



CENTRO UNIVERSITÁRIO SENAI CIMATEC

**PROGRAMA DE PÓS-GRADUAÇÃO EM GESTÃO E TECNOLOGIA
INDUSTRIAL**

Doutorado em Gestão e Tecnologia Industrial

Tese de Doutorado

**Sistema computacional integrado a modelos digitais
para reconstrução de estruturas subaquáticas com
alta definição**

Apresentada por: Leonardo de Paula Nardy
Orientador: Prof. Dr. Herman Augusto Lepikson
Co-orientador: Prof. Dr. Oberdan Rocha Pinheiro

2022

Leonardo de Paula Nardy

**Sistema computacional integrado a modelos digitais
para reconstrução de estruturas subaquáticas com
alta definição**

Tese de Doutorado apresentada ao Programa de Pós-graduação em Gestão e Tecnologia Industrial, Curso de Doutorado em Gestão e Tecnologia Industrial do CENTRO UNIVERSITÁRIO SENAI CIMATEC, como requisito parcial para a obtenção do título de **Doutor em Gestão e Tecnologia Industrial**.

Área de conhecimento: Interdisciplinar

Orientador: Prof. Dr. Herman Augusto Lepikson

Coorientador: Prof. Dr. Oberdan Rocha Pinheiro

Salvador

2022

Ficha catalográfica elaborada pela Biblioteca do Centro Universitário SENAI CIMATEC

N223s Nardy, Leonardo de Paula

Sistema computacional integrado a modelos digitais para reconstrução de estruturas subaquáticas com alta definição / Leonardo de Paula Nardy. – Salvador, 2022.

122 f. : il. color.

Orientador: Prof. Dr. Herman Augusto Lepikson.
Coorientador: Prof. Dr. Oberdan Rocha Pinheiro.

Tese (Doutorado em Gestão e Tecnologia Industrial) – Programa de Pós-Graduação,
Centro Universitário SENAI CIMATEC, Salvador, 2022.
Inclui referências.


1. Reconstrução 3D. 2. Fusão de sensores. 3. Fusão de softwares. 4. Robótica subaquática. I. Centro Universitário SENAI CIMATEC. II. Lepikson, Herman Augusto. III. Pinheiro, Oberdan Rocha. IV. Título.

CDD 620.1

CENTRO UNIVERSITÁRIO SENAI CIMATEC**Doutorado em Gestão e Tecnologia Industrial**

A Banca Examinadora, constituída pelos professores abaixo listados, aprova a Defesa de Doutorado, intitulada “**Sistema computacional integrado a modelos digitais para reconstrução de estruturas subaquáticas com alta definição**” apresentada no dia 24 de novembro de 2022, como parte dos requisitos necessários para a obtenção do Título de Doutora em Gestão e Tecnologia Industrial.

Orientador:



Assinado eletronicamente por:
Herman Augusto Lepikson
CPF: ***.543.375-2*

Data: 25/11/2022 15:57:02 -03:00

Prof. Dr. Herman Augusto Lepikson
SENAI CIMATEC

Coorientador:

Assinado eletronicamente por:
Oberdan Rocha Pinheiro
CPF: ***.073.705-**
Data: 25/11/2022 09:04:14 -03:00

Prof. Dr. Oberdan Rocha Pinheiro
SENAI CIMATEC

Membro Interno:

Assinado digitalmente por:
Thiago Barros Murari
CPF: ***.984.148-7*

Data: 25/11/2022 09:16:19 -03:00

Prof. Dr. Thiago Barros Murari
SENAI CIMATEC

Membro Interno:

Assinado eletronicamente por:
Aloisio Santos Nascimento Filho
CPF: ***.459.665-**
Data: 25/11/2022 10:04:36 -03:00

Prof. Dr. Aloisio Santos Nascimento Filho
SENAI CIMATEC

Membro Externo:

Assinado eletronicamente por:
Josemar Rodrigues de Souza
CPF: ***.040.605-**
Data: 29/11/2022 10:35:23 -03:00

Prof. Dr. Josemar Rodrigues de Souza
UNEB

Membro Externo:

Electronically signed by:
Marcelo Henrique Souza Bomfim
CPF: ***.646.836-**
Date: 11/25/2022 1:36:10 PM -03:00

Prof. Dr. Marcelo Henrique Souza Bomfim
IFMG

Dedico este trabalho a todos que se propõem a trabalhar pelo desenvolvimento científico e tecnológico, pela geração e disseminação de conhecimento e que acreditam que a educação é o principal instrumento para proporcionar o bem-estar de um povo e para alcançar sua prosperidade.

Agradecimentos

Ao Prof. Oberdan agradeço pela orientação na realização deste trabalho, pelo incentivo constante, pela paciência, pelas sugestões, pelas condições proporcionadas para a realização deste trabalho e pela confiança depositada em mim ao longo dos últimos anos.

Ao Prof. Herman, pela orientação e apoio a este trabalho, pelas sugestões valiosas e pela condução, sempre coerente, da investigação e da elaboração desta tese de doutorado.

Aos professores do SENAI CIMATEC, pelo incentivo, críticas e questionamentos que contribuíram para a pesquisa e desenvolvimento deste trabalho.

Aos professores membros da banca examinadora, por aceitarem participar deste trabalho e pelas contribuições valiosíssimas na revisão de todo conteúdo.

A toda equipe de engenheiros e pesquisadores do SENAI CIMATEC que atuaram no projeto.

Ao SENAI CIMATEC, por proporcionar e incentivar a realização desta jornada.

À Petrobras/CENPES, por propiciarem a execução do projeto tecnológico e a inserção de minha tese em seu desenvolvimento.

Aos meus familiares, pelo apoio, incentivo e compreensão, ante a minha ausência e necessidade de dedicação em muitos momentos ao longo destes anos.

Agradeço em especial à minha esposa, Giulianna, e aos meus filhos, Clara e Antônio, pela paciência, compreensão, apoio, incentivo e por estarem sempre ao meu lado.

Por fim, a Deus, por ter me dado força, coragem e saúde para o cumprimento de mais esta etapa em minha vida.

Resumo

O desenvolvimento de pesquisas voltadas à inspeção de equipamentos submarinos e à digitalização desses ativos vem ganhando maior importância na medida em que a exploração de petróleo e gás evolui em águas cada vez mais profundas, além disto, há várias outras demandas para o uso de robôs e sistemas autônomos, como exploração arqueológica subaquática, monitoramento dos oceanos, estudos da biologia marinha e geológicos. Muitas vezes, a falta de informações geométricas precisas sobre os equipamentos submarinos utilizados pela indústria de petróleo e gás impõe dificuldades na realização das operações de inspeção e manutenção. Um desafio presente é o de reconstituição geometricamente precisa do ambiente onde se encontram esses equipamentos submarinos. Observa-se, a partir dos estudos sobre digitalização 3D subaquática, uma evolução nos métodos com o objetivo de melhorar a eficiência desse tipo de sistema, porém são necessários novos esforços de pesquisa, uma vez que estes sistemas ópticos são complexos. As nuvens de pontos formadas a partir de digitalizadores ópticos contêm pontos desorganizados e *outliers*, os quais prejudicam o processo de reconstrução da cena. Isto justifica o interesse no desenvolvimento de novas soluções tecnológicas para superar as dificuldades à obtenção de informação por meio do uso de métodos de aquisição ópticos. O trabalho desenvolvido nesta tese busca consolidar estas nuvens de pontos desorganizadas, além de fornecer uma representação tridimensional precisa de estruturas dos equipamentos subaquáticos. Para isso, um método de fusão de *softwares* baseados em *Structure from Motion* - SfM e *Multiview Stereo* é desenvolvido com o objetivo de minimizar erros introduzidos nas nuvens de pontos, em decorrência das perturbações de movimentação que o *Unmanned Underwater Vehicle* - UUV sofre no ambiente de operação. Os resultados alcançados durante os testes em águas profundas evidenciam que o sistema apresenta bons resultados, com erros de medição entre 0,47 e 4,12 milímetros para distâncias de medição entre 5 a 6 metros, com uma incerteza que varia entre +/- 0,81 a +/- 6,50 milímetros. A viabilidade de implantação deste sistema é confirmada pelos mapas de equipamentos digitalmente reconstruídos, obtidos durante a fase de testes. Portanto, esta tese traz uma nova e relevante contribuição para reconstrução 3D dessas estruturas com a exatidão dimensional necessária para as operações robotizadas e considerando as questões inerentes impostas por este ambiente.

Palavras-chaves: Reconstrução 3D; Fusão de sensores; Fusão de softwares; Robótica subaquática.

Abstract

The development of research focused on the inspection and scanning of subsea equipment is becoming increasingly important as oil and gas exploration evolves in deeper waters, in addition, there are several other demands for the use of robots and autonomous systems, such as underwater archaeological exploration, ocean monitoring, marine biology and geological studies. Frequently, the lack of accurate geometric information on the subsea equipment imposes difficulties in performing operations related to inspection and maintenance. A challenge presented is the geometrically accurate reconstitution of the environment where these subsea equipment are located. From studies on underwater 3D scanning, an evolution in methods is observed with the aim of improving the efficiency of this type of system, but new research efforts are needed, since these optical systems are complex. Point clouds formed from optical scanners contain noise and outliers, which hinder the scene reconstruction process. This justifies the interest in developing new technological solutions to overcome the difficulties in obtaining information through the use of optical acquisition methods. This thesis seeks to consolidate point clouds cluttered and provide an accurate three-dimensional representation of underwater equipment structures. For this purpose, a new data fusion method of merging software based on Structure from Motion - SfM and Multiview Stereo is developed with the goal of minimizing errors introduced into the point clouds as a result of motion disturbances in the operating environment. The results achieved during the deep-water tests highlight that the system presents good results, with measurement errors between 0.47 and 4.12 millimeters for measurement distances between 5 to 6 meters, and an uncertainty between +/- 0.81 and +/- 6.50 millimeters. The feasibility of implementing this system is confirmed by the digitally reconstructed equipment maps obtained during the testing phase. Therefore, this thesis brings a new and relevant contribution to the 3D reconstruction of these structures with the dimensional precision necessary for robotic operations and considering the inherent issues imposed by this environment.

Keywords: 3D reconstruction; fusion of sensors; fusion of softwares; underwater robotics.

Sumário

1	Introdução	1
1.1	Definição do problema	5
1.2	Objetivo Geral	9
1.3	Objetivos Específicos	9
1.4	Importância da pesquisa	9
1.4.1	Contribuições científicas do presente trabalho e o seu ineditismo	10
1.5	Aspectos Metodológicos	10
1.6	Organização da Tese de Doutorado	11
2	Revisão da Literatura	12
2.1	Digitalização 3D Subaquática	12
2.2	Calibração	19
2.3	Tecnologias de Sensores Ópticos Aplicados à Digitalização e Reconstrução 3D	22
2.3.1	Métodos de reconstrução 3D com câmeras	22
2.3.2	Métodos de reconstrução 3D com <i>laser</i>	24
2.3.3	Desempenho dos diferentes tipos de sensores	25
2.3.4	Avaliação sobre tecnologias atuais disponíveis	27
2.4	Algoritmos para reconstrução 3D	31
2.5	Considerações - correlação dos trabalhos e identificação das limitações	41
3	Projeto e desenvolvimento da pesquisa	44
3.1	Sistema Computacional - DIGISUB	44
3.1.1	Módulo de aquisição de imagem	48
3.1.2	Módulo de Detecção de <i>Features</i>	51
3.1.3	Módulo de reconstrução 3D	55
3.1.4	Arquitetura de <i>software</i>	56
3.2	Implementação	58
4	Avaliação e resultados experimentais	60
4.1	Testes na Baía de Aratu - ambiente relevante	60
4.1.1	Preparação dos testes em ambiente relevante	60
4.1.2	Resultados dos testes em ambiente relevante	64
4.2	Testes na Bacia de Campos - ambiente operacional	67
4.2.1	Preparação dos testes em ambiente operacional	67
4.2.2	Avaliação da reconstrução	70
4.2.3	Desempenho computacional	75
4.3	Análise do sistema híbrido DIGISUB comparado à referência estabelecida na pesquisa sistemática	79
4.4	Impacto do resultado da pesquisa em aplicações industriais	80
5	Considerações Finais	83
5.1	Conclusões	83
5.2	Contribuições	86
5.3	Recomendação de atividades para futuras pesquisas	87

Referências	88
A Projeto detalhado de <i>hardware</i> do protótipo	99
A.0.1 Projeto Eletroeletrônico	102
A.0.2 Projeto dos Encapsulamentos	103

Lista de Tabelas

2.1	Pesquisa Sistemática	12
2.2	Escâneres 3D ópticos ativos subaquáticos atuais.	30
2.3	Trabalhos selecionados	42
4.1	Erro de medição do sistema alcançado nos testes da Baía de Aratu.	66
4.2	Dimensões do painel de <i>back-up</i> da ANM.	70
4.3	Comparação dimensões do painel de <i>back-up</i> da ANM.	71
4.4	Comparação resultados das cotas Baía de Aratu X Bacia de Campos.	72
4.5	Resultado do desempenho computacional.	75
4.6	Comparação resultados entre vistas parcial e total da ANM.	79
A.1	Componentes de hardware do protótipo.	102

Lista de Figuras

1.1	Ilustração da operação de um ROV em águas ultraprofundas	2
1.2	Fotografia da operação de um ROV em águas ultraprofundas	2
1.3	Árvore de Natal Molhada sendo montada em um fabricante	3
1.4	Organização dos projetos de desenvolvimento tecnológico demandado pela Petrobras e o desenvolvimento da pesquisa desta tese.	5
2.1	Classificação das Técnicas de Digitalização 3D.	15
2.2	Espectro de absorção da luz na água.	19
2.3	Refração dos raios de luz nas portas plana e de domo, respectivamente. . .	21
2.4	Esquema de um sensor ToF 3D.	23
2.5	Mudança na direção de uma linha de <i>laser</i> em função do ângulo de incidência na água e da dupla refração.	26
2.6	Ângulo de incidência da janela de visualização do <i>laser</i>	27
2.7	<i>Pipeline</i> de detecção e descritores de <i>features</i>	36
3.1	Arquitetura do sistema computacional - DIGISUB	46
3.2	Fluxograma do sistema computacional - DIGISUB	47
3.3	Processo de calibração e retificação da imagem - Pinax	51
3.4	Descrição do passo Detecção de <i>Features</i>	53
3.5	Arquitetura de <i>software</i> do sistema.	56
3.6	Esquema de comunicação.	57
3.7	Visão de pacotes e subsistemas da aplicação.	58
3.8	Visão de implantação da aplicação.	59
4.1	Embarcação rebocador utilizado para testes na Baía de Aratu.	60
4.2	Mapa Baía de Aratu, BA	61
4.3	<i>Mock-up</i> da cabeça de poço	61
4.4	Protótipo montado	62
4.5	Visão do mergulhador no cenário de testes na Baía de Aratu.	62
4.6	Cenário de execução dos testes na Baía de Aratu.	63
4.7	Preparação dos testes do sistema.	63
4.8	Avarias provocadas no <i>mock-up</i>	64
4.9	Nuvens de pontos esparsas.	64
4.10	Nuvens de pontos densas.	65
4.11	Erro absoluto de medição do sistema	66
4.12	Erro relativo de medição do sistema	67
4.13	Encapsulamentos do protótipo instalados no ROV.	68
4.14	ROV FCV3000.	68
4.15	RSV Parcel do Bandolim.	69
4.16	Cenário de execução dos testes na Bacia de Campos.	69
4.17	Desenho técnico com as principais cotas da ANM digitalizada.	70
4.18	Nuvem de pontos gerada pelo método DIGISUB.	71
4.19	Reconstrução completa da parte frontal ANM pelo método DIGISUB. . . .	74
4.20	Imagem da câmera esquerda do protótipo.	75
4.21	Nuvem de pontos gerada pelo método KLT.	76

4.22	Nuvem de pontos gerada pelo método DDA.	76
4.23	Nuvem de pontos gerada pelo método DIGISUB.	77
4.24	Resultados dos métodos KLT, DDA e DIGISUB.	78
A.1	Conceito preliminar da solução.	99
A.2	Detalhe dos componentes do protótipo.	100
A.3	Dimensões do protótipo montado.	101
A.4	Arquitetura Geral do Protótipo.	101
A.5	Arquitetura Eletroeletrônica.	103
A.6	Encapsulamento óptico.	104
A.7	Encapsulamento de processamento e potência.	104
A.8	Encapsulamento da câmera.	104
A.9	Desenho técnico de construção do encapsulamento de processamento.	105
A.10	Desenho técnico de construção do encapsulamento de projeção de luz estruturada.	106
A.11	Desenho técnico de construção do encapsulamento da câmera.	107

Lista de Abreviaturas e Siglas

3D	Tridimensional ou Terceira Dimensão
ADC	<i>Analog to Digital Converter</i>
ANP	Agência Nacional do Petróleo, Gás Natural e Biocombustíveis
ANM	Árvore de Natal Molhada
AOD	<i>Acousto-Optic Deflector</i>
ASIFT	<i>Affine-SIFT</i>
CAD	<i>Computer-aided design</i>
CCD	<i>Charge-coupled device</i>
CMOS	<i>Complementary metal-oxide-semiconductor</i>
DAQ	<i>Unidade de Aquisição de Dados</i>
DC	<i>Direct Current</i>
DDA	<i>Detect-Describe-Associate</i>
DOE	<i>Diffraction Optical Element</i>
DoG	<i>Difference of Gaussian</i>
DTU	<i>Database Transaction Unit</i>
EMBRAPII .	Empresa Brasileira de Pesquisa e Inovação Industrial
EOD	<i>Electro-Optic Reflector</i>
FAST	<i>Features from Accelerated Segment Test</i>
FPGA	<i>Field Programmable Gate Arrays</i>
GPU	<i>Graphics Processing Unit</i>
ICP	<i>Iterative Closest Point</i>
IMU	<i>Inertial Measurement Unit</i>
INS	<i>Inertial Navigation System</i>
IP	<i>Interesting Points</i>
KLT	<i>Kanade-Lucas-Tomasi</i>
LED	<i>Light-emitting diode</i>
LLS	<i>Laser Line Scan</i>
LM	<i>Levenberg-Marquardt</i>
LoG	<i>Laplacian of Gaussian</i>
LTS	<i>Laser Triangulation Sensor</i>
MEMS	<i>Microelectromechanical System</i>
MVS	<i>Multiview Stereo</i>
N/A	Não Aplicável
NUC	<i>Computador Central Embarcado</i>
O&G	Petróleo & Gás, do inglês <i>Oil & Gas</i>
OPA	<i>Optical Phased Array</i>
PCA	<i>Principal Component Analysis</i>
PCL	<i>Point Cloud Library</i>
PCCS	<i>Point Cloud Semantic Segmentation</i>
PLLS	<i>Pulsed Laser Line Scan</i>
PLY	<i>Polygon File Format</i>
PoE	<i>Power over Ethernet</i>
RGB	<i>Red, Green, Blue</i>
ROS	<i>Robotics Operational System</i>
ROV	<i>Remotely Operated Underwater Vehicle</i>
RSV	<i>ROV Support Vessel</i>

S.A.	Sociedade Anônima
SfM	<i>Structure from Motion</i>
SGM	<i>Semi-global Matching</i>
SIFT	<i>Scale Invariant Feature Transform)</i>
SSD	<i>Sum of Squared Differences</i>
SURF	<i>Speeded-Up Robust Features</i>
SLAM	<i>Simultaneous Localization and Mapping</i>
ToF	<i>Time of Flight</i>
USB	<i>Universal Serial Bus</i>
UT	<i>Ultrasonic Testing</i>
UUV	<i>Unmanned Underwater Vehicle</i>
Vca	Tensão em corrente alternada
Vcc	Tensão em corrente contínua
WCT	<i>Wet Christmas Tree</i>

Introdução

Os modelos digitais de alta definição tridimensionais (3D) de equipamentos e estruturas subaquáticos, embarcados em robôs e sistemas autônomos, são úteis para as empresas que operam equipamentos que necessitam de intervenções regulares, como é o caso da indústria de petróleo & gás (O&G). Além destas aplicações, há várias outras demandas para o emprego desta tecnologia, como exploração arqueológica subaquática, monitoramento dos oceanos, estudos da biologia marinha e geológicos. As instalações *offshore*, em particular, apresentam desafios específicos, consideradas as grandes profundidades, as distâncias entre os equipamentos no fundo do mar, a limitação da presença de mergulhadores à profundidade de 300 m e sua mobilidade limitada (CASTILLÓN *et al.*, 2019). Essas condições impõem que as atividades de inspeção e manutenção confiável em equipamentos instalados em águas profundas¹ e ultraprofundas², por exemplo, só sejam possíveis com o apoio de *Underwater Unmanned Vehicles* (UUVs). UUVs abrangem o grupo de veículos subaquáticos não tripulados, onde se enquadram os AUVs (*Autonomous Underwater Vehicles*) e os ROVs³ (*Remotely Operated Vehicles*), apropriados e viáveis para executar tais tarefas (YU *et al.*, 2019). As Figuras 1.1 e 1.2 ilustram o ambiente de operação de um ROV em águas ultraprofundas.

A instalação dos equipamentos no poço e no solo marinho e a operação dos equipamentos e dutos que executam a extração, o controle e a condução dos fluxos de petróleo e gás no mar são dificultadas pela completa invisibilidade⁴ das operações (LERØEN, 2005). A utilização de veículos são essenciais nas práticas rotineiras de inspeção à detecção de pontos de corrosão, fissuras e danos em equipamentos, com vistas à realização de reparos. Porém, estes veículos têm um custo operacional relevante, estimado entre 100 mil e 150 mil dólares por dia (ZAGATTI *et al.*, 2018) e qualquer tecnologia que permita uma redução ou otimização do uso destes equipamentos podem trazer impactos positivos sobre o custo operacional para as empresas do setor, além de confiabilidade e segurança.

¹Águas profundas: Profundidade de 401 a 1.500 metros.

²Águas ultraprofundas: Profundidade acima de 1.500 metros.

³Os ROVs são conectados a uma embarcação por um cabo umbilical (*tether*) para transmissão de potência e controle e são comandados à distância por técnicos especializados nas salas de operação de embarcações específicas para prestar tal tipo de serviço.

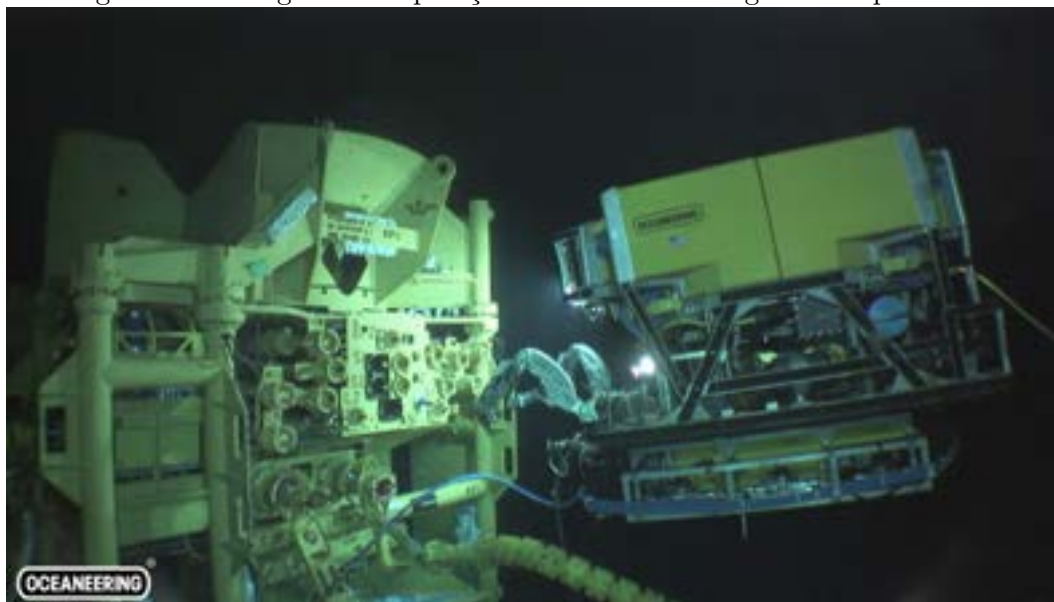
⁴A invisibilidade é decorrente da falta da luz solar em áreas profundas e da turbidez da água, em razão da presença de materiais em suspensão.

Figura 1.1: Ilustração da operação de um ROV em águas ultraprofundas



Fonte: [Morais \(2013\)](#).

Figura 1.2: Fotografia da operação de um ROV em águas ultraprofundas



Fonte: [Oceaneering](#).

Já a Figura 1.3 mostra uma Árvore de Natal Molhada (ANM), ou em inglês *Wet Christmas Tree* (WCT), um dos principais equipamentos utilizados na exploração de petróleo *offshore*, durante o processo de montagem em um fabricante.

Figura 1.3: Árvore de Natal Molhada sendo montada em um fabricante



Fonte: PetroNotícias.

Nos últimos anos a inspeção subaquática obteve um grande avanço no desenvolvimento tecnológico dos sistemas de visão, que têm se tornado cada vez mais sofisticados por meio de novas técnicas de calibração (SHORTIS, 2015) e desenvolvimento de câmeras para aplicação subaquática (PALOMER et al., 2019). Técnicas de visão computacional baseadas em abordagens estéreo têm oferecido métodos para reconstruções 3D com alta precisão e resolução para aplicações terrestres. No entanto, em ambiente subaquático, a turbidez, o retroespalhamento, a degradação da imagem causada por absorção da luz, a iluminação irregular causada por iluminação artificial, além da própria calibração da câmera trazem uma série de desafios que normalmente não estão presentes para os sistemas de visão no ar. Segundo (WANG et al., 2020), estes fatores contribuem para degradação da qualidade das imagens captadas impactando na precisão do processo de reconstrução 3D, tornando-se, então, um desafio a enfrentar. Além disso, apesar de o sistema de visão estéreo ser um método de fácil aplicação, por possuir uma configuração de *hardware* simples e o processo de calibração estéreo ser direto, no ambiente subaquático ele não pode ser aplicado a superfícies metálicas planas com limitadas informações de textura (KIM; CHOI; KO, 2020; KONOLIGE, 2010), o que exige sua combinação com luz estruturada, que cria formas geométricas na superfície alvo. Outra característica é que este método (visão estéreo mais luz estruturada) demanda um grande esforço computacional e gera uma nuvem de pontos com alta densidade, especialmente em cenários de baixo contraste (CASTILLÓN et al., 2019). Portanto, os sistemas atuais utilizados apresentam muitas limitações relacionadas à resolução adequada e aos métodos de tratamento dos dados, o que será visto em maiores

detalhes no Capítulo 2.

A partir dos estudos sobre digitalização 3D subaquática, pode-se verificar uma evolução nos métodos com o objetivo de melhorar a eficiência desse tipo de sistema, porém são necessários novos esforços de pesquisa, uma vez que estes sistemas ópticos são complexos. Segundo Berger et al. (2016), nuvens de pontos criadas provenientes de digitalizador a *laser* contêm pontos desorganizados com ruídos e *outliers*, sem estrutura inerente ou informações de orientação, o que prejudica o processo de reconstrução do objeto digitalizado. Isto justifica o interesse no desenvolvimento de novas soluções tecnológicas para superar as dificuldades à obtenção de informação por meio do uso de métodos de aquisição ópticos. O desenvolvimento apresentado nesta tese busca consolidar nuvens de pontos desorganizadas por ruídos e *outliers* e que forneça uma representação tridimensional precisa de estruturas dos equipamentos subaquáticos.

Este trabalho de investigação está inserido no projeto de Pesquisa, Desenvolvimento e Inovação (P&D) intitulado: "Automação de Operações com ROV: Digitalização Submarina e Análise de Viabilidade", executado pelo SENAI CIMATEC sob demanda da Petrobras e com recursos financeiros provenientes da empresa a partir da obrigatoriedade de investimentos em pesquisa determinada pela ANP - Agência Nacional de Petróleo, Gás Natural e Biocombustíveis, com registro nº 20948-6, além de recursos da EMBRAPPII - Empresa Brasileira de Pesquisa e Inovação Industrial. O objetivo estipulado para tal projeto foi: "Desenvolver protótipo que forneça nuvem de pontos de alta resolução para identificar, com maior exatidão, estruturas submarinas, por meio da aquisição e cruzamento de dados de sensores diversos, além de verificar a viabilidade técnica-econômica de automatizar operações com manipuladores de ROV". Do objetivo foi definido o acrônimo para o projeto, DIGISUB, proveniente de **D**igitalizador **S**ubmarino.

Portanto, a definição de um conceito de um sistema de digitalização, elaboração do projeto detalhado e construção de um protótipo e execução de testes em ambiente relevante foram as etapas principais do referido projeto, culminado com a geração de nuvens de ponto brutas. Então, o propósito desta tese busca ampliar o escopo do projeto e, por meio de uma nova rota para obtenção das nuvens brutas, gerar a reconstrução 3D de estruturas instaladas no leito do oceano, com execução de testes e geração de *dataset* em ambiente operacional. Abaixo segue uma ilustração que demonstra a relação entre as duas frentes de desenvolvimento (Figura 1.4).

Figura 1.4: Organização dos projetos de desenvolvimento tecnológico demandado pela Petrobras e o desenvolvimento da pesquisa desta tese.



Fonte: Autor.

1.1 Definição do problema

Devido às grandes distâncias entre os equipamentos no fundo do mar e as plataformas de petróleo, sabendo-se que a presença humana por meio do trabalho de mergulhadores está limitada à profundidade de 300 metros e com todos os riscos inerentes envolvidos, UUVs têm sido cada vez mais utilizados para atividades de operação, inspeção ou intervenção em equipamentos. Já as atividades e operações em profundidades maiores que 300 metros são possíveis somente com o apoio de UUVs. Em todas essas operações com UUVs o conhecimento prévio detalhado e preciso (ordem de grandeza: milimétrica) das instalações e equipamentos subaquáticos, considerando que estão inseridos em um ambiente em constante mutação, é altamente relevante para o adequado planejamento e execução das atividades necessárias a estas operações. As empresas do setor de petróleo têm, portanto, grande interesse em realizar tarefas de digitalização tridimensional dos equipamentos subaquáticos, com eficiência e exatidão, em lugares onde a presença humana é difícil, arriscada ou até mesmo impossível. Este interesse está relacionado ao monitoramento estrutural dos equipamentos subaquáticos como forma de aumentar a integridade, prolongar o tempo de vida útil e melhorar a eficiência da produção. As informações dimensionais adquiridas durante as operações de inspeção permitem avaliar o estado dos equipamentos, bem como identificar possíveis falhas e evitar consequências de maior gravidade. Embora a maioria dos equipamentos utilizados pelas indústrias *offshore* tenha sido projetada usando técnicas de CAD (*computer-aided design*), na maior parte dos casos o modelo inicial representa um *design* funcional, e não a situação final construída (*as-built*). Além disso, esses equipamentos são instalados em ambientes muito dinâmicos (fundo do mar), onde adequações e ajustes são avaliados constantemente de forma a ampliar a segurança, aumentar a eficiência e evitar as emissões e/ou vazamentos perigosos, de acordo com a regulamentação ambiental. Ademais, por conta da instalação gradual de novos equipamentos e sistemas, a localização relativa desses ativos torna-se imprescindível para que operações com UUV sejam efetivas e seguras. Como resultado, após algum tempo

passa a existir diferenças significativas entre os modelos projetados e a situação como construída. Para equipamentos antigos instalados há anos a situação é ainda pior, pois a maioria deles foi projetada utilizando técnicas antigas, muitas ainda em papel, não havendo modelo 3D disponível. Há ainda que se considerar que a instalação de equipamentos implica em anexar acessórios, componentes de ligação, conectores, sistemas de proteção, entre outros, os quais impactam no conjunto, principalmente em como os equipamentos de inspeção podem se aproximar e se relacionar com as instalações. Consequentemente, em qualquer dos casos, para que novas alterações ou intervenções possam ser implementadas, são necessárias informações tridimensionais precisas desses equipamentos. Nesse sentido, as tecnologias de digitalização de alta definição capazes de adquirir perfis 3D precisos dos equipamentos apresentam-se como uma solução necessária para suprir esta lacuna. Portanto, destaca-se a demanda premente da indústria de petróleo *offshore* por tais soluções tecnológicas.

Os robôs subaquáticos requerem normalmente um grande número de módulos, cada um deles destinado a cumprir um requisito diferente: IMU (*Inertial Measurement Unit*) para operar navegação, motores para empuxo e deslocamento, modems acústicos ou ópticos para comunicação, entre outros. Além de tudo isso, uma habilidade chave de um sistema robótico autônomo é a capacidade de sentir seu ambiente (*perception*) (CASTILLÓN *et al.*, 2019). Para UUVs é especialmente importante adquirir dados 3D de seu entorno para realizar tarefas como reconhecimento de objetos, inspeção, manipulação e/ou navegação. Segundo Menna, Agrafiotis e Georgopoulos (2018), a maioria das técnicas utilizadas para digitalização 3D subaquática é baseada em sensores ópticos e acústicos, adaptados para atender às necessidades de determinadas aplicações, normalmente para águas rasas, profundas e ultraprofundas. De acordo com Wang *et al.* (2019), um sistema óptico oferece um resultado e informações mais detalhadas de uma estrutura imersa do que um sistema acústico. Diante disto, propôs um sistema de visão estéreo⁵ subaquática para aplicação em SLAM (do inglês *Simultaneous Localization and Mapping*) de um AUV, além de outras aplicações oceânicas. Castillón *et al.* (2019) ratificam que a maioria dos sistemas de detecção 3D na literatura são baseados em sinais acústicos (SONAR) ou de luz (LiDAR). Os sonares podem funcionar a um alcance muito maior, de até alguns milhares de metros, e não são afetados pela turbidez da água. Por outro lado, os sensores ópticos fornecem uma resolução lateral e taxa de atualização muito maiores, porém em distâncias curtas. Seu curto alcance, normalmente alguns metros, não limita o desempenho do UUV para tarefas de intervenção, uma vez que o robô precisa se aproximar do alvo. Para Skarlatos *et al.* (2016), a digitalização 3D e técnicas de modelagem representam um inestimável conjunto de ferramentas eficazes para reconhecimento, documentação e monitoramento de ativos.

⁵Visão estéreo é o ramo da visão computacional que analisa o problema da reconstrução da informação tridimensional de objetos a partir de um par de imagens capturadas simultaneamente, mas com um pequeno deslocamento lateral.

Teste ultrassônico, ou do inglês *Ultrasonic Testing* - UT, é uma das técnicas mais populares que têm sido utilizadas para inspeção de estruturas subaquáticas em O&G. Os sistemas UT são baseados em sensores que podem emitir uma onda de ultrassom e recebê-la após sua propagação no material. A diferença de tempo, entre os sinais emitidos e recebidos, é utilizada para identificar corrosão ou falhas na região inspecionada (LEDEZMA et al., 2015). Esta tecnologia fornece resultados de alta precisão e também pode ser utilizada para medir a espessura da parede, no entanto, apenas um único ponto pode ser inspecionado por vez e defeitos ao longo das linhas de onda não podem ser identificados (YU et al., 2019). Mergulhadores comumente realizam inspeções UT, mas o desenvolvimento de ferramentas que podem ser conectadas a um ROV de classe de inspeção é uma realidade e sua utilização está se tornando mais frequente para operações na zona de águas próximas à superfície (TECCHIO et al., 2021).

Caimi et al. (2008) afirmam que obter informações satisfatórias de objetos e instalações submarinas tem sido difícil devido às propriedades de absorção e dispersão da luz na água do mar. Daí, a atenuação desses efeitos tem sido foco de pesquisas há algum tempo e os avanços em métodos de *hardware*, *software* e algoritmos permitiram melhorias consideráveis desse tipo de sistema. Para Castellón et al. (2019) um dos principais desafios da imagem subaquática é que a luz é fortemente atenuada pela água. Somado a isso existem outros fenômenos que degradam essas imagens, como tremulação e maior retrodifusão devido à presença de partículas suspensas na água. Os sistemas de visão são geralmente fechados dentro de uma caixa selada (*housing*) com uma janela de visualização transparente. Isso implica que a luz sofre o processo de refração duas vezes antes de chegar à câmera. Isso traz ainda mais dificuldade para o cálculo da posição 3D do objeto observado (CASTILLÓN et al., 2019). Outra questão fundamental para sistemas de visão é processo de calibração (CASTILLÓN et al., 2019), o qual tem sido amplamente estudado na literatura. Shortis (2015) apresentou uma pesquisa bastante completa sobre técnicas de calibração e considerações para sistemas fotogramétricos subaquáticos. Dadas as soluções que vêm sendo discutidas e estudadas na literatura para aplicação de sistemas de visão ao meio subaquático, especificamente para a indústria de petróleo *offshore*, a combinação de precisão e distância entre o sensor (embarcada num veículo) e o objeto mostra-se um importante fator para a viabilidade técnica da execução de intervenções nos equipamentos. Castellón et al. (2019) apresentam uma comparação entre os resultados alcançados pelos estudos mais recentes (a partir de 2015), assim como pelos equipamentos já disponíveis comercialmente no mercado. É percebido nos diversos trabalhos analisados que a realização de testes em condições e ambientes de maior relevância, ou seja, mais próxima das condições reais de operação, tem se mostrado como o próximo passo recomendado por Skinner, Iscar e Johnson-Roberson (2017), como forma de dar robustez e confiabilidade aos métodos propostos até então.

