



UNIVERSIDADE FEDERAL DE SANTA CATARINA
CENTRO TECNOLÓGICO
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA MECÂNICA

Alice Weza Fava Bilbáo

**SISTEMA DE VISÃO ESTÉREO ATIVA COM PROJEÇÃO *LASER*
PSEUDOALEATÓRIA PARA RECONSTRUÇÃO 3D**

Florianópolis
2026

Alice Weza Fava Bilbáo

**SISTEMA DE VISÃO ESTÉREO ATIVA COM PROJEÇÃO *LASER*
PSEUDOALEATÓRIA PARA RECONSTRUÇÃO 3D**

Dissertação submetida ao Programa de Pós-Graduação em Engenharia Mecânica da Universidade Federal de Santa Catarina como requisito parcial para a obtenção do título de Mestra em Engenharia Mecânica.

Orientador(a): Prof. Tiago Loureiro Figaro da Costa Pinto, Dr. Eng.

Florianópolis

2026

Bilbáo, Alice Weza Fava

Sistema de visão estéreo ativa com projeção laser pseudoaleatória para reconstrução 3D / Alice Weza Fava Bilbáo ; orientador, Tiago Loureiro Figaro da Costa Pinto, 2026.

100 p.

Dissertação (mestrado) - Universidade Federal de Santa Catarina, Centro Tecnológico, Programa de Pós-Graduação em Engenharia Mecânica, Florianópolis, 2026.

Inclui referências.

1. Engenharia Mecânica. 2. Visão Computacional. 3. Reconstrução 3D. 4. Visão Estéreo. 5. Luz Estruturada. I. Pinto, Tiago Loureiro Figaro da Costa. II. Universidade Federal de Santa Catarina. Programa de Pós-Graduação em Engenharia Mecânica. III. Título.

Alice Weza Fava Bilbáo

**SISTEMA DE VISÃO ESTÉREO ATIVA COM PROJEÇÃO *LASER*
PSEUDOALEATÓRIA PARA RECONSTRUÇÃO 3D**

O presente trabalho em nível de Mestrado foi avaliado e aprovado, em 18/12/2025 pela banca examinadora composta pelos seguintes membros:

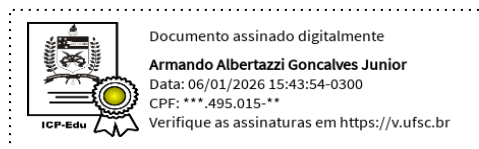
Prof. Armando Albertazzi Gonçalves Júnior, Dr. Eng.

Universidade Federal de Santa Catarina

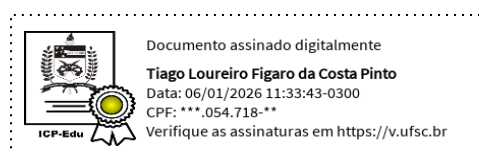
Prof. Maurício Edgar Stivanello, Dr. Eng.

Instituto Federal de Santa Catarina

Certificamos que esta é a versão original e final do trabalho de conclusão que foi julgado adequado para obtenção do título de Mestra em Engenharia Mecânica.



Coordenação do Programa de Pós-Graduação



Prof. Tiago Loureiro Figaro da Costa Pinto, Dr. Eng

Orientador

Florianópolis, 2026.

Dedico este trabalho à minha irmã Ana Carolina, minha maior inspiração acadêmica, profissional e pessoal.

AGRADECIMENTOS

Agradeço primeiramente ao meu orientador, Prof. Tiago, por toda a orientação, paciência e confiança depositada neste trabalho. Agradeço à CAPES, pelo apoio financeiro que possibilitou a realização deste mestrado.

Agradeço a todos os colegas de LABMETRO, em especial aos membros do Projeto VORIS, pelo suporte durante toda a execução do meu mestrado e desta dissertação.

Agradeço à minha família, minha maior motivação. Meus pais, Rosangela e Henrique, que sempre me mostraram a importância da educação. Meus irmãos, Ana Carolina e Henrique Marcos, que sempre foram meus exemplos. Minhas afilhadas, Ana Maria e Carolina, que me fazem querer ser motivo de orgulho. Meu namorado, Samuel, que foi a minha base nos momentos de dificuldade. Eu não seria nada sem vocês.

Agradeço por fim à Dynamox SA, pela flexibilidade e suporte, fazendo com que fosse possível conciliar as atividades do trabalho do dia a dia com a finalização deste mestrado.

RESUMO

O presente trabalho detalha o desenvolvimento e a validação metrológica de um sistema de visão estéreo ativa para reconstrução 3D, visando uma arquitetura aberta e flexível, projetada para atender aos requisitos de inspeção visual em ambientes subaquáticos. O objetivo geral consiste em desenvolver e avaliar um sistema baseado na projeção *laser* de um padrão de pontos pseudoaleatórios. A arquitetura de *software* proposta fundamenta-se no Método da Triangulação Inversa, mas implementa uma métrica de correspondência baseada na correlação espaço-temporal ($N \times N \times T$). Esta abordagem processa a variação do padrão *laser*, controlada por prismas contrarrotativos e capturada em uma sequência de imagens, para estabelecer a correspondência 3D com alta robustez. O *software* foi implementado em linguagem Python com a biblioteca PyTorch para processamento paralelo em GPU. Uma otimização sistemática dos parâmetros de execução (número de pares de imagens, tamanho do *kernel* de correlação e sensibilidade do *grid* adaptativo) foi conduzida para maximizar o desempenho. Os resultados demonstram que o sistema otimizado atinge um desvio padrão de 0,07 mm na medição de um plano de referência. A validação metrológica, realizada com esferas padrão calibradas, permitiu a caracterização dos erros de medição do sistema. Adicionalmente, a robustez do método de reconstrução foi avaliada qualitativamente em objetos de forma livre, demonstrando a capacidade do sistema em resolver geometrias complexas e gerar nuvens de pontos densas e íntegras. A análise identificou um erro sistemático de -0,345 mm na medição de distância entre centros. A principal conclusão do trabalho é de que o gargalo de desempenho metrológico do sistema não reside no ruído aleatório do algoritmo de correlação (desvio padrão médio de 0,15 mm), mas sim no erro sistemático de escala proveniente da calibração estéreo.

Palavras-chave: Visão Computacional; Reconstrução 3D; Visão Estéreo; Metrologia; Luz Estruturada.

ABSTRACT

This work details the development and metrological validation of an active stereo vision system for 3D reconstruction, aiming for an open and flexible architecture designed to meet visual inspection requirements in underwater environments. The general objective is to develop and evaluate a system based on the projection of a pseudo-random laser dot pattern. The proposed software architecture is based on the Inverse Triangulation Method but implements a matching metric based on spatiotemporal correlation ($N \times N \times T$). This approach processes the laser pattern variation, controlled by counter-rotating prisms and captured in an image sequence, to establish the 3D correspondence with high robustness. The software was implemented in Python using the PyTorch library for parallel processing on a GPU. A systematic optimization of the execution parameters (number of image pairs, correlation kernel size and adaptive grid sensitivity) was conducted to maximize performance. The results demonstrate that the optimized system achieves a standard deviation of 0.07 mm when measuring a reference plane. The metrological validation, performed using calibrated standard spheres, allowed for the characterization of the system's measurement errors. Additionally, the robustness of the reconstruction method was qualitatively evaluated on free-form objects, demonstrating the system's capability to resolve complex geometries and generate dense and integral point clouds. The analysis identified a systematic error of -0.345 mm in the distance-between-centers measurement. The main conclusion of this work is that the system's metrological performance bottleneck does not lie in the random noise of the correlation algorithm (mean standard deviation of 0.15 mm), but rather in the systematic scale error stemming from the stereo calibration.

Keywords: Computer Vision; 3D Reconstruction; Stereo Vision; Metrology; Structured Light.

LISTA DE FIGURAS

Figura 1 - Esquema da câmera pinhole.....	23
Figura 2 – Modelo matemático da câmera pinhole.....	23
Figura 3 – Sistema de visão estéreo.....	28
Figura 4 – Padrão de franjas senoidais.....	30
Figura 5 – Padrão de banda limitada.....	31
Figura 6 – Padrão de pontos pseudoaleatórios.....	32
Figura 7 – Padrões de calibração.....	33
Figura 8 – Correlação espaço-temporal.....	43
Figura 9 – Diagrama de blocos do sistema de aquisição.....	45
Figura 10 – Diagrama de blocos do sistema de processamento.....	45
Figura 11 – Fluxograma do processo de reconstrução 3D.....	47
Figura 12 – Sistema de projeção.....	50
Figura 13 – Modelo 3D da estrutura de fixação.....	51
Figura 14 – Montagem final.....	52
Figura 15 – Placa de interface eletrônica.....	52
Figura 16 – Exemplo de par de imagens do padrão de calibração.....	63
Figura 17 – Exemplo de aquisições do plano.....	66
Figura 18 – Exemplo de aquisições obtidas do plano.....	66
Figura 19 – Exemplo de aquisições obtidas das esferas.....	67
Figura 20 – Reconstrução ótima de $T = 10$ no plano.....	69
Figura 21 – PNR de $T = 10$ no plano.....	69
Figura 22 – Reconstrução ótima de $T = 10$ nas esferas.....	71
Figura 23 – PNR de $T = 10$ nas esferas.....	72
Figura 24 – Perfis de correlação reais (ZNCC x Profundidade Z).....	74
Figura 25 – Reconstrução melhor caso plano.....	78
Figura 26 – Reconstrução do plano inclinado.....	81
Figura 27 – Reconstrução melhor caso esferas.....	83
Figura 28 – PNR melhor caso esferas.....	83
Figura 29 – Exemplo de aquisições da calota.....	84
Figura 30 – Reconstrução calota.....	85
Figura 31 – PNR calota.....	86
Figura 32 - Objeto 1: réplica de crânio humano.....	87

Figura 33 - Reconstrução da réplica de crânio humano.	88
Figura 34 – Objeto 2: cabeça de manequim.	88
Figura 35 - Reconstrução da cabeça de manequim.	89

LISTA DE TABELAS

Tabela 1 - Parâmetros intrínsecos do sistema estéreo.....	64
Tabela 2 - Parâmetros extrínsecos do par estéreo (câmera direita em relação à esquerda).	64
Tabela 3 - Resultados da otimização do número de imagens (T) no plano..	67
Tabela 4 – Resultados da otimização do número de imagens (T) nas esferas.	70
Tabela 5 – Resultados da otimização do kernel no plano.	72
Tabela 6 - Resultados da otimização do kernel nas esferas.	75
Tabela 7 – Resultados da otimização da sensibilidade.	76
Tabela 8 – Parâmetros de execução do software de reconstrução.	78
Tabela 9 – Resultados da análise da distância de trabalho.....	79
Tabela 10 – Resultados comparação dos padrões de luz.	82
Tabela 11 – Resultados medição das esferas.	84

LISTA DE GRÁFICOS

Gráfico 1 - Resultados da otimização do número de imagens (T) no plano.	68
Gráfico 2 - Resultados da otimização do número de imagens (T) nas esferas.	70
Gráfico 3 - Resultados da otimização do kernel no plano.....	73
Gráfico 4 - Resultados da otimização do kernel nas esferas.....	76
Gráfico 5 - Resultados da otimização da sensibilidade.	77
Gráfico 6 - Resultados da análise da distância de trabalho.....	79

LISTA DE ABREVIATURAS E SIGLAS

1D Unidimensional

2D Bidimensional

3D Tridimensional

BLP Padrão de Banda Limitada (do inglês, *Band Limited Pattern*)

BRIEF *Binary Robust Independent Elementary features*

CMOS Semicondutor de Óxido Metálico Complementar (do inglês, *Complementary Metal-Oxide-Semiconductor*)

CPU Unidade Central de Processamento (do inglês, *Central Processing Unit*)

DOE Elemento Óptico Difractivo (do inglês, *Diffraction Optical Element*)

FAST *Features from Accelerated Segment Test*

FPS Quadros por Segundo (do inglês, *Frames Per Second*)

GPU Unidade de Processamento Gráfico (do inglês, *Graphics Processing Unit*)

MMC Máquina de Medição por Coordenadas

NCC Correlação Cruzada Normalizada (do inglês, *Normalized Cross-Correlation*)

ORB *Oriented FAST and Rotated BRIEF*

PNR Relação Pico-Ruído (do inglês, *Peak-to-Noise Ratio*)

ROV Veículo Operado Remotamente (do inglês, *Remotely Operated Vehicle*)

RPP *Random Pattern Projector*

SfM Estrutura a partir do Movimento (do inglês, *Structure from Motion*)

SIFT *Scale-Invariant Feature Transform*

SURF *Speeded Up Robust Features*

USB Barramento Serial Universal (do inglês, *Universal Serial Bus*)

VRAM Memória de Vídeo de Acesso Aleatório (do inglês, *Video Random Access Memory*)

ZNCC Correlação Cruzada Normalizada a Zero (do inglês, *Zero-Normalized Cross-Correlation*)

LISTA DE EQUAÇÕES

Equação 1	24
Equação 2	24
Equação 3	24
Equação 4	24
Equação 5	24
Equação 6	25
Equação 7	25
Equação 8	29
Equação 9	29
Equação 10	36
Equação 11	60
Equação 12	62
Equação 13	62

SUMÁRIO

1	INTRODUÇÃO.....	15
1.1	CONTEXTUALIZAÇÃO.....	15
1.2	PROBLEMA DE PESQUISA.....	17
1.3	JUSTIFICATIVA.....	17
1.4	OBJETIVOS.....	18
1.5	ESTRUTURA DA DISSERTAÇÃO.....	19
2	FUNDAMENTAÇÃO TEÓRICA.....	21
2.1	MODELO DE CÂMERA.....	22
2.2	VISÃO COMPUTACIONAL.....	26
2.3	VISÃO ESTÉREO.....	27
2.3.1	Princípios da estereoscopia.....	28
2.3.1.1	<i>Projeção de luz.....</i>	29
2.3.2	Calibração de sistemas estéreo.....	33
2.4	TÉCNICAS DE RECONSTRUÇÃO 3D.....	34
2.5	TRABALHOS CORRELATOS.....	38
3	MATERAIS E MÉTODOS.....	41
3.1	ARQUITETURA DO SISTEMA.....	41
3.2	COMPONENTES DE <i>HARDWARE</i>	48
3.2.1	Eletrônica de controle e sincronização.....	52
3.3	DESENVOLVIMENTO DO <i>SOFTWARE</i> DE RECONSTRUÇÃO.....	53
3.3.1	Calibração do sistema estéreo.....	54
3.3.2	Aquisição e pré-processamento de imagens.....	55
3.3.3	Triangulação e geração da nuvem de pontos.....	56
3.3.4	Pós-processamento.....	58
3.4	METODOLOGIA DE VALIDAÇÃO.....	59
3.4.1	Métrica de qualidade da correlação: Relação Pico-Ruído (PNR).....	59
3.4.2	Avaliação do desempenho metrológico.....	60
4	RESULTADOS E DISCUSSÃO.....	63
4.1	RESULTADOS DA CALIBRAÇÃO.....	63
4.2	DEFINIÇÃO E OTIMIZAÇÃO DOS PARÂMETROS.....	65
4.3	RECONSTRUÇÃO DE OBJETOS PADRÃO.....	80
4.4	RECONSTRUÇÃO DE OBJETOS DE FORMA LIVRE.....	86

4.5	DISCUSSÃO	90
5	CONCLUSÃO	93
5.1	RECAPITULAÇÃO DO TRABALHO	93
5.2	CONTRIBUIÇÕES	94
5.3	LIMITAÇÕES DO TRABALHO.....	95
5.4	SUGESTÃO DE TRABALHOS FUTUROS	96
	REFERÊNCIAS	98

1 INTRODUÇÃO

A Engenharia Mecânica moderna está intrinsecamente ligada à capacidade de transitar entre o mundo físico e o digital. A obtenção de modelos virtuais fiéis de componentes e sistema, os conhecidos gêmeos digitais, transformou a indústria e os seus ciclos de projeto, análise e fabricação. Inseridas diretamente nesse contexto, a metrologia dimensional, ciência da medição de comprimentos, ângulos e formas geométricas, em conjunto com a metrologia óptica, área da metrologia que utiliza a luz para efetuar as medições, assumem um papel fundamental. A obtenção de dados tridimensionais precisos de um objeto físico é a base para a engenharia reversa, a inspeção da qualidade, o controle dimensional de processos de fabricação e a prototipagem rápida.

1.1 CONTEXTUALIZAÇÃO

Tradicionalmente, a medição dimensional de alta exatidão, caracterizada por baixas incertezas de medição, era dominada por sistemas de contato, como as máquinas de medição por coordenadas (MMC). Apesar de seu bom desempenho metrológico, tais sistemas apresentam limitações quanto à velocidade da medição e à capacidade de capturar geometrias complexas e de forma livre de maneira densa (MATACHE et al., 2015). Como alternativa, as tecnologias de digitalização 3D sem contato, especialmente as baseadas em óptica, ganharam grande destaque nas últimas décadas.

Estes sistemas ópticos podem ser classificados em duas categorias principais: passivos e ativos. Os métodos passivos, como a fotogrametria, reconstroem a geometria a partir da correspondência de características em imagens de uma cena sob iluminação ambiente, contudo, apresentam dificuldades em superfícies com pouca textura ou sob condições de iluminação não controladas (LUO et al., 2024). Por outro lado, os métodos ativos superam essa limitação ao projetarem ativamente um padrão de luz (que podem ser linhas, pontos ou franjas) sobre o objeto. Isso cria uma textura artificial na superfície fazendo com que o problema da correspondência seja simplificado e deixando a medição independente da textura natural da cena.

Dentro das técnicas ativas, a triangulação *laser* é uma das mais robustas e difundidas (GENG, 2011), assim como outras abordagens baseadas em luz estruturada, como a projeção de franjas. O princípio da triangulação *laser* consiste em projetar linhas ou pontos de *laser* sobre a superfície do objeto e capturar imagens com uma ou mais câmeras a partir de ângulos conhecidos. A variação da posição dos pontos projetados nas imagens permite, através da triangulação, calcular as coordenadas 3D dos milhares de pontos, gerando assim uma nuvem de pontos que representa a superfície do objeto.

Para superar as limitações de densidade da triangulação *laser* convencional e, ao mesmo tempo, evitar as restrições de *hardware* dos sistemas de projeção de franjas (que demandam projetores digitais de maior porte e maior profundidade de foco), este trabalho propõe uma arquitetura alternativa. O método desenvolvido emprega um sistema de visão estéreo ativo baseado em um *laser*, projetando um padrão de pontos pseudoaleatórios. Ao contrário dos métodos estáticos (*single-shot*), o padrão é movido dinamicamente gerando uma sequência de imagens, o que permite aplicar uma métrica de correlação espaço-temporal que combina a qualidade da análise de séries temporais típica da projeção de franjas com a versatilidade física de um projetor *laser*, resultando em um sistema compacto e robusto para cenas complexas.

Atualmente, existem diversos sistemas comerciais de digitalização 3D que empregam essa tecnologia, oferecendo alta resolução e exatidão. No entanto, o custo de aquisição desses equipamentos usualmente é elevado, representando uma barreira para laboratórios de pesquisa, pequenas e médias empresas e dificultando a sua integração em projetos de orçamento limitado. Além disso, o fato de a tecnologia por trás desses sistemas muitas vezes não ser disponibilizada para o usuário gera uma restrição de flexibilidade e customização, criando assim uma lacuna para o desenvolvimento de sistemas de digitalização 3D de arquitetura aberta, de forma mais acessível e bom desempenho metrológico. Adicionalmente, esta abordagem permite que sejam feitos projetos de sistemas otimizados que atendem a demandas específicas, como aplicações em ambientes com restrições ou aplicações subaquáticas, requisitos que sistemas comerciais genéricos não necessariamente atendem plenamente.

1.2 PROBLEMA DE PESQUISA

Diante do contexto apresentado, observa-se uma lacuna tecnológica na disponibilidade de sistemas de digitalização 3D que sejam, simultaneamente, de arquitetura aberta e adaptáveis às restrições de veículos subaquáticos. As soluções comerciais existentes muitas vezes operam como “caixas-pretas”, impedindo a customização necessária para inspeções em ambientes confinados ou sob condições ópticas adversas típicas do cenário *offshore*. Portanto, esse trabalho busca responder à seguinte questão: como desenvolver e validar um sistema de visão estéreo ativa, baseado na projeção *laser* de um padrão de pontos pseudoaleatórios, capaz de realizar a reconstrução 3D de cenas com incerteza submilimétrica, visando a futura integração ao Projeto VORIS?

1.3 JUSTIFICATIVA

A resposta ao problema de pesquisa proposto possui relevância sob as perspectivas tecnológica e científica. Do ponto de vista tecnológico e industrial, o desenvolvimento de um sistema de digitalização 3D de arquitetura aberta se justifica pela necessidade de flexibilidade e controle sobre os parâmetros de aquisição, características frequentemente ausentes em soluções comerciais fechadas. Embora existam equipamentos de mercado com alto desempenho, a impossibilidade de acesso e customização de seus algoritmos de controle e *hardware* representa uma barreira técnica para aplicações em ambientes não estruturados ou hostis. A presente pesquisa visa superar essa limitação, propondo uma arquitetura modular que permite a otimização dedicada do sistema para requisitos específicos de volume, geometria e óptica.

Nesse contexto, este trabalho atende diretamente a uma demanda do Projeto VORIS, do Laboratório de Metrologia e Automatização (LABMETRO) da Universidade Federal de Santa Catarina (UFSC). A tecnologia aqui desenvolvida tem como aplicação final a inspeção subaquática de estruturas *offshore* na indústria de óleo e gás, na qual a reconstrução 3D de cenas em distâncias operacionais compatíveis com ROVs, no caso desse sistema de até 3 m, é um requisito fundamental, permitindo a identificação e monitoramento de corrosão, trinca em soldas, amassados estruturais e bioincrustação. Isso demonstra não apenas uma

aplicação prática e imediata para a pesquisa, mas também a sua inserção em um contexto de alto valor tecnológico e industrial.

A conexão com o Projeto VORIS e a possibilidade de uma posterior integração do sistema desenvolvido ao seu sistema guiou a seleção de componentes chave. A utilização de um projetor de pontos pseudoaleatórios de laser azul (comprimento de onda de aproximadamente 450 nm), por exemplo, foi uma decisão técnica justificada pela aplicação final, uma vez que o seu comprimento de onda mais curto apresenta melhor propagação em ambientes subaquáticos. Essa característica é fundamental para superar os desafios da reconstrução 3D sob a água, um tema investigado por autores como Zhuang et al. (2023).

Sob a ótica científica e acadêmica, a contribuição deste trabalho reside no detalhamento do ciclo de desenvolvimento de um sistema de digitalização 3D, desde a seleção de componentes até a sua validação. A análise de desempenho e a análise quantitativa dos erros do sistema em um caso de medição, apresentado no Capítulo 4, estabelecem um referencial sobre a incerteza alcançável com os componentes de *hardware* escolhidos, servindo como guia para futuros pesquisadores que busquem desenvolver soluções de visão estéreo customizadas para ambientes complexos.

1.4 OBJETIVOS

O objetivo geral desta dissertação é desenvolver e validar metrologicamente um sistema de visão estéreo ativa com projeção *laser* estruturada de uma matriz de pontos pseudoaleatórios, projetado para reconstrução 3D de estruturas e componentes subaquáticos da indústria de óleo e gás, visando atender aos requisitos de inspeção visual automatizada.

Para que esse objetivo geral seja atingido, devem ser atingidos os seguintes objetivos específicos:

- Especificar e selecionar os componentes de *hardware* (câmeras, lentes, filtros e projetor *laser*) adequados para a digitalização de cenas a distâncias de operação compatíveis com a navegação de ROVs (neste caso de até 3 metros), considerando a futura aplicação subaquática;

- Integrar os sistemas óptico e mecânico, desenvolvendo uma estrutura de montagem rígida e compacta que garanta a estabilidade da calibração estereoscópica;
- Implementar o controle de variação temporal do padrão *laser* utilizando prismas contrarrotativos, visando maximizar a robustez da correspondência estéreo;
- Desenvolver e otimizar a rotina computacional baseada na correlação espaço-temporal, ajustando os parâmetros de reconstrução para garantir uma densidade de pontos adequada à detecção de falhas;
- Avaliar o desempenho metrológico do sistema desenvolvido por meio da medição de objetos padrão, verificando se a incerteza de medição obtida é compatível com o requisito de exatidão submilimétrica;
- Validar a robustez da reconstrução em objetos de forma livre que simulem a complexidade geométrica de componentes reais de plataformas, avaliando a qualidade da nuvem de pontos em situações de oclusão e variação de curvatura.

1.5 ESTRUTURA DA DISSERTAÇÃO

Este trabalho está estruturado em cinco capítulos.

O Capítulo 2 apresenta a fundamentação teórica necessária para a compreensão do projeto. São abordados os princípios do modelo de câmera *pinhole*, os fundamentos da visão computacional e da visão estéreo, com ênfase nas técnicas de reconstrução 3D ativas por luz estruturada e nos métodos de calibração. Ao final do capítulo, são discutidos trabalhos correlatos da literatura no estado da arte.

O Capítulo 3 descreve os materiais e métodos empregados. Nele são detalhados a arquitetura do sistema, os componentes de *hardware* utilizados e a metodologia para o desenvolvimento do *software* de reconstrução, que engloba as etapas de calibração, aquisição de imagens, triangulação e pós-processamento da nuvem de pontos. A metodologia de validação para análise dos erros de medição do sistema também é apresentada.

O Capítulo 4 expõe os resultados obtidos e promove a sua discussão. São apresentados os resultados da calibração, a reconstrução de objetos padrão e cenas complexas. A análise quantitativa dos erros de medição e a avaliação da qualidade metrológica do sistema é o foco central deste capítulo.

Finalmente, o Capítulo 5 apresenta a conclusão do trabalho, recapitulando a pesquisa realizada, destacando as principais contribuições, apontando as limitações do sistema desenvolvido e, por fim, sugerindo direções para trabalhos futuros.

2 FUNDAMENTAÇÃO TEÓRICA

O objetivo deste capítulo é revisitar os principais conceitos e técnicas para a reconstrução tridimensional de cenas a partir de imagens bidimensionais, um problema central em visão computacional. A abordagem aqui adotada, baseada em visão estéreo ativa com luz estruturada, é consolidada na metrologia óptica, cujos fundamentos e aplicações são extensivamente cobertos em tutoriais como o de Geng (2011), mas que continua evoluindo com novas arquitetura de processamento.

O problema central da reconstrução 3D reside na correspondência estéreo: identificar, em diferentes vistas, os *pixels* que representam o mesmo ponto físico no espaço (SCHARSTEIN; SZELISKI, 2002). Em cenas com regiões de baixa textura ou iluminação homogênea, métodos clássicos de correlação falham devido à falta da informação visual distinta. Historicamente, estratégias de normalização como a Correlação Cruzada Normalizada (NCC) e a Correlação Cruzada Normalizada a Zero (ZNCC) foram fundamentais para lidar com variações de ganho e exposição entre câmeras (LEWIS, 1995).

Mais recentemente, abordagens baseadas em aprendizado profundo (*deep learning*) têm dominado o estado da arte em *benchmarks* de visão estéreo passiva. Redes neurais convolucionais são treinadas para “aprender” métricas de similaridade complexas, superando métodos manuais em cenários desafiadores (ZBONTAR; LECUN, 2016). Contudo, essas abordagens dependem de grandes volumes de dados de treinamento anotados e podem sofrer com a generalização em domínios específicos. Em contraste, os métodos geométricos e fotométricos clássicos, foco deste trabalho, oferecem previsibilidade, rastreabilidade metrológica e não dependem de treinamento prévio.

