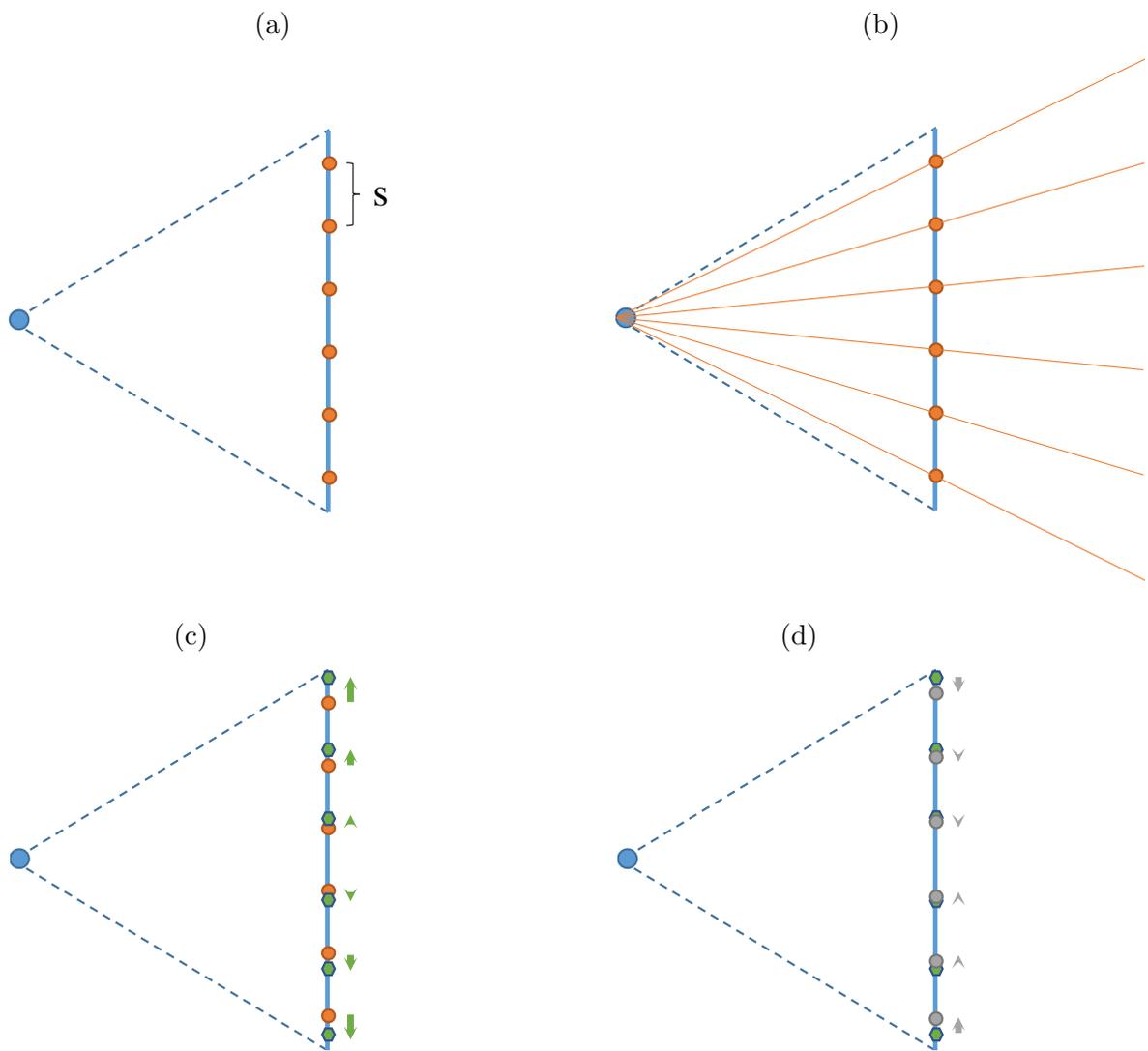
Analisando a [Equação 4.15](#), verificamos que o fator de distorção, definido por

$$\left(\frac{\eta_2}{\eta_1} \left(\frac{1}{\sqrt{\left(\frac{\|\mathbf{p}_i - \mathbf{c}_0\|}{f} \right)^2 \left(1 - \left(\frac{\eta_2}{\eta_1} \right)^2 \right) + 1}} \right) \right), \quad (4.18)$$

pode assumir três casos particulares. O primeiro, quando $\eta_2 = \eta_1$, o fator possui valor igual a 1, ou seja, como os meios envolvidos possuem o mesmo índice de refração, nenhuma perturbação é aplicada aos pontos. Nas demais configurações, a [Equação 4.18](#) apresenta uma característica radial, devido ao termo $\|\mathbf{p}_i - \mathbf{c}_0\|$, produzindo um aumento ou diminuição do fator, a medida que \mathbf{p}_i encontra-se mais afastado do centro da imagem. Na segunda configuração, onde $\eta_2 > \eta_1$, o resultado apresentado pelo fator possui valor mínimo quando $\mathbf{p}_i = \mathbf{c}_0$, aumentando à medida que o ponto se distancia do centro. Enquanto a última configuração, em que $\eta_2 < \eta_1$, o fator diminui conforme \mathbf{p}_i distancia-se de \mathbf{c}_0 .

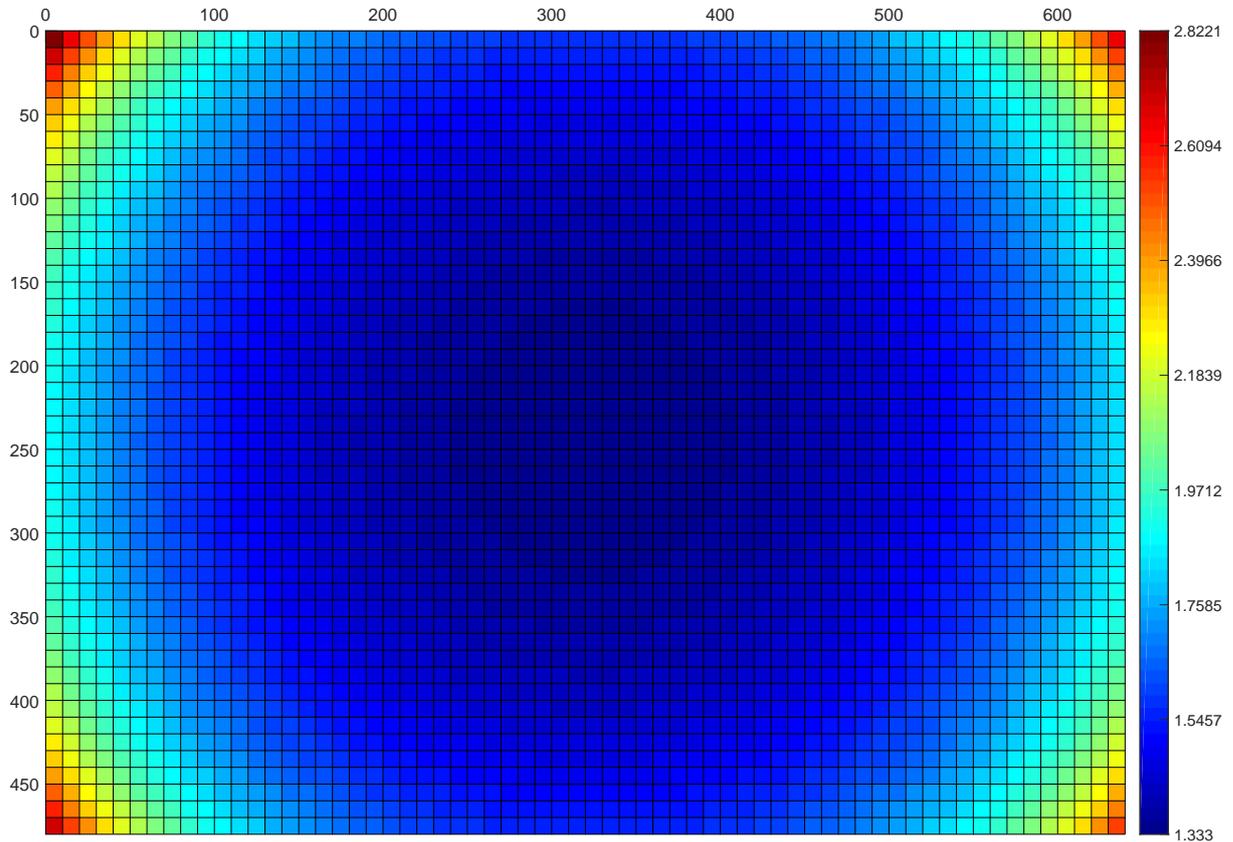
A segunda configuração será detalhada, uma vez que os cenários normais de aplicações utilizando câmera submersa possui $\eta_2 > \eta_1$, ou seja, o índice de refração do meio onde a câmera está imersa é menor do que o recinto onde está inserida. Para analisarmos o comportamento do fator de distorção devido a refração entre os meios, utilizaremos um exemplo hipotético onde uma câmera, que possui $f = 400$ e $\mathbf{c}_0 = [320, 240]^T$, está inserida em uma caixa de estanque preenchida por ar ($\eta_1 = 1,0$) que é submergida no meio

Figura 19 – Representação gráfica das quatro primeiras etapas da algoritmo CSAT. Amostragem de pontos laranjas no plano de imagem (a), retas laranjas retroprojetadas através das amostras (b), distorção causada pela refração (c), produzindo os pontos, representados pelo hexágono verde, para, posteriormente, aplicar a perturbação ocorrida devido a lente da câmera (d), gerando os pontos em cinza. As setas, apresentadas em (c) e (d), identificam as diferentes intensidades de distorção aplicadas aos pontos.



Fonte: Elaborada pelo próprio autor.

Figura 20 – Mapa de calor dos valores que representam o fator de distorção, ocasionado pela mudança de meio, em cada ponto amostrado na imagem. Nesse exemplo, foram considerados os meios mais comuns nas aplicações, câmera no ar ($\eta_1 = 1,0$) sendo submersa no meio aquático ($\eta_2 = 1,333$). A configuração da câmera utilizada foi de $f = 400$ e $c_0 = [320, 240]^T$, produzindo imagem com dimensão de $w = 640$ e $h = 480$, sendo amostrado um ponto a cada 10 *pixels* ($s = 10$).



Fonte: Elaborada pelo próprio autor.

aquático ($\eta_2 = 1,333$). A imagem gerada por essa câmera possui dimensão de 640×480 , sendo amostrada uniformemente a cada 10 *pixels*.

Os fatores obtidos para cada ponto amostrado são apresentados na [Figura 20](#) através de um mapa de calor, onde os pontos que possuem menor fator são representados pela cor azul, valores intermediários pela cor verde e os de maiores medidas pela cor vermelha. Como pode ser visualizado, os pontos centrais possuem valores menores, assemelhando-se a aproximação proposta em (LAVEST; RIVES; LAPRESTÉ, 2003), onde o fator aplicado é igual a η_2/η_1 , que nesse cenário é 1,333. A medida em que os pontos amostrados distanciam-se radialmente do centro da imagem, o fator de distorção aumenta consideravelmente, alcançando valores próximos a 2,8 nos extremos.

4.3.4 Etapa 4: Adição da distorção de lente

Após aplicar a distorção causada pela refração ocasionada na mudança de meio, a quarta etapa é responsável por adicionar a perturbação produzida na formação da imagem devido a construção da lente da câmera e ao alinhamento dos seus elementos internos. Os coeficientes dos polinômios que descrevem como remover esse tipo de distorção (representados por \mathbf{a} e \mathbf{p}) são previamente calculados durante a calibração realizada no meio 1, conforme o modelo apresentado na [Subseção 2.3.3](#). A representação gráfica desse processo é exposta, em uma visão 2D, na [Figura 19d](#).

Para cada ponto obtido na etapa anterior, uma perturbação é adicionada através de uma aproximação da função inversa a apresentada na [Equação 2.11](#), definida para remover a distorção. Como uma solução analítica dessa inversão não é trivial, recomenda-se o uso de soluções numéricas para o problema, aplicando assim um método não linear de mínimo quadrados como o de Gauss-Newton, para determinar o ponto distorcido.

4.3.5 Etapa 5: Otimização dos parâmetros do modelo

Dispondo dos pontos 3D $\{\mathbf{X}_i\}$, calculados durante a etapa de retroprojeção das amostras obtidas na imagem, e os pontos 2D $\{\mathbf{u}_i\}$, determinados após a inserção das distorções da refração e lente da câmera, essa etapa tem por finalidade determinar os parâmetros que descrevem o comportamento da câmera no meio em que será submersa. Como o modelo utilizado não é linear, não há uma forma analítica capaz de definir as suas variáveis. Por esse motivo, é necessário empregar um processo iterativo de otimização não-linear para determinar os parâmetros do modelo.

A inicialização do modelo que descreve a câmera no ambiente submerso, descrito por \mathbf{K}_2 , \mathbf{a}_2 e \mathbf{p}_2 , é realizada considerando os valores obtidos no meio 1. Posteriormente, esses parâmetros são refinados otimizando-se

$$\arg \min_{\mathbf{K}_2, \mathbf{a}_2, \mathbf{p}_2} \sum_i^n d(\delta(\mathbf{u}_i, \mathbf{a}_2, \mathbf{p}_2), \mathbf{K}_2 \mathbf{X}_i), \quad (4.19)$$

onde d é a função de distância euclidiana.

A [Equação 4.19](#) corresponde ao problema de ajuste não linear dos parâmetros de câmera no espaço de imagem sem distorção, podendo ser resolvida através do algoritmo de otimização de Levenberg-Marquardt. Os parâmetros \mathbf{K}_2 , \mathbf{a}_2 e \mathbf{p}_2 são alterados buscando minimizar o erro da comparação entre os pontos projetados ($\mathbf{K}_2 \mathbf{X}_i$) em relação aos obtidos após a remoção da distorção dos pontos de imagem ($\delta(\mathbf{u}_i, \mathbf{a}_2, \mathbf{p}_2)$), conforme descrito na [Subseção 2.3.3](#).

Uma variação da técnica CSAT também foi proposta, sendo denominada de CSAT-M. Essa modificação tem por objetivo diminuir a quantidade de parâmetros otimizados, realizando uma mescla entre a adaptação apresentada por Lavest et al. e o algoritmo CSAT, proposto nesta dissertação. Para tal, apenas os parâmetros de distorção (\mathbf{a}_2 e \mathbf{p}_2)

são refinados, sendo o K_2 obtido através da multiplicação do foco da calibração do meio 1 pelo fator η_2/η_1 e a manutenção do ponto principal da imagem $[u_0, v_0]^T$, assim como descrito em (LAVEST; RIVES; LAPRESTÉ, 2003).

4.4 CONSIDERAÇÕES FINAIS

Esse capítulo apresentou o algoritmo de adaptação proposto CSAT e sua variação CSAT-M, com a finalidade de evitar o processo de calibração em ambientes submersos. Tal objetivo foi alcançado com o intuito de evitar procedimentos mais complexos, como uso de calibração em ambientes que possam oferecer restrições e interferir na imagem, aumentando assim o tempo e a dificuldade para a determinação de parâmetros que irão descrever o comportamento da câmera submersa. Diferenciando-se das adaptações apresentadas na literatura, este trabalho forneceu uma aproximação para descrever a perturbação causada através da distorção introduzida pela mudança de meio no caminho do raio de luz no ambiente submerso.

No [Capítulo 5](#) serão apresentadas análises e comparações dos algoritmos propostos, CSAT e CSAT-M, com duas técnicas da literatura.

5 EXPERIMENTOS E ANÁLISE DOS RESULTADOS

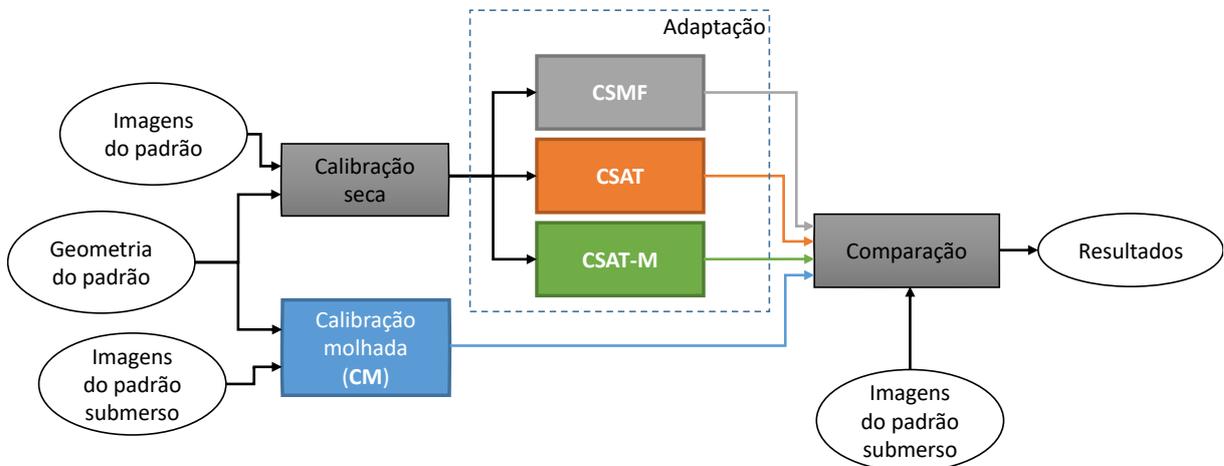
Após a investigação de trabalhos correlatos e a discussão teórica sobre a técnica de adaptação da calibração para ambiente subaquático, este capítulo descreve a parte experimental desta dissertação. A [Seção 5.1](#) expõe as configurações gerais dos experimentos realizados, como informações de hardware, da implementação e metodologia empregada para avaliação comparativa entre os algoritmos. Na [Seção 5.2](#), as técnicas propostas, CSAT e CSAT-M, são avaliadas em diferentes cenários sintéticos com a finalidade de verificar a influência de alguns fatores inerentes às câmeras. A [Seção 5.3](#) relata experimentos em ambientes reais de aplicação subaquática, sendo realizadas validações em um tanque de água, relatados na [Subseção 5.3.1](#), comparando os algoritmos através da reconstrução de um objeto com geometria conhecida, enquanto a [Subseção 5.3.2](#) detalha os testes realizados em um cenário real de operação no ambiente submarino. Por fim, a [Seção 5.4](#) apresenta as argumentações finais do capítulo.

5.1 CONFIGURAÇÕES DOS EXPERIMENTOS

As calibrações realizadas no escopo desta dissertação, tanto a seca quanto molhada, baseiam-se no algoritmo de [Heikkila e Silven \(1997\)](#), aumentando a quantidade de coeficientes que representam a distorção de lente da câmera. Esse aumento foi realizado devido à grande distorção presente nas lentes utilizadas pelas câmeras nos cenários sintéticos e reais. A técnica proposta para adaptar uma calibração seca, determinando novos parâmetros que modelam a câmera no ambiente subaquático, é denominada como CSAT, que significa calibração seca com adaptação de todos os parâmetros.

Durante os experimentos foram desempenhadas comparações entre a técnica proposta e os algoritmos de calibração molhada (CM), calibração seca com multiplicação do foco (CSMF) proposta por [Lavest, Rives e Lapresté \(2003\)](#) e um terceiro algoritmo criado a partir da combinação entre CSAT e CSMF. Esse terceiro algoritmo, denominado como CSAT-M, realiza a multiplicação do parâmetro de distância focal e refina os coeficientes de distorção para obter uma calibração válida para ambiente subaquático, como definido no [Capítulo 4](#). Um resumo do fluxo utilizado para a comparação desses algoritmos é ilustrado na [Figura 21](#), onde as cores utilizadas para representar cada técnica são as mesmas encontradas nos gráficos de resultados apresentados pelos experimentos. As calibrações molhada e seca necessitam de imagens de um objeto com geometria previamente conhecida no ambiente subaquático e terrestre, respectivamente. Esse objeto é comumente denominado de padrão de calibração. Detalhes sobre materiais e métodos de avaliações comuns entre os experimentos serão relatados nas próximas subseções.

Figura 21 – Fluxograma da comparação entre as técnicas de calibração para o ambiente subaquático.



Fonte: Elaborada pelo próprio autor.

5.1.1 Implementação

O computador utilizado para a execução dos algoritmos possui um processador Intel Core i7-3960X com 3,30 GHz, 24GB de memória RAM e placa gráfica NVIDIA GeForce GTX 560 Ti com 1GB de memória dedicada. Os testes foram executados sobre o sistema operacional Windows, instalado na versão 10 com 64 bits. Todos os algoritmos utilizados neste trabalho, tanto para calibrar as câmeras como para adaptar os parâmetros para ambientes subaquáticos, foram implementados na linguagem C/C++ utilizando como plataforma de desenvolvimento o *software* Visual Studio 2015.

O código desenvolvido emprega as bibliotecas Ceres Solver (AGARWAL; MIERLE et al., 2010), Eigen (GUENNEBAUD; JACOB et al., 2010) e OpenCV (BRADSKI, 2000) como suporte. Ceres Solver é uma biblioteca C++ de código aberto usada para modelar e resolver problemas grandes e complexos no contexto de otimização. Na implementação, essa biblioteca forneceu suporte para otimizar os parâmetros durante a calibração e a adaptação através do algoritmo de Levenberg-Marquardt. O Eigen também é uma biblioteca C++, que possui suporte e algoritmos de álgebra linear, sendo utilizada na implementação para representar matrizes, vetores e suas operações. Além disso, essa biblioteca serve de suporte para a Ceres Solver. Por fim, a biblioteca OpenCV foi usada para detectar os pontos do padrão de calibração nas imagens capturadas e fornecer inicialização para os parâmetros do modelo de câmera.

5.1.2 Método de avaliação comparativa

A avaliação de técnicas para calibração de câmeras é uma questão que merece uma maior discussão. Na literatura não há um consenso acerca de qual a melhor métrica para comparar diferentes calibrações, principalmente se houver divergência entre os modelos adotados

por essas técnicas. Alguns trabalhos utilizam a distância entre os pontos 3D projetados pelo modelo e seus respectivos pontos 2D detectados pela câmera como métrica de análise, calculando a raiz do valor quadrático médio (RMS - *Root Mean Square*) dessas distâncias. Porém, essa métrica é muito dependente do modelo e, principalmente, do conjunto de imagens fornecidas para a avaliação. Outro problema de utilizar o RMS é que muitos autores utilizam nos experimentos comparativos o mesmo conjunto de imagens usadas para a calibração, ou seja, verificam apenas qual o algoritmo que melhor minimiza este erro no espaço de imagem.

Outra abordagem usual para validação das técnicas de calibração é utilizar um algoritmo de reconstrução 3D, usando pelo menos duas imagens com pontos de vista diferentes, de um objeto com geometria previamente conhecida. Essa geometria fornece o *ground-truth* necessário para a elaboração da métrica de comparação. O objeto proveniente da reconstrução é comparado com o *ground-truth*, normalmente verificando-se as posições de seus pontos característicos. A diferença obtida nessa comparação entre a geometria esperada e a calculada é denominada erro de reconstrução. Como o resultado da calibração influencia diretamente na qualidade da reconstrução, que é um dos algoritmos de maior importância para aplicativos de visão computacional, é coerente utilizá-la como o método de avaliação entre as técnicas de calibração.

Cada ponto reconstruído \mathbf{X}_i , possui um erro de reprojeção associado, definido por

$$\varepsilon_i = d(\boldsymbol{\delta}(\mathbf{x}_i, \mathbf{a}, \mathbf{p}), \mathbf{K}(\mathbf{R} \mathbf{X}_i + \mathbf{t})), \quad (5.1)$$

onde a função $\boldsymbol{\delta}$, definida na [Equação 2.11](#), remove a distorção do ponto de imagem \mathbf{x}_i (correspondente ao ponto 3D \mathbf{X}_i) através dos parâmetros de distorção radial (\mathbf{a}) e tangencial (\mathbf{p}) do modelo. Por outro lado, a função d calcula a distância euclidiana entre o ponto de imagem sem distorção e o ponto 3D projetado. Os coeficientes da matriz \mathbf{R} e o vetor \mathbf{t} , que representam a orientação e posição do objeto em relação à câmera, são calculados em conjunto com os pontos 3D $\{\mathbf{X}_i\}$, otimizando-se

$$\arg \min_{\mathbf{R}, \mathbf{t}, \mathbf{X}} \sum_i \varepsilon_i^2. \quad (5.2)$$

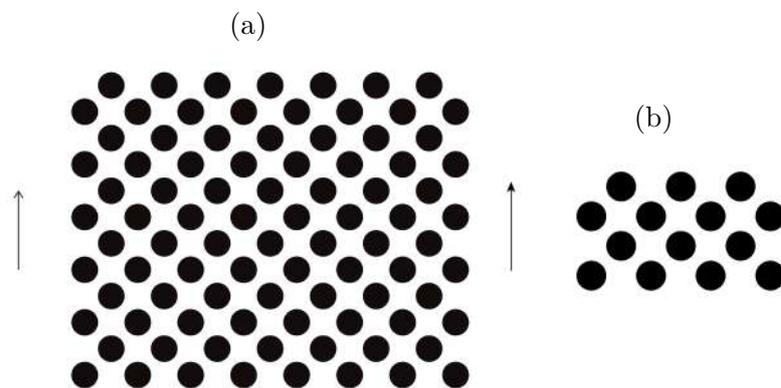
O conjunto de pontos $\{\mathbf{X}_i\}$ do objeto de validação não possuem informação sobre a escala da cena, uma vez que a reconstrução foi realizada sem nenhuma informação métrica do objeto. Por isso, conhecendo as distâncias entre a distribuição dos pontos característicos reconstruídos, uma correção no fator de escala é realizada para, posteriormente, calcular o erro de reconstrução através da distância 3D entre os valores de $\{\mathbf{X}_i\}$ e o *ground-truth* fornecido sobre o objeto de validação.

Na [Seção 5.2](#) e [Subseção 5.3.1](#) foram utilizados um objeto, comumente encontrado em processos de calibração, com pontos circulares dispostos de maneira assimétrica sobre uma superfície plana. Esse objeto é formado por 14 pontos circulares pretos que, em

dimensões reais, possuem 2 cm cada ponto e estão localizados na região central de uma superfície plana branca com 24×18 cm de dimensão, dispendo de uma borda com 5 cm. No experimento do ambiente submarino, a reconstrução é realizada em um duto cilíndrico usado para extração de gás e petróleo e, por não possuir um *ground-truth* de sua geometria, a métrica utilizada não será o erro de reconstrução. Mais detalhes sobre as informações do experimento no ambiente submarino serão fornecidos na [Subseção 5.3.2](#).

Os procedimentos de calibração, seca e molhada, foram realizados usando um padrão de geometria igual ao de validação, porém com 90 pontos característicos e superfície de 40×34 cm, com os mesmos 5 cm de borda. Os objetos utilizados para a calibração e para a validação das técnicas são apresentados na [Figura 22a](#) e [Figura 22b](#), respectivamente.

Figura 22 – Padrões utilizados para calibração (a) e validação nos experimentos (b).



Fonte: Elaborada pelo próprio autor.

5.2 EXPERIMENTOS SINTÉTICOS

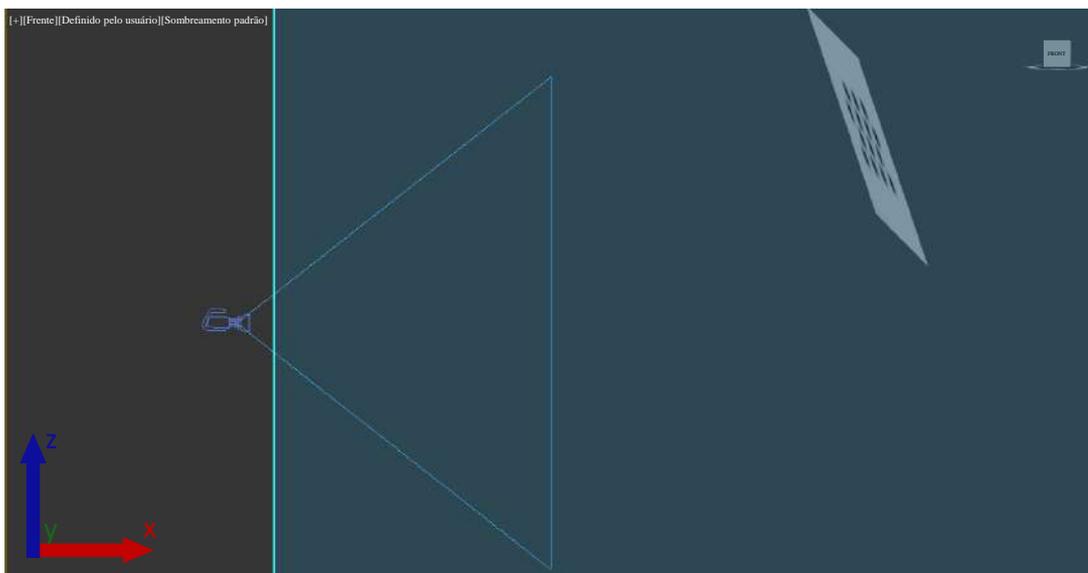
Todos os cenários sintéticos foram elaborados utilizando o *software* 3DS Max, na versão 2018 estudante ([AUTODESK, 2017](#)). Essa ferramenta de computação gráfica, desenvolvida pela Autodesk, permite a modelagem, renderização e animação de objetos tridimensionais, sendo comumente aplicada para criação de jogos, animações em filmes e comerciais gráficos, tal como em simulações e testes da mesma maneira que desenvolvida neste trabalho.