Nos últimos anos, houve um grande avanço no desenvolvimento tecnológico desses siste-

mas, que têm se tornado cada vez mais sofisticados, por exemplo: técnicas de calibração (SHORTIS, 2015) e câmeras para aplicação subaquática (PALOMER et al., 2019). Técnicas de visão computacional baseadas em abordagens estéreo oferecem métodos para reconstruções 3D com alta precisão e resolução para aplicações terrestres. No entanto, a turbidez, retroespalhamento, degradação da imagem causada por absorção, iluminação irregular causada por iluminação artificial e calibração da câmera e sistema trazem uma série de desafios que normalmente não são apresentados nos sistemas de visão no ar. Segundo Wang et al. (2020), estes fatores contribuem para degradação da qualidade das imagens captadas impactando na precisão do processo de reconstrução 3D. De acordo com Kim, Choi e Ko (2020), o sistema de visão estéreo é um método de fácil aplicação, pois possui uma configuração de *hardware* simples e o processo de calibração estéreo para determinar os parâmetros extrínsecos entre as duas câmeras é direto. No entanto, no ambiente subaquático ele não pode ser aplicado a superfícies metálicas planas que apresentam falta de informações de textura, pois os pontos característicos aplicáveis não podem ser encontrados e rastreados para fazer a correspondência correta (KONOLIGE, 2010; KIM; CHOI; KO, 2020). Ainda de acordo com Kim, Choi e Ko (2020), o método usando a visão estéreo e luz estruturada pode ser usado para a medição de forma 3D de áreas sem informações de textura porque a luz estruturada projetada cria formas geométricas na superfície alvo. Além disso, a luz estruturada é usada apenas para medir a disparidade entre as duas câmeras e, portanto, não há necessidade de corrigir qualquer refração causada por ela atravessar vários meios diferentes quando submersa (ar, lente, água). Conseqüentemente, esta abordagem não requer calibração extrínseca entre o projetor de padrões e o sistema estéreo, sendo necessária apenas a calibração entre as duas câmeras estéreo. Outra característica é que este método (visão estéreo e luz estruturada) demanda um grande esforço computacional e gera uma nuvem de pontos com alta densidade, especialmente em cenários de baixo contraste (CASTILLÓN et al., 2019).

Observa-se, portanto, a partir dos estudos sobre digitalização 3D subaquática, uma evolução nos métodos com o objetivo de melhorar a eficiência desse tipo de sistema, porém são necessários novos esforços de pesquisa, uma vez que estes sistemas ópticos são complexos. As nuvens de pontos formadas a partir de digitalizadores ópticos contêm pontos desorganizados e *outliers*, os quais prejudicam o processo de reconstrução da cena. Isto justifica o interesse no desenvolvimento de novas soluções tecnológicas para superar as dificuldades à obtenção de informação por meio do uso de métodos de aquisição ópticos. O desenvolvimento proposto aqui pretende consolidar nuvens de pontos desorganizadas por ruídos e *outliers*. Para isso, um modelo de sistema computacional baseado em SfM (*Structured from Motion*) e MSV (*Multiview Stereo*) é proposto e desenvolvido com o objetivo de minimizar erros introduzidos nas nuvens de pontos decorrentes das perturbações de movimentação que o UUV sofre no ambiente de operação.

1.2 *Objetivo Geral*

Desenvolver um modelo apto a fornecer representação tridimensional de estruturas dos equipamentos subaquáticos, com precisão milimétrica, para aplicação focada nas demandas do setor de exploração e produção de petróleo e gás *offshore*.

1.3 *Objetivos Específicos*

Os objetivos específicos são:

- Definir um modelo eficaz para o mapeamento tridimensional de estruturas subaquáticas com erro milimétrico.
- Propor e implementar algoritmos para o mapeamento tridimensional proposto.
- Implementar uma arquitetura modular de *software* adequada à implementação da metodologia para o sistema de digitalização.
- Produzir um *dataset* composto por nuvens de pontos de estruturas representativas em ambiente subaquático.
- Testar e validar o modelo proposto em ambiente operacional relevante.

1.4 *Importância da pesquisa*

O desenvolvimento de pesquisas voltadas para inspeção de equipamentos subaquáticos vem ganhando importância devido à alta demanda por soluções necessárias para suporte à exploração e produção de petróleo e gás em águas profundas e ultraprofundas. A utilização de UUVs, principalmente ROVs, destaca-se como meio para execução de tarefas de inspeção subaquática, no entanto não é trivial obter dados confiáveis da condição dos equipamentos instalados no fundo do oceano, isto em função das características do ambiente que envolve esse tipo de operação, como: a turbidez, o retroespalhamento, a degradação da imagem causada por absorção da luz, a iluminação irregular causada por iluminação artificial, além da própria calibração da câmera. Também contribuem para a complexidade do desafio o fato das superfícies metálicas planas apresentarem baixa textura e demandarem o emprego de luz estruturada, o que incrementa fortemente o esforço computacional para processamento das informações obtidas.

A atenuação desses efeitos tem sido foco de pesquisa já há algum tempo, mas os recentes avanços nos métodos de *hardware*, *software* e algoritmos passaram a viabilizar melhorias

notáveis no alcance operacional desse tipo de sistema. Porém, os sistemas atuais apresentam muitas limitações relacionadas à resolução adequada e aos métodos de tratamento dos dados, o que será visto em maiores detalhes no capítulo seguinte.

1.4.1 *Contribuições científicas do presente trabalho e o seu ineditismo*

Pelos motivos mencionados acima, o desenvolvimento de um sistema computacional que forneça uma representação em alta resolução das estruturas dos equipamentos subaquáticos, com erro milimétrico e obtida a maiores distâncias que àquelas alcançadas nos trabalhos recentes levantados ganha relevância e prioridade tanto para a indústria, quanto para o meio acadêmico. A aquisição e cruzamento de dados de sensores ópticos além da fusão de *softwares* mostram-se promissoras para uso na inspeção e intervenção em equipamentos subaquáticos (ativos, válvulas, dutos/*risers* dentre outros) contribuindo para uma melhor gestão da integridade destes sistemas.

Esta pesquisa, portanto, traz novas contribuições para o estado da arte e representa um avanço tecnológico na área de investigação. Busca superar as dificuldades à obtenção de informação por meio do uso de métodos de aquisição atualmente em uso. O desenvolvimento proposto aqui tem o objetivo de consolidar nuvens de pontos desorganizadas por ruídos e *outliers* e, para isso, propõe um modelo de fusão de dados para minimizar os erros introduzidos nas nuvens de pontos, em decorrência das perturbações de movimentação comuns ao ambiente de operação.

1.5 *Aspectos Metodológicos*

Durante esta pesquisa foram selecionados prioritariamente artigos publicados a partir de 2015 a 2020 nas seguintes bases de dados: ACM Digital Library (ACM), IEEE Xplore Digital Library (IEEE Xplore), MDPI Sensors, Scientific Electronic Library Online (SciELO), Springer Link (Springer). Foram utilizados como descritores: Reconstrução 3D, fusão de sensores, fusão de software e robótica subaquática, não havendo restrições em relação aos idiomas espanhol, inglês e português. O detalhamento da pesquisa sistemática pode ser encontrado no Capítulo 2.

Inicialmente buscou-se a definição de sistemas de digitalização subaquática, suas aplicações e identificação dos principais desafios tecnológicos relacionados ao tema. Em seguida, foi trazido o estado da arte caracterizando os sistemas desenvolvidos em diversas pesquisas, seus desafios relacionados, suas diferentes configurações e arranjos de componentes, além das soluções já disponíveis comercialmente.

Posteriormente à concepção, projeto e montagem do protótipo, deu-se início às etapas de testes de integração e experimentos relacionados aos algoritmos de controle e comando, seguido da geração de *dataset*, o que, oportunamente, serviu de plataforma de testes da proposta de modelo de algoritmo para reconstrução 3D.

Uma vez superada a etapa de testes em laboratório, partiu-se para as etapas de testes em ambiente relevante e testes em ambiente operacional. No que tange a esta pesquisa, ambas proporcionaram condições singulares para teste do protótipo em si, da aplicação e do aprimoramento do modelo de algoritmo de reconstrução 3D, bem como da geração de *dataset* específico para desenvolvimento da pesquisa.

A demonstração dos resultados foi realizada a partir dos dados obtidos em cada uma das etapas de testes, proporcionando a comparação entre os resultados obtidos com diferentes técnicas e equipamentos.

1.6 Organização da Tese de Doutorado

O conteúdo dos capítulos seguintes da tese encontra-se organizado da seguinte forma:

- Capítulo 2: é apresentado o estado da arte em relação à digitalização 3D subaquática, passando pelas diferentes configurações e arranjo de componentes, técnicas de calibração, métodos de reconstrução 3D, seus algoritmos e uma avaliação das tecnologias já disponíveis comercialmente.
- Capítulo 3: são apresentados os projetos de *hardware* e *software* do protótipo desenvolvido e, principalmente, do sistema computacional do algoritmo para reconstrução 3D propriamente dito.
- Capítulo 4: trata as sequências de testes e validações em diferentes ambientes e condições, além de uma discussão sobre os resultados obtidos na pesquisa comparando-os com os resultados de diferentes técnicas de reconstrução 3D a partir do *dataset* gerado ao longo da pesquisa.
- Capítulo 5: traz as considerações finais sobre o trabalho, encerra as discussões técnicas da pesquisa, trazendo os principais ganhos que a proposta oferece, seus limites e limitações, pontos de melhorias, além de sugerir atividades para futuras pesquisas.

Revisão da Literatura

Neste capítulo são apresentados os detalhes e resultados da pesquisa sobre o estado da arte em relação à digitalização 3D subaquática, passando pelas formas construtivas e integração de componentes, pelas técnicas de calibração de câmeras e dispositivos, métodos de reconstrução 3D e seus algoritmos, seguido de uma avaliação das tecnologias já disponíveis comercialmente e, por fim, uma correlação dos trabalhos mais relevantes e identificação das limitações.

Durante este estudo foram selecionados prioritariamente artigos publicados entre 2015 e 2020 nas bases de dados: ACM Digital Library (ACM), IEEE Xplore Digital Library (IEEE Xplore), MDPI Sensors, Scientific Electronic Library Online (SciELO), Springer Link (Springer). Foram utilizados como descritores: Reconstrução 3D, Fusão de sensores, Fusão de softwares e Robótica subaquática, não havendo restrições aos idiomas espanhol, inglês e português. A Tabela 2.1 apresenta o detalhamento quantitativo dos estudos relacionados durante a busca nas bases citadas, totalizando 254 estudos selecionados. Foram acrescentados mais 30 estudos publicados anteriormente a 2015, em função da relevância e aderência ao tema. Também foram inseridos pelo mesmo motivo 72 estudos publicados em bases diferentes daquelas consideradas inicialmente.

Tabela 2.1: Pesquisa Sistemática

Plataforma	Número Artigos
ACM Digital Library	30
IEEE Xplore	72
MDPI Sensors	34
SciELO	12
Springer Link	34
Outros	72
Total	254

Fonte: Autor.

2.1 Digitalização 3D Subaquática

A partir do levantamento sobre o estado da arte foram relacionados os trabalhos de maior relevância e conexão com o objetivo desta pesquisa, sendo possível identificar os diferentes e principais métodos aplicados para digitalização 3D subaquática. Nesta seção encontram-se destacados aqueles com maior relação ao tema e que contribuíram para o

avanço das pesquisas, seja pelos resultados alcançados, seja pelos métodos propostos, e que subsidiaram o direcionamento desta tese.

De acordo com [Skarlatos et al. \(2016\)](#), o escaneamento digital 3D e técnicas de modelagem representam um inestimável conjunto de ferramentas eficazes para reconhecimento, documentação e monitoramento de ativos.

A maioria das técnicas utilizadas para digitalização 3D subaquática é baseada em sensores ópticos e acústicos adaptados para atender às necessidades de determinadas aplicações, normalmente para águas rasas, profundas e/ou ultraprofundas ([CASTILLÓN et al., 2019](#); [MENNA; AGRAFIOTIS; GEORGOPOULOS, 2018](#); [MASSOT-CAMPOS; OLIVER-CODINA, 2015](#)). Da mesma forma, [Wang et al. \(2019\)](#) também mencionam que um sistema óptico oferece um resultado e informações mais detalhadas de um estrutura imersa do que um sistema acústico e propuseram um sistema de visão estéreo subaquática para aplicação em SLAM de um AUV.

De acordo com [Menna, Agrafiotis e Georgopoulos \(2018\)](#), os sonares podem trabalhar em uma faixa muito mais longa (de alguns milhares de metros) e não são afetados pela turbidez da água. Por outro lado, os sensores ópticos fornecem uma resolução e taxa de atualização muito mais altas e seu curto alcance (normalmente alguns metros) não limita o desempenho do sistema em tarefas de inspeção, porém é necessário aproximar o sistema de digitalização do alvo a ser inspecionado.

A digitalização 3D de objetos em terra usa técnicas ou dispositivos de imagem ativa ou passiva, na maioria dos casos baseada em ondas eletromagnéticas ([STYLIANIDIS; REMONDINO, 2016](#)). Apesar destas técnicas terem sido usadas com sucesso em ambiente terrestre há décadas, as características físicas da água são tais que a modelagem 3D de ambientes subaquáticos torna-se muito mais complexa ou até mesmo inviável em certos casos ([MENNA; AGRAFIOTIS; GEORGOPOULOS, 2018](#)). A água é um meio diferente do ar e a primeira diferença essencial consiste na densidade média. A água do mar é quase 800 vezes mais densa que o ar e sua densidade não é constante através da profundidade, variando em função da temperatura, salinidade e pressão ([MENNA; AGRAFIOTIS; GEORGOPOULOS, 2018](#)). Segundo [Morel \(1974\)](#), a atenuação da luz na água é regida pela dispersão e absorção.

As águas salgadas compõem um ambiente opticamente puro quando totalmente isento de partículas suspensas ([MOREL, 1974](#)). Um ambiente opticamente puro é definido como um ambiente totalmente isento de quaisquer partículas em suspensão. Esta definição não implica que o ambiente seja um composto quimicamente puro, pode ser uma mistura ou mesmo uma solução. Esta definição implica apenas que as propriedades ópticas, especialmente espalhamento e absorção, são determinadas apenas por moléculas ou íons.

Embora os problemas de dispersão e absorção de água pura pertençam à físico-química, também são de interesse da oceanografia óptica. Ainda segundo [Morel \(1974\)](#), as águas do mar aberto, particularmente as águas profundas, são de grande pureza. Dadas as suas características, longos comprimentos de ondas visíveis, como a luz vermelha, são absorvidos primeiro, comprimentos de ondas curtas visíveis, como azul, por último. Por esta razão, apenas 1% da luz que entra no mar atinge 100 m de profundidade ([MENNA; AGRAFIOTIS; GEORGOPOULOS, 2018](#)). Comprimentos de onda de luz com baixa absorção debaixo d'água podem se propagar por distâncias maiores. Esses comprimentos de onda correspondem ao verde ou azul, mas as fontes de *laser* verde são geralmente preferidas pois apresentam menor custo e são mais eficientes em termos de energia ([HILDEBRANDT et al., 2008](#); [CASTILLÓN et al., 2019](#)).

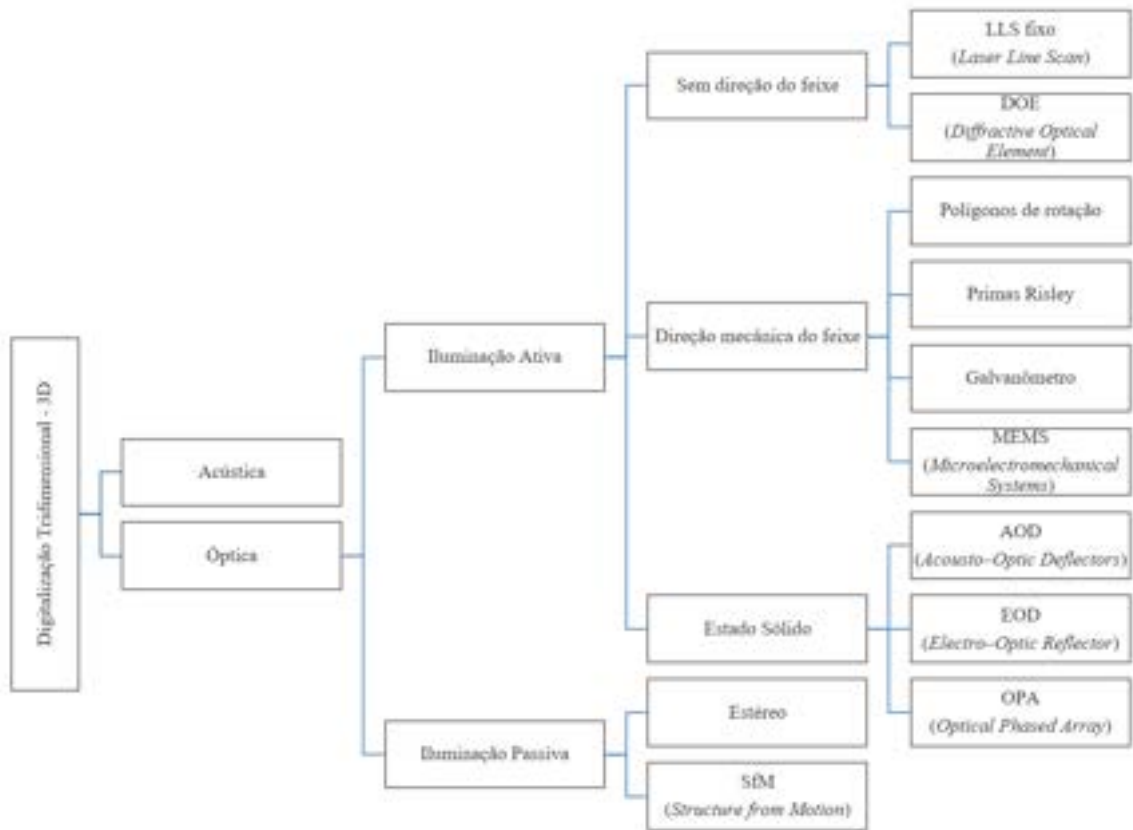
Algumas pesquisas sobre sistemas de imagem subaquática foram publicadas nas últimas duas décadas. [Jaffe et al. \(2001\)](#) resumiram a história da imagem óptica subaquática e sua relação com outros campos da óptica oceânica, concentrando-se nos avanços tecnológicos na última década do século XX. [Kocak e Caimi \(2005\)](#) revisaram o progresso histórico das imagens subaquáticas, com foco especial no período de 2000 a 2005. [Massot-Campos e Oliver-Codina \(2015\)](#) apresentaram uma revisão sobre a reconstrução 3D óptica subaquática, incluindo uma análise quantitativa da comparação de critérios de desempenho. [Castillón et al. \(2019\)](#) fizeram uma ampla revisão do estado da arte dos escâneres (do inglês, *scanners*) 3D ópticos para aplicações subaquáticas e dos sensores ópticos, focada especialmente nos ativos, bem como apresentação de uma revisão da literatura relacionada às tecnologias de sensores e projeção de luz. Neste mesmo trabalho de [Castillón et al. \(2019\)](#) também foram relacionadas as pesquisas mais recentes referentes ao tema, bem como os dispositivos já disponibilizados comercialmente.

[Bruno et al. \(2011\)](#) testaram uma técnica óptica baseada na projeção de padrões de iluminação estruturados adquiridos por um sistema de visão estéreo. Os testes subaquáticos foram realizados em um tanque de água com diferentes condições de turbidez e seus resultados evidenciaram que a qualidade da reconstrução 3D é aceitável¹ mesmo com altos valores de turbidez, apesar da forte presença de efeitos de espalhamento e absorção.

Na Figura 2.1 é apresentada uma classificação das tecnologias de digitalização 3D ([CASTILLÓN et al., 2019](#)). Primeiramente uma classificação entre ativas e passivas, seguido de um detalhamento das técnicas ativas que não realizam o direcionamento do feixe. Posteriormente, as tecnologias de direção a *laser* são classificadas de acordo com o fato de envolverem ou não elementos mecânicos.

¹O termo de aceitável no referido trabalho tomou como referência a aplicação da reconstrução 3D para monitoramento do crescimento de recifes de coral e reconstrução de peças arqueológicas submarinas.

Figura 2.1: Classificação das Técnicas de Digitalização 3D.



Fonte: adaptação de [Castillón et al. \(2019\)](#)

Outros pesquisadores também abordam tal classificação das tecnologias de digitalização 3D, especificamente entre as técnicas ativa e passiva, como apresentado a seguir.

Os sensores ópticos 3D podem ser classificados como ativos ou passivos. De acordo com [Bruno et al. \(2011\)](#), um sensor subaquático usa luz ativa (ou estruturada) quando projeta padrões de luz na cena de maneira controlada. Esses padrões podem ser um ponto, uma linha ou formas mais complexas. Nas técnicas ativas, as informações fornecidas pela estrutura do padrão são essenciais para reconstruir a cena em 3D. Vale notar que a estrutura do padrão não se limita ao domínio espacial, mas também pode ser temporal, que é a base para os sensores de tempo de voo (ToF, do inglês *time of flight*). Técnicas ativas determinam a posição 3D dos pontos no ambiente, por ToF ou por princípios de triangulação. Por outro lado, a iluminação passiva depende apenas da luz do ambiente para iluminar a cena, embora a luz difusa artificial possa ser usada em ambientes escuros ([BRUNO et al., 2011](#)). As técnicas passivas em ambientes subaquáticos geralmente usam visão estéreo ou "estruturada a partir do movimento" (SfM²).

² *Structure from motion* é uma técnica para estimar estruturas tridimensionais a partir de seqüências de imagens bidimensionais, as quais podem ser acoplados com locais específicos ao sinal de movimento.

De acordo com [Palomer et al. \(2019\)](#), as técnicas de visão computacional podem ser divididas em dois grupos, dependendo da fonte de luz utilizada. Existem técnicas passivas que usam uma fonte de luz que ilumina toda a cena da forma mais homogênea possível, por meio da própria luz ambiente ou luz artificial, como lâmpadas e holofotes. Por outro lado, as técnicas ativas utilizam um padrão projetado na cena para triangular os diferentes pontos, produzindo reconstruções com melhores resultados do que as técnicas passivas, especialmente nas áreas sem características definidas, onde o 3D das técnicas passivas é calculado assumindo a continuidade das áreas com determinadas características.

Segundo [Fujimura et al. \(2018\)](#), as principais vantagens de sensores passivos, como sistemas estéreo passivos e estéreo fotométrico, são seu baixo preço e sua alta resolução, que é limitada principalmente pela luz dispersa ([BIANCO et al., 2013](#)). Suas principais desvantagens são o fato de exigirem alto poder computacional e dependerem muito da textura do alvo para extrair *features* da imagem ([DALGLEISH et al., 2009](#); [PALOMER et al., 2019](#)). No entanto, eles são amplamente utilizados pela comunidade de pesquisa marinha ([JOHNSON-ROBERSON et al., 2017](#); [JOHNSON-ROBERSON et al., 2010](#); [MENNA et al., 2013](#); [YAMAFUNE; TORRES; CASTRO, 2017](#)).

Por outro lado, os sensores 3D ativos normalmente atingem uma densidade de nuvens de pontos maior, especialmente em cenários de baixo contraste ([ROMAN; INGLIS; RUTTER, 2010](#)). Uma de suas principais desvantagens é que seu desempenho diminui sob a luz do sol. No entanto, como a luz do sol é rapidamente atenuada pela água do oceano, isso geralmente não se torna um grande problema em ambientes subaquáticos relativamente profundos (a partir de algumas dezenas de metros).

[Yang et al. \(2013\)](#) utilizaram *Laser Line Scan* - LLS para obtenção da reconstrução 3D. Uma relação de mapeamento entre a coordenada de pix 2D e a coordenada 3D foi estabelecida, a qual permitiu ser aplicada para reconstrução de objetos 3D tendo como fonte dados de digitalização 2D. Para o processo de calibração da câmera, o *laser* e a câmera digital foram configurados em uma guia linear e a placa de calibração foi colocada no plano do *laser*. A relação de coordenada entre os dados de amostragem 2D e 3D foi estabelecida de forma que cada ponto de calibração possuía uma coordenada pix na imagem adquirida pela câmera digital. Tanto as coordenadas 2D pix quanto as coordenadas 3D de todos os pontos de calibração foram armazenadas como matrizes, as quais puderam ser consideradas para determinar a relação de mapeamento ponto a ponto. Os experimentos deste trabalho de investigação foram realizados em uma piscina.

[Bobkov et al. \(2018\)](#) aplicaram a técnica onde os conjuntos de pontos 3D das diversas vistas individuais são unidos em um sistema de coordenadas como modelo de ponto global vinculado à próxima visualização simplificada da cena, utilizando "referências" de superfície. Este modelo de pontos é formado por algoritmo que elimina os pontos repetidos

e duvidosos, o que aumenta consideravelmente a velocidade de construção e visualização para a cena do modelo 3D, de acordo com os resultados apresentados pelos autores.

Li et al. (2019) utilizaram um sistema estéreo para reconstrução 3D de objetos numa condição de não luminosidade no ambiente subaquático. Um conjunto de LED na cor branca foi utilizado para iluminação artificial. O conjunto óptico foi encapsulado numa *housing* com interface plana. A opção pela cor branca deve-se ao fato de que sem um equilíbrio de branco adequado, as fotos tiradas pelas câmeras podem apresentar uma tonalidade de cor indesejável (HE et al., 2012).

Skinner, Iscar e Johnson-Roberson (2017) em seu trabalho sobre correção automática de cores para reconstrução 3D subaquática propuseram um método para alcançar a estimativa simultânea dos coeficientes de atenuação para correção de cor durante a recuperação da estrutura de uma cena subaquática, integrando esta estimativa diretamente na etapa de ajuste do feixe. Faz parte do método a utilização da técnica de triangulação entre as câmeras estéreo para reconstrução da cena a partir de múltiplas tomadas.

Skinner e Johnson-Roberson (2015) utilizaram o método Simple Linear Iterative Clustering (SLIC) para calcular os *superpixels* e segmentar as imagens para identificação de rochas em sítios arqueológicos submersos. Os *superpixels* resultam em uma grande segmentação da imagem e que fornece contornos de segmento iniciais.

Weidner et al. (2017) apresentaram resultados de reconstrução de uma caverna subaquática usando uma abordagem que utiliza a iluminação artificial da cena como uma ferramenta para mapear os limites. Mas sugere novas pesquisas para o desenvolvimento para caracterização de sombras provocadas outras fontes de iluminação artificial e eliminação das áreas escuras devido à absorção de luz.

Onmek et al. (2017) propuseram um método baseado em uma mono-câmera para reconstrução 3D de objetos arqueológicos submersos. Para o processamento de imagem foram aplicados filtros já conhecidos para melhorar a qualidade das imagens subaquáticas, selecionando as características de interesse entre pares de imagens. Posteriormente, o método RANSAC³ foi aplicado para rejeitar pontos de *outliers*. Uma triangulação dos *inliers*⁴ encontrados anteriormente e as matrizes de projeção é usada para criar a reconstrução 3D. A triangulação "3D Delaunay" foi usada para criar uma textura de malha, bem como renderizar no modelo 3D.

Mansi e Hussein (2018) propuseram uma técnica de imageamento estéreo para aumentar

³O algoritmo RANSAC foi usado para verificar as correspondências entre pontos-chave. Seu conceito baseia-se na interpretação de um método e encontrar seus parâmetros com um valor de "N" subconjuntos de um determinado número "n" de dados aleatórios.

⁴*Inliers* são pontos que têm uma correspondência correta com os pares de imagens iniciais.

a precisão na reconstrução de objetos 3D. Já a técnica descrita por [Drap et al. \(2015\)](#) foi a fotogrametria, onde três câmeras de alta resolução foram sincronizadas e controladas por um computador, todas montadas em um veículo subaquático operado remotamente (ROV), utilizando LEDs para iluminação adequada do cenário. Este trabalho teve como objetivo principal não o de fazer um levantamento automático do local, mas de oferecer ao piloto do ROV um controle do processo de levantamento por meio de sua visualização.

[Himri et al. \(2018\)](#) utilizou a técnica ICP, do inglês *Iterative Closest Point*, que consiste numa técnica em que um algoritmo é empregado para minimizar a diferença entre duas nuvens de pontos. O ICP é frequentemente usado para reconstruir superfícies 2D ou 3D de diferentes varreduras, para localizar robôs e para alcançar um planejamento ideal de trajetória. Trata-se de um algoritmo de registro iterativo, que é aplicado para refinar o resultado do alinhamento inicial. No trabalho de [Himri et al. \(2018\)](#) foi utilizado o resultado anterior como ponto de partida inicial para alinhar as nuvens de pontos combinadas e as nuvens de pontos digitalizadas.

[Palomer et al. \(2019\)](#) desenvolveram um sensor baseado em uma câmera, *laser* e galvanômetro. Nesse trabalho de pesquisa, foi levada em consideração a distorção introduzida pela refração dupla através das janelas de visualização planas. Este sensor foi usado com sucesso para manipulação, bem como para reconhecimento de objetos e tarefas de Localização e Mapeamento Simultâneos - SLAM.

Após análise dos métodos e resultados obtidos nos diversos trabalhos analisados, tem-se as seguintes indicações:

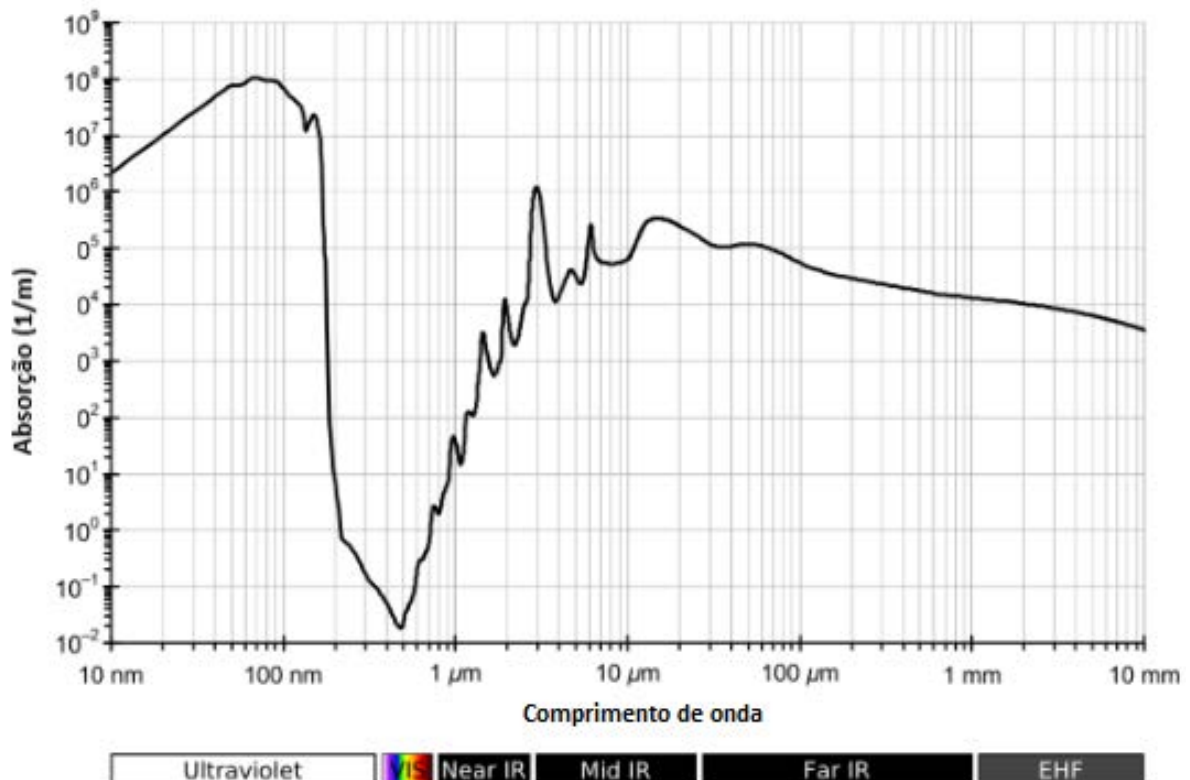
- A realização de testes em condições de maior relevância, ou seja, simulações e testes em condições mais próximas das reais onde sistemas de digitalização são necessários, é um passo recomendado ([SKINNER; ISCAR; JOHNSON-ROBERSON, 2017](#)) como forma de dar robustez e confiabilidade aos métodos propostos.
- Não foi identificado nos trabalhos analisados a realização de testes em condições operacionais extremas em altas profundidades. Portanto, isto corrobora e reforça um dos objetivos propostos nesta tese de realizar testes do protótipo em ambiente relevante perante a condição operacional da indústria de O&G.
- Já [Onmek et al. \(2017\)](#), recomendaram como trabalhos futuros a utilização de um sistema estéreo para obtenção de melhores resultados numa reconstrução 3D.
- A utilização de técnicas de projeção de linhas *laser* permitem obter informação precisa, sendo possível construir modelos de estruturas submersas 3D com alta resolução.
- Estes fatores reforçam a opção de utilização de um sistema estéreo combinado com

projeção estruturada como solução viável para implementação.

2.2 Calibração

Um dos principais desafios na obtenção de imagens subaquáticas é que a luz é fortemente atenuada pela água e este processo depende do comprimento de onda (Figura 2.2). O espectro visível pode viajar até algumas centenas de metros antes de ser completamente absorvido pela água, já os comprimentos de onda de infravermelho não se propagam além de 30 cm (SMITH; BAKER, 1981). Imagens capturadas em águas rasas (com profundidade inferior a 10 m) são menos afetadas pela atenuação da luz na água, pois a partir de poucas dezenas de metros a luz do sol é rapidamente atenuada pela água do oceano. Porém, segundo Garcia et al. (2017) e Castellón et al. (2019), existem outros fenômenos que degradam essas imagens, como a tremulação e a retrodispersão elevada provocada pela presença de partículas em suspensão. Tais partículas refletem a luz projetada de volta para o sensor, gerando ofuscamento.

Figura 2.2: Espectro de absorção da luz na água.



Fonte: Castellón et al. (2019)

Além disso, os sistemas de visão geralmente são enclausurados dentro de um invólucro selado com uma janela de visualização transparente. Isso implica que a luz sofre o processo

de refração duas vezes antes de chegar à câmera da cena (HECHT, 2016; CASTILLÓN et al., 2019), o que traz ainda mais complexidade ao cálculo da posição 3D do objeto observado.

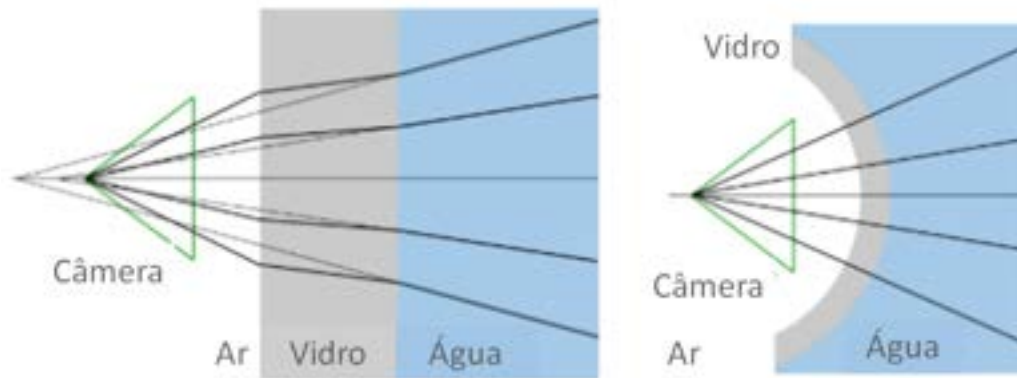
A calibração por si só é uma etapa fundamental em qualquer sistema de visão que visa adquirir dados não distorcidos, precisos, confiáveis. Geralmente compreende duas etapas:

- Primeiro, os parâmetros intrínsecos da câmera, incluindo a lente, devem ser calculados.
- Segundo, a posição e a orientação da câmera, chamados parâmetros extrínsecos, em relação ao projetor a *laser* (no caso de um sistema de triangulação a *laser*) ou em relação à outra câmera (no caso de visão estéreo) também devem ser determinadas.

Métodos de calibração de câmera subaquática têm sido amplamente estudados na literatura (CASTILLÓN et al., 2019). Shortis (2015) apresentou uma pesquisa muito completa sobre técnicas de calibração e considerações para sistemas fotogramétricos subaquáticos. Sedlazeck e Koch (2012) compararam modelos de câmera em perspectiva e não-perspectiva em imagens subaquáticas. Os parâmetros de calibração de um sistema de visão mudam dependendo das condições externas: profundidade, temperatura e salinidade, que alteram o índice de refração da água (SEDLAZECK; KOCH, 2012; LUCZYŃSKI; PFINGSTHORN; BIRK, 2017b). Além disso, o formato da caixa da câmera é propenso a deformações em pressões crescentes (MENNA et al., 2013).

Segundo Shortis (2015), a refração provoca uma distorção, o que faz com que os maiores erros de reconstrução apareçam nas bordas do alvo. Devido à natureza simétrica desse efeito, ele pode ser absorvido pelo componente de distorção da lente radial dos parâmetros de calibração. Um método prático para calibrar uma câmera para digitalização a *laser* subaquática é apresentado em Wang e Cheng (2007). No entanto, o efeito de refração envolve erros sistemáticos, uma vez que a suposição de um único centro de projeção para a câmera (modelo de câmera do ponto de vista único) não se sustenta (Figura 2.3). Para tratar esse problema, pode-se adotar uma abordagem mais complexa por meio do rastreamento dos raios de luz pelas interfaces de refração (PALOMER et al., 2017).

Figura 2.3: Refração dos raios de luz nas portas plana e de domo, respectivamente.



Fonte: Sedlazeck e Koch (2012).

As câmeras subaquáticas usam principalmente dois tipos de portas: plana ou em forma de cúpula. À custa de um processo mais oneroso e de difícil fabricação e montagem, as portas das cúpulas podem, em princípio, reduzir o efeito de refração, pois existe um alinhamento teórico entre a interface normal e os raios recebidos. No entanto, devido a pequenos desalinhamentos, essa redução geralmente não é total (SEDLAZECK; KOCH, 2012). Palomer et al. (2017) demonstraram que um cone elíptico é uma geometria melhor do que um plano para descrever a deformação de uma linha de *laser* por meio de uma porta plana em água, especialmente quando o ângulo de incidência aumenta entre o *laser* e a porta. Usar um cone elíptico em vez de um plano, no entanto, torna o processo de reconstrução 3D mais exigente em termos computacionais.