Para superar a dependência da textura natural da cena, os métodos ativos projetam padrões de luz controlados sobre o objeto ou cena analisada (Salvi et al., 2004). A escolha do padrão define as capacidades do sistema, códigos binários, como por exemplo o *gray code*, oferecem alta robustez contra ruídos, mas exigem a projeção de uma longa sequência de imagens (*multi-shot*) para decodificar a posição (HORN; KIRYATI, 1999), tornando-os limitados a cenas estáticas. Já os padrões de projeção de franjas, que analisam a defasagem de fase, proporcionam a mais alta resolução espacial e precisão (sub-*pixel*), mas também exigem a técnica *multi-shot*.

Em outra vertente, os padrões pseudoaleatórios (de pontos ou *speckle*) se destacam pela versatilidade. Podem ser utilizados em abordagens *single-shot* (uma única captura), permitindo a reconstrução de cenas dinâmicas à custa de menor resolução espacial (Dekiff et al., 2010), ou em abordagens temporais para aumentar a densidade da nuvem de pontos (PINTO; ALBERTAZZI, 2013). A distinção entre métodos de disparo único (*single-shot*) e múltiplos disparos (*multi-shot*) é crítica, enquanto os primeiros priorizam a velocidade para capturar objetos em movimento, os últimos priorizam a densidade e a qualidade metrológica, exigindo, em contrapartida, a estabilidade temporal da cena durante a sequência de captura.

O capítulo estrutura-se da seguinte forma: o ponto de partida é a compreensão do modelo de câmera, que descreve matematicamente o processo de formação da imagem. A partir disso, aborda-se o problema da visão estéreo em suas três etapas: calibração e reconstrução. Por fim, detalham-se as técnicas de luz estruturada e apresenta-se uma análise de trabalhos correlatos para contextualizar a presente pesquisa no estado da arte.

2.1 MODELO DE CÂMERA

O processo de formação da imagem em uma câmera digital é um processo óptico tridimensional que é mapeado em um plano bidimensional. Para que se possa extrair informação métrica a partir de imagens, é indispensável o uso de um modelo matemático que descreva essa transformação de forma precisa. O modelo fundamental e mais amplamente utilizado em visão computacional é o modelo de câmera *pinhole*, também conhecido como câmera escura, que mesmo sendo relativamente simples, descreve com eficácia a relação geométrica entre um ponto no espaço 3D e sua projeção no plano da imagem (HARTLEY; ZISSERMAN, 2004).

O modelo idealiza a câmera como uma caixa opaca ou ambiente escuro, que tenha um pequeno orifício, denominado centro de projeção ou centro óptico, por onde a luz passa e forma uma imagem na parede oposta, denominado plano da imagem, com isso é percebida a formação de uma imagem invertida da cena externa formada através desse orifício. Um exemplo dessa formação de imagem é apresentado na Figura 1.

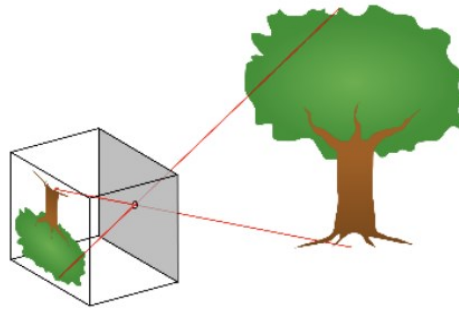


Figura 1 - Esquema da câmera pinhole.
Fonte: Mellish, 2025.

O modelo *pinhole* físico, ilustrado na Figura 1, resulta em uma imagem geometricamente invertida. No entanto, o modelo matemático utilizado na visão computacional é comumente representado por um plano de imagem virtual posicionado à frente do centro óptico, a uma distância focal f , conforme ilustrado na Figura 2.

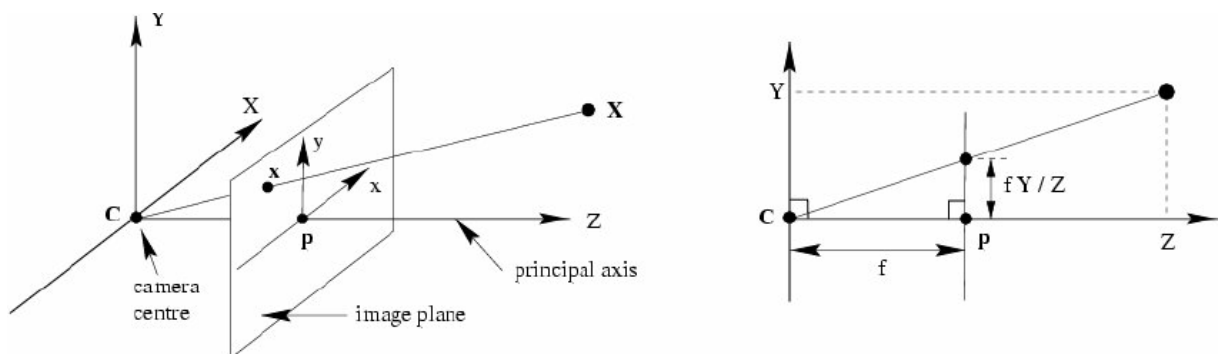


Figura 2 – Modelo matemático da câmera pinhole.
Fonte: Forsman, 2010.

Hartley e Zisserman (2004) explicam que este modelo de “plano frontal” resulta em uma imagem projetada com a mesma orientação da cena real. Essa abordagem é válida pois reflete como a imagem é efetivamente salva pelo sensor: o *hardware* da câmera ou o *software* já desfaz a inversão física, entregando uma imagem orientada como se estivessemos observando a cena.

A derivação do modelo completo envolve a relação entre diferentes sistemas de coordenadas: do mundo, da câmera, do plano imagem e o de *pixels*. Um ponto P_e no espaço, descrito por suas coordenadas no sistema de referência do mundo $P_e = [X_e, Y_e, Z_e]^T$, precisa ser expresso em relação ao sistema de coordenadas da

câmera $P_c = [X_c, Y_c, Z_c]^T$. Essa transformação é uma operação de corpo rígido, envolvendo uma rotação e uma translação, e é definida pelos parâmetros extrínsecos da câmera. A transformação é dada por:

$$P_c = R \cdot P_e + t \quad (\text{Equação 1})$$

Onde:

- R é uma matriz de rotação 3×3 que alinha os eixos do sistema do mundo com a câmera;
- t é um vetor de translação 3×1 que descreve a posição da origem do sistema do mundo em relação à origem do sistema da câmera.

Os três parâmetros obtidos para rotação combinados com os três parâmetros obtidos para a translação definem a posição e a orientação da câmera no espaço. Uma vez que o ponto está no referencial da câmera (P_c), sua projeção no plano imagem, com coordenadas (x, y) , pode ser calculada por semelhança de triângulos. Este mapeamento é definido pelos parâmetros intrínsecos, que são características internas da câmera combinadas com a distância focal (f) da câmera.

$$x = f \frac{X_c}{Z_c} \quad (\text{Equação 2})$$

$$y = f \frac{Y_c}{Z_c} \quad (\text{Equação 3})$$

Onde:

- (x, y) são as coordenadas métricas do ponto projetado no plano imagem;
- (X_c, Y_c, Z_c) são as coordenadas 3D do ponto no referencial da câmera.

As coordenadas (x, y) no plano imagem são métricas, portanto, precisam ser convertidas para as coordenadas do *pixel* (u, v) , cujo sistema geralmente tem origem no canto superior esquerdo da imagem. Essa conversão é uma transformação afim que envolve o ponto principal e o tamanho dos *pixels*.

$$u = c_x + \frac{x}{s_x} \quad (\text{Equação 4})$$

$$v = c_y + \frac{y}{s_y} \quad (\text{Equação 5})$$

Onde:

- (c_x, c_y) são as coordenadas do ponto principal, que é a projeção do centro óptico no plano imagem (idealmente o centro da imagem);

- s_x e s_y são os tamanhos físicos de um *pixel* nas direções horizontal e vertical, respectivamente, variáveis que fazem a conversão de unidades métricas para *pixels*.

Para unificar essas etapas em uma única equação matricial, utilizam-se coordenadas homogêneas. A relação completa entre um ponto 3D $P_e = [X_e, Y_e, Z_e, 1]^T$ e sua projeção em *pixels* $p = [u, v, 1]^T$ é dada pela matriz de projeção da câmera P :

$$s \cdot p = P \cdot P_e = K [R|t] \cdot P_e \quad (\text{Equação 6})$$

Onde:

- s é um fator de escala arbitrário;
- K é a matriz de parâmetros intrínsecos, definida como:

$$K = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \quad (\text{Equação 7})$$

Com $f_x = f/s_x$ e $f_y = f/s_y$ sendo as distâncias focais em *pixels*;

- $[R|t]$ é a matriz de parâmetros extrínsecos. 3×4 .

O modelo *pinhole* é uma idealização, lentes reais introduzem distorções não lineares na projeção da imagem, que precisam ser modeladas para aplicações de alta precisão. As distorções mais significativas são a radial e a tangencial (ZHANG, Z., 2000). A distorção radial é uma aberração geométrica decorrente da variação na ampliação transversal da lente em função da distância de uma câmera *pinhole*, onde a ampliação é constante, em lentes reais a luz que incide nas bordas da lente sofre um desvio angular distinto daquela que incide no centro. Isso resulta em um deslocamento não linear dos pontos na imagem, manifestando-se tipicamente como distorção tipo barril, onde a ampliação diminui nas bordas, ou tipo almofada, onde a ampliação aumenta nas bordas. Essa distorção é comumente modelada por um polinômio que tem como componentes os parâmetros k_1 , k_2 e k_3 .

Já a distorção tangencial ocorre quando a lente e o plano do sensor de imagem não estão perfeitamente paralelos, essa distorção é modelada por dois parâmetros, p_1 e p_2 . Todos esses parâmetros envolvidos com a distorção são considerados parte do conjunto de parâmetros intrínsecos da câmera, por serem

características internas do conjunto lente-sensor. A determinação dos parâmetros é o objetivo do processo de calibração, que será abordado na secção 2.3.2.

2.2 VISÃO COMPUTACIONAL

A Visão Computacional é um campo interdisciplinar da ciência da computação e da inteligência artificial que busca capacitar os computadores a “ver”, interpretar e compreender o mundo visual. De maneira análoga ao que acontece com a visão humana, a área busca extrair informações e derivar conclusões a partir de imagens digitais, vídeos ou outras formas de dados visuais de alta dimensionalidade (SZELISKI, 2010). O seu objetivo não é apenas processar as imagens, mas também tomar decisões ou até mesmo construir modelos descritivos a partir delas.

Uma vasta gama de tarefas e aplicações é abrangida, que podem ser categorizadas de acordo com o tipo de informação que se deseja extrair. Entre as principais áreas de pesquisa e aplicação destacam-se:

- Reconhecimento de objetos: identificar e classificar objetos específicos dentro de uma imagem, como por exemplo um carro ou uma face humana;
- Segmentação de imagens: particionar uma imagem em múltiplas regiões ou conjuntos de *pixels*, geralmente para isolar objetos de interesse em relação ao fundo;
- Detecção de características: encontrar pontos de interesse específicos em uma imagem, como cantos, bordas ou texturas salientes, que são verificados com mudanças de iluminação, escala e rotação;
- Análise de movimento e reconstrução 3D: determinar a estrutura tridimensional de uma cena e/ou o movimento da câmera a partir de uma ou mais imagens bidimensionais.

O desafio central da Visão Computacional se encontra no fato de que ela busca resolver um problema inverso e mal condicionado, conhecido como *ill-posed problem*. O processo de formação da imagem, como descrito na secção 2.1, projeta o mundo 3D em um plano 2D, perdendo de forma inerente a informação de profundidade. A tarefa passa a ser inferir as propriedades do mundo 3D, a causa, a

partir de suas projeções 2D, o efeito, um processo que é fundamentalmente ambíguo quando se utiliza uma única imagem (GONZALEZ; WOODS, 2010).

Para superar essa ambiguidade e recuperar a informação de profundidade perdida, diversas técnicas foram desenvolvidas. Dentre as mais fundamentais e eficazes, destaca-se a Visão Estéreo, que se inspira diretamente no sistema visual humano ao utilizar duas ou mais vistas da mesma cena para inferir a estrutura tridimensional.

2.3 VISÃO ESTÉREO

Conforme estabelecido anteriormente, uma única imagem 2D é insuficiente para inferir de forma inequívoca a profundidade de uma cena. A Visão Estéreo, também conhecida como Estereoscopia, contorna essa limitação fundamental inspirada no sistema visual binocular humano. Ao utilizar duas ou mais imagens de uma mesma cena, capturadas de diferentes e conhecidos pontos de vista, é possível triangular a posição de pontos no espaço e reconstruir a sua geometria tridimensional (MARR; POGGIO, 1979).

O problema central da visão estéreo pode ser decomposto em três etapas sequenciais:

1. Calibração do sistema: determinação precisa dos parâmetros intrínsecos de cada câmera e dos parâmetros extrínsecos, rotação e translação, que definem a relação geométrica entre elas;
2. Estabelecimento de correspondência: para um ponto de interesse em uma imagem, encontrar o ponto correspondente na outra imagem que seja a projeção do mesmo ponto 3D no mundo. Este é conhecido como o problema da correspondência, um desafio central da estereoscopia cuja taxonomia e avaliação de algoritmos foram detalhadamente explorados no trabalho de Scharstein e Szeliski (2002);
3. Reconstrução 3D (triangulação): utilizando os parâmetros do sistema calibrado e os pontos correspondentes, calcular as coordenadas 3D do ponto no espaço.

2.3.1 Princípios da estereoscopia

O princípio fundamental por trás da reconstrução através da visão estéreo é a triangulação, um processo que calcula a posição 3D de um ponto a partir de suas projeções em duas ou mais imagens. Conforme apresentado na Figura 3, o processo de reconstrução da triangulação clássica ocorre na direção 2D para a 3D.

Primeiro, o algoritmo de correspondência encontra os pontos p_1 (imagem esquerda) e p_2 (imagem direita) que são projeções do mesmo ponto 3D no mundo (P). Em seguida, conhecendo os parâmetros de calibração, duas retas 3D são traçadas, uma partindo do centro óptico da câmera esquerda (O_1) e passando por p_1 , e outra partindo do centro óptico da câmera direita (O_2) e passando por p_2 . Idealmente, a interseção dessas duas retas no espaço revela a localização 3D original do ponto P .

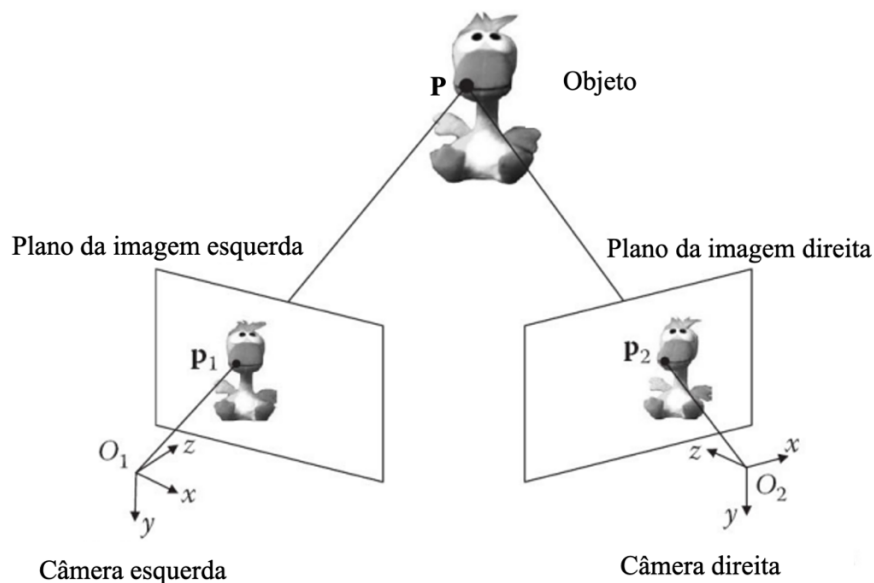


Figura 3 – Sistema de visão estéreo.
Fonte: Zhang, S. (2013) adaptado pela autora (2025).

A busca por pontos correspondentes entre as duas imagens, conhecida como o problema da correspondência, seria computacionalmente custoso se fosse necessário procurar por toda a segunda imagem. A solução para isso é encontrada na geometria epipolar, que restringe a busca de correspondência a uma única linha. Dados dois centros de câmera O_1 e O_2 , e um ponto P_1 na imagem 1, o ponto correspondente P_2 na imagem 2 deve, obrigatoriamente, pertencer a uma linha específica chamada de linha epipolar. Essa linha é a projeção, na imagem 2, do raio

que conecta o ponto 3D P ao centro óptico O_1 . Essa restrição geométrica reduz drasticamente a complexidade da busca por correspondência de um problema 2D para um problema 1D (HARTLEY; ZISSERMAN, 2004).

Para simplificar ainda mais o processo, as imagens de um par estéreo são geralmente retificadas. A retificação da imagem é uma transformação projetiva que alinha os planos das duas imagens de forma que todas as linhas epipolares se tornem paralelas e horizontais, coincidindo com as linhas de varredura das imagens. Após a retificação, para um ponto (u_1, v_1) na imagem 1, o ponto correspondente (u_2, v_2) na imagem 2 terá a mesma coordenada vertical, ou seja, $v_1 = v_2$. Com isso, a busca fica, então, restrita à mesma linha horizontal. A diferença nas coordenadas horizontais entre os pontos correspondentes é chamada de disparidade (d):

$$d = u_1 - u_2 \text{ (Equação 8)}$$

A disparidade é inversamente proporcional à profundidade (Z) do ponto na cena. Pontos mais próximos da câmera terão uma disparidade maior, enquanto pontos muito distantes terão uma disparidade que tende a zero. Para este caso idealizado de configuração canônica (eixos paralelos), ou após a aplicação da retificação das imagens, a relação geométrica entre a profundidade (Z) e a disparidade (d) torna-se linear e inversamente proporcional, sendo descrita por:

$$Z = \frac{f \cdot B}{d} \text{ (Equação 9)}$$

Onde:

- f é a distância focal das câmeras (em *pixels*);
- B é a *baseline*, ou seja, a distância entre os centros ópticos das duas câmeras;
- d é a disparidade calculada.

2.3.1.1 *Projeção de luz*

A escolha do tipo de projeção de luz, sua cor (comprimento de onda) e método de identificação é uma decisão de projeto crucial para definir a velocidade, resolução e incerteza do sistema de medição. Os padrões de luz estruturada podem

ser classificados em diversas categorias, mas a divisão principal se dá entre métodos temporais (que exigem múltiplas aquisições) e espaciais (que funcionam com uma única aquisição).

No caso dos padrões temporais (*multi-shot*), por exemplo, estes métodos projetam uma sequência de padrões ao longo do tempo sobre a cena. Cada *pixel* na imagem da câmera recebe uma espécie de “código” temporal único. Exemplos clássicos incluem a projeção de códigos binários (*gray code*) e os métodos de projeção de franjas senoidais (*phase-shift*) (Figura 4), em que são projetados tipicamente três ou mais padrões de franjas com deslocamento de fase controlados. Embora sejam muito robustos e precisos, exigem que a cena e o sistema permaneçam estáticos durante toda a sequência de projeção, tornando-os inadequados quando há movimento relativo entre o sistema e a cena medida (Salvi et al., 2004).

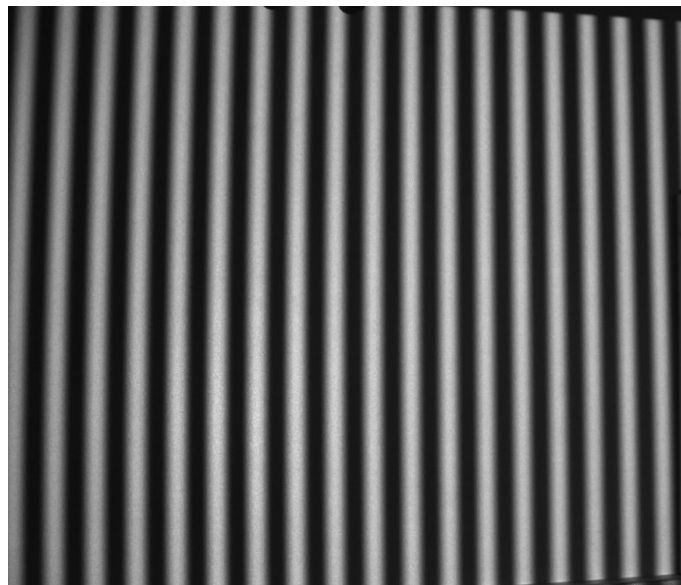


Figura 4 – Padrão de franjas senoidais.
Fonte: A Autora.

Já no caso dos padrões codificados espacialmente (*single-shot*), para superar a limitação dos métodos temporais, os padrões codificados espacialmente incorporam toda a formação de correspondência em um único padrão. Isso permite a reconstrução 3D a partir de uma única aquisição de imagens, viabilizando a digitalização de objetos sob movimentação. Nesta categoria se destacam os padrões de banda limitada (BLP) e os padrões de pontos pseudoaleatórios (DOE).

A técnica dos padrões de banda limitada (BLP) projeta um padrão contínuo, não repetitivo, que se assemelha a manchas (Figura 5). O padrão é gerado algoritmicamente para que seu conteúdo de frequência espacial seja controlado, a sua alta aleatoriedade local o torna ideal para algoritmos de correlação espacial.

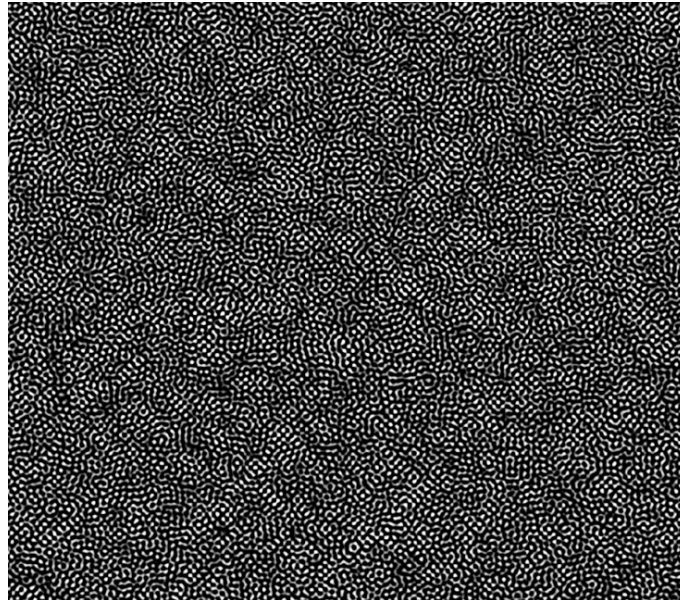


Figura 5 – Padrão de banda limitada.
Fonte: A Autora.

Já na técnica dos padrões de pontos pseudoaleatórios (DOE), abordagem utilizada neste trabalho, o sistema utiliza um elemento óptico difrativo (DOE) para dividir um único feixe de *laser* em uma matriz de milhares de pontos (Figura 6). O padrão gerado parece aleatório, porém na realidade é pseudoaleatório, por ser um padrão determinístico e conhecido, sua principal propriedade é que cada “vizinhança” de pontos na matriz é matematicamente única. Um exemplo comercial conhecido que popularizou esta técnica é o sensor Kinect de primeira geração, que projetava uma matriz de pontos em infravermelho.

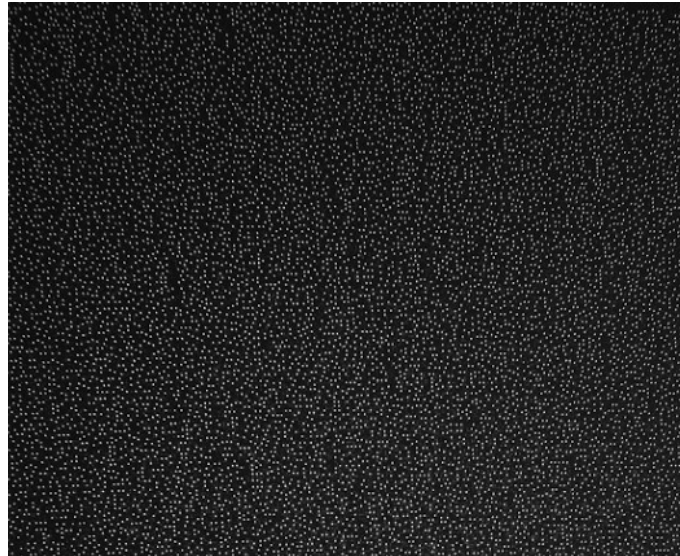


Figura 6 – Padrão de pontos pseudoaleatórios.
Fonte: A Autora.

O método para identificar o padrão projetado está intrinsecamente ligado ao tipo de padrão. No caso dos padrões pseudoaleatórios por exemplo, o problema da correspondência é resolvido por meio da correlação, que será apresentada na seção 2.4. A escolha da luz projetada é uma escolha de engenharia que impacta diretamente a qualidade da medição, de acordo com o ambiente e superfície em que será realizada.

A luz infravermelha é frequentemente utilizada em aplicações de consumo, por ser invisível ao olho humano e menos suscetível a interferências da iluminação ambiente visível. A luz vermelha é utilizada em casos que priorizam o baixo custo, pela alta disponibilidade de diodos de *laser* vermelho. Já a luz azul é considerada um padrão-ouro para aplicações de metrologia industrial de alta precisão, por possuir um comprimento de onda mais curto (entre 405 e 450 nm) em comparação com a luz vermelha (635 a 670 nm), permitindo assim melhor foco e gerando projeções mais finas e nítidas. Além disso, a luz azul sofre menos dispersão subsuperficial, resultando em medições mais precisas e com menos ruídos (CHIANG et al., 2011).

Adicionalmente, o comprimento de onda azul apresenta menor atenuação e absorção na água em comparação com comprimentos de onda mais longos, como o vermelho, sendo ideal para aplicações de medição subaquáticas (LIU et al., 2024).

2.3.2 Calibração de sistemas estéreo

Para que a análise das capturas realizadas com sistemas estéreos seja feita de forma válida e precisa os parâmetros que são utilizados nos cálculos devem ser conhecidos com exatidão. A calibração do sistema é o processo que permite determinar não apenas os parâmetros intrínsecos de cada câmera individualmente (matriz K e coeficientes de distorção, como visto em 2.1), mas também a relação geométrica entre elas.

Essa relação é descrita pelos parâmetros extrínsecos do par estéreo: uma matriz de rotação R e um vetor de translação t , que transformam o sistema de coordenadas da câmera 1 no sistema de coordenadas da câmera 2. A partir do vetor de translação t é possível determinar o *baseline* (B) do sistema.

O método de calibração mais comum, popularizado por Zhang, Z. (2000), tornou-se o padrão da indústria e da academia devido à sua flexibilidade e precisão. Ele utiliza um objeto padrão com um padrão geométrico bem definido, como os apresentados na Figura 7, que é capturado em múltiplas posições e orientações.



Figura 7 – Padrões de calibração.
Fonte: MATLAB, 2025.

A robustez desse método levou à sua implementação em diversas ferramentas de *software*, incluindo bibliotecas de código aberto como a OpenCV e pacotes de computação técnica como o *Stereo Camera Calibrator App*, parte da *Image Processing and Computer Vision Toolbox* do MATLAB. Essas ferramentas automatizam a detecção dos pontos de referência no padrão, o cálculo dos

parâmetros intrínsecos e extrínsecos e a estimativa das distorções da lente, com base nas informações de passo dos artefatos no padrão, que pode ser pré-calibrado.