Os cenários elaborados possuem uma câmera, centrada na posição $[0,0,0]^T$ da cena, um cubo com dimensões extensas posicionado a frente da câmera, simulando um ambiente preenchido com água, além de um plano com a imagem do objeto empregado para calibrar ou validar as técnicas, posicionado no interior da caixa. A face do cubo mais próxima à câmera simula a interface de vidro da caixa de estanque, necessária para a submersão do equipamento, que fica entre os meios: ar e água. A câmera criada no 3DS Max é um dispositivo físico sem alvo direcionado e com campo de visão no ar definido com 93° de abertura horizontal, simulando uma máquina fotográfica grande angular. As imagens

foram produzidas usando o renderizador Mental Ray (LIVNY, 2008), disponível no software 3DS Max, com resolução de 640 de largura por 480 de altura. Os *pixels* das imagens geradas são perfeitamente quadrados, ou seja, relação de aspecto (largura e comprimento) é 1.

Na Figura 23 é possível visualizar os objetos que compõem os cenários de validação e calibração através de uma vista frontal da cena fornecida pelo 3DS Max, onde também pode ser observado o sistema de coordenadas empregado pela ferramenta. Os eixos x, y e z representam, respectivamente, a profundidade, deslocamento horizontal e vertical do objeto em relação à câmera e possuem as cores vermelho, verde e azul. Esse mesmo sistema de cores é utilizado para representar as translações e rotações realizadas em cada um dos eixos, como será apresentado nas descrições dos cenários desenvolvidos.

Figura 23 – Cenário elaborado na ferramenta 3DS Max 2018 para a validação das técnicas de calibração. Possui uma câmera que visualiza o objeto de validação no interior de um cubo que simula o meio aquático. Essa cena é semelhante à utilizada para a técnica CM, modificando apenas o padrão de calibração.



Fonte: Elaborada pelo próprio autor.

Diversas calibrações foram executadas durante os experimentos sintéticos. Esse fato ocorreu devido à necessidade de uma nova calibração seca e molhada cada vez que uma modificação era aplicada nos parâmetros da câmera, como a alteração do grau de distorção existente na lente. Também era necessário uma nova calibração molhada a cada deslocamento efetuado na caixa de estanque, uma vez que, como mencionado anteriormente, esse tipo de calibração assume o conjunto câmera-interface como um único objeto e qualquer deslocamento modifica esse conjunto.

O processo de calibração foi realizado capturando 36 imagens de posições e orientações diferentes do padrão de calibração. O movimento e orientação do padrão durante o procedimento de uma calibração seca é exemplificado na Figura 24a e Figura 24b, respectivamente. Basicamente, o padrão foi transladado com o intuito de ser amostrado no

centro e nas bordas das imagens formadas pela câmera. Esse deslocamento tornava-se maior a cada modificação aplicada para distanciar o padrão em relação à câmera. Aliadas às translações, foram aplicadas rotações no padrão em cada um dos eixos. Esses movimentos e rotações foram ajustados para a calibração molhada pelo fato de o objeto ser visualizado mais próximo à câmera.

Cada cenário de validação retratado nessa seção foi elaborado para explorar a influência de determinadas variáveis existentes na cena, como a distância da interface para a câmera, em cada uma das técnicas comparadas. As próximas subseções descreverão detalhadamente qual a variável explorada e a metodologia utilizada para obter os resultados apresentados em cada um dos experimentos sintéticos, assim como cada um dos detalhes que os diferem dos demais.

5.2.1 Cenário 01

O primeiro cenário foi produzido para verificar a interferência ocasionada pela variação da distância entre a interface e a câmera no erro de reconstrução do padrão. Mais precisamente, a análise foi efetuada através da relação entre as distâncias câmera-interface e câmera-objeto, uma vez que, quanto menor a diferença entre essas distâncias, menor a representatividade da aproximação realizada pela técnica de adaptação proposta. Recordando que, como definido no [Capítulo 4](#), a técnica de adaptação proposta nesse trabalho parte do pressuposto que a profundidade do objeto (z) em relação a câmera é muito maior que a distância câmera-interface (d), ou seja, $z \gg d$. Nesse cenário, uma nova calibração molhada foi exercida para cada deslocamento aplicado no posicionamento da interface. Nesse contexto, o experimento foi desenvolvido através de imagens obtidas com o objeto de validação fixo a uma distância de 500 mm e 1000 mm em relação a câmera.

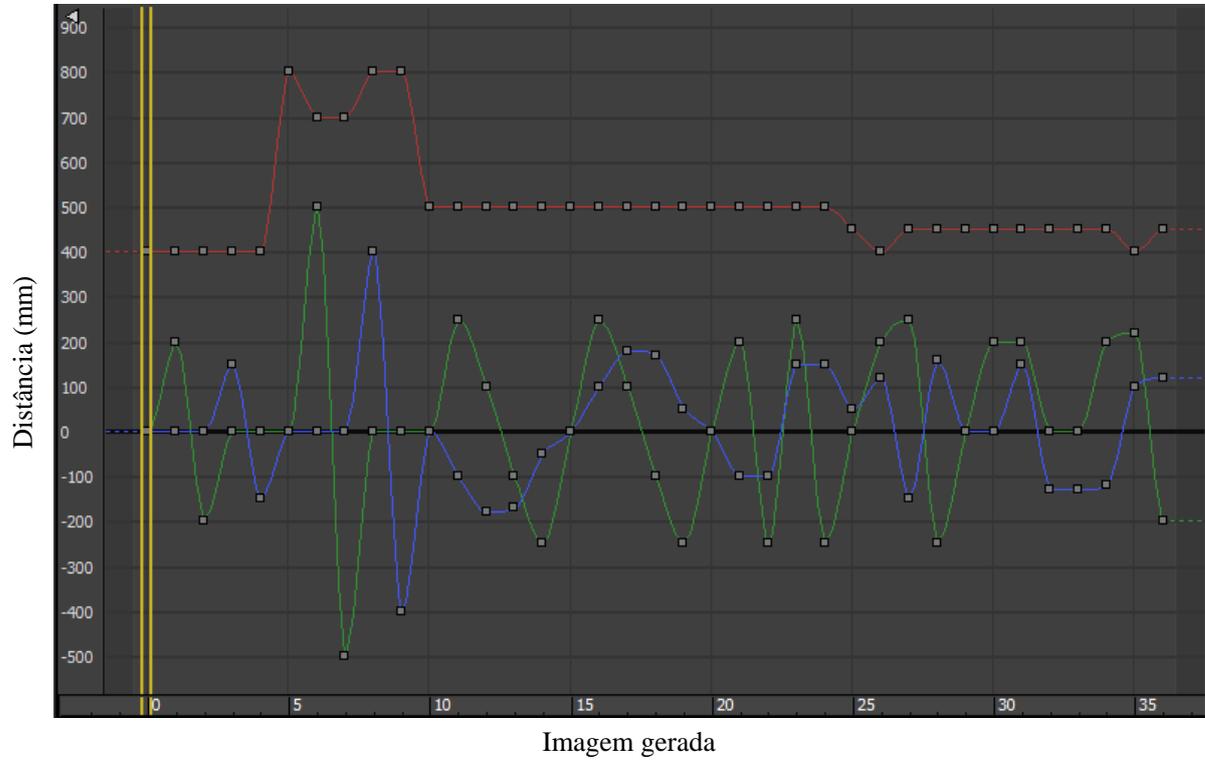
5.2.1.1 Objeto de validação fixo à 500 mm da câmera

A cena elaborada possui um conjunto de 56 imagens, sendo que a cada duas imagens geradas, utilizadas para uma reconstrução do objeto de validação, a interface é distanciada da câmera em 5 mm. Como apresentado na [Figura 25](#), a distância câmera-interface, que inicialmente é de 5 mm, é gradativamente aumentada até alcançar 140 mm. Nessa análise, o objeto de validação foi posicionado a 500 mm da câmera, o que, em termos percentuais, significa uma variação de 1% a 28% na relação entre as distâncias câmera-interface e câmera-objeto. Esses valores dificilmente são encontrados em cenários reais, porém a ideia foi extrapolar os valores reais para investigar o comportamento obtido pelas técnicas de calibração em cenários mais extremos, assim como verificar o efeito causado nas imagens geradas nesses casos.

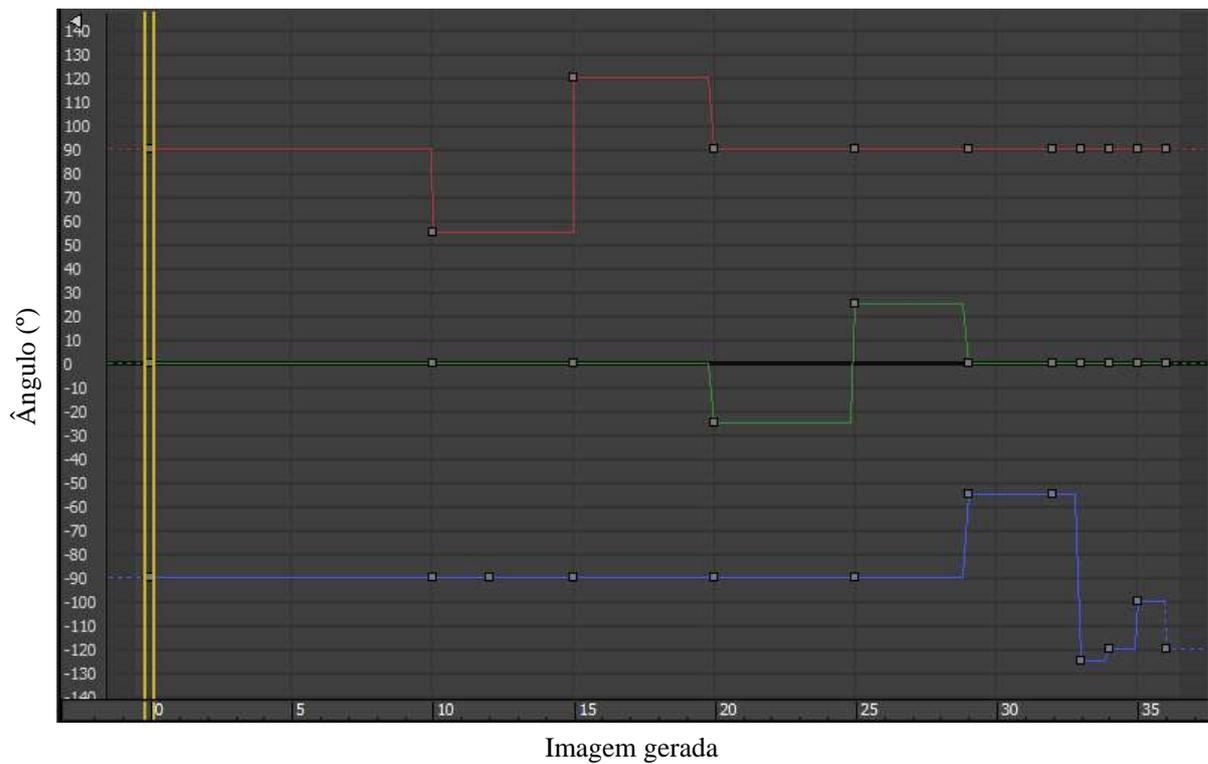
A pose do objeto de validação, apresentado na [Figura 26](#), consiste basicamente em uma mesma transformação sendo repetida a cada duas imagens. Para o objeto localizar-se no canto superior esquerdo e no canto inferior direito da imagem, é efetuada uma translação

Figura 24 – Amostragem do movimento (a) e orientação (b) do padrão em um dos cenários de calibração seca. As cores vermelha, verde e azul representam as translações e rotações, respectivamente, nos eixos X, Y e Z.

(a) Translação do padrão

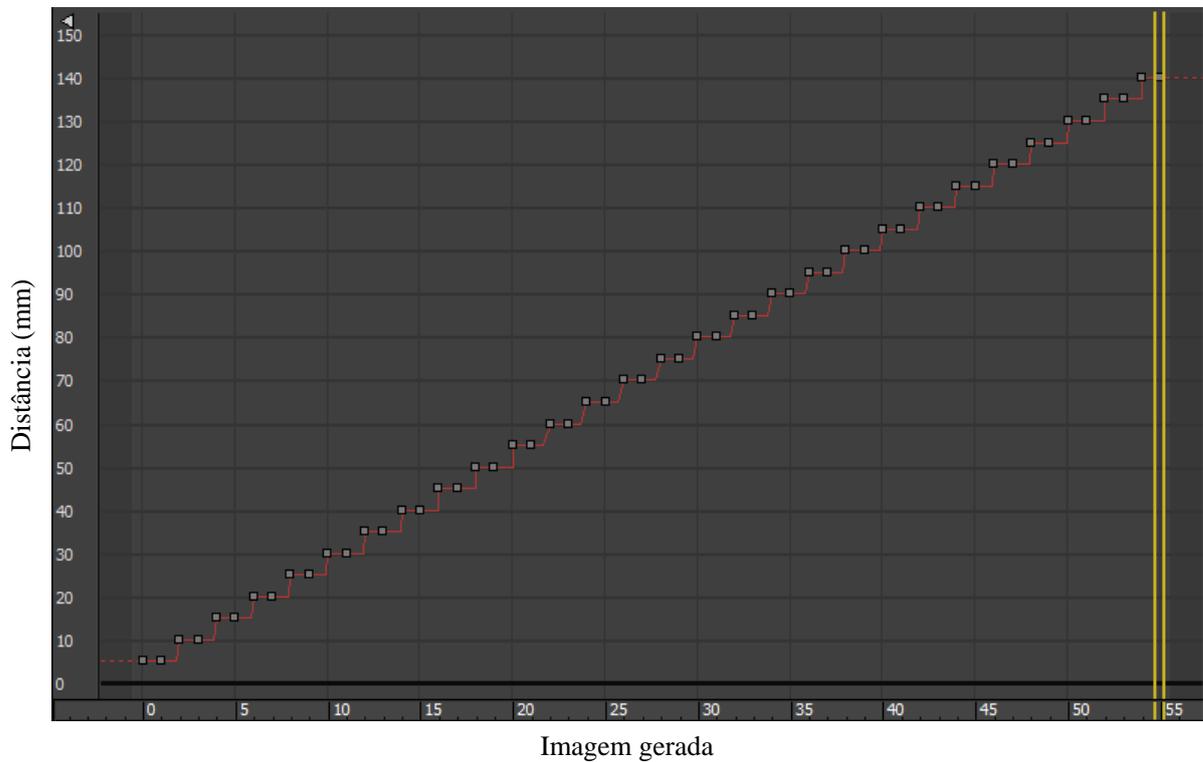


(b) Rotação do padrão



Fonte: Elaborada pelo próprio autor.

Figura 25 – Representação do 3DS Max sobre a distância entre a interface e a câmera, variando de 5 mm até 140 mm.



Fonte: Elaborada pelo próprio autor.

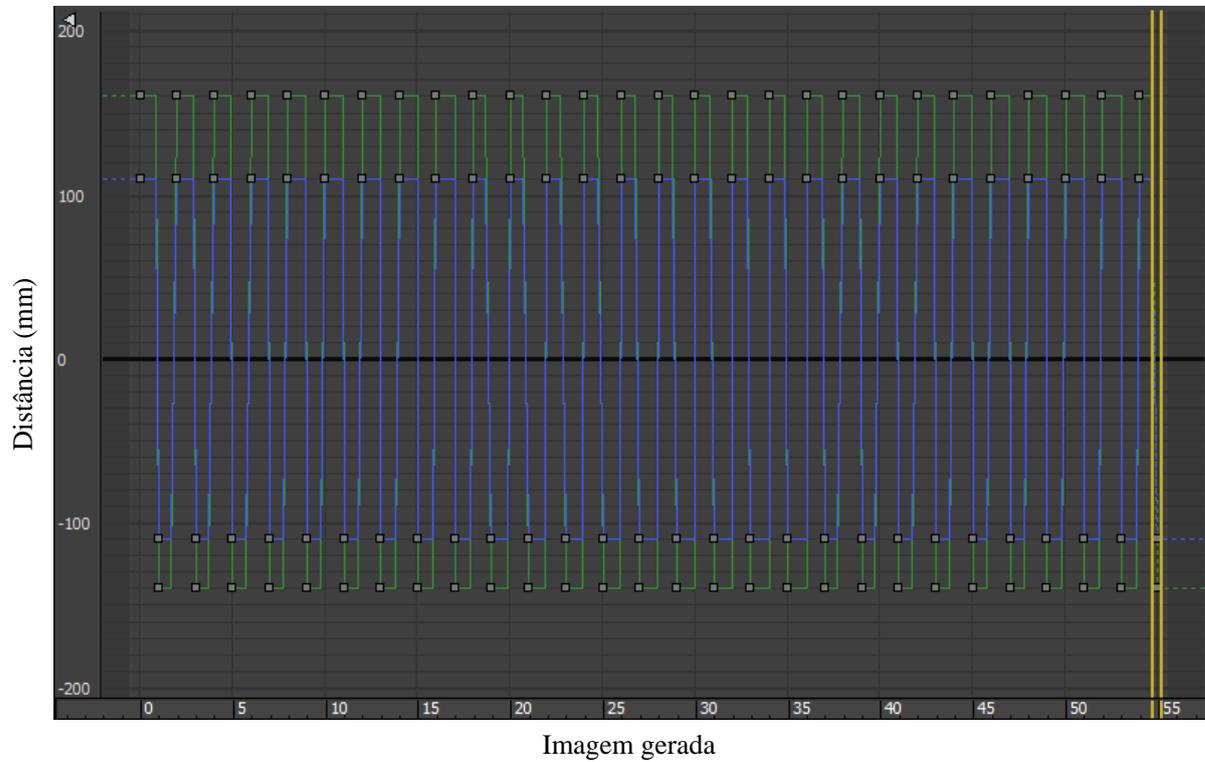
de 160 mm e -140 mm na direção do eixo y e em 110 mm e -110 mm no eixo z , como exibido na [Figura 26a](#). A fim de variar a orientação do padrão, o objeto é rotacionado nas duas direções em 20° no eixo x e 10° nos eixos y e z , como exposto na [Figura 26b](#).

Ainda que nesse cenário a lente da câmera não possuísse nenhum tipo de distorção, as imagens geradas encontram-se distorcidas devido à mudança de meio durante o caminho de luz do objeto, submerso em água, para a câmera, que habita o interior da caixa de estanque preenchida por ar. Esse fenômeno pode ser visualizado na [Figura 27](#), que apresenta algumas das imagens obtidas para a validação do cenário 01 com o objeto fixo a 500 mm. Cada coluna de imagem representa um conjunto utilizado para a reconstrução dos pontos presentes no objeto de validação. O erro médio de reconstrução obtido por conjunto é calculado comparando os pontos 3D adquiridos com os seus correspondentes no gabarito do padrão, previamente conhecido. Um fato interessante de se observar na [Figura 27](#) é que quanto mais próximo a interface está da câmera, mais distorcida encontra-se a borda da imagem.

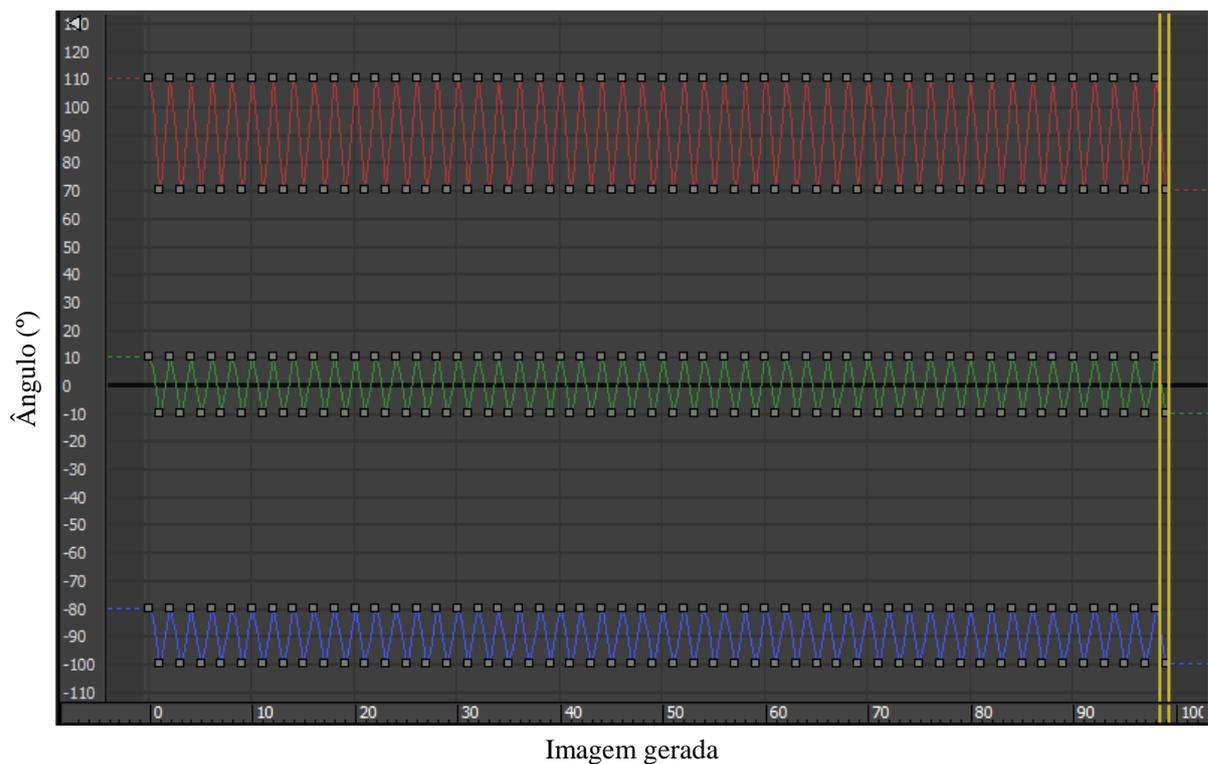
A [Figura 28](#) apresenta o efeito de distorção formado na imagem devido à mudança de meio e exemplifica a diferença gerada pelo deslocamento da interface. Uma visão geral é exposta na [Figura 28a](#) sobre a projeção de um ponto do objeto no plano de imagem em duas caixas de estanque diferentes, de linha contínua (interface 1) e linha pontilhada

Figura 26 – Representação do 3DS Max sobre o posicionamento (a) e orientação (b) do padrão no cenário 01 com distância para a câmera fixa em 500 mm. As cores vermelha, verde e azul representam as translações e rotações, respectivamente, nos eixos X, Y e Z.

(a) Translação do padrão

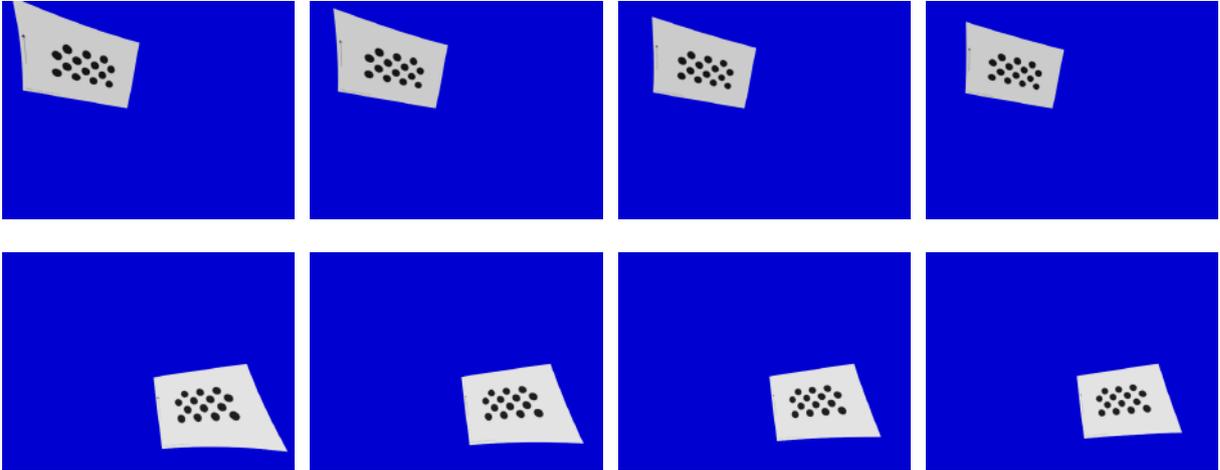


(b) Rotação do padrão



Fonte: Elaborada pelo próprio autor.

Figura 27 – Exemplos das imagens do conjunto de validação do cenário 01 com objeto fixo a 500 mm. Cada coluna refere-se a um conjunto usado para a reconstrução do objeto e possui, da esquerda para a direita, distância da interface para a câmera de 5 mm, 50 mm, 100 mm e 140 mm, respectivamente.



Fonte: Elaborada pelo próprio autor.

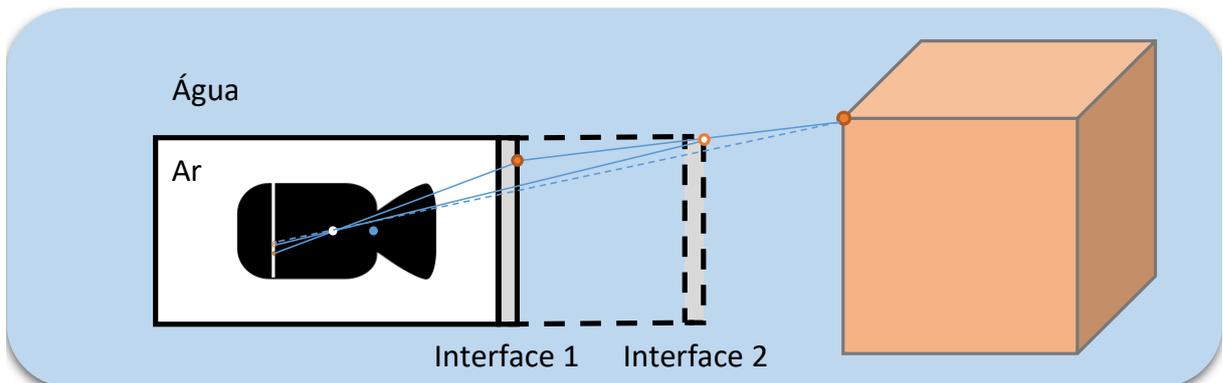
(interface 2), além de exibir a projeção do ponto, através do caminho de luz pontilhado, como se o objeto estivesse no ar. A fim de expor com mais detalhes a localização do ponto no plano de imagem, a [Figura 28b](#) destaca apenas a parte interna da câmera com os três pontos projetados no plano de imagem. O ponto em cinza corresponde à projeção do objeto como se estivesse no ar, enquanto os pontos em branco com borda laranja e o totalmente laranja são obtidos através do raio de luz proveniente da água interceptando as interfaces 2 e 1, respectivamente.

O erro médio de reconstrução dos pontos obtidos na variação da distância da interface para a câmera com o objeto à 500 mm é apresentado no gráfico presente na [Figura 29](#). No primeiro momento, quando o distanciamento da interface representa apenas 1% em relação ao do objeto, as técnicas propostas CSAT e CSAT-M obtiveram resultados (0,3093 mm e 0,5768 mm, respectivamente) muito mais próximos ao da técnica CM (0,0192 mm) do que a adaptação da literatura (CSMF, que apresentou erro de 2,9043 mm). De modo geral, as técnicas de adaptação CSAT e CSAT-M adquiriram erro abaixo de 1 mm até aproximadamente 25% de relação entre as distâncias câmera-interface e câmera-objeto. Após esses 25%, o erro assemelha-se ao encontrado pela CSMF, que não alcançou nenhum valor abaixo de 1,45 mm.

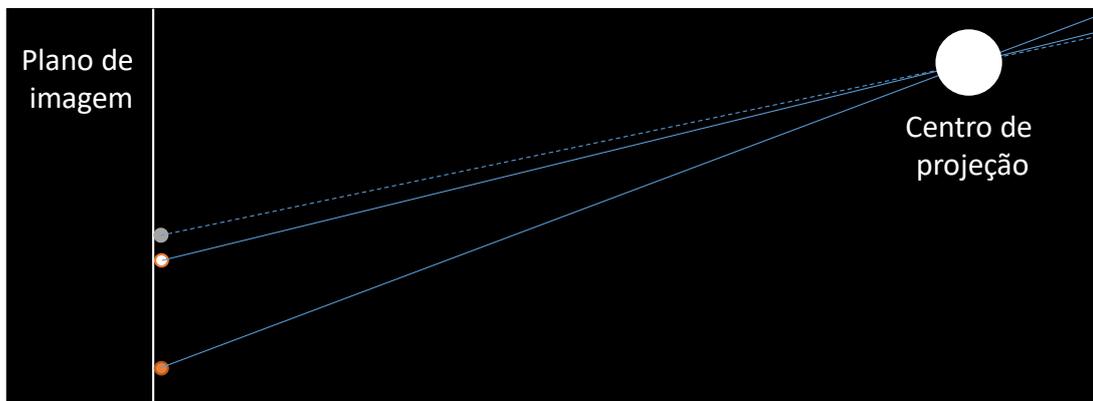
Um fato notável na [Figura 29](#) é que o método CM foi capaz de modelar satisfatoriamente o comportamento da câmera com distância da interface variável, alcançando erro de apenas 0,3181 mm com a interface à 140 mm da câmera. Esse erro assemelha-se ao menor erro obtido pela técnica CSAT, quando a interface está mais próxima da câmera. Porém, é importante ressaltar que nesses cenários sintéticos não existe a presença

Figura 28 – Esquema de exemplificação da inserção de distorção causada pela mudança de meio. A figura é dividida em duas partes: (a) apresenta uma visão geral de diferentes projeções para um ponto pertencente a um objeto submerso em água; (b) retrata apenas a parte interna da câmera, destacando a projeção dos pontos no plano de imagem em cinza, branco com borda laranja e o laranja que representam, respectivamente, os pontos resultantes da projeção simulando o objeto no ar (linha azul tracejada), objeto na água refratando na interface 2 e na interface 1.