Segundo Luczyński, Pfingsthorn e Birk (2017a), as janelas de tela plana usadas nas caixas seladas (*housings*) causam distorções não triviais baseadas em refração na interface ar-vidro-água. Diante disto, desenvolveram um novo modelo de calibração de câmeras subaquáticas chamado Pinax, que combina aspectos de um modelo de câmera axial e *pinhole* para obter uma solução tão precisa quanto possível e ao mesmo tempo computacionalmente viável. A princípio, havia sido projetado para câmeras com eixos ópticos ortogonais ao painel de vidro, porém mostrou-se viável também para sistemas estéreo.

Em outro trabalho, Luczyński, Pfingsthorn e Birk (2017b) combinaram aspectos de um modelo *pinhole* virtual com a função de projeção do modelo de câmera axial. Nele o modelo Pinax permite o pré-processamento de uma tabela de consulta para correção de refração da tela plana com alta precisão e de forma rápida. O modelo leva em consideração os índices de refração da água, principalmente no que diz respeito à salinidade (mesmo com valores de salinidade apenas estimados), sendo, portanto, suficiente calibrar a câmera subaquática apenas uma vez no ar.

As principais desvantagens dos sistemas de visão subaquática são o alcance limitado, a

necessidade de boa iluminação da cena observada e o fenômeno de refração que influencia a imagem observada. A refração, como principal fonte de distorção da imagem, é o maior obstáculo no caminho para o uso de algoritmos comuns de visão 3D subaquática. Isso pode ser atenuado pelo uso de portas de dome, mas elas precisam ser especificamente projetadas para se adequar aos parâmetros da câmera e sua integração não é trivial. [Luczyński, Pfingsthorn e Birk \(2017a\)](#) desenvolveram, a partir da combinação de aspectos do modelo *pinhole* com o modelo axial, um novo processo de calibração e retificação epipolar para câmeras subaquáticas, o qual é muito rápido e eficiente, exigindo apenas uma calibração única no ar. Este método de calibração e o novo modelo de câmera, denominado Pinax ([LUCZYŃSKI; PFINGSTHORN; BIRK, 2017b](#)), podem ser aplicados a câmeras estéreo de maneira direta se ambas as câmeras estiverem em caixas separadas ou se estiverem paralelas atrás de um único painel plano.

2.3 Tecnologias de Sensores Ópticos Aplicados à Digitalização e Reconstrução 3D

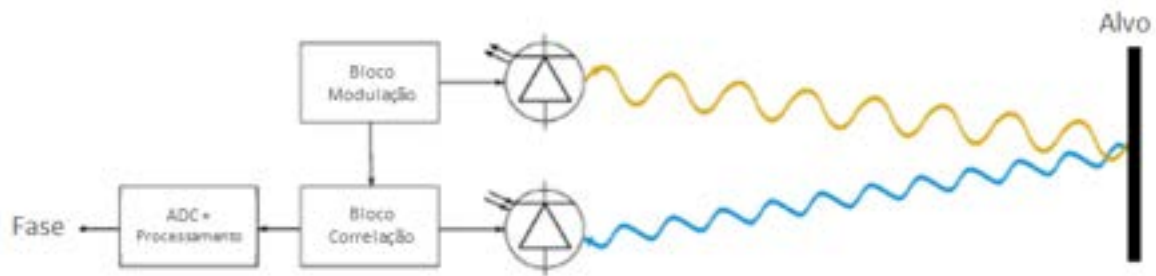
Nesta seção trata-se sobre os dois principais tipos de sensores e técnicas para obtenção da digitalização e reconstrução 3D de estruturas subaquáticas. Primeiro aborda-se os métodos com câmeras e, em seguida, os métodos com a aplicação de *laser*. Por fim, uma análise de desempenho entre os diferentes sensores é apresentada.

2.3.1 Métodos de reconstrução 3D com câmeras

A maioria dos dispositivos na extremidade receptora de sensores 3D ópticos ativos subaquáticos na literatura é baseada em métodos de ToF (*Time of FLight*) ou triangulação ([ZANUTTIGH et al., 2016](#); [CASTILLÓN et al., 2019](#)).

Os sensores ToF calculam a profundidade de um ponto medindo o tempo desde a emissão de um raio de luz até a sua recepção. A maioria deste sensores compartilha uma estrutura básica (Figura 2.4) composta por: uma fonte de luz (transmissor), um detector correspondente (receptor), um circuito adequado para fornecer a fonte de alimentação e os sinais de controle necessários ao transmissor e receptor, além de um conversor analógico-digital (ADC).

Figura 2.4: Esquema de um sensor ToF 3D.



Fonte: Zanuttigh et al. (2016).

Diferentes tecnologias são utilizadas na extremidade receptora dos sensores ToF. As câmeras de alcance limitado têm a vantagem de serem mais robustas aos efeitos de retroespalhamento (CHUA et al., 2017). Mais especificamente, intensificadores fechados acoplados a sensores de imagem de dispositivo de carga acoplada (CCD) permitem um ajuste preciso da faixa. No entanto, os *chips* complementares de óxido de metal semiconductor (CMOS) atingem altas taxas de quadros e alta resolução espacial e de profundidade, com uma menor complexidade do sistema (RISHOLM et al., 2018).

Um novo tipo de sensor, chamado câmera de eventos, está sendo gradualmente usado em sistemas de visão robótica 3D no ar para reconstrução e rastreamento de cenas (KIM et al., 2014a; BARDOW; DAVISON; LEUTENEGGER, 2016). Em vez de armazenar um quadro de imagem completo em intervalos definidos, as câmeras de eventos gravam uma sequência assíncrona de alterações de intensidade por pixel, cada uma contendo informações precisas do momento do evento (data/hora). Segundo Falanga, Kim e Scaramuzza (2019), o baixo consumo de energia das câmeras de eventos, a alta taxa de quadros e a ausência de desfoque de movimento postulam-nos como sensores apropriados para utilização em robôs. Bardow, Davison e Leutenegger (2016) propuseram um algoritmo para recuperar de forma simultânea o campo de movimento e a imagem de brilho, enquanto a câmera é movida em uma determinada cena. A abordagem emprega a minimização de uma função de custo que contém os dados do evento assíncrono, bem como a regularização espacial e temporal dentro de um intervalo de tempo de janela deslizante. No referido trabalho, Bardow, Davison e Leutenegger (2016) demonstraram bom resultado dos experimentos, inclusive em situações em que câmeras convencionais sofrem limitações de faixa dinâmica e desfoque de movimento.

Os sensores ativos de triangulação de luz encontram a posição 3D de um ponto na cena, combinando informações geométricas do emissor de luz e do receptor. Dessa forma, a posição no espaço do ponto digitalizado coincide com a interseção de dois raios de luz: aquele enviado pelo projetor e aquele que passa do ponto focal da câmera pelo *pixel* iluminado. Em relação às tecnologias aplicadas, a maioria dos escâneres baseados em triangulação subaquática usa câmeras regulares protegidas dentro de uma caixa selada.

As dificuldades que esse fato acarreta foram explicadas na seção 2.2, principalmente o fato de os sistemas de visão geralmente serem enclausurados dentro de um invólucro selado com uma janela de visualização transparente. Isso implica que a luz sofre o processo de refração duas vezes antes de chegar à câmera da cena (HECHT, 2016; CASTILLÓN et al., 2019).

2.3.2 Métodos de reconstrução 3D com *laser*

A utilização de tecnologia *laser* no ambiente subaquático pressupõe conhecimento básico do comportamento da luz neste meio e dos fenômenos que nele ocorrem. Segundo May, Pervoelz e Surmann (2007), as imagens captadas debaixo d'água têm má resolução e baixo contraste. Isso acontece devido à atenuação da luz, que é um processo caracterizado por fenômenos de absorção e de dispersão provocados não só pela água, mas também por partículas presentes nela (SCHETTINI; CORCHS, 2010). Apesar disso, o uso do *laser* para mapeamento de estruturas submersas continua a ser uma solução viável, pois permite examinar objetos num curto espaço de tempo e obter informação sobre eles com elevada precisão (HUACHUAN et al., 2014). Em função destas características, esta tecnologia é aplicada em tarefas como inspeção de dutos, auxílio à navegação de veículos subaquáticos autônomos, ao monitoramento da biologia marinha e à digitalização de objetos submersos.

Normalmente estes sistemas são constituídos por uma fonte de luz (*laser*) e um sistema de aquisição de imagem (câmeras). De acordo com Chen et al. (2012), nas tecnologias que recorrem à projeção de linhas *laser* é comum haver alguma flexibilidade na definição dos parâmetros de *hardware* (potência do *laser*), por isso é possível encontrar sistemas *laser* com princípios de funcionamento semelhantes, mas com arquiteturas e características diferentes.

O *Laser Line Scan* (LLS) utiliza projeção de linhas *laser*, como sua própria denominação diz. Esta técnica utiliza a transmissão contínua do feixe *laser* com comprimentos de onda dentro da cor azul-verde que, posteriormente, é captada por sensores. Segundo Bechtold, Hohenstein e Schmidt (2013), esta estratégia é considerada como uma das melhores no que diz respeito ao alcance dentro das tecnologias de aquisição de imagens ópticas em ambiente subaquático. Seu funcionamento é semelhante aos escâneres, mas eles projetam um plano na cena. A interseção deste plano com um objeto cria uma linha iluminada que é registrada pela câmera. A posição 3D no espaço de cada linha em relação ao quadro de coordenadas da câmera pode ser calculada combinando as informações do plano do *laser* com os *pixels* da câmera, isso permite a criação de nuvens de pontos densas.

Já o *Pulsed Laser Line Scan* (PLLS) é uma técnica ativa usada na detecção e identificação de estruturas, capaz de produzir mapas de elevada qualidade. Esta tecnologia mantém

as vantagens do LLS, contudo a transmissão do *laser* é feita de forma intermitente. Ao associar esta tecnologia a sistemas de aquisição com tempo de exposição limitado, a qualidade e o contraste das imagens melhoram, pois apenas as reflexões provenientes do alvo são captadas e a quantidade de ruído introduzido pelo meio é reduzido (BECHTOLD; HOHENSTEIN; SCHMIDT, 2013).

A utilização de técnicas que recorrem à projeção de linhas *laser* permitem obter informação extremamente precisa, sendo possível construir nuvem de pontos de modelos de estruturas submersas 3D com elevada resolução, devido à alta densidade dos pontos coletados e, conseqüentemente, redundância na descrição discreta dos objetos, independentemente da sua dimensão ou forma (DIGUMARTI et al., 2016).

No trabalho desenvolvido por Palomer et al. (2019), em vez de uma triangulação baseada em *laser*, foram usados cones elípticos para acelerar a reconstrução sem aumentar o erro. O *laser* foi projetado em diferentes partes da cena, sendo direcionado por meio de um espelho acionado por um galvanômetro, que tem como uma de suas características principais uma inércia muito baixa e, por se mover no ar (interno ao dispositivo), sofre menos atrito. O resultado do trabalho foi usado com sucesso para manipulação, bem como para reconhecimento de objetos e tarefas de SLAM (Localização e Mapeamento Simultâneos).

2.3.3 Desempenho dos diferentes tipos de sensores

A principal diferença no desempenho entre os sensores ToF e triangulação diz respeito à faixa de varredura e à resolução da profundidade (disparidade). A resolução de profundidade de um escâner ToF depende da resolução do tempo ou da medição de fase, mas não da distância da digitalização, diferentemente dos escâneres de triangulação. Esse fato foi estudado por McLeod et al. (2013), que identificaram um limite de faixas (abaixo de 1 m) em que os sensores baseados em triangulação proporcionavam maior precisão de profundidade. Acima de 2,5 m, os sensores ToF apresentaram-se geralmente mais precisos.

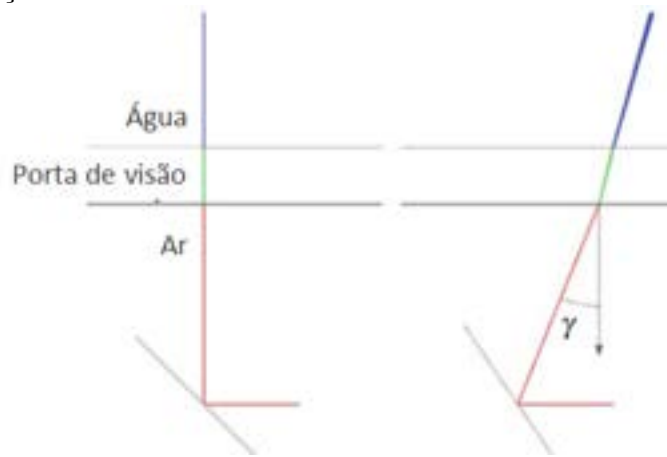
Além disso, a implementação de qualquer um desses métodos de medição em um sensor subaquático envolve vários desafios. Conforme explicado na subseção 2.2, a principal dificuldade para reconstruir com precisão a posição 3D de um ponto tem a ver com o fato de que cada raio de luz viaja através de três meios diferentes: o ar, o material da janela de visualização e a água. Isso afeta cada método de uma maneira diferente:

- Em sensores de triangulação subaquática usando janelas de visualização (*viewports*) planas, a direção dos raios de luz muda duas vezes devido à dupla refração (Figura 2.5), o que pode afetar a precisão da reconstrução. Em ângulos de incidência cres-

centes do *laser* na janela de visualização, o plano do *laser* se transforma em um cone elíptico (Figura 2.6), o que torna a reconstrução 3D mais exigente em termos computacionais (PALOMER et al., 2017).

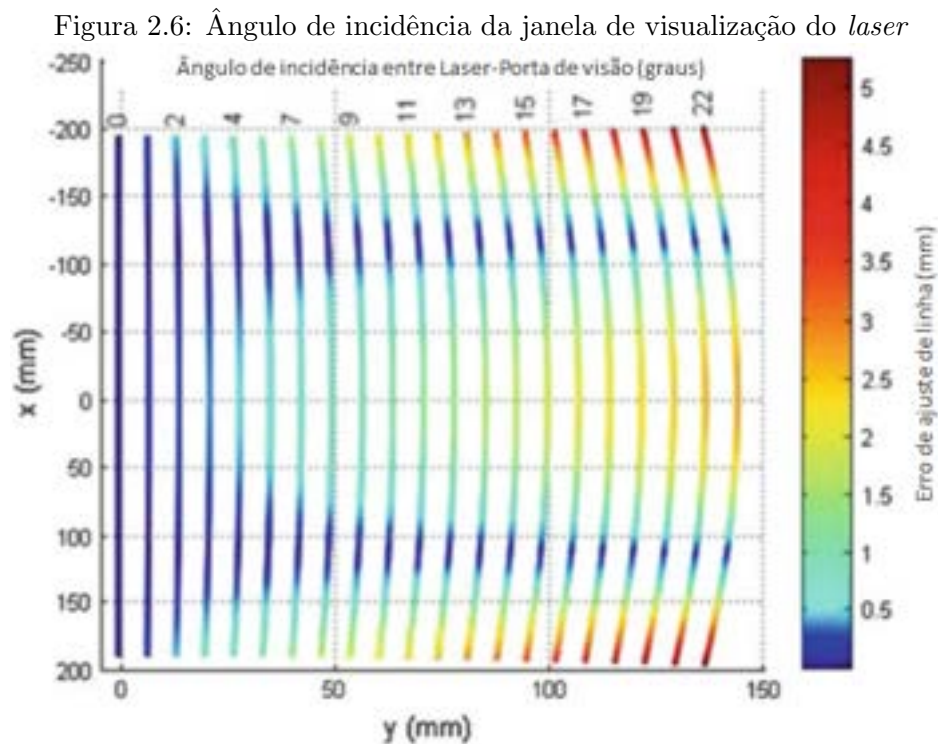
- Já os sensores ToF subaquáticos sofrem com o fato de que as diferentes velocidades de propagação da luz afetam o cálculo da distância total percorrida (DIGUMARTI et al., 2016).

Figura 2.5: Mudança na direção de uma linha de *laser* em função do ângulo de incidência na água e da dupla refração.



Fonte: Palomer et al. (2017).

A partir da Figura 2.6 verifica-se que as linhas de *laser* projetadas estão cada vez mais distorcidas para ângulos de incidência mais altos do *laser* na janela de exibição.



Fonte: Palomer et al. (2017).

2.3.4 Avaliação sobre tecnologias atuais disponíveis

Aqui são analisados sensores ópticos 3D subaquáticos ativos presentes na literatura. Consequentemente, nem os escâneres acústicos nem os de luz passivos são considerados para este estudo pelas razões já comentadas nas seções anteriores. Além disso, apenas os dispositivos desenvolvidos após 2015 foram levados em consideração nesta análise.

Risholm et al. (2018), Mariani et al. (2019) e Chua et al. (2017) utilizaram em seus trabalhos de pesquisas sensores ToF controlados por faixa. Isso significa que o receptor de luz abre com um atraso após o envio da luz. Isso é feito para torná-los mais robustos contra retroespalhamento. Risholm et al. (2018) desenvolveram um algoritmo de localização de pico que pode lidar com dispersão direta e, ao mesmo tempo, pode perceber picos distantes que são pouco mais altos do que picos causados por ruído. Eles mediram o efeito de espalhamento em uma piscina e a relação sinal-ruído versus resolução de profundidade no mar. Chua et al. (2017) introduziram um novo modelo de estimativa de intervalo para reduzir os efeitos induzidos pela distância, reflexão do alvo e distorção do intervalo com base na reconstrução por divisão de tempo e função de distribuição de reflexão bidirecional.

Massot-Campos e Oliver-Codina (2015) apresentaram um sensor *laser*, baseado em triangulação para reconstruções 3D subaquáticas de curto alcance, para operar em ambiente

de turbidez elevada. No trabalho desenvolvido por [Massot-Campos et al. \(2015\)](#), eles compararam com a visão estéreo. Os autores descobriram que uma reconstrução baseada em estéreo é mais adequada para levantamentos longos, desde que a cena tenha textura e luz suficientes. Por outro lado, sua reconstrução a partir de luz estruturada funcionou melhor em distâncias curtas, onde as dimensões precisas de um objeto ou estrutura eram necessárias.

[Matos, Buschinelli e Pinto \(2020\)](#) aplicaram sensores de triangulação a *laser* (LTS) para medições subaquáticas com o objetivo de atender à necessidade da indústria de petróleo e gás de medir equipamentos subaquáticos. Algoritmos de calibração e medição também foram desenvolvidos e avaliados por experimentos controlados. O método de calibração proposto está baseado em duas etapas: uma calibração no ar e algumas aquisições subaquáticas de um padrão simples. Essas aquisições subaquáticas são usadas para calibrar a distância da interface de refração do centro do orifício da câmera, permitindo assim o traçado do raio da câmera ao plano do *laser*, resultando em um ponto 3D medido. Objetos de diâmetro de 100 mm foram usados como alvos de inspeção e os experimentos foram executados em laboratório sob influência de luz natural. A qualidade da medição é influenciada pela qualidade da luz *laser*, ou seja, quão "plano" é o plano do *laser* projetado. Este erro pode ser minimizado alterando a superfície ajustada para uma equação que melhor representa a superfície projetada. Com este método foi alcançado um erro de medição de 0,48 mm.

[Kim, Choi e Ko \(2020\)](#) trabalharam no desenvolvimento de um sistema de medição 3D subaquático utilizando uma câmera estéreo, um *laser line* e um motor DC (*direct current*). Nesse trabalho, a calibração estéreo subaquática foi desenvolvida usando o método de Zhang ([ZHANG, 2000](#)). Diferentes fotografias foram tiradas debaixo d'água usando um tabuleiro de xadrez. Em seguida, cada foto do tabuleiro foi utilizada para obtenção dos parâmetros intrínsecos e extrínsecos das duas câmeras. Durante o processo de digitalização, imagens das duas câmeras foram armazenadas na memória do computador para serem processadas e, a partir delas, obter-se a nuvem de pontos. Esse processo levou cerca de 40 segundos para cobrir uma área de 60° no eixo horizontal. Cerca de 1.200 imagens foram gravadas para cada câmera. Em seguida, as imagens são decodificadas com o objetivo de segmentá-las para detectar as linhas de *laser*. Durante o processo de digitalização, apenas as linhas de *laser* se movem enquanto as câmeras são fixas. A distância entre as duas câmeras era de 60 cm e o gerador de linha a *laser*, colocado entre as elas, foi parametrizado para girar da esquerda para a direita enquanto projeta padrões de linha. Experimentos de medição de objetos foram executados em condições terrestres e subaquáticas e a reconstrução 3D foi realizada para analisar os resultados experimentais. O erro de distância entre os dois planos de referência foi encontrado em cerca de 4 a 9 mm.

Palomer et al. (2019) desenvolveram um sensor baseado em uma câmera, *laser* e galvanômetro. Nesse trabalho de pesquisa, foi levada em consideração a distorção introduzida pela refração dupla através das janelas de visualização planas. Em vez de uma triangulação baseada em *laser*, eles usaram cones elípticos para acelerar a reconstrução sem aumentar o erro. O método de calibração foi baseado na estimativa dos parâmetros intrínsecos da câmera e, em seguida, a posição relativa do laser e do espelho foi calculada. Essas duas etapas do processo de calibração foram feitas no ar e sem que as janelas de visualização estivessem montadas no sensor. Em seguida, tanto a câmera quanto as portas de visualização do *laser* foram montadas e suas posições foram estimadas debaixo d'água. Finalmente, um cone foi calculado para cada ângulo de espelho. Este sensor foi usado com sucesso para manipulação, bem como para reconhecimento de objetos e tarefas de SLAM.

Lopes et al. (2015) desenvolveram um escâner com *laser* de duas linhas (vermelho e verde). Os resultados com o *laser* vermelho foram supostamente mais precisos do que com o verde, embora nenhuma explicação das possíveis causas tenha sido fornecida. Seu objetivo era construir um protótipo de uso geral e baixo custo para indústrias subaquáticas, mas o custo final não foi informado. O sistema foi calibrado em ambientes secos e subaquáticos usando dois métodos diferentes, baseados no princípio da invariância da razão cruzada e no ajuste robusto da projeção da linha de *laser*, respectivamente. O segundo método de calibração produziu resultados mais precisos. Durante os testes, o escâner não foi submerso, então a varredura foi realizada de fora do tanque de água.

No trabalho desenvolvido por Digumarti et al. (2016), foi utilizado o Intel RealSense para desenvolver uma solução barata e compacta que permitia a varredura manual da vida marinha por mergulhadores. Eles escanearam o recife de coral no oceano a uma profundidade de 20 m. Os desempenhos relatados dos escâneres estão reunidos na Tabela 2.2.

Tabela 2.2: Escâneres 3D ópticos ativos subaquáticos atuais.

Referência	Precisão (mm)	Tipo	Cenário de teste
Matos, Buschinelli e Pinto (2020)	0,48	Triangulação	Execução de medições de pequenos objetos em laboratório (bancada) com muito foco na calibração de câmera
Kim, Choi e Ko (2020)	4	Triangulação	Desenvolvido um digitalizador com sistema estéreo, <i>laser</i> e motor CC. Os testes ocorreram numa piscina com as câmeras fixadas para digitalização de objetos cerâmicos pequenos
Palomer et al. (2019)	0,98	Triangulação	Trabalho baseado em sistema estéreo, <i>laser</i> e galvanômetro, onde o protótipo foi embarcado no veículo Girona; o erro atingido foi determinado a partir de um ponto de observação de 0,5 a 1,2 m e para objetos de até 100 mm
Chua et al. (2017)	12	ToF	Testes laboratoriais de pequenos objetos
Digumarti et al. (2016)	4	Não informado	Foram utilizados digitalizador com <i>laser</i> para varredura de recifes de coral numa profundidade de 20 m; não foi dado foco à reconstrução da cena para fins de intervenção
Lopes et al. (2015)	0,22	Triangulação	Buscou-se construir um protótipo de uso geral e baixo custo para indústrias subaquáticas e os testes foram realizados em tanque laboratorial, alcançando erro em torno de 2% de pequenos objetos, sem execução da reconstrução 3D
Massot-Campos e Oliver-Codina (2015)	2,6	Triangulação	Desenvolvido com sensor <i>laser</i> , baseado em triangulação para reconstruções 3D subaquáticas de curto alcance e realização de testes laboratoriais

Fonte: Autor.

A precisão dos valores apresentados foram mantidos conforme àqueles expostos em cada um dos trabalhos relacionados.

Há uma diferença no desempenho entre ToF e sensores de triangulação. **A maioria dos sensores de triangulação analisados alcançam erros de medição menores.** Os melhores resultados são obtidos por [Lopes et al. \(2015\)](#), [Matos, Buschinelli e Pinto \(2020\)](#) e [Palomer et al. \(2019\)](#). Por outro lado, os sensores ToF podem trabalhar em alcances

mais longos, o que torna seu erro de profundidade relativo em distâncias mais longas, comparáveis aos escâneres de triangulação trabalhando em alcances curtos (geralmente menos de 2 m).

2.4 Algoritmos para reconstrução 3D

No âmbito da **percepção**, comparada a outros sensores, as observações visuais são extremamente ricas, o que necessariamente implica numa enorme quantidade de dados produzida (RUSSELL; NORVIG, 2013).

Ainda de acordo com Russell e Norvig (2013), três abordagens podem ser caracterizadas para tratar esse problema:

- Extração de características, com cálculos simples aplicados diretamente às observações do sensor.
- Reconhecimento, como a distinção entre os objetos encontrados com base nas informações visuais.
- Reconstrução, trata-se da construção de um modelo geométrico do cenário ou do objeto a partir de imagem ou conjunto de imagens.

Portanto, nesta seção, são apresentados algoritmos relacionados à **reconstrução 3D** por meio de técnicas computacionais com aplicação de filtros para tratamento das imagens.

De acordo com Wolff et al. (2016), conjuntos de pontos gerados por técnicas de reconstrução 3D baseadas em imagem são geralmente muito mais ruidosos do que aqueles obtidos usando técnicas ativas como varredura a *laser*. Os autores apresentaram um método para remover ruídos e *outliers* de tais conjuntos de pontos, por meio do uso de imagens de entrada e os mapas de profundidade correspondentes para remover *pixels* que se apresentam inconsistentes com a superfície colorida implícita na entrada. Isso permite que os métodos de reconstrução de superfície padrão (como a reconstrução de superfície de Poisson) realizem menos suavização e, assim, obtenham superfícies de maior qualidade com mais *features*. Wolff et al. (2016) também demonstraram que o algoritmo pode ser aplicado em combinação com uma variedade de métodos de reconstrução de profundidade e superfície de última geração.

Rakotosaona et al. (2019) propuseram um método baseado em dados para remover *outliers* e reduzir o ruído em nuvens de pontos não ordenadas. A abordagem foi baseada em uma arquitetura de aprendizado profundo adaptada do PCPNet proposta para estimar

propriedades de forma 3D local em nuvens de pontos. O método primeiro classifica e descarta amostras discrepantes e, em seguida, estima vetores de correção que projetam pontos ruidosos nas superfícies originais limpas. Seus resultados indicam que o método é capaz de lidar com grandes nuvens de pontos densamente amostradas, porém ainda dependem de pares de nuvens de pontos livres de ruídos no conjunto de treinamento (*dataset*).

Fu et al. (2014) propuseram uma abordagem para aprimorar um única imagem subaquática baseada em retinex⁵, composta por três etapas: primeiro, feita a correção de cores simples; segundo, aplicada uma estrutura variacional para retinex para decompor a refletância e a iluminação; terceiro, a refletância e a iluminação são aprimoradas por diferentes estratégias para resolver o problema de subexposição e de difusão. A imagem final aprimorada é obtida combinando o uso da refletância e iluminação aprimoradas. Além disso, a abordagem proposta é um método geral que pode aprimorar outros tipos de imagem degradada.

Ahn et al. (2013) propuseram uma nova técnica de mapeamento de tons para imagens de alta faixa dinâmica baseada na teoria do retinex. O algoritmo consiste em duas etapas: adaptação global e adaptação local do sistema visual humano. No processo de adaptação local, o filtro gaussiano dos algoritmos retinex é substituído por um filtro guiado para reduzir efeitos de aréola. Para garantir uma boa reprodução e compressão de faixa dinâmica, foi proposto um fator de aprimoramento de contraste com base nos valores de luminância da cena. Além disso, um deslocamento adaptativo de não linearidade foi introduzido para lidar com a força da não linearidade da função logarítmica. O algoritmo apresentou resultados satisfatórios, preservando detalhes e reduzindo efeitos de aréola.

A maioria dos algoritmos de desembaçamento de imagem requer, para sua operação, o vetor de luz atmosférica, A , que descreve a luz ambiente na cena. Sulami et al. (2014) apresentaram um novo método automático para recuperar o vetor de luz atmosférica em cenas nebulosas a partir de uma única imagem de entrada. O método recupera a orientação do vetor de luz atmosférica, explorando a abundância de pequenos trechos de imagem em que a transmissão da cena e o albedo da superfície são aproximadamente constantes. A partir daí, deriva-se um modelo de formação reduzida que descreve a distribuição dos *pixels* dentro de tais manchas como linhas no espaço RGB, as quais são usadas para extrair A de forma robusta. A questão chave está em identificar nas imagens nebulosas onde o nível de intensidade dos *pixels* mais brilhantes é aproximadamente independente de seu valor de transmissão. Para explorar esta propriedade é derivada uma expressão analítica para a dependência que uma magnitude errada introduz, minimizando

⁵Retinex: o filtro retinex é baseado na teoria da percepção de imagem de Land, proposta para explicar a constância de cor percebida de objetos sob diferentes condições de iluminação. Existem várias abordagens para implementar os princípios do retinex, entre eles o *multiscale retinex* com algoritmo de restauração de cores (MSRCR) que combina constância de cor com aprimoramento de contraste local para que as imagens sejam renderizadas de maneira semelhante à forma como a visão humana funciona.

este tipo particular de dependência.

As imagens subaquáticas normalmente exibem distorção de cor e baixo contraste como resultado do decaimento exponencial que a luz sofre à medida que viaja. Além disso, cores associadas a diferentes comprimentos de onda possuem diferentes taxas de atenuação, sendo o comprimento de onda vermelho aquele que atenua mais rapidamente. Para restaurar imagens subaquáticas, [Galdran et al. \(2015\)](#) propuseram um método Red Channel, no qual as cores associadas a comprimentos de onda curtos podem ser recuperadas e, conforme esperado para imagens subaquáticas, levam ao restabelecimento do contraste perdido. Tal técnica pode ser aplicada em áreas iluminadas artificialmente e proporcionam uma correção de cor natural e uma melhoria de visibilidade.

[Ancuti et al. \(2012\)](#) propuseram uma estratégia para aprimorar vídeos e imagens subaquáticas que, a partir dos princípios da fusão, deriva as entradas e as medidas de peso apenas da versão degradada da imagem. Para superar as limitações do meio subaquático, foram definidas duas entradas que representam versões com correção de cor e contraste aprimorado da imagem/quadro original subaquático, mas também quatro mapas de peso que visam aumentar a visibilidade dos objetos distantes degradados devido ao meio dispersão e absorção, resultando na redução de ruído e preservação de borda.

Um algoritmo de desembaçamento rápido e otimizado para imagens e vídeos nebulosos foi proposto por [Kim et al. \(2014b\)](#). Com base na observação de que uma imagem nebulosa apresenta baixo contraste em geral, ela foi restaurada e teve seu contraste aumentado. Essa ação porém, trunca os valores dos *pixels* e causa perda de informações. Daí, a técnica propôs uma função de custo que consiste no termo de contraste e no termo de perda de informação. Ao minimizar a função de custo, o algoritmo proposto aumenta o contraste e preserva a informação de forma otimizada, além de ser suficientemente rápido para aplicações de desembaçamento em tempo real.

De acordo com [He et al. \(2012\)](#), a temperatura de cor em uma imagem nebulosa pode não ser detectada corretamente, o que resulta em distorção de cor. Além disso, toda a cena parece mais brilhante, o que leva à subexposição da imagem. Diante disto, [He et al. \(2012\)](#) propuseram um método para lidar com essas duas questões aplicando a correção do balanço de branco e decompondo uma imagem em duas imagens componentes, imagem de luminosidade reflexa e imagem de iluminação ambiente. Foi desenvolvido um algoritmo baseado em canal escuro aprimorado para desembaçar a imagem de luminosidade reflexa e o ajuste de exposição é estimado a partir da imagem de iluminação ambiente.

[Humenberger et al. \(2010\)](#) propuseram um algoritmo de correspondência estéreo baseado na transformada Census⁶. O algoritmo propõe a adaptação e otimização explícita da

⁶A transformada Census consiste em um operador de imagem que associa a cada *pixel* de uma imagem em

conhecida transformada Census em relação a sistemas de tempo real embarcados em *software*. Isto se dá com a utilização de uma máscara de Census esparsa que reduz pela metade o tempo de processamento com qualidade de correspondência quase inalterada. Segundo [Humenberger et al. \(2010\)](#), isso deve-se ao fato de que grandes máscaras de Census esparsas têm um desempenho melhor do que pequenas máscaras densas com o mesmo esforço de processamento. No mesmo trabalho foi proposto também um sistema completo de correspondência estéreo com seu algoritmo de núcleo baseado em correlação, que permite tratar áreas com baixa textura. O sistema apresentou-se robusto, fácil de parametrizar e oferece alta flexibilidade.

[Hirschmuller \(2008\)](#) apresentou o método estéreo Semi-Global Matching (SGM). Ele usa um custo de correspondência baseado em informações mútuas, *pixelwise*, para compensar as diferenças radiométricas das imagens de entrada. A correspondência *pixelwise* é suportada por uma restrição de suavidade que geralmente é expressa como uma função de custo global. O SGM realiza uma aproximação rápida por otimizações de caminhos de todas as direções. Também trata a detecção de oclusão, refinamento de *subpixel* e correspondência de várias linhas de base. Além disso o SGM propõe etapas de pós-processamento para remoção de *outliers*, recuperação de problemas específicos de ambientes estruturados e interpolação de *gaps*.

Segundo [Spangenberg et al. \(2014\)](#), o SGM é amplamente utilizado para visão estéreo em tempo real no contexto automotivo. Apesar de sua popularidade, apenas implementações usando hardware reconfigurável (FPGA) ou hardware gráfico (GPU) atingem taxas de quadros altas o suficiente para veículos inteligentes. As implementações em tempo real existentes para computadores de uso geral usam subamostragem de imagem e disparidade em detrimento da qualidade correspondente. [Spangenberg et al. \(2014\)](#) propuseram um método para melhorar a eficiência do SGM em computadores de uso geral, por meio de paralelização refinada e uso de vários núcleos. O sistema apresenta a vantagem de ser escalável para o número de núcleos disponíveis e portátil para processadores embarcados.

Considerando que a visão estéreo é uma tarefa que requer avanços para aplicação em sistemas robóticos, [Haller e Nedeveschi \(2010\)](#) propuseram um sistema de reconstrução estéreo denso em tempo real. Foi selecionado o método SGM como base do sistema devido à sua alta qualidade e reduzida complexidade computacional. A transformação Census foi selecionada como a métrica de correspondência pois os resultados indicavam que ela poderia reduzir os erros de correspondência para imagens de tráfego em comparação com as soluções clássicas. Duas modificações no algoritmo Semi-Global original foram propostas para melhorar a precisão dos *subpixels* e o tempo de execução.

A correspondência estéreo em tempo real de alta qualidade tem o potencial de permi-

escala de cinza a uma sequência binária, codificando se o *pixel* tem intensidade menor que cada um de seus vizinhos, um para cada bit ([WOODFILL; HERZEN, 1997](#)).

tir várias aplicações de visão computacional, incluindo cirurgia robótica semiautomática, teleimersão (no âmbito da realidade virtual e aumentada) e vigilância por vídeo tridimensional (KOWALCZUK; PSOTA; PEREZ, 2013). Um método de correspondência estéreo em tempo real foi proposto por Kowalczuk, Psota e Perez (2013) e utilizou uma aproximação de duas passagens de agregação de peso de suporte adaptativo e uma técnica de refinamento de disparidade iterativa de baixa complexidade. Tal método produziu uma aproximação precisa dos pesos de suporte, reduzindo bastante a complexidade da agregação. A técnica de refinamento, construída usando uma estrutura probabilística, incorporou um termo aditivo na minimização de custo correspondente e facilitou o processamento iterativo para melhorar a precisão do mapa de disparidade.

Herb (2015) propôs uma implementação em tempo real do algoritmo de correspondência SGM com extensões algorítmicas de aplicações automotivas em uma plataforma de hardware reconfigurável buscando um baixo consumo de energia (inferior a 3 Watts).

Banz et al. (2010) trouxeram uma nova arquitetura e a implementação correspondente de um sistema de visão estéreo que abrange todo o processo de visão estéreo, incluindo redução de ruído, retificação, estimativa de disparidade e visualização. A estimativa de disparidade densa é realizada usando a transformação de classificação não paramétrica e correspondência SGM. Apesar de suas dependências de dados complexas e não alinhadas à varredura, o processamento baseado em fluxo do SGM é obtido por uma arquitetura escalável baseada em matriz sistólica. Essa arquitetura atende às demandas de aplicações do mundo real em relação à taxa de quadros, resolução de profundidade e baixo uso de recursos.

A correspondência de imagens baseada em *features* (características) é um método para solucionar o problema de correspondência entre duas ou mais imagens (CHEN; ROTTENS-TEINER; HEIPKE, 2021). Os pontos conjugados são um requisito para a estimativa dos parâmetros de orientação da imagem (também chamados de parâmetros de pose), que é um pré-requisito de todas as aplicações geométricas em fotogrametria, robótica e visão computacional envolvendo três dimensões. A reconstrução 3D a partir de múltiplas imagens, Localização e Mapeamento Simultâneos (SLAM), Estrutura de Movimento ou *Structure from Motion* (SfM) e a geração de mosaicos de imagens dependem de coordenadas de imagem de *features* conjugados correspondentes. Portanto, a qualidade dos algoritmos de correspondência é vital para a estabilidade e qualidade da solução desses problemas.