2.4 TÉCNICAS DE RECONSTRUÇÃO 3D

As abordagens para a reconstrução tridimensional a partir de imagens são diversas, mas podem ser categorizadas de forma geral com base na maneira como a informação sobre a cena é adquirida. A divisão mais clássica e fundamental se dá entre métodos passivos e métodos ativos (BESL, 1988). A principal diferença entre elas está na dependência ou não de uma fonte de energia externa e controlada para iluminar o objeto e permitir a extração de suas características geométricas.

Os métodos passivos são aqueles que utilizam apenas a iluminação ambiente para capturar imagens da cena. Eles dependem inteiramente das características visuais intrínsecas dos objetos, como bordas, textura e cor para realizar a reconstrução. As duas técnicas mais proeminentes são a estereoscopia convencional e a fotogrametria.

Na estereoscopia convencional, conforme descrito na seção 2.3.1, utiliza-se um par de câmeras com geometria fixa e pré-calibrada. A reconstrução é baseada no cálculo da disparidade a partir da correspondência de pontos em um único par de imagens retificadas. A fotogrametria, por sua vez, é a ciência de obter informações confiáveis sobre objetos físicos e o meio ambiente através de processos de gravação, medição e interpretação de imagens fotográficas (LUHMANN et al., 2014).

Embora seja um termo amplo, no contexto da reconstrução 3D moderna, a fotogrametria está frequentemente associada à técnica de Estrutura a partir do Movimento (*Structure from Motion* – SfM). O SfM permite reconstruir a geometria 3D de uma cena a partir de um conjunto de imagens não ordenadas e capturadas de múltiplos pontos de vista, sem a necessidade de conhecer previamente a posição das câmeras. O algoritmo resolve simultaneamente a estrutura 3D da cena e os parâmetros de pose (posição e orientação) de cada câmera (SZELISKI, 2010).

O desafio central de todos os métodos passivos é o problema da correspondência. Antes de medir a similaridade entre regiões de imagens diferentes, no entanto, é necessário determinar quais pontos são suficientemente distintos para serem encontrados de forma confiável em outra vista. Essa tarefa é realizada por detectores e descritores de características. Algoritmos clássicos como SIFT (*Scale-*

Invariant Feature Transform) foram marcos na área, capazes de encontrar pontos de interesse robustos a mudanças de escala e rotação (Lowe, 2004). Pouco depois, o SURF (*Speeded Up Robust Features*) foi desenvolvido como uma alternativa mais rápida, mantendo sua performance comparável (BAY et al., 2008).

Contudo, o uso de ambos foi historicamente restrito por patentes. Como alternativa, o algoritmo ORB (*Oriented FAST and Rotated BRIEF*) surgiu como uma solução de alto desempenho e livre de patentes, tornando-se amplamente adotado em aplicações de tempo real (RUBLEE et al., 2011). O ORB fundamenta-se nessa mesma estrutura de dois estágios, primeiramente utiliza o algoritmo *FAST* para a detecção rápida de pontos-chave (cantos ou bordas) e, na sequência, emprega uma versão modificada do descritor binário *BRIEF* para gerar um vetor descritor de características que codifica a aparência local de cada um desses pontos.

Uma vez que os descritores para os pontos de interesse em ambas as imagens são calculados, a correspondência é estabelecida encontrando os pares de descritores mais similares entre si. A principal ferramenta para validar e refinar essa correspondência é a correlação. A correlação é uma medida estatística que quantifica a similaridade entre dois sinais ou conjuntos de dados. No contexto da visão estereó, ela é utilizada para encontrar a correspondência entre uma região de *pixels* em uma imagem e a região mais similar na outra.

Dependendo da natureza dos dados de entrada, a técnica de correlação pode ser explorada de diferentes maneiras. A correlação espacial parte do princípio de que áreas geograficamente próximas em uma imagem tendem a ter características semelhantes, ou seja, *pixels* vizinhos costumam possuir valores parecidos. No contexto da estereoscopia, essa premissa é utilizada para encontrar correspondências. Para cada par de pontos correspondentes, uma pequena janela de *pixels* (matriz) ao seu redor é comparada com janelas candidatas na outra imagem. A métrica de Correlação Cruzada Normalizada a Zero (ZNCC) é particularmente robusta para essa comparação, pois é insensível a variações lineares de brilho e contraste entre as imagens (LEWIS, 1995). Seu uso para encontrar correspondências é bem estabelecido na literatura (DI STEFANO et al., 2005).

A ZNCC, que é matematicamente equivalente ao coeficiente de correlação de Pearson, é calculada pela seguinte equação:

$$\rho = \frac{\sum_{k=1}^M (g_k - \bar{g})(g'_k - \bar{g}')}{\sqrt{\sum_{k=1}^M (g_k - \bar{g})^2} \sqrt{\sum_{k=1}^M (g'_k - \bar{g}')^2}} \quad (\text{Equação 10})$$

Onde:

- g e g' são os vetores contendo os valores de intensidade de cinza extraídos dos *patches* espaço-temporais das imagens esquerda e direita, respectivamente;
- $\bar{g}$ e $\bar{g}'$ são os valores médios de intensidade de cinza para cada um desses vetores;
- M é o número total de pontos no *patch* espaço-temporal.

O cálculo resulta em um valor entre -1 e 1. A interpretação desses valores é fundamental para a estereoscopia:

- -1 (correlação perfeita negativa): indica um padrão invertido, onde um é claro e outro é escuro. Em visão estéreo, é tratado como uma correspondência incorreta (*mismatch*);
- 0 (sem correlação linear): não há semelhança útil entre os *patches*, sendo um resultado similar ao ruído aleatório;
- +1 (correlação perfeita positiva): indica que os *patches* são idênticos, mesmo considerando ganho ou *offset* linear. Este é o melhor resultado (*match*) e sinaliza uma correspondência correta.

Já na correlação temporal são analisados valores de uma mesma variável ao longo do tempo para identificar tendências e padrões. Em visão computacional, isso se traduz no rastreamento de características visuais em sequências de vídeo, sendo a base da técnica de SfM. A correspondência é estabelecida assumindo que a aparência de um ponto de interesse muda pouco entre quadros consecutivos. Na área de reconstrução 3D com luz estruturada, esta técnica foi explorada, por exemplo, por Pinto e Albertazzi (2013), que utilizaram a variação da intensidade de *pixels* individuais ao longo de uma sequência de projeções de padrões aleatórios para estabelecer a correspondência.

Existe também uma abordagem que, conceitualmente, combina as duas anteriores, que é a correlação espaço-temporal. Nela é possível analisar como os valores e características de uma cena mudam tanto no espaço (vizinhança do *pixel*) quanto no tempo (ao longo de uma sequência de imagens). Essa combinação permite obter mais informações sobre os dados e é aplicável a sistemas com

múltiplas câmeras capturando uma cena dinâmica, permitindo uma reconstrução mais robusta (ZHANG; CURLESS; SEITZ, 2003).

As técnicas de reconstrução também podem ser diferenciadas pela sua abordagem fundamental à triangulação. A metodologia clássica, descrita na Seção 2.3.1, parte da identificação de características 2D nas imagens (por exemplo, usando ORB) e, após estabelecer a correspondência 2D via correlação espacial, calcula a posição 3D do ponto.

Uma abordagem alternativa é o Método da Triangulação Inversa, também descrito como uma “organização orientada ao objeto”. Esta abordagem, formalizada por Pinto et al. (2012), inverte a lógica da triangulação clássica. No processo, conforme descrito pelos autores, primeiramente, define-se uma malha regular virtual no espaço do objeto, com coordenadas (X, Y) pré-definidas. A tarefa passa a ser encontrar a coordenada Z correspondente para cada par (X, Y) da malha, isso é feito através de uma varredura virtual do valor de Z ao longo de um intervalo ($Z_{\min}$, $Z_{\max}$).

Cada ponto 3D (X, Y, Z) testado durante a varredura é projetado matematicamente nos planos de imagem, ou mapas de fase, de ambas as câmeras. A métrica de correspondência utilizada é a diferença de fase absoluta entre os valores de fase extraídos das projeções em cada câmera, o ponto que apresentar a menor diferença absoluta de fase é considerado como pertencente à superfície, e sua coordenada Z é atribuída ao ponto (X, Y) da malha que está sendo avaliado. Este método permite que o ponto 3D e seus pontos homólogos nas imagens sejam determinados simultaneamente e o resultado é uma nuvem de pontos intrinsecamente estruturada em uma malha regular.

A eficácia de todos esses métodos, no entanto, está intrinsicamente ligada à premissa de que a cena possui textura suficiente para que os detectores de características encontrem pontos distintos e as métricas de correlação produzam resultados não ambíguos. Em contraste com os métodos passivos, os métodos ativos não dependem da textura do objeto, pois projetam luz estruturada na cena. Esta abordagem supera a principal limitação dos métodos passivos, tornando-os ideais para aplicações industriais onde as superfícies são difíceis de serem trabalhadas e analisadas, simplificando o problema da correspondência.

2.5 TRABALHOS CORRELATOS

A reconstrução 3D por meio de sistemas de visão estéreo ativa, especialmente aqueles que utilizam projeção *laser* em uma configuração estéreo, tem sido objeto de extensa pesquisa e desenvolvimento nas últimas décadas. Esta seção revisa alguns dos trabalhos mais pertinentes, a fim de contextualizar a presente pesquisa no estado da arte e destacar suas contribuições específicas.

Uma vertente significativa da pesquisa acadêmica tem se concentrado na democratização da tecnologia de digitalização 3D, propondo sistemas construídos com componentes simples ou de prateleira. Um marco nesse sentido foi o sistema “DAVID Laserscanner” descrito por Winkelback et al. (2006), que utilizava uma *webcam* e um *laser* de linha de mão, sendo considerado um método inovador de calibração que facilitava o seu uso.

Outro grupo de trabalho foca na minimização das incertezas, onde a precisão na detecção do padrão de luz é um ponto crítico. Métodos temporais, que projetam uma sequência de padrões foram amplamente explorados para garantir uma correspondência robusta, como estabelecido no trabalho de Horn e Kiryati (1999) sobre o uso de códigos binários.

Em um nível de sistema mais complexo, a tese de Stivanello (2013) aborda o complexo desafio da medição 3D de superfícies em ambientes dinâmicos, propondo um sistema híbrido baseado na projeção de franjas senoidais. Independentemente do padrão utilizado, a pesquisa de Steger (1998) continua sendo marco, ao propor um método para a extração de linhas e bordas com precisão sub-*pixel* baseado na matriz Hessiana, oferecendo maior robustez do que simples cálculos de centro de massa.

Enquanto muito sistemas focavam na triangulação clássica, outras pesquisas exploram paradigmas de reconstrução alternativos, como o Método da Triangulação Inversa apresentado conceitualmente na Seção 2.4. Um trabalho central que aplicou esta abordagem foi o de Pinto et al. (2012). Em sua implementação, os autores utilizaram a projeção de franjas senoidais, onde a correspondência de fase permitia reconstruir uma nuvem de pontos 3D já estruturada em um *grid* regular.

Dando continuidade a essa linha de pesquisa Pinto e Albertazzi (2013) exploraram o uso de padrões aleatórios e *laser speckle* para medição em malha

regular. Neste trabalho, os autores compararam o desempenho do Método da Triangulação Inversa com projeção de franjas senoidais, projeção de *laser speckle* e, com um padrão aleatório de banda limitada (BLP).

Para os padrões não determinísticos, a correspondência foi estabelecida por meio da correlação temporal, que analisa a variação da intensidade de cinza de um *pixel* ao longo de uma sequência de imagens. Os resultados validaram a técnica, demonstrando que a projeção de padrão aleatório (BLP) apresentou desempenho superior ao *laser speckle*, alcançando um desvio padrão de erro inferior a 0,01 mm na medição de um plano de referência ao utilizar uma sequência de 15 imagens.

A projeção de *speckle*, por sua simplicidade e capacidade de gerar texturas densas, também foi aplicada com sucesso em sistema de aquisição rápida (*single-shot*) para aplicações médicas, como demonstrado por Khan et al. (2018), e para aquisição 3D em geral, como no trabalho de Dekiff et al. (2010).

Paralelamente, o estado da arte na correspondência para sistemas estéreo passivos têm sido influenciado por abordagens de aprendizado profundo. Um trabalho seminal nesta área foi o de Zbontar e LeCun (2016), que treinaram redes neurais convolucionais para medir a similaridade entre regiões de imagens, superando métodos clássicos em cenários com textura adequada.

A presente dissertação se posiciona de forma distinta frente a essas abordagens. Ao empregar um sistema de visão estéreo ativa com a projeção de um padrão de pontos pseudoaleatórios variável no tempo, o complexo problema da correspondência que os métodos de aprendizado profundo buscam resolver é contornado de forma determinística.

Inserido diretamente nesta linha de pesquisa, o trabalho de Regner et al. (2025), do qual a autora desta dissertação é coautora, propôs uma abordagem que combina a técnica da triangulação inversa com a correlação espaço-temporal. A principal inovação do referido artigo foi o uso de um projetor de padrão pseudoaleatório variável, que utiliza prismas contrarrotativos para garantir a variação da textura ao longo de uma sequência de imagens. Esse trabalho validou a eficácia do método e seu bom desempenho metrológico, servindo como base conceitual e ponto de partida para a metodologia aprofundada e a análise de erros de medição desenvolvida na presente dissertação.

Em suma, a análise dos trabalhos correlatos, tanto clássicos quanto recentes, revela que, embora a construção de sistemas de digitalização e reconstrução 3D seja um tema consolidado, existem lacunas a serem preenchidas, inserindo assim a presente dissertação.

Neste capítulo, foi apresentada a fundamentação teórica que alicerça o desenvolvimento do sistema proposto. Partindo dos modelos matemáticos do modelo de câmera, foram revisados os conceitos de visão computacional e estereoscopia, com ênfase nas técnicas de reconstrução 3D por luz estruturada. A análise dos trabalhos correlatos permitiu contextualizar a pesquisa no estado da arte, identificando a lacuna e a contribuição específica desta dissertação. Com a base conceitual e contextual estabelecida, o capítulo seguinte se dedicará a descrever em detalhes os materiais e métodos empregados para concepção, implementação e validação do sistema desenvolvido.

3 MATERIAIS E MÉTODOS

Após a apresentação da fundamentação teórica sobre os princípios da visão computacional e da reconstrução 3D no Capítulo 2, este capítulo se dedica a detalhar a metodologia empregada para atingir o objetivo geral desta dissertação: o desenvolvimento e a validação de um sistema de visão estéreo ativa *laser* para reconstrução 3D de cenas. O propósito aqui é descrever de forma pormenorizada todos os procedimentos, equipamentos e algoritmos utilizados, de modo a garantir a clareza e a reprodutibilidade da pesquisa.

O capítulo está estruturado da seguinte forma: inicialmente na seção 3.1 é apresentada a arquitetura geral do sistema, que ilustra a integração entre os componentes de *hardware* e o fluxo de processamento do *software*. A seção 3.2 descreve em detalhes cada um dos componentes de *hardware* selecionados para o projeto, justificando as escolhas técnicas. Em seguida, a seção 3.3 aborda o desenvolvimento do *software* de reconstrução, detalhando as etapas de calibração do sistema, aquisição e processamento de imagens, triangulação dos pontos e reconstrução da superfície 3D. Por fim, a seção 3.4 apresenta a metodologia de validação, delineando os procedimentos experimentais e as métricas definidas para avaliar o desempenho metrológico do sistema desenvolvido.

3.1 ARQUITETURA DO SISTEMA

A arquitetura do sistema desenvolvido nesta dissertação baseia-se no Método da Triangulação Inversa, cujos fundamentos foram apresentados na Seção 2.4. No entanto esta arquitetura se diferencia fundamentalmente da abordagem original, que utiliza mapas de fase, ao propor e implementar uma métrica de correspondência alternativa: a correlação espaço-temporal.

Esta abordagem foi escolhida por ser ideal para a projeção de padrões não determinísticos, como o padrão de pontos pseudoaleatórios variável no tempo utilizado neste trabalho. A lógica desta métrica é ilustrada na Figura 8. O processo baseia-se em discretizar um volume de busca tridimensional em uma grade regular de pontos. De maneira análoga a um *pixel* (elemento de imagem), que representa a menor unidade de uma imagem 2D, cada ponto nesta grade é um *voxel* (elemento de volume).

Em vez de comparar um único valor (como a fase) em cada projeção 2D, a correspondência é avaliada analisando a variação da textura (intensidade) de uma vizinhança de pontos ao longo de uma sequência T de imagens. O processo, que é o núcleo da contribuição desta arquitetura segue os seguintes passos:

1. Para cada *voxel* 3D (X, Y, Z) sendo testado no *grid* de busca o algoritmo considera sua vizinhança espacial na forma de uma janela planar (um *patch* 2D de $N \times N$ pontos), esta janela é definida no espaço do objeto, centrada em (X, Y) e situada no plano de profundidade constante Z ;
2. Este *kernel* 3D (o ponto central e seus vizinhos) é projetado matematicamente para os planos 2D das câmeras esquerda e direita, usando os parâmetros de calibração;
3. A intensidade de cada um dos $N \times N$ pontos projetados é extraída via interpolação ao longo de uma sequência T de imagens que foram capturadas enquanto o padrão *laser* varia;
4. Dois vetores de dados (chamados de assinaturas) de tamanho $(N \times N \times T)$ são gerados, um para cada câmera, e a similaridade entre os dois vetores espaço-temporais é calculada no domínio da imagem, utilizando a métrica ZNCC (equação 10);
5. Para cada posição (X, Y) da grade, o algoritmo avalia os resultados ao longo da varredura no eixo Z e determina a coordenada que maximiza o *score* de correlação como sendo a superfície.

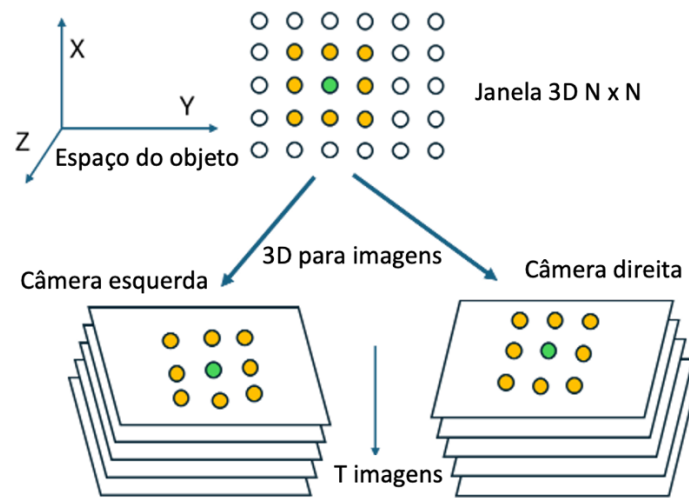


Figura 8 – Correlação espaço-temporal.
 Fonte: Regner et. al., 2025 adaptado pela autora (2025).

A aplicação desta abordagem espaço-temporal ($N \times N \times T$) em conjunto com o Método da Triangulação Inversa, embora robusta, impõe desafios conceituais e computacionais significativos que precisam ser resolvidos. O primeiro é o custo de memória computacional. O *grid* 3D de *voxels* (X, Y, Z) por si só já consome uma quantidade substancial de memória. A métrica espaço-temporal multiplica essa demanda, pois o algoritmo precisa, idealmente, manter todos os dados da T imagens de ambas as câmeras, além dos *kernels* $N \times N$, na memória da GPU para processamento paralelo. O que torna a reconstrução de cenas grandes em um único passo inviável.

Paralelamente a isso a complexidade do algoritmo é alta. O processo envolve múltiplos laços alinhados: iterar sobre o *grid* X - Y , varrer o eixo Z , processar o *kernel* $N \times N$ e analisar a sequência de T imagens, realizando projeções e cálculos ZNCC para cada *voxel* candidato. O desafio mais crítico refere-se à discretização do *grid* 3D. Uma abordagem “simples” utilizando um passo de varredura fixo em milímetros no espaço do objeto, como por exemplo um *voxel* a cada 1 mm nos eixos X, Y e Z , não é tão eficiente.

Em regiões distantes, um passo de 1 mm no objeto pode ser projetado em um deslocamento de 0,2 *pixel* na imagem, o que gera uma superamostragem, em que vários *voxels* 3D são projetados no mesmo *pixel*, “desperdiçando” processamento. Já em regiões próximas, esse mesmo passo de 1 mm pode ser

projetado em um deslocamento de 3 *pixels* na imagem, o que gera uma subamostragem, em que o algoritmo “pula” *pixels* da imagem, perdendo resolução e detalhes da textura.

Para contornar essas limitações, a arquitetura foi projetada com duas estratégias de otimização principais:

1. Processamento em lotes (*tiling*): para resolver os desafios de custo de memória e computacional foi implementada uma estratégia de processamento em lotes. O volume de interesse global é sistematicamente dividido em sub-volumes menores (*tiles*) com uma área de sobreposição definida para garantir a continuidade da reconstrução. Cada *tile* é processado de forma independente, permitindo a reconstrução de áreas arbitrariamente grandes com uso de memória controlada;
2. Discretização adaptativa do *grid*: para resolver o desafio da distorção de amostragem, a segunda e mais relevante otimização foi implementada. Ao invés de utilizar um passo de varredura fixo em milímetros, o algoritmo calcula um passo adaptativo para cada *tile*. O sistema determina a sensibilidade do *grid* no centro do *tile* e ajusta os passos nos eixos X, Y e Z para que correspondam a um deslocamento alvo no plano da imagem, como por exemplo 0,8 *pixel*. Essa abordagem garante que a densidade de amostragem seja consistente do ponto de vista do sensor (próximo de 1 *voxel* por *pixel*), otimizando a robustez da correlação em todo o campo de visão.

A arquitetura computacional foi projetada para suportar a alta demanda de processamento deste método, sendo o uso de uma unidade de processamento gráfico (GPU) com suporte a CUDA um pilar central. O *software*, desenvolvido em Python com a biblioteca PyTorch, explora o paralelismo massivo da GPU para viabilizar os cálculos de correlação e da métrica de qualidade PNR em um tempo de execução factível. Os diagramas de blocos das Figuras 9 e 10 ilustram a integração geral entre os componentes de *hardware* e o fluxo de processamento de dados do *software*.

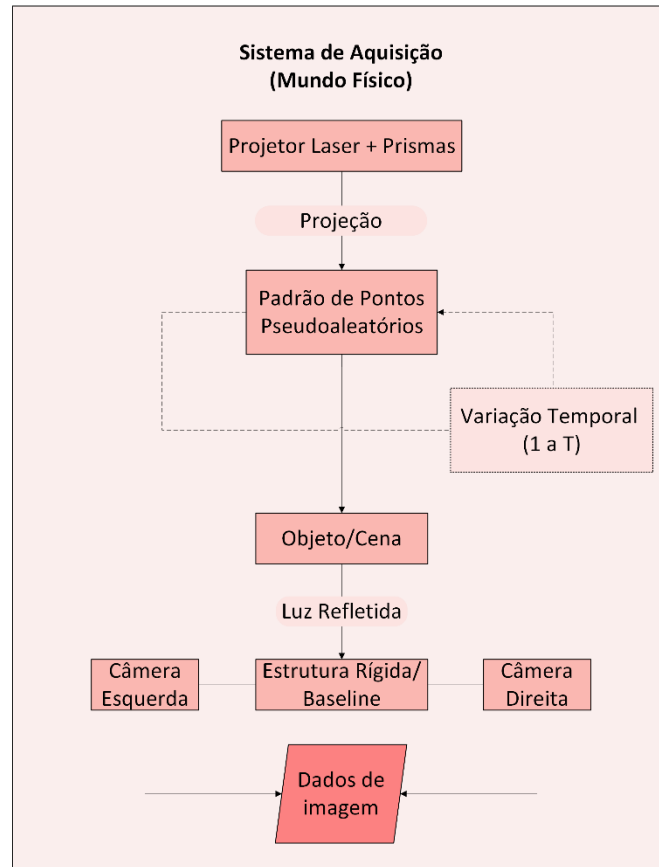


Figura 9 – Diagrama de blocos do sistema de aquisição.
Fonte: A Autora (2025).

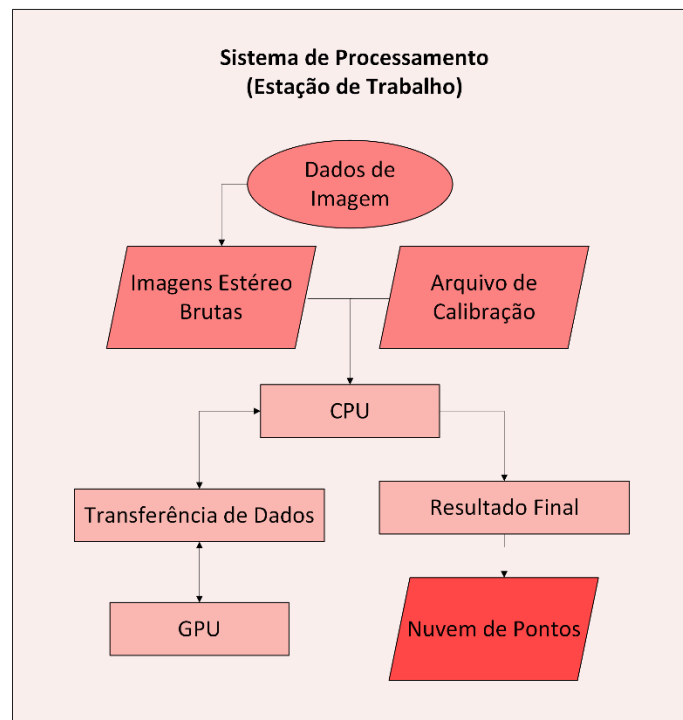


Figura 10 – Diagrama de blocos do sistema de processamento.
Fonte: A Autora (2025).

O fluxo de processamento dos dados, esquematizado na Figura 11, reflete a lógica da arquitetura otimizada. A execução não ocorre em um único volume monolítico, mas sim em um laço que itera sobre múltiplos lotes (*tiles*). Que subdividem a área de interesse global.

Para cada *tile*, o processo se inicia com o cálculo de um *grid* 3D adaptativo, cujos passos são ajustados para garantir uma amostragem consistente no espaço da imagem. Com o volume de busca local definido, o núcleo do algoritmo de correlação espaço-temporal é executado: cada ponto do *grid* é projetado nos planos das imagens (cujos dados já foram pré-carregados e pré-processados), vetores espaço-temporais ($N \times N \times T$) são extraídos por interpolação e a similaridade fotométrica é quantificada via ZNCC.

A nuvem de pontos bruta do *tile* é formada pelos pontos que maximizam a correlação em cada coordenada (X, Y), com a coordenada Z sendo refinada por interpolação para precisão sub-voxel. Finalmente, a nuvem de pontos do tile passa por um *pipeline* de filtragem para remover pontos de baixa confiança e *outliers* geométricos, e o resultado final daquele lote é salvo em disco.