(a) Visão geral



(b) Zoom na parte interna da câmera



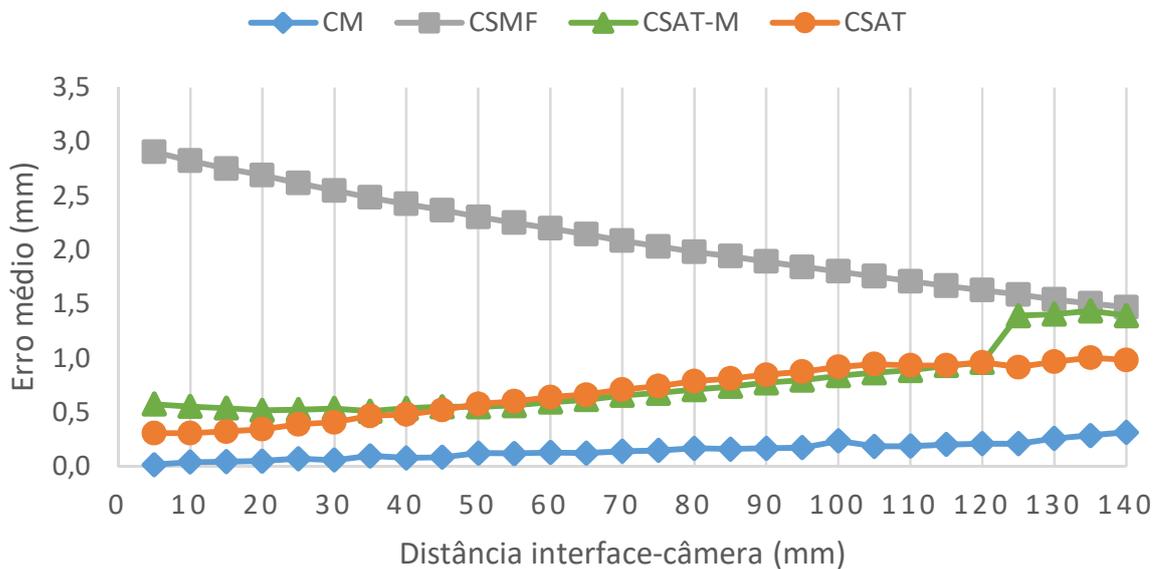
Fonte: Elaborada pelo próprio autor.

de nenhum tipo de ruído, fato muito diferente dos cenários reais, principalmente para ambientes subaquáticos onde a técnica de CM é aplicada. Todos os valores apresentados no gráfico da [Figura 29](#), assim como os dados de RMS, erro máximo e desvio padrão, estão evidenciados na [Tabela 1](#) presente no [Apêndice A](#).

5.2.1.2 Objeto de validação fixo à 1000 mm da câmera

Outra cena foi elaborada para continuar a avaliação acerca da influência que a variação presente na distância da interface para a câmera ocasiona no erro de reconstrução. A metodologia adotada é semelhante a descrita na [Subseção 5.2.1.1](#), e por isso, apenas as diferenças serão ressaltadas. A distância câmera-interface possui inicialmente 5 mm, sendo

Figura 29 – Gráfico de erro médio da reconstrução dos pontos do padrão obtido no Cenário 01 com objeto fixo sem 500 mm de distância da câmera.



Fonte: Elaborada pelo próprio autor.

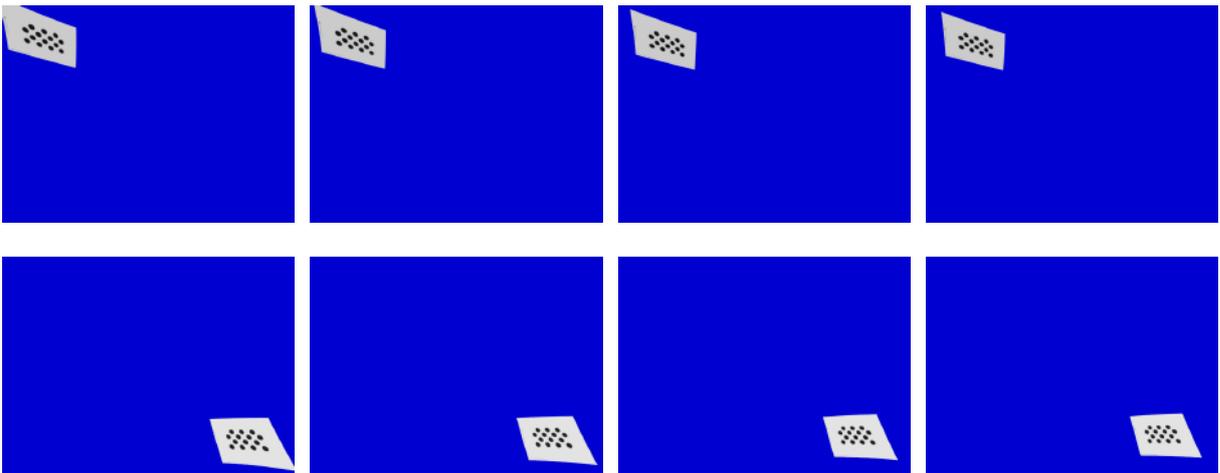
gradativamente espaçada até alcançar 125 mm, produzindo um conjunto de 50 imagens. A análise é realizada com o objeto de validação posicionado a 1000 mm da câmera, o que, em termos percentuais, significa uma variação de 0,5% a 12,5% na relação entre as distâncias câmera-interface e câmera-objeto. Assim como na análise anterior, os valores da distância entre a interface e o centro de câmera foram extrapolados, porém, essa distância do objeto para a câmera é frequentemente encontrada em aplicações de visão computacional. Um exemplo de aplicação onde o distanciamento da câmera para o objeto é elevado será exposta na [Subseção 5.3.2](#), onde há valores superiores a 15m.

Do mesmo modo que na [Subseção 5.2.1.1](#), a pose do objeto de validação consiste na mesma transformação sendo repetida a cada duas imagens. Os valores de rotação são os mesmos apresentados na [Figura 26b](#), com variação de 20° no eixo x e 10° nos eixos y e z. Como o objeto foi distanciado da câmera, de 500 mm para 1000 mm, os valores de translação no plano paralelo ao de projeção da câmera necessitaram ser reajustados para que o objeto continuasse no canto superior esquerdo e inferior direito da imagem, sendo modificado para 480 mm e -450 mm na direção do eixo y e em 350 mm e -350 mm no z.

Assim como demonstrado anteriormente, mesmo a lente da câmera não possuindo nenhum tipo de distorção, as imagens obtidas para a validação do cenário 01 com o objeto fixo a 1000 mm apresentam deformações causadas pela mudança de meio água-ar, como pode ser visualizado na [Figura 30](#). Essa deformação faz com que o objeto de validação encontre-se mais próximo à borda da imagem quando a distância interface-câmera é menor, corroborando com a justificativa apresentada de forma esquemática na

Figura 28. Em razão de o objeto de validação estar fixado a uma distância maior da câmera, todos os seus pontos característicos estão mais próximos da borda da imagem. Consequentemente, devido a maior distorção existente nas extremidades da imagem, os pontos característicos possuem mais perturbação no seu posicionamento.

Figura 30 – Exemplos das imagens do conjunto de validação do cenário 01 com objeto fixo a 1000 mm. As imagens selecionadas possuem, da esquerda para a direita, distância da interface para a câmera de 5, 50, 100 e 125 mm, sendo cada coluna referente a um conjunto usado para a reconstrução do objeto.



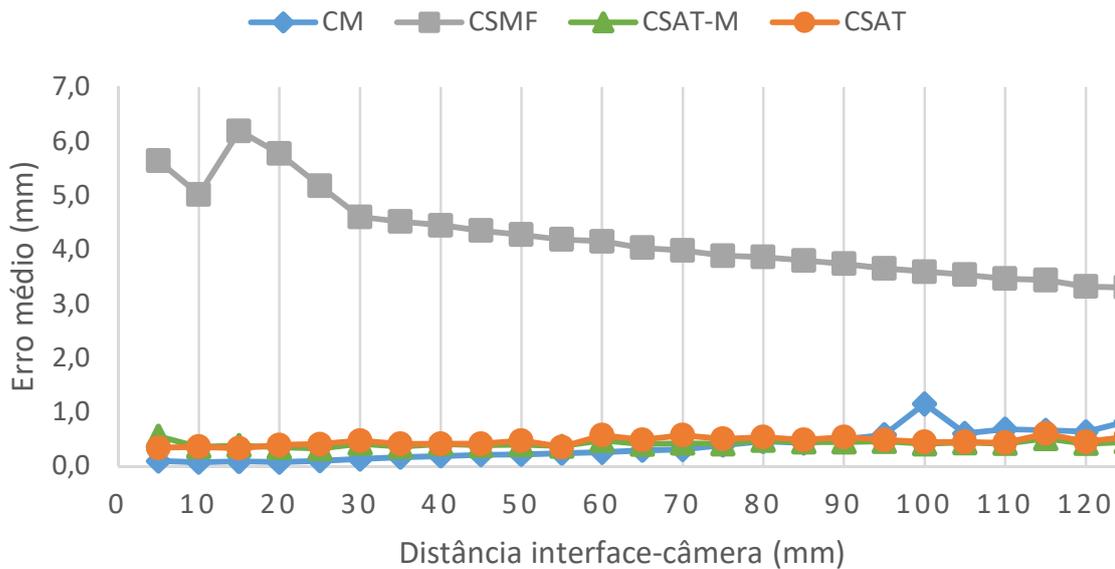
Fonte: Elaborada pelo próprio autor.

O resultado, graficamente ilustrado, dos erros médios de reconstrução dos pontos obtido nesse cenário é apresentado na **Figura 31**. Como a cada variação na distância são realizadas novas calibrações, os resultados obtidos possuem algumas variações (por exemplo, em 100 mm para técnica CM e em 5 até 15 mm para CSMF), possivelmente devido ao processo de otimização que não alcança o mínimo global da solução. Por não modelar a perturbação ocasionada pela refração existente na mudança de meio, a técnica CSMF apresentou erros de reconstrução mais elevados, uma vez que, o posicionamento dos pontos característicos do objeto localizam-se mais próximos das bordas da imagem, posição onde existe maior distorção pela refração. Em geral, a CSMF obteve erro médio superior a 3 mm, chegando a alcançar 6,1902 mm. Analisando o erro máximo dos pontos, a técnica variou de 5,0675 mm até 11,8108 mm e, apesar da discrepância em relação às demais técnicas nos erros de reconstrução, os valores de RMS alcançados são semelhantes.

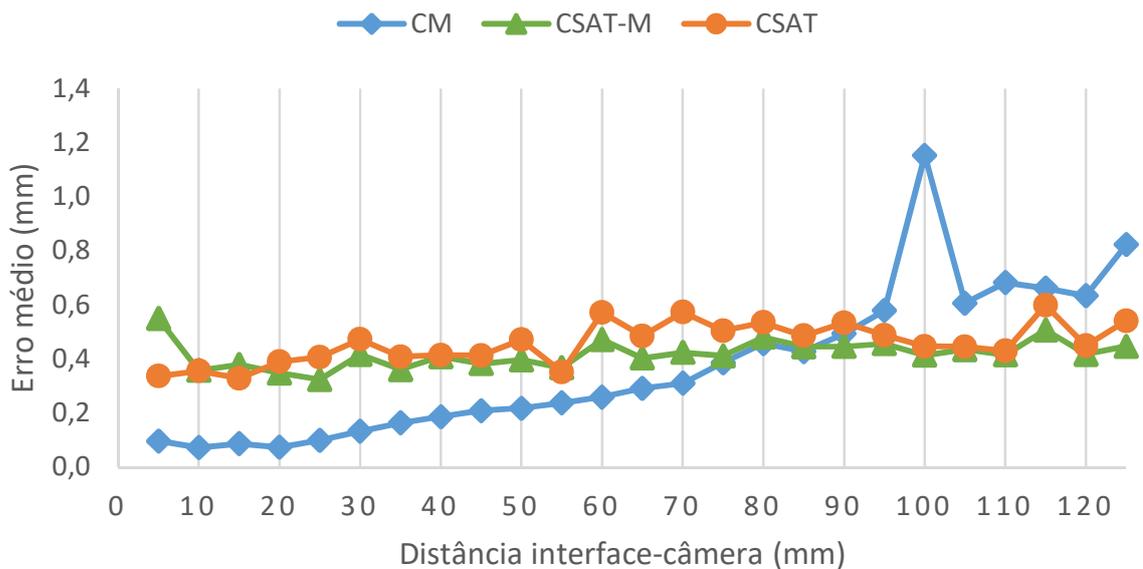
Para uma melhor análise das técnicas CM, CSAT e CSAT-M, a **Figura 31b** expõe o gráfico do erro médio apenas desses métodos, excluindo a CSMF. Um fato relevante de ser ressaltado nesse gráfico é que as técnicas CSAT e CSAT-M praticamente conservaram o erro obtido durante a variação da distância interface-câmera, variando entre os intervalos fechados $[0,32965; 0,59981]$ e $[0,3251; 0,54981]$, respectivamente. Em contrapartida, a CM apresentou uma degradação de seu resultado, variando entre o intervalo $[0,07295; 1,15241]$,

Figura 31 – Gráfico de erro médio da reconstrução dos pontos do padrão obtido no Cenário 01 com objeto fixo em 1000 mm de distância da câmera. O resultado obtido pelas técnicas é apresentado na totalidade (a) e excluindo a técnica CSMF (b), que, devido ao seu elevado erro, impede uma visualização mais detalhada da performance das demais técnicas.

(a) Resultado completo



(b) Resultado sem a técnica CSMF



Fonte: Elaborada pelo próprio autor.

ultrapassando o erro obtido pelas técnicas CSAT e CSAT-M quando a distância interface-câmera é de 95 mm, ou 9,5% de relação entre as distâncias. Essa degradação indica que a calibração realizada foi menos representativa à medida que a interface se distanciou da câmera. Todos os valores do gráfico presentes na [Figura 29](#), assim como os dados de RMS, erro máximo e desvio padrão, estão relatados na [Tabela 2](#) contido no [Apêndice A](#).

5.2.2 Cenário 02

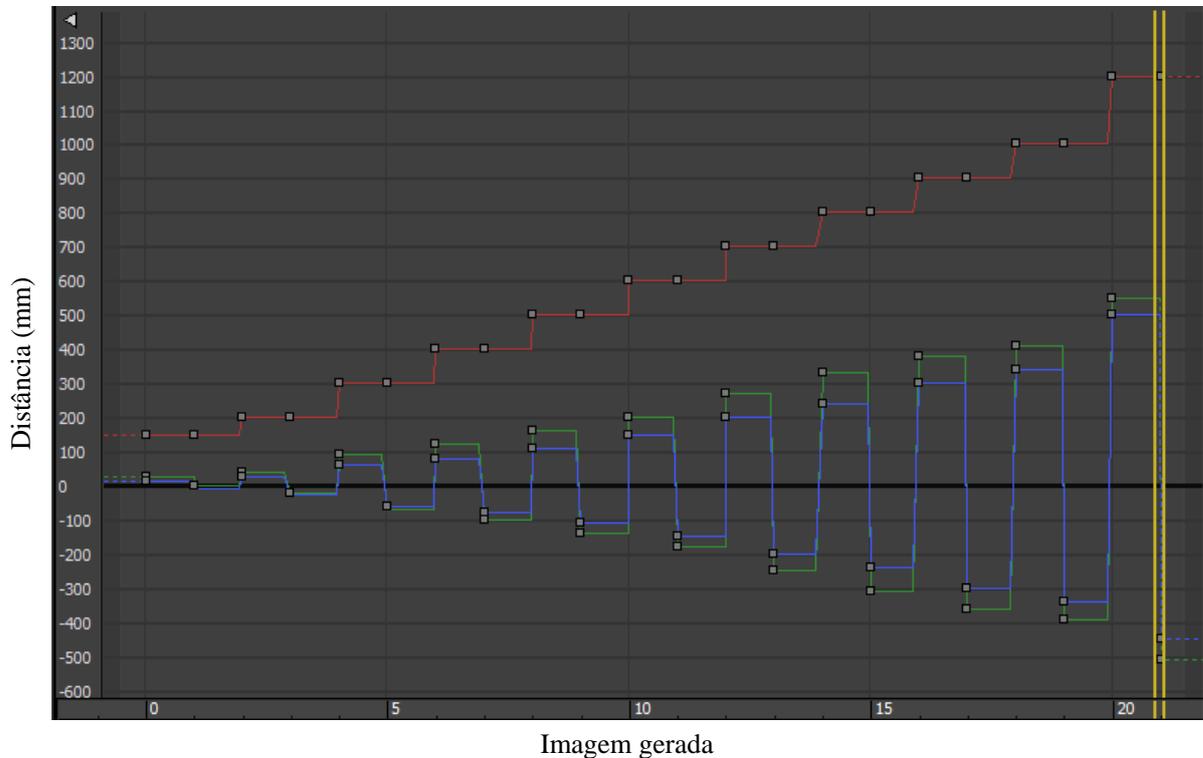
O segundo cenário foi elaborado para verificar a interferência ocasionada pela variação da distância entre o padrão e a câmera no erro de reconstrução do objeto de validação. Nesse cenário, a calibração molhada foi executada em uma única ocasião, em razão da distância câmera-interface estar fixada em cada caso de avaliação em 30 mm. Essa distância é coerente com as encontradas nas câmeras utilizadas em ambientes subaquáticas e será adotada para os demais experimentos sintéticos. Assim como mencionado no primeiro cenário, o aumento da distância do objeto diminuirá a relação entre as distâncias câmera-interface e câmera-objeto, aumentando a representatividade da aproximação realizada pela técnica de adaptação proposta. Nessas circunstâncias, o experimento foi desenvolvido através de imagens obtidas com o objeto de validação sendo deslocado para a borda da imagem a cada vez que era afastado e, em uma segunda abordagem, fixo no centro da imagem. Essas duas abordagens foram adotadas para verificar o desempenho das técnicas nas posições da imagem com mais e menos deformação, respectivamente, na borda e no centro da imagem.

5.2.2.1 Objeto de validação na borda da imagem

A cena produzida possui um conjunto de 22 imagens exibindo o objeto de validação sendo distanciado da câmera. Como apresentado na [Figura 32](#), o objeto, que inicialmente está a 150 mm, é afastado da câmera até a posição de 1200 mm. Essa distância inicial foi determinada por ser o menor posicionamento em que o objeto possuía todos os seus pontos característicos visíveis. Os deslocamentos no plano paralelo ao de projeção da câmera (eixos y e z) foram realizados de modo que o objeto estivesse posicionado no canto superior esquerdo e no canto inferior direito da imagem, e, por isso, a cada distanciamento esses deslocamentos foram aumentados. A rotação presente no objeto de validação é a mesma descrita no cenário 01 com o objeto sendo rotacionado nas duas direções em 20° no eixo x e 10° nos eixos y e z.

Um subconjunto de imagens utilizadas para a validação nesse cenário é apresentado na [Figura 33](#). A distorção presente nas imagens é exclusivamente obtida pela mudança do meio durante o caminho de luz do objeto para a câmera, habitada no interior da caixa de estanque. Um acontecimento importante a ser ressaltado nas imagens é que os pontos circulares, reconstruídos para a validação das técnicas, localizam-se mais próximos da extremidade à medida que o objeto é distanciado da câmera. Como mencionado anterior-

Figura 32 – Representação do 3DS Max sobre a translação do objeto de validação no cenário 02 nas 22 imagens produzidas. As cores vermelha, verde e azul representam os deslocamentos, respectivamente, nos eixos X, Y e Z.



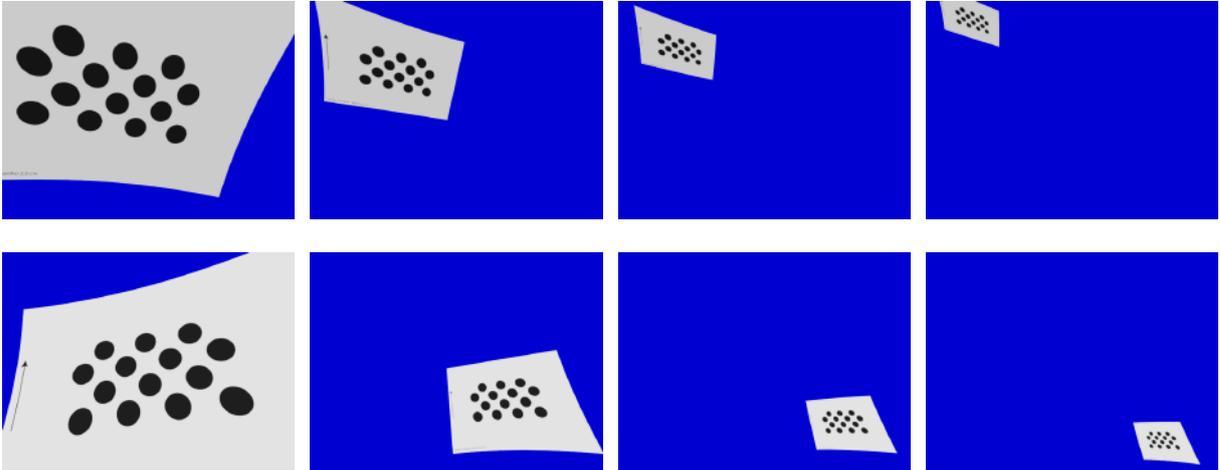
Fonte: Elaborada pelo próprio autor.

mente, as extremidades apresentam distorção mais acentuada, influenciando diretamente no aumento do erro em técnicas que não modelam essa perturbação corretamente.

Os erros médios dos pontos reconstruídos, através de duas imagens do objeto de validação com orientações e translação no plano paralelo ao de projeção diferentes, são apresentados no gráfico exposto na [Figura 34](#). Cada erro está associado a uma distância diferente do objeto para a câmera, com a interface fixada a 30 mm da câmera. No momento inicial, todas as técnicas obtiveram erros abaixo de 1,4 mm, porém dada a localização do objeto de validação, muito próximo à câmera, geraram mais pontos próximos ao centro da imagem do que na borda. Analisando o erro máximo obtido no primeiro momento, CSMF apresentou 3,18 mm de erro no ponto mais próximo à borda, enquanto as técnicas CSAT e CSAT-M possuíram erros de 1,39 mm e 1,65 mm, respectivamente, e CM com 0,86 mm de erro. Todos os valores de RMS, erro máximo e desvio padrão, assim como os de erro médio indicados no gráfico da [Figura 34](#), obtidos durante a reconstrução do objeto de validação estão relatados na [Tabela 3](#) exposta no [Apêndice A](#).

Observando o gráfico da [Figura 34](#), nota-se que a técnica de adaptação da literatura CSMF apresenta uma propensão ao aumento do erro à medida que os pontos característicos do objeto estão mais próximo à extremidade da imagem, o que ocorre à medida que o objeto está sendo distanciado da câmera. Os erros médios apresentados por CSMF estão

Figura 33 – Exemplos das imagens do conjunto de validação do cenário 02 com objeto deslocado para a borda. As imagens selecionadas possuem distância do padrão para a câmera de 150 mm, 400 mm, 800 mm e 1200 mm, sendo cada coluna referente a um conjunto usado para a reconstrução do objeto.



Fonte: Elaborada pelo próprio autor.

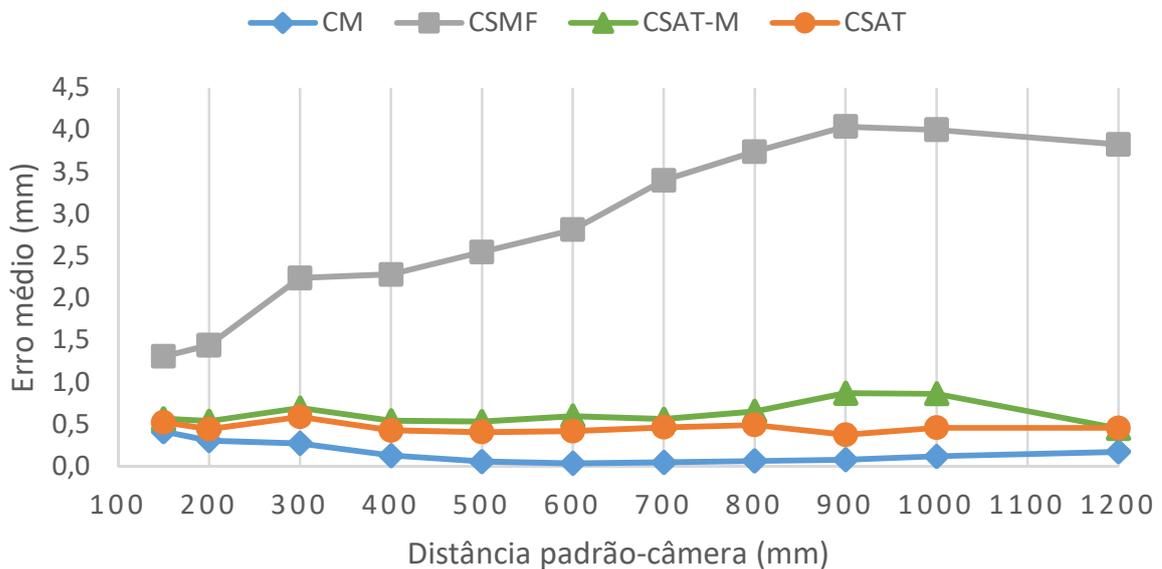
contidos no intervalo de $[1,3083; 4,0369]$, alcançando um erro máximo de até 6,09 mm. As técnicas CSAT e CSAT-M mantiveram valores no entorno de 0,5 mm de erro médio, variando no intervalo de $[0,3802; 0,5920]$ e $[0,4544; 0,8733]$, respectivamente. Analisando o erro máximo das reconstruções, as técnicas CSAT e CSAT-M atingiram valores máximos de 1,3897 e 1,6505 mm, respectivamente, justamente no primeiro momento, quando o padrão está mais próximo à câmera.

Investigando a calibração molhada (CM), dois fatos relevantes são destacados: o primeiro é que essa técnica apresentou os melhores resultados, principalmente pela total ausência de ruído no ambiente subaquático utilizado na calibração; o segundo é que os menores erros foram obtidos exatamente na profundidade em que mais o padrão de calibração foi amostrado para a câmera, entre 550 mm e 700 mm. Esse segundo fato indica que, para alcançar uma reconstrução mais acurada, é necessário calibrar a câmera amostrando o padrão de calibração principalmente próximo da área em que o objeto a ser reconstruído estará posicionado.

5.2.2.2 Objeto de validação no centro da imagem

Continuando a verificação da interferência ocasionada pela variação da distância entre o objeto de validação e a câmera, foi elaborada uma nova cena onde o objeto está posicionado sempre no centro da imagem. Essa cena, produzida com um conjunto de 20 imagens, tem como finalidade comparar as técnicas realizando a reconstrução do objeto localizado no espaço com menor perturbação da imagem, devido à distorção da lente ou mudança de meio. Assim como na [Subseção 5.2.2.1](#), o objeto de validação é posicionado inicialmente

Figura 34 – Gráfico de erro médio da reconstrução dos pontos do padrão obtido no Cenário 02 com objeto deslocado para a borda.



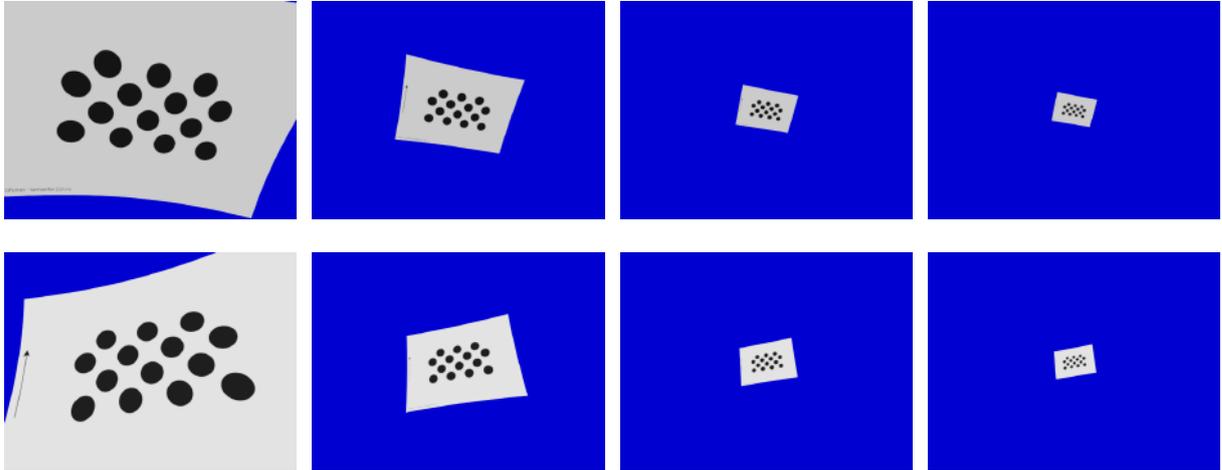
Fonte: Elaborada pelo próprio autor.

a 150 mm da câmera, sendo distanciada até 1000 mm. A rotação aplicada no objeto é a mesma adotada nos outros cenários de validação, com 20° no eixo x e 10° nos eixos y e z, sendo aplicada nas duas direções. Um subconjunto das imagens dessa cena é apresentada na [Figura 35](#), onde pode ser observado que quanto mais próximo o padrão está da câmera, mais espaçado pela imagem estão os seus pontos característicos.