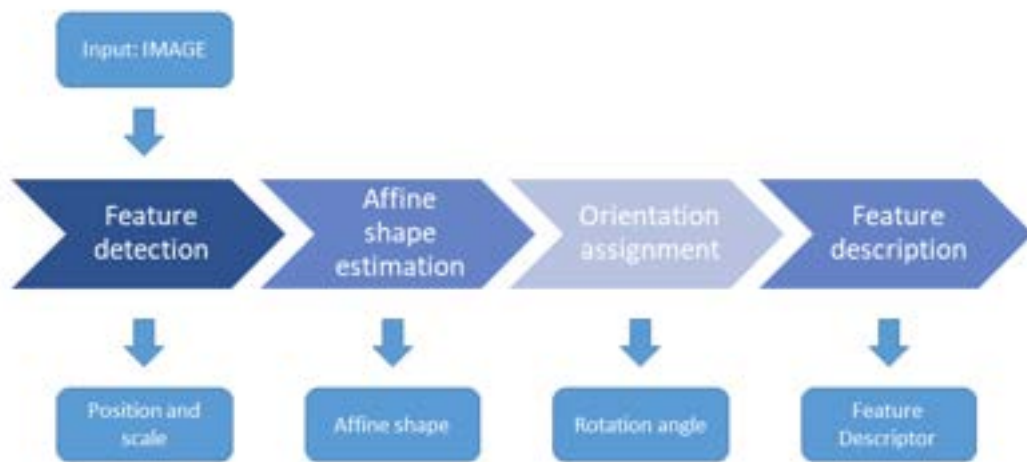
A correspondência de imagem baseada em *features* consiste em cinco etapas: detecção de *features*, estimativa de forma afim, atribuição de orientação, descrição de *features* e correspondência dos descritores (SZELISKI, 2022). O *pipeline*⁷ básico de extração de

⁷Em engenharia de software, uma canalização ou *pipeline*, em inglês, consiste de uma cadeia elementos de processamento, organizados de forma que a saída de cada elemento seja a entrada do próximo. O nome é uma analogia a qualquer canalização física.

features e descritores de imagens locais é mostrado na Figura 2.7. Antes de fornecer uma discussão mais detalhada de cada etapa, segue uma visão geral da correspondência de imagens baseada em *features*.

Os operadores de interesse em várias escalas detectam *features* também em várias escalas e, em seguida, os combinam entre estas escalas (CHEN; ROTTENSTEINER; HEIPKE, 2021). Um método mais avançado é analisar as características usando a teoria do espaço de escala, que descreve o espaço de escala em alguma escala t como uma convolução da imagem original com a função gaussiana bidimensional com variância t . Ao alterar t continuamente em vez de em etapas discretas, a escala torna-se uma variável de uma função que mapeia a imagem original para o espaço de escala. A soma das derivadas de segunda ordem da função gaussiana nas direções x e y , ou seja, o Laplaciano de Gaussiano (LoG), normalizado pela variância do Gaussiano, é usado para calcular os extremos locais no espaço de escala e esses extremos são selecionados como *features*, que agora são invariáveis em escala.

Figura 2.7: *Pipeline* de detecção e descritores de *features*



Fonte: Chen, Rottensteiner e Heipke (2021)

A transformada de característica invariante de escala SIFT (do inglês *Scale Invariant Feature Transform*) é um método para detectar e combinar pontos-chave robustos, que foi desenvolvido por Lowe (2004). A singularidade do SIFT é que esses *features* são extremamente distintos e podem ser combinados com sucesso entre imagens com mudanças de iluminação, rotação, ponto de vista e escala muito diferentes. Sua alta repetibilidade e alta taxa de correspondência em condições muito desafiadoras fizeram do SIFT um dos melhores detectores de *features* (CHEN; ROTTENSTEINER; HEIPKE, 2021; SIEGWART; NOURBAKSH; SCARAMUZZA, 2004). Foram encontradas muitas aplicações em reconhecimento

de objetos, mapeamento e navegação robótica, costura de imagens (*image stitching*), por exemplo, panoramas, mosaicos, modelagem 3D, reconhecimento de gestos, rastreamento de vídeo e reconhecimento de rosto. A principal vantagem das características do SIFT é que trata-se de um descritor calculado a partir da região ao redor do ponto de interesse, que descreve distintamente a informação transportada pela feição. Nele, o LoG (*Laplacian of Gaussian*) normalizado é aproximado pela Diferença de Gaussiana (DoG, do inglês *Difference of Gaussian*) e a localização subpixel é obtida ajustando uma quadrática 3D local à vizinhança dos extremos no espaço de escala. Este descritor é um vetor que representa a distribuição local dos gradientes da imagem ao redor do ponto de interesse. Conforme comprovado por seu [Lowe \(2004\)](#), na verdade é esse descritor que torna o SIFT robusto à rotação e pequenas mudanças de iluminação, escala e ponto de vista. Atualmente, o SIFT é um dos operadores mais conhecidos para detecção de *features* e funciona bem na correspondência de imagens com mudança de escala. Ele também pode tolerar uma certa quantidade de *affine distortion*.

Um canto em uma imagem pode ser definido como a interseção de duas ou mais arestas. Os cantos são *features* com alta repetibilidade. Um dos primeiros detectores de canto ou *corner detectors* foi proposto por Moravec ([MORAVEC, 1977](#); [MORAVEC, 1979](#)). Ele definiu uma curva como um ponto onde há uma grande variação de intensidade em todas as direções.

[Harris e Stephens \(1988\)](#) melhoraram o detector de canto Moravec considerando as derivadas parciais do SSD (*Sum of Squared Differences*) em vez de usar janelas deslocadas. Eles propuseram que um valor de canto, que é calculado como $Det(M) - INFINITO Trace(M)$, onde INFINITO é um determinante e traço de balanceamento variável, no qual os pontos de interesse são encontrados comparando o canto calculado com um limiar. Desta forma, ao invés de calcular os autovalores de M, é necessário apenas avaliar o determinante e o traço de M. O valor do determinante deve ser definido empiricamente. A última etapa do detector de cantos de Harris consiste em extrair os máximos locais da função de cantos, usando supressão não máxima. Finalmente, apenas os máximos locais que estão acima de um determinado limite são retidos. Os resultados comprovaram que a repetibilidade do detector Harris sob rotações e pequenas mudanças de iluminação é elevada.

Em [Mikolajczyk e Schmid \(2004\)](#) um mecanismo de seleção de escala é adicionado ao detector de canto de Harris. O LoG sobre a escala é avaliado em cada ponto de Harris detectado e os pontos para os quais o LoG é identificado como um extremo são preservados. Isso é seguido por um refinamento iterativo opcional para escala e posição.

Já o detector Hessian Laplace ([MIKOLAJCZYK; SCHMID, 2004](#)) é semelhante a esse último método, porém inicia-se a partir de pontos detectados usando a matriz Hessiana, na qual os valores de cinza em uma janela local também pode ser usada para detectar caracte-

terísticas (LINDBERG, 1998). Com base no determinante e traço da matriz Hessiana, são derivados critérios de seleção de características semelhantes aos aplicados à matriz de autocorrelação.

O detector FAST (Features from Accelerated Segment Test) é um detector de canto (*corner detector*) e foi desenvolvido por Rosten, Porter e Drummond (2010). O FAST compara pixels apenas em um círculo de 16 pixels ao redor do canto selecionado e usa aprendizado de máquina para acelerar o processo de comparação. Isso resulta em um detector muito eficiente que é até trinta vezes mais rápido que o Harris e é um dos detectores de *features* mais eficiente disponível do ponto de vista computacional. No entanto não é robusto em altos níveis de ruído, já que as comparações são executadas apenas em pixels discretos e a precisão da localização não pode ser refinada para o nível de subpixel. É empregado principalmente em aplicações nas quais a velocidade é essencial, mas não é necessária uma alta precisão de localização.

Devido às diferentes condições de iluminação, o formato 3D da superfície do objeto e as funções de reflexão tornam-se potencialmente complexos, portanto analisar as diferenças de valores de cinza entre imagens usando transformações matemáticas explícitas ou características de *design* de maneira intuitiva pode ser inviável. Uma alternativa é considerar a detecção de *features* como uma tarefa de aprendizado de máquina (CHEN; ROTTENSTEINER; HEIPKE, 2021).

O principal fundamento de tais métodos é mapear uma imagem de entrada para um mapa de pontuação em que o valor (pontuação) para cada pixel pode ser interpretado como a probabilidade de ser uma característica distintiva. Os parâmetros das funções de mapeamento usadas neste processo são determinados a partir de dados de treinamento por técnicas de aprendizado de máquina, amplamente utilizadas em fotogrametria e sensoriamento remoto atualmente (HEIPKE; ROTTENSTEINER, 2020).

O detector LIFT (*Learned Invariant Feature Transform*) (YI et al., 2016) prevê um mapa de pontuação cujos máximos são pontos com alta repetibilidade sob mudanças de iluminação desafiadoras. As características que se espera serem mais estáveis contra a mudança de iluminação são extraídas por supressão não máxima. O LIFT foi projetado para melhor discriminar entre pares de *features* combinados e não combinados de maneira global. A ideia central do detector covariante é que os *features* detectados no *patch*⁸ da imagem original e que são transformados em outro sistema de coordenadas, usando alguma transformação geométrica, devem ter as mesmas posições nesse sistema de coordenadas que os *features* detectados após a aplicação da transformação geométrica no *patch* da imagem original.

⁸*Patch* trata-se de função para “remendar” ou corrigir algo que, por algum motivo, não está funcionando do jeito que deveria dentro de um determinado software.

Rosten, Porter e Drummond (2010) usam aprendizado de máquina para melhorar a velocidade e a repetibilidade da detecção de *features* com base em uma comparação de valores de cinza de pixels na vizinhança.

Após a detecção das *features*, a orientação de um *feature* pode ser estimada calculando uma direção principal usando os gradientes calculados em uma janela local ao redor do *features* detectado. Em SIFT (LOWE, 2004), um histograma de gradientes orientados é calculado. A caixa com a contagem máxima é então escolhida e a direção correspondente é refinada, ajustando-se uma parábola ao pico e às duas caixas adjacentes. Outros *bins* com valores altos, ou seja, maiores que 80% do valor máximo do bin, são retidos como orientações principais secundárias.

Features com vários picos na janela de suporte podem ser melhor combinados dessa maneira. Em SURF (*Speeded-Up Robust Feature*) (BAY et al., 2008), as respostas da *wavelet* de Haar nas direções horizontal e vertical dentro de uma janela circular ao redor do *feature* detectado são computadas e plotadas como pontos em 2D. As respostas dentro de um cone rotativo de tamanho π igual a 3 são somadas e a direção principal é atribuída à direção do cone com o resultado mais alto. Além disso, o gradiente médio em uma pequena janela ao redor do *feature* detectado provou ser útil no alinhamento de *features* (BROWN; SZELISKI; WINDER, 2005).

Em (YI et al., 2016), a orientação é estimada por aprendizado profundo. A direção principal para um *patch* de entrada em torno de um *feature* detectado é prevista por uma Rede Neural Convolutiva Siamesa (CNN), que maximiza a similaridade dos descritores calculados para pares de *patches* de *features* de entrada conjugados. Uma ideia semelhante é utilizada em Mishkin, Radenovic e Matas (2018) e Chen, Rottensteiner e Heipke (2016) para aprender a orientação de *features* locais. Esta estratégia apresenta um desempenho significativamente melhor do que os métodos mencionados acima com base em *features* "artesanal".

No entanto, esta transformação não é suficiente para modelar as transformações geométricas entre manchas de imagem locais no caso de grandes mudanças no ponto de vista e direção de visão entre as imagens (CHEN; ROTTENSTEINER; HEIPKE, 2021). Portanto, as mudanças de perspectiva, que para pequenas janelas podem ser compensadas por uma transformação *affine*, também devem ser estimadas e levadas em consideração antes da descrição das *features*.

Detectar feições locais no espaço de escala e atribuir uma orientação a elas é basicamente equivalente a normalizar a rotação e o dimensionamento de feições locais antes da descrição. Entretanto, uma quantidade considerável de *features* é removida após o algoritmo de adaptação *affine* iterativo. De acordo com Mikolajczyk e Schmid (2004), apenas 20-

30% dos *features* detectados inicialmente são preservados para posterior correspondência. Para resolver esse problema, a estimativa de forma *affine* baseada em uma rede neural profunda é proposta em [Mishkin, Radenovic e Matas \(2018\)](#), onde os parâmetros de forma são estimados minimizando a distância entre os descritores de pares de pontos correspondentes. Um modelo para prever a forma *affine* também é aprendido em [Chen, Rottensteiner e Heipke \(2021\)](#). Além disso, é usado para combinar imagens tiradas de câmeras aéreas oblíquas, resultando em uma notável melhoria de desempenho em comparação ao uso de algoritmos "artesanal" para remover a distorção *affine*.

A ASIFT (*Affine SIFT*) ([MOREL; YU, 2009](#)) primeiro simula diferentes versões da imagem de entrada aplicando diferentes conjuntos de parâmetros de transformação *affine*. Em uma segunda etapa, as *features* DoG e os descritores SIFT detectados em cada imagem transformada são combinados para correspondência de descritores, um processo que torna o ASIFT, computacionalmente, onerosa. A ASIFT, portanto, não toma medidas algorítmicas para estimar a forma *affine* local para cada *feature*, mas torna a correspondência mais invariável em relação a distorções *affine*, aplicando algoritmos padrão a diferentes visualizações simuladas.

Os rastreadores de *features* mais populares podem ser classificados entre detectar-descrever-associado (DDA, do inglês *detect-describe-associate*) ou detectar-rastro (DT, do inglês *detect-track*) ([ABELES, 2013](#)). Os rastreadores baseados em DDA operam detectando localizações e características de *features*, descrevendo cada *feature* com informações de imagem local e associando *features* entre imagens. Numerosos detectores de características (Harris, Shi-Tomasi, FAST, MSER) e descritores de características locais (SIFT, SURF, BRIEF) estão disponíveis para rastreadores DDA. A associação é feita selecionando pares de descritores de características entre duas imagens que minimizam uma métrica de erro. Os rastreadores DDA são robustos a obstruções temporárias e mudanças abruptas de visão, mas são computacionalmente custosos.

[Abeles \(2013\)](#) apresentou um trabalho onde analisa e compara o desempenho de detectores de *features* do tipo DDA e DT, além do tipo híbrido que combina ações destes dois tipos de forma conjunta. Em seu trabalho de investigação utilizou cenas reais de diferentes ambientes (no ar, não submarinos) variando as condições de iluminação e vistas. O resultado obtido ratifica o comportamento de cada detector, citado no parágrafo anterior, demonstra a capacidade do tipo híbrido em reconstruir a cena e compara o desempenho computacional de cada um deles, onde o tipo híbrido apresenta-se 1,5 vezes mais lento que o DT e 7,5 vezes mais rápido que o DDA, por meio da utilização da variante FH-SURF-KLT. Informações sobre precisão e capacidade de geração de nuvens de pontos densas e reconstrução 3D do objeto não foram abordadas no trabalho.

O algoritmo de rastreamento KLT é usado para seleção de *features* e rastreamento de

objetos (MSTAFA; ELLEITHY, 2016; TOMASI; KANADE, 1991). Ele opera encontrando bons pontos de *features* (detector Harris) na área facial a partir do primeiro quadro. Esses pontos de *features* são rastreados em todos os quadros de vídeo. Cada ponto de *features* tem um ponto correspondente entre duas imagens consecutivas. Com isso, o deslocamento dos pares de pontos correspondentes pode ser calculado como vetores de movimento.

O algoritmo de rastreamento de *features* KLT apresentado no trabalho de (DEVASAGAYAM; BOSMA; COLLIER, 2022) mostrou-se altamente confiável e computacionalmente eficiente.

Multiview stereo (MVS) busca reconstruir uma representação 3D completa de um objeto ou cena a partir de uma série de imagens tiradas de pontos de vista de câmeras conhecidos. O surgimento de um grande número de ferramentas de aquisição de dados promoveu a aplicação desta tecnologia de reconstrução 3D e atende a diferentes necessidades práticas Zhang et al. (2021).

2.5 Considerações - correlação dos trabalhos e identificação das limitações

O projeto de um sensor 3D para aplicações subaquáticas depende fortemente das características da tarefa a ser realizada. A decisão sobre seguir com um ToF ou um escâner de triangulação é determinada principalmente pelo *trade-off*⁹ entre erro de medição de profundidade e alcance. Um sensor ToF é mais adequado para alcances de até algumas dezenas de metros, enquanto um sensor de triangulação pode ser capaz de alcançar erros de medição milimétricos ao trabalhar em espaços mais curtos, sendo portanto mais adequado ao propósito definido para a tese. Onmek et al. (2017) recomendaram como trabalhos futuros a utilização de um sistema estéreo para obtenção de melhores resultados numa reconstrução 3D.

Outras necessidades de projeto diz respeito à resolução, tempo de aquisição e simplicidade estrutural. Os sistemas de disparo único adquirem toda a cena de uma vez, o que os torna mais adequados para cenas altamente dinâmicas. Os escâneres de linha de direção alcançam nuvens de pontos muito densas, mas se eles estiverem conectados a um ROV, o movimento do veículo pode distorcer os dados resultantes. Daí, a utilização de técnicas que levem este fato em consideração torna-se um caminho de solução a ser avaliado por meio de SfM.

⁹ *Trade-off* é o termo da língua inglesa que define uma situação em que há conflito de escolha.

As principais desvantagens dos sistemas de visão subaquática são o alcance limitado, a necessidade de boa iluminação da cena observada e o fenômeno de refração que influencia a imagem observada. A refração, como principal fonte de distorção da imagem, é o maior obstáculo no caminho para o uso de algoritmos comuns de visão 3D subaquática.

A utilização de técnicas de projeção de linhas *laser* permitem obter informação precisa, sendo possível construir modelos de estruturas submersas 3D com alta resolução, independentemente da sua dimensão ou forma. Devido ao fato de utilizarem a projeção de linhas *laser* em vez de ondas acústicas, estes sistemas, quando otimizados, permitem reduzir o tempo necessário para aquisição de imagem, assumindo que as condições do meio são favoráveis para a propagação da luz.

A caracterização do sistema estéreo combinado com projeção de *laser* em cenas de baixa textura, diante da classificação proposta por [Castillón et al. \(2019\)](#) e apresentada na seção 2.1, combina características de um sistema ativo - projeção de *laser*, com as de um sistema passivo - câmeras estéreo.

Ressalta-se que sistemas estéreo passivos e estéreo fotométrico têm como principais vantagens o seu baixo custo e sua alta resolução, que é limitada principalmente pela luz dispersa. Já suas principais desvantagens são o fato de exigirem alto poder computacional e dependerem muito da textura do alvo para extrair *features* da imagem.

Desta forma, considerando as robustez dos resultados apresentados, o trabalho desenvolvido por [Palomer et al. \(2019\)](#) foi identificado como uma referência adequada como base do trabalho para definição do *hardware* e visão estéreo do protótipo, devido, principalmente, à simplicidade e robustez de sua construção. Já no que diz respeito às técnicas aplicadas à combinação de *softwares* e soluções para reconstrução 3D, foram tomadas outras referências, as quais podem ser verificadas logo abaixo (Tabela 2.3).

Tabela 2.3: Trabalhos selecionados

Trabalho	Método
Palomer et al. (2019)	Visão estéreo
Łuczyński, Pfingsthorn e Birk (2017b)	Calibração - PINAX
Özyeşil et al. (2017)	Reconstrução 3D - SfM
Lowe (2004)	Detecção de <i>features</i> - SIFT
Shi e Tomasi (1994)	Detecção de <i>features</i> - DDA
Abeles (2013)	Detecção de <i>features</i> - Híbrido
Mstafa e Elleithy (2016)	Detecção de <i>features</i> - KLT
Zhang et al. (2021)	Reconstrução 3D - MVS

Fonte: Autor.

Levando-se em consideração o fato do ROV não conseguir manter-se estável (ou imóvel) no momento em que as inspeções (tomadas) são realizadas e pela possibilidade de utilização de apenas imagens RGB, os métodos de reconstrução 3D *Structure from Motion* (SfM) (ÖZYEŞIL et al., 2017) e *Multi-view Stereo* (MVS) (ZHANG et al., 2021) foram identificados como um possível caminho de desenvolvimento.

Já para a etapa de calibração, foi tomado como referência o modelo Pinax (LUCZYŃSKI; PFINGSTHORN; BIRK, 2017b), o qual leva em consideração os índices de refração da água, especialmente no que diz respeito à salinidade. Este modelo permite que a calibração da câmera subaquática seja feita apenas uma única vez ainda no ar, antes da imersão.

Novos algoritmos vêm sendo propostos para calcular os pontos de interesse - IPs (do inglês *interesting points*, por exemplo, *Features from Accelerated Segment Test* ou FAST (ROSTEN; PORTER; DRUMMOND, 2010). Os algoritmos frequentemente mencionados em trabalhos de pesquisa como o SIFT (LOWE, 2004; ÖZYEŞIL et al., 2017; CHEN; ROTTENSTEINER; HEIPKE, 2021), o SURF (*Speeded-Up Robust Features*) (BAY; TUYTELAARS; GOOL, 2006; BAY et al., 2008; CHEN; ROTTENSTEINER; HEIPKE, 2021) e o ASIFT (*Affine-SIFT*) (MOREL; YU, 2009) baseiam-se em variantes dos detectores Hessian e Harris.

Neste trabalho de investigação, é proposto um modelo híbrido baseado nos detectores de *features Detect-Describe-Associate* - DDA (SHI; TOMASI, 1994) e *Kanade-Lucas-Tomasi* - KLT (MSTAFSA; ELLEITHY, 2016), buscando referências nos resultados apresentados por Abeles (2013) em seu trabalho e, a partir deste último principalmente, desenvolvendo toda proposta de um sistema computacional capaz de reconstruir objetos e equipamentos utilizados na indústria de O&G.

No entanto, a exploração de novas técnicas e algoritmos, associados à necessidade de realizar operações de digitalização para inspeção de equipamentos subaquáticos, promovem o crescimento da exploração subaquática a partir da utilização de sistemas robóticos, tornando-os cada vez mais robustos e versáteis corroborando, portanto, com o objetivo do desenvolvimento aqui proposto.

Outro fato observado, após análise dos métodos e resultados obtidos nos diversos trabalhos analisados, é que a realização de testes em condições de maior relevância, ou seja, execução de testes em condições mais próximas das reais onde sistemas de digitalização são necessários, é um passo recomendado (SKINNER; ISCAR; JOHNSON-ROBERSON, 2017) como forma dar robustez e confiabilidade aos métodos propostos. Tal etapa de atividades faz parte dos objetivos específicos definidos na seção 1.3.

Projeto e desenvolvimento da pesquisa

Este capítulo apresenta o projeto, arquitetura e implementação do sistema computacional integrado a modelos digitais para reconstrução de estruturas subaquáticas com alta definição¹, desenvolvido no âmbito deste trabalho de investigação. O sistema computacional foi projetado com características modulares para facilitar a utilização de parte ou de todos os componentes do sistema.

A definição de um conceito de um sistema de digitalização, elaboração do projeto detalhado e construção de um protótipo e execução de testes em ambiente relevante foram as etapas principais do projeto demandado pela Petrobras, culminado com a geração de nuvens de ponto brutas. Conforme já mencionado no Capítulo 1, o propósito desta tese visa ampliar o escopo e, por meio de uma nova rota para obtenção das nuvens brutas, gerar a reconstrução 3D de estruturas e equipamentos instalados no leito do oceano, com execução de testes e geração de um *dataset* em ambiente operacional.

Daí, neste capítulo, é dado maior ênfase ao sistema computacional proposto, sendo o desenvolvimento do projeto de *hardware* inserido no Apêndice A.

O sistema computacional desenvolvido neste trabalho de investigação utilizou as imagens brutas de uma das câmeras (esquerda) do protótipo digitalizador, as quais foram obtidas durante as fases de testes em ambientes relevante e operacional.

3.1 Sistema Computacional - DIGISUB

O sistema computacional recebeu a mesma denominação do projeto, DIGISUB, e está ilustrado na Figura 3.1. Foi projetado adotando premissas técnicas de reconstrução 3D baseadas nos algoritmos de *Structure from Motion* (SfM) (ÖZYEŞİL et al., 2017) e *Multi-view Stereo* (MVS) (BROEKMAN; GRÄBE, 2020; BAILER; FINCKH; LENSCH, 2012; SEITZ et al., 2006).

A proposta do DIGISUB é baseada nos rastreadores de *features Detect-Describe-Associate* - DDA (SHI; TOMASI, 1994) e *Kanade-Lucas-Tomasi* - KLT (DEVASAGAYAM; BOSMA; COLLIER, 2022; MSTAFA; ELLEITHY, 2016). Esta abordagem herda os benefícios de cada uma

¹O termo alta definição neste trabalho refere-se a um nível de qualidade da reconstrução do objeto que permita a identificação de seus detalhes e características e que, ao mesmo tempo, permita obter as dimensões de suas cotas com precisão milimétrica.

delas, as quais foram a base para sua elaboração e, possivelmente, também seus defeitos. Mas por meio de uma atuação híbrida (ABELES, 2013), que é descrita em detalhes na sequência, busca uma reconstrução completa da cena com menor esforço computacional.

Ressalta-se que anteriormente, no âmbito da pesquisa e trabalhos investigados, nenhuma outra abordagem foi aplicada e submetida a testes em um cenário operacional subaquático para digitalizar equipamentos complexos e de grandes dimensões, como uma Árvore de Natal Molhada (ANM) e embarcada num ROV em operação em águas profundas.

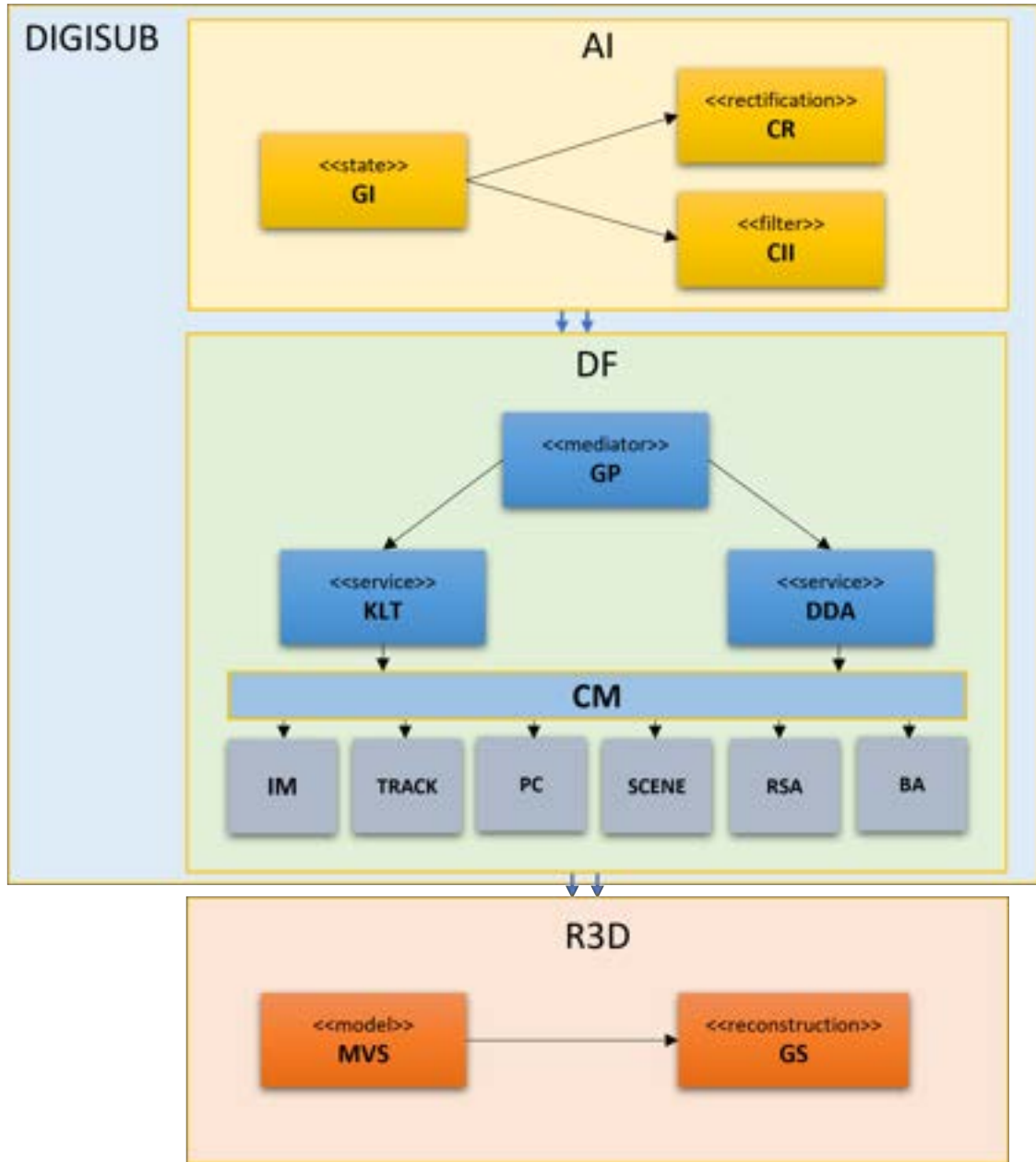
Esta estratégia foi adotada levando-se em consideração dois motivos principais:

- Primeiro, devido ao fato do ROV, portador do protótipo do sistema, não conseguir manter-se estável (ou imóvel) no momento em que as inspeções para tomadas das imagens eram realizadas, isto em função das perturbações ocasionadas pelo próprio ambiente operacional.
- E o segundo, pelos métodos de reconstrução 3D, SfM e MVS, fazerem uso de imagens RGB bidimensionais sem a necessidade do uso de *hardware* específico.

Inicialmente, as imagens brutas são lidas e armazenadas na memória do computador, caracterizando a etapa de **Aquisição de Imagem (AI)**. Logo em seguida ocorre a **Detecção de *Features* (DF)**, com o objetivo de reconstruir a geometria da cena como uma nuvem de pontos de alta definição. Dado um conjunto de imagens, que retratam a mesma cena de diferentes pontos de vista, pode-se recuperar o posicionamento das câmeras diante da cena (pose) utilizado para aquisição e reconstrução da geometria do equipamento inspecionado por meio de uma nuvem de pontos esparsa. Então, para a **Reconstrução 3D (R3D)** propriamente dita, o método MVS utiliza a saída do método SfM para gerar um modelo 3D denso de alta definição do objeto inspecionado.

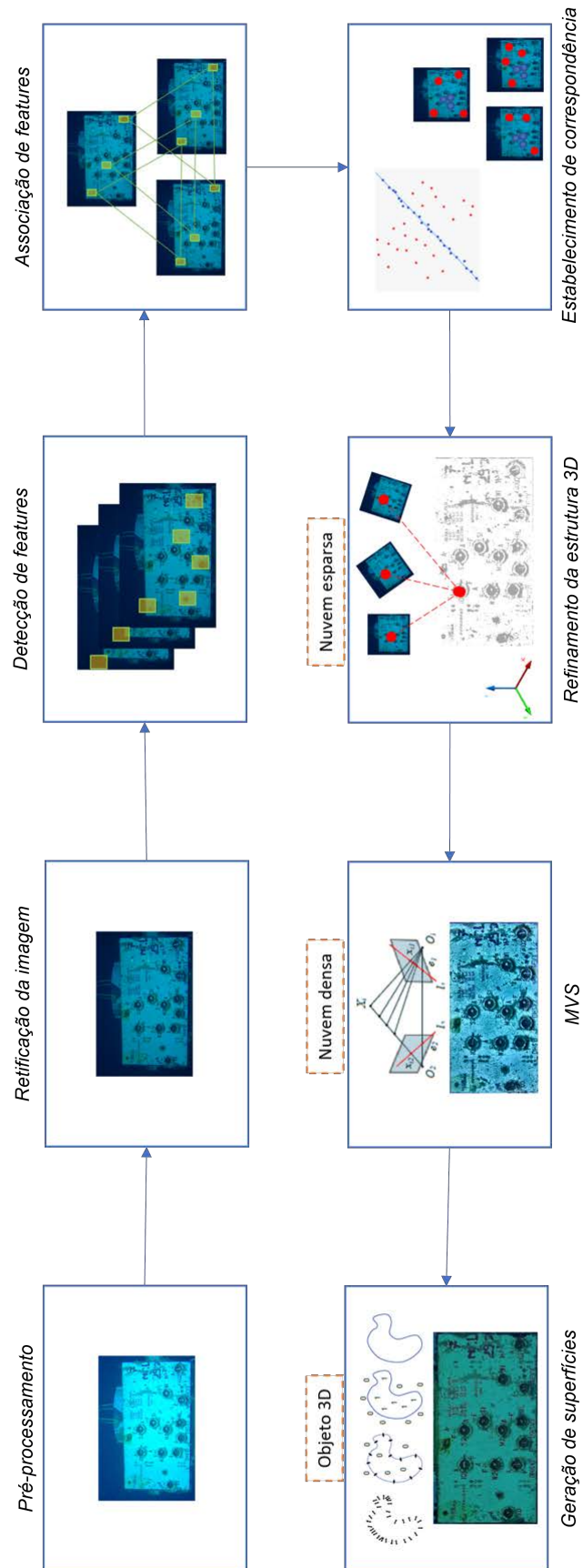
O processo completo do DIGISUB é composto por diferentes etapas de processamento e os componentes do sistema computacional possuem funções específicas, descritas na Figura 3.1, que representa sua arquitetura, além da Figura 3.2, que apresenta o fluxograma com a sequência de execução destas etapas. Uma descrição detalhada de cada uma delas é encontrada nos tópicos seguintes.

Figura 3.1: Arquitetura do sistema computacional - DIGISUB



Fonte: Autor.

Figura 3.2: Fluxograma do sistema computacional - DIGISUB



Fonte: Autor.

3.1.1 Módulo de aquisição de imagem

O módulo AI é composto de três componentes que cooperam para prover os mecanismos de pré-processamento e retificação das imagens. O componente GI (Gestão de Imagens) é responsável pela manutenção dos componentes CR (Correção de Refração) e CII (Compensação de Inhomogeneidades de Iluminação).

A aquisição de imagens foi realizada por meio do protótipo DIGISUB. O protótipo utiliza porta plana, que causa distorções que não são triviais e estão baseadas em refração na interface ar-vidro-água, mas são de mais simples construção, mais baratas e podem adaptar-se a diferentes tipos de câmeras. Porém apresenta a desvantagem de introduzir distorções devido à refração ar-vidro-água. A maneira para lidar com isto é utilizar um modelo de projeção de perspectiva padrão e calibração adequada. O cálculo da compensação está inserido no algoritmo de calibração que, no caso do DIGISUB, foi previamente realizada em terra, antes do protótipo ser submergido, por meio do método disponibilizado por [Kaehler e Bradsky \(2017\)](#). Nele, as matrizes de calibração, contendo os parâmetros intrínsecos de cada uma das câmeras assim como do par, foram disponibilizadas uma a uma. Daí a possibilidade de utilizar-se apenas as imagens de uma das câmeras.

No processo de aquisição de imagens ocorre interferência na geração da imagem devido à introdução de ruídos pela exposição excessiva à iluminação artificial do ROV e do próprio protótipo². A intensidade e a direção da luz também podem trazer mais complexidade às condições de iluminação do cenário. Com isto, a aparência do objeto pode mudar drasticamente em função da variação de iluminação provocada pela projeção de luz proveniente do veículo. Sendo assim, para resolver este problema, o método de pré-processamento de imagem proposto por [Michálek et al. \(2010\)](#) foi implementado no componente CII e aplicado às imagens de entrada. Isso faz com que as etapas de processamento subsequentes sejam mais robustas em relação às variações de iluminação.

Para remover as distorções introduzidas pela refração, que ocorre quando a luz passa pelas janelas de vidro das caixas das câmeras subaquáticas, foi utilizado o modelo Pinax ([LUCZYŃSKI; PFINGSTHORN; BIRK, 2017b](#)), implementado no componente CR. Este modelo leva em consideração os índices de refração da água, especialmente no que diz respeito à salinidade, e também permite que a calibração da câmera subaquática ocorra apenas uma única vez ainda no ar, antes de sua imersão.

Dadas as características construtivas do protótipo, pode-se assumir que as condições e premissas, definidas em [Luczyński, Pfingsthorn e Birk \(2017b\)](#) e citadas abaixo, foram atendidas conforme detalhamento apresentado na Seção do Apêndice [A.0.2](#), mais especi-

²No protótipo DIGISUB foram instaladas duas lâmpadas de LED de 36 W com capacidade de emissão de 5.000 lúmens cada uma.

ficamente na Figura A.11, na qual podem ser observadas as distâncias e posicionamento da câmera em relação à janela de visualização.

- O plano refrativo normal e, portanto, o eixo da câmera, para o modelo axial, é paralelo ao eixo óptico da câmera (parâmetro intrínseco).
- O eixo óptico da câmera é perpendicular à superfície de vidro (parâmetro intrínseco).
- A configuração típica de câmeras subaquáticas, em que a extremidade frontal das lentes é instalada o mais próximo possível da janela frontal. A distância entre a janela de vidro e o centro de projeção da câmera física é pequena, ou seja, concretamente assume-se que está na faixa de 0 a 10mm (parâmetro intrínseco).
- A espessura do vidro da janela e seu índice de refração são conhecidos (parâmetro intrínseco).
- O índice de refração da água está dentro de parâmetros conhecidos (ROSWELL; HALIKAS, 2010).