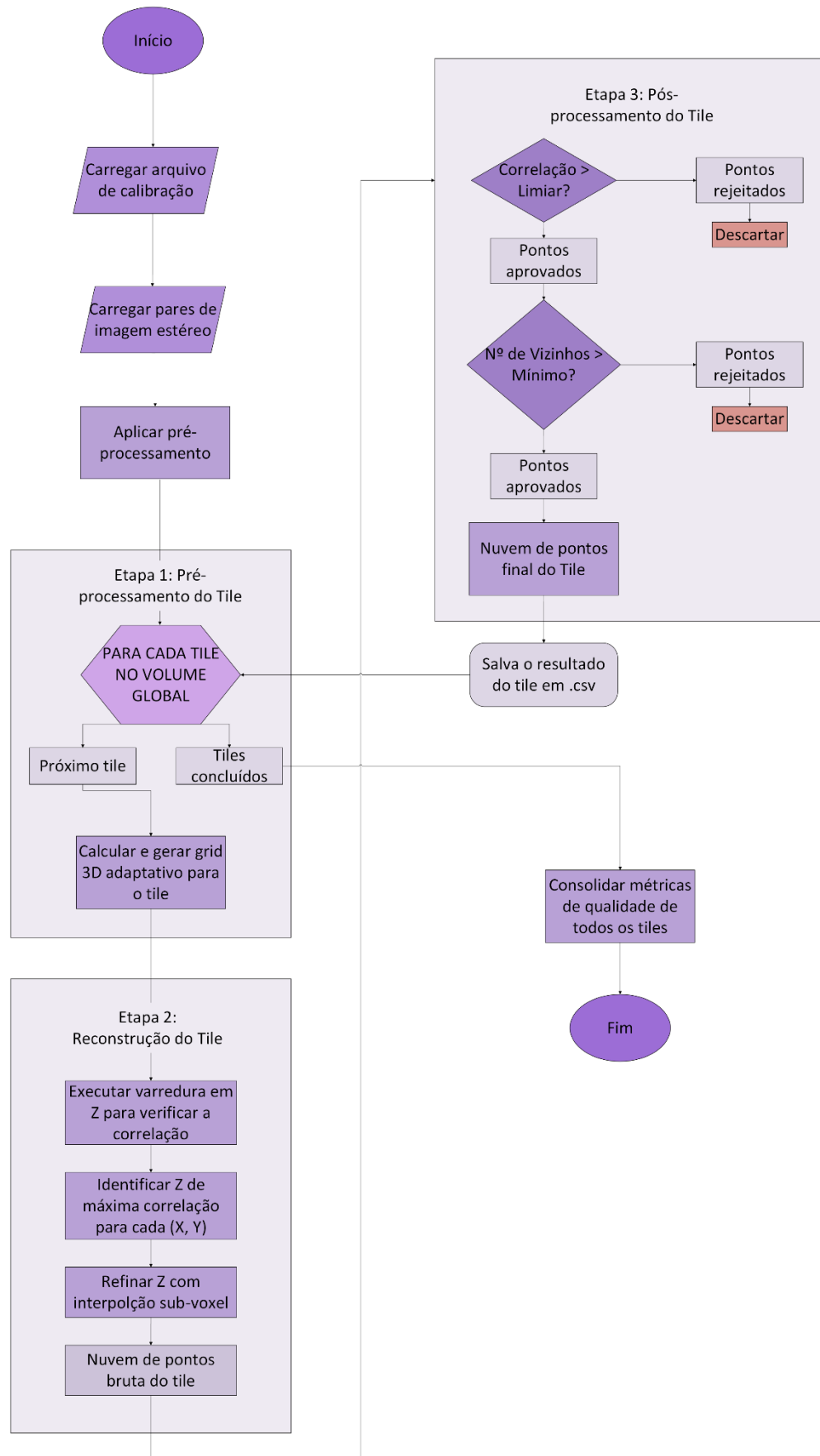


Figura 11 – Fluxograma do processo de reconstrução 3D.
Fonte: A Autora (2025).

3.2 COMPONENTES DE *HARDWARE*

A seleção dos componentes de *hardware* foi uma etapa crucial do projeto, guiada pelos objetivos de se alcançar um bom desempenho metrológico e garantir a adequação do sistema para a aplicação final no Projeto VORIS. Nesta seção, são detalhados os principais componentes que constituem o sistema de aquisição de dados: o par de câmeras e lentes, o sistema de projeção de luz estruturada e a estrutura de montagem.

O coração do sistema estéreo é o par de câmeras responsável por capturar as imagens da cena. Para este trabalho, de acordo com a disponibilidade do Projeto VORIS, foram selecionadas duas câmeras industriais do modelo Backfly S USB3 (BFS-U3-51S5M-C), da fabricante Teledyne FLIR.

Este modelo foi escolhido por sua alta resolução e por ter características ideais para aplicações de visão computacional. As câmeras são equipadas com um sensor monocromático CMOS Sony IMX250 de 5.0 *megapixels* (resolução de 2445 x 2048 *pixels*), capaz de fornecer imagens de alta resolução, o que é fundamental para detecção dos pontos de laser. A interface USB 3.0 garante a alta taxa de transferência de dados necessária para a operação, suportando uma taxa de captura de até 75 Hz em resolução máxima.

Acoplada a cada câmera, foi utilizada uma lente de distância focal fixa de 12 mm, modelo 33-303 UC, da fabricante Edmund Optics. Para otimizar a detecção do *laser* e minimizar a interferência da luz ambiente, cada lente foi equipada com um filtro óptico passa-banda com comprimento de onda central de 465 nm e largura de banda de 70 nm.

Esta configuração é fundamental para a aplicação, pois o filtro bloqueia seletivamente a iluminação indesejada, permitindo que apenas o comprimento de onda específico do *laser* atinja o sensor da câmera. A combinação de lentes de distância focal fixa, que minimiza a distorção, com um filtro passa-banda, que isola o sinal do *laser*, visa otimizar a detecção do padrão projetado, resultando em uma menor interferência na calibração e nas medições.

Conforme foi discutido na fundamentação teórica, o uso de um método ativo de medição simplifica o problema da correspondência em superfícies com pouca textura. O sistema de projeção de luz estruturada, da fabricante Osela, é peça central da abordagem ativa. Ele utiliza um *laser* de luz azul com comprimento de

onda de 450 nm e um elemento óptico difrativo (DOE) para gerar uma matriz com cerca de 40.000 pontos num campo de visão de $30^\circ \times 30^\circ$. Essa abordagem baseada em DOE para sensoriamento 3D preciso é explorada em trabalhos como o de Song et al. (2019).

De acordo com as especificações técnicas fornecidas pela fabricante, o componente emprega a tecnologia *Random Pattern Projector (RPP)*, apresentando uma razão entre *pixels* iluminados e escuros de aproximadamente 15%. A matriz é construída de forma que cada ponto brilhante seja cercado por pelo menos 8 *pixels* escuros, garantindo o isolamento dos pontos e facilitando a segmentação. A distribuição é pseudoaleatória, porém determinística e simétrica em relação ao centro. Adicionalmente, essa tecnologia permite a formação de mosaicos (*tiling*) com múltiplos projetores para cobrir áreas maiores mantendo a continuidade do padrão.

Sua utilização combinada com o conjunto de prismas contrarrotativos, também conhecido como par de Risley, cujo princípio de funcionamento é ilustrado na Figura 12, permite uma variação dinâmica da posição do padrão projetado, o que é fundamental para o algoritmo de correlação espaço-temporal. O mecanismo baseia-se na refração do feixe óptico ao atravessar dois prismas em forma de cunha. Individualmente, cada prisma desvia a luz em um ângulo fixo, contudo, ao realizar a rotação de um em relação ao outro os vetores de desvio se combinam, permitindo direcionar o padrão para diferentes posições dentro de um cone de varredura, sem a necessidade de mover o projetor.

Além disso, o fato de os prismas serem acionados por um motor de passo, faz com que o deslocamento do padrão seja preciso e controlado, garantindo a melhor qualidade dos resultados.

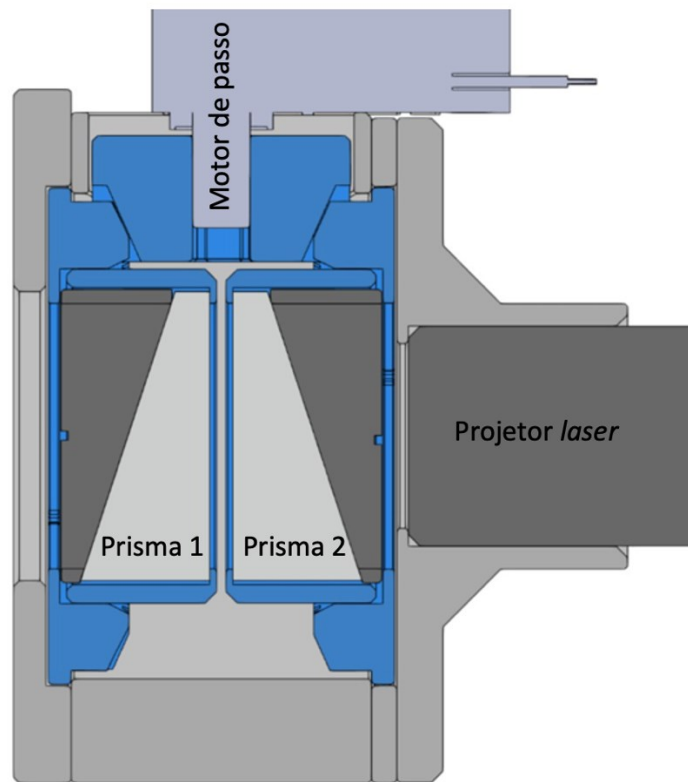


Figura 12 – Sistema de projeção.

Fonte: Regner et. al., 2025 adaptado pela autora (2025).

A preferência pela luz azul é justificada pelo fato de que o comprimento de onda mais curto permite gerar projeções mais finas e nítidas, além de sofrer menos dispersão superficial, resultando em medições com menor ruído e maior precisão. Adicionalmente, a escolha do comprimento de onda azul visa à futura integração do sistema ao sistema do Projeto VORIS, pois esta cor possui melhor propagação em sistemas subaquáticos.

Para garantir a estabilidade da geometria estéreo, os componentes foram fixados em uma estrutura de montagem rígida, composta por componentes obtidos através de impressão 3D para fixação em uma barra de metal própria para manter a posição e a orientação relativa entre as câmeras e o projetor constantes durante a operação, conforme modelo apresentado na Figura 13.

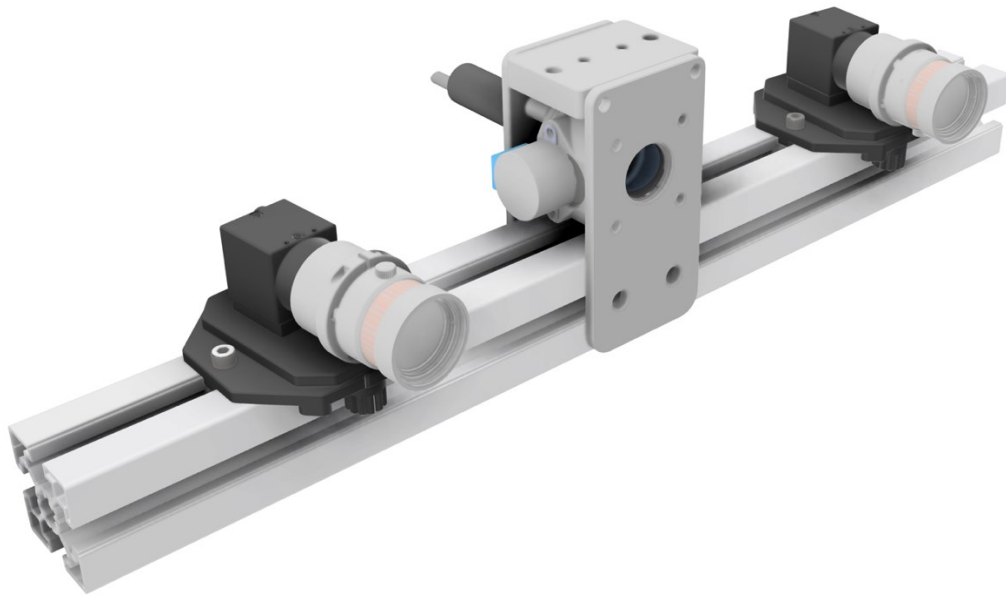


Figura 13 – Modelo 3D da estrutura de fixação.
Fonte: A Autora (2025).

É importante destacar que a montagem das câmeras na estrutura não segue a configuração estéreo canônica, onde os eixos ópticos permanecem paralelos. Em vez disso, optou-se por uma configuração em que as câmeras foram fixadas com uma rotação angular intencional em torno do eixo vertical, de modo que seus eixos ópticos converjam para o centro do volume de medição. Essa disposição geométrica foi adotada para maximizar a área de sobreposição entre os campos de visão esquerdo e direito, otimizando o volume útil onde a triangulação é possível.

A distância entre os centros ópticos das câmeras (*baseline*), de aproximadamente 420 mm, foi definida de modo a otimizar a relação entre o campo de visão e a resolução em profundidade para a distância de trabalho planejada. O sistema também incorpora componentes para controle e automação, incluindo o motor de passo para a movimentação dos prismas e uma placa controladora compatível com Arduino, responsável pela interface de comunicação do sistema com o computador. A montagem final do aparato experimental é apresentada na Figura 14.

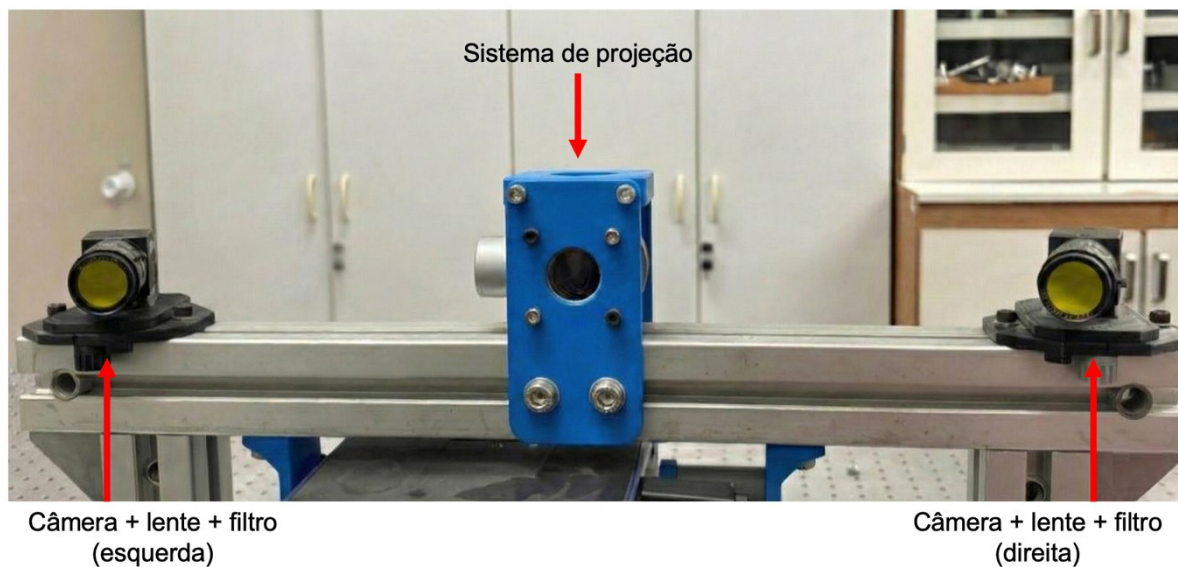


Figura 14 – Montagem final.
Fonte: A Autora (2025).

3.2.1 Eletrônica de controle e sincronização

A integração entre os sistemas óptico, mecânico e de aquisição/processamento é realizada por uma unidade de controle dedicada, baseada na plataforma NVIDIA Jetson Orin Nano (8GB). Esta unidade coordena a placa de interface eletrônica (Figura 15), que gerencia o acionamento do *laser*, a movimentação dos prismas e o sincronismo das câmeras.

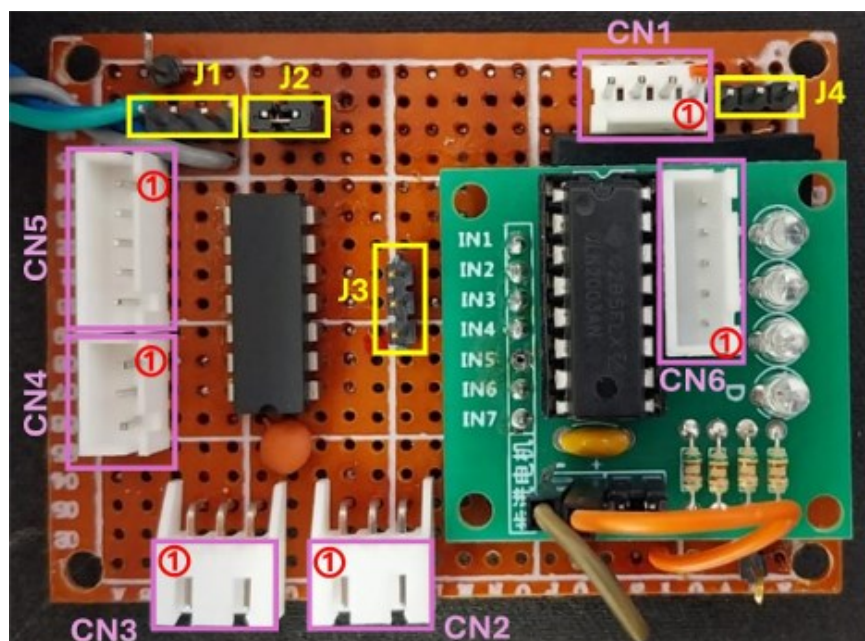


Figura 15 – Placa de interface eletrônica.
Fonte: A Autora (2025).

O coração da sincronização de *hardware* é um circuito integrado Schmitt Trigger, que recebe sinais digitais da Jetson Orin e os distribui para as câmeras através dos conectores CN2 e CN3. Este método garante que o *trigger* de aquisição seja simultâneo para ambos os sensores, eliminando atrasos inerentes a comandos via *software*. A placa foi projetada para centralizar todas as conexões do sistema, conforme detalhado abaixo:

1. Gerenciamento do *laser*: a conexão com a fonte *laser* é feita via CN1. O sinal de controle vindo da Jetson (CN4) opera de forma invertida: um nível lógico baixo (0 V) liga o *laser*, enquanto um nível alto (3,3 V) o desliga;
2. Controle do motor de passo: através do conector CN5, a Jetson envia sinais para quatro canais digitais que, após amplificação, acionam o motor de passo via CN6. O motor opera no regime *Stop-and-Go*, permitindo a estabilização do padrão de pontos antes de cada captura;
3. Ajustes via *jumper*s: a placa permite configurações físicas rápidas, como a seleção da tensão de alimentação do *laser* entre 5V e 12,5 V (J1/J2), a habilitação do nível lógico de ativação do *laser* (J3) e o controle de potência da fonte *laser* (J4).

3.3 DESENVOLVIMENTO DO SOFTWARE DE RECONSTRUÇÃO

O *software* de reconstrução, responsável por processar as imagens pré-capturadas e gerar a nuvem de pontos 3D, foi integralmente desenvolvido em linguagem Python. A escolha da plataforma foi motivada pela necessidade de executar um algoritmo de correlação computacionalmente intensivo, o que demandou o uso da biblioteca PyTorch para aproveitar a capacidade de processamento paralelo de unidades de processamento gráfico (GPU).

Dentre as principais bibliotecas utilizadas, destacam-se:

- PyTorch: conforme mencionada anteriormente, é utilizada para implementar o núcleo de correlação na GPU;
- OpenCV: utilizado para as tarefas de pré-processamento de imagem;
- NumPy: utilizada para operações matriciais gerais;
- SciPy: utilizada para a implementação do filtro espacial de densidade.

Além disso, foram utilizados os códigos para comunicação com os componentes disponibilizados pelas fabricantes, todos os arquivos de acesso público estão disponíveis no seguinte Git: [Projeto-Voris/stereo_zscan](https://github.com/Projeto-Voris/stereo_zscan). O processo de reconstrução implementado pode ser decomposto em quatro etapas principais:

1. Calibração do sistema estéreo;
2. Aquisição de imagens e pré-processamento;
3. Triangulação e geração da nuvem de pontos;
4. Pós-processamento da nuvem de pontos e reconstrução da malha em software comercial.

Tanto o desenvolvimento do código quanto o processamento das reconstruções foi feito em computador equipado com um processador Intel Core i7-14700F, uma placa de vídeo NVIDIA GeForce RTX 4060 Ti (16 GB), modelo XLR8 da fabricante PNY, e 32 GB de memória RAM DDR5.

3.3.1 Calibração do sistema estéreo

Conforme estabelecido na seção 2.3.2, a calibração é um pré-requisito indispensável para a reconstrução 3D métrica. Para este trabalho, o processo foi realizado utilizando a aplicação *Stereo Camera Calibrator*, parte da *Image Processing and Computer Vision Toolbox* do MATLAB. Esta ferramenta oferece um ambiente interativo para estimar os parâmetros da câmera a partir de um conjunto de imagens de calibração, implementando o método proposto por Zhang, Z. (2000).

O procedimento foi executado da seguinte forma:

1. Aquisição de imagens: foram capturados aproximadamente 60 pares de imagens estéreo de um padrão de tabuleiro de xadrez de 16 x 11 quadrados, sendo que o lado de cada foi medido e definido como 25 mm. O padrão foi posicionado em diferentes orientações em relação ao par de câmeras, buscando cobrir todo o campo de visão;
2. Processamento na *toolbox*: os pares de imagens da câmera esquerda e direita foram carregados na aplicação, o *software* detectou automaticamente os cantos do tabuleiro de xadrez em cada imagem e descartou os pares considerados de má qualidade;

3. Estimação dos parâmetros: a toolbox utilizou as correspondências entre os pontos 3D conhecidos do padrão e suas projeções 2D nas imagens para estimar, em um primeiro passo, os parâmetros intrínsecos de cada câmera de forma independente. O modelo de câmera considerou um conjunto de cinco coeficientes de distorção, três para a distorção radial (k_1 , k_2 e k_3) e dois para a distorção tangencial (p_1 e p_2);
4. Calibração estéreo: em um segundo passo, a ferramenta calculou os parâmetros extrínsecos do par estéreo, ou seja, a matriz de rotação R e o vetor de translação (t) que descrevem a transformação de corpo rígido entre o sistema de coordenadas da câmera direita e o da esquerda.

A qualidade da calibração foi avaliada iterativamente através do erro médio de reprojeção, que a ferramenta exibe graficamente. Pares de imagens com alto erro de reprojeção foram inspecionados e removidos para refinar o resultado final, que culminou em um erro médio de aproximadamente 0,4 *pixels*. Ao final do processo, os parâmetros foram exportados como um objeto *stereoParameters* do MATLAB, contendo todas as informações necessárias (matrizes intrínsecas, coeficientes de distorção, posição relativa entre as câmeras) que foram salvas em um arquivo de formato “yaml” para serem utilizadas pelo *software* de reconstrução.

3.3.2 Aquisição e pré-processamento de imagens

A primeira etapa do processo de medição consiste na aquisição de uma quantidade T de pares de imagens estéreo da cena, devidamente sincronizadas. A sincronização é fundamental para garantir que ambas as câmeras capturem o objeto no mesmo instante, o que é especialmente crítico para cenas dinâmicas. A captura de 10 pares de imagens sincronizadas, por exemplo, leva cerca de 10 segundos. Embora projetado para futuramente funcionar com o ROV em movimento, os testes experimentais dessa pesquisa foram realizados com o sistema parado, garantindo a estabilidade necessária para a validação inicial do método de correlação espaço-temporal.

Antes de iniciar a reconstrução 3D, as imagens brutas capturadas pelas câmeras passam por uma etapa de pré-processamento. Esta etapa consiste primariamente na correção da distorção da lente, utilizando os coeficientes calibrados visando a linearidade geométrica das projeções. Adicionalmente, nesta fase é definido e gerado o volume de busca tridimensional (*grid*), que delimita a região do espaço onde a correlação será calculada, dispensando a necessidade de detecção prévia de características nas imagens.

Para a correção da distorção introduzida pelo conjunto óptico, é implementada a função `convert_images`. Nela, cada imagem é submetida à remoção da distorção através da função `cv2.undistort`, utilizando os coeficientes de calibração obtidos previamente para corrigir as distorções radiais e tangenciais. A criação do *grid* volumétrico corresponde ao núcleo da metodologia, que consiste em discretizar um volume de busca tridimensional em uma grade regular de pontos. Este volume, definido por limites nos eixos X, Y e Z, encapsula a região do espaço onde se espera que a superfície do objeto esteja. A nuvem de pontos final será um subconjunto dos pontos (ou voxels) desta grade. A resolução da reconstrução é diretamente controlada pela densidade de pontos (passos `xy_step` e `z_step`).

3.3.3 Triangulação e geração da nuvem de pontos

Diferente das abordagens clássicas que extraem características 2D, determinam a correspondência no plano imagem e depois realizam a triangulação para determinação do ponto 3D, o método implementado opera de forma inversa: ele testa cada ponto do *grid* 3D como uma hipótese de pertencer à superfície do objeto. Esta abordagem, conforme definido na Seção 3.1, é executada para cada *voxel* 3D do *grid* de busca.

O processo para encontrar a coordenada Z de uma determinada coluna (X, Y) do *grid* segue os seguintes passos:

1. Projeção do *kernel* 3D: para cada *voxel* da grade (X, Y) na profundidade Z sendo testada, define-se uma janela plana virtual de $N \times N$ pontos, centrada em (X, Y) e situada no mesmo plano de profundidade Z. Este *kernel* planar é então projetado matematicamente nos planos de imagem 2D das câmeras esquerda e

direita utilizando as matrizes de projeção da calibração. Esta etapa é realizada pela função *transform_gcs2ccs*;

2. Extração da assinatura espaço-temporal: o resultado da projeção é um conjunto de $N \times N$ coordenadas 2D para cada câmera, a intensidade em cada uma dessas coordenadas 2D projetadas é extraída por interpolação bilinear (*torch.nn.functional.grid.sample*). Esse processo é repetido para toda a sequência de T imagens capturadas e o resultado são dois vetores de dados (a “assinatura”) de tamanho $(N \times N \times T)$, um para cada câmera;
3. Cálculo da similaridade: a correspondência entre as duas assinaturas espaço-temporais (esquerda e direita) para um mesmo ponto 3D é quantificada pela métrica de correlação cruzada normalizada a zero (ZNCC), implementada na função *zncc_correlation*. Este cálculo é repetido para cada *voxel* ao longo do eixo Z daquela coluna (X, Y).

A etapa final de reconstrução é realizada em duas fases, para maximizar a precisão. Primeiramente, para cada coluna (X, Y) do *grid*, o algoritmo “varre” todas as profundidades (eixo Z), calculando o *score* ZNCC em cada *voxel*. O *voxel* do *grid* que apresentar o valor máximo de correlação é identificado como a correspondência inicial. Contudo, se o processo fosse encerrado neste ponto, a resolução de profundidade do sistema estaria limitada pela discretização do *grid* (*z_step*).

Para que a resolução da medição não seja limitada pelo passo do *grid*, uma segunda fase de refinamento sub-*voxel* é executada, que se refere a uma técnica que permite estimar uma posição com precisão mais fina do que a própria grade de *voxels*. O método implementado assume que a curva de correlação em torno de seu valor máximo se comporta de forma contínua e suave, podendo ser aproximada por um polinômio. Para encontrar o “verdadeiro” pico de forma computacionalmente eficiente, o algoritmo utiliza uma interpolação parabólica.

O algoritmo utiliza o valor de correlação do pico e os valores de seus dois vizinhos imediatos ao longo do eixo Z para ajustar a parábola. O vértice desta parábola, que representa a estimativa do pico de correlação no espaço contínuo, é então calculado analiticamente. A coordenada Z deste vértice é adotada como a profundidade final e com a maior precisão para o ponto (X, Y) da superfície. A união de todos esses pontos refinados forma a nuvem de pontos bruta.

3.3.4 Pós-processamento

A nuvem de pontos bruta, gerada pelo processo de correlação, contém a geometria da superfície, mas também pode incluir pontos espúrios resultantes de oclusões, ambiguidades de correspondência ou áreas classificadas como de “baixa textura”. No contexto de um sistema de visão estéreo ativa, onde a textura é projetada artificialmente, o termo “baixa textura” refere-se a condições ópticas locais que degradam a qualidade ou o contraste do padrão *laser* na imagem capturada, inviabilizando a correlação robusta.