O resultado do erro médio de reconstrução dos pontos obtidos a cada duas imagens do padrão de validação, centralizado na imagem com orientações diferentes, são apresentados graficamente na [Figura 36](#). Como os pontos característicos estão mais espaçados pela imagem quando o objeto está mais próximo às câmeras, a análise da interferência ocasionada pela variação da distância do objeto, localizado no centro da imagem, torna-se desviada. A partir da distância de 400 mm, como é apresentado na [Figura 35](#), os pontos característicos começam a ocupar a região central da imagem. Por esse motivo, a avaliação do resultado ressaltará principalmente os valores obtidos após o objeto de validação atingir a região central da imagem, considerando as amostras anteriores a 400 mm menos relevantes para essa análise.

Analisando o gráfico do erro médio de reconstrução do objeto de validação, observa-se um erro abaixo dos demonstrados nos cenários anteriores. Esse fato deve-se à menor perturbação causada nos pontos, localizados próximos ao centro, durante a formação da imagem pela mudança de meio. A técnica CSMF obteve erros pertencentes ao intervalo de $[0,0567; 0,6271]$ e quando examinado apenas a partir do objeto à 400 mm o seu erro médio máximo diminuiu para 0,1077 mm. Apesar da diferença encontrada no erro de reconstrução

Figura 35 – Exemplos das imagens do conjunto de validação do cenário 02 com objeto fixo no centro da imagem. As imagens selecionadas possuem distância do padrão para a câmera de 150 mm, 400 mm, 800 mm e 1000 mm, sendo cada coluna referente a um conjunto usado para a reconstrução do objeto.



Fonte: Elaborada pelo próprio autor.

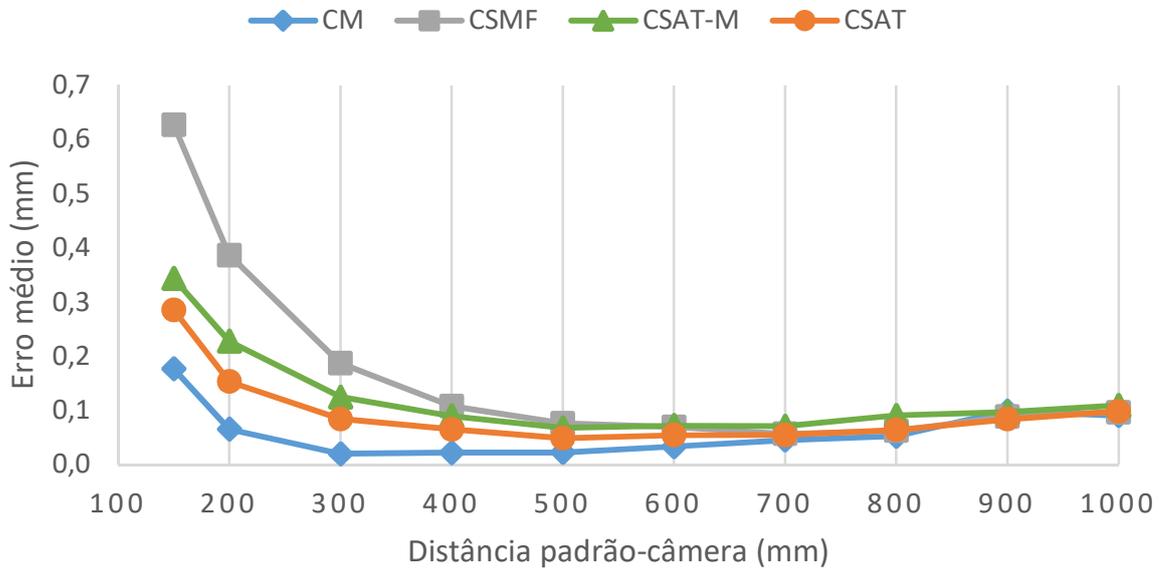
médio, os valores de RMS obtidos são todos muito próximos e a técnica CSMF possui o menor valor de RMS quando a distância é de 150 mm, fato que mostra-se o oposto no caso da reconstrução 3D. Todos os valores apresentados no gráfico da [Figura 36](#), assim como os dados de RMS, erro máximo e desvio padrão, estão evidenciados na [Tabela 4](#) presente no [Apêndice A](#).

A medida que os pontos ocuparam a região mais central da imagem, a amplitude dos erros obtidos (diferença entre o maior e menor erro) pelas técnicas diminuem significativamente. Essa amplitude é inicialmente de 0,4505 mm, decrescendo para 0,0854 mm quando o objeto está a 400 mm de distância para a câmera e para 0,0152 mm a 900 mm. Esta análise também pode ser aplicada verificando a amplitude dos erros máximos das reconstruções, que inicialmente é de 1,0156 mm, desce para 0,1949 mm quando o objeto está a 400 mm e para 0,0221 mm a 900mm. A tendência dos valores de amplitude, apresentada no gráfico de escala logarítmica na [Figura 37](#), indica que o uso das quatro técnicas de calibração na reconstrução nesse cenário podem ser equivalentes.

5.2.3 Cenário 03

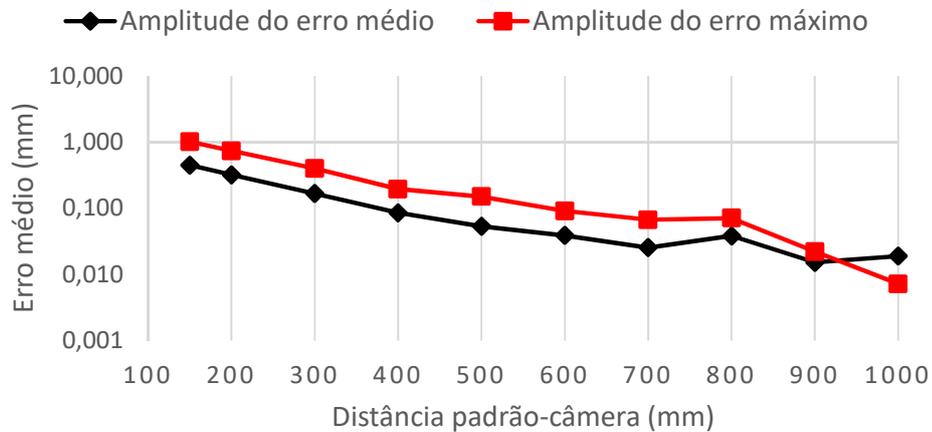
Como as câmeras utilizadas em cenários subaquáticos, normalmente, possuem lentes grande angulares com distorção, o terceiro cenário foi elaborado para verificar a interferência proporcionada pela variação no grau de distorção existente na lente dessa câmera. A fim de produzir o efeito de deformação na lente da câmera simulada pelo 3DS Max, foi utilizada a adição de distorção cúbica fornecida pela ferramenta. Para o funcionamento correto desse efeito foi necessário utilizar o renderizador Mental Ray, que simula física-

Figura 36 – Gráfico de erro médio da reconstrução dos pontos do padrão obtido no Cenário 02 com objeto fixo no centro da imagem.



Fonte: Elaborada pelo próprio autor.

Figura 37 – Gráfico das amplitudes dos erros médios e máximos obtidos pelas técnicas no Cenário 02 com objeto fixo no centro da imagem..



Fonte: Elaborada pelo próprio autor.

mente os efeitos causados na formação da imagem. Nesse cenário, a interface continuou sendo posicionada a frente da câmera com uma distância de 30 mm. A cada mudança no grau de distorção, foi necessário executar uma nova calibração seca e molhada, devido à mudança nos parâmetros físicos que modelam a câmera. Nessas conjunturas, o experimento foi desenvolvido através de imagens obtidas com o objeto de validação sendo deslocado para a borda da imagem, cuja movimentação depende do grau de distorção, e também fixo em uma posição onde fosse detectado por todas as diferentes câmeras.

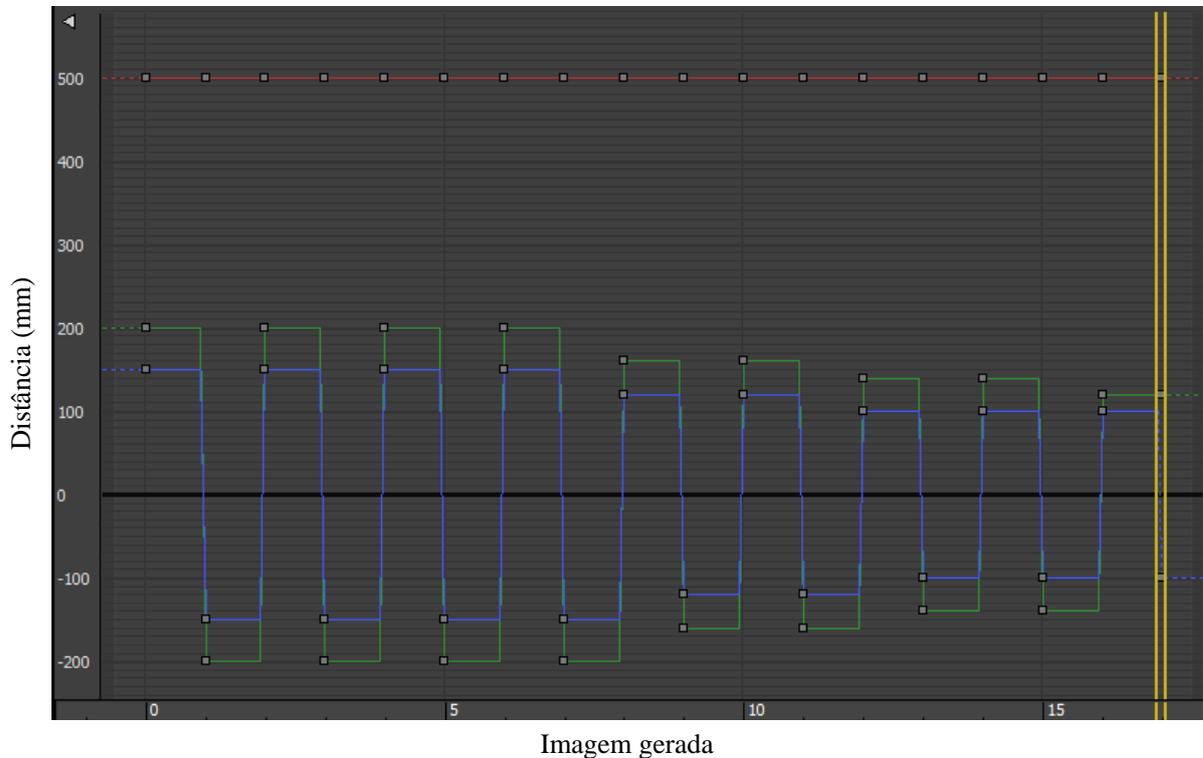
5.2.3.1 Objeto de validação na borda da imagem

A primeira cena, elaborada para verificar a interferência da variação do grau de distorção da lente no erro de reconstrução, possui um conjunto de 18 imagens geradas. A cada duas imagens, usadas para a reconstrução do objeto de validação, o grau de deformidade da lente é alterado, simulando assim uma nova câmera com outros parâmetros intrínsecos. O valor de distorção cúbica, iniciado em zero, foi gradativamente aumentado por um incremento de 0,04 até alcançar o valor de 0,32. O objeto de validação foi movimentado de maneira a localizar-se próximo aos cantos superior esquerdo e inferior direito da imagem, com o objeto a uma distância de 500 mm da câmera, assim como apresentado na [Figura 38](#). À medida que o grau de distorção aumenta, as translações diminuem para que os pontos característicos do padrão continuem visíveis na imagem. A rotação aplicada no objeto de validação seguem as mesmas definidas anteriormente, rotacionando-o nas duas direções em 20° no eixo x e 10° nos eixos y e z.

Um subconjunto das imagens utilizadas para a validação nesse cenário, formado apenas por aquelas onde o objeto está localizado no canto superior esquerdo, é apresentado na [Figura 39](#). Um fato importante que pode ser visualizado nessas imagens é a mudança do tipo de distorção, presente na formação da imagem, à medida que o grau de perturbação causado pela lente aumenta. Inicialmente, nesse cenário, a lente não introduz nenhuma perturbação na imagem e, portanto, a deformação identificada é proveniente unicamente da mudança de meio, que produz uma distorção do tipo almofada. Como a lente da câmera simulada representa uma grande angular, frequentemente utilizada em aplicações subaquáticas, uma deformação do tipo barril é inserida, magnificando mais o centro da imagem do que suas bordas.

O fato da lente possuir distorção do tipo oposto à introduzida pela mudança do meio faz com que a perturbação do tipo almofada seja gradativamente anulada, à medida que o grau de distorção da lente aumenta, até que a característica da distorção seja completamente modificada. O ponto de mudança do tipo de perturbação existente na imagem encontra-se, aproximadamente, entre os graus 0,08 e 0,16 da distorção cúbica utilizada pelo 3DS Max, como pode ser verificado nas imagens. Essa conclusão foi obtida através da análise visual das imagens produzidas nesse cenário e auxiliará na avaliação dos resultados obtidos.

Figura 38 – Representação do 3DS Max sobre a translação do objeto de validação no cenário 03 nas 18 imagens produzidas. As cores vermelha, verde e azul representam os deslocamentos, respectivamente, nos eixos X, Y e Z.



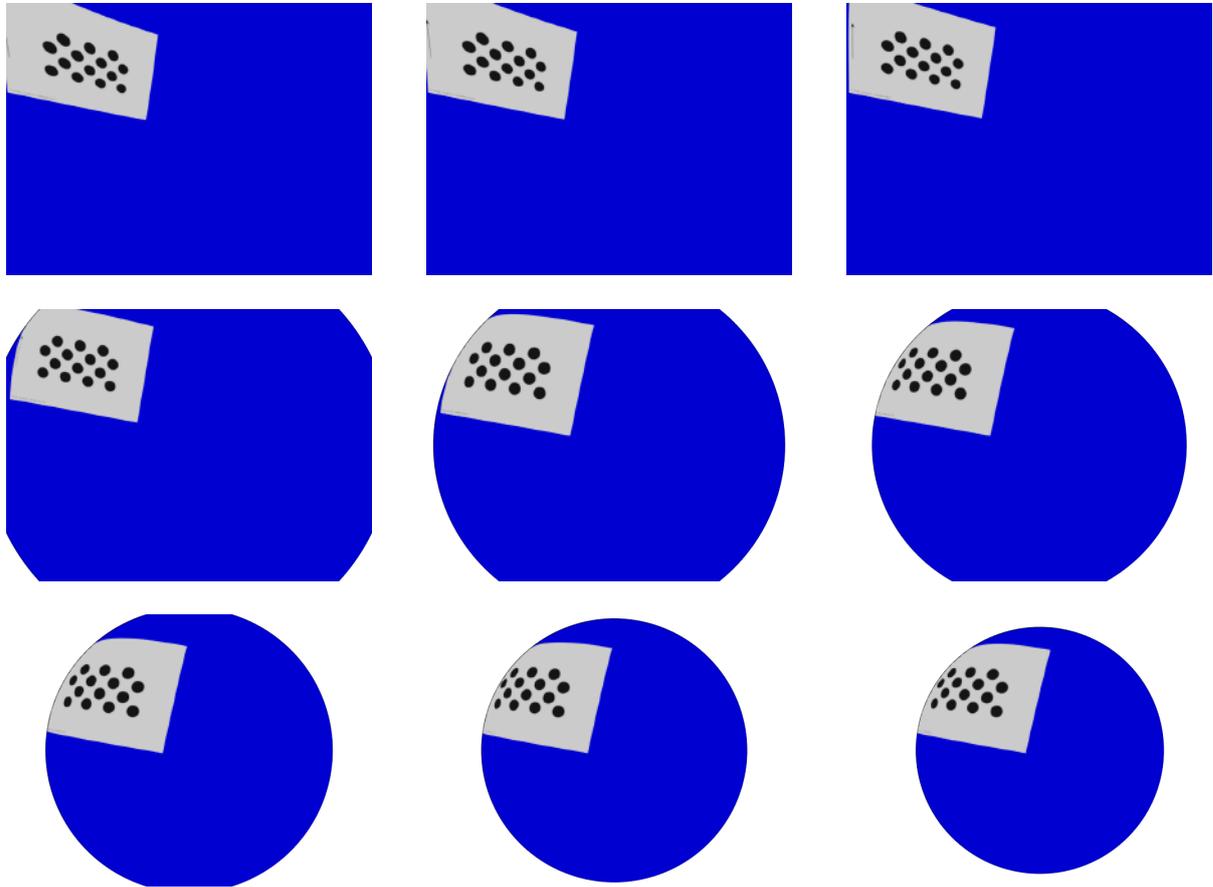
Fonte: Elaborada pelo próprio autor.

Os erros médios dos pontos reconstruídos, através de duas imagens do objeto de validação com orientações e translação diferentes, são indicados no gráfico apresentado na Figura 40. Cada erro está associado a uma câmera simulada com lente de grau de distorção diferente, com a interface fixada a 30 mm e objeto a 500 mm da câmera. No momento inicial, quando a distorção é representada pelo tipo almofada, as técnicas CM, CSAT e CSAT-M obtiveram os seus menores valores de erro médio, dispostos nos intervalos de $[0,0055; 0,0073]$, $[0,02218; 0,03022]$ e $[0,02249; 0,02989]$, respectivamente. Em contrapartida, o CSMF obteve os seus maiores erros, situados no intervalo de $[1,5092; 3,7304]$, fato que foi reiterado na análise do erro máximo dos pontos reconstruídos.

Observando os valores de RMS, a técnica CSMF obteve comportamento inverso ao apresentado no erro de reconstrução, com valores menores (cerca de 0,04) para distorção do tipo almofada e maiores (em torno de 0,14) para o tipo barril. A técnica da literatura obteve valor mínimo (1,5092 mm) no erro de reconstrução justamente próximo da anulação entre os efeitos de perturbação inseridos pela mudança do meio e da lente da câmera. Após o valor de 0,12 do grau de distorção, os valores obtidos do erro médio para a técnica CSMF passa a variar menos, dentro do intervalo $[1,87739; 2,83098]$ de amplitude 0,9536 mm.

Em uma análise mais geral, as técnicas propostas CSAT e CSAT-M obtiveram valores de erro médio da reconstrução do objeto muito próximos, variando entre $[0,2925; 2,4395]$

Figura 39 – Exemplos das imagens do conjunto de validação do cenário 03 com objeto deslocado para borda. As imagens selecionadas foram obtidas por uma câmera com distorção de lente cúbica, fornecida pelo 3DS Max, variando de grau 0 até 0,32 aumentado gradativamente com incrementos de 0,04.



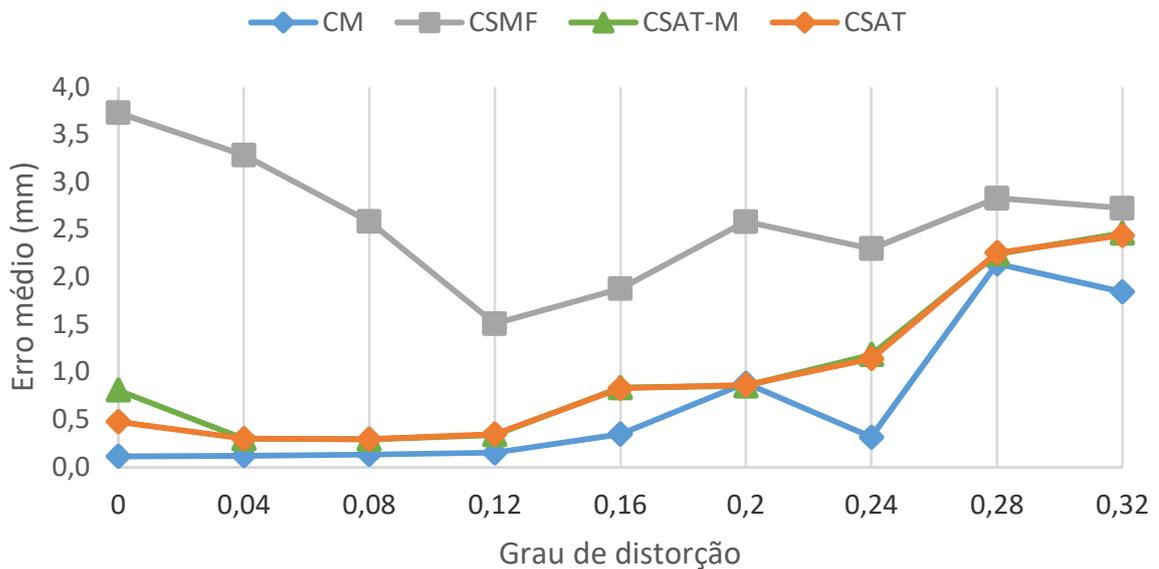
Fonte: Elaborada pelo próprio autor.

e $[0,29242; 2,46228]$, respectivamente. Todos os valores indicados no gráfico da [Figura 40](#), assim como os dados de RMS, erro máximo e desvio padrão, estão relatados na [Tabela 5](#) apresentada no [Apêndice A](#).

5.2.3.2 Objeto de validação com posição fixa

Com a finalidade de refinar a validação da interferência ocasionada pela variação do grau de distorção na lente da câmera, foi elaborada uma nova cena onde a posição 3D do objeto é a mesma para todos os níveis de distorção, ou seja, o objeto não foi deslocado para a borda da imagem a cada alteração realizada na lente. A posição do objeto de validação utilizada foi a mesma aplicada para a câmera com lente de grau 0,32 de distorção da cena anterior, posicionando-o a 500 mm de distância para a câmera e deslocando para o canto superior esquerdo e inferior direito com o deslocamento de ± 120 mm e ± 100 mm. Essa cena também possui as mesmas 18 imagens produzidas com o valor de distorção

Figura 40 – Gráfico de erro médio da reconstrução dos pontos do padrão obtido no Cenário 03 com objeto deslocado para borda.



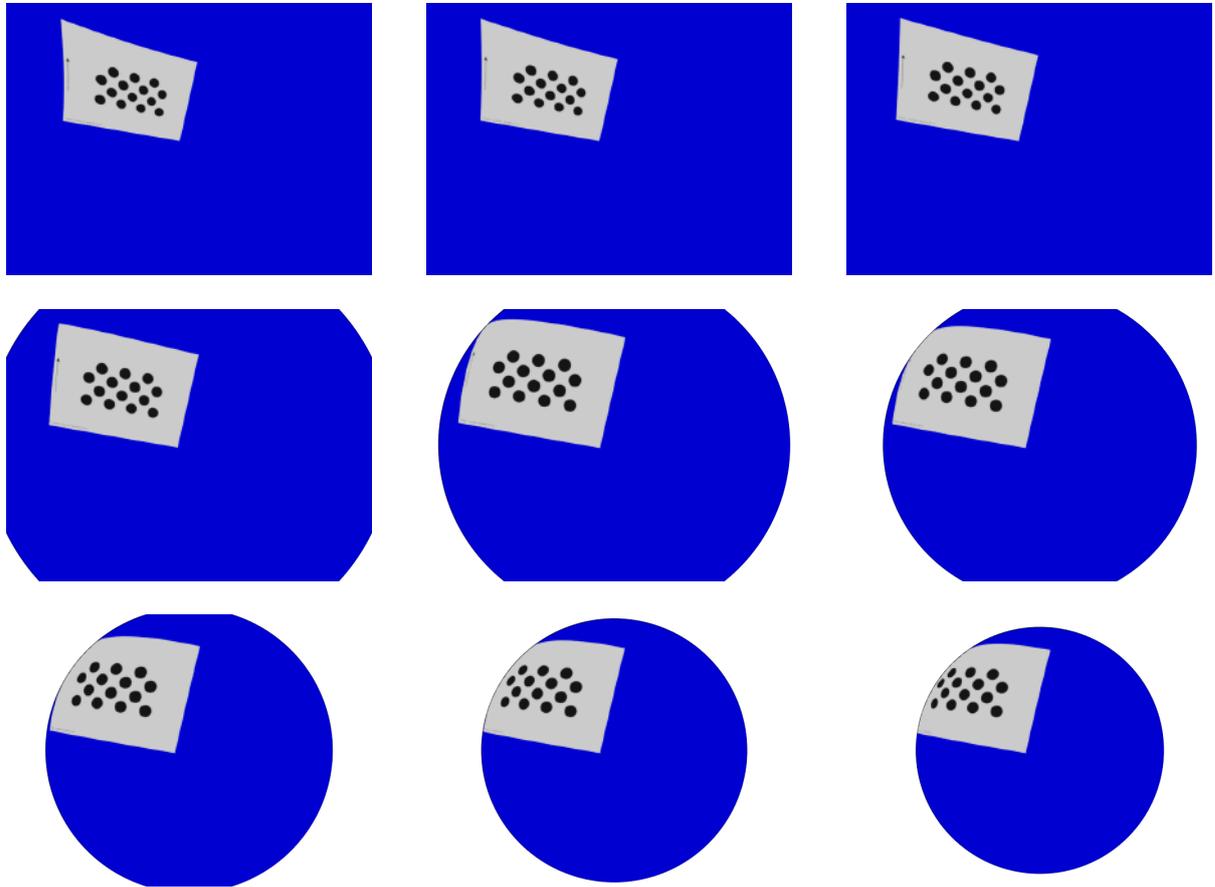
Fonte: Elaborada pelo próprio autor.

cúbica inserido pelo 3DS Max na lente da câmera, simulada variando de 0 até 0,32 com um incremento de 0,04 e a interface da caixa de estanque fixada a 30 mm da câmera.

Um subconjunto das imagens utilizadas para a validação nesse cenário é apresentado na [Figura 41](#), onde estão presentes apenas as imagens produzidas em que o objeto está localizado no canto superior esquerdo. Em razão do objeto de validação estar fixado para todos os valores de distorção, os pontos característicos do padrão situam-se mais próximos do centro quando a câmera possui menos distorção, pois a perturbação aplicada nos pontos é maior à medida que o coeficiente de distorção cúbica do 3DS aumenta. Em termos numéricos, a diferença entre a distância do centro da imagem, $[320, 240]^T$, para o primeiro ponto à esquerda do padrão, na parte inferior, é maior que 30 *pixels* quando comparado à câmera sem distorção com a de grau 0,32 presente em sua lente.

O gráfico que apresenta os erros médios de reconstrução dos pontos característicos do objeto de validação desse cenário é exposto na [Figura 42](#). Analisando o gráfico de forma geral, nota-se um comportamento muito semelhante ao apresentado na cena anterior, onde o objeto de validação foi deslocado para a borda. As principais diferenças é a suavização do resultado e a diminuição do erro obtido quando o grau de distorção é pequeno. O fato da suavização se explica devido à diminuição das variáveis analisadas, uma vez que nessa cena o objeto não é movimentado. O único fator que modifica a reconstrução do objeto de validação é o grau de distorção causado na imagem decorrente da lente da câmera. Enquanto o fato da diminuição do erro apresentado nos momentos iniciais deve-se à localização em que o objeto de validação está posicionado, que nessa cena é mais

Figura 41 – Exemplos das imagens do conjunto de validação do cenário 03 com objeto fixo. As imagens selecionadas foram obtidas por uma câmera com distorção de lente cúbica, fornecida pelo 3DS Max, variando de grau 0 até 0,32 aumentado gradativamente com incrementos de 0,04.



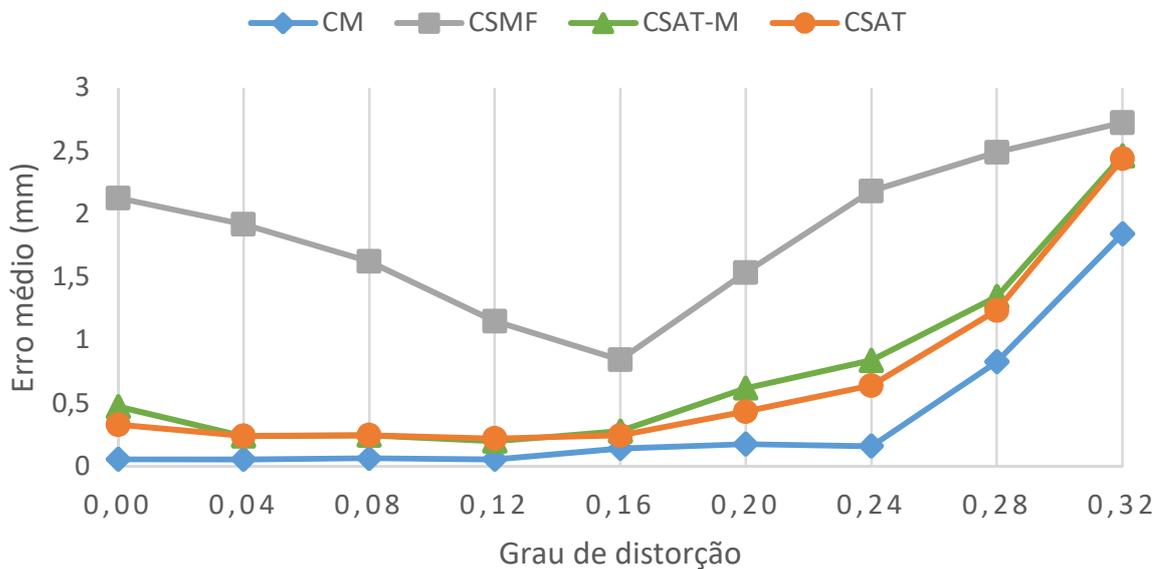
Fonte: Elaborada pelo próprio autor.

distante da borda do que na anterior.