As premissas adotadas sobre a configuração do *hardware* permitem tratar uma câmera física como se estivesse na posição ideal em relação ao painel de vidro, sem causar erros significativos. Essas suposições e o conhecimento adicional dos parâmetros intrínsecos e extrínsecos do sistema completo de câmeras físicas no ar permitem calcular os parâmetros intrínsecos e extrínsecos virtuais para aplicação subaquática. Para visão estéreo, as imagens devem ser corrigidas para remover a distorção e refração da lente e retificadas, tornando possível a aplicação do método Pinax, desenvolvido por [Luczyński, Pfingsthorn e Birk \(2017b\)](#). As imagens (direita e esquerda) obtidas são então submetidas ao método, que corrige os efeitos da refração e as disponibiliza novamente para sequência do processamento. Lembrando que o DIGISUB utiliza as imagens de apenas uma das câmeras, no caso, a esquerda. A vantagem de se usar uma única câmera dá-se por: simplificação do *hardware*; sistema, uma vez baseado em SfM, ser mais adequado às características da operação, em que o ROV não se estabiliza completamente enquanto faz as tomadas das imagens (ele não permanece imóvel, fixo a um local, e 3); por sempre ocorrer pequenos movimentos na tentativa de mantê-lo estabilizado diante do objeto, proporcionando rotações nas imagens e dando condições de execução do SfM; também há uma vantagem comum ao SfM e ao MVS de usarem imagens 2D sem a necessidade de um *hardware* específico. A desvantagem ocorre principalmente pelo tempo de processamento. A resolução não é afetada, pois o DIGISUB busca a correção por meio das demais implementações de software incorporadas ao modelo e detalhadas neste capítulo. Além disto, uma vez calibrada a câmera e mantida as tomadas das imagens não há alteração na interferência sobre a incerteza de medição já determinada anteriormente por meio da etapa de Calibração. O sistema estéreo, usado no desenvolvimento do projeto Petrobras, apresentou dificuldades na reconstrução justamente pela movimentação do ROV e dificuldade na determinação

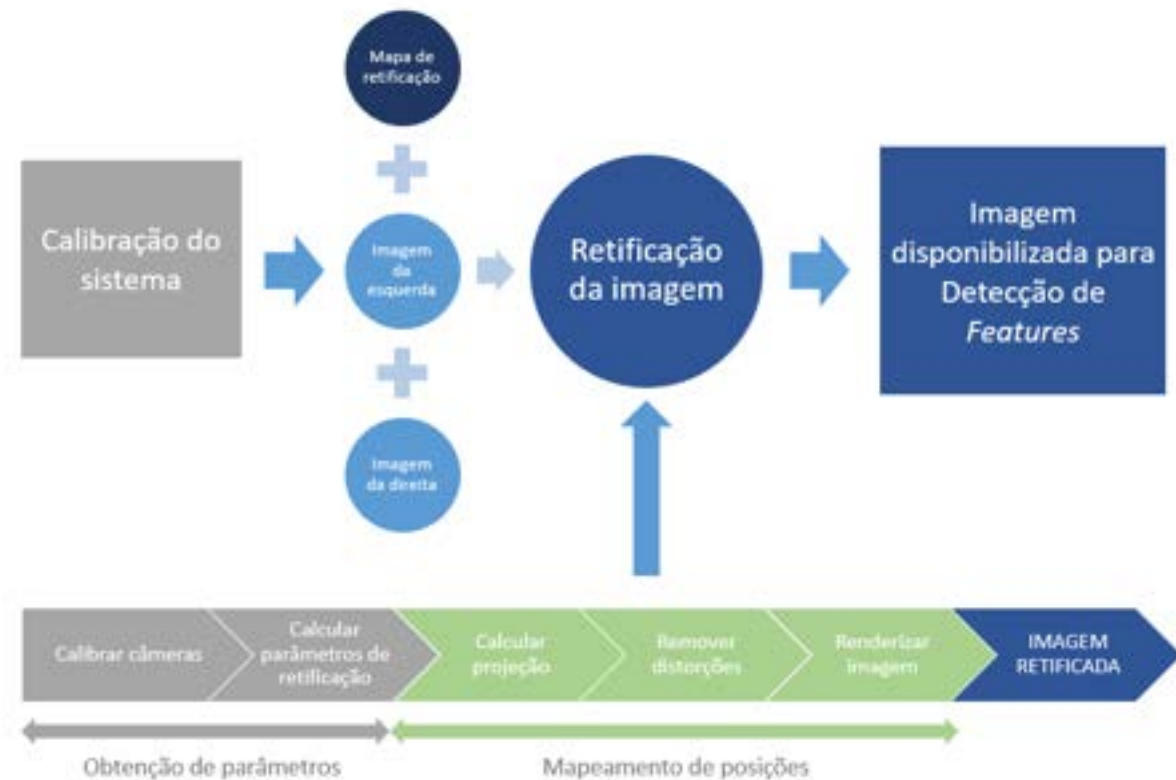
da localização por meio de compensação pela INS. Mas apresenta a vantagem de, se a referência da câmera, no caso o ROV, estiver fixada, imóvel, traz resultados precisos e com menor custo computacional.

A obtenção de parâmetros para calibração do sistema estéreo e determinação dos parâmetros de retificação está descrita abaixo e pode ser visualizada pelo fluxograma mostrado na Figura 3.3.

- Primeiro, a rotação entre as duas câmeras reais do protótipo e as duas virtuais deve ser encontrada (tanto para câmera direita, quanto para a esquerda). Para tanto, ambas as câmeras virtuais são definidas como tendo eixos ópticos ortogonais ao painel de vidro e pode ser, então, aplicado o modelo Pinax.
- Seguindo o processo para utilização do Pinax, são inseridos dados com as condições do ambiente em que serão realizados as tomadas das imagens. Como base nestas informações, o Pinax recorre ao trabalho de [Roswell e Halikas \(2010\)](#) para determinar o índice de correção pela salinidade a ser utilizado. Tais parâmetros são inseridos no cálculo da incerteza de medição, que será abordado na Subseção 4.2.2 e permitem a minimização do erro referente a tal correção, conforme descrito por [Luczyński, Pfingsthorn e Birk \(2017b\)](#).
- Entretanto, existe uma rotação conhecida (parâmetro extrínseco) entre essas câmeras que pode ser determinada pela calibração no ar. O eixo x do sistema de câmera estéreo final é definido para estar ao longo do vetor de translação t entre as câmeras. Isso garante que os eixos x da câmera resultantes estejam alinhados. O eixo z é considerado a média dos dois eixos ópticos da câmera no ar. O eixo y é definido de acordo com os eixos x e z.
- Utilizando o método de [Luczyński, Pfingsthorn e Birk \(2017b\)](#) a câmera subaquática é calibrada uma vez no ar, ou seja, sua matriz intrínseca K^3 ([SIEGWART; NOUR-BAKHSI; SCARAMUZZA, 2004](#)) é determinada usando um processo de calibração padrão ([KAEHLER; BRADSKY, 2017](#)).
- As refrações ar-vidro e vidro-ar só levam a uma mudança de escala nas imagens, que faz parte dos parâmetros extrínsecos e as relações geométricas relativas entre os pontos da cena são preservadas. Os resultados obtidos são a matriz intrínseca da câmera, os coeficientes de distorção, os vetores de rotação e os vetores de translação. Os dois primeiros constituem os parâmetros intrínsecos da câmera e os dois últimos são as medidas extrínsecas, as quais determinam onde os objetos foram encontrados e quais eram suas orientações.

³A matriz K, também denominada matriz de calibração, armazena os parâmetros intrínsecos da câmera, que estão relacionados à forma como a câmera realiza a projeção.

Figura 3.3: Processo de calibração e retificação da imagem - Pinax



Fonte: Autor.

3.1.2 Módulo de Detecção de *Features*

Após a aquisição de imagens, a segunda etapa do modelo, Detecção de *Features* (DF) pode ser executada. Este estágio consiste em vários passos sequenciais de processamento e que são melhores descritos a seguir.

Passo 1 - Detecção de *features*

A detecção de *features* é a primeira etapa de muitas aplicações relacionadas à visão computacional, como reconhecimento de objetos. Nas últimas décadas, uma ampla variedade de detectores de *features* foi desenvolvida. Além de sua eficácia, eles variam muito em complexidade computacional e no tipo de *features* que detectam. Embora existam abordagens que detectam bordas e regiões de interesse, as características de imagem usadas na maioria das abordagens SfM compreendem pontos de interesse (do inglês, *interesting points* - IPs).

IPs representam locais da imagem que são, de certa forma, excepcionais e encontram-se cercados localmente por uma textura distinta. Além disso, eles devem ser definidos de

forma estável na imagem e nos espaços de escala, além de reproduzíveis em diferentes condições de imagem. Os IPs devem ter uma alta repetibilidade, o que significa que devem ser invariantes sob qualquer mudança na iluminação, ruído da imagem e transformações geométricas básicas, como escala, translação e rotação. Para a reconstrução 3D, o algoritmo detector de características deve ser robusto o suficiente para encontrar os mesmos IPs em imagens diferentes. Portanto, a qualidade dos algoritmos de rastreadores de *features* é vital para a estabilidade e bom resultado da solução desses problemas.

Nos últimos anos, novos algoritmos foram propostos para calcular esses IPs, por exemplo, *Features from Accelerated Segment Test* ou FAST (ROSTEN; PORTER; DRUMMOND, 2010). Outros algoritmos frequentemente mencionados e utilizados em pesquisas recentes como o SIFT (LOWE, 2004; ÖZYEŞIL et al., 2017; CHEN; ROTTENSTEINER; HEIPKE, 2021), o SURF (*Speeded-Up Robust Features*) (BAY; TUYTELAARS; GOOL, 2006; BAY et al., 2008; CHEN; ROTTENSTEINER; HEIPKE, 2021) e o ASIFT (*Affine-SIFT*) (MOREL; YU, 2009) baseiam-se em variantes dos detectores Hessian e Harris.

Para o desenvolvimento do DIGISUB componentes foram projetados para atender às necessidades do sistema computacional proposto.

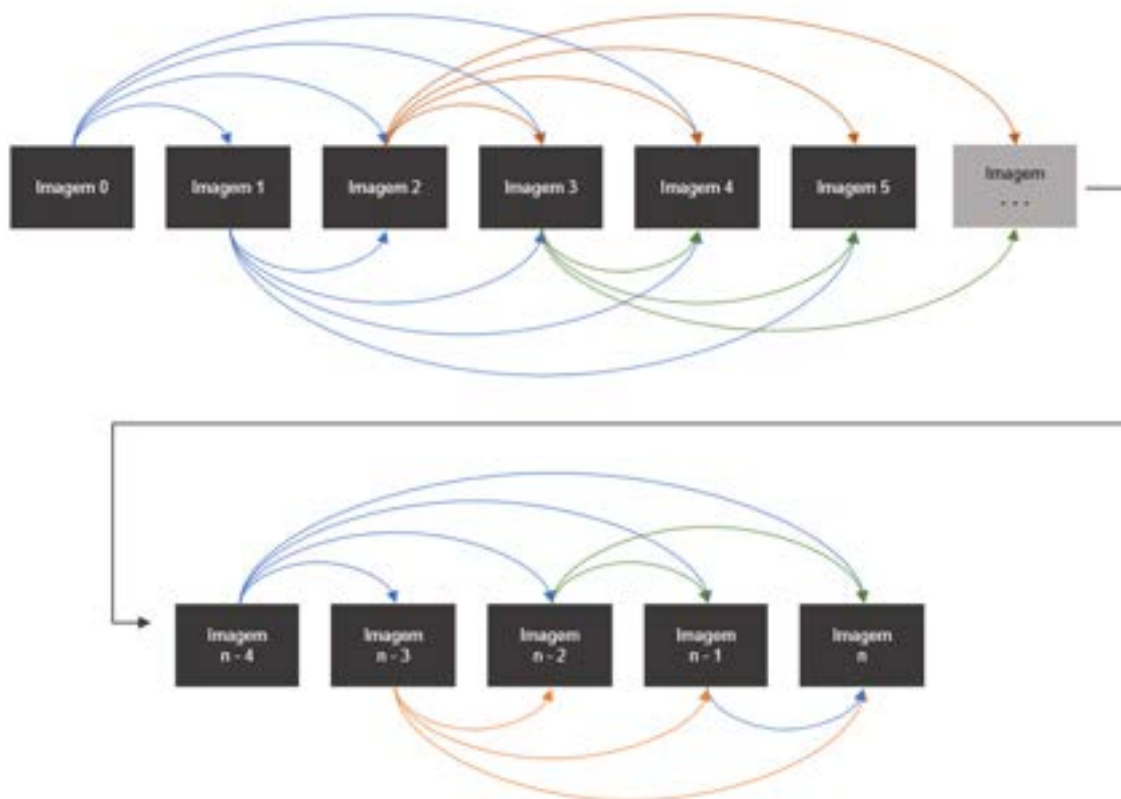
- O componente GP é responsável pela gestão do processamento, atua no escalonamento dos serviços (*services*) KLT e DDA.
- O CM (Componente Modular) serve para encontrar, carregar e descarregar os demais componentes do DIGISUB.
- O componente IM é responsável por guardar as informações da imagem, a matriz intrínseca e extrínseca e os IPs.
- O *TRACK* representa um ponto no espaço. Esse componente armazena uma lista de objetos do tipo IM que corresponde ao ponto que o *track* representa.
- Já o componente PC, representa um par do componente IM. Contém as duas imagens que formam o par e uma lista de *tracks* formadas pelos *matches* entre os IPs dos dois objetos do tipo IM.
- O *SCENE* é o componente responsável por manter informações de uma cena formada pela lista de *tracks*.
- O RSA é o componente que implementa o algoritmo de detecção de *outlier* (*Random Sample Consensus*).
- O componente BA (*Bundle Adjustment*) é responsável pelo refinamento da estrutura 3D.

No DIGISUB, o rastreamento inicia-se ao gerar novas *tracks* a partir de *features* detectadas. Uma *feature track* é definida por uma localização 2D e possui dois tipos de descrições: descrição DDA e descrição KLT.

- A descrição DDA é imutável e a descrição KLT é atualizada após cada imagem ser processada. Quando a próxima imagem na sequência é gerada, os *tracks* são atualizados primeiro por meio do rastreador KLT. Caso ocorra uma falha na atualização de KLT, o *track* é armazenado em um *buffer*.
- Os *tracks* armazenados não são atualizados à medida que novas imagens são geradas. Quando acionado, novos *tracks* podem ser gerados e os *tracks* antigos armazenados no *buffer* são reativados. Um gatilho foi projetado para aumentar a eficiência computacional e é acionado periodicamente após o processamento de N *frames*.

A Figura 3.4 ilustra de forma simplificada o processo de atuação do DIGISUB no gerenciamento entre os métodos KLT e DDA, descrito nos passos 1 e 2.

Figura 3.4: Descrição do passo Detecção de *Features*



Fonte: Autor.

Passo 2 - Associação de *features* e *tracks*

- O primeiro passo para gerar novos *tracks* ou reativar *tracks* do *buffer* é detectar as *features* das imagem.
- Depois que as *features* da imagem são detectadas, elas são, então, associadas a todos os *tracks* ativos e a todos os *tracks* armazenados no *buffer*.
- Se uma *feature* de imagem não estiver associada a um *track*, uma nova *feature* deve ser gerada a partir dela.
- Se uma *feature* do *buffer* de armazenamento estiver associada a uma *feature* de imagem, ela é reativada removendo-a do *buffer* de armazenamento e definindo sua localização para a *feature* e então, sua descrição de KLT é atualizada.
- Um algoritmo guloso do vizinho mais próximo é usado para associação nesta implementação.

No algoritmo DIGISUB, uma *feature* só é transferida para a abordagem DDA quando o rastreador KLT detecta uma falha. Os rastreadores KLT normalmente empregam métodos para detectar falhas, mas são suscetíveis à mudança gradual da cena. É possível que uma sequência de imagens, que se altera lentamente, produza um desempenho pior do que uma com mudanças abruptas. A principal diferença entre as duas abordagens envolve a forma como os *tracks* são atualizados. O DDA realiza uma pesquisa global custosa por *features*. Já o KLT realiza uma busca local para cada *track* de forma independente.

Passo 3 - Estabelecimento de correspondência

Ao contrário dos rastreadores DDA, a velocidade de execução dos rastreadores KLT depende do número de *tracks* e não, do tamanho da imagem. Geralmente, a estabilidade do *track* é melhorada à medida que há a atualização da descrição realizada após cada *frame*. A desvantagem de atualizar o descritor é que o ruído da imagem pode levar a prováveis correspondências falsas. Diante disto, um algoritmo de detecção de *outlier* robusto, como RANSAC (*Random Sample Consensus*) (FISCHLER; BOLLES, 1981; ONMEK et al., 2017), garante a rejeição de prováveis correspondências falsas, testando-as quanto à consistência. Isso é feito para todos os pares de imagens possíveis, verificando se suas correspondências atendem à restrição de geometria epipolar, ou seja, se os deslocamentos de IPs são um resultado possível apenas do movimento da câmera entre as duas imagens.

Passo 4 - Refinamento da estrutura 3D

Após a etapa de correspondência, os pares de imagens são combinados para formar o bloco completo de imagens e produzir a estrutura da cena. Desta forma, os resultados obtidos não estão otimizados, porque nem todas as imagens sobrepostas são usadas ao mesmo tempo e as discrepâncias na estrutura, causadas por pequenos erros durante a

etapa de detecção de *features*, não são minimizadas de forma otimizada. A partir disto, para superar esse problema, o estágio final da maioria dos algoritmos SfM é o *bundle adjustment*. Este algoritmo otimiza de forma iterativa a estrutura 3D de todas as imagens simultaneamente, realizando uma minimização não linear robusta dos erros de medição, também conhecidos na literatura como erros de re-projeção (TRIGGS et al., 2000). O erro de re-projeção é a distância euclidiana entre a posição em que o ponto 3D foi projetado na imagem e a posição da *feature* que originou esse ponto. O algoritmo utilizado para resolver esta questão foi o Levenberg-Marquardt (LM) (LEVENBERG, 1944) adaptado no trabalho desenvolvido por Wu et al. (2011). Este algoritmo atua solucionando uma série de aproximações lineares regularizadas da questão original.

3.1.3 Módulo de reconstrução 3D

Nesta etapa, uma nuvem de pontos 3D esparsa, baseada apenas no conjunto reconstruído de IPs, da cena que contém o equipamento inspecionado é disponibilizada. No entanto, com a orientação agora conhecida das imagens torna-se possível criar um modelo 3D denso (densidade tal quando tomada como referência as nuvens esparsas) com mapeamento de textura, cujo passo essencial neste processo é o cálculo deste modelo 3D. O módulo R3D é composto de dois componentes descritos a seguir.

Multi-view Stereo - MVS

Segundo Seitz et al. (2006), pode-se interpolar o conjunto esparsa de pontos 3D para geração do modelo 3D denso, porém isto produziria um resultado longe do ideal. Daí, foi utilizado um algoritmo baseado em MVS (ZHANG et al., 2021) para calcular uma estimativa da densidade da geometria da superfície da cena observada. Esta etapa adicional permite a geração de modelos 3D detalhados (nuvens de pontos densas) a partir das nuvens de pontos esparsas calculadas inicialmente, permitindo a reprodução dos detalhes finos presentes na cena.

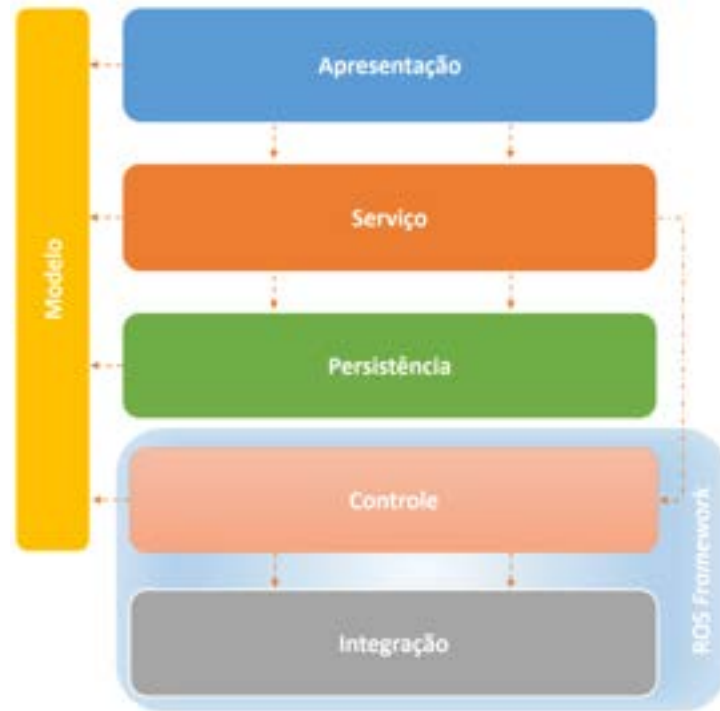
Geração de Superfícies - GS

Essa etapa é responsável pela conexão entre os pontos da nuvem de pontos 3D densa gerada pelo MVS. O método *Poisson* foi utilizado para reconstrução de superfícies. Sua escolha foi baseada no trabalho desenvolvido por Masson e Petry (2017), onde foi investigado o uso de algoritmos de geração de malha e nuvem de pontos para construir superfície 3D. O método *Poisson* apresentou malhas mais precisas ao comparar medidas quantitativas e menor consumo de tempo na geração das malhas.

3.1.4 Arquitetura de *software*

A Figura 3.5 ilustra a arquitetura de *software* do sistema DIGISUB, baseada numa divisão em camadas com características modulares para facilitar a utilização de parte ou de todos os componentes do sistema.

Figura 3.5: Arquitetura de *software* do sistema.



Fonte: Autor.

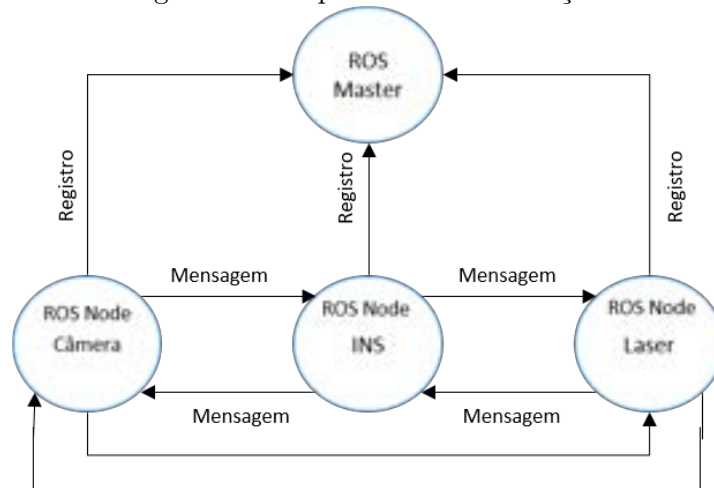
A camada de Apresentação contém os componentes de UI (*User Interface*) responsáveis pela interação do sistema com o usuário da aplicação. Esta camada comunica-se com as camadas de Serviço e Modelo. Já a camada de Serviço provê as funcionalidades referentes à lógica da aplicação, gerenciamento das classes de negócio e controle, sendo independente da camada de Apresentação. A camada de Persistência tem como objetivo implementar uma solução de acesso aos dados, provendo uma interface de métodos à camada de Serviço para consulta e gravação de objetos. A camada de Modelo contém as abstrações que representam as classes de entidades do sistema e que são utilizadas para transferir dados entre as camadas da aplicação.

A camada de Controle, por sua vez, é composta por classes responsáveis pelo gerenciamento do *hardware* do sistema. Nesta camada encontra-se o ROS⁴ *Master*, responsável por registrar e coordenar a comunicação entre os diversos nós (ROS Nodes) do sistema.

⁴*Robot Operating System* (Sistema Operacional de Robôs) é uma coleção de *frameworks* de *software* para desenvolvimento de sistemas robóticos, que fornece a funcionalidade de um sistema operacional em um *cluster* de computadores heterogêneo. Disponível em: <https://www.ros.org/>.

A Figura 3.6 ilustra o esquema de comunicação contendo quatro componentes do sistema: Câmeras, INS e *Laser*, além do *Master*. Cada nó comunica-se com outros nós por meio de mensagens, que são atribuídas a um endereço de memória (local padrão daquele nó, no qual ele recebe ou envia mensagens). Assim, este nó pode ler a mensagem em um endereço e realizar uma ação.

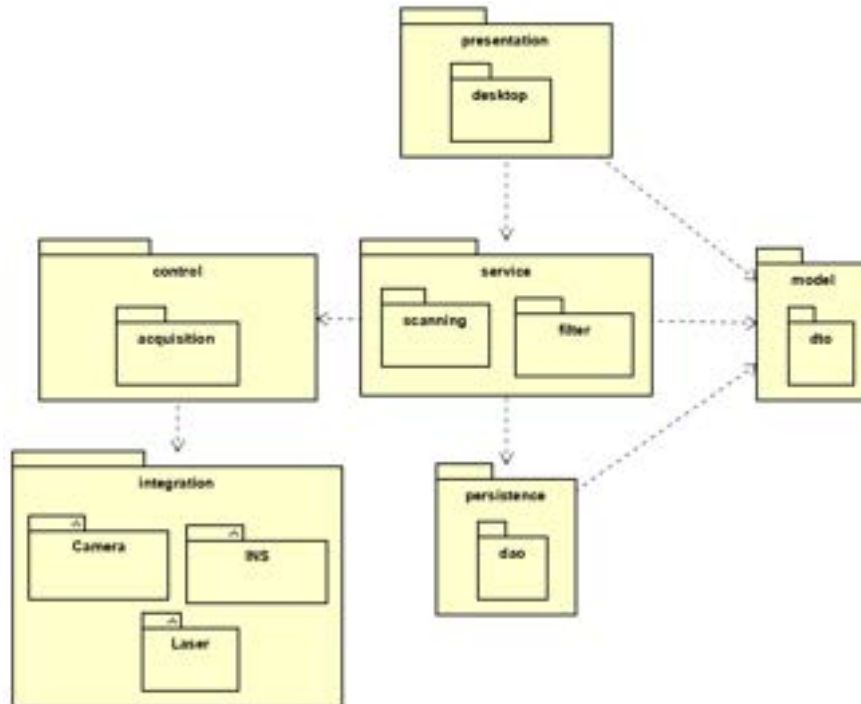
Figura 3.6: Esquema de comunicação.



Fonte: Autor.

Já a camada de Integração implementa classes, as quais oferecem serviços de abstração do *hardware* e controle de baixo nível de dispositivos (*drivers*). As classes dessa camada devem abstrair e encapsular todo acesso ao *hardware* do sistema, deixando transparente à camada de Controle o modo como os dados são obtidos, mantendo a implementação de acesso aos dados isolado e facilitando sua manutenção. A estrutura de pacotes e subsistemas da aplicação segue o formato presente no diagrama da Figura 3.7.

Figura 3.7: Visão de pacotes e subsistemas da aplicação.



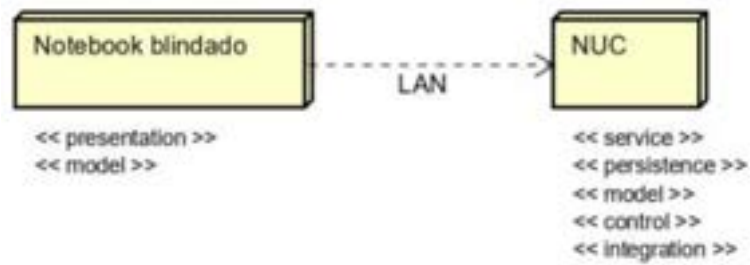
Fonte: Autor.

3.2 Implementação

Os componentes relacionados às camadas Apresentação, Serviço, Persistência e Modelo foram implementados utilizando a linguagem de programação Java. A escolha dessa linguagem é justificada por ser multiplataforma, ou seja, é independente de sistema operacional. Assim, não será preciso reescrever os códigos do sistema quando for instalá-los em outro sistema operacional, pois a conversão é feita pela Máquina Virtual do Java. Já as classes relacionadas às camadas Controle e Integração foram implementadas utilizando a linguagem de programação C++. Nesse caso, a linguagem C++ foi escolhida em razão da utilização do ROS.

A Figura 3.8 apresenta a visão de implantação da arquitetura nos dois nós físicos para as configurações do sistema. Também apresenta a alocação dos pacotes para os nós físicos.

Figura 3.8: Visão de implantação da aplicação.



Fonte: Autor.

A arquitetura proposta suporta os requisitos de uma arquitetura cliente/servidor. A seguir descrição dos aspectos dinâmicos da aplicação, existentes em cada nó, a saber:

- *Notebook* blindado - local onde é hospedado os componentes de *software* que trata das funcionalidades relacionadas à interação entre o usuário e o sistema do protótipo digitalizador. Contém a versão do Java SE *Development Kit* 8u241 e o Ubuntu 16.04.6 LTS.
- NUC - local onde é hospedado o conteúdo dinâmico da aplicação, abrangendo os componentes de *software* das camadas de Serviço, Persistência, Modelo, Controle e Integração. Contém a versão do ROS Kinetic Kame e o Ubuntu 16.04.6 LTS.

Avaliação e resultados experimentais

Neste capítulo serão apresentados os experimentos realizados em ambientes relevante (Baía de Aratu) e operacional (Bacia de Campos) com o objetivo principal de avaliar o comportamento do sistema computacional proposto, além de gerar um *dataset* representativo composto por nuvens de pontos de estruturas de aplicação operacional instaladas em ambiente subaquático.

4.1 Testes na Baía de Aratu - ambiente relevante

O objetivo desses experimentos foi avaliar o comportamento do sistema quando os dados fornecidos provêm de um ambiente relevante e, ao mesmo tempo, avaliar o desempenho do sistema na geração das nuvens de pontos e de seu erro de medição.

4.1.1 Preparação dos testes em ambiente relevante

Para avaliação do sistema proposto foi utilizada uma embarcação para serviços de apoio marítimo aos testes do protótipo DIGISUB. A Figura 4.1 apresenta o rebocador (*Tug vessel*, de 23,50 m de comprimento, com guincho de proa 10 t) da empresa GEOMAR, utilizado nos testes em águas abrigadas na Baía de Aratu (Figura 4.2), situada no estado da Bahia.

Figura 4.1: Embarcação rebocador utilizado para testes na Baía de Aratu.



Fonte: Autor

A Baía de Aratu é uma grande enseada que abriga a Base Naval de Aratu, o Porto de Aratu, um estaleiro, duas marinas (Aratu Iate Clube e Marina Aratu) que oferecem apoio e serviços aos navegantes, e também empresas que operam em escala industrial. Localizada a cerca de 20 quilômetros ao norte de Salvador, BA.

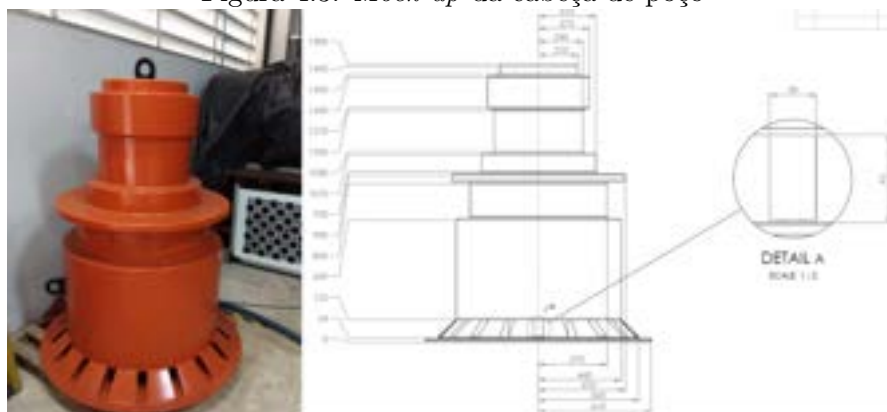
Figura 4.2: Mapa Baía de Aratu, BA



Fonte: Google

A Figura 4.3 ilustra o *mock-up* da cabeça de poço e suas dimensões, construído especificamente para representar o equipamento subaquático a ser digitalizado durante os testes, com massa aproximada de 200 kg.

Figura 4.3: *Mock-up* da cabeça de poço

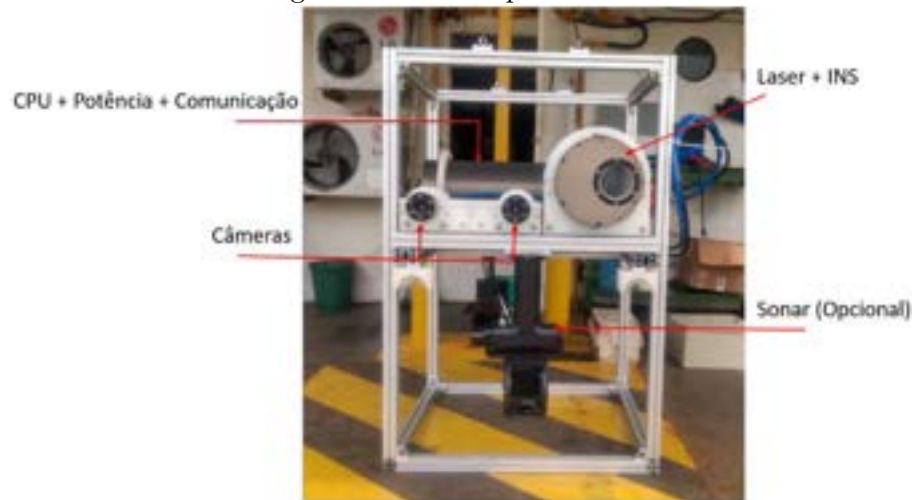


Fonte: Equipe Projeto DIGISUB

Na Figura 4.4 é possível observar os encapsulamentos em que o protótipo foi subdividido, compostos por uma unidade de processamento e potência, um sistema óptico baseado em projeção de linhas por laser, duas câmeras RGB e um sonar. A massa do protótipo, no ar, é 66 kg. Já na água, é de aproximadamente 32 kg. Neste projeto de pesquisa, utilizou-se

apenas as imagens provenientes da câmera esquerda do protótipo.

Figura 4.4: Protótipo montado



Fonte: Equipe Projeto DIGISUB

Para formação da base de dados das imagens subaquáticas do *mock-up*, o qual foi mergulhado a uma profundidade entre 15 e 17 m, buscou-se variar a distância do protótipo em relação ao *mock-up*, a partir de 1,5 m¹ com incremento de 1,0 m até atingir 10 m, isto como forma de simular diferentes condições de visibilidade do fundo mar. Porém, dadas as condições de turbidez da água na região onde os testes foram executados, foi possível visualizar o *mock-up*, objeto da cena, somente até a distância de 8 m (Figura 4.5). A Figura 4.6 mostra o cenário de testes executado na Baía de Aratu.

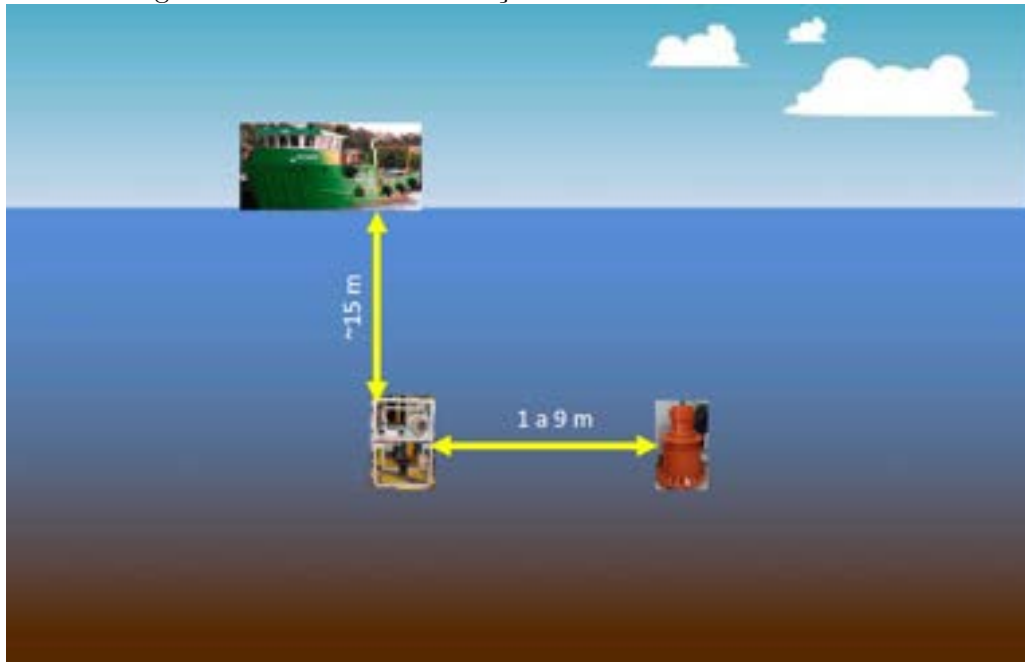
Figura 4.5: Visão do mergulhador no cenário de testes na Baía de Aratu.



Fonte: Equipe Projeto DIGISUB

¹Esta distância foi definida a partir de requisitos de segurança, buscando evitar uma possível colisão entre protótipo e *mock-up* e de manter o campo de visão para a câmera.

Figura 4.6: Cenário de execução dos testes na Baía de Aratu.



Fonte: Equipe Projeto DIGISUB

A Figura 4.7 demonstra o ambiente na embarcação durante a preparação dos testes e lançamentos dos equipamentos na água. O *mock-up* foi lançado no mar por um guindaste hidráulico da embarcação e com auxílio de um mergulhador o protótipo DIGISUB foi conduzido aos locais especificados pela equipe técnica.

Figura 4.7: Preparação dos testes do sistema.



Fonte: Equipe Projeto DIGISUB

A partir dos dados das imagens geradas durante os experimentos, nuvens de pontos foram geradas e em seguida a erro de medição do sistema foi avaliado. A Figura 4.8 mostra o *mock-up* modificado (introdução de novos elementos) para avaliar o processo de geração das nuvens de pontos.