As principais causas físicas para essa degradação incluem a reflexão especular, onde o feixe incidente é refletido para fora da lente ou diretamente para o sensor, a absorção elevada, que atenuam o sinal da luz reduzindo a relação sinal-ruído e a desfocagem, que ocorre quando regiões do objeto se encontram nos limites da profundidade do campo, resultando em um padrão de pontos borrado e com perda de conteúdo. Para mitigar esses erros e garantir a confiabilidade dos dados métricos, uma sequência de filtragem em duas etapas é aplicada para refinar o resultado.

O primeiro passo consiste em uma filtragem baseada na confiança da correspondência fotométrica. Cada ponto da nuvem de ponto bruta possui um *score* de correlação associado. Um limiar de correlação, definido no código como *CORR_THRESHOLD*, é aplicado, e todos os pontos cujo *score* de correlação seja inferior a este limiar são descartados. Esta etapa remove eficazmente os pontos gerados em regiões onde a correspondência entre as imagens esquerda e direita foi considerada fraca ou não confiável.

Após a filtragem por correlação, a nuvem de pontos ainda pode conter *outliers* geométricos – pontos isolados que não pertencem à superfície contínua do objeto. Para removê-los, é aplicado um filtro espacial de densidade, conforme implementado na função *filter_sparse_points*. O algoritmo opera da seguinte forma: para cada ponto ele realiza uma busca espacial e conta quantos pontos vizinhos existem dentro de um raio (*SPATIAL_FILTER_RADIUS*) pré-definido. Se o número de vizinhos for menor que um limiar mínimo (*SPATIAL_FILTER_MIN_NEIGHBORS*), o ponto é classificado como um *outlier* e eliminado da nuvem de pontos. Este método de busca por vizinhos é executado de forma computacionalmente eficiente através do uso de uma estrutura de dados *k-d tree* (*scipy.spatial.cKDTree*).

O resultado deste *pipeline* de processamento e filtragem é uma nuvem de pontos final, densa e limpa, que representa a geometria do objeto digitalizado com melhor fidelidade. Os dados (coordenadas X, Y e Z e o *score* de correlação de cada ponto) são então salvos em um arquivo de formato CSV, permitindo que posteriores reconstruções de malha triangular e análises dimensionais sejam realizadas em *softwares* comerciais de metrologia.

3.4 METODOLOGIA DE VALIDAÇÃO

Após a descrição da arquitetura do sistema e do *software* de reconstrução, este capítulo conclui com a apresentação da metodologia de validação. O objetivo deste procedimento é avaliar quantitativamente o desempenho do sistema de digitalização 3D desenvolvido. A metodologia foi dividida em duas etapas complementares: primeiramente, uma avaliação da confiabilidade de cada ponto individualmente, baseada na qualidade da correlação estéreo; e em segundo lugar, uma avaliação do desempenho metrológico global do sistema, realizada através da medição de artefatos calibrados ou com geometria conhecida.

3.4.1 Métrica de qualidade da correlação: Relação Pico-Ruído (PNR)

O valor da correlação cruzada normalizada (ZNCC), embora seja um indicador primário, pode ser insuficiente para garantir uma correspondência inequívoca, especialmente em regiões de baixa textura. Um valor de correlação acima do limiar pode ser obtido mesmo quando múltiplos picos de correlação de magnitude semelhante existem ao longo do eixo de busca (eixo Z), indicando uma correspondência ambígua.

Visando mitigar essa ambiguidade e avaliar a qualidade de cada ponto, foi adotada a métrica da relação pico-ruído, do inglês *peak-to-noise ratio* (PNR). Esta métrica não avalia apenas o valor absoluto do pico de correlação, mas sim a sua proeminência em relação à média dos demais valores de correlação ao longo do mesmo eixo de busca. Um PNR elevado significa que o pico de correlação é distinto e se sobressai do ruído de fundo, indicando uma correspondência 3D de alta confiança.

A PNR para cada ponto (x, y) da grade de busca foi calculada conforme a equação 11:

$$PNR = \frac{C_{max}}{\frac{1}{N-1}(\sum_{i=1}^N C_i - C_{max})} \text{ (Equação 11)}$$

Onde:

- C_{max} é o valor máximo de correlação (o pico) encontrado ao longo do eixo Z;
- C_i é o valor de correlação em cada um dos N pontos de busca ao longo do eixo Z;
- O denominador representa a média de todos os outros valores de correlação, excluindo-se o pico.

O cálculo da PNR foi integrado ao *pipeline* de *software*, sendo executado para cada ponto 3D candidato. Para a avaliação do desempenho global do sistema, a distribuição dos valores de PNR de todos os pontos válidos foi visualizada através de um histograma consolidado. Durante a análise, observou-se que a presença de uma pequena quantidade de *outliers* com valores extremamente negativos tornava a média aritmética uma métrica de tendência central não representativa do desempenho típico do sistema. Por essa razão, optou-se pela mediana dos valores de PNR como indicador estatístico principal. Sendo uma medida de estatística robusta, a mediana não é influenciada pela magnitude de valores espúrios, fornecendo uma avaliação metrologicamente mais fidedigna da qualidade de correspondência alcançada para a maioria dos pontos da nuvem 3D.

3.4.2 Avaliação do desempenho metrológico

A avaliação do desempenho metrológico visa determinar quantitativamente a capacidade do sistema de medição. Para tal, foram definidos um conjunto de experimentos baseados na medição de objetos de referência com geometria e dimensões conhecidas. A análise dos resultados permitirá quantificar o erro e a precisão do sistema.

Os artefatos de validação e seus respectivos objetivos de medição são:

- Superfície plana: um plano de referência foi utilizado para avaliar o erro de planeza nas medições do sistema. A análise consiste em ajustar um plano por mínimos quadrados à nuvem de pontos medida e calcular os desvios de todos os pontos em relação a este plano. A qualidade da reconstrução é quantificada tanto pela planeza (*flatness*), definida como a distância pico-vale dos desvios em relação ao plano ajustado por mínimos quadrados, quanto pelo desvio padrão dos resíduos, uma métrica estatística que oferece uma avaliação menos sensível a eventuais pontos espúrios (*outliers*) na nuvem;
- Esferas: um par de esferas de referência foi utilizado para determinar o desempenho do sistema em medir dimensões e distâncias. A partir da nuvem de pontos, são avaliados o erro de apalpação (*probing error*), que corresponde ao erro de forma na superfície da esfera, e o erro de escala (*scale error*), calculado a partir da medição do diâmetro de cada esfera e da distância entre seus centros;
- Objetos de forma livre: artefatos de geometria complexa foram utilizados para uma avaliação qualitativa da reconstrução. O objetivo destes ensaios é verificar a robustez do algoritmo de correlação em situações de oclusão, variações abruptas de profundidade e diferentes condições de reflectância superficial.

Para garantir a consistência e a validade estatística dos resultados, todos os ensaios de validação seguiram um procedimento experimental padronizado, realizado sob condições ambientais controladas, para minimizar fontes de luz e temperatura externas. O procedimento é descrito a seguir:

1. O objeto de referência é posicionado na área de trabalho do sistema, a uma distância representativa da aplicação final;
2. Para cada objeto, é realizado um conjunto de aquisição. Este conjunto consiste na captura de uma sequência de T pares de imagens, capturadas enquanto o padrão de projeção *laser* é variado dinamicamente pelos prismas. Esta variação do padrão ao longo da sequência é o que garante a robustez da correlação espaço-temporal;
3. Cada conjunto T de aquisições gera um conjunto de nuvens de pontos, que é salvo para análise posterior.

Após a aquisição, um algoritmo de ajuste de mínimos quadrados é aplicado a cada conjunto de nuvem de pontos para determinar a primitiva geométrica que melhor a representa (plano ou esfera). A partir da geometria ajustada, os parâmetros dimensionais de interesse e os desvios de forma são calculados. Com base nesses dados, as seguintes métricas de desempenho são utilizadas para caracterizar o sistema:

- Erro de medição: define a proximidade da média das medições em relação ao valor verdadeiro do mensurado. Para o diâmetro da esfera, por exemplo, o erro (E) é calculado como a diferença entre o diâmetro médio medido (D_{medido}) e o diâmetro nominal calibrado ($D_{nominal}$):

$$E = D_{medido} - D_{nominal} \text{ (Equação 12)}$$

- Erro de forma: mede o quão bem a nuvem de pontos adquirida se ajusta à sua forma geométrica ideal. Para o plano, é calculado como o desvio padrão (σ_{plano}) das P distâncias ortogonais (d_i) de todos os pontos da nuvem em relação ao plano ajustado por mínimos quadrados. Para a esfera, o erro é avaliado de forma análoga, como a variação dos raios medidos (desvio padrão dos raios) em relação ao centro da esfera ajustada. Nos ensaios que envolvem a medição simultânea de múltiplos artefatos (como o par de esferas), a métrica de desempenho reportada para o sistema corresponde à média aritmética dos desvios padrão individuais calculados para cada objeto.

$$\sigma_{plano} = \sqrt{\frac{\sum_{i=1}^P (d_i - \bar{d})^2}{P-1}} \text{ (Equação 13)}$$

Os resultados detalhados a partir da aplicação desta metodologia são apresentados e discutidos no Capítulo 4.

4 RESULTADOS E DISCUSSÃO

Este capítulo apresenta os resultados experimentais obtidos com o sistema de visão estéreo ativa desenvolvido, seguidos de uma análise quantitativa e discussão sobre o seu desempenho. A avaliação foi conduzida com base na metodologia de validação descrita na Seção 3.4.

4.1 RESULTADOS DA CALIBRAÇÃO

A calibração do sistema estéreo é um pré-requisito fundamental para a reconstrução 3D métrica, pois determina os parâmetros geométricos e ópticos do conjunto. Conforme descrito na metodologia (Seção 3.3.1), o processo foi realizado utilizando a aplicação *Stereo Camera Calibrator* do MATLAB, a partir da aquisição de 60 pares de imagens de um padrão de tabuleiro de xadrez posicionado em distâncias entre 700 mm e 1000 mm do sistema. A Figura 16 apresenta um exemplo de par de imagens utilizado.

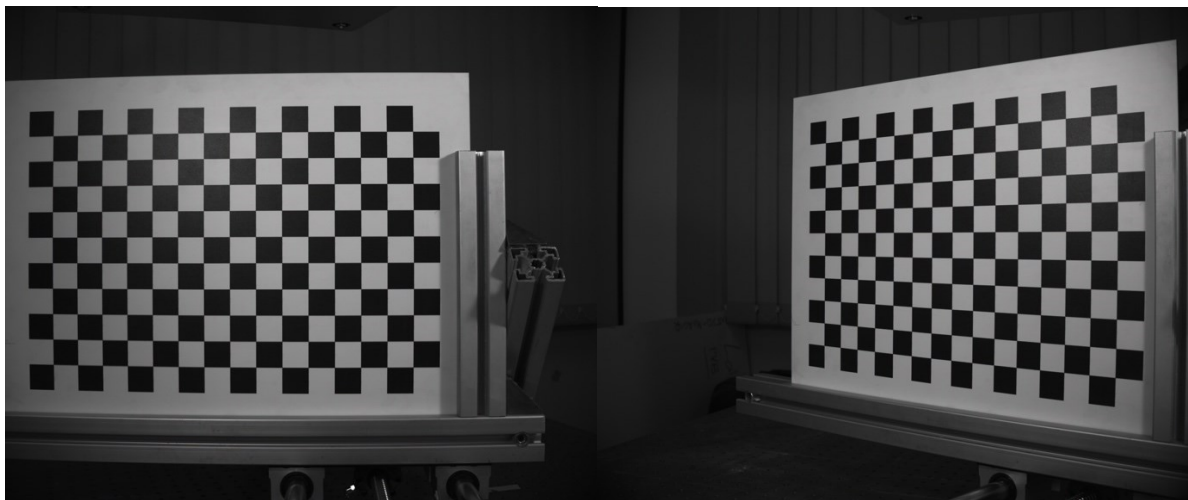


Figura 16 – Exemplo de par de imagens do padrão de calibração.
Fonte: A Autora (2025).

A qualidade do ajuste foi avaliada pelo erro médio de reprojeção, que representa a distância média em *pixels* entre os cantos detectados nas imagens e os pontos do padrão projetados utilizando os parâmetros calculados. O processo resultou em um erro médio de reprojeção de aproximadamente 0,4 *pixels*, indicando um ajuste consistente do modelo geométrico da câmera aos dados observados.

Os parâmetros intrínsecos, que descrevem as características internas de cada conjunto câmera-lente, são apresentados na Tabela 1.

Tabela 1 - Parâmetros intrínsecos do sistema estéreo

Parâmetro	Câmera Esquerda	Câmera Direita
Matriz da Câmera (K)	$\begin{bmatrix} 3501,6 & 0 & 1219,0 \\ 0 & 3497,0 & 1000,7 \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 3514,7 & 0 & 1201,8 \\ 0 & 3508,3 & 960,6 \\ 0 & 0 & 1 \end{bmatrix}$
Distorção Radial (k_1 ; k_2 ; k_3)	-0,2408; 0,1362; 0,3490	-0,2341; 0,0718; 0,6035
Distorção Tangencial (p_1 ; p_2)	0,0008; -0,0003	-0,0016; 0,0004

Fonte: A Autora (2025).

Os parâmetros extrínsecos, que definem a relação de posição e orientação (transformação de corpo rígido) entre as duas câmeras, são detalhados na Tabela 2. O vetor de translação (t) é de particular importância, pois define a linha de base (*baseline*) do sistema.

Tabela 2 - Parâmetros extrínsecos do par estéreo (câmera direita em relação à esquerda).

Parâmetro	Valor
Matriz de Rotação (R)	$\begin{bmatrix} 0,7763 & -0,0064 & 0,6304 \\ 0,0029 & 1,0000 & 0,0066 \\ -0,6304 & -0,0033 & 0,7763 \end{bmatrix}$
Vetor de Translação (t) [mm]	$\begin{bmatrix} -392,74 \\ -0,84 \\ 152,51 \end{bmatrix}$
<i>Baseline</i> calculado a partir do vetor de translação [mm]	421,31

Fonte: A Autora (2025).

Os parâmetros obtidos foram salvos em um arquivo de configuração e são utilizados pelo *software* de reconstrução, para corrigir a distorção das imagens e

para realizar a triangulação dos pontos 3D, garantindo a validade métrica de todas as medições apresentadas nas seções seguintes.

4.2 DEFINIÇÃO E OTIMIZAÇÃO DOS PARÂMETROS

O método de reconstrução adotado, baseado na correlação-espço temporal apresentada na Seção 3.1, exige que o padrão de pontos pseudoaleatórios projetado varie ao longo da sequência de imagens capturadas. Essa variação da textura ao longo do tempo é fundamental para garantir a robustez da correlação e a correta identificação da superfície. Se o padrão permanecesse estático, por exemplo, o componente temporal (T) da assinatura ($N \times N \times T$) seria redundante, fazendo com que o algoritmo falhasse em superfícies com textura ambígua.

O deslocamento do padrão é controlado pelo sistema de prismas contrarrotativos acionados por um motor de passo apresentado na Seção 3.2. Conforme estabelecido em estudo prévio da autora e colaboradores (Regner et al., 2025), a definição deste parâmetro baseou-se na análise do Desvio Absoluto Médio (MAD) das imagens. O critério de otimização consistiu em identificar a configuração que maximizasse a variabilidade da textura projetada, visando uma maior robustez na correlação, ao mesmo tempo em que minimizasse o deslocamento angular total dos prismas. A análise concluiu que 10 passos do motor era o valor ótimo. Este valor, que corresponde a uma rotação de 5° dos prismas, foi, portanto, definido como o parâmetro padrão para todas as aquisições realizadas e analisadas nesta dissertação.

A Figura 17 ilustra o momento de uma das aquisições realizadas, mostrando o sistema completo com o projetor *laser* ativo e projetado sobre um dos objetos em análise, o plano.

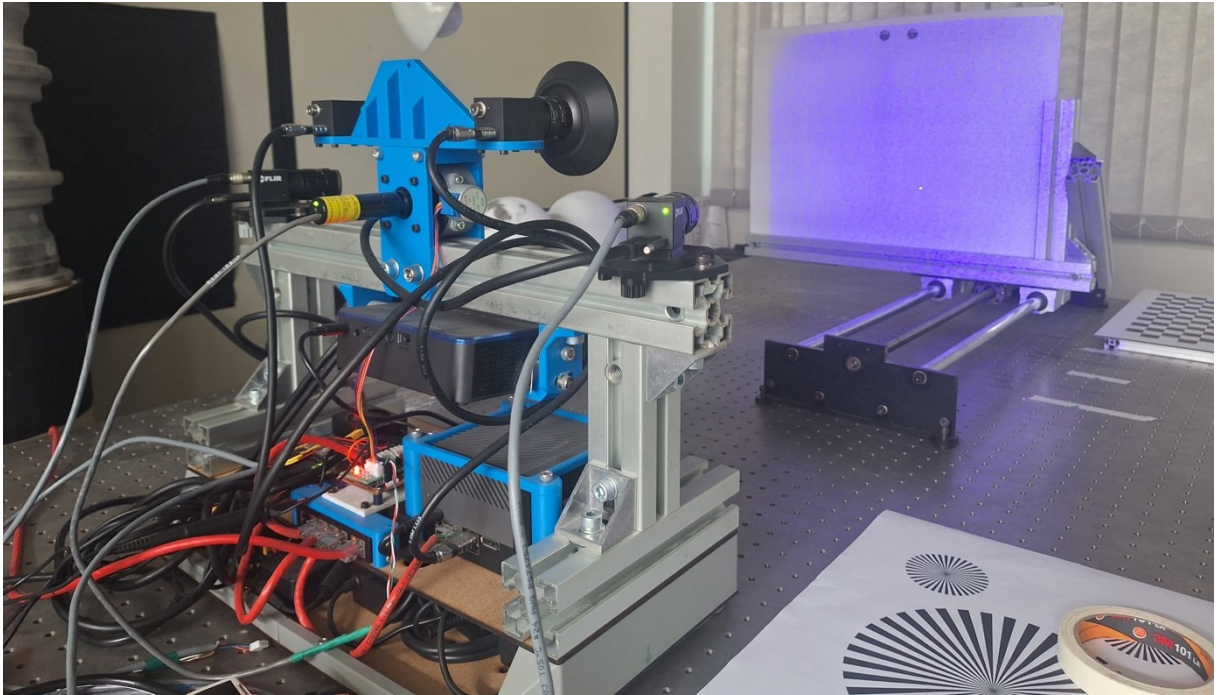


Figura 17 – Exemplo de aquisições do plano.
Fonte: A Autora (2025).

As Figuras 18 e 19 apresentam exemplos das aquisições feitas para o plano e a esfera, respectivamente.

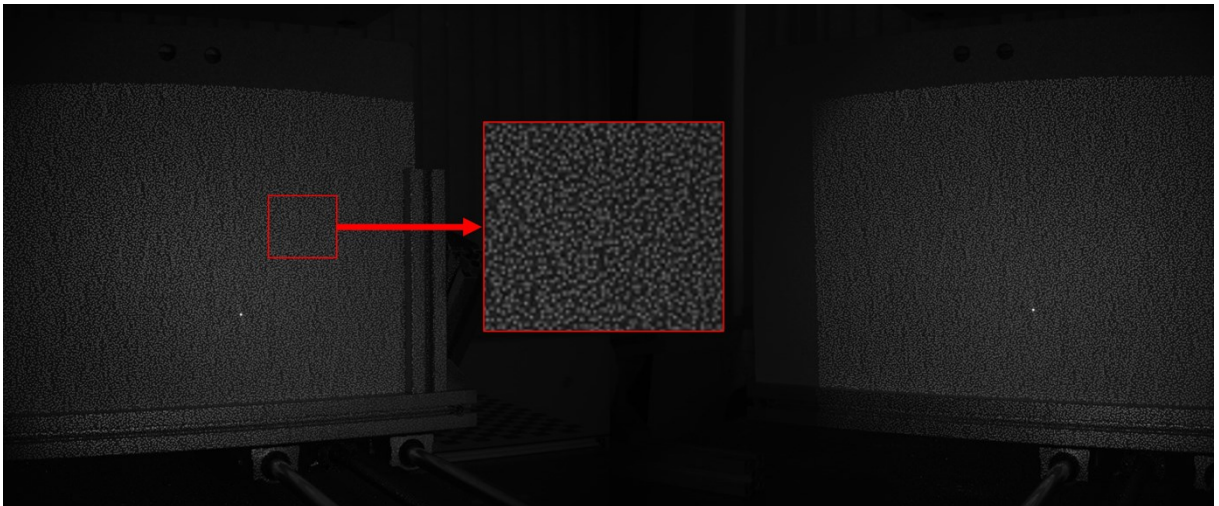


Figura 18 – Exemplo de aquisições obtidas do plano.
Fonte: A Autora (2025).



Figura 19 – Exemplo de aquisições obtidas das esferas.
Fonte: A Autora (2025).

Após a definição do parâmetro de projeção (passos do motor), a etapa seguinte consistiu em otimizar os parâmetros do algoritmo de reconstrução. O primeiro é o número de pares de imagens (T) utilizados na correlação espaço-temporal, que impacta diretamente no tempo de aquisição, na robustez da correlação e no custo computacional do processamento. Para determinar o valor ótimo de T , foram realizados ensaios em um plano de referência posicionado a 850 mm do sistema, variando o número de imagens entre 3, 5, 7, 10 e 15 pares.

Para isolar a influência deste parâmetro, o tamanho do *kernel* de correlação ($N \times N$) foi fixado em 5×5 pixels e a análise estatística foi restrita a uma região de interesse central com dimensões de 200 mm x 150 mm, visando evitar efeitos de borda. Os resultados consolidados, comparando o erro de forma (desvio padrão das distâncias) e a confiabilidade da correlação (mediana PNR), são apresentados na Tabela 3 e sumarizados visualmente no Gráfico 1.

Tabela 3 - Resultados da otimização do número de imagens (T) no plano.

Número de imagens (T)	Desvio Padrão [mm]	Mediana PNR
3	0,26	22,67
5	0,10	28,71
7	0,08	34,57
10	0,08	43,81
15	0,08	58,73

Fonte: A Autora (2025).

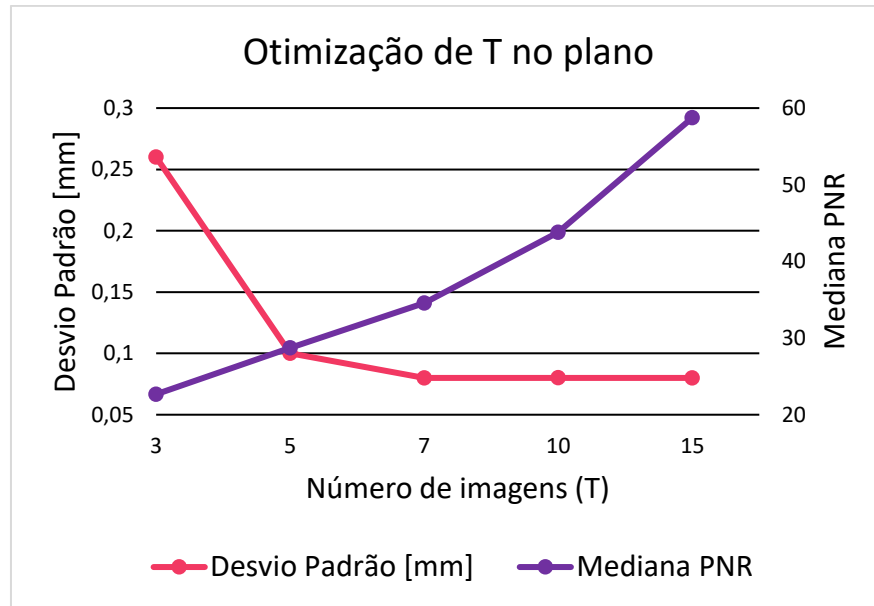


Gráfico 1 - Resultados da otimização do número de imagens (T) no plano.
Fonte: A Autora (2025).

A análise dos resultados demonstrou uma clara relação entre o número de imagens e a qualidade da reconstrução. Com $T = 3$, o sistema apresentou alto ruído (desvio padrão de 0,26 mm) e baixa confiança de correlação (PNR 22,67). O desvio padrão melhorou com $T = 5$ (0,10 mm) e atingiu seu valor mínimo com $T = 7$, estabilizando em um desvio padrão de 0,08 mm. Este valor se manteve para $T = 10$ e $T = 15$, indicando o limite de precisão do sistema sob essas condições.

Contudo, a análise da Relação Pico-Ruído (PNR) revelou que a robustez da correlação continuava a melhorar. Ao passar de 7 (PNR 34,57) para 10 (PNR 43,81) pares de imagens, observou-se um ganho de 26,7%, indicando picos de correlação mais definidos. Embora o uso de 15 pares de imagens tenha elevado o PNR para 58,73, o ganho de robustez não foi considerado suficiente para justificar o custo computacional envolvido. Portanto, definiu-se $T = 10$ como o parâmetro ótimo, por representar o melhor equilíbrio entre a precisão geométrica máxima (desvio padrão de 0,08 mm), alta robustez de correlação (PNR 43,81) e custo computacional factível.

A Figura 20 apresenta a reconstrução do plano obtida para a configuração ótima de $T = 10$. Para esta reconstrução, o volume de medição foi segmentado em uma grade de 9 *tiles* (lotes). A variação da tonalidade observada na imagem, formando um padrão de mosaico, é característica desse método de processamento em lotes, onde cada sub-volume é reconstruído de forma independente. A análise

estatística da superfície resultou em um desvio padrão de 0,08 mm, conforme metodologia definida na Seção 3.4.2. Já a Figura 21 apresenta o PNR obtido.

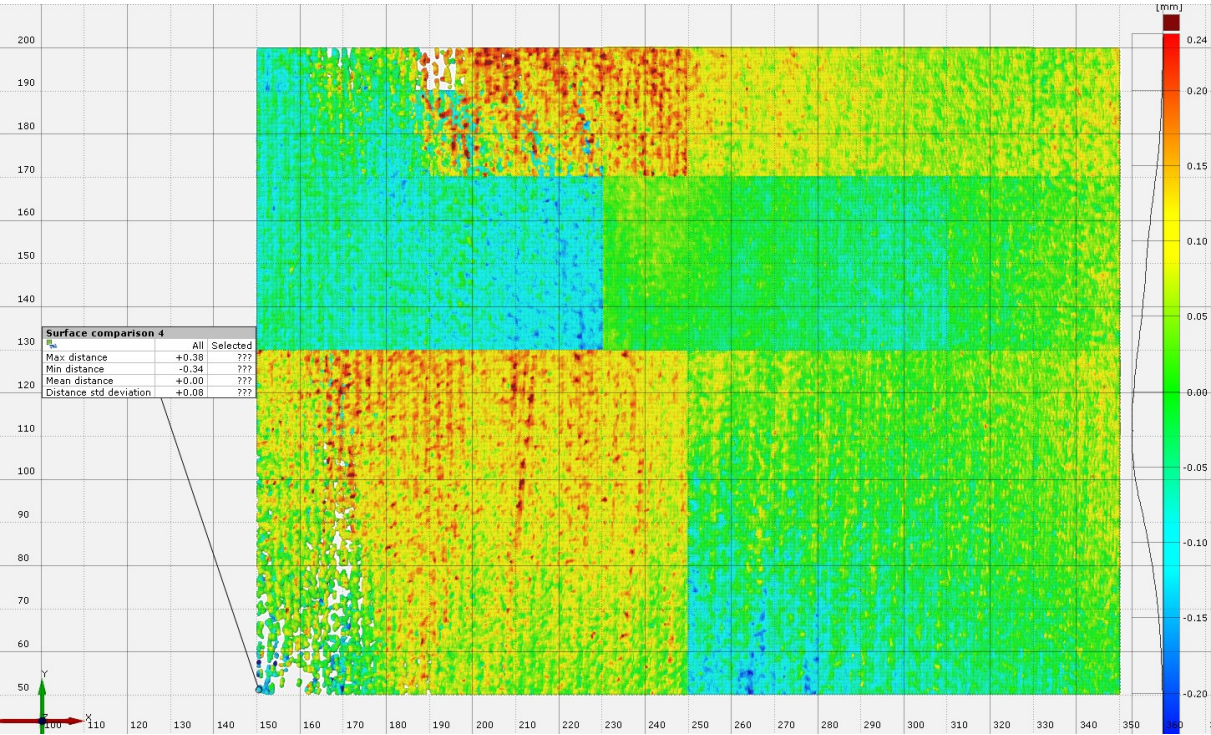


Figura 20 – Reconstrução ótima de T = 10 no plano.
Fonte: A Autora (2025).