Avaliando o gráfico, é possível visualizar que a técnica CM possui um leve aumento do erro quando o grau de distorção alcança o valor de 0,16 e que seu erro degrada após o grau de 0,24. No primeiro momento, até o grau de 0,16 de distorção, a técnica CM obteve valores no intervalo de [0,0581; 0,0667], no segundo momento seus resultados passam a integrar o espaço de [0,1441; 0,1795], no fim seus valores são de 0,8309 mm e 1,8442 mm para os graus de 0,28 e 0,32, respectivamente. A técnica proposta, CSAT, apresentou tendência semelhante ao CM, porém com apenas dois momentos: no primeiro o erro variou no intervalo [0,3329; 0,2229], onde o maior erro obtido foi justamente quando a única distorção presente era causada pela mudança de meio; o segundo momento foi o de degradação, a partir do grau de 0,20 de distorção, onde o erro passa de 0,4367 até 2,439mm. CSAT-M obteve tendência muito similar à apresentada por CSAT, porém com erros pouco mais acentuados.

As conclusões obtidas para a técnica CSMF na cena anterior foram reiteradas nessa

Figura 42 – Gráfico de erro médio da reconstrução dos pontos do padrão obtido no Cenário 03 com objeto de fixo.



Fonte: Elaborada pelo próprio autor.

cena, porém o valor mínimo de erro obtido nessa cena foi com grau de distorção de 0,16, enquanto na anterior foi no grau de 0,12. Essa mudança no valor mínimo também refletiu nos valores de RMS, que mantiveram a mesma tendência, modificando apenas o ponto em que o RMS aumenta drasticamente. Todas as demais técnicas obtiveram valores de RMS similares, com exceção de quando as lentes possuíam grau de distorção entre 0,20 e 0,28, onde ocorreu uma maior variação. Todos os valores de RMS e erro médio indicados nessa cena, assim como os dados de desvio padrão e erro máximo, estão relatados na [Tabela 6](#) contida no [Apêndice A](#).

5.3 EXPERIMENTOS EM AMBIENTES REAIS

Com o propósito de ratificar as conclusões obtidas nos cenários sintéticos, foram desenvolvidos experimentos em ambientes subaquáticos, onde diversas aplicações de visão computacional são executadas atualmente. Os experimentos foram conduzidos em um tanque de água e no ambiente submarino, sendo suas particularidades descritas nas próximas subseções. Em ambos os casos foi utilizada a câmera da marca Kongsberg Maritime de modelo OE15-101c cuja interface é plana, ilustrada na [Figura 43](#). Essa câmera, que possui uma caixa de estanque de material em liga leve de titânio, possui lente grande angular, com campo de visão de 84° na diagonal quando submersa, e sensibilidade a iluminação de $1,3 \times 10^{-3}$ LUX. Por essas suas características, é comumente utilizada em operações submarinas de alta profundidade, uma vez que suporta pressões elevadas equivalente a (profundidades de até 3.000 metros) e adquire imagens mesmo em cenas com

baixa iluminação.

Figura 43 – Câmera subaquática, com caixa de estanque de interface planar, utilizada nos experimentos reais desta dissertação.



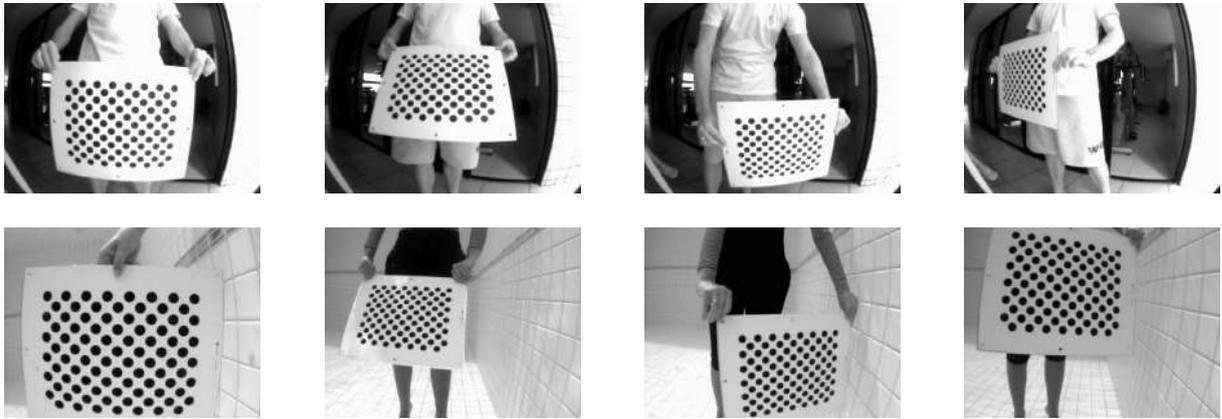
Fonte: Kongsberg Maritime.

5.3.1 Tanque de água

O primeiro experimento real foi realizado em um tanque de água em uma profundidade inferior a 1 m. Os problemas desse cenário em relação ao ruído e a variação de iluminação são mínimos, uma vez que a água não possui resíduos, não está turva e, estando o objeto muito próximo à superfície, a iluminação proveniente do sol apresenta-se uniformemente ao longo da imagem gerada. Para realizar a validação nesse cenário foram executadas duas vezes cada processo de calibração seca e molhada, capturando aproximadamente 40 imagens seguindo as especificações do cenário sintético de movimentação e orientação do padrão de calibração. A [Figura 44](#) apresenta um subconjunto das imagens utilizadas nas calibrações seca e molhada para esse cenário.

As imagens usadas para validação das técnicas foram produzidas com o objeto a aproximadamente 95 cm e 130 cm de distância da câmera, sendo separadas em duas categorias: objeto no centro e na borda. Essa divisão de categorias baseou-se no cenário sintético 02, descrito na [Subseção 5.2.2](#), buscando corroborar com as conclusões obtidas anteriormente. Os quatro conjuntos de validações são apresentados na [Figura 45](#), onde: o primeiro e segundo conjunto representam o objeto distanciado da câmera em aproximadamente 95 cm, posicionados, respectivamente, no centro ([Figura 45a](#)) e nas bordas superior esquerda e inferior direita da imagem ([Figura 45b](#) e [Figura 45c](#)); enquanto o terceiro e quarto conjunto consistem em uma separação de aproximadamente 130 cm com o objeto localizado no centro ([Figura 45d](#)) e nas bordas ([Figura 45e](#) e [Figura 45f](#)), respectivamente. A determinação da maior distância de validação com 130 cm ocorreu devido ao fato de, após esse distanciamento, os pontos não serem mais detectados.

Figura 44 – Subconjunto das imagens utilizadas nas calibrações seca e molhada para a validação no tanque de água. A primeira linha de imagens correspondem ao vídeo de calibração seca, enquanto a segunda ao de calibração molhada.



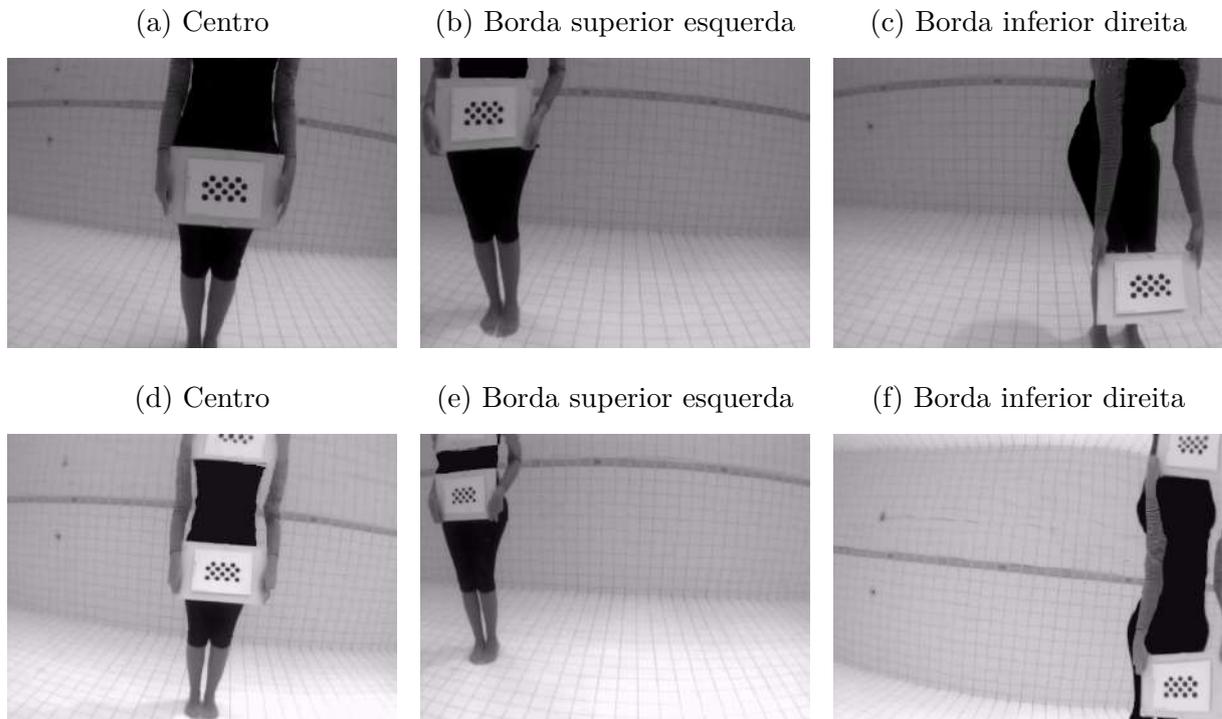
Fonte: Elaborada pelo próprio autor.

Para reconstrução do objeto, mesmo necessitando de apenas dois pontos de vista, foram selecionadas automaticamente quatro imagens, igualmente distribuídas ao longo do vídeo, para formar um conjunto de validação. Essa prática foi utilizada para diminuir o erro de reconstrução obtido, uma vez que, como não foi desenvolvido nenhum critério de seleção para utilizar pontos de vistas com orientações diferentes, conjuntos com apenas duas imagens (como utilizado no caso sintético) não forneceram informações necessárias para o algoritmo de reconstrução, retornando assim erros discrepantes que poderiam prejudicar a análise dos resultados. Como pode ser visualizado nas figuras 45d e 45f, o reflexo do objeto de validação apareceu em alguns trechos dos vídeos, reproduzindo-o na parte superior da imagem com dimensões modificadas. Para que esse fato não influenciasse o resultado obtido, uma vez que esses pontos também foram detectados, foi necessário um pré-processamento no vídeo antes da detecção para remover essa anomalia.

Os vídeos de validações foram gravados com cerca de 90 segundos, obtendo em torno de 2700 imagens. Se o objeto fosse detectado em toda a sequência de quadros, seria alcançado um número superior a 650 conjuntos de validações, porém devido às falhas na detecção, os vídeos “Centro 95 cm”, “Centro 130 cm”, “Borda 95 cm” e “Borda 130 cm” atingiram 544, 178, 421 e 137 conjuntos, respectivamente. O fato dos vídeos com o objeto a 130 cm possuírem um conjunto menor deve-se a maior quantidade de falhas na detecção do padrão, uma vez que os pontos característicos do objeto na imagem são representados por uma pequena quantidade de *pixels*, prejudicando assim a performance do algoritmo de detecção. Os resultados, com os dois conjuntos de calibrações obtidos, estão apresentados nos gráficos das figuras 46 e 47.

Assim como foi apresentado no cenário 02 do experimento sintético, na [Subseção 5.2.2](#), todas as técnicas obtiveram resultados equivalentes quando o objeto reconstruído localizou-

Figura 45 – Imagens dos vídeos utilizados para a validação no tanque de água. A primeira fileira de imagens corresponde às obtidas com o objeto de validação a aproximadamente 95 cm de distância, enquanto a segunda está com 130 cm.



Fonte: Elaborada pelo próprio autor.

se no centro da imagem. Nos vídeos “Centro 95 cm” e “Centro 130 cm”, as reconstruções obtidas utilizando as diferentes calibrações apresentaram erro médio que variaram de 0,12 a 0,13 mm e de 0,18 a 0,19 mm, respectivamente, considerando os dois conjuntos de calibrações. Essas pequenas diferenças entre as técnicas são menores do que os desvios padrões obtidos ao longo do conjunto de validação, com valores superiores a 0,035 mm.

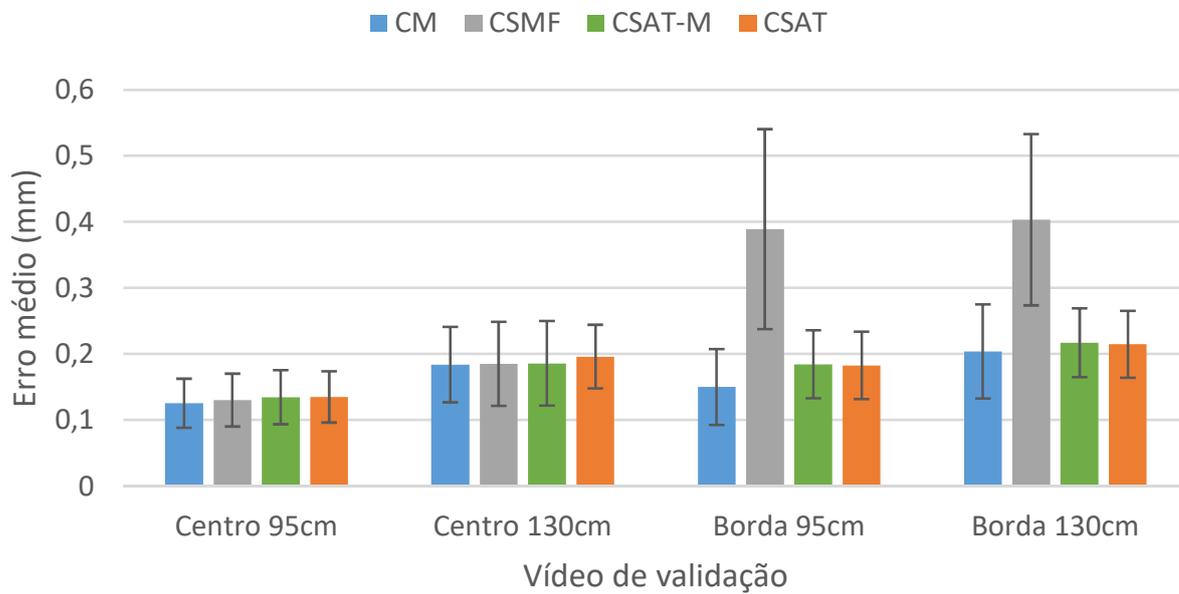
Quando o objeto reconstruído localiza-se na borda da imagem, é verificado uma discrepância no erro alcançado pela técnica da literatura CSMF, onde o resultado do erro médio alcançado encontra-se próximo ao dobro dos resultados obtidos pelas demais técnicas. Mesmo considerando-se os desvios padrões, que no caso da técnica CSMF é alto, o erro obtido pelo CSMF não se equivale aos demais resultados. Por outro lado, considerando os desvios padrões, as técnicas CM, CSAT e CSAT-M possuem resultados equivalentes em todos os cenários de validação.

Todos os dados de erro médio, desvio padrão e erro máximo apresentados como resultados nessa subseção estão relatados nas tabelas 7 e 8, presentes no [Apêndice B](#).

5.3.2 Ambiente submarino

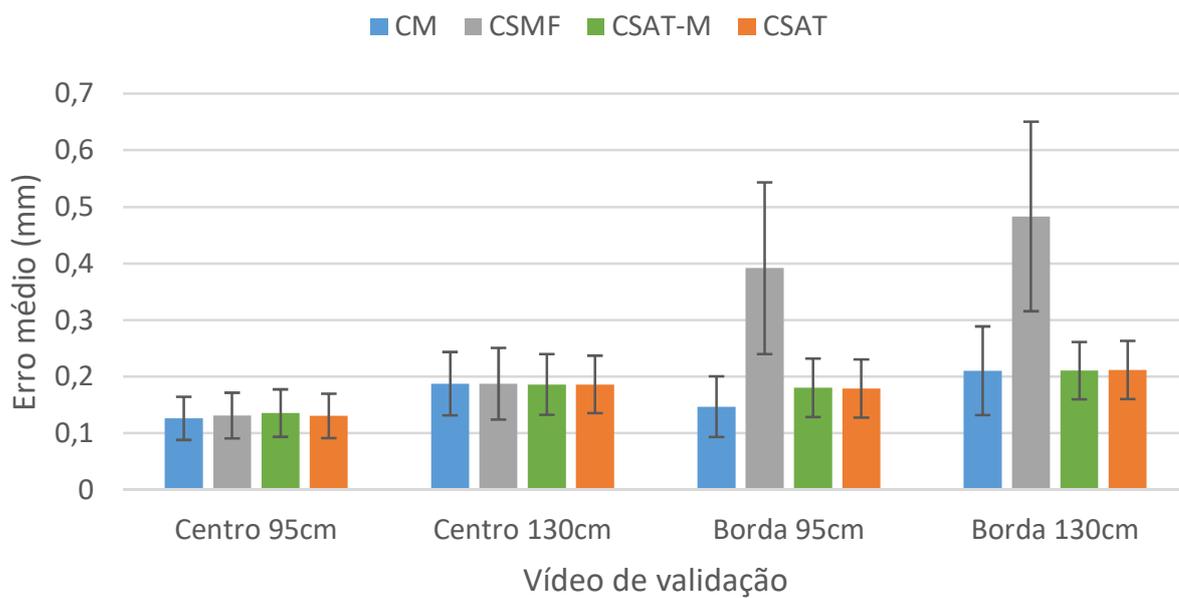
Com o intuito de validar a técnica de adaptação proposta nesse trabalho em cenários mais complexos, foram desenvolvidos experimentos em operações reais de instalação de dutos

Figura 46 – Gráfico de erro médio da reconstrução no tanque de água com o primeiro conjunto de calibrações. Considerando desvio padrão obtido em cada um dos vídeos de validação.



Fonte: Elaborada pelo próprio autor.

Figura 47 – Gráfico de erro médio da reconstrução no tanque de água com o segundo conjunto de calibrações. Considerando desvio padrão obtido em cada um dos vídeos de validação.



Fonte: Elaborada pelo próprio autor.

submarinos para extração de petróleo e gás. Para tal, utilizou-se o *software* de suporte a operações de instalações submarinas (SOIS) proposto em (SANTOS et al., 2015), que monitora o raio de curvatura do duto através da reconstrução 3D de pontos detectados na imagem. Esses pontos são localizados a partir da identificação de um padrão de pintura preestabelecida na região de interesse do duto, intercalando regiões de colorações brancas e pretas. As regiões brancas, que possuem aproximadamente o dobro do comprimento das pretas, são pintadas na extensão similar ao diâmetro do duto.

No navio de operações, que realiza as instalações submarinas e controla o ROV, são recebidas e processadas as imagens através da ferramenta do SOIS. Essa ferramenta utiliza duas câmeras, fixadas na estrutura do ROV, para realizar a reconstrução 3D dos pontos do duto detectados nas imagens. Durante a operação, o ROV posiciona-se a uma distância média de 5 a 15 metros do duto, observando-o em diferentes pontos de vista. Devido à resistência mecânica da água, os movimentos realizados são lentos e suaves, e em razão disso, as imagens produzidas em momentos subsequentes apresentam pouco deslocamento do duto. O sistema de câmeras utilizado nessas operações é composto por duas câmeras submarinas de baixa iluminação, iguais às descritas anteriormente. As câmeras foram posicionadas com eixos ópticos paralelos, distanciadas horizontalmente em 1,06 m e fixadas na estrutura do ROV de forma a evitar o movimento relativo entre elas.

A luz natural dificilmente alcança profundidades superiores a 30 m em águas costeiras e, por isso, nesse cenário é necessário a utilização de fontes artificiais de iluminação. Como os experimentos foram realizados em águas profundas do oceano, a iluminação presente no ambiente foi gerada exclusivamente por iluminadores equipados no ROV próximos às câmeras, facilitando a visibilidade, mas causando iluminação não uniforme no cenário. Devido a essas condições, as imagens produzidas pelas câmeras apresentam o objeto de interesse mais claro do que o fundo da cena, fazendo com que os *pixels* pertencentes ao duto possuam maior intensidade do que a sua vizinhança, facilitando assim sua detecção.

Dado que o duto é uma região cilíndrica flexível, não foi possível obter um gabarito para os pontos 3D reconstruídos, assim como, também não foi factível atingir uma calibração para a câmera que pudesse ser considerada como verdade absoluta em uma eventual comparação entre as técnicas. Por esse fato, a métrica utilizada nesse experimento é diferente das anteriores, empregando a geometria da visão estéreo, geometria epipolar, para realizar as comparações.

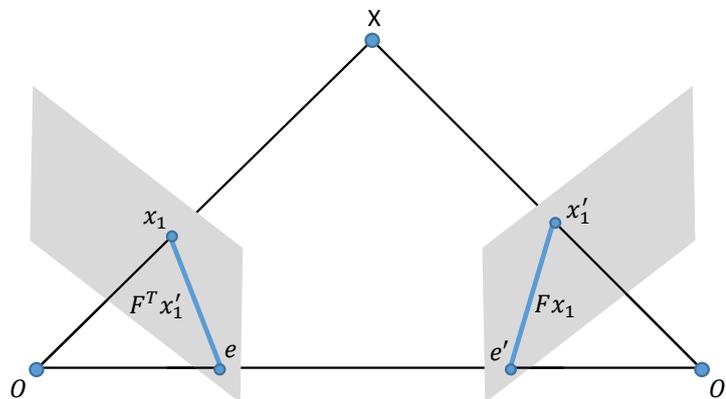
A geometria epipolar descreve as propriedades geométricas existentes entre duas imagens da mesma cena, a partir dos parâmetros intrínsecos (matrizes de projeção K e K') e extrínsecos (através de uma matriz de rotação R e um vetor de translação \mathbf{t}) das câmeras. Essas propriedades são encapsuladas em uma matriz F , de dimensão 3×3 , denominada de matriz fundamental, correlacionando os pontos x_1 e x'_1 (projeções do ponto 3D X pelas

câmeras de centro O e O' , respectivamente) através da relação

$$x'_1 F x_1 = 0. \quad (5.3)$$

Uma importante propriedade da matriz fundamental é que a reta obtida por $l' = F x_1$ (reta epipolar) contém o ponto X projetado em x'_1 na outra imagem, assim como a reta $l = F^\top x'_1$ contém o ponto x_1 . Portanto, se x'_1 corresponde corretamente a x_1 , a distância de x'_1 para a reta epipolar l' deve ser igual a zero. Este cenário é apresentada na [Figura 48](#).

Figura 48 – Geometria epipolar.



Fonte: Elaborada pelo próprio autor.

O erro epipolar é definido por:

$$\epsilon = \sum (d(x_1, F^\top x'_1)^2 + d(x'_1, F x_1)^2), \quad (5.4)$$

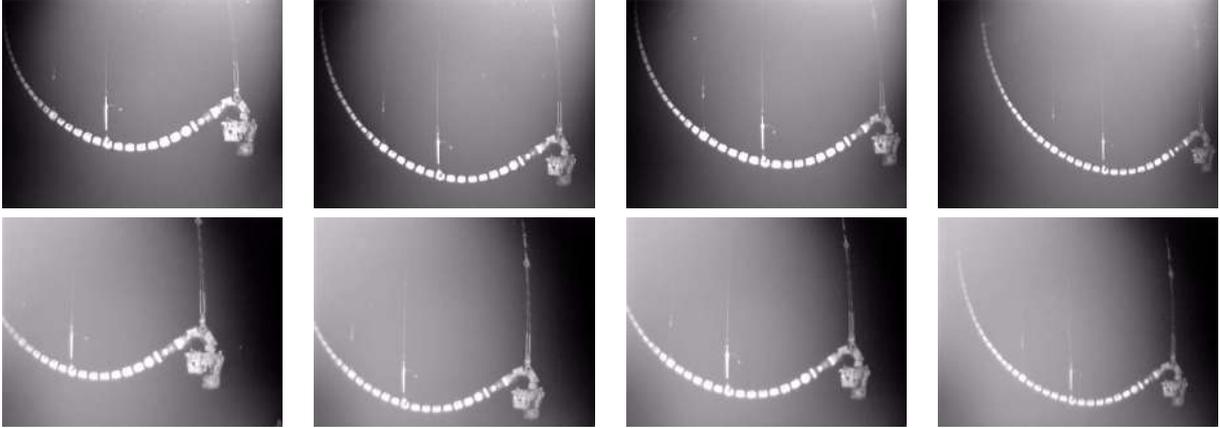
onde d é a distância euclidiana entre os pontos x_1 e x'_1 com suas respectivas retas epipolares $F^\top x'_1$ e $F x_1$. Essa métrica de erro é definida como as médias das distâncias euclidianas, sendo expressa em *pixels*. Para maiores informações sobre geometria epipolar e matriz fundamental consultar [Hartley e Zisserman \(2004\)](#).

As operações em estudo realizaram a conexão de um duto flexível no equipamento submarino a uma profundidade superior a 2000 metros, sendo acompanhadas nos dias 19 (Vídeo 1 e 2) e 21 (Vídeo 3 e 4) de junho de 2017. Nas próximas subseções serão detalhados os resultados obtidos nessas operações para as técnicas CSMF, CSAT e CSAT-M, uma vez que não foi possível realizar a calibração molhada no local de operação.

5.3.2.1 Vídeo 01

O primeiro vídeo investigado possui um total de 4755 quadros do duto em posição horizontal, com um subconjunto das imagens obtidas pelas câmeras ilustradas na [Figura 49](#). É possível visualizar nessas imagens a iluminação não uniforme causada pela luz gerada unicamente pelo iluminadores acoplados no ROV. Esse é um dos fatores que tornam esse

Figura 49 – Subconjunto das imagens do Vídeo 01 analisado durante o experimento no ambiente submarino. A primeira fileira de imagens obtidas corresponde à câmera posicionada na esquerda do ROV e a segunda às adquiridas pela câmera à direita, referindo-se aos quadros 1, 1500, 3000 e 4500 do vídeo.



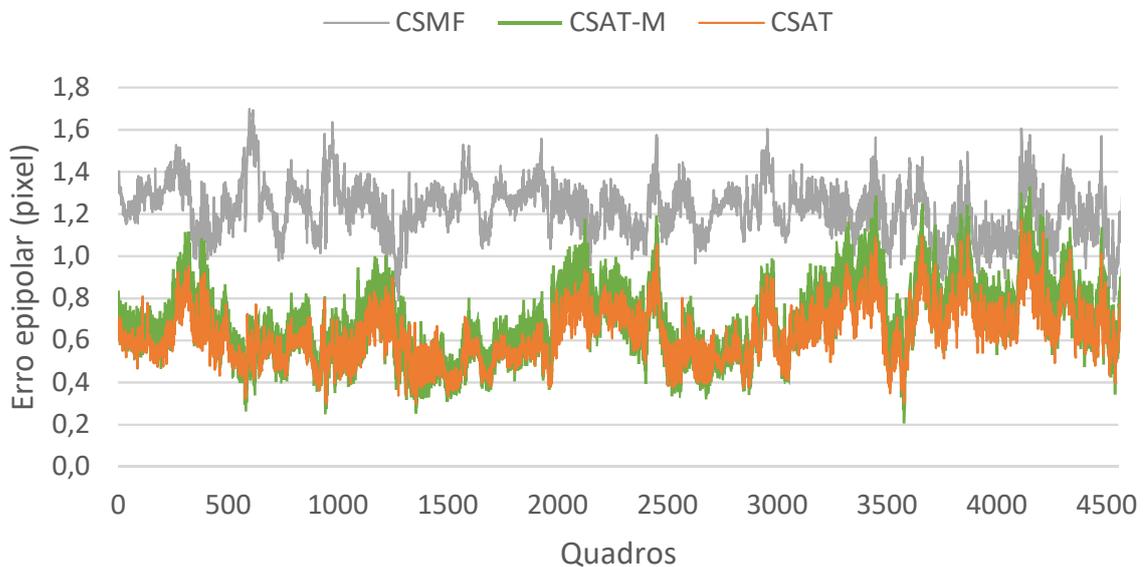
Fonte: Petrobras.

ambiente mais desafiador para aplicações de visão computacional. Os pontos utilizados na reconstrução, com a finalidade de comparar as técnicas, são os centroides das marcações brancas existentes ao longo do duto. Esses pontos são reconstruídos após uma etapa de correspondências entre as detecções obtidas nas imagens geradas pelas duas câmeras utilizadas. Portanto, é importante que o duto esteja visível simultaneamente nas duas visões, sendo análogo à reconstrução realizada nos experimentos anteriores, onde apenas uma câmera foi utilizada com pelo menos duas imagens obtidas de posição diferente.

Nesse vídeo de validação, o objeto de interesse apresentou-se, predominantemente, na região inferior da imagem, cobrindo praticamente toda sua extensão horizontal. Devido ao fato do duto ainda não estar acoplado no poço de extração, seu movimento tornou-se mais livre, possuindo uma maior variação na disposição dos pontos detectados. Ao longo de toda a sequência de quadros, o duto foi apropriadamente localizado nas imagens geradas por ambas as câmeras, variando apenas a quantidade de pontos detectados em cada momento. Utilizando uma média das reconstruções dos pontos, verificou-se que a distância do duto para a câmera variou entre 10,5 e 17,5 metros.