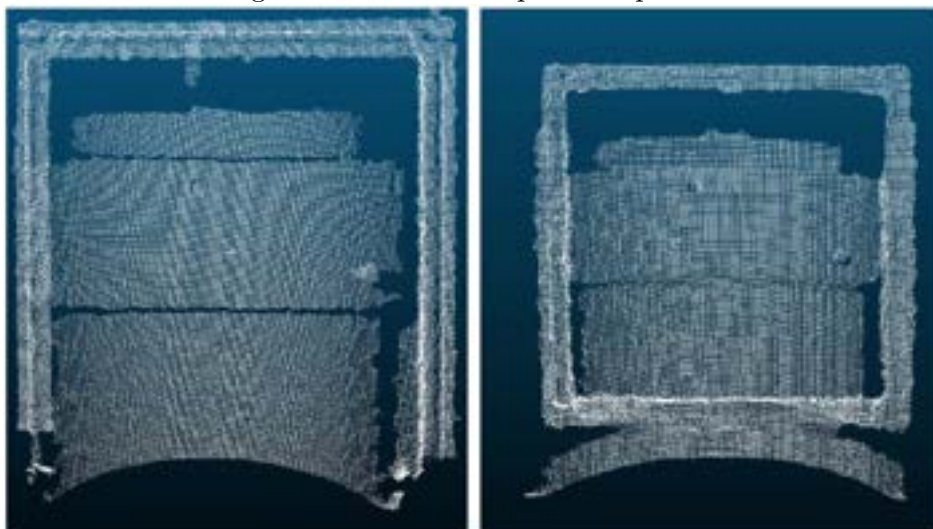
Figura 4.8: Avarias provocadas no *mock-up*

Fonte: Equipe Projeto DIGISUB

4.1.2 Resultados dos testes em ambiente relevante

A Figura 4.9 apresenta as nuvens de pontos esparsas geradas pelo sistema proposto. Observa-se que as nuvens de pontos apresentam lacunas em sua formação, isto pelo fato de alguns pontos não poderem ser encontrados e rastreados para fazer a correta correspondência, devido à falta de informações de textura do objeto submerso em ambiente subaquático. O protótipo DIGISUB utiliza LEDs para iluminação do ambiente de operação, o que ajudaria ao sistema obter informações de textura das formas geométricas na superfície do objeto alvo. Contudo, a utilização da iluminação artificial foi prejudicada pela incidência de luz solar em águas rasas (neste caso, até 17 metros de profundidade).

Figura 4.9: Nuvens de pontos esparsas.

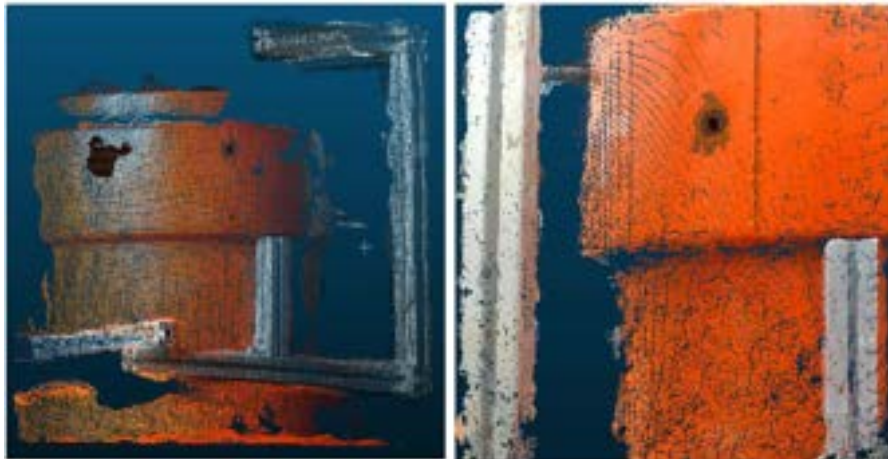


Fonte: Autor

Após a geração das nuvens de pontos esparsas, o passo seguinte é calcular uma recons-

trução densa. Os quadros principais são selecionados internamente para executar e registrar a reconstrução a partir do estéreo com várias linhas de base e, em seguida, suas nuvens de pontos resultantes são combinadas em uma única nuvem. A Figura 4.10 apresenta o resultado final da reconstrução. A malha é convertida em um arquivo PLY, o qual é gravado no disco rígido do *notebook* e que pode ser importado diretamente para um *software* de processamento de malha 3D. Uma nuvem de pontos densa é criada usando uma reconstrução esparsa computada anteriormente e uma implementação básica de estéreo *multiview*. Esta abordagem funciona por meio da identificação de visualizações centrais que têm o melhor conjunto de vizinhos para cálculos estéreos. Em seguida, uma nuvem de pontos global é criada a partir da visualização central das imagens de disparidade, tomando-se os cuidados necessários para evitar a adição de pontos duplicados.

Figura 4.10: Nuvens de pontos densas.



Fonte: Autor

O *software* CloudCompare foi utilizado para extração das medidas do objeto reconstruído. Trata-se de uma ferramenta de *design* de protótipo baseada em malhas de triângulos dinâmicos de alta resolução.

A Tabela 4.1 apresenta os resultados das medições. As médias das medições da cota do *mock-up* estão estabelecidas como referência para avaliação do desempenho do sistema, por meio erro absoluto e relativo de medição. Durante as tomadas, a distância entre sistema protótipo e objeto alvo partiram de 1,5 m até atingir 8 m, que foi o limite onde o protótipo foi capaz de perceber o alvo, isto em função das condições do ambiente e da turbidez da água no local (o planejamento inicial era de alcançar 10 m).

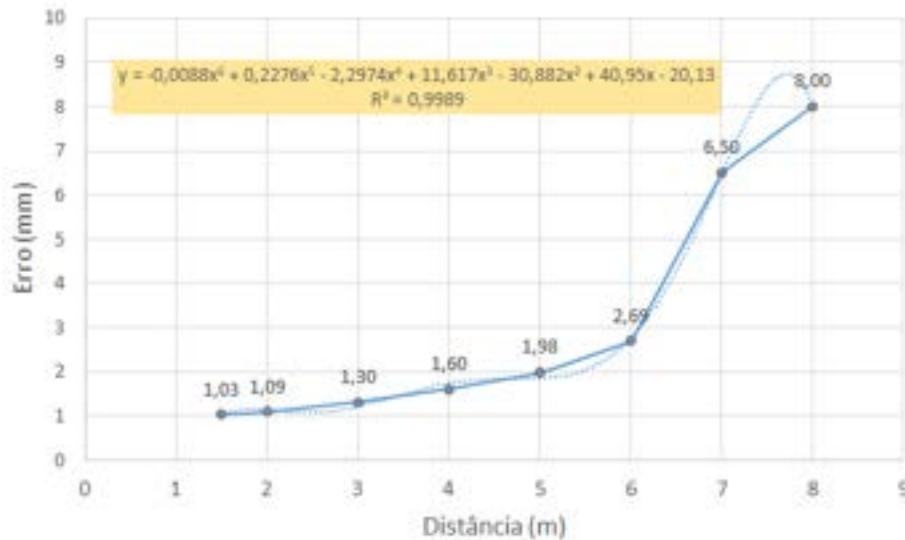
Tabela 4.1: Erro de medição do sistema alcançado nos testes da Baía de Aratu.

Distância	Cota	Média Medição	Erro	Erro Absoluto	Erro relativo
1,5	160,00	161,03	1,03	1,03	0,64%
2	160,00	158,91	-1,09	1,09	0,68%
3	160,00	158,70	-1,30	1,30	0,81%
4	160,00	161,60	1,60	1,60	1,00%
5	160,00	161,98	1,98	1,98	1,24%
6	160,00	162,69	2,69	2,69	1,68%
7	160,00	166,50	6,50	6,50	4,06%
8	160,00	168,00	8,00	8,00	5,00%

Valores dados em milímetros (mm), exceto aqueles indicados em percentual (%).

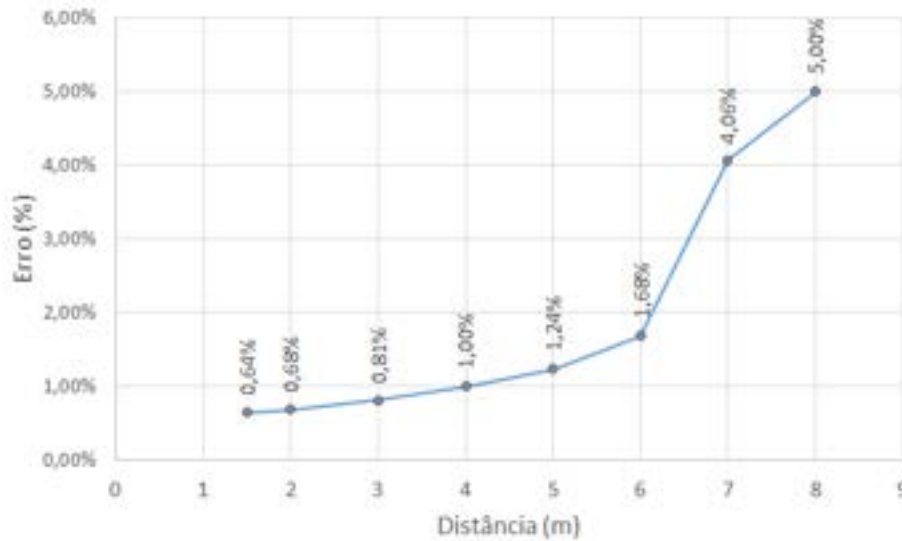
Os gráficos apresentados pelas Figuras 4.11 e 4.12 foram criados a partir dos registros da Tabela 4.1. Percebe-se um crescimento do Erro Absoluto com comportamento de um polinômio de 6ª ordem. Entre as medições realizadas de 1,5 m e 5 m mostra-se mais constante, havendo um salto para as demais, o que pode ser explicado justamente pelo aumento da distância entre protótipo e objeto alvo e, conseqüentemente, pela perda de resolução das imagens capturadas.

Figura 4.11: Erro absoluto de medição do sistema



Fonte: Autor

Figura 4.12: Erro relativo de medição do sistema



Fonte: Autor

O sistema apresenta um erro de medição de aproximadamente 1 mm para distâncias de até 3 m e de 2 mm para distância de até 5 m. Observa-se que para distâncias acima de 5 m, o erro de medição atinge até 8 mm. Portanto, o sistema proposto, mesmo em condições diferentes daquelas para as quais foi concebido (ausência de iluminação natural e sem turbidez que são as condições esperadas para águas profundas e ultraprofundas), atende ao objetivo de gerar uma reconstrução 3D com erro milimétrico. Com isto, tornando-se apto a ser testado em ambientes de maior relevância ou já operacionais.

4.2 Testes na Bacia de Campos - ambiente operacional

A segunda parte dos testes teve como objetivo avaliar o desempenho computacional do sistema e realizar medições em um equipamento real utilizado na exploração de petróleo em águas profundas.

4.2.1 Preparação dos testes em ambiente operacional

Para realização dos testes, o protótipo DIGISUB (Figura 4.13) foi instalado no ROV FCV3000, operado pela empresa Fugro (Figura 4.14), lançado por meio da embarcação RSV (*ROV Support Vessel*) Parcel do Bandolim, de propriedade da empresa OceanPact. Trata-se de uma embarcação de grande porte (dimensões 91 m x 19 m) para apoio às atividades diversas realizadas para indústria de O&G *offshore*. (Figura 4.15).

Os testes foram realizados na Bacia de Campos, situada entre a costa norte do estado do Rio de Janeiro e o sul do estado do Espírito Santo, e para isso, uma equipe do SENAI CIMATEC esteve embarcada no referido RSV por um período de quinze dias.

Figura 4.13: Encapsulamentos do protótipo instalados no ROV.



(a) Encapsulamentos óptico e câmeras

(b) Encapsulamento de processamento e potência

Fonte: Equipe Projeto DIGISUB

Figura 4.14: ROV FCV3000.



Fonte: Fugro

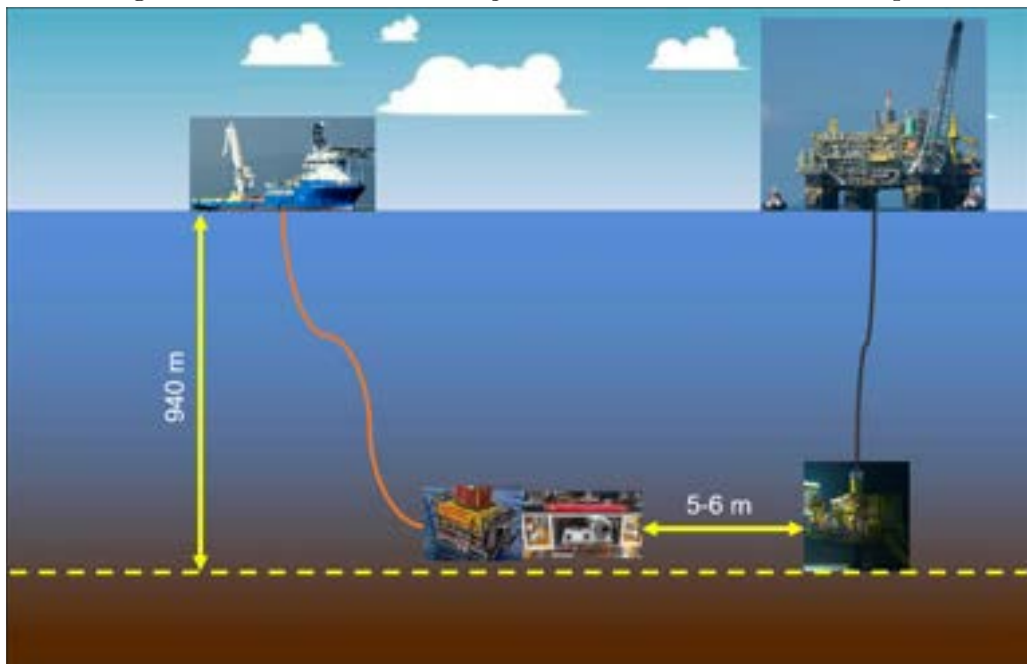
Figura 4.15: RSV Parcel do Bandolim.



Fonte: OceanPact

O cenário dos testes é ilustrado na Figura 4.16, onde é possível identificar a conexão entre o RSV e o ROV que carrega o protótipo a 940 m de profundidade, inspecionando e digitalizando a ANM, conectada à plataforma de petróleo.

Figura 4.16: Cenário de execução dos testes na Bacia de Campos.

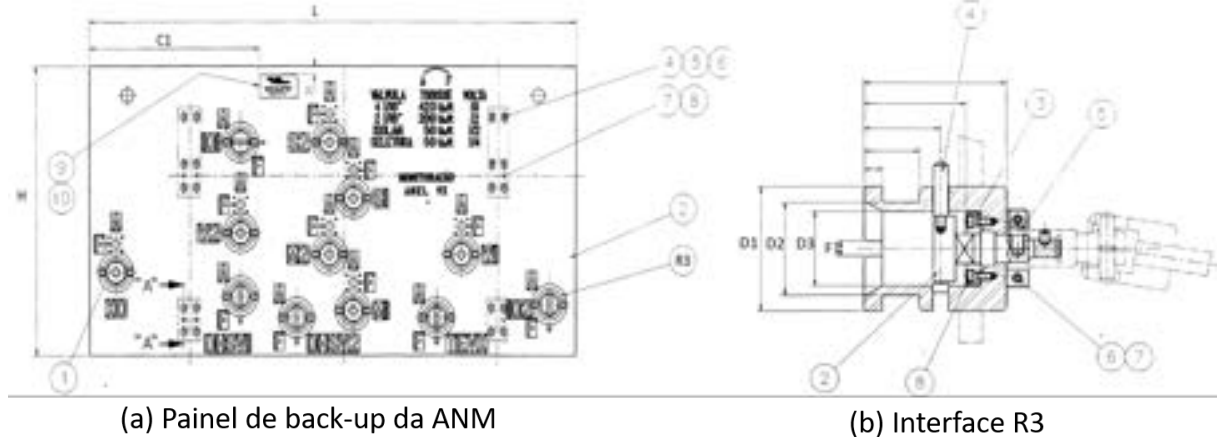


Fonte: Autor

Para avaliação do desempenho do sistema, o painel de *back-up* de uma ANM (Árvore de Natal Molhada) de dimensões conhecidas foi digitalizado em seu local de instalação, a

940 m de profundidade, por meio do sistema DIGISUB proposto. Os valores registrados foram obtidos de forma indireta após o processamento das informações de maneira *off-line* e comparados com as dimensões extraídas do desenho técnico do equipamento fornecido pela Petrobras, proprietária do ativo. Na Figura 4.17 e na Tabela 4.2 são apresentadas as dimensões das cotas para comparação com as medidas obtidas pelo sistema computacional DIGISUB.

Figura 4.17: Desenho técnico com as principais cotas da ANM digitalizada.



Fonte: Petrobras

Tabela 4.2: Dimensões do painel de *back-up* da ANM.

Cota	Dimensão (mm)
H	1.230
L	2.370
C1	825
D1	170
D2	127
D3	122
F	20

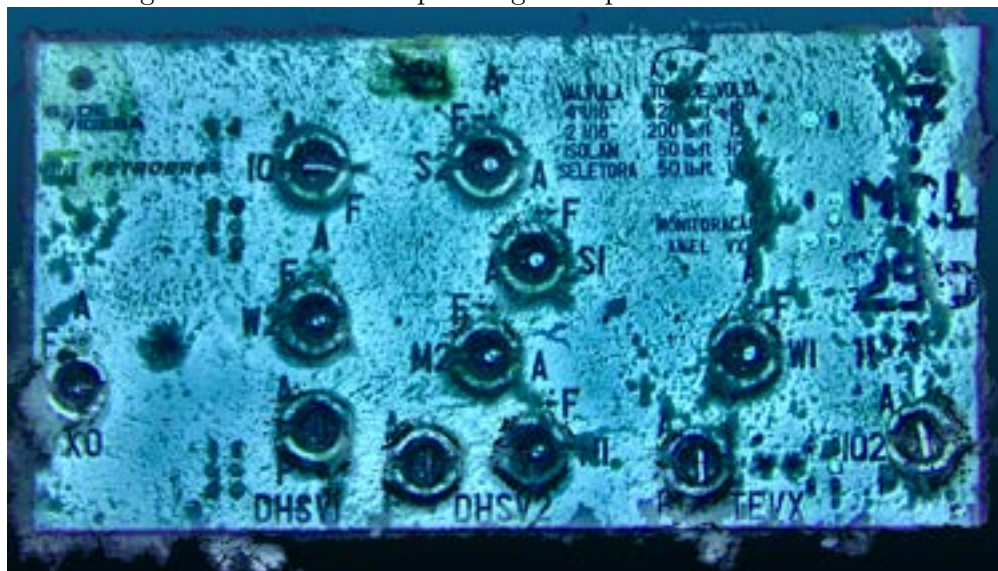
Os resultados apresentados a seguir são abordados sob dois aspectos: avaliação da capacidade do DIGISUB em reconstruir a cena, especificamente a ANM objeto de inspeção propriamente dita; e o desempenho computacional demandado pelo processamento das informações.

4.2.2 Avaliação da reconstrução

Nesse experimento, são apresentados os resultados dos testes de avaliação dos erros de medição das medidas da nuvem de pontos densa (Figura 4.18) gerada pelo sistema. As

medições foram realizadas com a utilização do *software* CloudCompare v2.11.3 e os resultados foram catalogados na Tabela 4.3.

Figura 4.18: Nuvem de pontos gerada pelo método DIGISUB.



Fonte: Autor.

Tabela 4.3: Comparação dimensões do painel de *back-up* da ANM.

Cota	H	C1	L	D1	D2
Valor da Cota	1.230,00	825,00	2.370,00	170,00	127,00
Medição 1	1.225,75	835,21	2.365,40	171,10	131,10
Medição 2	1.228,77	827,17	2.369,46	169,51	130,50
Medição 3	1.226,17	826,95	2.370,89	170,86	129,66
Medição 4	1.230,77	827,21	2.356,34	171,04	128,81
Medição 5	1.228,19	829,05	2.375,10	169,84	129,72
Média Medições	1.227,93	829,12	2.367,44	170,47	129,96
Erro de Medição	-2,07	4,12	-2,56	0,47	2,96
Erro relativo	-0,17%	0,50%	-0,11%	0,28%	2,33%
Incerteza A	+/- 1,83	+/- 3,14	+/- 6,36	+/- 0,66	+/- 0,78
Incerteza B	+/- 0,15	+/- 0,15	+/- 0,15	+/- 0,15	+/- 0,15
Incerteza Total	+/- 1,97	+/- 3,28	+/- 6,50	+/- 0,81	+/- 0,93

Valores dados em milímetros (mm), exceto aqueles indicados em percentual (%).

Considerando a impossibilidade de se efetuar por métodos convencionais as medições do painel objeto de inspeção, representado pela imagem da Figura 4.20, uma vez que se encontra instalado em águas profundas, foi tomada como referência as cotas do desenho técnico que o representa (Tabela 4.2).

O erro encontrado é um comparativo da Média das Medições pelo processo de repetibilidade (OLIVEIRA, 2012) em relação ao Valor da Cota, assumindo-se incerteza desprezível para estas cotas. Observa-se a partir dos dados obtidos pelo sistema do protótipo que o Erro de Medição do sistema varia de 0,47 mm até 4,12 mm e uma Incerteza Total de +/- 0,81 a +/- 6,50 mm, considerando que a distância do sistema até o objeto alvo foi mantida em aproximadamente 5 a 6 m.

A Repetibilidade é a execução de medições efetuadas num determinado conjunto de condições, as quais incluem o mesmo procedimento de medição, os mesmos operadores, o mesmo sistema de medição, as mesmas condições de operação e o mesmo local, assim como medições repetidas no mesmo objeto ou em objetos similares durante um curto período de tempo. Os valores obtidos diferem entre si por causa de efeitos aleatórios que ocorrem durante as medições.

A Incerteza A é caracterizada pela variabilidade dos valores medidos (Medições 1 a 5), isto é, a dispersão dos mesmos em torno do valor médio. Já a Incerteza B está associada ao processo de calibração e ao processo de determinação dos pontos da nuvem no espaço tridimensional (FELICIANO; SOUZA; LETA, 2005). O valor para esta incerteza foi obtido a partir do resultado apresentado ao final da etapa de processamento onde é gerada e determinada a nuvem esparsa.

Não foi possível realizar as medições das cotas D3 e F, pois os pontos não tiveram definição adequada na reconstrução do painel de *back-up* da ANM.

A partir dos resultados obtidos pelos testes na Baía de Aratu e testes na Bacia de Campos, são identificadas duas cotas utilizadas para análise de desempenho da precisão do sistema, de dimensões e distâncias de tomadas das imagens semelhantes, conforme apresentado na Tabela 4.4.

Tabela 4.4: Comparação resultados das cotas Baía de Aratu X Bacia de Campos.

Local do testes	Cota	Distância	Medição	Erro
Baía de Aratu	160 mm	6 m	162,69 mm	2,69 mm
Bacia de Campos	170 mm	6 m	170,47 mm	0,47 mm

O erro quase seis vezes é explicado pelas condições do cenário em que os testes foram realizados na Baía de Aratu, as quais foram desfavoráveis em função da turbidez do local e influência da iluminação natural, que podem ser visualizadas na imagem da Figura 4.5. O sistema computacional DIGISUB foi concebido para atuar em águas profundas, que têm como características principais a ausência de partículas suspensas (baixa turbidez) e sem iluminação natural.

Portanto, pode-se concluir que:

- O sistema atende ao objetivo de gerar uma reconstrução 3D com erro milimétrico de um equipamento instalado em ambiente operacional de grandes dimensões (aproximadamente, 6 m de altura total), com toda severidade presente neste cenário.
- O sistema proposto tem potencial para ser aplicado no auxílio às atividades de inspeção de integridade de ativos (mais sobre isto será debatido na Seção 4.4), bem como ser incorporado a UUVs no futuro para auxiliá-los na execução de atividades de intervenção.
- Os resultados apresentados são melhores que aqueles obtidos por trabalhos de pesquisa anteriores, relacionados na Subseção 2.3.4.

Figura 4.19: Reconstrução completa da parte frontal ANM pelo método DIGISUB.



Fonte: Autor.

4.2.3 Desempenho computacional

Nesse experimento, as nuvens de pontos do painel de *back-up* da ANM foram processadas e reconstruídas variando os métodos KLT, DDA e DIGISUB (Híbrido). Foram utilizadas 50 imagens para geração das nuvens de pontos. A Figura 4.20 mostra o exemplo de uma das imagens capturada pela câmera esquerda do sistema proposto durante os testes em condições submarinas.

Figura 4.20: Imagem da câmera esquerda do protótipo.



Fonte: Autor.

Os resultados do desempenho computacional² para os diferentes métodos são observados na Tabela 4.5. O método KLT foi o mais rápido para geração da nuvem de pontos e com um consumo de memória RAM muito inferior (duas a três vezes) aos outros métodos, contudo, nota-se que o KLT obteve uma menor quantidade de pontos reconstruídos (três a quatro vezes), em função da menor aceitação dos pontos da triangulação no momento do rastreamento das *features* e, portanto, não foi capaz de realizar a reconstrução da cena.

Tabela 4.5: Resultado do desempenho computacional.

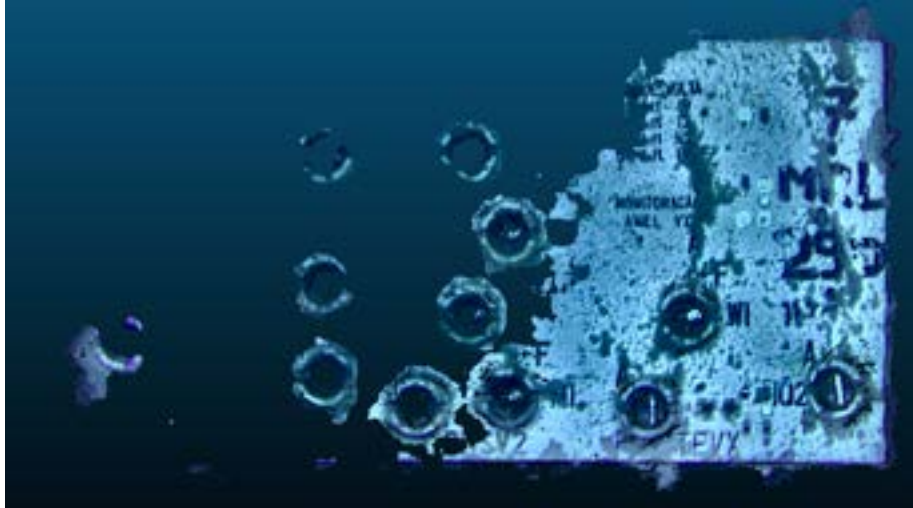
Método	Tempo	Número de pontos	Memória RAM
KLT	4 min 16 s	181.908.000	2,8 GB
DDA	1 h 13 min 14 s	708.520.000	9,2 GB
DIGISUB	34 min 39 s	594.330.000	6,8 GB

A Figura 4.21 apresenta a nuvem de pontos gerada a partir método KLT. Nela é possível constatar a ausência de pontos (lacunas) necessários para formação completa da nuvem

²O computador utilizado para processamento e obtenção dos resultados para análise foi o modelo Avell A65 144 Hz, processador Intel® Core™ i7-8750H Coffee Lake, 9 MB Cache (2.2 GHz até 4.1 GHz com Intel® Turbo Boost) e memória 16 GB [2x 8GB - Dual Channel].

de pontos, ratificando a afirmação mencionada acima de que este método, KLT, para esta cena especificamente foi incapaz, por si só, de gerar sua reconstrução completa.

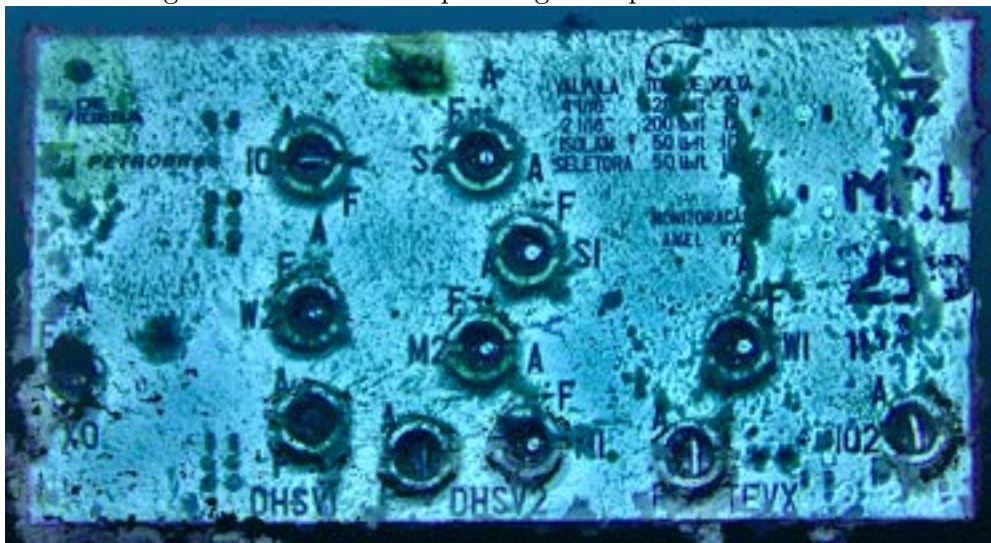
Figura 4.21: Nuvem de pontos gerada pelo método KLT.



Fonte: Autor.

Já o método DDA apresentou o tempo de processamento de 1h 13min 14s, sendo o mais lento para geração da nuvem de pontos e com um consumo de memória RAM (9,2GB) superior aos outros métodos, isto em função deste método utilizar a busca gulosa para encontrar correspondência entre *features* presentes nas imagens. Apesar do maior consumo de memória RAM e maior tempo de processamento para geração da nuvem de pontos, ressalta-se que o método DDA obteve uma maior quantidade de pontos reconstruídos (total de 708.520.000) em comparação com os outros métodos avaliados (Figura 4.22) e que foi capaz de reconstruir a cena completa.

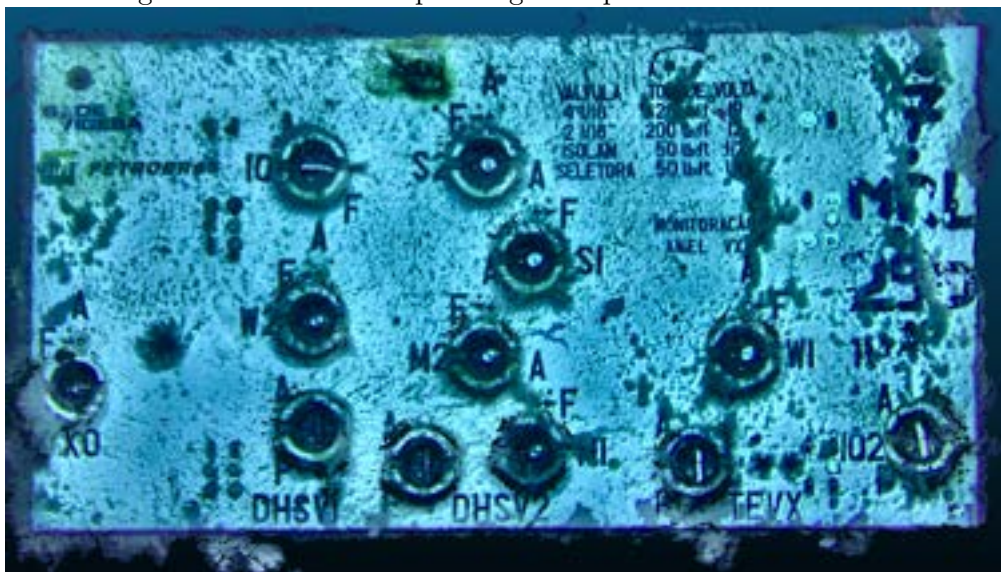
Figura 4.22: Nuvem de pontos gerada pelo método DDA.



Fonte: Autor.

O método DIGISUB (híbrido) apresentou um tempo de processamento de 34min 39s (53% menor que o tempo do DDA), 594.330.000 pontos reconstruídos (apenas 18% a menos que o DDA) e consumo de 6,8 GB de memória RAM no processo de geração da nuvem de pontos. Em comparação ao método DDA, o método DIGISUB apresenta ganhos em relação ao tempo de processamento (cerca de duas vezes mais rápido) e consumo de memória RAM (2,4 GB menor), mantendo uma alta densidade de pontos reconstruídos, que além de dar condições para reconstrução de toda a cena, também permite a obtenção de medições do equipamento para comparação com as cotas do desenho técnico. A Figura 4.23 apresenta a nuvem de pontos gerada a partir método do DIGISUB.

Figura 4.23: Nuvem de pontos gerada pelo método DIGISUB.

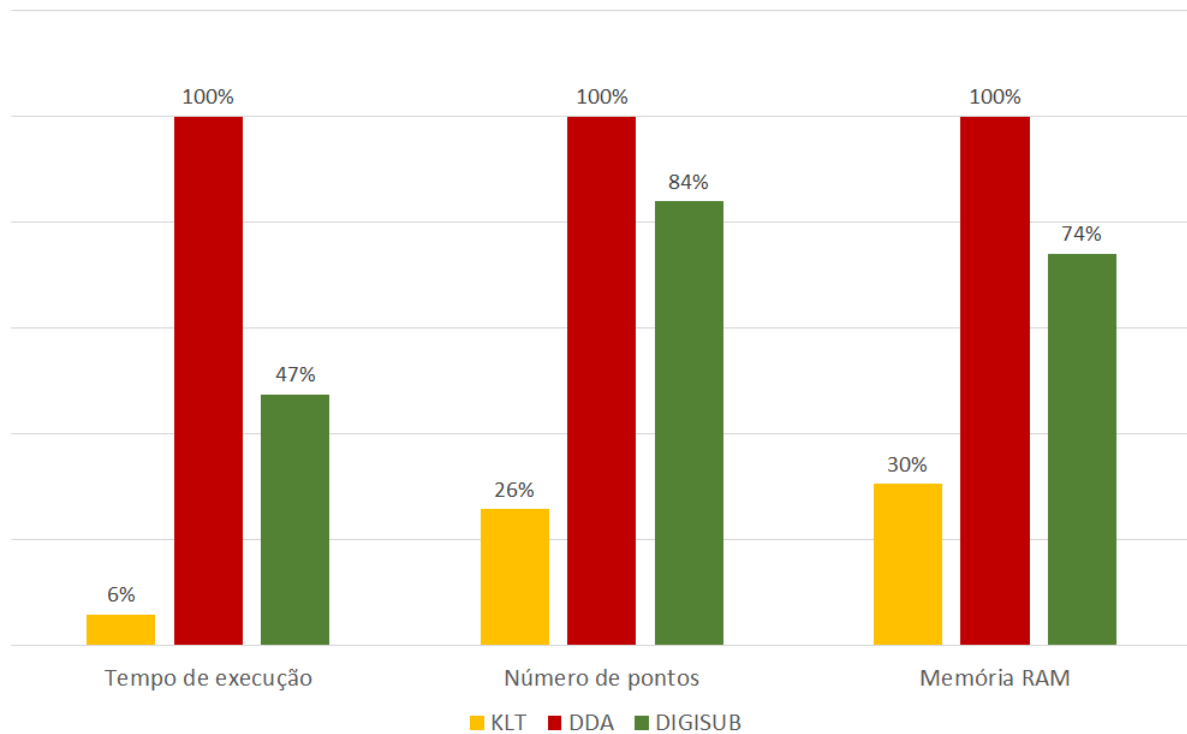


Fonte: Autor.

Após a conclusão dos experimentos e de acordo com os resultados observa-se que o método KLT apresenta melhor desempenho computacional, porém tem problemas de rastreamento, causado pela falta de suavidade entre os *frames* das imagens decorrentes da variação de iluminação ou rotação introduzidas pela instabilidade do ROV. O DDA é o método que apresenta o melhor resultado em encontrar correspondência entre *features*, obtendo maior quantidade de pontos reconstruídos, contudo, é o método que mais consome recursos computacionais. Já o método DIGISUB alterna seu comportamento entre os dois métodos KTL e DDA, ajustando-se de acordo com a necessidade. Ou seja, quando ocorre uma falha do método KLT, o DIGISUB altera sua atuação de forma a assumir as características do rastreador DDA, recuperando o desempenho em relação à reconstrução 3D do objeto inspecionado.

Os resultados apresentados na Tabela 4.5 também podem ser observados na Figura 4.24, logo abaixo.

Figura 4.24: Resultados dos métodos KLT, DDA e DIGISUB.



Fonte: Autor.

Utilizando as imagens geradas do painel frontal da ANM durante os testes realizados, foi possível realizar a reconstrução da ANM pelo método DIGISUB e o resultado da digitalização pode ser observado na Figura 4.19. O tempo total de processamento foi de 12h 04min 06s, gerando mais de quatro bilhões de pontos reconstruídos (4.063.040.000 pontos) e consumo de 15,6 GB de memória RAM.

Em relação à análise dos resultados considerando a vista parcial, representada pela Figura 4.23 gerada pelo método DIGISUB, e vista total, representada pela 4.19 e também gerada pelo método DIGISUB, pode-se estabelecer a comparação apresentada na Tabela 4.6, a partir da qual conclui-se que:

- Não é possível estabelecer uma relação linear entre o tempo de processamento e o número de pontos gerados.
- O tempo de processamento está associado ao número de comutações entre os algoritmos KLT e DDA ocorridas durante a execução do método DIGISUB.
- O tempo de processamento também está relacionado às características encontradas, qualitativamente e quantitativamente, em cada parte da cena e não das dimensões de cada uma das vistas da cena.

Tabela 4.6: Comparação resultados entre vistas parcial e total da ANM.

Vista	Tempo (s)	Pontos gerados	Pontos gerados / seg.
Parcial (Figura 4.23)	2.079	594.330.000	285.873
Total (Figura 4.19)	43.446	4.063.040.000	93.519

4.3 Análise do sistema híbrido DIGISUB comparado à referência estabelecida na pesquisa sistemática

Em Abeles (2013) forma combinados métodos DDA e DT por meio da proposta de um método híbrido. A evolução deste método é o núcleo desta tese, a qual propõe uma forma de rastreamento e detecção de *features* numa dada sequência de imagens.