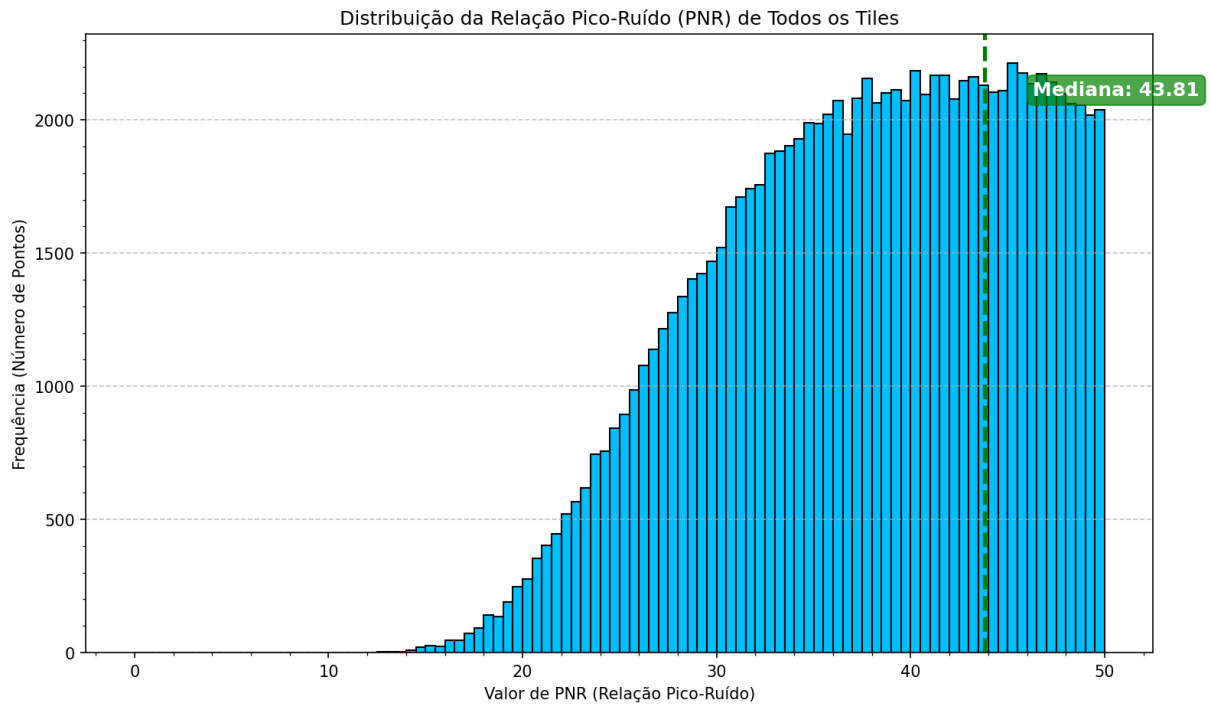


Figura 21 – PNR de T = 10 no plano.
Fonte: A Autora (2025).

Para validar a escolha de $T = 10$ em uma geometria mais complexa, os mesmos ensaios de variação do número de imagens foram replicados utilizando um par de esferas de referência. Os resultados, sumarizados na Tabela 4 e no Gráfico 2, corroboram com as conclusões obtidas no plano.

Tabela 4 – Resultados da otimização do número de imagens (T) nas esferas.

Número de imagens (T)	Desvio Padrão [mm]	Mediana PNR
3	0,57	17,12
5	0,19	24,83
7	0,17	27,92
10	0,15	30,43
15	0,14	36,52

Fonte: A Autora (2025).

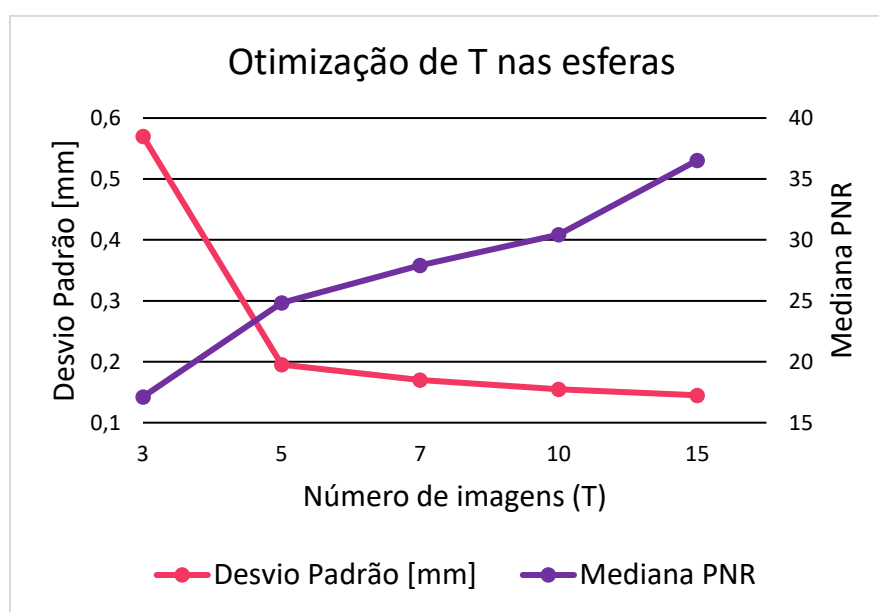


Gráfico 2 - Resultados da otimização do número de imagens (T) nas esferas.

Fonte: A Autora (2025).

É importante ressaltar que os valores de desvio padrão apresentados na Tabela 4 correspondem à média aritmética dos resultados obtidos para as duas esferas. A análise individualizada, que será ilustrada mais adiante na Figura 22, revela que a esfera da direita apresentou um desvio de 0,08 mm, enquanto a da esquerda obteve 0,23 mm, resultando na média global de 0,15 mm reportada. Essa discrepância de qualidade deve-se à posição dos objetos no volume de medição, embora as esferas estejam em profundidades similares, a esfera da esquerda situa-

se em uma região mais periférica do campo de visão das lentes. Nessa região a incidência de distorções ópticas residuais e a menos uniformidade da iluminação do padrão *laser* tendem a elevar o erro de reconstrução em comparação à região central.

No caso das esferas, observou-se a mesma tendência do plano: o desvio padrão apresenta ganhos baixos após $T = 7$, indicando a estabilização da precisão geométrica. No entanto o ganho de robustez (mediana PNR) de $T = 7$ (27,92) para $T = 10$ (30,43) foi novamente considerado fator decisivo. Em geometrias complexas como esferas, que possuem superfícies curvas e bordas que naturalmente dificultam a correlação é possível verificar que resultam em um PNR geral mais baixo do que o do plano. A Figura 22 apresenta a reconstrução obtida para a configuração ótima de $T = 10$ nas esferas, já a Figura 23 apresenta o PNR obtido.

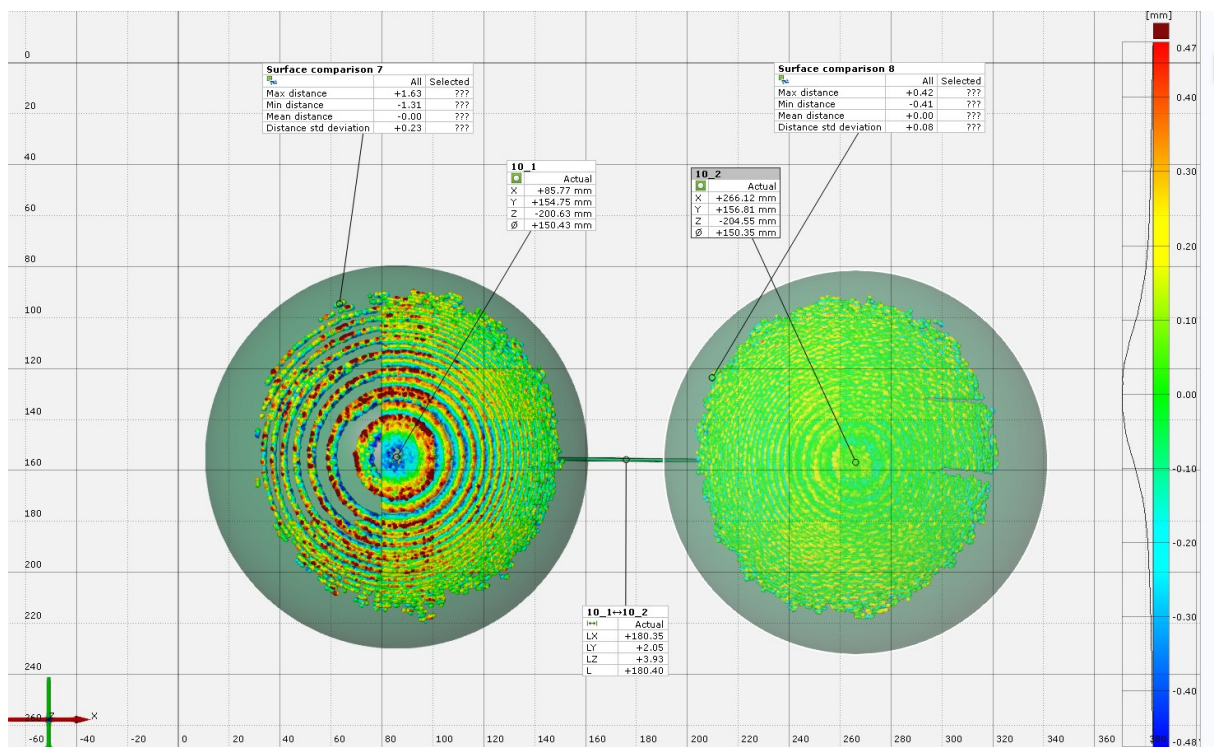


Figura 22 – Reconstrução ótima de $T = 10$ nas esferas.

Fonte: A Autora (2025).

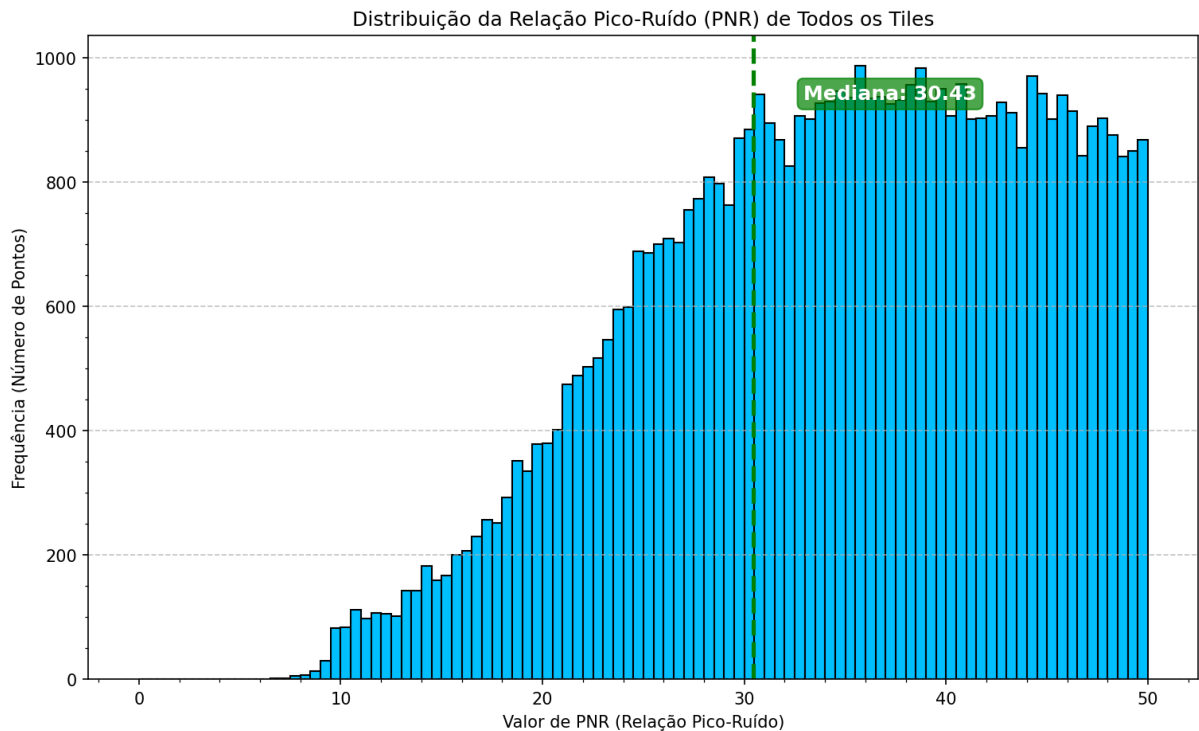


Figura 23 – PNR de $T = 10$ nas esferas.
Fonte: A Autora (2025).

Os dados das esferas confirmam que $T = 10$ oferece o melhor equilíbrio entre precisão e custo computacional, sendo assim definido como parâmetro padrão para os testes subsequentes.

Com o número ótimo de imagens definido em $T = 10$, a etapa seguinte foi otimizar o tamanho da janela de correlação espacial, ou *kernel* ($N \times N$), definido no código como *patch_size*. Um *kernel* maior captura mais informação de textura, o que pode aumentar a robustez da correlação, mas também aumenta o custo computacional e pode “suavizar” detalhes finos da geometria.

Foram realizados ensaios no plano de referência (a 850 mm, com $T = 10$) variando o kernel entre 3×3 , 5×5 e 7×7 . Os resultados estão sumarizados na Tabela 5 e no Gráfico 3.

Tabela 5 – Resultados da otimização do kernel no plano.

Tamanho do <i>Kernel</i>	Desvio Padrão [mm]	Mediana PNR
3×3	0,09	28,56
5×5	0,08	43,81
7×7	0,07	59,87

Fonte: A Autora (2025).

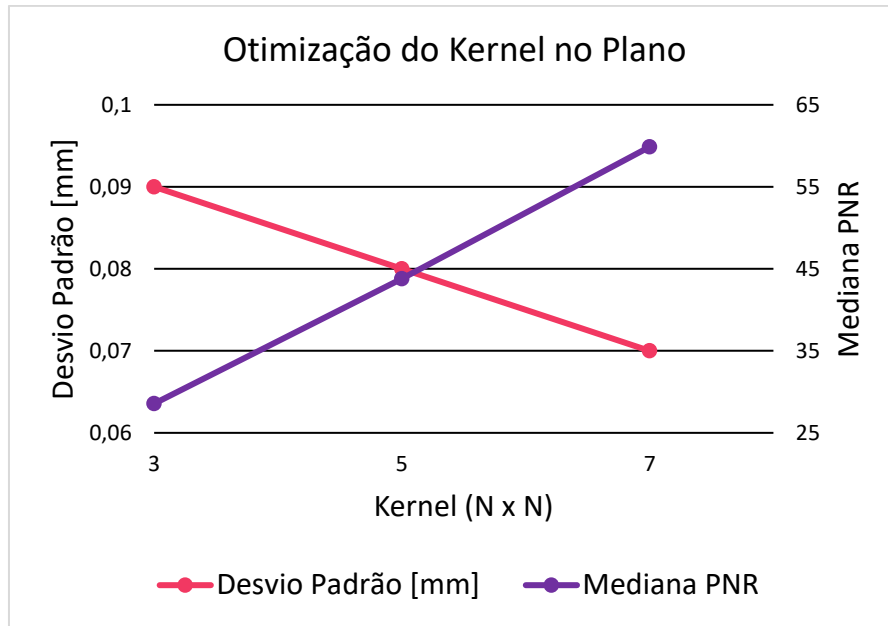


Gráfico 3 - Resultados da otimização do kernel no plano.
Fonte: A Autora (2025).

Os dados mostram que um *kernel* 3 x 3 é insuficiente, resultando em um desvio padrão de 0,09 mm e uma confiança de correlação relativamente baixa (PNR 28,56). Ao aumentar o *kernel* para 5 x 5, o desvio padrão atingiu 0,08 mm e um PNR 43,81. Já ao aumentar o *kernel* para 7 x 7 os valores seguiram melhorando em ambas as métricas, o desvio padrão melhorou para 0,07 mm e a mediana PNR atingiu 59,87. O ganho de desempenho foi considerado superior ao custo computacional adicional, justificando a escolha do *kernel* 7 x 7 como parâmetro ótimo.

Para validar empiricamente a escolha do parâmetro de *kernel* 7 x 7 sugerida pelos dados estatísticos da Tabela 5, a Figura 24 apresenta os perfis de correlação reais extraídos de um ponto da superfície, variando-se a dimensão da janela espacial (N x N).

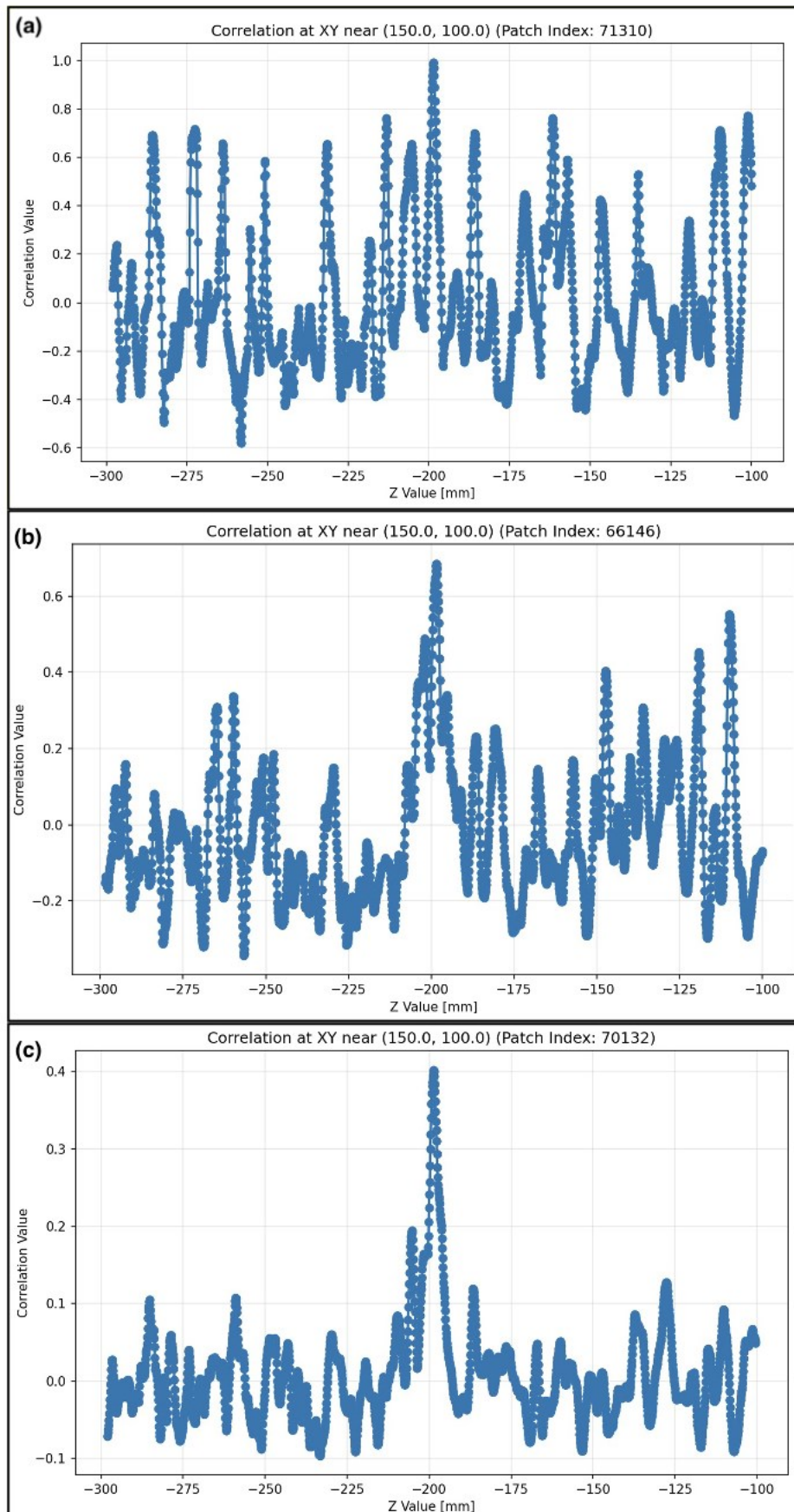


Figura 24 – Perfis de correlação reais (ZNCC x Profundidade Z).
Fonte: A Autora (2025).

No primeiro cenário (a), utilizando um *kernel* unitário de 1 x 1, o perfil de correlação apresenta-se extremamente ruidoso. Observa-se que, além do pico verdadeiro, existem dezenas de picos secundários com magnitude elevada (acima de 0,6), competindo diretamente pela correspondência. Isso caracteriza um cenário de alta ambiguidade, onde a probabilidade de o algoritmo selecionar um falso positivo é crítica. Ao expandir a janela do *kernel* para 3 x 3, conforme ilustrado em (b), nota-se uma melhoria significativa na relação sinal-ruído. O pico principal torna-se mais distinto, contudo, os ruídos continuam apresentando oscilações consideráveis, que podem gerar instabilidade na detecção da profundidade (eixo Z) em regiões de textura repetitiva.

Por fim, no cenário otimizado (c), com *kernel* 7 x 7, o perfil de correlação atinge a condição ideal. O pico de correspondência destaca-se isoladamente, enquanto o ruído de fundo é suprimido para valores próximos de zero. Este comportamento comprova que a informação espacial adicional contida na vizinhança 7 x 7 é suficiente para eliminar a ambiguidade estocástica do padrão pseudoaleatório, garantido a unicidade da solução metrológica.

De forma similar ao realizado na análise da quantidade de imagens, para validar a escolha do *kernel* 7 x 7 em uma geometria mais complexa, os mesmos ensaios de variação do *kernel* foram replicados utilizando um par de esferas de referência. Os resultados, sumarizados na Tabela 6 e no Gráfico 4, corroboram com as conclusões obtidas no plano.

Tabela 6 - Resultados da otimização do kernel nas esferas.

Tamanho do <i>Kernel</i>	Desvio Padrão [mm]	Mediana PNR
3 x 3	0,18	24,81
5 x 5	0,15	30,43
7 x 7	0,14	33,96

Fonte: A Autora (2025).

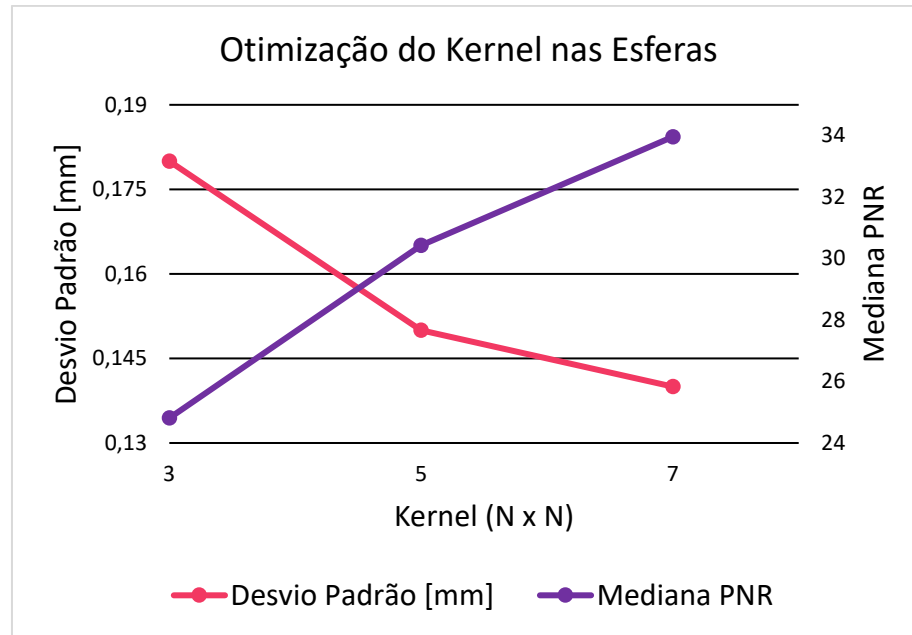


Gráfico 4 - Resultados da otimização do kernel nas esferas.
Fonte: A Autora (2025).

Finalmente, o último parâmetro otimizado foi a sensibilidade do *grid* adaptativo, identificado no código como `pixel_target`. Este valor define o passo de amostragem alvo no espaço da imagem (sensor). Um valor de 1,0, por exemplo, representa uma amostragem 1:1 (um *voxel* por *pixel*), enquanto valores menores representam uma sobreamostragem. Foram realizados ensaios no plano de referência (a 850 mm, com $T = 10$ e $K = 7 \times 7$), variando o `pixel_target` entre 1,0, 0,9 e 0,8. Os resultados estão sumarizados na Tabela 7 e no Gráfico 5.

Tabela 7 – Resultados da otimização da sensibilidade.

Sensibilidade do <i>grid</i>	Desvio Padrão [mm]	Mediana PNR
1,0	0,10	72,09
0,9	0,08	67,54
0,8	0,07	59,87

Fonte: A Autora (2025).

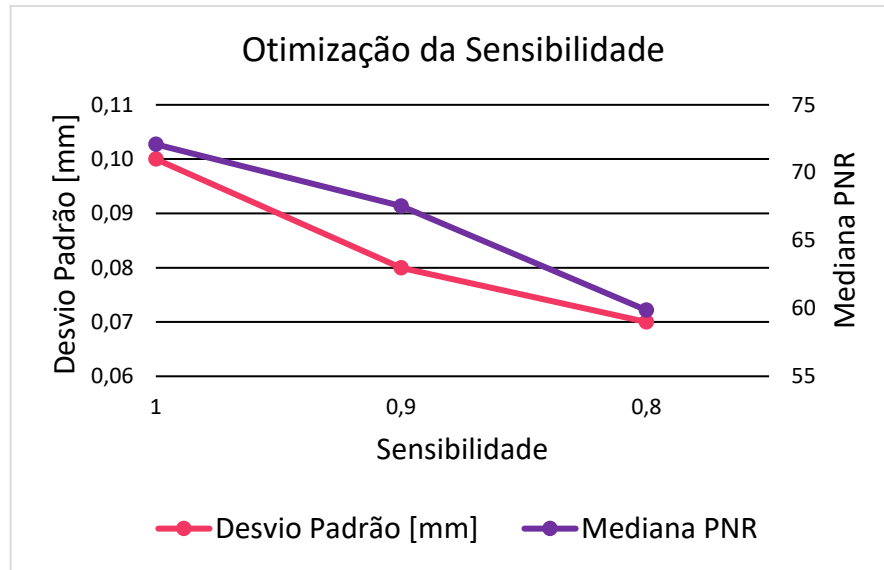


Gráfico 5 - Resultados da otimização da sensibilidade.
Fonte: A Autora (2025).