O erro epipolar obtido ao longo da sequência de quadros do vídeo é apresentado no gráfico da [Figura 50](#). As técnicas CSAT e CSAT-M obtiveram, respectivamente, valores de erro epipolar médio próximos a 0,6342 e 0,6891 *pixels*, variando nos intervalos de [0,2987; 1,1683] e [0,2668; 1,2715]. Por outro lado, a técnica da literatura, CSMF, alcançou erro epipolar médio próximo ao dobro da técnica proposta, com valor de 1,2179 *pixels*, apresentando uma variação entre o intervalo de [0,7856; 1,6990], sendo o seu erro mínimo adquirido superior à média das demais técnicas comparadas.

Figura 50 – Gráfico do erro epipolar obtido no Vídeo 01 do experimento no ambiente submarino, utilizando as técnicas CSMF, CSAT e CSAT-M.



Fonte: Elaborada pelo próprio autor.

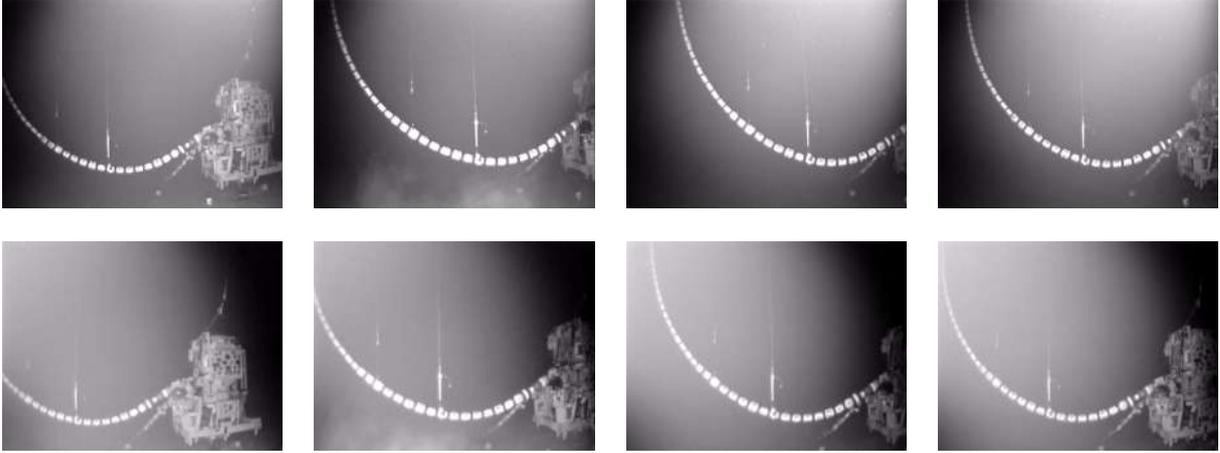
5.3.2.2 Vídeo 02

O vídeo 02 estudado nesse cenário dispõe de um total de 23172 quadros do objeto de interesse no ambiente submarino, encontrando-se ainda em posição horizontal. Amostras das imagens utilizadas para a validação desse cenário, obtidas pelas câmeras fixadas no ROV, são apresentadas na [Figura 51](#). Nas imagens, podem ser visualizados dois fatores que são capazes de prejudicar a avaliação desse cenário, sendo eles a presença de dois dutos ao mesmo tempo e partículas suspensas devido à proximidade ao fundo do oceano, onde encontra-se o poço de extração. Apesar da interferência que pode haver com a existência de dois dutos na mesma imagem, o SOIS capta apenas um, que está em primeiro plano e que possui mais pontos detectáveis, e por isso não ocorreu distorção na avaliação dos dados nesse estudo.

Assim como no vídeo 01, o duto apresentou-se, predominantemente, na região inferior da imagem, cobrindo praticamente toda extensão horizontal. Porém, a variação na disposição dos pontos detectados decresceu em razão do movimento mais restrito do duto, pelo fato do mesmo encontrar-se acoplado ao poço de extração durante toda a extensão do vídeo. Nessa avaliação, observou-se que as câmeras realizaram movimentos mais próximos ao objeto de interesse, sendo medidas distâncias com valores entre 9 e 16 metros, calculado através da média das reconstruções dos pontos ao longo da análise, aplicando cada calibração fornecida pelas técnicas de adaptação avaliadas.

O erro epipolar atingido durante a reconstrução utilizando a técnica CSMF variou entre 0,5583 e 1,6942 *pixels*, alcançando uma média de 1,1736 *pixels*. Por outro lado, a

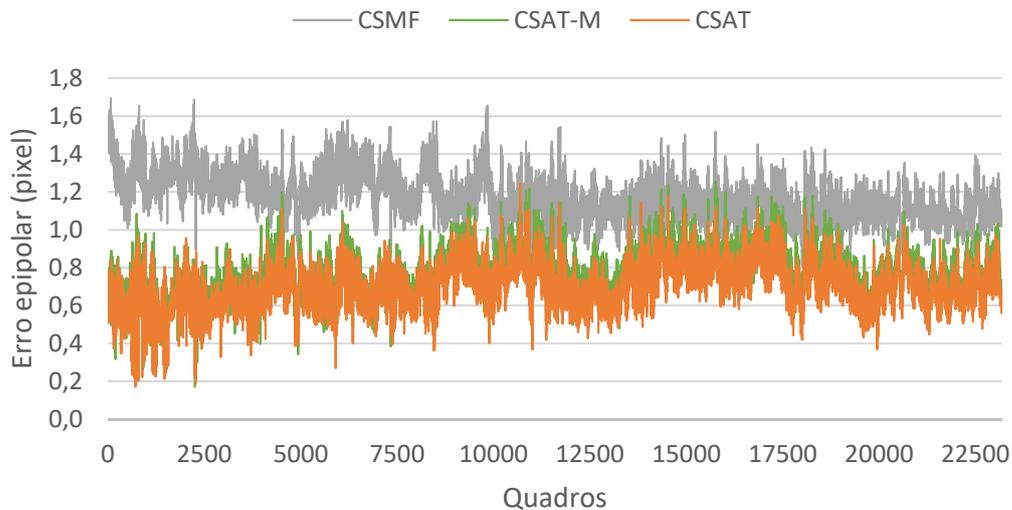
Figura 51 – Subconjunto das imagens do Vídeo 02 analisado durante o experimento no ambiente submarino. A primeira fileira de imagens obtidas corresponde à câmera posicionada na esquerda do ROV e a segunda as adquiridas pela câmera à direita, referindo-se aos quadros 1, 7500, 15000 e 22500 do vídeo.



Fonte: Petrobras.

técnica proposta, CSAT, apresentou uma média de 0,7019 *pixel*, com valores mínimo e máximo de 0,1708 e 1,2429 *pixels*, respectivamente. Já a CSAT-M (que combina CSMF e CSAT) obteve erros pertencentes ao intervalo de [0,1693; 1,2537] e média de 0,7649 *pixels*. A representação gráfica dos resultados de erro epipolar alcançados por toda a extensão do vídeo de validação são apresentados na [Figura 52](#).

Figura 52 – Gráfico do erro epipolar obtido no Vídeo 02 do experimento no ambiente submarino, utilizando as técnicas CSMF, CSAT e CSAT-M.

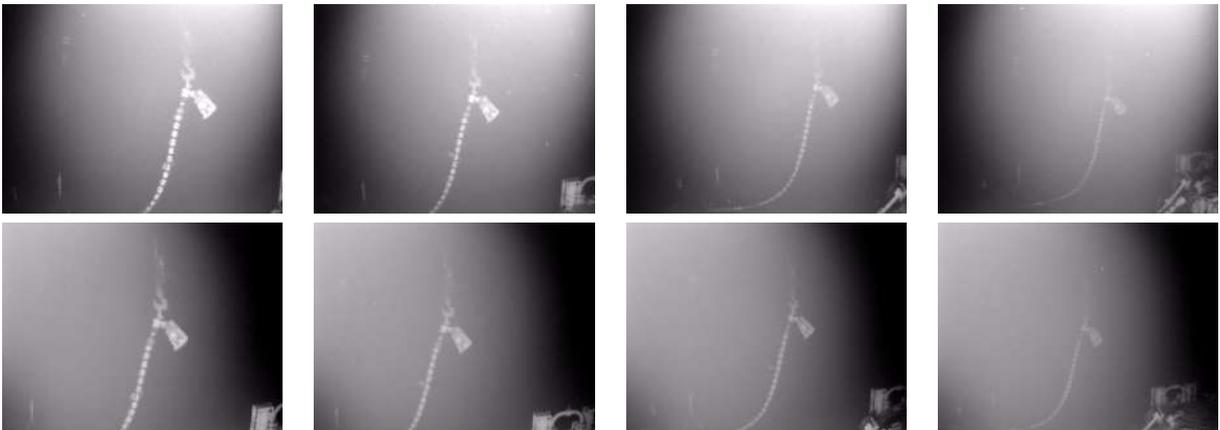


Fonte: Elaborada pelo próprio autor.

5.3.2.3 Vídeo 03

O terceiro vídeo analisado em ambiente submarino é composto por 18063 quadros, diferenciando-se dos vídeos anteriores devido à orientação com que o objeto de estudo apresenta-se nas imagens, estando na direção vertical. A [Figura 53](#) referencia um subconjunto da sequência de quadros fornecidos pelas câmeras fixadas no ROV e que serão utilizadas para a validação desse cenário. Como pode ser verificada na amostragem do vídeo, o duto apresenta uma predominância espacial na região central das imagens, com poucos pontos apresentando-se na parte inferior central da imagem. Outro fator a ser destacado é que a extremidade analisada do objeto de estudo não encontra-se fixada ao poço de extração. Além disso, nessa avaliação, percebeu-se que as câmeras realizaram movimentos mais longínquos ao duto, sendo medidas distâncias entre 10,5 e 26 metros, determinadas pela média das reconstruções dos pontos, com os resultados de calibração proporcionado pelas técnicas de adaptação avaliadas.

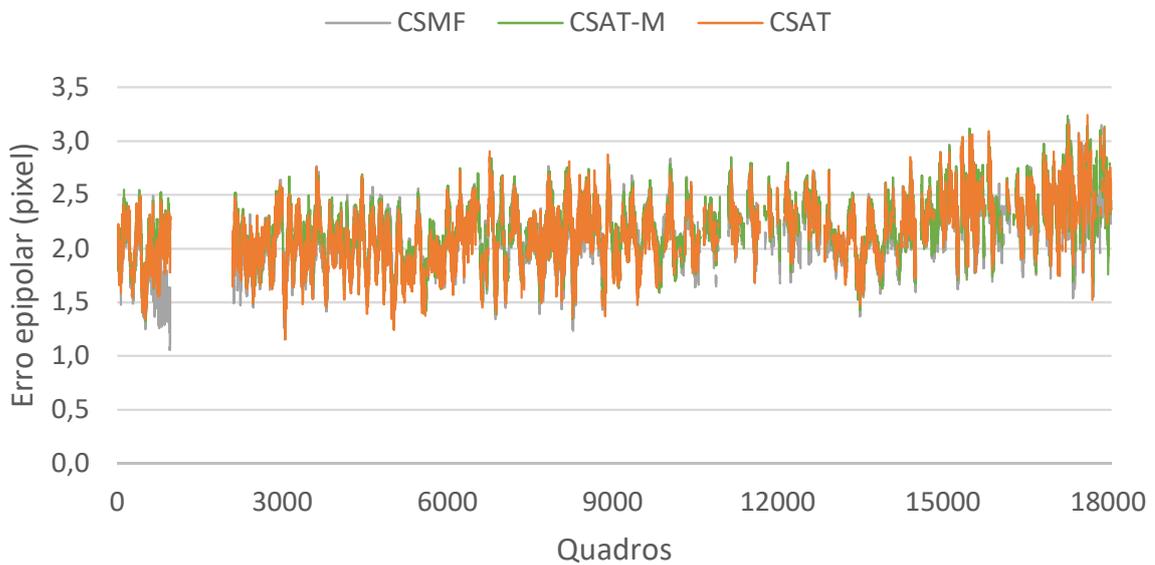
Figura 53 – Subconjunto das imagens do Vídeo 03 analisado durante o experimento no ambiente submarino. A primeira fileira de imagens obtidas corresponde à câmera posicionada na esquerda do ROV e a segunda as adquiridas pela câmera à direita, referindo-se aos quadros 1, 6000, 12000 e 18000 do vídeo.



Fonte: Petrobras.

O erro epipolar médio adquirido no decorrer da reconstrução do duto empregando o resultado da técnica CSMF foi de 2,1101 *pixels*, com variações mínima e máxima de 1,0602 e 3,2036, respectivamente. Com o uso da técnica CSAT-M, obteve-se média de 2,1591 *pixels* e variações dentro do intervalo fechado [1,2161 3,2340]. Enquanto os resultados adquiridos por CSAT variaram entre 1,1578 e 3,2410 (mínima e máxima), obtendo-se erro epipolar médio de 2,1226 *pixels*. Todos os resultados são interpretados do gráfico apresentado na [Figura 54](#), onde é possível visualizar algumas lacunas devido a falha na reconstrução do duto por falta de enquadramento ou a não detecção em uma das câmeras, por causa do ruído presente na imagem e o distanciamento para a câmera.

Figura 54 – Gráfico do erro epipolar obtido no Vídeo 03 do experimento no ambiente submarino, utilizando as técnicas CSMF, CSAT e CSAT-M.



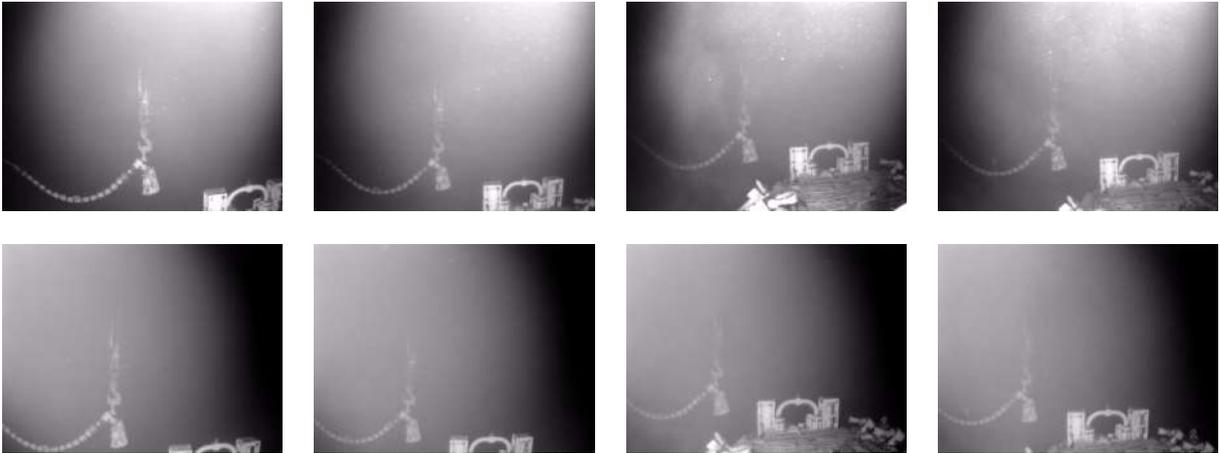
Fonte: Elaborada pelo próprio autor.

5.3.2.4 Vídeo 04

O último vídeo avaliado para cenário submarino possui cerca de apenas 40 segundos, sendo composto por um total de 1114 quadros. Assim como nos dois primeiros vídeos, o duto encontra-se em posição horizontal, como apresentado no subconjunto da sequência de quadros na [Figura 55](#). A distribuição dos pontos característicos do objeto de interesse apresentaram-se, predominantemente, no canto inferior esquerdo da imagem, não apresentando grande variação em sua disposição devido à baixa amostragem de quadros analisados. Nessa análise, constatou-se que as câmeras realizaram movimentos, variando a distância para o duto entre as medidas de 14,5 e 19 metros. Assim como nos demais vídeos, esses valores de intervalo foram determinados através da média das reconstruções dos pontos, empregando todas as técnicas de adaptação avaliadas nesse cenário.

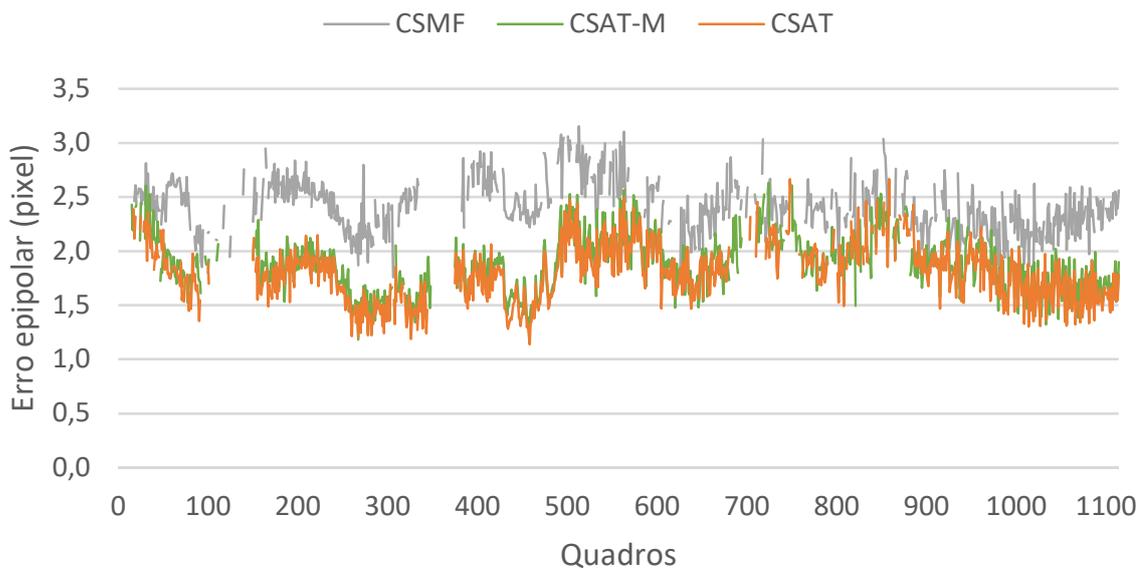
Como demonstrado no gráfico da [Figura 56](#), a técnica CSMF obteve erro epipolar variando entre [1,7588; 3,1528], apresentando a média de 2,4159 *pixels*. Para a técnica proposta CSAT-M, os valores mínimo e máximo foram de 1,1830 e 2,6301 *pixels*, desempenhando um valor médio de 1,8680 *pixels*. Calculou-se, para a técnica proposta CSAT, valores enquadrados no intervalo de 1,1383 a 2,7018 *pixels*, possuindo um erro epipolar médio de 1,7972 *pixels*. As falhas apresentadas na reconstrução do duto nesse cenário ocorreram, diferentemente do vídeo 03, exclusivamente à não detecção em uma das câmeras envolvidas, devido ao acúmulo do ruído presente na imagem e a extensa distância para o objeto.

Figura 55 – Subconjunto das imagens do Vídeo 04 analisado durante o experimento no ambiente submarino. A primeira fileira de imagens obtidas corresponde à câmera posicionada na esquerda do ROV e a segunda as adquiridas pela câmera à direita, referindo-se aos quadros 1, 350, 700 e 1050 do vídeo.



Fonte: Petrobras.

Figura 56 – Gráfico do erro epipolar obtido no Vídeo 04 do experimento no ambiente submarino, utilizando as técnicas CSMF, CSAT e CSAT-M.

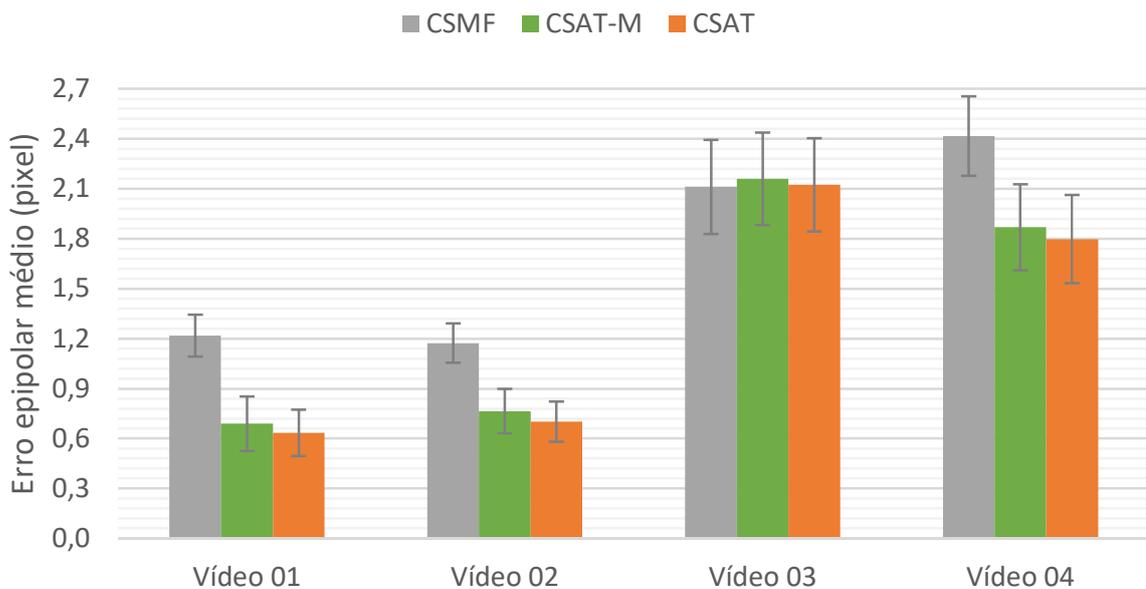


Fonte: Elaborada pelo próprio autor.

5.3.2.5 Comparação dos resultados

Sumarizando os resultados obtidos nos quatro vídeos apresentados para o ambiente submarino, podemos realizar uma comparação nos dados processados. Na [Figura 57](#) são apresentados os valores dos erros epipolares médios alcançados, assim como o desvio padrão de cada amostra, durante a reconstrução do duto, em cada um dos vídeos, empregando as técnicas de adaptação utilizadas nesta dissertação. Os vídeos 01 e 02 apresentaram comportamentos semelhantes, onde a adaptação da literatura, CSMF, possui erro médio próximo ao dobro das demais técnicas, sendo o procedimento proposto, CSAT, ligeiramente melhor do que CSAT-M, composição entre CSMF e CSAT. Quando é considerado o desvio padrão das amostras, as técnicas CSAT e CSAT-M denotam equivalência nos erros obtidos e permanecem com desempenho superiores à CSMF.

Figura 57 – Gráfico comparativo entre as técnicas de adaptação usando erro epipolar médio, considerando desvio padrão. Vídeos analisados dos experimentos em ambiente submarino.



Fonte: Elaborada pelo próprio autor.

Nos vídeos 03 e 04 são apresentados erros epipolares médios mais elevados do que nos demais analisados. Esse fato ocorreu, possivelmente, devido a dois fatores: acurácia da calibração seca e distância do duto para a câmera. Como esses vídeos foram realizados em dias diferente dos demais, a calibração seca efetuada antes da operação pode ter estimado parâmetros menos corretos para descrever a relação entre as câmeras envolvidas no processo de reconstrução. O outro fator que influencia no aumento desse erro é o incremento na distância da câmera para o objeto a ser reconstruído, uma vez que, quanto maior essa distância mais, imprecisa é a detecção dos pontos característicos do objeto. Como apresentado anteriormente, nos vídeos 03 e 04 as distâncias são superiores

às encontradas nos vídeos 01 e 02, causando assim um aumento no erro epipolar médio alcançado independente da técnica utilizada.

Independentemente do erro epipolar médio mais elevado, o comportamento apresentado no vídeo 04 se mantém o mesmo indicado pelos vídeos 01 e 02, onde o método CSMF possui erro maior do que os obtidos pelas adaptações CSAT e CSAT-M. Apesar de uma discrepância menor entre a diferença alcançada entre os procedimentos de adaptação e um maior desvio padrão envolvido nas amostras, não é possível considerar a técnica CSMF equivalente às demais. Em contrapartida, o vídeo 03 apresenta uma conduta diferente dos outros, uma vez que todos os procedimentos de adaptação expressam erros extremamente similares e, quando consideramos o desvio padrão das amostras, todas as técnicas revelam-se equivalentes.

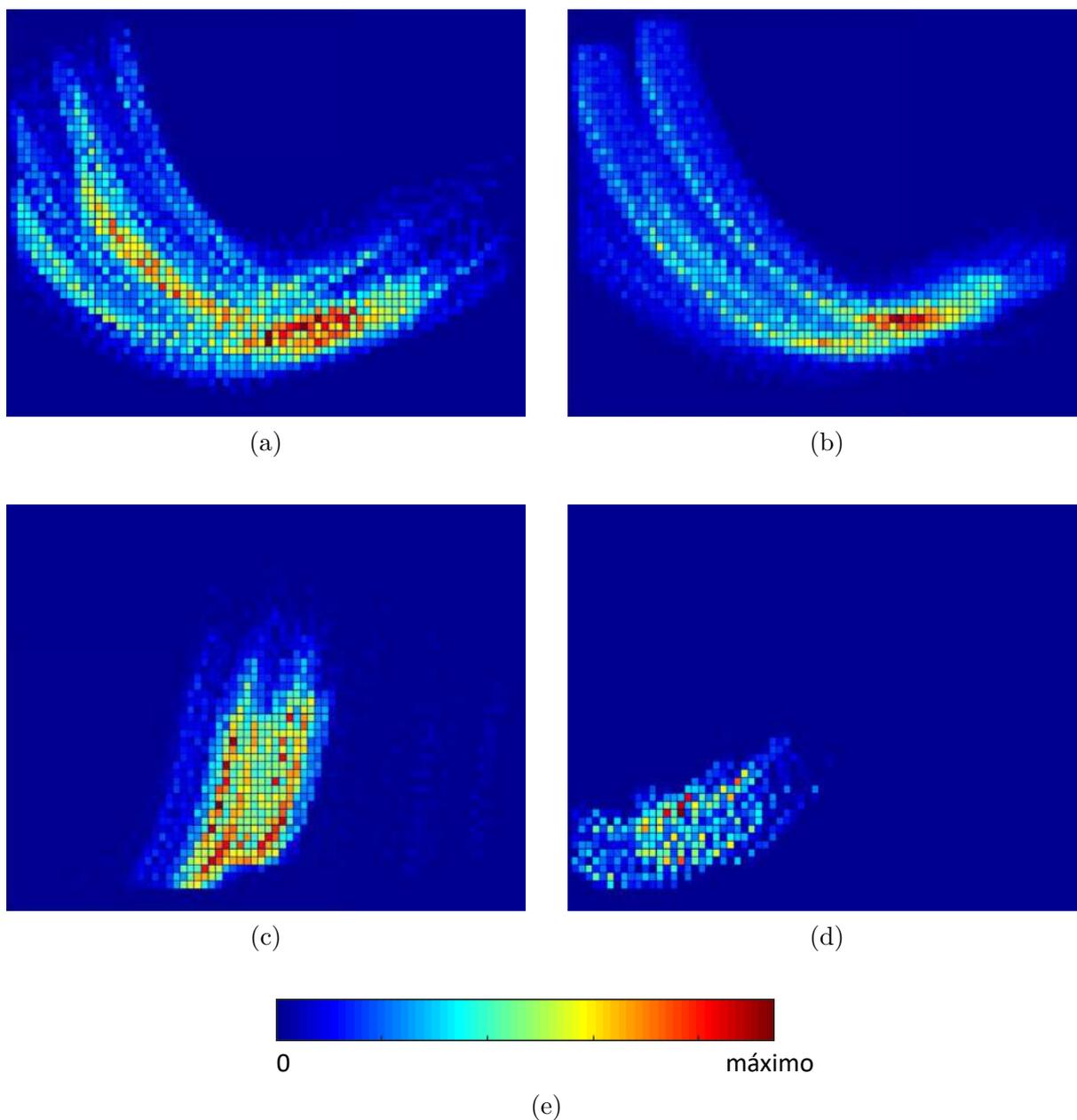
O comportamento encontrado no vídeo 03 corrobora com as conclusões obtidas nos experimentos sintéticos e ratificada no cenário real do tanque de água, embora tenha se utilizado uma métrica diferente. Isso porque, nesse vídeo, o objeto de validação encontra-se predominantemente na região central da imagem, espaço onde o efeito causado pela distorção, tanto pela lente da câmera quanto a refração no caminho de luz, é mínimo. As localizações dos pontos detectados no duto ao longo de cada vídeo analisado para o ambiente submarino são apresentadas na [Figura 58](#) através de um mapa de calor, onde o menor valor no mapa possui cor azul escuro, o maior valor cor vermelho brilhante e os valores medianos tons de verde (como representado na [Figura 58e](#)).

Todos os dados de erro epipolar médio, mínimo e máximo, além dos valores de desvio padrão apresentados como resultados ao longo dos experimentos no ambiente submarino estão relatados nas [Tabela 9](#), presente no [Apêndice B](#).

5.4 CONSIDERAÇÕES FINAIS

Neste capítulo foram apresentados os dados obtidos durante o experimento desta dissertação, evidenciando os resultados adquiridos a partir de experimentos realizados nos ambientes sintético e reais, com o intuito de comparar a técnica proposta para adaptar uma calibração seca para o ambiente subaquático (CSAT) com as existentes na literatura (CM e CSMF), além de um procedimento criado a partir da combinação entre CSMF e CSAT, denominado de CSAT-M. Os experimentos sintéticos buscaram verificar a influência de determinadas variáveis existentes no cenário onde uma câmera, protegida por uma caixa de estanque em um ambiente seco, é submersa no meio aquático. Por outro lado, os experimentos reais foram investigados com a finalidade de corroborar com as conclusões encontradas nos cenários sintéticos em ambientes onde soluções de visão computacional são aplicadas.