O modelo híbrido DDA-DT proposto por Abeles (2013) apresenta as seguintes características:

- Testado em ambiente terrestre, buscando a reconstrução de cenas do cotidiano (parede de tijolos e tapeçaria com variação de detalhes).
- Avaliado sob influência de diferentes condições de iluminação natural.
- Sem preocupar-se com a precisão da reconstrução 3D.
- Nele o foco foi direcionado para o desempenho computacional e para a estabilidade dos rastreadores.

Tomando-se os resultados obtidos por Abeles (2013), a combinação de algoritmo FH-SURF-KLT apresentou melhor desempenho dentre às que foram testadas. Portanto, o DIGISUB partiu deste algoritmo com base da etapa de Detecção de *Features* proposto em seu *pipeline*.

O DIGISUB tem como características:

- Busca a precisão milimétrica da reconstrução 3D com a utilização dos rastreadores de *features*.
- Avalia a forma como estes algoritmos se comportam num ambiente não controlado e operacional, embarcado num ROV atuando em águas profundas.
- Realização de testes sob influência dos efeitos da salinidade nos índices de refração da água, também sob influência de luminosidade artificial, sem influência de iluminação natural e sob alta pressão pela coluna d'água (águas profundas).

- Propõe para correção destes efeitos, a combinação de algoritmos e sua inserção *pipeline*, como demonstrado na Figura 3.2.

Portanto, o DIGISUB traz sua contribuição à ciência por meio da proposição de um novo sistema capaz de realizar a reconstrução 3D de objetos complexos (ANM) e instalados em ambiente severo (águas profundas), de forma a obter uma precisão milimétrica e obteve melhores resultados que os trabalhos destacados na Tabela 2.2, síntese da pesquisa sistemática.

4.4 Impacto do resultado da pesquisa em aplicações industriais

Desafiado a permanecer competitivo em um ambiente de baixo preço do petróleo enquanto navega pela transição energética, o setor de O&G está agora buscando maneiras de evoluir as operações por meio da digitalização. A robótica já possibilitou uma evolução da indústria ao possibilitar operações em águas profundas por meio da implantação de ROVs. Agora, procura reduzir os custos operacionais unitários das instalações *offshore*, melhorando a segurança e o desempenho ambiental (ZAGATTI et al., 2018).

Os custos com operações *offshore* receberam limitada atenção dos meios acadêmico e corporativo. Em boa parte, os custos com a produção com petróleo e gás ocupam uma pequena porção orçamentária, além dos altos preços das commodities gerarem grandes fluxos de caixa, aumentando a produção e vendas por consequência (KAISER, 2019).

Diferentemente dos custos de capital para prospecção de petróleo e seus derivados, os custos operacionais são menos transparentes, com menor disponibilidade de dados e informações pertinentes para avaliação, vindo de diferentes formas e aplicações. Apesar de tais obstáculos, há estudos de *benchmark* que propuseram diminuir a lacuna entre as informações obtidas e a avaliação resultante do processo.

Atualmente, a inspeção de equipamentos submarinos instalados no leito marinho é realizada por meio de ROVs de grande porte, que necessitam de barcos de suporte e equipes especializadas para sua operação. Estes barcos de suporte têm um custo aproximado de 100 mil dólares por dia e, portanto, têm uma significativa relevância nos custos de inspeção e operação das unidades submarinas de produção de óleo e gás (ZAGATTI et al., 2018).

O uso de soluções robóticas em meio submarino, aliadas às capacidades autônomas e inteligência artificial podem possibilitar operações mais seguras com impacto mínimo no ambiente marinho (KAMPMANN et al., 2018). Sistemas autônomos, em primeira instância,

podem melhorar a produtividade para reduzir diretamente o custo da inspeção submarina. Em segundo lugar, ao reduzir a quantidade de pessoal a bordo necessário para as operações de inspeção, haverá redução dos riscos relacionados à exposição humana (ZAGATTI et al., 2018).

Considerando os aspectos sobre os métodos aplicados pela indústria para gestão da integridade e manutenção de ativos, tornam-se necessárias inspeções preliminares para obtenção de informações que, posteriormente, subsidiarão todo planejamento de ações e recursos para as intervenções. Porém, dadas as limitações dos sistemas de inspeção visual muitas lacunas ocorrem neste planejamento que acabam exigindo missões de inspeção adicionais ou falha no provisionamento de sobressalentes.

Ter disponível um sistema de inspeção que permita uma visualização detalhada e, conseqüentemente, um planejamento refinado das missões de intervenção pode impactar positivamente na redução do tempo de operação de ROVs e de sua infraestrutura de suporte e claro, reduzindo os custos operacionais na proporção do valor acima mencionado.

Ao fazê-lo, a pesquisa demonstra o valor real da digitalização, em que a tecnologia desenvolvida para viabilizar a reconstrução 3D refinada dos objetos pode permitir outras aplicações. Como possibilidade de impacto do resultado da pesquisa é a automação de atividades que hoje são realizadas e controladas pelos operadores de ROVs, ou até mesmo em AUVs. No entanto, os AUVs com capacidade de intervenção ainda estão sujeitos a mais pesquisas, pois a incorporação de tecnologias necessárias para garantir a execução correta dessas tarefas ainda não está completamente desenvolvida (TECCHIO et al., 2021). Para tal aplicação, a condição central está em aproximar mais precisamente os tempos de execução das atividades presentes nos registros dos melhores condutores dos manipuladores. Por mais que estes tenham a experiência para uma boa condução, ainda há fatores (sejam eles de caráter endógeno ou exógeno) que possibilitam a presença de anomalias, enquanto com a automação estas seriam minimizadas. Da mesma forma, qualquer impacto positivo da redução do tempo de operação de um ROV é relevante em termos de custos para as empresas operadoras.

Algumas operações de intervenção rotineiras realizadas por ROVs em equipamentos instalados no fundo do mar, especificamente Árvores de Natal Molhadas, e que, uma vez automatizadas por meio de um sistema de digitalização como o DIGISUB teriam impacto positivo na redução de tempo de execução, seriam: abertura e fechamento de válvulas, colocação e retirada de *hot stab*, aplicação de ferramenta de torque, instalação de *tree cap*, conexão de *flying leads*, limpeza de *hub*, instalação de flanges.

Porém, ressalta-se que não foi objeto desta pesquisa elaborar um estudo detalhado dos possíveis impactos diretos do emprego de seu resultado, método DIGISUB, nas atividades

industriais. Caberia, portanto, uma recomendação de trabalho futuro tal análise.

Obviamente pela natureza da empresa demandante da pesquisa, Petrobras, ser do segmento de O&G, sistemas de digitalização e, especificamente o DIGISUB, são de ampla utilização em outros setores da indústria e contribuirão para viabilizar operações de mineração submarina, inspeção em portos, parques de geração de energia *offshore*, operação e manutenção de barragens, além de pesquisa arqueológicas e oceanográficas, dentre outras aplicações.

Considerações Finais

Este trabalho de investigação apresenta um sistema computacional integrado a modelos digitais para reconstrução com alta definição de estruturas subaquáticas em forma de nuvens de pontos, as quais por sua vez podem ser utilizadas para gerar modelos tridimensionais destas estruturas. A metodologia proposta é estabelecida por um processo (*pipeline*) composto pelas etapas aquisição de imagem, processamento e reconstrução 3D. A aquisição de imagem é realizada através do pré-processamento e da retificação das imagens. Já na etapa de Processamento, é realizada a detecção de *features*, correspondência e refinamento 3D. Na reconstrução 3D, ocorre a geração da nuvem de pontos densa e a geração de superfícies. Experimentos de digitalização conduzidos em condições subaquáticas foram realizados em ambientes relevante e operacional e a reconstrução 3D foi feita para permitir a análise dos resultados.

5.1 Conclusões

A pesquisa, inicialmente, abordou o desenvolvimento de um sistema de digitalização subaquática desde a concepção da ideia, seleção do *hardware*, definição do método de calibração e chegando à reconstrução 3D propriamente dita. Porém, ao longo da pesquisa, foram identificadas linhas de desenvolvimento bastante específicas, principalmente no que se refere à calibração do sistema e proposição de algoritmos para reconstrução 3D. Foi necessária, portanto, a delimitação da pesquisa sobre um desses temas, sendo então optada pela última mencionada, **algoritmo para reconstrução 3D**.

O DIGISUB traz sua contribuição à ciência por meio da proposição de um novo sistema capaz de realizar a reconstrução 3D de objetos complexos (ANM) e instalados em ambiente severo (águas profundas), de forma a obter uma precisão milimétrica. Isto foi possível por meio utilização de modelo híbrido de rastreadores de *features* e da inserção e combinação de algoritmos para correção dos efeitos da salinidade nos índices de refração da água e para correção de luminosidade artificial, além da possibilidade de execução de testes em um ambiente não controlado e operacional, acoplado a um ROV em operação.

Por meio dos experimentos realizados como detalhado nos Capítulos 3 e 4, o sistema computacional proposto nesta pesquisa mostrou ser possível realizar a reconstrução 3D de estruturas subaquáticas com alta definição. Conforme os objetivos estabelecidos inicialmente nesta pesquisa, um sistema foi projetado adotando premissas técnicas de recons-

trução 3D baseadas nos algoritmos de *Structure from Motion* e *Multiview Stereo*, uma arquitetura de *software* foi proposta para atender às necessidades deste sistema. As imagens brutas utilizadas durante as fases de testes foram catalogadas e disponibilizadas em um repositório público, no endereço eletrônico: <https://github.com/leonardonardy/digisub>.

Um fator limitante identificado ao longo da pesquisa foi a baixa disponibilidade de *datasets* de estruturas e de equipamentos instalados em águas profundas. Considerações e premissas foram adotadas ao longo do desenvolvimento e dos testes preliminares em ambiente relevante, mas que puderam ser validadas de fato somente com a realização dos testes em ambiente operacional, os quais foram cruciais para confirmar o desempenho do sistema, além de proporcionar a geração de um *dataset* específico para instalações desta natureza.

A capacidade do sistema proposto em fornecer representação tridimensional com exatidão milimétrica de estruturas dos equipamentos subaquático, minimizando os erros introduzidos nas nuvens de pontos decorrentes das perturbações de movimentação que o ROV sofre no ambiente de operação, foi avaliada nos experimentos conforme detalhado nas Seções 4.1 e 4.2.

Em ambiente relevante, o sistema apresentou um erro de medição de aproximadamente 1 mm para distâncias de até 3 m e de 2 mm para distâncias entre 3 m e 5 m. Nesta etapa da pesquisa não foram levadas em consideração as incertezas de medição inerentes ao sistema de digitalização por meio de câmeras.

Uma vez no ambiente operacional, o DIGISUB apresentou um erro de medição absoluto de 0,47 mm a 4,12 mm, para uma incerteza de até +/- 6,50 mm. A distância do sistema até o alvo foi de aproximadamente entre 5 e 6 m. A determinação da incerteza levou em conta a repetibilidade das medições e as características inerentes ao sistema de digitalização.

Relativo à capacidade de reconstrução da cena, especificamente o painel frontal da ANM, o método KLT atuando isoladamente não foi capaz de reconstruí-la. O método DDA reconstruiu a cena completa, gerando uma nuvem de pontos densa (3,9 vezes a densidade gerada pelo KLT), porém ao custo de um maior esforço computacional. Já o DIGISUB mostrou-se capaz de gerar uma nuvem também de alta densidade (84% da densidade do DDA), porém a um custo computacional menor (47% do tempo exigido pelo DDA).

Em relação ao desempenho computacional do sistema, o método DIGISUB apresentou resultado de 34 m 39 s para geração da nuvem de pontos, ou seja, praticamente duas vezes mais rápido que o DDA, com um consumo de memória RAM 6,8 GB. O método DDA apresentou resultado de 1 h 13 m 14 s, sendo o mais lento para geração da nuvem de pontos e com um consumo de memória RAM de 9,2 GB. Já o método KLT, foi o

que menos consumiu recursos computacionais, tempo de 4 m 16 s e consumo de memória RAM de 2,8 GB, contudo, o KLT obteve uma menor quantidade de pontos reconstruídos (181.908.000), o que impossibilitou a reconstrução completa do objeto alvo e, consequentemente, realização das medições posteriores de suas dimensões e comparação com as cotas do desenho técnico.

O método DIGISUB gerou 594.330.000 pontos. Em comparação com o DDA, o método proposto nesse trabalho de pesquisa apresenta ganhos em relação ao tempo de processamento (47% do tempo levado pelo DDA) e ao consumo de memória RAM 26% menor e, por fim, uma desvantagem a respeito da quantidade de pontos reconstruídos, 114.190.000 a menos, o que é equivalente a 16%, mas que não impactaram na qualidade do resultado. Portanto, a partir do processamento dos pontos originados pelo DIGISUB foi possível realizar a reconstrução do objeto alvo e avaliar os erros de medições.

Durante os testes realizados na Baía de Aratu, a distância entre o protótipo digitalizador e o objeto alvo foi de 1,5 m (não foi aproximado a uma distância menor em função de questões de segurança pela possibilidade de colisão entre ambos) até 8 m, que foi a distância em que ainda foi possível uma identificação mínima visual do objeto devido à interferência da turbidez encontrada nas águas daquele local. Neste mesmos testes, também houve interferência da iluminação natural decorrente da baixa profundidade da referida baía (foi possível mergulhar o objeto alvo até 17 m).

Já durante os testes em águas profundas, o protótipo digitalizador permaneceu a uma distância entre 5 e 6 m da ANM (objeto alvo). Esta variação da distância entre o sistema e o alvo ocorreu em função das limitações de estabilidade e controle do ROV no ambiente operacional. Outro fator também levado em consideração para determinação desta distância foi a necessidade de manter o campo de visão da câmera sobre o alvo, de forma a possibilitar um enquadramento que permitisse a identificação das características do objeto. A incidência de iluminação artificial, apesar de afetar o resultado da imagem capturada, veio a ser corrigida posteriormente na etapa de aquisição de imagem inserida no modelo DIGISUB.

Outra importante constatação do trabalho foi de ser possível realizar todo processo de geração de nuvens de pontos e reconstrução 3D do objeto alvo com imagens capturadas por apenas uma câmera, sem a necessidade de utilização de outros sensores e componentes como: uma segunda câmera, sistema de *laser*, INS, ou mesmo a possibilidade de não utilização da iluminação artificial do protótipo já que a iluminação do veículo ROV seria suficiente para esta necessidade, bem como a possibilidade também de redução das dimensões do encapsulamento de potência, controle e bateria.

As técnicas desenvolvidas nesta pesquisa são promissoras, os algoritmos implementados

foram capazes de realizar a digitalização e reconstrução de uma ANM em ambiente operacional com medidas milimétricas. O estudo elaborado neste trabalho de investigação, levando-se em consideração todas as etapas de processamento para desenvolvimento do sistema computacional, não é encontrado no âmbito da revisão de literatura realizada e tem valor para a ciência e para a tecnologia aplicada diretamente à indústria.

Por fim, os resultados obtidos pelo DIGISUB, em relação à capacidade de reconstrução 3D de objetos complexos bem como à precisão atingida, são melhores que os trabalhos destacados na Tabela 2.2, síntese da pesquisa sistemática.

5.2 Contribuições

As principais contribuições proporcionadas por este trabalho de investigação à ciência e à tecnologia estão descritas a seguir. Ressaltando sua possibilidade de rápida inserção no mercado e aplicação aos processos industriais, seja pelo segmento de O&G ou por outros que demandem soluções de digitalização com precisão milimétrica de ativos.

Desenvolvimento de um sistema computacional, DIGISUB, para criação de nuvens de pontos 3D que, por sua vez, podem ser utilizadas para gerar modelos tridimensionais de equipamentos e estruturas subaquáticas digitalizadas com precisão milimétrica. Isto vai ao encontro da tendência do segmento de O&G de digitalização de suas operações e remanejamento de seus processos da superfície das plataformas de petróleo, *topside*, para o meio submarino, *subsea*, dando condições reais da criação das chamadas *subsea factories*.

Outra contribuição relevante é a utilização de apenas uma única câmera como fonte de dados para efetuar a reconstrução 3D, pois isto está relacionado diretamente à diminuição significativa da quantidade de equipamentos necessários para este tipo de operação envolvendo ROV, impactando diretamente na maior disponibilidade da capacidade de carga do veículo (*payload*), trazendo ganhos na fluidez e navegabilidade do ROV, bem como redução do consumo de energia demandada para sua operação, o que, considerando a possibilidade de embarcar a tecnologia num AUV, torna-se ainda mais relevante. Também cabe destacar a redução do custo de aquisição e montagem do sistema com apenas uma câmera frente a todos componentes relacionados originalmente para o protótipo, o que pode ser verificado no Apêndice A.

O método proposto, uma vez aplicado, permitirá acesso à informação dimensional precisa do equipamento inspecionado dando condições de realização do planejamento de intervenção e manutenção de sua integridade de maneira refinada e assertiva sobre a instalação, impactando em menor de custo destas operações assumindo-se que serão ne-

cessárias realizar menos missões preliminares de inspeção para execução do planejamento das intervenções. O DIGISUB demonstra-se, portanto, apto ao auxílio e subsídio de informações para gestão da integridade de ativos da indústria de O&G.

Da mesma forma, o DIGISUB apresenta-se com alto potencial para aplicação de controle de manipuladores seja em ROVs ou até mesmo em AUVs, sendo que este último encontra-se no limite da fronteira tecnológica para atividades de intervenção e é uma tendência do segmento de O&G para aumentar o nível de segurança destas operações com possibilidade de reduzir a demanda por pessoal embarcado no *offshore*.

Por fim, uma constatação da pesquisa foi a falta de *dataset* de imagens de equipamentos submarinos quando instalados em condições reais de operação. Portanto, a geração e disponibilização de um *dataset* contendo 2.847 imagens de um equipamento subaquático real, no caso uma Árvore de Natal Molhada - ANM, e que possam ser utilizadas para novas pesquisas futuras, proporcionando condição de novos direcionamentos no desenvolvimento de sistemas desta natureza, também é uma contribuição deste trabalho de investigação para a ciência.

5.3 Recomendação de atividades para futuras pesquisas

Como proposta de atividade futura desta pesquisa, uma questão importante a ser tratada é em relação à diminuição do tempo para geração das nuvens de pontos. A utilização de algoritmos baseados em paralelismo para rastreamento em GPU (*Graphics Processing Unit*) reduziria o tempo de busca das *features* e, conseqüentemente, o tempo total para geração das nuvens de pontos 3D, o que traria vantagens em termos de desempenho. A busca pelo processamento *on-line* viabilizaria o emprego do DIGISUB também para controle autônomo de atividades de intervenção realizadas por ROVs ou AUVs em instalações subaquáticas.

Outra proposta de atividade futura é a utilização de métodos baseados em *deep learning* para geração de nuvens de pontos 3D. De acordo com Li et al. (2020), a geração de nuvens de pontos 3D baseadas em rede neural convolucional gráfica (GCNN, do inglês *graph convolutional neural network*) apresenta resultados promissores, mas ainda continua sendo um desafio para novos trabalhos e melhorias.

Também é uma proposta de novos trabalhos o desenvolvimento de um sistema de autocalibração e que incorpore adicionalmente a funcionalidade de compensação dos efeitos da salinidade da água do mar, porém sem a necessidade de uma segunda câmera. Função esta, encontrada no método Pinax (LUCZYŃSKI; PFINGSTHORN; BIRK, 2017b) e que foi utilizada no DIGISUB.

Referências

- ABELES, P. Examination of hybrid image feature trackers. In: *Advances in Visual Computing*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013. v. 8034, p. 552–561. ISBN 10.1007/978-3-642-41939-3₅₄.
- AHN, H.; KEUM, B.; KIM, D.; LEE, H. S. Adaptive local tone mapping based on retinex for high dynamic range images. In: *2013 IEEE International Conference on Consumer Electronics (ICCE)*. [S.l.: s.n.], 2013. p. 153–156.
- ANCUTI, C.; ANCUTI, C. O.; HABER, T.; BEKAERT, P. Enhancing underwater images and videos by fusion. In: *2012 IEEE Conference on Computer Vision and Pattern Recognition*. [S.l.: s.n.], 2012. p. 81–88.
- BAILER, C.; FINCKH, M.; LENSCH, H. P. A. Scale robust multi view stereo. In: FITZGIBBON, A.; LAZEBNIK, S.; PERONA, P.; SATO, Y.; SCHMID, C. (Ed.). *Computer Vision – ECCV 2012*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012. p. 398–411.
- BANZ, C.; HESSELBARTH, S.; FLATT, H.; BLUME, H.; PIRSCH, P. Real-time stereo vision system using semi-global matching disparity estimation: Architecture and fpga-implementation. In: *2010 International Conference on Embedded Computer Systems: Architectures, Modeling and Simulation*. [S.l.: s.n.], 2010. p. 93–101.
- BARDOW, P.; DAVISON, A. J.; LEUTENEGGER, S. Simultaneous optical flow and intensity estimation from an event camera. In: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. [S.l.: s.n.], 2016. p. 884–892.
- BAY, H.; ESS, A.; TUYTELAARS, T.; GOOL, L. V. Speeded up robust features (surf). *Computer Vision and Image Understanding*, v. 110, n. 3, p. 346–359, 2008.
- BAY, H.; TUYTELAARS, T.; GOOL, L. V. Surf: Speeded up robust features. *9th European Conference on Computer Vision - EVVC*, v. 3951, n. 1, p. 404–417, 2006.
- BECHTOLD, P.; HOHENSTEIN, R.; SCHMIDT, M. Evaluation of disparate laser beam deflection technologies by means of number and rate of resolvable spots. *Opt. Lett.*, OSA, v. 38, n. 16, p. 2934–2937, Aug 2013. Disponível em: <<http://ol.osa.org/abstract.cfm?URI=ol-38-16-2934>><http://ol.osa.org/abstract.cfm?URI=ol-38-16-2934>.
- BERGER, M. et al. A survey of surface reconstruction from point clouds. *Computer Graphics Forum*, v. 36, p. n/a–n/a, 03 2016.
- BIANCO, G.; GALLO, A.; BRUNO, F.; MUZZUPAPPA, M. A comparative analysis between active and passive techniques for underwater 3d reconstruction of close-range

objects. *Sensors*, v. 13, n. 8, p. 11007–11031, 2013. ISSN 1424-8220. Disponível em: <<https://www.mdpi.com/1424-8220/13/8/11007>><https://www.mdpi.com/1424-8220/13/8/11007>.

BOBKOV, V. A.; KUDRYASHOV, A. P.; MELMAN, S. V.; MAY, V. P. Image-based navigation of autonomous underwater robot and 3d reconstruction of environment. In: *2018 3rd Russian-Pacific Conference on Computer Technology and Applications (RPC)*. [S.l.: s.n.], 2018. p. 1–4.

BROEKMAN, A.; GRÄBE, P. J. Pasmvs: A perfectly accurate, synthetic, path-traced dataset featuring specular material properties for multi-view stereopsis training and reconstruction applications. *Data in Brief*, v. 32, p. 106219, 2020. ISSN 2352-3409. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S2352340920311136>><https://www.sciencedirect.com/science/article/pii/S2352340920311136>.

BROWN, M.; SZELISKI, R.; WINDER, S. Multiimage matching using multi-scale oriented patches. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, v. 1, p. 510–517, 2005.

BRUNO, F.; BIANCO, G.; MUZZUPAPPA, M.; BARONE, S.; RAZIONALE, A. Experimentation of structured light and stereo vision for underwater 3d reconstruction. *ISPRS Journal of Photogrammetry and Remote Sensing*, v. 66, n. 4, p. 508 – 518, 2011. ISSN 0924-2716. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0924271611000414>><http://www.sciencedirect.com/science/article/pii/S0924271611000414>.

CAIMI, F. M.; KOCAK, D. M.; DALGLEISH, F.; WATSON, J. Underwater imaging and optics: Recent advances. In: *OCEANS 2008*. [S.l.: s.n.], 2008. v. 2008-Supplement, p. 1–9.

CASTILLÓN, M.; PALOMER, A.; FOREST, J.; RIDAO, P. State of the art of underwater active optical 3d scanners. *Sensors*, MDPI AG, v. 19, n. 23, p. 5161, Nov 2019. ISSN 1424-8220. Disponível em: <<http://dx.doi.org/10.3390/s19235161>><http://dx.doi.org/10.3390/s19235161>.

CHEN, L.; ROTTENSTEINER, F.; HEIPKE, C. Invariant descriptor learning using a siamese convolutional neural network. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, v. 3, n. 3, p. 11–18, 2016.

CHEN, L.; ROTTENSTEINER, F.; HEIPKE, C. Feature detection and description for image matching: from hand-crafted design to deep learning. *Geo-spatial Information Science*, v. 24, n. 1, p. 58–74, 2021.

CHEN, Y.; YANG, K.; ZHANG, X.; XIA, M.; LI, W. Modelling of beam propagation and its applications for underwater imaging. *Frontiers of Optoelectronics in China*, v. 4, 12 2012.

- CHUA, S. Y.; GUO, N.; TAN, C. S.; WANG, X. Improved range estimation model for three-dimensional (3d) range gated reconstruction. *Sensors*, v. 17, n. 9, 2017. ISSN 1424-8220. Disponível em: <<https://www.mdpi.com/1424-8220/17/9-/2031>><https://www.mdpi.com/1424-8220/17/9/2031>.
- DALGLEISH, F. R.; CAIMI, F. M.; BRITTON, W. B.; ANDREN, C. F. Improved LLS imaging performance in scattering-dominant waters. In: HOU, W. W. (Ed.). *Ocean Sensing and Monitoring*. SPIE, 2009. v. 7317, p. 89 – 100. Disponível em: <<https://doi.org/10.1117/12.820836>><https://doi.org/10.1117/12.820836>.
- DEVASAGAYAM, J.; BOSMA, R.; COLLIER, C. M. A velocity program using the kanade–lucas–tomasi feature-tracking algorithm with demonstration for pressure and electroosmosis conditions. *ELECTROPHORESIS*, v. 43, n. 7-8, p. 865–878, 2022.
- DIGUMARTI, S. T. et al. Underwater 3d capture using a low-cost commercial depth camera. In: *2016 IEEE Winter Conference on Applications of Computer Vision (WACV)*. [S.l.: s.n.], 2016. p. 1–9.
- DRAP, P. et al. The rov 3d project: Deep-sea underwater survey using photogrammetry: Applications for underwater archaeology. *J. Comput. Cult. Herit.*, Association for Computing Machinery, New York, NY, USA, v. 8, n. 4, ago. 2015. ISSN 1556-4673. Disponível em: <<https://doi.org/10.1145/2757283>><https://doi.org/10.1145/2757283>.
- FALANGA, D.; KIM, S.; SCARAMUZZA, D. How fast is too fast? the role of perception latency in high-speed sense and avoid. *IEEE Robotics and Automation Letters*, v. 4, n. 2, p. 1884–1891, 2019.
- FELICIANO, F.; SOUZA, I.; LETA, F. Visão computacional aplicada à metrologia dimensional automatizada: Considerações sobre sua exatidão. *Engevista*, v. 7, 01 2005.
- FISCHLER, M. A.; BOLLES, R. C. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Commun. ACM*, Association for Computing Machinery, New York, NY, USA, v. 24, n. 6, p. 381–395, jun 1981. ISSN 0001-0782. Disponível em: <<https://doi.org/10.1145/358669-358692>><https://doi.org/10.1145/358669.358692>.
- FU, X. et al. A retinex-based enhancing approach for single underwater image. In: *2014 IEEE International Conference on Image Processing (ICIP)*. [S.l.: s.n.], 2014. p. 4572–4576.
- FUJIMURA, Y.; IYAMA, M.; HASHIMOTO, A.; MINOH, M. Photometric stereo in participating media considering shape-dependent forward scatter. *IEEE Conference on Computer Vision and Pattern Recognition*, v. 18, n. 23, p. 7445–7453, Jan. 2018.
- GALDRAN, A.; PARDO, D.; PICÓN, A.; ALVAREZ-GILA, A. Automatic red-channel underwater image restoration. In: *Journal of Visual Communication and Image Representation*. [S.l.: s.n.], 2015. v. 26, p. 132–145.

- GARCIA, R. et al. Exploring the seafloor with underwater robots. In: _____. *Computer Vision in Vehicle Technology*. John Wiley & Sons, Ltd, 2017. cap. 4, p. 75–99. ISBN 9781118868065. Disponível em: <<https://onlinelibrary.wiley.com/doi/abs/10.1002/9781118868065.ch4>><https://onlinelibrary.wiley.com/doi/abs/10.1002/9781118868065.ch4>.
- HALLER, I.; NEDEVSKI, S. Gpu optimization of the sgm stereo algorithm. In: *International Conference on Intelligent Computer Communication and Processing*. [S.l.: s.n.], 2010. p. 197–202.
- HARRIS, C.; STEPHENS, M. A combined corner and edge detector. *Proceedings of Alvey Vision Conference*, v. 15, p. 147–151, 1988.
- HE, R.; WANG, Z.; XIONG, H.; FENG, D. D. Single image dehazing with white balance correction and image decomposition. In: *2012 International Conference on Digital Image Computing Techniques and Applications (DICTA)*. [S.l.: s.n.], 2012. p. 1–7.
- HECHT, E. *Optics*. 5th ed. ed. London, UK: Pearson, 2016.
- HEIPKE, C.; ROTTENSTEINER, F. Deep learning for geometric and semantic tasks in photogrammetry and remote sensing. *Geo-spatial Information Science*, v. 23, n. 1, p. 10–19, 2020.
- HERB, M. Computing the stereo matching cost with a convolutional neural network seminar recent trends in 3 d computer vision. In: . [S.l.: s.n.], 2015.
- HILDEBRANDT, M.; KERDELS, J.; ALBIEZ, J.; KIRCHNER, F. A practical underwater 3d-laserscanner. In: *OCEANS 2008*. [S.l.: s.n.], 2008. p. 1–5.
- HIMRI, K. et al. Object recognition and pose estimation using laser scans for advanced underwater manipulation. In: *2018 IEEE/OES Autonomous Underwater Vehicle Workshop (AUV)*. [S.l.: s.n.], 2018. p. 1–6.
- HIRSCHMULLER, H. Stereo processing by semiglobal matching and mutual information. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 30, n. 2, p. 328–341, 2008.
- HUACHUAN, H. et al. High quality underwater imaging platform with laser range gated technique combining with image denoising and restoration. *SPIE*, v. 9273, 10 2014.
- HUMENBERGER, M.; ZINNER, C.; WEBER, M.; KUBINGER, W.; VINCZE, M. A fast stereo matching algorithm suitable for embedded real-time systems. In: *Computer Vision and Image Understanding*. [S.l.: s.n.], 2010. v. 114, p. 1180–1202.
- JAFFE, J. S.; MOORE, K. D.; MCLEAN, J.; STRAND, M. P. Underwater optical imaging: Status and prospects. *Oceanography*, 2001.

- JOHNSON-ROBERSON, M. et al. High-resolution underwater robotic vision-based mapping and three-dimensional reconstruction for archaeology. *Journal of Field Robotics*, v. 34, n. 4, p. 625–643, 2017.
- JOHNSON-ROBERSON, M.; PIZARRO, O.; WILLIAMS, S. B.; MAHON, I. Generation and visualization of large-scale three-dimensional reconstructions from underwater robotic surveys. *Journal of Field Robotics*, v. 27, n. 1, p. 21–51, 2010.
- KAEHLER, A.; BRADSKY, G. *Learning OpenCV 3: Computer Vision in C++ with the OpenCV Library*. [S.l.]: O’Reilly Media; 1st edition, 2017, 2017.
- KAISER, M. J. The role of factor and activity-based models in offshore operating cost estimation. *Journal of Petroleum Science and Engineering - Elsevier*, v. 174, p. 1062–1092, 2019.
- KAMPMANN, P. et al. How ai and robotics can support marine mining. In: . [S.l.: s.n.], 2018.
- KIM, H.; HANDA, A.; BENOSMAN, R.; IENG, S.; DAVISON, A. Simultaneous mosaicing and tracking with an event camera. In: *Proceedings of the British Machine Vision Conference*. [S.l.]: BMVA Press, 2014.
- KIM, J.-H.; JANG, W.-D.; SIM, J.-Y.; KIM, C.-S. Optimized contrast enhancement for real-time image and video dehazing. In: *Journal of Visual Communication and Image Representation*. [S.l.: s.n.], 2014. v. 24, n. 3, p. 410–425.
- KIM, T.; CHOI, Y.; KO, Y. Development of underwater 3d shape measurement system with improved radiation tolerance. *Nuclear Engineering and Technology*, 2020. ISSN 1738-5733.
- KOCAK, D. M.; CAIMI, F. M. The current art of underwater imaging with a glimpse of the past and vision of the future. *Marine Technology Society Journal*, v. 39, n. 3, p. 5–26, 2005. ISSN 0025-3324.
- KONOLIGE, K. Projected texture stereo. In: *2010 IEEE International Conference on Robotics and Automation*. [S.l.: s.n.], 2010. p. 148–155.
- KOWALCZUK, J.; PSOTA, E. T.; PEREZ, L. C. Real-time stereo matching on cuda using an iterative refinement method for adaptive support-weight correspondences. *IEEE Transactions on Circuits and Systems for Video Technology*, v. 23, n. 1, p. 94–104, 2013.
- LEDEZMA, F. D. et al. A market survey of offshore underwater robotic inspection technologies for the oil and gas industry. In: . [S.l.: s.n.], 2015.
- LERØEN, B. V. Invisible technology: from tommeliten to snøhvit. *Norwegian petroleum technology: a success story*, KEILEN, H, 2005.

- LEVENBERG, K. A method for the solution of certain non-linear problems in least squares. In: *Quarterly of Applied Mathematics*. [S.l.: s.n.], 1944. p. 164–168.
- LI, J. et al. 3d large-scale point cloud semantic segmentation using optimal feature description vector network: Ofdv-net. *IEEE Access*, v. 8, p. 226285–226296, 2020.
- LI, T. et al. Optical sources optimization for 3d reconstruction based on underwater vision system. In: *2019 IEEE Underwater Technology (UT)*. [S.l.: s.n.], 2019. p. 1–4.
- LINDBERG, T. Feature detection with automatic scale selection. *International Journal of Computer Vision*, v. 30, n. 2, p. 79–116, 1998.
- LOPES, F.; SILVA, H.; ALMEIDA, J. M.; MARTINS, A.; SILVA, E. Structured light system for underwater inspection operations. p. 1–6, 2015.
- LOWE, D. G. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, v. 60, p. 91–110, 2004.
- MANSI, A.; HUSSEIN, W. Enhanced stereovision processing for improved 3d object reconstruction. In: *Proceedings of the 2018 International Conference on Sensors, Signal and Image Processing*. New York, NY, USA: Association for Computing Machinery, 2018. (SSIP 2018), p. 73–81. ISBN 9781450366205.
- MARIANI, P. et al. Range-gated imaging system for underwater monitoring in ocean environment. *Sustainability*, v. 11, n. 1, 2019. ISSN 2071-1050. Disponível em: <<https://www.mdpi.com/2071-1050/11/1/162>><https://www.mdpi.com/2071-1050/11/1/162>.
- MASSON, J. E. N.; PETRY, M. R. Comparison of mesh generation algorithms for railroad reconstruction. *2017 IEEE International Conference on Autonomous Robot Systems and Competitions (ICARSC)*, p. 266–271, 2017.
- MASSOT-CAMPOS, M.; OLIVER-CODINA, G. Optical sensors and methods for underwater 3d reconstruction. *Sensors*, v. 15, n. 12, p. 31525–31557, 2015. ISSN 1424-8220. Disponível em: <<https://www.mdpi.com/1424-8220/15/12/29864>><https://www.mdpi.com/1424-8220/15/12/29864>.
- MASSOT-CAMPOS, M.; OLIVER-CODINA, G.; KEMAL, H.; PETILLOT, Y.; BONIN-FONT, F. Structured light and stereo vision for underwater 3d reconstruction. In: *OCEANS 2015 - Genova*. [S.l.: s.n.], 2015. p. 1–6.
- Matos, G.; Buschinelli, P.; Pinto, T. Underwater laser triangulation sensor model with flat refractive interfaces. *IEEE Journal of Oceanic Engineering*, v. 45, n. 3, p. 937–945, 2020.
- MAY, S.; PERVOELZ, K.; SURMANN, H. 3d cameras: 3d computer vision of wide scope. In: OBINATA, G.; DUTTA, A. (Ed.). *Vision Systems*. Rijeka: IntechOpen, 2007. cap. 11. Disponível em: <<https://doi.org/10.5772/4988>><https://doi.org/10.5772/4988>.

- MCLEOD, D.; JACOBSON, J.; HARDY, M.; EMBRY, C. Autonomous inspection using an underwater 3d lidar. In: *2013 OCEANS - San Diego*. [S.l.: s.n.], 2013. p. 1–8.
- MENNA, F.; AGRAFIOTIS, P.; GEORGOPOULOS, A. State of the art and applications in archaeological underwater 3d recording and mapping. *Journal of Cultural Heritage*, v. 33, p. 231 – 248, 2018. ISSN 1296-2074. Cultural heritage in times of armed conflicts in the Middle East: Much more than material damage?
- MENNA, F.; NOCERINO, E.; TROISI, S.; REMONDINO, F. A photogrammetric approach to survey floating and semi-submerged objects. In: REMONDINO, F.; SHORTIS, M. R.; BEYERER, J.; LEÓN, F. P. (Ed.). *Videometrics, Range Imaging, and Applications XII; and Automated Visual Inspection*. SPIE, 2013. v. 8791, p. 117 – 131. Disponível em: <<https://doi.org/10.1117/12.2020464>><https://doi.org/10.1117/12.2020464>.
- MICHÁLEK, J.; CAPEK, M.; MAO, X. W.; KUBÍNOVÁ, L. Application of morphology filters to compensation of lateral illumination inhomogeneities in confocal microscopy images. In: . [S.l.: s.n.], 2010.
- MIKOLAJCZYK, K.; SCHMID, C. Scale affine invariant interest point detectors. *International Journal of Computer Vision*, v. 60, p. 63–86, 2004.
- MISHKIN, D.; RADENOVIC, F.; MATAS, J. Repeatability is not enough: Learning affine regions via discriminability. *Proceedings of the European Conference on Computer Vision*, p. 284–300, 2018.
- MORAIS, J. M. de. *Petróleo em águas profundas: uma história tecnológica da PETROBRAS na exploração e produção offshore*. [S.l.]: Instituto de Pesquisa Econômica Aplicada - IPEA, 2013.
- MORAVEC, H. P. Towards automatic visual obstacle avoidance. In: *Proceedings of 5th International Joint Conference on Artificial Intelligence*. [S.l.: s.n.], 1977.
- MORAVEC, H. P. Visual mapping by a robot rover. In: *International Joint Conference on Artificial Intelligence*. [S.l.: s.n.], 1979.
- MOREL, A. Optical properties of pure water and pure sea water. p. 1–24, 1974. Cited By 508.
- MOREL, J.-M.; YU, G. Asift: A new framework for fully affine invariant image comparison. *SIAM J. Imaging Sci.*, v. 2, p. 438–469, 2009.
- MSTAFA, R.; ELLEITHY, K. A video steganography algorithm based on kanade-lucas-tomasi tracking algorithm and error correcting codes. *Multimed Tools Appl*, v. 75, p. 10311–10333, 2016.
- OLIVEIRA, J. C. V. d. *Vocabulário Internacional de Metrologia: Conceitos fundamentais e gerais e termos associados (VIM 2012)*. Duque de Caxias, Brasil: INMETRO, 2012. ISBN 978-85-86920-09-7.