Os dados revelaram um claro *trade-off* entre a precisão geométrica representada pelo desvio padrão e a robustez da correlação representada pela mediana PNR. À medida que o passo de amostragem diminuiu, aumentando a sobreamostragem, o desvio padrão melhora consistentemente, passando de 0,10 mm (*target* = 1,0) para 0,07 mm (*target* = 0,8). Tentativas de utilizar um *target* ainda menor, 0,7 por exemplo, foram realizadas, mas se mostraram inviáveis. O aumento exponencial no número de *voxels* a serem processados levou ao esgotamento da memória da GPU, causando falhas na execução do software.

Desta forma, o *pixel_target* = 0,8 foi definido como parâmetro ótimo, pois apresenta a melhor precisão geométrica que pôde ser alcançada dentro das limitações do sistema. Embora a mediana PNR tenha diminuído em relação aos valores de *target* maiores, o resultado de 59,87 foi considerado suficiente para a medição.

Por fim, os parâmetros de limiar de correlação (*CORR_THRESHOLD*) e do filtro espacial (*SPATIAL_FILTER_RADIUS* e *SPATIAL_FILTER_MIN_NEIGHBORS*) foram ajustados para remover ruídos de forma eficaz sem “corroer” as superfícies válidas do objeto medido. Os parâmetros-chave, definidos para os ensaios de validação metrológica apresentados nas seções seguintes, estão listados de forma resumida na Tabela 8.

Tabela 8 – Parâmetros de execução do software de reconstrução.

Parâmetro	Símbolo no código	Valor	Unidade
Número de Pares de Imagens (T)	n_images	10	(adimensional)
Sensibilidade do Grid Adaptativo	pixel_target	0,8	<i>pixel</i>
Janela de Correlação (<i>Kernel</i>)	patch_size	7x7	<i>pixels</i>
Limiar de Correlação Mínima	CORR_THRESHOLD	0,75	(adimensional)
Raio do Filtro Espacial	SPATIAL_FILTER_RADIUS	3,0	mm
Mínimo de Vizinhos (Filtro)	SPATIAL_FILTER_MIN_NEIGHBORS	20	(adimensional)

Fonte: A Autora (2025).

A aplicação dos parâmetros resultou na reconstrução do plano apresentado na Figura 25.

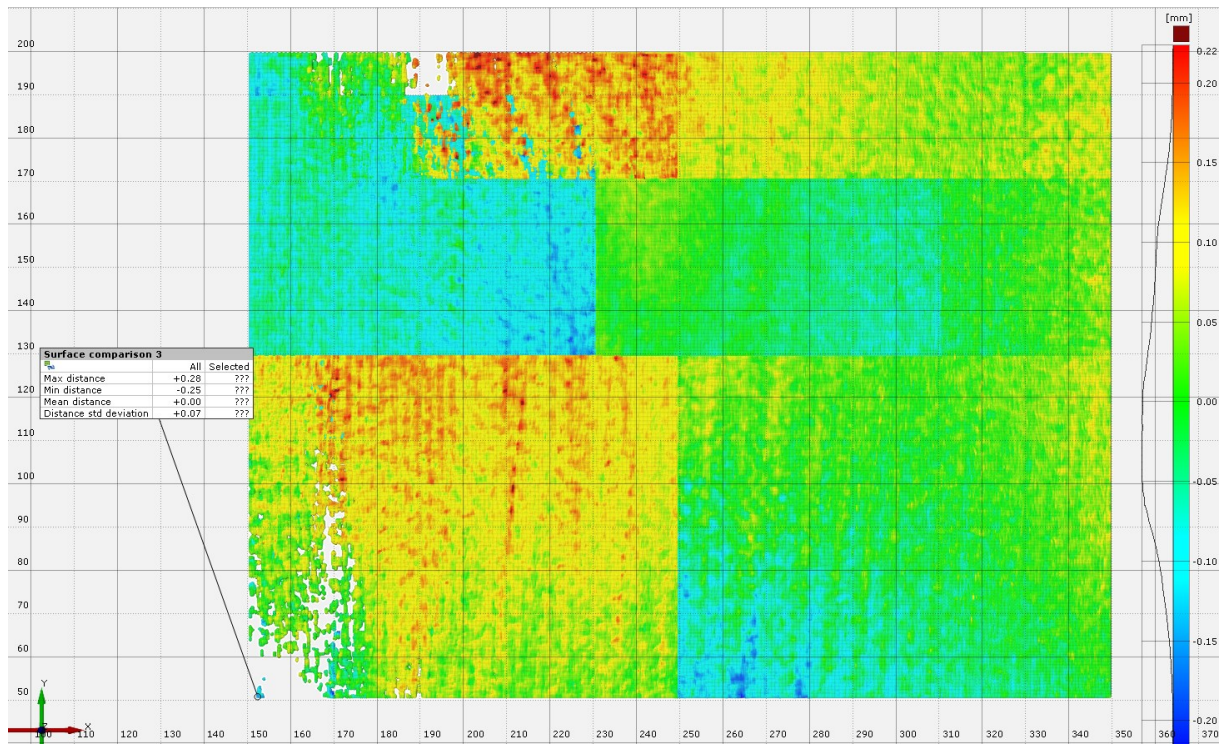


Figura 25 – Reconstrução melhor caso plano.
Fonte: A Autora (2025).

A análise estatística desta aquisição, que representa a condição ideal de operação do sistema, confirmou o desvio padrão mínimo de 0,07 mm. Adicionalmente, avaliou-se a planeza (*flatness*) da superfície reconstruída através da distância pico-vale dos resíduos. Conforme observado na caixa de estatísticas da Figura 24, os desvios variaram entre +0,28 mm e -0,25 mm, resultando em uma planeza global de 0,53 mm. A visualização do mapa de cores evidencia que os maiores desvios se concentram nas regiões de transição entre os *tiles*.

Com todos os parâmetros do algoritmo otimizados, foi realizado um último ensaio para caracterizar o desempenho do sistema em diferentes distâncias de trabalho. O plano de referência foi medido a 700 mm, 850 mm e 1000 mm. Para garantir a consistência da análise em todas as configurações, o volume de busca tridimensional (*grid*) foi delimitado por uma área de varredura de 200 mm (eixo X) por 150 mm (eixo Y) e uma profundidade total de 400 mm (eixo Z). Os resultados estão sumarizados na Tabela 9 e no Gráfico 6.

Tabela 9 – Resultados da análise da distância de trabalho.

Distância de trabalho [mm]	Desvio Padrão [mm]	Mediana PNR
700	0,15	38,63
850	0,07	59,87
1000	0,25	60,59

Fonte: A Autora (2025).

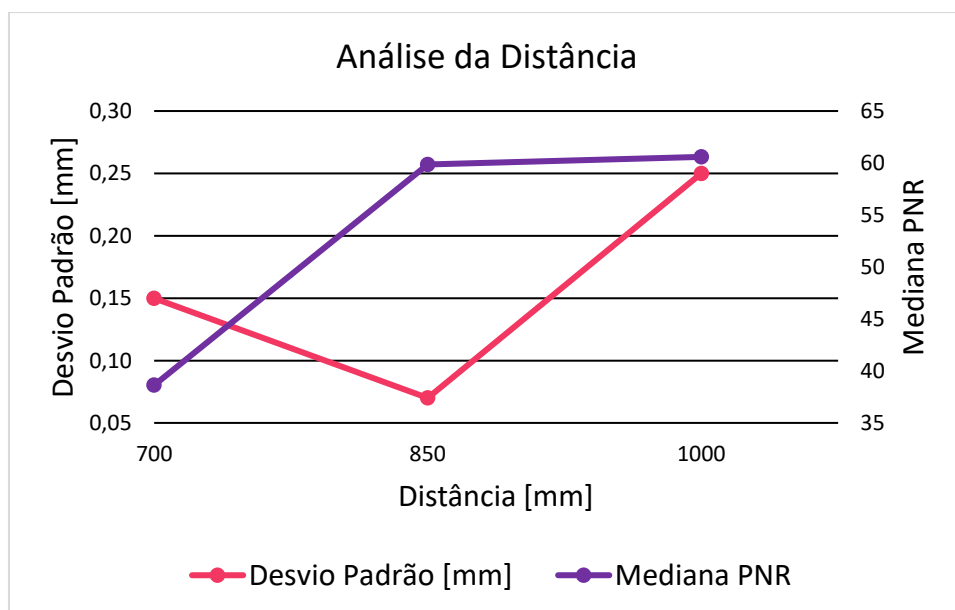


Gráfico 6 - Resultados da análise da distância de trabalho.

Fonte: A Autora (2025).

Os dados demonstram o sistema é dependente da distância, com um ponto ótimo claro em 850 mm, onde o desvio padrão atingiu seu mínimo, de 0,07 mm, e a mediana PNR ficou muito próxima do valor máximo, 59,87. O comportamento verificado é característico da visão estéreo, onde a incerteza da medição de profundidade aumenta de acordo com a distância, amplificando pequenos erros de correspondência. Além disso, a melhor performance em 850 mm é esperada, uma vez que o ajuste de foco das lentes e a calibração do sistema foram otimizados priorizando o centro do volume de medição.

4.3 RECONSTRUÇÃO DE OBJETOS PADRÃO

Com os parâmetros ótimos do sistema definidos na seção anterior ($T = 10$, $K = 7 \times 7$, $\text{pixel_target} = 0,8$) e a distância de trabalho ideal estabelecida em 850 mm, esta seção avalia o desempenho do sistema em diferentes cenários de medição. O objetivo é validar a capacidade do sistema em reconstruir geometrias variadas e comparar o desempenho do método de projeção *laser* pseudoaleatório com outras técnicas de projeção ativa.

A primeira validação geométrica foi a reconstrução de um plano inclinado. Este ensaio avalia o desempenho do sistema em uma superfície que não é perpendicular ao eixo óptico, simulando ângulos de incidência mais desafiadores. A Figura 26 apresenta a reconstrução do ensaio.



Figura 26 – Reconstrução do plano inclinado.

Fonte: A Autora (2025).

A análise dos resultados revelou que, embora mediana PNR tenha se mantido em um nível similar ao caso ótimo, a precisão geométrica proveniente do desvio padrão degradou-se para 0,17 mm, em comparação a 0,07 mm do caso ótimo. Estes dados indicam que o aumento do ruído não se deveu a uma falha do algoritmo, pois o PNR se manteve, mas sim à propagação natural das incertezas da triangulação estéreo, que são amplificadas em superfícies angulares.

A hipótese de que o erro observado no lado esquerdo da reconstrução apresentada na Figura 26 seria causado por falhas no passo adaptativo foi analisada e descartada através da verificação quantitativa dos dados de execução. Para manter a resolução alvo de 0,8 *pixels* no plano da imagem ao longo de toda a extensão do plano inclinado, o *software* ajustou dinamicamente o passo de varredura (*z_step*). Os registros confirmam a compensação geométrica projetiva, apresentando um passo mínimo de 0,74 mm na região mais próxima e um máximo de 4,00 mm na região mais distante, resultando em uma média de 1,35 mm considerando todos os dados. Portanto, o aumento do ruído observado deve-se preponderantemente ao ângulo de incidência da superfície, ou seja, a inclinação acentuada em relação ao eixo óptico provoca uma distorção perspectiva do padrão

de pontos que difere para cada câmera, degradando a qualidade da correspondência, e consequentemente da correlação, localmente.

Dando sequência à validação foi realizada uma comparação de desempenho do método de projeção *laser* pseudoaleatório com um padrão alternativo de luz estruturada (banda limitada), ambos medindo o plano de referência a 850 mm. Os resultados estão sumarizados na Tabela 10.

Tabela 10 – Resultados comparação dos padrões de luz.

Método de projeção	Desvio Padrão [mm]	Mediana PNR
<i>Laser</i> pseudoaleatório	0,07	59,87
Banda limitada	0,07	19,83

Fonte: A Autora (2025).

A análise revelou que tanto o método proposto quanto o padrão de banda limitada permitiram ao algoritmo de correlação atingir a mesma precisão geométrica de 0,07 mm. A diferença fundamental reside na robustez da correlação. O sistema *laser* pseudoaleatório da proposta apresentou uma mediana PNR de 59,87, já o padrão de banda limitada apresentou uma mediana de 19,83, valor 3 vezes inferior.

A validação do erro de medição do sistema foi realizada através da medição de um par de esferas de referência com diâmetro aproximado de 150 mm cada. A Figura 27 apresenta a reconstrução obtida para a melhor configuração nas esferas, já a Figura 28 apresenta o PNR obtido.

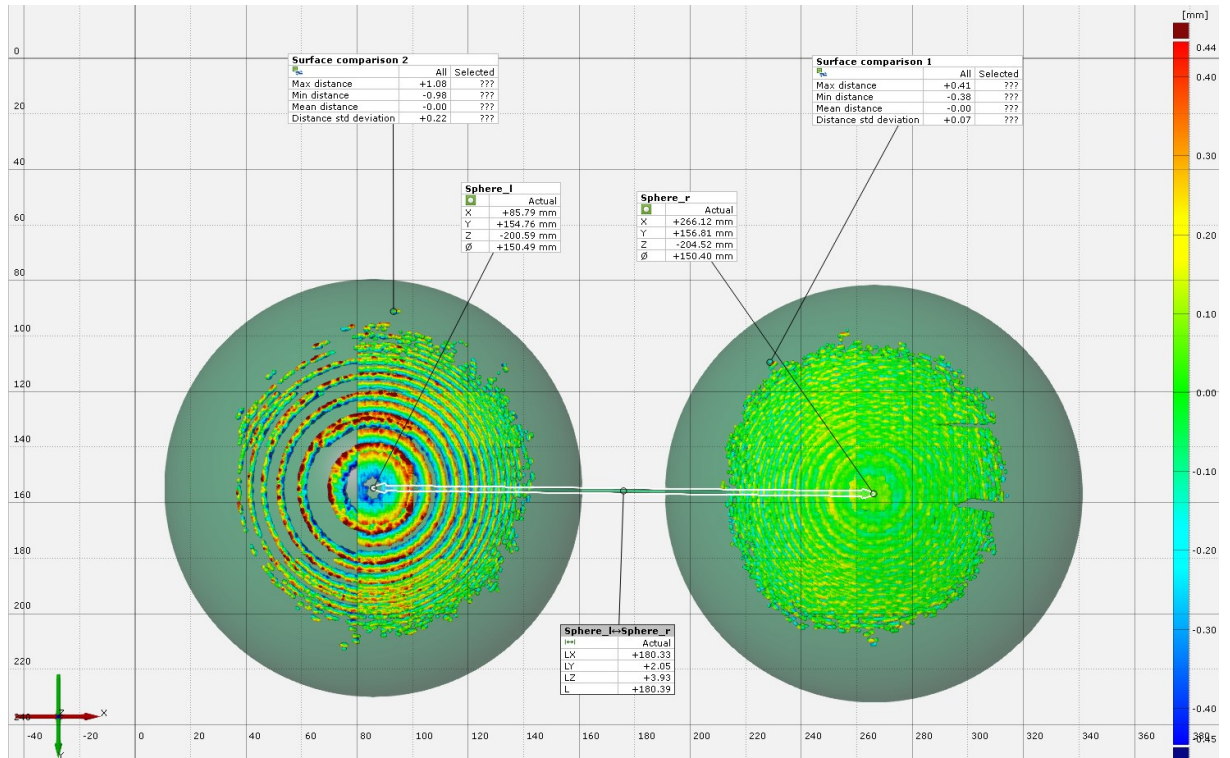


Figura 27 – Reconstrução melhor caso esferas.
Fonte: A Autora (2025).

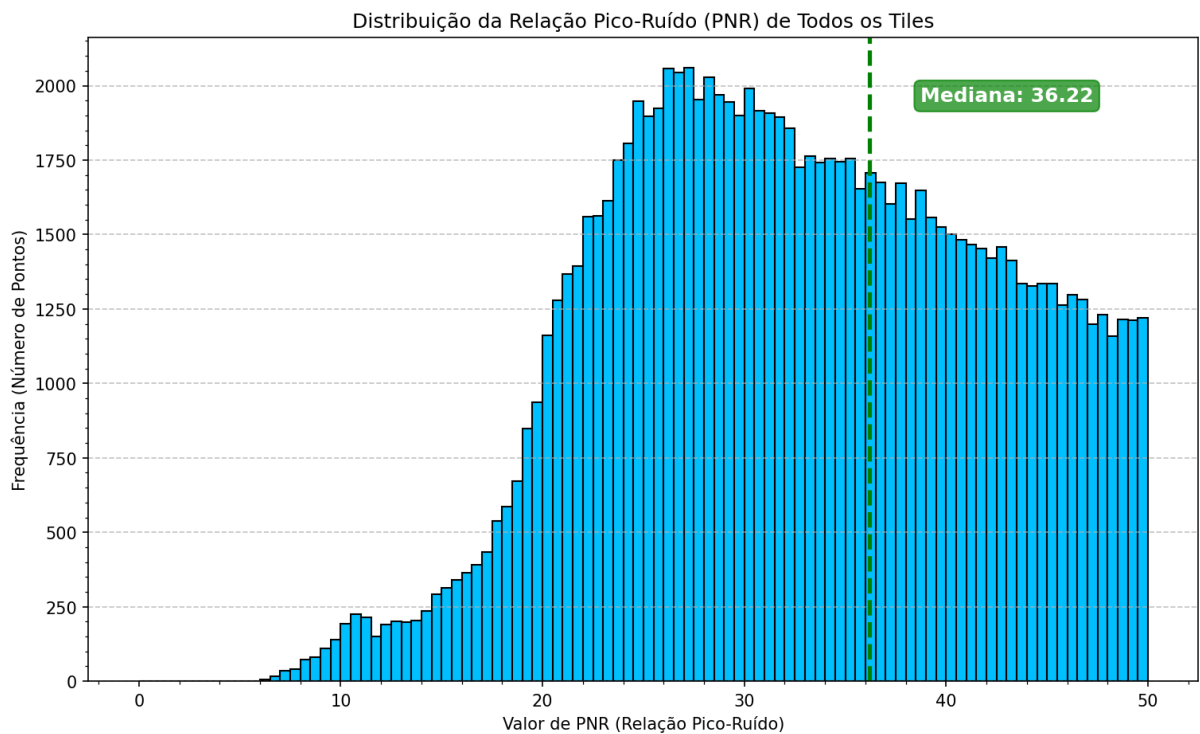


Figura 28 – PNR melhor caso esferas.
Fonte: A Autora (2025).

A análise foi dividida em erro de forma e erro de medição. O erro de forma, análogo ao erro de apalpação de uma MMC, foi avaliado pelo desvio padrão dos

pontos medidos em relação à geometria esférica ideal ajustada. Os valores obtidos foram de 0,07 mm para a Esfera 1 (direita) e de 0,22 mm para a Esfera 2 (esquerda), resultando em um erro de forma médio de 0,145 mm. A robustez da medição se manteve relativamente alta, com uma mediana PNR de 36,22.

O erro de medição foi avaliado comparando os diâmetros medidos (D1 e D2) e a distância entre os centros (L) com os valores nominais do certificado de calibração. Os resultados estão sumarizados na Tabela 11.

Tabela 11 – Resultados medição das esferas.

Parâmetro	Valor nominal [mm]	Valor medido [mm]	Erro [mm]
D1	150,056	150,40	+0,344
D2	150,066	150,49	+0,424
L	180,735	180,39	-0,345

Fonte: A Autora (2025).

O sistema apresentou um erro de medição sistemático positivo nos diâmetros, medindo-os maiores, e um erro de escala negativo na distância entre centros, medindo-a menor.

A validação final do desempenho foi realizada na reconstrução de uma cena complexa, um artefato cilíndrico escalonado (calota), conforme aquisições apresentadas na Figura 29.



Figura 29 – Exemplo de aquisições da calota.

Fonte: A Autora (2025).

A nuvem de pontos adquirida foi comparada a um modelo de referência calibrado em sistema de medição de referência e os resultados são apresentados nas Figuras 30 e 31.

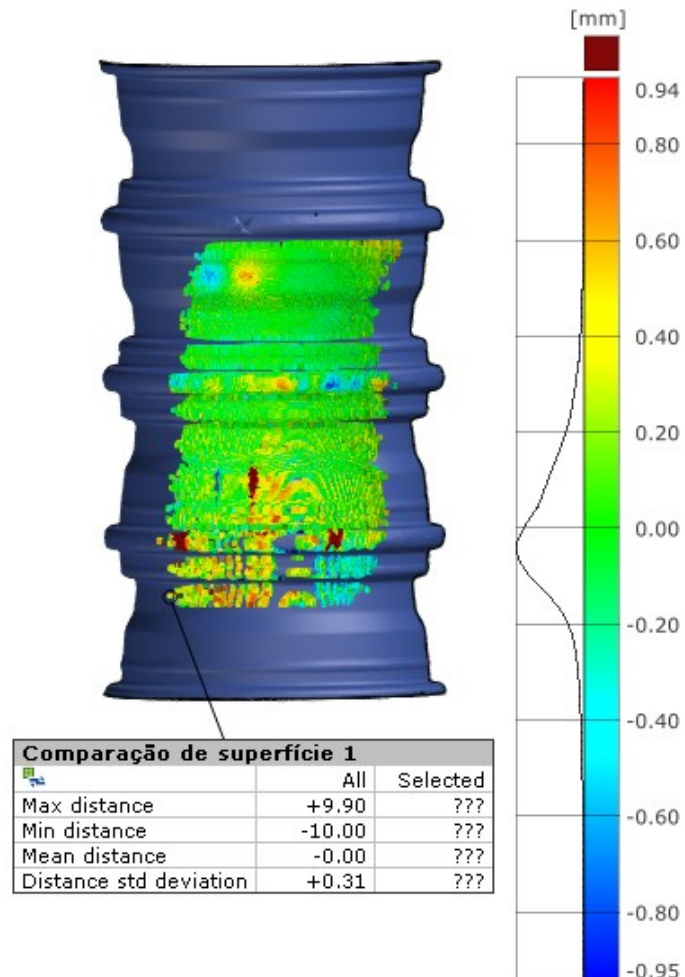


Figura 30 – Reconstrução calota.
Fonte: A Autora (2025).

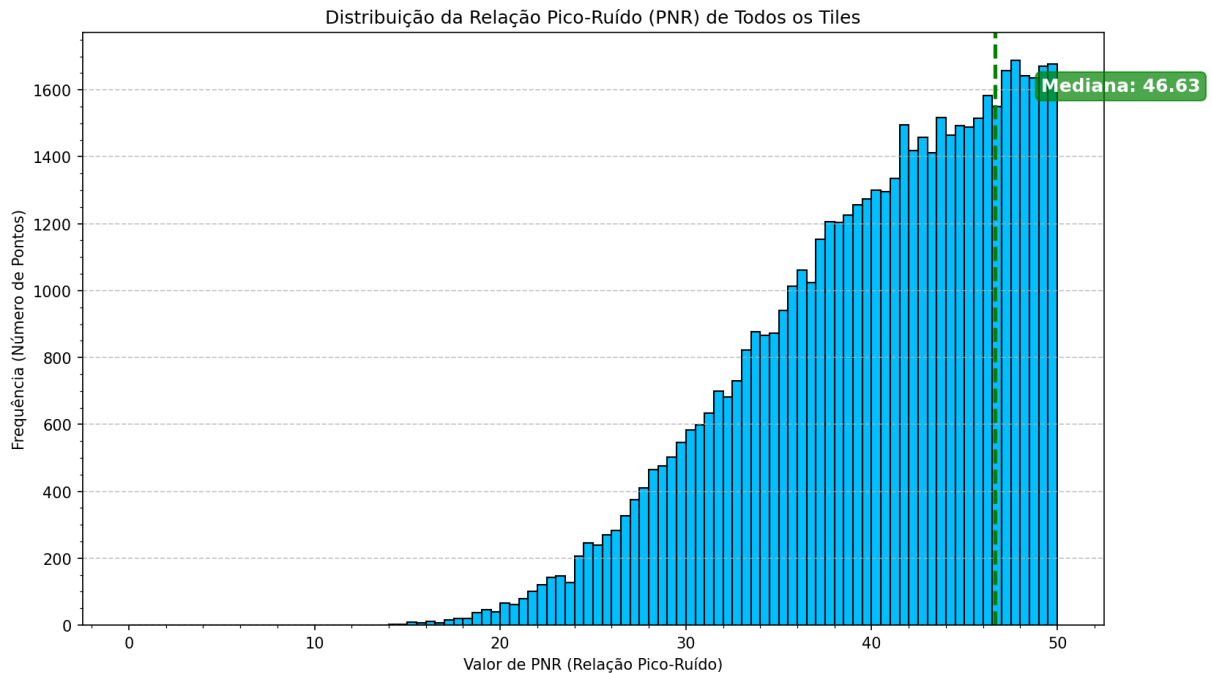


Figura 31 – PNR calota.
Fonte: A Autora (2025).

O sistema demonstrou capacidade de reconstruir a geometria complexa com boa fidelidade, apresentando um desvio padrão dos resíduos de 0,31 mm na comparação com o modelo de referência. Notavelmente, a robustez da correlação (mediana PNR) manteve-se em um nível elevado de 46,63.

4.4 RECONSTRUÇÃO DE OBJETOS DE FORMA LIVRE

Após a validação metrológica do sistema utilizando artefatos de geometria conhecida, buscou-se avaliar o desempenho qualitativo da reconstrução em objetos de forma livre. Esta etapa permite demonstrar a versatilidade da técnica de correlação espaço-temporal em cenários mais próximos de aplicações reais, onde as superfícies apresentam curvaturas complexas, variações abruptas de profundidade e oclusões que não ocorrem em primitivas geométricas como planos e esferas.

Para esta avaliação foram selecionados dois objetos com características distintas. O primeiro objeto selecionado, apresentado na Figura 32, consiste em uma réplica de crânio humano manufaturada em polímero. A escolha deste artefato justifica-se pela sua topologia desafiadora, que intercala superfícies de curvatura

suave na caixa craniana com regiões de alta descontinuidade geométrica e concavidades profundas.



Figura 32 - Objeto 1: réplica de crânio humano.
Fonte: A Autora (2025).

A reconstrução resultante, apresentada na Figura 33, demonstra uma nuvem de pontos de alta densidade, com considerável integridade geométrica em toda a superfície visível. O sistema manteve a capacidade de resolver geometria com variações abruptas de profundidade (eixo Z), sendo que a definição obtida nas regiões de concavidade é consistente e nítida. O algoritmo de correlação espaço-temporal demonstrou robustez suficiente para discriminar detalhes finos na região da maxila e da mandíbula.