Figura 58 – Mapa de calor da posição dos pontos detectados ao longo de cada vídeo analisado para o ambiente submarino. Apresenta a distribuição obtida para os vídeos 01 (a), 02 (b), 03 (c) e 04 (d), assim como a representação da escala de cor considerada (e).



Fonte: Elaborada pelo próprio autor.

6 CONCLUSÃO

Essa dissertação propôs uma técnica de calibração capaz de adaptar os parâmetros de câmera para ambientes com diferentes índices de refração, sendo projetado e representado, matematicamente, um modelo de adaptação que incorpora a distorção ocasionada pela mudança de meio. Em conjunto com a técnica proposta, foi realizado um estudo investigativo sobre as técnicas de calibração para o ambiente subaquático com o objetivo de sumarizar o desenvolvimento da área, facilitando a expansão e reutilização dessa pesquisa. Nessa investigação, foram apresentadas duas abordagens comumente utilizadas pela comunidade científica para resolver o problema existente sobre a calibração em ambientes subaquáticos.

A primeira abordagem consiste em um modelo mais simples que considera a câmera com um único centro de projeção, enquanto a segunda considera diversos centros de projeções para formar uma única imagem, aumentando assim a complexidade do modelo utilizado. Devido a relação de custo-benefício, a primeira abordagem foi selecionada para o estudo comparativo existente nessa dissertação. Para tal, a técnica proposta, denominada CSAT, foi comparada com dois outros algoritmos já existentes (CM e CSMF) e com uma variação, também proposta nessa dissertação, que combina o CSAT com uma adaptação da literatura CSMF, denominada de CSAT-M.

Durante as análises dos resultados apresentados pelos experimentos sintéticos e reais, pode-se concluir que a reconstrução de objetos localizados no centro da imagem utilizando os parâmetros fornecidos por qualquer uma das técnicas de calibração produzem erros similares. Uma vez que neste espaço da imagem existe baixa intensidade de distorção, tanto produzida pela lente da câmera como pela mudança de meio. Além disso, foi avaliado que com o objeto sendo apresentado nas bordas da imagem, a técnica de adaptação da literatura CSMF expressou erros mais elevados do que CM e as técnicas propostas CSAT e CSAT-M. Este aumento do erro apresentado por CSMF é justificado pela ausência da modelagem mais completa dos efeitos de perturbação causados pela refração do raio de luz durante a formação da imagem.

Nos cenários sintéticos, onde nenhum tipo de ruído foi inserido, a técnica CM apresentou resultados melhores aos das adaptações propostas, sendo observado ainda que para CM produzir resultados de reconstrução melhores, é necessário realizar o procedimento de calibração apresentando o padrão próximo a localização 3D de onde o objeto estará durante a aplicação. Em contrapartida, quando os experimentos reais foram realizados, os erros obtidos nas reconstruções foram similares, principalmente considerando o desvio padrão obtido em cada avaliação. Este fato indica que os ruídos presentes no ambiente subaquático degradam o resultado do procedimento de calibração efetuado nesse meio.

6.1 LIMITAÇÕES

Por consequência do desvio ocorrido no percurso realizado pelo raio de luz estar diretamente relacionado com o formato da superfície em que é refratado, a modelagem proposta para a adaptação dos parâmetros de câmera para diferentes meios de aplicação limitou-se a caixas de estaque que possuem interface plana, não incorporando em seu desenvolvimento a interface no formato de domo. Outra limitação deste trabalho a ser considerada relaciona-se aos cenários em que o objeto de estudo da aplicação está localizado muito próximo a câmera, infringindo o pressuposto assumido durante a modelagem proposta sobre a relação entre as distâncias interface-câmera e objeto-câmera, causando aumento nos erros produzidos.

6.2 CONTRIBUIÇÕES

As principais contribuições constatadas no decorrer desta dissertação foram:

1. Apanhado literário acerca dos trabalhos acadêmicos que realizam calibrações em ambientes subaquáticos, categorizando-as em técnicas que utilizam um modelo de câmera com ponto de vista único (SVP) e múltiplos pontos de vista (nSVP).;
2. A formulação matemática de um modelo genérico para adaptação dos parâmetros de câmera para diferentes ambientes de aplicação, capaz de incorporar o efeito de distorção introduzida pela refração dos raios de luz ocorrida durante a formação da imagem;
3. Os algoritmos de adaptação propostos para obter parâmetros que descrevam o comportamento da câmera em meio diferente ao que originalmente foi calibrado. Essas técnicas possuem como vantagem executar a adaptação sem requerer a aquisição de novas imagens, utilizando apenas informações sobre os índices de refração dos meios envolvidos no processo;
4. Implementação de calibrações existentes na literatura e a criação de cenários sintéticos capazes de explorar a influência de algumas variáveis existentes em ambientes subaquáticos;
5. A avaliação comparativa entre as técnicas propostas e as existentes na literatura, que empregam o modelo SVP, em cenários relevantes para aplicações subaquáticas de visão computacional, tanto simulado sinteticamente quanto em ambientes reais de aplicações.

6.3 TRABALHOS FUTUROS

Os principais objetivos desta dissertação foram alcançados com sucesso, contudo, ainda existem algumas contribuições que podem ser desenvolvidas no futuro. Destacando-se a expansão do modelo de adaptação para utilizá-lo também em câmeras que possuam caixa de estanque com interface no formato de domo, ampliando a generalidade do modelo proposto através da incorporação de diferentes superfícies onde o raio de luz é refratado. Outro ponto refere-se ao desenvolvimento de experimentos para as técnicas que utilizam a abordagem nSVP, realizando uma análise sobre a complexidade desse modelo e quais os efeitos causados nas aplicações de visão computacional, comparando-as com as técnicas apresentadas nesta dissertação.

REFERÊNCIAS

- ABDEL-AZIZ, Y. I.; KARARA, H. M. Direct linear transformation from comparator coordinates into object space coordinates in close-range photogrammetry. In: *Proceedings of the Symposium on Close-Range photogrammetry*. [S.l.: s.n.], 1971. v. 1, p. 18.
- AGARWAL, S.; MIERLE, K.; OTHERS. *Ceres Solve*. 2010. <<http://ceres-solver.org>>. Acessado em 05 de Novembro de 2017.
- AGRAWAL, A.; RAMALINGAM, S.; TAGUCHI, Y.; CHARI, V. A theory of multi-layer flat refractive geometry. In: IEEE. *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*. [S.l.], 2012. p. 3346–3353.
- ARYA, S.; MOUNT, D. M.; NETANYAHU, N. S.; SILVERMAN, R.; WU, A. Y. An optimal algorithm for approximate nearest neighbor searching fixed dimensions. *Journal of the ACM (JACM)*, ACM, v. 45, n. 6, p. 891–923, 1998.
- AUTODESK. *3ds Max - 3d modeling, animation and rendering software*. 2017. <<https://www.autodesk.com/products/3ds-max/overview>>. Acessado em 08 de Novembro de 2017.
- BALDWIN, R. An underwater photogrammetric measurement system for structural inspection. *Int. Arch. Photogramm*, v. 25, p. 9–18, 1984.
- BIANCO, G.; GALLO, A.; BRUNO, F.; MUZZUPAPPA, M. A comparison between active and passive techniques for underwater 3d applications. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci*, v. 34, p. 357–363, 2011.
- BINGHAM, B.; FOLEY, B.; SINGH, H.; CAMILLI, R.; DELAPORTA, K.; EUSTICE, R.; MALLIOS, A.; MINDELL, D.; ROMAN, C.; SAKELLARIOU, D. Robotic tools for deep water archaeology: Surveying an ancient shipwreck with an autonomous underwater vehicle. *Journal of Field Robotics*, Wiley Online Library, v. 27, n. 6, p. 702–717, 2010.
- BOSCH, J.; GRACIAS, N.; RIDAO, P.; RIBAS, D.; ISTENIČ, K.; GARCIA, R.; ROSSI, I. R. Immersive touring for marine archaeology. application of a new compact omnidirectional camera to mapping the gnalić shipwreck with an auv. In: SPRINGER. *Iberian Robotics conference*. [S.l.], 2017. p. 183–195.
- BOUGUET, J. Y. *MATLAB calibration tool*. 2004. <http://www.vision.caltech.edu/bouguetj/calib_doc/>. Acessado em 20 de janeiro de 2015.
- BRADSKI, G. The OpenCV Library. *Dr. Dobb's Journal of Software Tools*, 2000.
- BRANDOU, V.; ALLAIS, A.-G.; PERRIER, M.; MALIS, E.; RIVES, P.; SARRAZIN, J.; SARRADIN, P.-M. 3d reconstruction of natural underwater scenes using the stereovision system iris. In: IEEE. *OCEANS 2007-Europe*. [S.l.], 2007. p. 1–6.
- BRÄUER-BURCHARDT, C.; HEINZE, M.; SCHMIDT, I.; KÜHMSTEDT, P.; NOTNI, G. Underwater 3d surface measurement using fringe projection based scanning devices. *Sensors*, Multidisciplinary Digital Publishing Institute, v. 16, n. 1, p. 13, 2015.

- BRYANT, M.; WETTERGREEN, D.; ABDALLAH, S.; ZELINSKY, A. et al. Robust camera calibration for an autonomous underwater vehicle. In: *Proc. Australian Conf. on Robotics and Autom.* [S.l.: s.n.], 2000. p. 111–116.
- CERVERI, P.; BORGHESE, N.; PEDOTTI, A. Complete calibration of a stereo photogrammetric system through control points of unknown coordinates. *Journal of biomechanics*, Elsevier, v. 31, n. 10, p. 935–940, 1998.
- CESERACCIU, E.; SAWACHA, Z.; FANTOZZI, S.; CORTESI, M.; GATTA, G.; CORAZZA, S.; COBELLI, C. Markerless analysis of front crawl swimming. *Journal of biomechanics*, Elsevier, v. 44, n. 12, p. 2236–2242, 2011.
- CHARI, V.; STURM, P. Multiple-view geometry of the refractive plane. In: THE BRITISH MACHINE VISION ASSOCIATION (BMVA). *BMVC 2009-20th British Machine Vision Conference*. [S.l.], 2009. p. 1–11.
- CHEN, X.; YANG, Y.-H. A closed-form solution to single underwater camera calibration using triple wavelength dispersion and its application to single camera 3d reconstruction. *IEEE Transactions on Image Processing*, IEEE, v. 26, n. 9, p. 4553–4561, 2017.
- DOLEREIT, T.; KUIJPER, A. Converting underwater imaging into imaging in air. In: IEEE. *Computer Vision Theory and Applications (VISAPP), 2014 International Conference on*. [S.l.], 2014. v. 1, p. 96–103.
- DOLEREIT, T.; LUKAS, U. F. von; KUIJPER, A. Underwater stereo calibration utilizing virtual object points. In: IEEE. *OCEANS 2015-Genova*. [S.l.], 2015. p. 1–7.
- DOUCETTE, J. S.; HARVEY, E. S.; SHORTIS, M. R. Stereo–video observation of nearshore bedforms on a low energy beach. *Marine Geology*, Elsevier, v. 189, n. 3, p. 289–305, 2002.
- DRAP, P.; SEINTURIER, J.; SCARADOZZI, D.; GAMBONI, P.; LONG, L.; GAUCH, F. Photogrammetry for virtual exploration of underwater archeological sites. In: *Proceedings of the 21st international symposium, CIPA*. [S.l.: s.n.], 2007. p. 1e6.
- DUANE, C. B. Close-range camera calibration. *Photogramm. Eng.*, v. 37, n. 8, p. 855–866, 1971.
- EUSTICE, R.; PIZARRO, O.; SINGH, H. Visually augmented navigation for autonomous underwater vehicles. *Oceanic Engineering, IEEE Journal of*, v. 33, n. 2, p. 103–122, April 2008. ISSN 0364-9059.
- FERREIRA, R.; COSTEIRA, J. P.; SANTOS, J. A. Stereo reconstruction of a submerged scene. In: SPRINGER. *Iberian Conference on Pattern Recognition and Image Analysis*. [S.l.], 2005. p. 102–109.
- FORESTI, G. L. Visual inspection of sea bottom structures by an autonomous underwater vehicle. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, IEEE, v. 31, n. 5, p. 691–705, 2001.
- FRYER, J.; FRASER, C. On the calibration of underwater cameras. *The Photogrammetric Record*, Wiley Online Library, v. 12, n. 67, p. 73–85, 1986.

-
- GEDGE, J. *Underwater stereo matching and its calibration*. Dissertação (Mestrado) — University of Alberta Libraries, EUA, 2011.
- GEDGE, J.; GONG, M.; YANG, Y.-H. Refractive epipolar geometry for underwater stereo matching. In: IEEE. *Computer and Robot Vision (CRV), 2011 Canadian Conference on*. [S.l.], 2011. p. 146–152.
- GEOMETRIC. *Vision Measurement System*. 2001. Acessado em 06-10-2017. Disponível em: <<http://www.geomsoft.com/VMS/>>.
- GUENNEBAUD, G.; JACOB, B. et al. *Eigen v3*. 2010. <<http://eigen.tuxfamily.org>>. Acessado em 05 de Novembro de 2017.
- HANSEN, N.; OSTERMEIER, A. Completely derandomized self-adaptation in evolution strategies. *Evolutionary computation*, MIT Press, v. 9, n. 2, p. 159–195, 2001.
- HARTLEY, R. I.; ZISSERMAN, A. *Multiple View Geometry in Computer Vision*. Second. [S.l.]: Cambridge University Press, ISBN: 0521540518, 2004.
- HARVEY, E.; CAPPO, M.; SHORTIS, M.; ROBSON, S.; BUCHANAN, J.; SPEARE, P. The accuracy and precision of underwater measurements of length and maximum body depth of southern bluefin tuna (*thunnus maccoyii*) with a stereo–video camera system. *Fisheries Research*, Elsevier, v. 63, n. 3, p. 315–326, 2003.
- HATZE, H. High-precision three-dimensional photogrammetric calibration and object space reconstruction using a modified dlt-approach. *Journal of biomechanics*, Elsevier, v. 21, n. 7, p. 533–538, 1988.
- HEIKKILA, J.; SILVEN, O. A four-step camera calibration procedure with implicit image correction. In: IEEE. *Computer Vision and Pattern Recognition, 1997. Proceedings., 1997 IEEE Computer Society Conference on*. [S.l.], 1997. p. 1106–1112.
- HU, H. *Euclidean reconstruction of natural underwater scenes using optic imagery sequence*. Dissertação (Mestrado) — University of New Hampshire, EUA, 2015.
- JESUS, K. de; JESUS, K. de; FIGUEIREDO, P. A.; GONÇALVES, P.; PAULO, J. Effects of fatigue on kinematical parameters during submaximal and maximal. *Journal of Applied Biomechanics 2012 Human Kinetics, Inc*, p. 2, 2012.
- JONSSON, P. *Underwater video systems*. 2013. Acessado em 29-09-2017. Disponível em: <http://www.coastalwiki.org/wiki/Underwater_video_systems>.
- JORDT-SEDLAZECK, A.; KOCH, R. Refractive calibration of underwater cameras. *Computer Vision–ECCV 2012*, Springer, p. 846–859, 2012.
- KANG, L.; WU, L.; YANG, Y.-H. Experimental study of the influence of refraction on underwater three-dimensional reconstruction using the svp camera model. *Applied optics*, Optical Society of America, v. 51, n. 31, p. 7591–7603, 2012.
- LAVEST, J.-M.; RIVES, G.; LAPRESTÉ, J.-T. Dry camera calibration for underwater applications. *Machine Vision and Applications*, Springer, v. 13, n. 5, p. 245–253, 2003.
- LEVENBERG, K. A method for the solution of certain non-linear problems in least squares. *Quarterly Journal of Applied Mathematics*, II, n. 2, p. 164–168, 1944.

- LI, W.; GEE, T.; DELMAS, P.; FRIEDRICH, H. Evaluating using gopro cameras and tsai's calibration for video-based submerged river-bed reconstruction. In: IEEE. *Image and Vision Computing New Zealand (IVCNZ), 2016 International Conference on*. [S.l.], 2016. p. 1–6.
- LIVNY, B. *MENTAL RAY, FOR MAYA, 3DS MAX, AND XSI: A 3D ARTIST'S GUIDE TO RENDERING (With CD)*. [S.l.]: John Wiley & Sons, 2008.
- LOWE, D. G. Object recognition from local scale-invariant features. In: IEEE. *Computer vision, 1999. The proceedings of the seventh IEEE international conference on*. [S.l.], 1999. v. 2, p. 1150–1157.
- MARQUARDT, D. W. An algorithm for least-squares estimation of nonlinear parameters. *Journal of the Society for Industrial and Applied Mathematics*, Society for Industrial and Applied Mathematics, v. 11, n. 2, p. pp. 431–441, 1963. ISSN 03684245. Disponível em: <<http://www.jstor.org/stable/2098941>>.
- MCCABE, C. B.; PSYCHARAKIS, S.; SANDERS, R. Kinematic differences between front crawl sprint and distance swimmers at sprint pace. *Journal of Sports Sciences*, Routledge, v. 29, n. 2, p. 115–123, 2011.
- MELINE, A.; TRIBOULET, J.; JOUVENCEL, B. A camcorder for 3d underwater reconstruction of archeological objects. In: IEEE. *OCEANS 2010*. [S.l.], 2010. p. 1–9.
- MIP, CAU Kiel, Germany. *BIAS: Basic Image AlgorithmS Library*. 2015. Acessado em 17-10-2017. Disponível em: <<http://www.mip.informatik.uni-kiel.de/BIAS>>.
- NEGAHDARIPOUR, S.; FIROOZFAM, P. An roV stereovision system for ship-hull inspection. *IEEE Journal of oceanic engineering*, IEEE, v. 31, n. 3, p. 551–564, 2006.
- O'BYRNE, M.; PAKRASHI, V.; SCHOEFS, F.; GHOSH, B. A comparison of image based 3d recovery methods for underwater inspections. In: *EWSHM-7th European Workshop on Structural Health Monitoring*. [S.l.: s.n.], 2014.
- OLEARI, F.; KALLASI, F.; RIZZINI, D. L.; ALEOTTI, J.; CASELLI, S. An underwater stereo vision system: from design to deployment and dataset acquisition. In: IEEE. *OCEANS 2015-Genova*. [S.l.], 2015. p. 1–6.
- ONMEK, Y.; TRIBOULET, J.; DRUON, S.; MELINE, A.; JOUVENCEL, B. Evaluation of underwater 3d reconstruction methods for archaeological objects: Case study of anchor at mediterranean sea. In: IEEE. *Control, Automation and Robotics (ICCAR), 2017 3rd International Conference on*. [S.l.], 2017. p. 394–398.
- OSBOROUGH, C.; DALY, D.; PAYTON, C. Effect of swim speed on leg-to-arm coordination in unilateral arm amputee front crawl swimmers. *Journal of sports sciences*, Routledge, v. 33, n. 14, p. 1523–1531, 2015.
- PESSEL, N.; OPDERBECKE, J.; ALDON, M.-J. Camera self-calibration in underwater environment. UNION Agency, 2003.
- PETRELL, R.; SHI, X.; WARD, R.; NAIBERG, A.; SAVAGE, C. Determining fish size and swimming speed in cages and tanks using simple video techniques. *Aquacultural Engineering*, Elsevier, v. 16, n. 1-2, p. 63–84, 1997.

- PUEL, F.; MORLIER, J.; AVALOS, M.; MESNARD, M.; CID, M.; HELLARD, P. 3d kinematic and dynamic analysis of the front crawl tumble turn in elite male swimmers. *Journal of biomechanics*, Elsevier, v. 45, n. 3, p. 510–515, 2012.
- RAHMAN, T.; ANDERSON, J.; WINGER, P.; KROUGLICOF, N. Calibration of an underwater stereoscopic vision system. In: IEEE. *Oceans-San Diego, 2013*. [S.l.], 2013. p. 1–6.
- ROMAN, C.; INGLIS, G.; RUTTER, J. Application of structured light imaging for high resolution mapping of underwater archaeological sites. In: IEEE. *OCEANS 2010 IEEE-Sydney*. [S.l.], 2010. p. 1–9.
- ROSEN, S.; JÖRGENSEN, T.; HAMMERSLAND-WHITE, D.; HOLST, J. C. Deepvision: a stereo camera system provides highly accurate counts and lengths of fish passing inside a trawl. *Canadian Journal of Fisheries and Aquatic Sciences*, NRC Research Press, v. 70, n. 10, p. 1456–1467, 2013.
- SANTANA-GARCON, J.; NEWMAN, S. J.; HARVEY, E. S. Development and validation of a mid-water baited stereo-video technique for investigating pelagic fish assemblages. *Journal of Experimental Marine Biology and Ecology*, Elsevier, v. 452, p. 82–90, 2014.
- SANTOS, I.; VARDARO, E.; GOES, E.; LOPES, V.; VAILLANT, A.; PALMEIRO, A.; KELNER, J.; CESAR, V.; PESSOA, S.; REIS, B. Real time radius of curvature measurement during dvc operations based on flexible pipe 3d reconstruction. In: OFFSHORE TECHNOLOGY CONFERENCE. *OTC Brasil*. [S.l.], 2015.
- SCHECHNER, Y. Y.; KARPEL, N. Clear underwater vision. In: IEEE. *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*. [S.l.], 2003. v. 1, p. I–I.
- SEDLAZECK, A.; KOCH, R. Calibration of housing parameters for underwater stereo-camera rigs. In: *BMVC*. [S.l.: s.n.], 2011. p. 1–11.
- SEDLAZECK, A.; KOCH, R. Perspective and non-perspective camera models in underwater imaging—overview and error analysis. In: *Outdoor and Large-Scale Real-World Scene Analysis*. [S.l.]: Springer, 2012. p. 212–242.
- SHORTIS, M. Calibration techniques for accurate measurements by underwater camera systems. *Sensors*, Multidisciplinary Digital Publishing Institute, v. 15, n. 12, p. 30810–30826, 2015.
- SHORTIS, M. R.; HARVEY, E. S. Design and calibration of an underwater stereo-video system for the monitoring of marine fauna populations. *International Archives of Photogrammetry and Remote Sensing*, INTERNATIONAL SOCIETY FOR PHOTOGRAMMETRY & REMOTE, v. 32, p. 792–799, 1998.
- SILVA, A. F.; FIGUEIREDO, P.; SEIFERT, L.; SOARES, S.; VILAS-BOAS, J. P.; FERNANDES, R. J. Backstroke technical characterization of 11-13 year-old swimmers. *Journal of sports science & medicine*, Dept. of Sports Medicine, Medical Faculty of Uludag University, v. 12, n. 4, p. 623, 2013.

-
- SILVATTI, A. P.; DIAS, F. A. S.; CERVERI, P.; BARROS, R. M. Comparison of different camera calibration approaches for underwater applications. *Journal of biomechanics*, Elsevier, v. 45, n. 6, p. 1112–1116, 2012.
- SILVATTI, A. P. et al. Calibração de câmeras para análises subaquáticas de movimentos. [sn], 2013.
- TELEM, G.; FILIN, S. Photogrammetric modeling of underwater environments. *ISPRS Journal of Photogrammetry and Remote Sensing*, Elsevier, v. 65, n. 5, p. 433–444, 2010.
- TOMASI, C.; KANADE, T. Detection and tracking of point features. School of Computer Science, Carnegie Mellon Univ. Pittsburgh, 1991.
- TRAFFELET, L.; EPPENBERGER, T.; MILLANE, A.; SCHNEIDER, T.; SIEGWART, R. Target-based calibration of underwater camera housing parameters. In: IEEE. *Safety, Security, and Rescue Robotics (SSRR), 2016 IEEE International Symposium on*. [S.l.], 2016. p. 201–206.
- TREIBITZ, T.; SCHECHNER, Y.; KUNZ, C.; SINGH, H. Flat refractive geometry. *IEEE transactions on pattern analysis and machine intelligence*, IEEE, v. 34, n. 1, p. 51–65, 2012.
- TSAI, R. A versatile camera calibration technique for high-accuracy 3d machine vision metrology using off-the-shelf tv cameras and lenses. *Robotics and Automation, IEEE Journal of*, v. 3, n. 4, p. 323–344, August 1987.
- YAU, T.; GONG, M.; YANG, Y.-H. Underwater camera calibration using wavelength triangulation. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. [S.l.: s.n.], 2013. p. 2499–2506.
- ZHANG, Z. A flexible new technique for camera calibration. *IEEE Trans. Pattern Anal. Mach. Intell.*, IEEE Computer Society, Washington, DC, USA, v. 22, n. 11, p. 1330–1334, nov. 2000. ISSN 0162-8828. Disponível em: <<http://dx.doi.org/10.1109/34.888718>>.

APÊNDICE A – RESULTADOS DOS CENÁRIOS SINTÉTICOS

Este apêndice apresenta os resultados complementares, em tabelas, obtidos nos seis cenários sintéticos criados neste trabalho. A primeira coluna de cada tabela descreve os valores das variáveis analisadas em cada cena, onde:

- As tabelas 1 e 2 indicam as distâncias entre a câmera e a interface da caixa de estanque;
- As tabelas 3 e 4 apresentam as distâncias entre a câmera e o objeto alvo de análise;
- As tabelas 5 e 6 retratam os valores de distorção cúbica aplicados na lente da câmera utilizada no *software* 3DS Max para a criação do cenário.

As demais colunas apresentam os resultados de RMS, erro médio de reconstrução (E_μ), desvio padrão do erro (E_σ) e erro máximo (E_{max}) obtidos pela calibração molhada (CM), técnica de adaptação proposta por Lavest et al. (CSMF) e as técnicas propostas nessa dissertação (CSAT e CSAT-M). Os resultados de RMS são demonstrados em unidade de *pixel*, enquanto as demais informações métricas em milímetros.

Tabela 1 – Resultado do cenário sintético 01 com objeto de validação fixo à 500 mm da câmera

| Distância | CM | | | CSMF | | | CSAT-M | | | CSAT | | | | | | |
|-----------|---------|---------|------------|--------|---------|------------|--------|---------|------------|--------|---------|------------|---------|--------|--------|--------|
| | RMS | E_μ | E_σ | RMS | E_μ | E_σ | RMS | E_μ | E_σ | RMS | E_μ | E_σ | | | | |
| 5 | 0,00278 | 0,0192 | 0,0129 | 0,0441 | 0,03215 | 2,9043 | 0,9976 | 4,4421 | 0,00744 | 0,5768 | 0,2688 | 1,1860 | 0,00228 | 0,3093 | 0,1564 | 0,6174 |
| 10 | 0,00226 | 0,0433 | 0,0181 | 0,0776 | 0,03050 | 2,8241 | 0,9652 | 4,3256 | 0,00690 | 0,5544 | 0,2572 | 1,1864 | 0,00335 | 0,3090 | 0,1723 | 0,6580 |
| 15 | 0,00301 | 0,0459 | 0,0159 | 0,0805 | 0,02968 | 2,7521 | 0,9442 | 4,2171 | 0,00723 | 0,5395 | 0,2440 | 1,1271 | 0,00519 | 0,3253 | 0,1812 | 0,7060 |
| 20 | 0,00273 | 0,0510 | 0,0215 | 0,1004 | 0,02801 | 2,6906 | 0,9191 | 4,1194 | 0,00637 | 0,5200 | 0,2282 | 1,0507 | 0,00579 | 0,3450 | 0,1811 | 0,7388 |
| 25 | 0,00327 | 0,0747 | 0,0293 | 0,1252 | 0,02819 | 2,6159 | 0,8907 | 4,0209 | 0,00623 | 0,5250 | 0,2372 | 1,0389 | 0,00598 | 0,3916 | 0,1891 | 0,8043 |
| 30 | 0,00288 | 0,0616 | 0,0224 | 0,1049 | 0,02539 | 2,5491 | 0,8642 | 3,9196 | 0,00752 | 0,5372 | 0,2436 | 1,0198 | 0,00773 | 0,4101 | 0,1978 | 0,8454 |
| 35 | 0,00288 | 0,1002 | 0,0463 | 0,1804 | 0,02627 | 2,4840 | 0,8427 | 3,8307 | 0,00663 | 0,5118 | 0,2457 | 0,9381 | 0,00835 | 0,4676 | 0,2006 | 0,8809 |
| 40 | 0,00235 | 0,0805 | 0,0277 | 0,1130 | 0,02576 | 2,4242 | 0,8185 | 3,7322 | 0,00775 | 0,5381 | 0,2699 | 1,0201 | 0,00940 | 0,4822 | 0,2139 | 0,9352 |
| 45 | 0,00361 | 0,0861 | 0,0370 | 0,1341 | 0,02365 | 2,3672 | 0,8000 | 3,6409 | 0,00745 | 0,5526 | 0,2838 | 1,0782 | 0,01073 | 0,5216 | 0,2271 | 0,9797 |
| 50 | 0,00226 | 0,1257 | 0,0528 | 0,2391 | 0,02488 | 2,3043 | 0,7786 | 3,5669 | 0,00782 | 0,5506 | 0,2844 | 1,0873 | 0,01125 | 0,5788 | 0,2418 | 1,0307 |
| 55 | 0,00281 | 0,1213 | 0,0370 | 0,1753 | 0,02297 | 2,2520 | 0,7615 | 3,4788 | 0,00915 | 0,5650 | 0,2942 | 1,1346 | 0,01265 | 0,6025 | 0,2466 | 1,0574 |
| 60 | 0,00340 | 0,1302 | 0,0542 | 0,2149 | 0,02159 | 2,1999 | 0,7425 | 3,4104 | 0,01066 | 0,5962 | 0,2981 | 1,1774 | 0,01479 | 0,6420 | 0,2636 | 1,1011 |
| 65 | 0,00350 | 0,1256 | 0,0407 | 0,1867 | 0,02208 | 2,1440 | 0,7243 | 3,3153 | 0,01187 | 0,6172 | 0,3061 | 1,2152 | 0,01600 | 0,6646 | 0,2657 | 1,1369 |
| 70 | 0,00335 | 0,1417 | 0,0459 | 0,1992 | 0,02025 | 2,0853 | 0,7005 | 3,2298 | 0,01264 | 0,6592 | 0,3147 | 1,2695 | 0,01692 | 0,7118 | 0,2817 | 1,1844 |
| 75 | 0,00430 | 0,1487 | 0,0561 | 0,2353 | 0,02060 | 2,0331 | 0,6825 | 3,1539 | 0,01402 | 0,6779 | 0,3226 | 1,3056 | 0,01863 | 0,7429 | 0,2993 | 1,2436 |
| 80 | 0,00475 | 0,1700 | 0,0656 | 0,2796 | 0,01945 | 1,9826 | 0,6673 | 3,0661 | 0,01554 | 0,7117 | 0,3270 | 1,3328 | 0,02032 | 0,7902 | 0,3161 | 1,3167 |
| 85 | 0,00360 | 0,1624 | 0,0609 | 0,2588 | 0,01920 | 1,9422 | 0,6490 | 3,0126 | 0,01523 | 0,7388 | 0,3331 | 1,3664 | 0,01992 | 0,8139 | 0,3223 | 1,3490 |
| 90 | 0,00414 | 0,1699 | 0,0617 | 0,2699 | 0,01859 | 1,8935 | 0,6330 | 2,9273 | 0,01704 | 0,7742 | 0,3418 | 1,4173 | 0,02183 | 0,8486 | 0,3328 | 1,3942 |
| 95 | 0,00474 | 0,1744 | 0,0687 | 0,2843 | 0,01816 | 1,8443 | 0,6144 | 2,8435 | 0,01768 | 0,7984 | 0,3575 | 1,4451 | 0,02268 | 0,8757 | 0,3494 | 1,4327 |
| 100 | 0,00562 | 0,2388 | 0,0846 | 0,3959 | 0,01807 | 1,8014 | 0,5967 | 2,8164 | 0,01741 | 0,8384 | 0,3605 | 1,4924 | 0,02199 | 0,9203 | 0,3595 | 1,4952 |
| 105 | 0,00425 | 0,1885 | 0,0710 | 0,3021 | 0,01608 | 1,7575 | 0,5867 | 2,7475 | 0,01965 | 0,8662 | 0,3674 | 1,5372 | 0,02407 | 0,9455 | 0,3689 | 1,5254 |
| 110 | 0,00449 | 0,1897 | 0,0835 | 0,3231 | 0,01591 | 1,7119 | 0,5612 | 2,6474 | 0,02034 | 0,8890 | 0,3882 | 1,5874 | 0,03402 | 0,9359 | 0,5156 | 2,3829 |
| 115 | 0,00420 | 0,2045 | 0,0740 | 0,3296 | 0,01673 | 1,6692 | 0,5555 | 2,5976 | 0,02184 | 0,9317 | 0,3875 | 1,5949 | 0,03495 | 0,9356 | 0,5224 | 2,4010 |
| 120 | 0,00514 | 0,2114 | 0,0812 | 0,3419 | 0,01493 | 1,6280 | 0,5408 | 2,5270 | 0,02320 | 0,9601 | 0,3974 | 1,6335 | 0,03578 | 0,9633 | 0,5293 | 2,4557 |
| 125 | 0,00483 | 0,2118 | 0,0789 | 0,3121 | 0,01455 | 1,5891 | 0,5265 | 2,4762 | 0,03112 | 1,3946 | 0,7618 | 3,6544 | 0,03602 | 0,9183 | 0,4840 | 2,2804 |
| 130 | 0,00455 | 0,2606 | 0,1034 | 0,4511 | 0,01565 | 1,5440 | 0,5058 | 2,3959 | 0,03130 | 1,4073 | 0,7752 | 3,6806 | 0,03646 | 0,9679 | 0,5349 | 2,4630 |
| 135 | 0,00591 | 0,2878 | 0,1102 | 0,4893 | 0,01451 | 1,5069 | 0,4975 | 2,3422 | 0,03084 | 1,4383 | 0,7962 | 3,7830 | 0,03609 | 1,0059 | 0,5625 | 2,5907 |
| 140 | 0,00529 | 0,3181 | 0,1184 | 0,5072 | 0,01504 | 1,4718 | 0,4806 | 2,2818 | 0,03118 | 1,3944 | 0,7675 | 3,6665 | 0,03614 | 0,9853 | 0,5444 | 2,5306 |

Fonte: Elaborada pelo próprio autor.

Tabela 2 – Resultado do cenário sintético 01 com objeto de validação fixo à 1000 mm da câmera

| Distância | CM | | | CSMF | | | CSAT-M | | | CSAT | | | | | | |
|-----------|---------|---------|------------|--------|---------|------------|--------|---------|------------|--------|---------|------------|-----------|--------|--------|--------|
| | RMS | E_μ | E_σ | RMS | E_μ | E_σ | RMS | E_μ | E_σ | RMS | E_μ | E_σ | E_{max} | | | |
| 5 | 0,00349 | 0,0973 | 0,0408 | 0,1916 | 0,03409 | 5,6405 | 2,5953 | 10,7851 | 0,00886 | 0,5498 | 0,2050 | 0,9526 | 0,00349 | 0,3378 | 0,1560 | 0,6476 |
| 10 | 0,00343 | 0,0729 | 0,0258 | 0,1433 | 0,01622 | 5,0146 | 1,8365 | 7,5398 | 0,01189 | 0,3583 | 0,1304 | 0,6897 | 0,00400 | 0,3569 | 0,1702 | 0,6863 |
| 15 | 0,00370 | 0,0882 | 0,0390 | 0,1478 | 0,03365 | 6,1902 | 2,8258 | 11,8108 | 0,00966 | 0,3800 | 0,1398 | 0,6783 | 0,00259 | 0,3297 | 0,1574 | 0,6373 |
| 20 | 0,00205 | 0,0745 | 0,0294 | 0,1372 | 0,03328 | 5,7724 | 2,6965 | 10,9565 | 0,01143 | 0,3485 | 0,1120 | 0,6130 | 0,00450 | 0,3904 | 0,1850 | 0,7406 |
| 25 | 0,00346 | 0,1002 | 0,0303 | 0,1831 | 0,03130 | 5,1783 | 2,3528 | 9,9645 | 0,01184 | 0,3252 | 0,0968 | 0,4657 | 0,00473 | 0,4076 | 0,1871 | 0,7290 |
| 30 | 0,00551 | 0,1334 | 0,0452 | 0,1945 | 0,01438 | 4,6027 | 1,6672 | 6,9212 | 0,01169 | 0,4173 | 0,1295 | 0,6591 | 0,00358 | 0,4730 | 0,2078 | 0,8419 |
| 35 | 0,00484 | 0,1650 | 0,0550 | 0,2570 | 0,01393 | 4,5189 | 1,6434 | 6,8195 | 0,01206 | 0,3615 | 0,0994 | 0,5507 | 0,00581 | 0,4094 | 0,1815 | 0,7236 |
| 40 | 0,00511 | 0,1873 | 0,0548 | 0,2683 | 0,01575 | 4,4474 | 1,6165 | 6,7133 | 0,01187 | 0,4093 | 0,1285 | 0,6848 | 0,00564 | 0,4148 | 0,1823 | 0,7618 |
| 45 | 0,00556 | 0,2099 | 0,0698 | 0,2922 | 0,01403 | 4,3488 | 1,5660 | 6,5629 | 0,01235 | 0,3835 | 0,1149 | 0,6126 | 0,00594 | 0,4137 | 0,1845 | 0,7801 |
| 50 | 0,00681 | 0,2186 | 0,0671 | 0,3191 | 0,01346 | 4,2752 | 1,5453 | 6,4797 | 0,01382 | 0,3972 | 0,1173 | 0,6366 | 0,00716 | 0,4720 | 0,1988 | 0,8258 |
| 55 | 0,00378 | 0,2388 | 0,0758 | 0,3632 | 0,01301 | 4,1874 | 1,5162 | 6,3437 | 0,01472 | 0,3685 | 0,1182 | 0,5582 | 0,00947 | 0,3532 | 0,1484 | 0,6888 |
| 60 | 0,00847 | 0,2598 | 0,0941 | 0,3932 | 0,01204 | 4,1520 | 1,5011 | 6,2784 | 0,01412 | 0,4722 | 0,1477 | 0,7807 | 0,00787 | 0,5724 | 0,2343 | 0,9407 |
| 65 | 0,00628 | 0,2925 | 0,0921 | 0,4132 | 0,01105 | 4,0305 | 1,4424 | 6,0898 | 0,01426 | 0,4040 | 0,1020 | 0,5921 | 0,00867 | 0,4858 | 0,1958 | 0,8852 |
| 70 | 0,00758 | 0,3108 | 0,1151 | 0,5033 | 0,01195 | 3,9845 | 1,4290 | 6,0613 | 0,01494 | 0,4245 | 0,1134 | 0,6718 | 0,00816 | 0,5735 | 0,2347 | 0,9681 |
| 75 | 0,00863 | 0,3870 | 0,1057 | 0,5276 | 0,01149 | 3,8860 | 1,3824 | 5,9323 | 0,01410 | 0,4123 | 0,1153 | 0,6556 | 0,00818 | 0,5050 | 0,1965 | 0,8797 |
| 80 | 0,00850 | 0,4580 | 0,1514 | 0,7206 | 0,01143 | 3,8625 | 1,3893 | 5,8859 | 0,01330 | 0,4811 | 0,1302 | 0,7650 | 0,00664 | 0,5355 | 0,2072 | 0,9190 |
| 85 | 0,00760 | 0,4282 | 0,1613 | 0,6943 | 0,01134 | 3,7984 | 1,3662 | 5,7943 | 0,01553 | 0,4459 | 0,1182 | 0,7093 | 0,00931 | 0,4869 | 0,1965 | 0,8715 |
| 90 | 0,00851 | 0,4941 | 0,1777 | 0,8344 | 0,00998 | 3,7365 | 1,3372 | 5,7160 | 0,01531 | 0,4464 | 0,1054 | 0,6929 | 0,00896 | 0,5348 | 0,1968 | 0,9105 |
| 95 | 0,00859 | 0,5797 | 0,2290 | 1,0420 | 0,01143 | 3,6502 | 1,3125 | 5,5744 | 0,01428 | 0,4573 | 0,1272 | 0,7337 | 0,00817 | 0,4887 | 0,1866 | 0,8636 |
| 100 | 0,00916 | 1,1524 | 0,4035 | 1,7698 | 0,01222 | 3,5913 | 1,2877 | 5,4867 | 0,01617 | 0,4140 | 0,1306 | 0,6855 | 0,01041 | 0,4475 | 0,1639 | 0,7941 |
| 105 | 0,00711 | 0,6068 | 0,2442 | 1,0447 | 0,01212 | 3,5364 | 1,2704 | 5,3990 | 0,01590 | 0,4355 | 0,1181 | 0,6842 | 0,01021 | 0,4463 | 0,1638 | 0,8028 |
| 110 | 0,00657 | 0,6831 | 0,2902 | 1,2574 | 0,01198 | 3,4616 | 1,2275 | 5,2735 | 0,01409 | 0,4158 | 0,1258 | 0,6432 | 0,00883 | 0,4302 | 0,1554 | 0,7898 |
| 115 | 0,00907 | 0,6626 | 0,2818 | 1,2029 | 0,01076 | 3,4380 | 1,2310 | 5,2565 | 0,01709 | 0,5082 | 0,1320 | 0,8026 | 0,01092 | 0,5998 | 0,2188 | 0,9834 |
| 120 | 0,00858 | 0,6342 | 0,2628 | 1,1412 | 0,01115 | 3,3143 | 1,1642 | 5,0675 | 0,01633 | 0,4182 | 0,1609 | 0,6770 | 0,01163 | 0,4498 | 0,1511 | 0,8244 |
| 125 | 0,00862 | 0,8236 | 0,3824 | 1,6722 | 0,00855 | 3,2996 | 1,1650 | 5,0825 | 0,01645 | 0,4472 | 0,1263 | 0,7551 | 0,01069 | 0,5420 | 0,2020 | 0,9512 |

Fonte: Elaborada pelo próprio autor.

Tabela 3 – Resultado do cenário sintético 02 com objeto de validação na borda da imagem

| Distância | CM | | | CSMF | | | CSAT-M | | | CSAT | | | | | | |
|-----------|---------|---------|------------|-----------|---------|---------|------------|-----------|---------|---------|------------|-----------|---------|---------|------------|-----------|
| | RMS | E_μ | E_σ | E_{max} |
| 150 | 0,04048 | 0,4182 | 0,2056 | 0,8592 | 0,05381 | 1,3083 | 0,7118 | 3,1838 | 0,04324 | 0,5692 | 0,3572 | 1,6505 | 0,05019 | 0,5271 | 0,2946 | 1,3897 |
| 200 | 0,01818 | 0,3120 | 0,1205 | 0,5708 | 0,05803 | 1,4422 | 0,5318 | 2,3178 | 0,02091 | 0,5413 | 0,2944 | 1,2126 | 0,02483 | 0,4451 | 0,2397 | 1,0046 |
| 300 | 0,01047 | 0,2776 | 0,1255 | 0,5367 | 0,05758 | 2,2407 | 0,7168 | 3,4993 | 0,01759 | 0,6978 | 0,3403 | 1,4232 | 0,02276 | 0,5920 | 0,2800 | 1,1757 |
| 400 | 0,00321 | 0,1320 | 0,0507 | 0,2443 | 0,03806 | 2,2827 | 0,7551 | 3,5427 | 0,00931 | 0,5462 | 0,2675 | 1,0514 | 0,00992 | 0,4312 | 0,2115 | 0,9126 |
| 500 | 0,00288 | 0,0616 | 0,0224 | 0,1049 | 0,02539 | 2,5491 | 0,8642 | 3,9196 | 0,00752 | 0,5372 | 0,2436 | 1,0198 | 0,00773 | 0,4101 | 0,1978 | 0,8454 |
| 600 | 0,00269 | 0,0402 | 0,0254 | 0,0918 | 0,01974 | 2,8118 | 0,9672 | 4,2897 | 0,00743 | 0,5994 | 0,2396 | 1,0652 | 0,00657 | 0,4231 | 0,1945 | 0,8132 |
| 700 | 0,00257 | 0,0531 | 0,0346 | 0,1277 | 0,01844 | 3,3986 | 1,1952 | 5,1836 | 0,00544 | 0,5653 | 0,2850 | 1,1552 | 0,00560 | 0,4681 | 0,2202 | 0,8709 |
| 800 | 0,00245 | 0,0687 | 0,0444 | 0,1771 | 0,01737 | 3,7404 | 1,3331 | 5,6693 | 0,00568 | 0,6566 | 0,2613 | 1,1295 | 0,00495 | 0,4964 | 0,2189 | 0,9097 |
| 900 | 0,00239 | 0,0837 | 0,0401 | 0,1616 | 0,01479 | 4,0369 | 1,4492 | 6,0902 | 0,01222 | 0,8733 | 0,3396 | 1,3944 | 0,00578 | 0,3802 | 0,1902 | 0,7836 |
| 1000 | 0,00244 | 0,1238 | 0,0563 | 0,2285 | 0,01075 | 4,0005 | 1,4388 | 6,0320 | 0,00862 | 0,8655 | 0,3249 | 1,3447 | 0,00465 | 0,4633 | 0,1956 | 0,8071 |
| 1200 | 0,00369 | 0,1776 | 0,0773 | 0,3415 | 0,05427 | 3,8250 | 1,4307 | 5,8908 | 0,00887 | 0,4544 | 0,1743 | 0,7117 | 0,00436 | 0,4597 | 0,2023 | 0,7703 |

Fonte: Elaborada pelo próprio autor.

Tabela 4 – Resultado do cenário sintético 02 com objeto de validação no centro da imagem

| Distância | CM | | | CSMF | | | CSAT-M | | | CSAT | | | | | | |
|-----------|---------|---------|------------|-----------|---------|---------|------------|-----------|---------|---------|------------|-----------|---------|---------|------------|-----------|
| | RMS | E_μ | E_σ | E_{max} |
| 150 | 0,01827 | 0,1766 | 0,1011 | 0,4503 | 0,01595 | 0,6271 | 0,3661 | 1,4658 | 0,01996 | 0,3434 | 0,2814 | 1,0194 | 0,01879 | 0,2852 | 0,1976 | 0,7877 |
| 200 | 0,00503 | 0,0654 | 0,0386 | 0,1472 | 0,00808 | 0,3864 | 0,2233 | 0,8877 | 0,01039 | 0,2277 | 0,1671 | 0,6743 | 0,00916 | 0,1532 | 0,1012 | 0,3946 |
| 300 | 0,00319 | 0,0200 | 0,0088 | 0,0354 | 0,00311 | 0,1872 | 0,1083 | 0,4362 | 0,00404 | 0,1247 | 0,0708 | 0,2998 | 0,00418 | 0,0841 | 0,0493 | 0,2103 |
| 400 | 0,00298 | 0,0223 | 0,0110 | 0,0459 | 0,00289 | 0,1077 | 0,0605 | 0,2407 | 0,00349 | 0,0896 | 0,0447 | 0,1822 | 0,00316 | 0,0650 | 0,0376 | 0,1658 |
| 500 | 0,00294 | 0,0226 | 0,0094 | 0,0367 | 0,00309 | 0,0759 | 0,0441 | 0,1867 | 0,00244 | 0,0679 | 0,0250 | 0,1120 | 0,00291 | 0,0488 | 0,0256 | 0,1030 |
| 600 | 0,00353 | 0,0331 | 0,0181 | 0,0706 | 0,00364 | 0,0695 | 0,0365 | 0,1615 | 0,00326 | 0,0719 | 0,0363 | 0,1218 | 0,00388 | 0,0544 | 0,0320 | 0,1145 |
| 700 | 0,00376 | 0,0453 | 0,0209 | 0,0736 | 0,00383 | 0,0567 | 0,0240 | 0,1027 | 0,00380 | 0,0707 | 0,0291 | 0,1405 | 0,00378 | 0,0546 | 0,0217 | 0,0928 |
| 800 | 0,00262 | 0,0526 | 0,0244 | 0,0967 | 0,00263 | 0,0625 | 0,0252 | 0,0996 | 0,00217 | 0,0909 | 0,0306 | 0,1679 | 0,00255 | 0,0639 | 0,0253 | 0,1029 |
| 900 | 0,00410 | 0,0988 | 0,0326 | 0,1435 | 0,00412 | 0,0885 | 0,0299 | 0,1420 | 0,00431 | 0,0967 | 0,0363 | 0,1641 | 0,00396 | 0,0836 | 0,0308 | 0,1469 |
| 1000 | 0,00361 | 0,0903 | 0,0361 | 0,1845 | 0,00361 | 0,0965 | 0,0388 | 0,1883 | 0,00371 | 0,1092 | 0,0464 | 0,1873 | 0,00351 | 0,0981 | 0,0358 | 0,1811 |

Fonte: Elaborada pelo próprio autor.

Tabela 5 – Resultado do cenário sintético 03 com objeto de validação na borda da imagem

| Distorção | CM | | | CSMF | | | CSAT-M | | | CSAT | | |
|-----------|---------|---------|------------|-----------|---------|---------|------------|-----------|---------|---------|------------|-----------|
| | RMS | E_μ | E_σ | E_{max} | RMS | E_μ | E_σ | E_{max} | RMS | E_μ | E_σ | E_{max} |
| 0,04 | 0,00548 | 0,1186 | 0,0487 | 0,1979 | 0,03733 | 3,2841 | 1,1071 | 5,0682 | 0,02368 | 0,2965 | 0,0961 | 0,4701 |
| 0,08 | 0,00728 | 0,1317 | 0,0481 | 0,2028 | 0,02928 | 2,5883 | 0,8447 | 4,1287 | 0,02610 | 0,2924 | 0,1027 | 0,4895 |
| 0,12 | 0,00730 | 0,1513 | 0,0566 | 0,2249 | 0,03502 | 1,5092 | 0,5194 | 2,6232 | 0,02989 | 0,3365 | 0,1138 | 0,5035 |
| 0,16 | 0,01917 | 0,3452 | 0,1306 | 0,6140 | 0,13477 | 1,8774 | 1,0573 | 4,2636 | 0,03574 | 0,8354 | 0,3469 | 1,5161 |
| 0,20 | 0,07491 | 0,8831 | 0,3997 | 1,8739 | 0,14010 | 2,5827 | 1,5288 | 6,4523 | 0,09818 | 0,8527 | 0,5496 | 1,9325 |
| 0,24 | 0,01735 | 0,3143 | 0,1404 | 0,5716 | 0,13611 | 2,2979 | 1,2837 | 5,3552 | 0,03261 | 1,1830 | 0,5033 | 2,1989 |
| 0,28 | 0,12527 | 2,1397 | 1,6128 | 6,1231 | 0,14036 | 2,8310 | 1,3934 | 6,4635 | 0,12895 | 2,2432 | 1,6758 | 6,4452 |
| 0,32 | 0,11663 | 1,8442 | 1,1919 | 4,2165 | 0,13381 | 2,7246 | 1,5870 | 6,7816 | 0,12899 | 2,4623 | 1,7482 | 6,3944 |

Fonte: Elaborada pelo próprio autor.

Tabela 6 – Resultado do cenário sintético 03 com objeto de validação com posição fixa

| Distorção | CM | | | CSMF | | | CSAT-M | | | CSAT | | |
|-----------|---------|---------|------------|-----------|---------|---------|------------|-----------|---------|---------|------------|-----------|
| | RMS | E_μ | E_σ | E_{max} | RMS | E_μ | E_σ | E_{max} | RMS | E_μ | E_σ | E_{max} |
| 0,00 | 0,00277 | 0,0581 | 0,0218 | 0,1061 | 0,01882 | 2,1266 | 0,7085 | 3,2180 | 0,00462 | 0,4753 | 0,2008 | 0,8275 |
| 0,04 | 0,00318 | 0,0587 | 0,0225 | 0,1068 | 0,01803 | 1,9186 | 0,6294 | 2,8919 | 0,00559 | 0,2426 | 0,0933 | 0,4338 |
| 0,08 | 0,00167 | 0,0667 | 0,0205 | 0,1031 | 0,01709 | 1,6246 | 0,5271 | 2,4523 | 0,00454 | 0,2489 | 0,0871 | 0,4206 |
| 0,12 | 0,00313 | 0,0592 | 0,0185 | 0,0818 | 0,01257 | 1,1518 | 0,3638 | 1,7731 | 0,00655 | 0,2007 | 0,0773 | 0,3383 |
| 0,16 | 0,00428 | 0,1441 | 0,0429 | 0,2084 | 0,01831 | 0,8465 | 0,3881 | 1,7592 | 0,00968 | 0,2824 | 0,1057 | 0,4640 |
| 0,20 | 0,00405 | 0,1795 | 0,0671 | 0,3102 | 0,11861 | 1,5374 | 0,6743 | 2,5904 | 0,03582 | 0,6213 | 0,2543 | 1,0966 |
| 0,24 | 0,01102 | 0,1625 | 0,0460 | 0,2424 | 0,12727 | 2,1843 | 1,1515 | 4,7395 | 0,05738 | 0,8413 | 0,3512 | 1,4440 |
| 0,28 | 0,05151 | 0,8308 | 0,3751 | 1,6654 | 0,13093 | 2,4913 | 1,4387 | 6,0172 | 0,03513 | 1,3461 | 0,5846 | 2,4911 |
| 0,32 | 0,11663 | 1,8442 | 1,1919 | 4,2165 | 0,13381 | 2,7246 | 1,5870 | 6,7816 | 0,12899 | 2,4623 | 1,7482 | 6,3944 |

Fonte: Elaborada pelo próprio autor.

APÊNDICE B – RESULTADOS DOS CENÁRIOS REAIS

Este apêndice apresenta os resultados complementares obtidos nos experimentos realizados no tanque de água (tabelas 7 e 8) e no ambiente submarino (tabela 9). A primeira coluna de cada tabela denomina o vídeo utilizado no experimento e a segunda coluna descreve o conjunto de validações aplicados em cada vídeo. Para os experimentos no tanque de água um conjunto é composto por quatro imagens de uma câmera em posições diferentes, enquanto na análise do ambiente submarino um conjunto é composto por duas imagens (uma para cada câmera do sistema).

As demais colunas apresentam os resultados obtidos pela técnica de adaptação de calibração proposta por Lavest et al. (CSMF) e as propostas nessa dissertação (CSAT e CSAT-M), onde o experimento do tanque de água indica os erros de reconstrução do objeto 3D e o submarino retrata o erro epipolar alcançado pelas técnicas. O erro médio (E_μ), o desvio padrão do erro (E_σ) e erro máximo (E_{max}) são demonstrado em milímetros e em *pixels* para o ambiente do tanque de água e submarino, respectivamente.

Tabela 7 – Resultado do experimento no tanque de água com o primeiro conjunto de calibração.

| Vídeo | # Reconstruções | CM | | | CSMF | | | CSAT-M | | | CSAT | | |
|--------------|-----------------|---------|------------|-----------|---------|------------|-----------|---------|------------|-----------|---------|------------|-----------|
| | | E_μ | E_σ | E_{max} |
| Centro 95cm | 544 | 0,12549 | 0,03711 | 0,24952 | 0,13029 | 0,04002 | 0,25661 | 0,13466 | 0,04094 | 0,26248 | 0,13508 | 0,03883 | 0,26192 |
| Centro 130cm | 178 | 0,18403 | 0,05710 | 0,39653 | 0,18527 | 0,06359 | 0,39787 | 0,18602 | 0,06396 | 0,39950 | 0,19610 | 0,04825 | 0,40911 |
| Borda 95cm | 421 | 0,15009 | 0,05740 | 0,26166 | 0,38902 | 0,15145 | 0,64143 | 0,18445 | 0,05156 | 0,31182 | 0,18285 | 0,05118 | 0,30917 |
| Borda 130cm | 137 | 0,20382 | 0,07128 | 0,34581 | 0,40340 | 0,12969 | 0,76434 | 0,21709 | 0,05215 | 0,37181 | 0,21489 | 0,05061 | 0,36926 |

Tabela 8 – Resultado do experimento no tanque de água com o segundo conjunto de calibração.

| Vídeo | # Reconstruções | CM | | | CSMF | | | CSAT-M | | | CSAT | | |
|--------------|-----------------|---------|------------|-----------|---------|------------|-----------|---------|------------|-----------|---------|------------|-----------|
| | | E_μ | E_σ | E_{max} |
| Centro 95cm | 544 | 0,12622 | 0,03810 | 0,25123 | 0,13116 | 0,04036 | 0,25920 | 0,13548 | 0,04194 | 0,26445 | 0,13065 | 0,03925 | 0,25673 |
| Centro 130cm | 178 | 0,18742 | 0,05607 | 0,40792 | 0,18748 | 0,06342 | 0,40779 | 0,18613 | 0,05351 | 0,40565 | 0,18635 | 0,05070 | 0,40542 |
| Borda 95cm | 421 | 0,14682 | 0,05370 | 0,25756 | 0,39146 | 0,15188 | 0,64373 | 0,18025 | 0,05163 | 0,30547 | 0,17904 | 0,05150 | 0,30391 |
| Borda 130cm | 137 | 0,21037 | 0,07834 | 0,34973 | 0,48291 | 0,16744 | 0,76029 | 0,21065 | 0,05068 | 0,36806 | 0,21173 | 0,05121 | 0,37005 |

Tabela 9 – Resultados obtidos nos experimentos realizados no ambiente submarino.

| Vídeo # | Quadros | CSMF | | | CSAT-M | | | CSAT | | | | | |
|---------|---------|---------|-----------|-----------|------------|---------|-----------|-----------|------------|---------|-----------|-----------|------------|
| | | E_μ | E_{min} | E_{max} | E_σ | E_μ | E_{min} | E_{max} | E_σ | E_μ | E_{min} | E_{max} | E_σ |
| 01 | 4755 | 1,2179 | 0,7856 | 1,6990 | 0,1250 | 0,6891 | 0,2668 | 1,2715 | 0,1635 | 0,6342 | 0,2987 | 1,1683 | 0,1391 |
| 02 | 23172 | 1,1736 | 0,5583 | 1,6942 | 0,1178 | 0,7649 | 0,1693 | 1,2537 | 0,1339 | 0,7019 | 0,1708 | 1,2429 | 0,1210 |
| 03 | 18063 | 2,1101 | 1,0602 | 3,2036 | 0,2820 | 2,1591 | 1,2161 | 3,2340 | 0,2770 | 2,1226 | 1,1578 | 3,2410 | 0,2797 |
| 04 | 1114 | 2,4159 | 1,7588 | 3,1528 | 0,2386 | 1,8680 | 1,1830 | 2,6301 | 0,2577 | 1,7972 | 1,1383 | 2,7018 | 0,2652 |