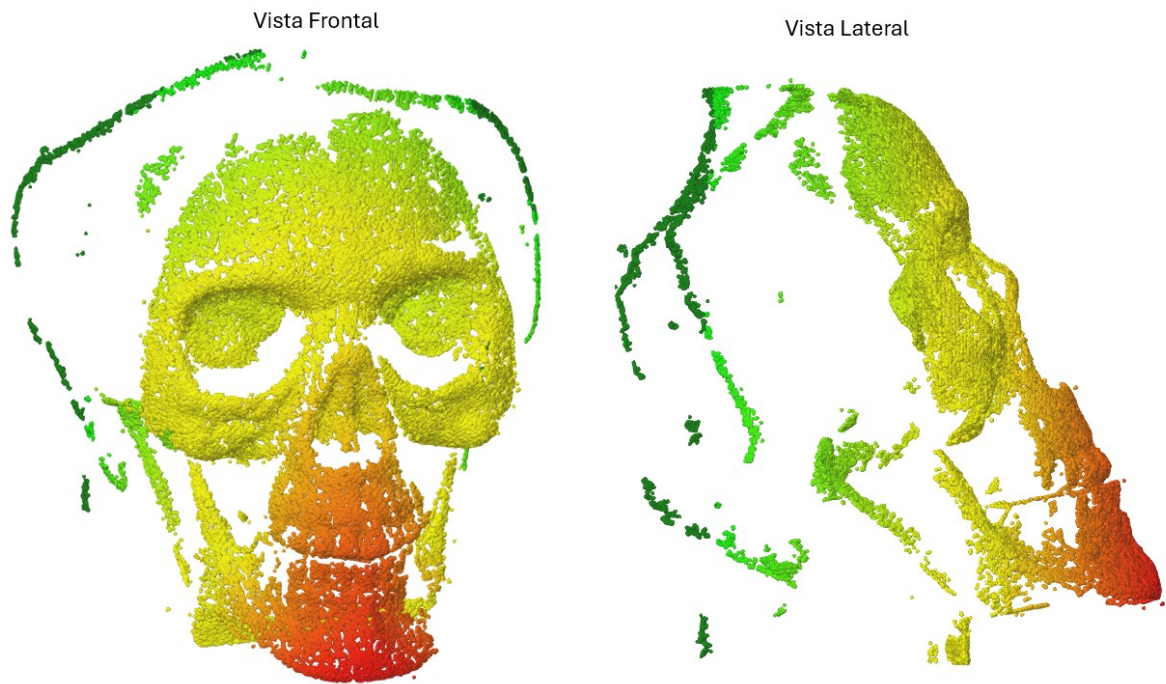


Figura 33 - Reconstrução da réplica de crânio humano.
Fonte: A Autora (2025).

O segundo objeto selecionado para a avaliação qualitativa, ilustrado na Figura 34, é uma cabeça de manequim padrão, caracterizada por superfícies suaves de curvatura contínua e um acabamento uniforme. Essa configuração permite avaliar o desempenho do sistema em um objeto com alta homogeneidade superficial.



Figura 34 – Objeto 2: cabeça de manequim.
Fonte: A Autora (2025).

O resultado da reconstrução é apresentado na Figura 35. A nuvem de pontos resultante é homogênea e densa em toda a superfície visível, demonstrando a capacidade do sistema em realizar o mapeamento de curvaturas. O sistema foi capaz de resolver detalhes da topografia fácil, como o perfil do nariz, lábios e contorno dos olhos. A integridade da nuvem de pontos na região central da face, que é a região com iluminação mais intensa, confirma a eficácia do sistema em capturar a textura do padrão projetado, garantindo uma reconstrução coesa.

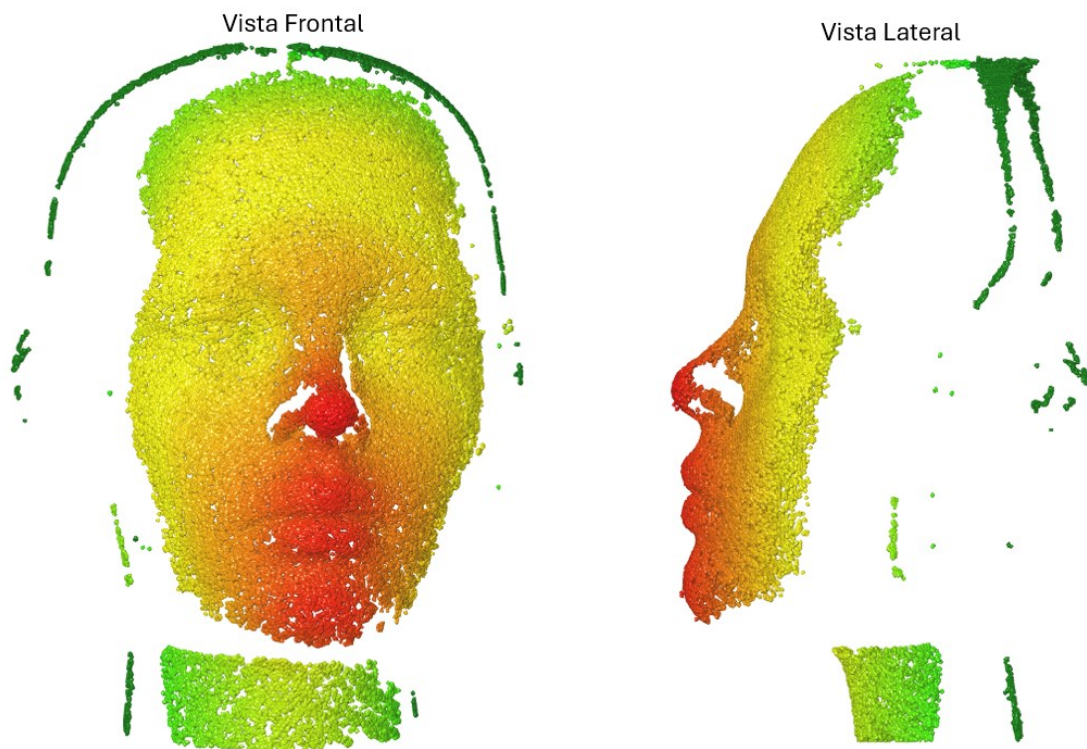


Figura 35 - Reconstrução da cabeça de manequim.
Fonte: A Autora (2025).

A avaliação qualitativa nos objetos de forma livre demonstrou a capacidade do sistema em resolver geometrias complexas e com grandes variações de curvatura. A densidade e a integridade geométrica alcançadas nas reconstruções apresentadas nas Figuras 33 e 35 validam a robustez do algoritmo de correlação espaço-temporal e a adequação do *grid* adaptativo para gerar nuvens de pontos de alta fidelidade em cenários reais.

4.5 DISCUSSÃO

A apresentação e análise dos resultados nas seções anteriores permitem uma discussão sobre o desempenho global do sistema de visão estéreo ativa desenvolvido. A discussão é centrada em quatro eixos principais: a validação da arquitetura de otimização, a avaliação do *hardware* de projeção, a caracterização dos componentes de erro do sistema e a análise da interação luz-matéria em objetos complexos.

A otimização de parâmetros da Seção 4.2 não apenas definiu a configuração ideal, mas também validou a robustez da arquitetura de *software* proposta. A análise do número de imagens (T) revelou que o desvio padrão de medição atinge um patamar de estabilidade (0,08 mm) a partir de 7 imagens. No entanto, a escolha de $T = 10$ se justificou pelo ganho expressivo na Relação Pico-Ruído (PNR), que subiu de 34,57 para 43,81. Esse comportamento demonstra que, embora o ganho em precisão geométrica sature, a robustez da correspondência continua a aumentar, garantindo maior confiabilidade em regiões de textura complexa sem inviabilizar o custo computacional.

A análise da sensibilidade do *grid* (*pixel_target*) evidenciou a importância de amostragem correta do volume de medição. O melhor desempenho (desvio padrão de 0,07 mm) foi obtido com o alvo de 0,8 *pixel*. Fisicamente, considerando a distância de trabalho de 850 mm e a ótica utilizada, um *pixel* no sensor corresponde a aproximadamente 0,25 mm na superfície do objeto. Ao fixar o alvo em 0,8 o algoritmo de discretização adaptativa força uma sobreamostragem (sub-pixel) consistente em todo o volume. Isso explica o desempenho superior em comparação ao alvo de 1,0 *pixel*, pois garante a extração máxima de detalhes da estrutura projetada, mitigando erros de *aliasing* na etapa de correlação.

A análise comparativa da Seção 4.3 validou a escolha do *hardware*. O sistema de projeção a *laser* com prismas contrarrotativos atingiu a mesma precisão geométrica (0,07 mm) que o padrão de banda limitada. Cabe ressaltar que, para esta comparação, utilizou-se um padrão de banda limitada gerado aleatoriamente, sem processos de otimização específicos. Contudo, a superioridade do sistema proposto reside na eliminação da ambiguidade: a variação dinâmica e controlada do padrão resultou em um PNR três vezes superior (59,87 contra 19,83). Isso confirma que a assinatura espaço-temporal ($N \times N \times T$) gerada pela movimentação dos

prismas oferece uma correspondência mais distinta do que a aleatoriedade puramente espacial.

A análise dos erros de medição permite segregar a influência dos componentes aleatórios e sistemáticos na incerteza global do sistema. O desvio padrão reduzido na reconstrução do plano (0,07 mm) e das esferas (0,15 mm) comprova que o método de correlação possui alta precisão e repetibilidade, sendo capaz de resolver a topografia com baixo ruído. Entretanto, o erro de medição na distância entre esferas (-0,345 mm) revela uma componente sistemática dominante. Isso indica que a limitação da exatidão atual não é intrínseca ao método de reconstrução proposto, mas sim uma consequência direta do modelo de calibração estéreo e da estimativa da *baseline*.

Em comparação com o estudo preliminar de Regner et al. (2025), a otimização sistemática aqui conduzida (ajustes de *T*, *kernel* e *grid*) resultou em uma redução do desvio padrão de 0,28 mm para 0,07 mm, uma melhoria de quatro vezes no desempenho. Embora trabalhos focados em interferometria ou pequenos volumes, como o de Pinto e Albertazzi (2013), reportem incertezas na ordem de 0,01 mm, deve-se considerar a escala de medição. O sistema desenvolvido neste trabalho manteve um desvio padrão abaixo de 0,1 mm em um volume de medição amplo, demonstrando sua adequação para a digitalização de componentes industriais de médio porte.

No entanto, ao analisar a topografia global da superfície otimizada, observa-se uma discrepância entre a alta precisão local (desvio padrão de 0,07 mm) e a planeza da superfície (0,53 mm). Este comportamento é atribuído aos artefatos gerados nas bordas de junção dos *tiles*. Como cada sub-volume é reconstruído independentemente, pequenas discontinuidades nas regiões de sobreposição acumulam erros de pico-vale, sem, contudo, degradar a confiabilidade estatística da nuvem de pontos como um todo.

Finalmente, a avaliação qualitativa nos objetos de forma livre da Seção 4.4 complementou a análise metrológica ao demonstrar a robustez operacional do sistema. Os resultados de reconstrução do crânio e do manequim comprovaram que, com os parâmetros de aquisição adequados, o sistema é capaz de gerar nuvens de pontos densas e íntegras em geometrias complexas e em diferentes condições de superfície. Isso valida a eficácia do método de correlação espaço-temporal mesmo em cenários desafiadores, confirmando que o sistema é adequado

para a digitalização de componentes industriais e a aplicação final no Projeto VORIS.

Em termos de performance computacional, o sistema gera nuvens de pontos densas, ultrapassando a marca de 1 milhão de pontos por reconstrução. Devido à estratégia de processamento em lotes, esses pontos distribuem-se de forma não simétrica entre os *tiles*, adaptando-se à complexidade da cena. O tempo total de processamento para uma nuvem completa varia entre 20 e 40 minutos, dependendo da profundidade do volume de busca e da densidade do *grid*. Ressalta-se que, nesta etapa da pesquisa, o foco deu-se à exatidão metrológica e robustez, não tendo sido aplicadas técnicas de otimização de código voltadas ao tempo real.

As limitações identificadas serão recapituladas no Capítulo 5, onde também serão propostos trabalhos futuros.

5 CONCLUSÃO

Este capítulo final apresenta a síntese e as conclusões da pesquisa desenvolvida. O propósito é consolidar os objetivos propostos, os métodos aplicados e os resultados obtidos, contextualizando as contribuições do trabalho frente ao estado da arte. A estrutura do capítulo inicia com uma recapitulação do escopo e dos principais achados. Em seguida, são detalhadas contribuições científicas e tecnológicas. Subsequentemente, é realizada uma análise crítica das limitações do sistema e da metodologia, que fundamenta, por fim, a sugestão de direcionamentos para trabalhos futuros.

5.1 RECAPITULAÇÃO DO TRABALHO

O objetivo geral desta dissertação foi o desenvolvimento e a validação metrológica de um sistema de visão estéreo ativa com projeção de luz estruturada, projetado para atender aos requisitos de inspeção visual automatizada no contexto do Projeto VORIS. A arquitetura proposta foi fundamentada no Método da Triangulação Inversa, utilizando uma métrica de correspondência baseada na correlação espaço-temporal ($N \times N \times T$) para processar as imagens adquiridas com a variação de um padrão *laser* de pontos pseudoaleatórios projetado.

Para atender aos objetivos específicos, foi implementado um *software* de reconstrução e conduzida uma otimização sistemática dos seus parâmetros de execução, apresentada na Seção 4.2. Esta análise determinou a configuração ótima ($T = 10$ pares de imagens, $kernel = 7 \times 7$, $pixel_target = 0,8$), que maximizou o equilíbrio entre a precisão geométrica e a robustez da correlação. O desempenho do sistema otimizado foi quantificado, atingindo um desvio padrão de 0,07 mm e uma planeza de 0,53 mm na medição de um plano de referência, uma melhoria de 4 vezes em relação ao trabalho base (Regner et al., 2025). Este resultado valida o sucesso do projeto frente ao requisito de incerteza submilimétrica estabelecido par a inspeção de estruturas *offshore*.

A validação metrológica, realizada com esferas padrão calibradas, permitiu a caracterização dos erros de medição do sistema. A análise quantificou os componentes de erro, identificando um erro sistemático de escala de -0,345 mm na medição de distâncias, superior ao erro aleatório observado de 0,15 mm,

proveniente do desvio padrão médio da reconstrução das esferas. Este erro sistemático é passível de correção via refinamento do processo de calibração.

Complementarmente, os testes qualitativos em objetos de forma livre validaram a capacidade do sistema em reconstruir detalhes finos e gerar nuvens de pontos densas e íntegras em geometrias complexas, conforme apresentado na Seção 4.4. A robustez demonstrada, aliada à validação do uso de *laser* azul, confirma a adequação da arquitetura desenvolvida para a futura integração em ROVs e aplicação em ambiente subaquático do Projeto VORIS, superando as limitações de flexibilidade dos sistemas comerciais.

5.2 CONTRIBUIÇÕES

O desenvolvimento desta pesquisa resultou em contribuições científicas e tecnológicas para a área de metrologia óptica e visão computacional, abrangendo desde a inovação na arquitetura de *software* até a validação experimental do sistema. Uma das principais contribuições deste trabalho reside na implementação de uma arquitetura de *software* de alto desempenho capaz de viabilizar a técnica de correlação espaço-temporal ($N \times N \times T$). Destacam-se as seguintes inovações na implementação:

- Processamento paralelo em GPU: o desenvolvimento de um núcleo de correlação otimizado utilizando a biblioteca PyTorch, permitindo o uso de paralelismo massivo em unidades de processamento gráfico (GPU). Essa abordagem foi fundamental para processar o alto volume de dados (milhares de *voxels* multiplicados por sequências de imagens) em tempo hábil, superando limitações de implementações convencionais em CPU. Testes comparativos preliminares indicaram que a implementação em GPU proporcionou uma redução da ordem de quatro vezes no tempo de execução;
- Estratégia de *tiling* para gestão de memória: a criação de um algoritmo de segmentação do volume baseado em lotes (*tiles*) independentes. Essa solução de engenharia de *software* permitiu a reconstrução de cenas não preparadas e geometricamente complexas. O método contorna as limitações físicas de memória de vídeo (VRAM) ao processar localmente sub-volumes com

sobreposição, permitindo a aplicação da correlação robusta em toda a extensão da cena, independentemente de sua topografia global;

- Discretização adaptativa do volume (busca em Z): a implementação de um *grid* de busca tridimensional com passo dinâmico e adaptativo. Diferente de abordagens com passo fixo, este algoritmo ajusta a densidade de *voxels* em função da distância da câmera para garantir uma amostragem uniforme no plano da imagem, assegurando que a resolução da construção seja consistente tanto em regiões próximas quanto distantes do sensor.

Em relação à metodologia de otimização de parâmetros, o trabalho contribuiu com um estudo sistemático dos parâmetros de correlação espaço-temporal. A análise dos *trade-offs* entre o número de imagens (T), o tamanho do *kernel* (K) e a sensibilidade do *grid* (pixel_target), apresentada na Seção 4.2, fornece um guia metodológico que validou a arquitetura e resultou em uma melhoria de 4 vezes na precisão do sistema (de 0,28 mm para 0,07 mm) em comparação a estudos preliminares.

Em relação à validação metrológica e de *hardware*, a pesquisa validou experimentalmente a eficácia do sistema de projeção baseado em prismas contrarrotativos. Os testes demonstraram que a variação dinâmica do padrão pseudoaleatória elimina ambiguidades de correspondência, resultando em uma relação pico-ruído três vezes superior à de padrões estáticos de banda limitada (59,87 contra 19,83). Adicionalmente, a caracterização metrológica identificou que o gargalo de desempenho atual do sistema é o erro sistemático de escala proveniente da calibração estéreo, e não o ruído aleatório do algoritmo de correlação, direcionando o foco para melhorias futuras na calibração.

5.3 LIMITAÇÕES DO TRABALHO

Uma análise crítica dos resultados revela os parâmetros de contorno e as limitações do sistema desenvolvido, que são fundamentais para a correta interpretação do seu desempenho:

1. Erro sistemático como fator dominante: conforme evidenciado na análise de erros com as esferas padrão, a exatidão global do sistema é limitada pelo erro de escala. O erro sistemático de -0,345 mm

identificado na medição de distância é a fonte de erro dominante. Isso indica que, embora o sistema apresente baixo ruído aleatório, sua exatidão final é comprometida pela qualidade da calibração do par estéreo;

2. Custo computacional e processamento *offline*: a arquitetura de correlação espaço-temporal é computacionalmente demandante, a implementação na linguagem Python com o uso da biblioteca PyTorch, embora ideal para a prototipagem e permitindo a flexibilidade da pesquisa, não viabiliza o processamento das imagens em tempo real. A otimização do *grid* é limitada pela memória da GPU, definindo uma restrição prática de *hardware*.
3. Dependência da distância de trabalho: a análise da Seção 4.2 demonstrou que o desempenho do sistema não é uniforme em todo o seu volume de medição. A precisão ótima (desvio padrão de 0,07 mm) foi alcançada com o plano posicionado a 850 mm, distância que corresponde ao centro do volume de medição calibrado. Observou-se uma degradação do desempenho nas extremidades do volume (desvio padrão de 0,15 mm a 700 mm e 0,25 mm a 1000 mm). Esse comportamento é uma limitação esperada para a arquitetura óptica adotada, uma vez que tanto o ajuste de foco das lentes quanto a minimização dos resíduos de calibração foram otimizados priorizando a região central de trabalho.

5.4 SUGESTÃO DE TRABALHOS FUTUROS

As limitações identificadas e as contribuições alcançadas fornecem uma base para direcionamentos futuros desta linha de pesquisa:

1. Mitigação da incerteza sistemática: sabendo que o gargalo metrológico do sistema é a calibração, sugere-se a pesquisa e investigação de métodos que minimizem o erro de escala da *baseline* (B). Isso pode ser feito utilizando artefatos de calibração 3D, como as barras esféricas (*ball bars*), ou a implementação de técnicas para refinar os parâmetros extrínsecos;

2. Otimização de desempenho para tempo real: visando viabilizar a aplicação do sistema em processos de inspeção dinâmica, é proposta a migração do núcleo de processamento da correlação espaço-temporal de Python para implementações consideradas de baixo nível, como C++. Esta otimização teria como objetivo reduzir o consumo de memória permitindo a reconstrução em tempo real;
3. Validação em ambiente subaquático: a justificativa inicial do projeto, que faz parte do escopo do Projeto VORIS, e a escolha do comprimento de onda relativo ao *laser* azul visam a aplicação subaquática. Um passo futuro é a validação experimental do sistema nesse ambiente, para quantificar o impacto da atenuação, do espalhamento e da refração na robustez da correlação, medida através do PNR, e na incerteza de medição;
4. Estudo da robustez da aquisição sob condições de vibração e movimento relativo simulando a dinâmica de flutuação de um ROV: recomenda-se investigar estratégias para a redução do número de quadros, sincronização de *hardware* mais robusta, otimizações computacionais, visando atingir taxas de reconstrução em tempo real.

REFERÊNCIAS

- BAY, H. et al. SURF: Speeded Up Robust Features. **Computer Vision and Image Understanding**, v. 110, n. 3, p. 346-359, 2008.
- BESL, P. J. Active, optical range imaging sensors. **Machine Vision and Applications**, 1(2), 127–152. doi:10.1007/bf01212277. 1988.
- CHIANG, J. Y.; CHEN, Y.; CHEN, Y. Underwater Image Enhancement: Using Wavelength Compensation and Image Dehazing (WCID). In: BLANC-TALON, Jacques et al. (Eds.). **Advanced Concepts for Intelligent Vision Systems**. Berlin: Springer, 2011. p. 372-383. (Lecture Notes in Computer Science, v. 6915).
- DEKIFF, M. et al. Single-shot 3D-acquisition with laser speckle projection. In: **PROCEEDINGS of the 7th International Conference on Computer Graphics, Virtual Reality, Visualisation and Interaction** in Africa, 2010, Franschhoek. Anais. New York: ACM, 2010. p. 77-80
- DI STEFANO, L.; MATTOCCIA, S.; TOMBARI, F., ZNCC-based template matching using bounded partial correlation, **Pattern Recognit Lett**, vol. 26, no. 14, pp. 2129–2134, Oct. 2005, doi: 10.1016/j.patrec.2005.03.02
- FORSMAN, M. **Point cloud densification**. Dissertação – Department of Physics, Umeå University, Umeå, 2010. Disponível em: <https://www.researchgate.net/publication/265190953_Point_cloud_densification>. Acesso em: 22. out. 2025.
- GENG, J., Structured-light 3D surface imaging: a tutorial, **Adv Opt Photonics**, vol. 3, no. 2, p. 128, Jun. 2011, doi: 10.1364/aop.3.000128.
- GONZALEZ, R. C.; WOODS, R. E. **Processamento Digital de Imagens**. 3. ed. São Paulo: Pearson Education do Brasil, 2010.
- HARTLEY, R; ZISSERMAN, A. **Multiple View Geometry in Computer Vision**. [S.l.]: Cambridge University Press, New York, 2004. (A Wiley-Interscience publication). ISBN 978-0-511-18618-9.
- HORN, E.; KIRYATI, N. Toward Optimal Structured Light Patterns. In: **INTERNATIONAL CONFERENCE ON COMPUTER VISION (ICCV)**, 7., 1999, Corfu. Proceedings. Corfu: IEEE, 1999. p. 166-173.
- KHAN, D.; SHIRAZI, M. A.; KIM, M. Y. Single shot laser speckle based 3D acquisition system for medical applications, **Opt Lasers Eng**, vol. 105, pp. 43–53, Jun. 2018, doi: 10.1016/j.optlaseng.2018.01.001.
- LEWIS, J. P. Fast normalized cross-correlation. In: **VISION INTERFACE**, 1995, Quebec. Proceedings. Quebec: Canadian Image Processing and Pattern Recognition Society, 1995.

LIU, T. et al. **Lightweight underwater image adaptive enhancement based on zero-reference parameter estimation network**. *Frontiers in Marine Science*. 11. 10.3389/fmars.2024.1378817. 2024.

LOWE, D. G. Distinctive image features from scale-invariant keypoints. **International Journal of Computer Vision**, v. 60, n. 2, p. 91-110, 2004.

LUHMANN, T. et al. **Close-Range Photogrammetry and 3D Imaging**, Berlin, Boston: De Gruyter, 2014. <https://doi.org/10.1515/9783110302783>

LUO, H. et al. Large-Scale 3D Reconstruction from Multi-View Imagery: A Comprehensive Review. **Remote Sensing** 16, no. 5: 773. <https://doi.org/10.3390/rs16050773>. 2024.

MATLAB. **Calibration Patterns**. Disponível em: <<https://www.mathworks.com/help/vision/ug/calibration-patterns.html>>. Acesso em: 13 ago. 2025.

MARR, D.; POGGIO, T. A computational theory of human stereo vision. **Proc. R. Soc. Lond. B**.204301–328. <http://doi.org/10.1098/rspb.1979.0029>. 1979.

MATACHE, G. et al. A Comparison between 3D Scanning and CMM Dimensional Inspection of Small Size Gas Turbine. **Advanced Materials Research**. 1128. 347-352. 10.4028/www.scientific.net/amr.1128.347. 2015.

MELLISH, B. **Principle of a pinhole camera**. Disponível em: <<https://commons.wikimedia.org/wiki/File:Pinhole-camera.png>>. Acesso em: 13 ago. 2025.

PINTO, T.; KOHLER, T. C.; ALBERTAZZI, A. Regular mesh measurement of large free form surfaces using stereo vision and fringe projection. **Optics and Lasers in Engineering**, Vol. 50, Issue 7, 2012, <http://dx.doi.org/10.1016/j.optlaseng.2012.03.003>.

PINTO, T.; ALBERTAZZI, A., 3d active stereo measurement in a regular mesh with random pattern and laser speckle projection, **ABCM Symposium Series in Mechatronics**, 2013.

REGNER, D., et al. Inverse Triangulation with Spatiotemporal Correlation using a Variable Pseudorandom Pattern Projector for 3D Stereo Measurement. In: **PHOTOMET**, 2025, MODENA.

RUBLEE, E. et al. ORB: An efficient alternative to SIFT or SURF. In: **IEEE INTERNATIONAL CONFERENCE ON COMPUTER VISION (ICCV)**, 2011, Barcelona. Proceedings. Barcelona: IEEE, 2011. p. 2564-2571.

SALVI, J.; PAGÈS, J.; BATTLE, J. (2004). Pattern codification strategies in structured light systems. **Pattern Recognition**, 37(4), 827–849.

SCHARSTEIN, D.; SZELISKI, R. A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms, In **International Journal of Computer Vision** (Vol. 47, Issue 3). 2002.

SONG, Z. et al. DOE-based structured-light method for accurate 3D sensing, **Opt Lasers Eng**, vol. 120, pp. 21–30, Sep. 2019, doi: 10.1016/j.optlaseng.2019.02.009.

STEGER, C. Unbiased extraction of curvilinear structures from 2D and 3D images. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 20, n. 2, p. 113-125, 1998.

STIVANELLO, M. E. **ALGORITMOS PARA MEDIÇÃO DE SUPERFÍCIES EM MOVIMENTO USANDO VISÃO 3D**. 2013. Tese (Doutorado em Engenharia de Automação e Sistemas) – Universidade Federal de Santa Catarina, Florianópolis, 2013.

SZELISKI, R. **Computer Vision: Algorithms and Applications**. London: Springer, 2010.

WINKELBACH, S.; MOLKENSTRUCK, S.; WAHL, F. M. Low-Cost Laser Range Scanner and Fast Surface Registration Approach. In: **EUROGRAPHICS WORKSHOP ON VIRTUAL REALITY AND AUGMENTED REALITY**, 3., 2006, Nicosia. Proceedings. Nicosia: Eurographics Association, 2006.

ZBONTAR, J; LECUN, Y. Stereo Matching by Training a Convolutional Neural Network to Compare Image Patches. **Journal of Machine Learning Research (JMLR)**, v. 17, p. 1-32, 2016.

ZHANG, L.; CURLESS, B.; SEITZ, S. M. Spacetime stereo: Shape recovery for dynamic scenes. In: **IEEE COMPUTER SOCIETY CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION (CVPR)**, 2003. Proceedings. Madison: IEEE, 2003. v. 2, p. 367-374.

ZHANG, S. **Handbook of 3D Machine Vision**. [S.l.]: Hoboken: Taylor and Francis, 2013. ISBN 1-4398-7219-8.

ZHANG, Z. A Flexible New Technique for Camera Calibration. *Pattern Analysis and Machine Intelligence*, **IEEE Transactions** on. 22. 1330 - 1334. 10.1109/34.888718. 2000.

ZHUANG, S. et al. The influence of active projection speckle patterns on underwater binocular stereo vision 3D imaging, **Opt Commun**, vol. 528, Feb. 2023, doi: 10.1016/j.optcom.2022.129014.