ONMEK, Y.; TRIBOULET, J.; DRUON, S.; MELINE, A.; JOUVENCEL, B. Evaluation of underwater 3d reconstruction methods for archaeological objects: Case study of anchor at mediterranean sea. In: *2017 3rd International Conference on Control, Automation and Robotics (ICCAR)*. [S.l.: s.n.], 2017. p. 394–398.

PALOMER, A.; RIDAO, P.; FOREST, J.; RIBAS, D. Underwater laser scanner: Ray-based model and calibration. *IEEE/ASME Transactions on Mechatronics*, v. 24, n. 5, p. 1986–1997, 2019.

PALOMER, A.; RIDAO, P.; RIBAS, D.; FOREST, J. Underwater 3d laser scanners: The deformation of the plane. In: _____. *ensing and Control for Autonomous Vehicles: Applications to Land, Water and Air Vehicles*. Cham: Springer International Publishing, 2017. p. 73–88. ISBN 978-3-319-55372-6. Disponível em: <https://doi.org/10.1007/978-3-319-55372-6_4>https://doi.org/10.1007/978-3-319-55372-6_4.

RAKOTOSAONA, M.-J.; BARBERA, V. L.; GUERRERO, P.; MITRA, N.; OVSJANIKOV, M. Pointcleannet : Learning to denoise and remove outliers from dense point clouds: Pointcleannet. *Computer Graphics Forum*, v. 39, 06 2019.

RISHOLM, P. et al. Real-time super-resolved 3d in turbid water using a fast range-gated cmos camera. *Appl. Opt.*, OSA, v. 57, n. 14, p. 3927–3937, May 2018. Disponível em: <<http://ao.osa.org/abstract.cfm?URI=ao-57-14-3927>><http://ao.osa.org/abstract.cfm?URI=ao-57-14-3927>.

ROMAN, C.; INGLIS, G.; RUTTER, J. Application of structured light imaging for high resolution mapping of underwater archaeological sites. In: . [S.l.]: OCEANS'10 IEEE, 2010. p. 24–27.

ROSTEN, E.; PORTER, R.; DRUMMOND, T. Faster and better: A machine learning approach to corner detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 32, n. 1, p. 105–119, 2010.

ROSWELL, A.; HALIKAS, G. *The Index of Refraction of Seawater: Technical Report*. [S.l.]: University of California, San Diego, Visibility Laboratory of the Scripps Institute of Oceanography, 1976, 2010.

RUSSELL, S.; NORVIG, P. *Inteligência Artificial*. [S.l.]: Elsevier, Rio de Janeiro, 2013, 2013.

SCHETTINI, R.; CORCHS, S. Underwater image processing: State of the art of restoration and image enhancement methods. *EURASIP Journal on Advances in Signal Processing*, v. 2010, 01 2010.

SEDLAZECK, A.; KOCH, R. erspective and non-perspective camera models in underwater imaging – overview and error analysis. In: DELLAERT, F.; FRAHM, J.-M.; POLLEFEYS, M.; LEAL-TAIXÉ, L.; ROSENHAHN, B. (Ed.). *Outdoor and Large-Scale*

Real-World Scene Analysis. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012. p. 212–242.

SEITZ, S.; CURLESS, B.; DIEBEL, J.; SCHARSTEIN, D.; SZELISKI, R. A comparison and evaluation of multi-view stereo reconstruction algorithms. In: *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06)*. [S.l.: s.n.], 2006. v. 1, p. 519–528.

SHI, J.; TOMASI. Good features to track. In: *1994 Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*. [S.l.: s.n.], 1994. p. 593–600.

SHORTIS, M. Calibration techniques for accurate measurements by underwater camera systems. *Sensors*, v. 15, n. 12, p. 30810–30826, 2015. ISSN 1424-8220. Disponível em: <<https://www.mdpi.com/1424-8220/15/12/29831>><https://www.mdpi.com/1424-8220/15/12/29831>.

SIEGWART, R.; NOURBAKHSI, I.; SCARAMUZZA, D. *Introduction to Autonomous Mobile Robots*. 2nd edition. ed. USA: Massachusetts Institute of Technology, 2004. ISBN 978-0-2626-01535-6.

SKARLATOS, D. et al. Project imareculture: Advanced vr, immersive serious games and augmented reality as tools to raise awareness and access to european underwater cultural heritage. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, v. 10058 LNCS, p. 805–813, 2016.

SKINNER, K. A.; ISCAR, E.; JOHNSON-ROBERSON, M. Automatic color correction for 3d reconstruction of underwater scenes. In: *2017 IEEE International Conference on Robotics and Automation (ICRA)*. [S.l.: s.n.], 2017. p. 5140–5147.

SKINNER, K. A.; JOHNSON-ROBERSON, M. Detection and segmentation of underwater archaeological sites surveyed with stereo-vision platforms. In: *OCEANS 2015 - MTS/IEEE Washington*. [S.l.: s.n.], 2015. p. 1–7.

SMITH, R. C.; BAKER, K. S. Optical properties of the clearest natural waters (200–800 nm). *Appl. Opt.*, OSA, v. 20, n. 2, p. 177–184, Jan 1981. Disponível em: <<http://ao.osa.org/abstract.cfm?URI=ao-20-2-177>><http://ao.osa.org/abstract.cfm?URI=ao-20-2-177>.

SPANGENBERG, R.; LANGNER, T.; ADFELDT, S.; ROJAS, R. Large scale semi-global matching on the cpu. In: *2014 IEEE Intelligent Vehicles Symposium Proceedings*. [S.l.: s.n.], 2014. p. 195–201.

STYLIANIDIS, E.; REMONDINO, F. *3D recording, documentation and management of cultural heritage*. 1. ed. Dunbeath, Caithness, Scotland, UK: Whittles Publishing, 2016. ISBN 978-184995-168-5.

- SULAMI, M.; GLATZER, I.; FATTAL, R.; WERMAN, M. Automatic recovery of the atmospheric light in hazy images. In: *2014 IEEE International Conference on Computational Photography (ICCP)*. [S.l.: s.n.], 2014. p. 1–11.
- SZELISKI, R. *Computer Vision: Algorithms and Applications*. 2nd ed. ed. London, UK: Springer Nature Switzerland AG 2022, 2022. ISBN 978-3-030-34372-9.
- TECCHIO, P. P. V. et al. Overview of robotic applications on offshore ogi imr: Splash zone. In: *OCEANS 2021: San Diego – Porto*. [S.l.: s.n.], 2021. p. 1–10.
- TOMASI, C.; KANADE, T. Detection and tracking of point features. 1991.
- TRIGGS, B.; MCLAUCHLAN, P. F.; HARTLEY, R. I.; FITZGIBBON, A. W. Bundle adjustment — a modern synthesis. In: TRIGGS, B.; ZISSERMAN, A.; SZELISKI, R. (Ed.). *Vision Algorithms: Theory and Practice*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2000. p. 298–372. ISBN 978-3-540-44480-0.
- WANG, C.; CHENG, M. Nonmetric camera calibration for underwater laser scanning system. *IEEE Journal of Oceanic Engineering*, v. 32, n. 2, p. 383–399, April 2007. ISSN 1558-1691.
- Wang, C. et al. Research and experiment of an underwater stereo vision system. In: *OCEANS 2019 - Marseille*. [S.l.: s.n.], 2019. p. 1–5.
- WANG, M.; WANG, X.; SUN, L.; YANG, Y.; ZHOU, Y. Underwater 3d deblurring-gated range-intensity correlation imaging. *Opt. Lett.*, OSA, v. 45, n. 6, p. 1455–1458, Mar 2020. Disponível em: <<http://ol.osa.org/abstract.cfm?URI=ol-45-6-1455>><http://ol.osa.org/abstract.cfm?URI=ol-45-6-1455>.
- WEIDNER, N.; RAHMAN, S.; LI, A. Q.; REKLEITIS, I. Underwater cave mapping using stereo vision. In: *2017 IEEE International Conference on Robotics and Automation (ICRA)*. [S.l.: s.n.], 2017. p. 5709–5715.
- WOLFF, J. et al. Point cloud noise and outlier removal for image-based 3d reconstruction. In: *2016 Fourth International Conference on 3D Vision (3DV)*. [S.l.: s.n.], 2016. p. 118–127.
- WOODFILL, J. I.; HERZEN, B. V. Real-time stereo vision on the parts reconfigurable computer. *Proceedings. The 5th Annual IEEE Symposium on Field-Programmable Custom Computing Machines Cat. No.97TB100186*), p. 201–210, 1997.
- WU, C.; AGARWAL, S.; CURLESS, B.; SEITZ, S. M. Multicore bundle adjustment. In: *CVPR 2011*. [S.l.: s.n.], 2011. p. 3057–3064.
- YAMAFUNE, K.; TORRES, R.; CASTRO, F. Multi-image photogrammetry to record and reconstruct underwater shipwreck sites. *Journal of Archaeological Method and Theory*, v. 24, n. 1, p. 703–725, 2017. Disponível em: <<https://link.springer.com/article/10.1007/s10816-016-9283-1>><https://link.springer.com/article/10.1007/s10816-016-9283-1>.

- YANG, Y. et al. 3d reconstruction for underwater laser line scanning. In: *2013 MTS/IEEE OCEANS - Bergen*. [S.l.: s.n.], 2013. p. 1–3.
- YI, K. M.; TRULLS, E.; LEPETIT, V.; FUA, P. Lift: Learned invariant feature transform. *Proceedings of the European Conference on Computer Vision*, p. 467–483, 2016.
- YU, L. et al. Inspection robots in oil and gas industry: a review of current solutions and future trends. In: *2019 25th International Conference on Automation and Computing (ICAC)*. [S.l.: s.n.], 2019. p. 1–6.
- ZAGATTI, R. et al. Flatfish resident auv: Leading the autonomy era for subsea oil and gas operations. *Offshore Technology Conference*, Abril 2018. Disponível em: <<https://doi.org/10.4043/28881-MS>><https://doi.org/10.4043/28881-MS>.
- ZANUTTIGH, P. et al. Time-of-flight and structured light depth cameras. In: _____. *Technology and Applications*. [S.l.]: Springer, Cham, 2016. p. 1–355. ISBN 978-3-319-30973-6.
- ZHANG, K.; LIU, M.; ZHANG, J.; DONG, Z. Pa-mvsnet: Sparse-to-dense multi-view stereo with pyramid attention. *IEEE Access*, v. 9, p. 27908–27915, 2021.
- ZHANG, Z. A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 22, n. 11, p. 1330–1334, 2000. Cited By 8546.
- ÖZYEŞİL, O.; VORONINSKI, V.; BASRI, R.; SINGER, A. A survey of structure from motion. *Acta Numerica*, Cambridge University Press, v. 26, p. 305–364, 2017.
- ŁUCZYŃSKI, T.; PFINGSTHORN, M.; BIRK, A. Image rectification with the pinax camera model in underwater stereo systems with verged cameras. In: *OCEANS 2017 - Anchorage*. [S.l.: s.n.], 2017. p. 1–7.
- ŁUCZYŃSKI, T.; PFINGSTHORN, M.; BIRK, A. The pinax-model for accurate and efficient refraction correction of underwater cameras in flat-pane housings. In: *Ocean Engineering*. [S.l.: s.n.], 2017. v. 133, p. 9–22.

Projeto detalhado de *hardware* do protótipo

Nesse trabalho de investigação, um sistema de digitalização baseado em visão estéreo e num gerador de padrão de feixe de *laser* é proposto, combinado com a aplicação de diferentes algoritmos de filtros para simplificar a nuvem de pontos de alta densidade gerada, característica deste tipo de sistema, além de reconstruir sua superfície. Inicialmente, foi proposto um método de fusão de dados com o objetivo de minimizar erros introduzidos nas nuvens de pontos em decorrência das perturbações de movimentação que o ROV sofre no ambiente de operação. Para isto, dados de uma INS (*Inertial Navigation System*) são utilizados para estimar a localização relativa (pose e orientação) do protótipo, provendo informações necessárias para realizar transformações geométricas sobre as nuvens de pontos.

A Fig. A.1 ilustra uma visão geral do conceito proposto. O usuário da aplicação acessa o sistema através de uma interface de operação disponível na estação de trabalho na embarcação de apoio às operações. Todo processamento e acionamento dos dispositivos de *hardware* está inserido no computador embarcado no protótipo, o qual está acoplado ao ROV. A comunicação entre estação de trabalho e o protótipo segue a arquitetura *Ethernet* através de uma rede local que utiliza o umbilical do ROV para envio de pacotes de dados. Multiplexadores (MUX) estão instalados na embarcação (*topside*) e no ROV, o qual é responsável pela alimentação do sistema de potência do protótipo e pela comunicação.

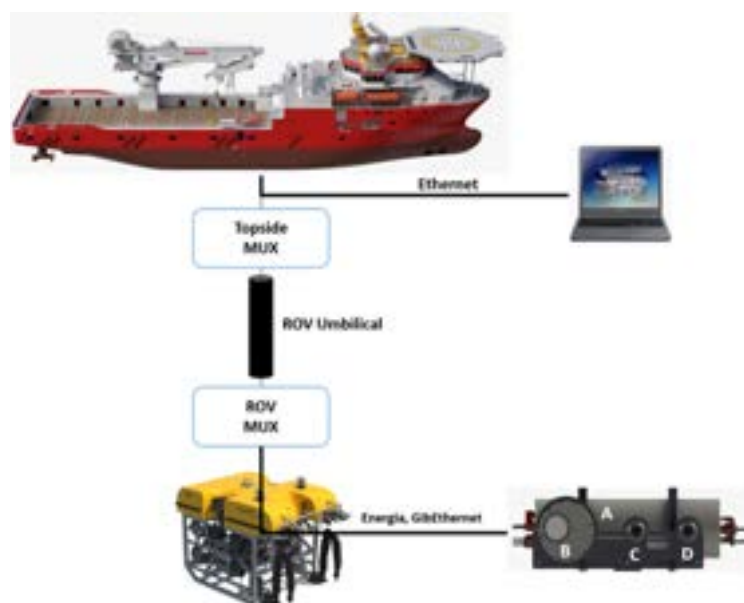


Figura A.1: Conceito preliminar da solução.

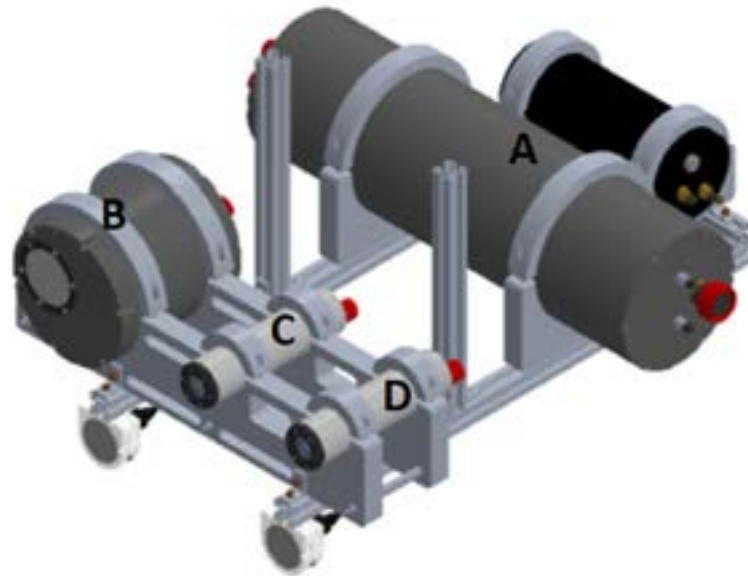


Figura A.2: Detalhe dos componentes do protótipo.

O protótipo é totalmente encapsulado e composto por:

- Uma unidade de processamento e potência (Fig. [A.2](#) detalhe **A**).
- Um sistema óptico (*laser*, Fig. [A.2](#) detalhe **B**).
- Duas câmeras RGB (*Red, Green, Blue*) que formam o sistema de visão estéreo (Fig. [A.2](#) detalhes **C** e **D**).
- Na parte central do protótipo encontra-se o invólucro, responsável pela conexão de energia entre o umbilical do ROV e os demais sistemas, bem como abrigar as unidades de processamento de dados adquiridos e de geração das nuvens de pontos (Fig. [A.2](#) detalhe **A**).
- O encapsulamento óptico abriga um sistema composto por um *laser* de linha, que é projetado sobre um espelho, rotacionado por um galvanômetro (Fig. [A.2](#) detalhe **B**).

As dimensões do protótipo montado podem ser identificadas na Figura [A.3](#).

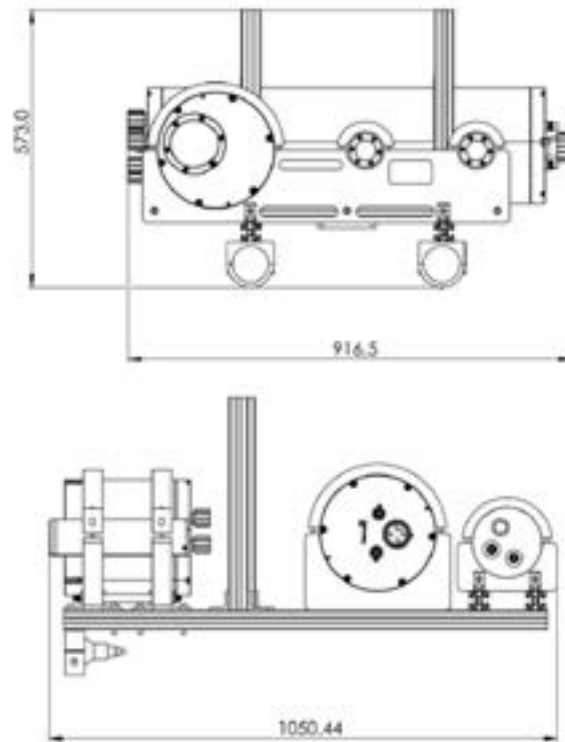


Figura A.3: Dimensões do protótipo montado.

A Fig. A.4 apresenta a arquitetura geral do protótipo e seus respectivos módulos funcionais. A arquitetura do protótipo foi projetada com características modulares para facilitar a utilização de todos os componentes do sistema.

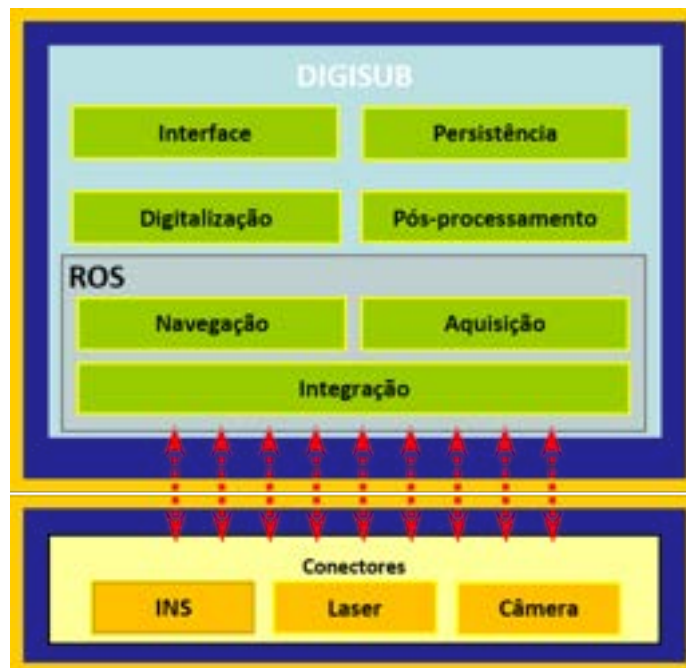


Figura A.4: Arquitetura Geral do Protótipo.

A.0.1 Projeto Eletroeletrônico

O projeto contempla um conjunto de sensores e atuadores independentes, que trabalham de forma sincronizada para realizar a digitalização de estruturas subaquáticas. Para que o processo de digitalização seja realizado de forma apropriada, uma fonte de alimentação provê energia para os sensores (INS e câmeras), dispositivo de iluminação (*laser*), assim como o galvanômetro, *switch*, unidade de aquisição de dados (DAQ), processador FPGA (*Field Programmable Gate Arrays*) e computador embarcado (NUC). Na Tab. A.1 são apresentados os componentes de *hardware* utilizados no protótipo.

Tabela A.1: Componentes de hardware do protótipo.

Componente	Modelo
Câmera RGB	Basler Ace acA1920-155uc
Lente	Fujinon CF16HA-1
Mini computador	Intel NUC NUC8i7HVK
Processador Stereo	Nerian SceneScan Pro
Switch	Ubiquiti Es-10xp-br Edge Switch 8-ports Gigabit
INS	VECTORNAV 200 Rugged GPS/INS
DAC	USB-1608G Series devices provide 16-bit
Laser	Uniform 532nm Green Line Laser Module
Galvanômetro	GVS211/M - Galvo de 1D 10mm / espelhos
Lâmpada LED	36 W e 5.000 lúmens

A tensão de alimentação para o protótipo é em corrente contínua e pode ser provida pelo ROV.

A comunicação entre o protótipo e o usuário é feito através de um *laptop* que utiliza um *enlace Ethernet* disponível na porta do *switch*. Os comandos são enviados à NUC, a qual funciona como central de aquisição e processamento. Os dados capturados são processados e, então, os resultados são enviados ao *laptop* através do cabo umbilical do ROV. A arquitetura eletroeletrônica completa é apresentada na Fig. A.5. Os protocolos de comunicação estão indicados junto às suas respectivas conexões.

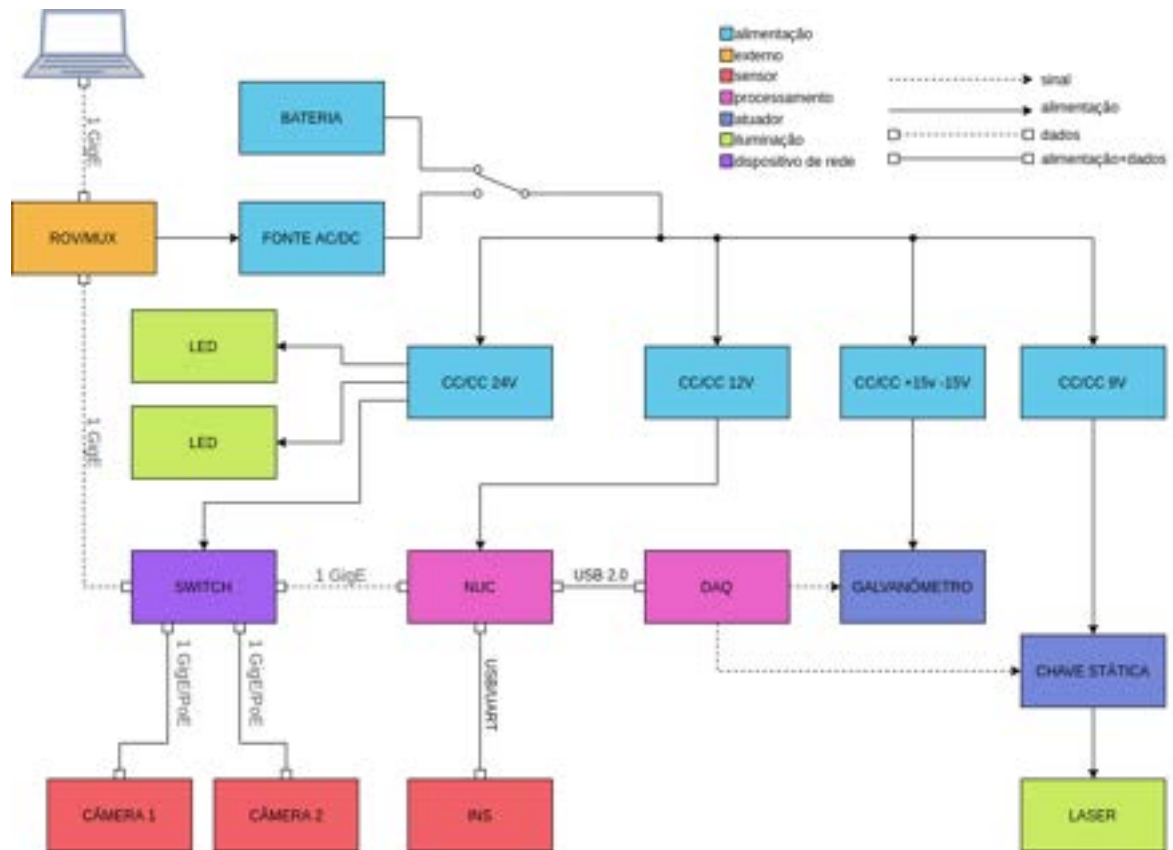


Figura A.5: Arquitetura Eletroeletrônica.

Os dispositivos precisam de fontes de corrente contínua que variam entre 5/ 9/ 12/ 15 e 24 V. Os dispositivos que funcionam a 5 V são alimentados via USB. As tensões de 9 V a 24 V são supridas pelo uso de conversores DC/DC. Os conversores, por sua vez, são alimentados pela fonte DC. Alguns são alimentados via *Power over Ethernet* (PoE) pelo *switch*. As conexões elétricas entre a unidade de processamento e os outros encapsulamentos, são realizadas através de cabos e conectores especiais para o uso específico em ambientes subaquáticos de altas profundidades e pressões, as quais o protótipo será exposto.

A.0.2 Projeto dos Encapsulamentos

O protótipo foi subdividido em quatro encapsulamentos, compostos por uma unidade de processamento e potência, duas câmeras RGB e um sistema óptico (*laser*), todos projetados para operar até uma profundidade 3.000 m de lâmina d'água, produzidos em material EN AW 6082 anodizado. A Fig. A.6 apresenta o *design* do encapsulamento óptico, massa de 17 kg, contendo uma janela de safira para projeção do laser de linha. A Fig. A.7 ilustra o *design* do encapsulamento de processamento e potência, massa de 37 kg. Já na

Fig. A.8 é apresentado o *design* do encapsulamento das câmeras RGB, massa de 1,9 kg.

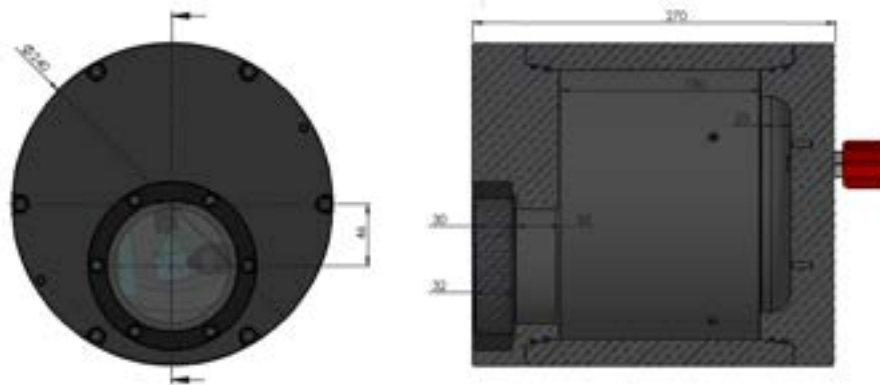


Figura A.6: Encapsulamento óptico.

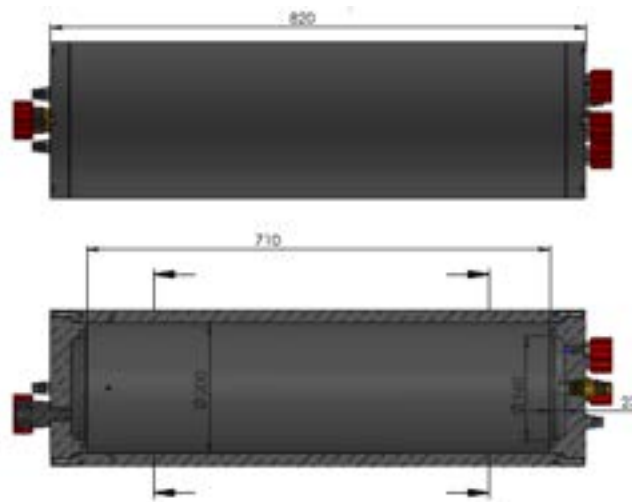


Figura A.7: Encapsulamento de processamento e potência.

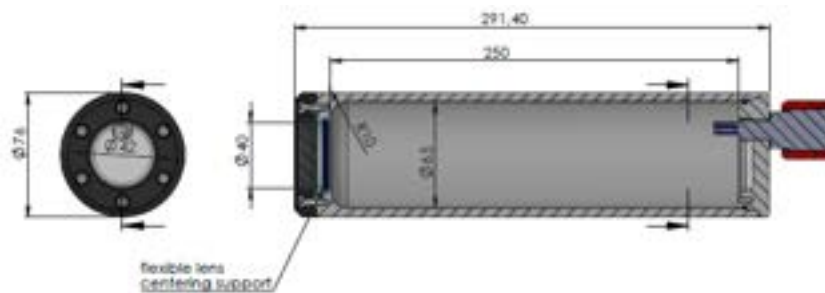


Figura A.8: Encapsulamento da câmera.

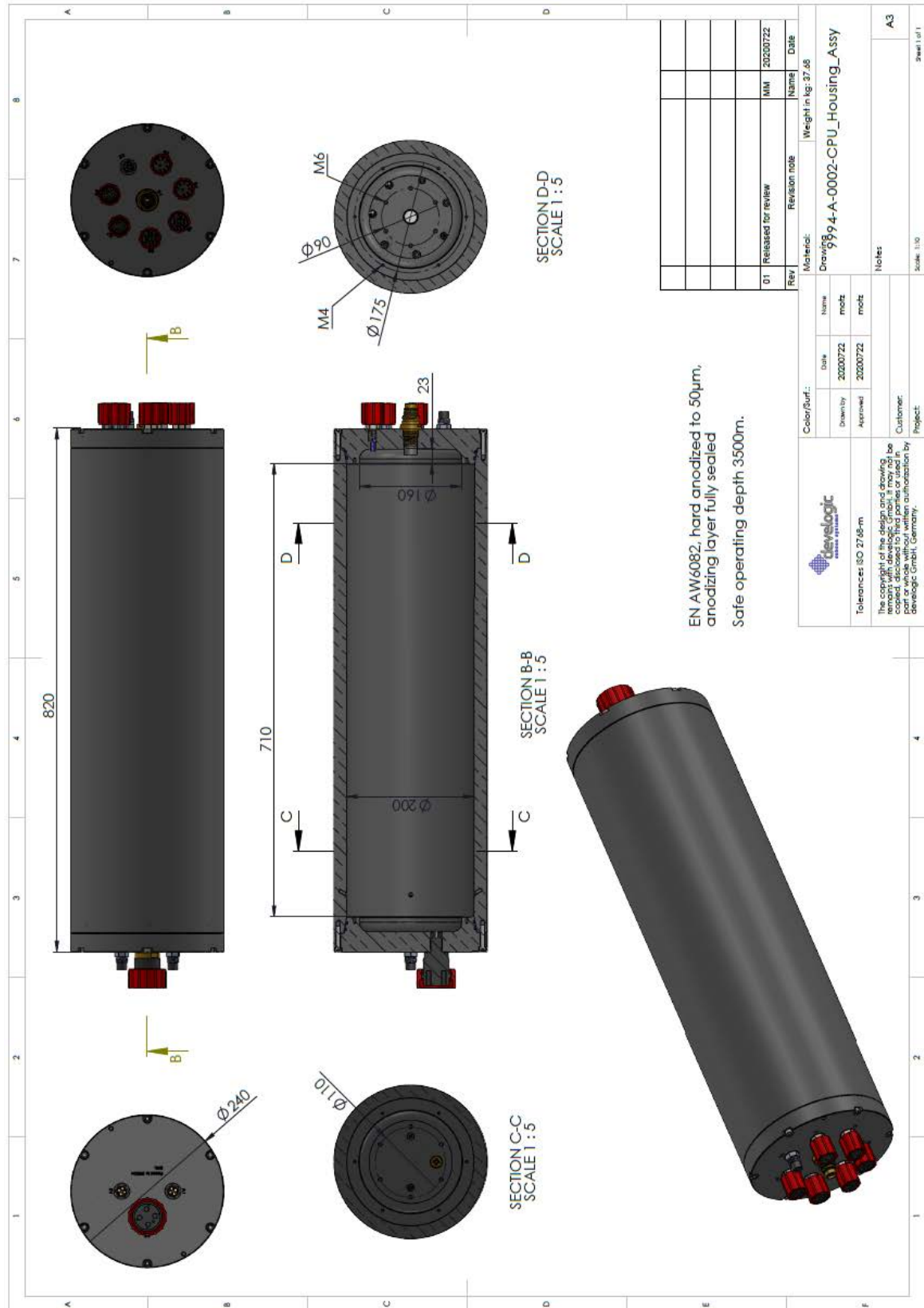


Figura A.9: Desenho técnico de construção do encapsulamento de processamento.

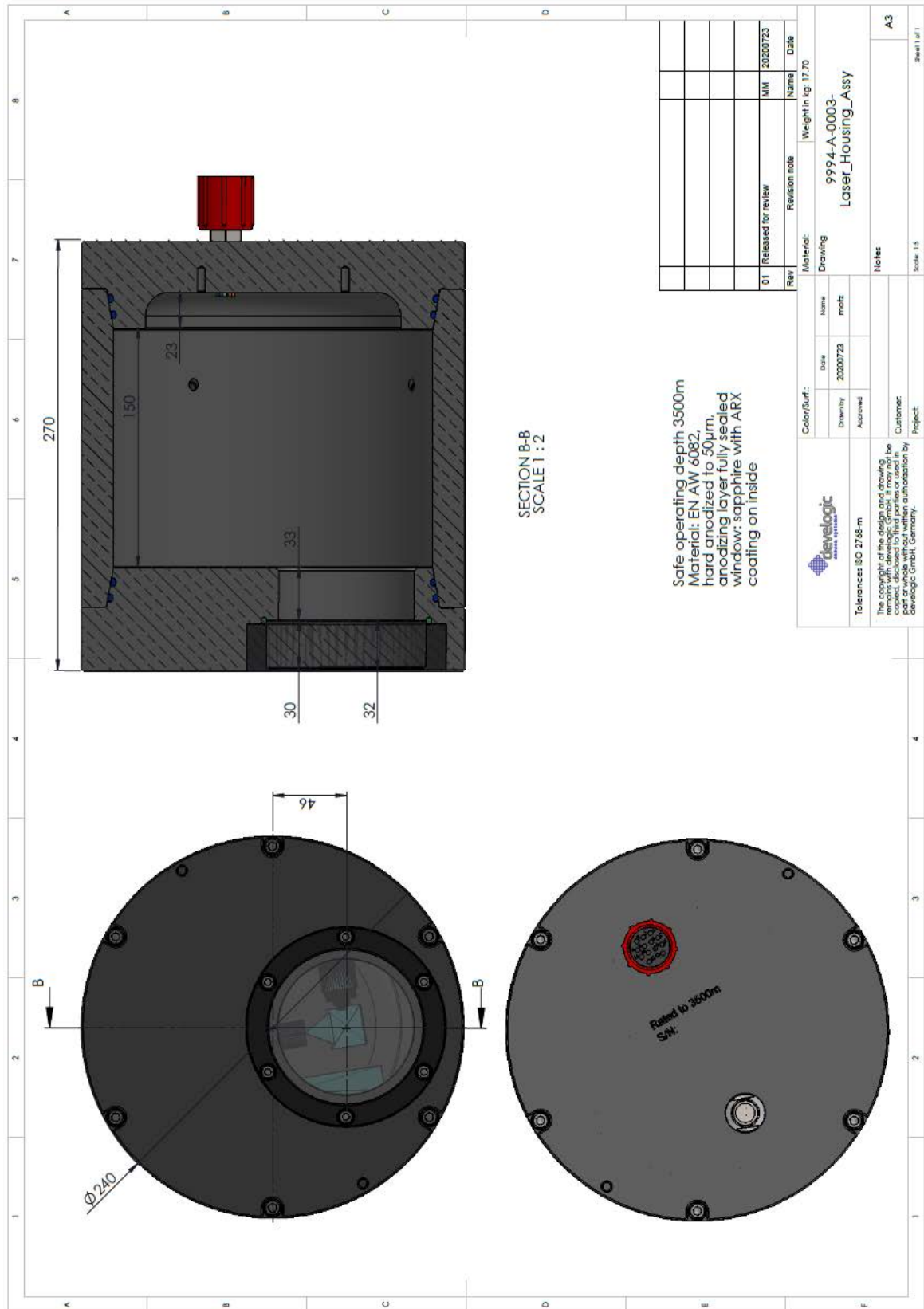


Figura A.10: Desenho técnico de construção do encapsulamento de projeção de luz estruturada.

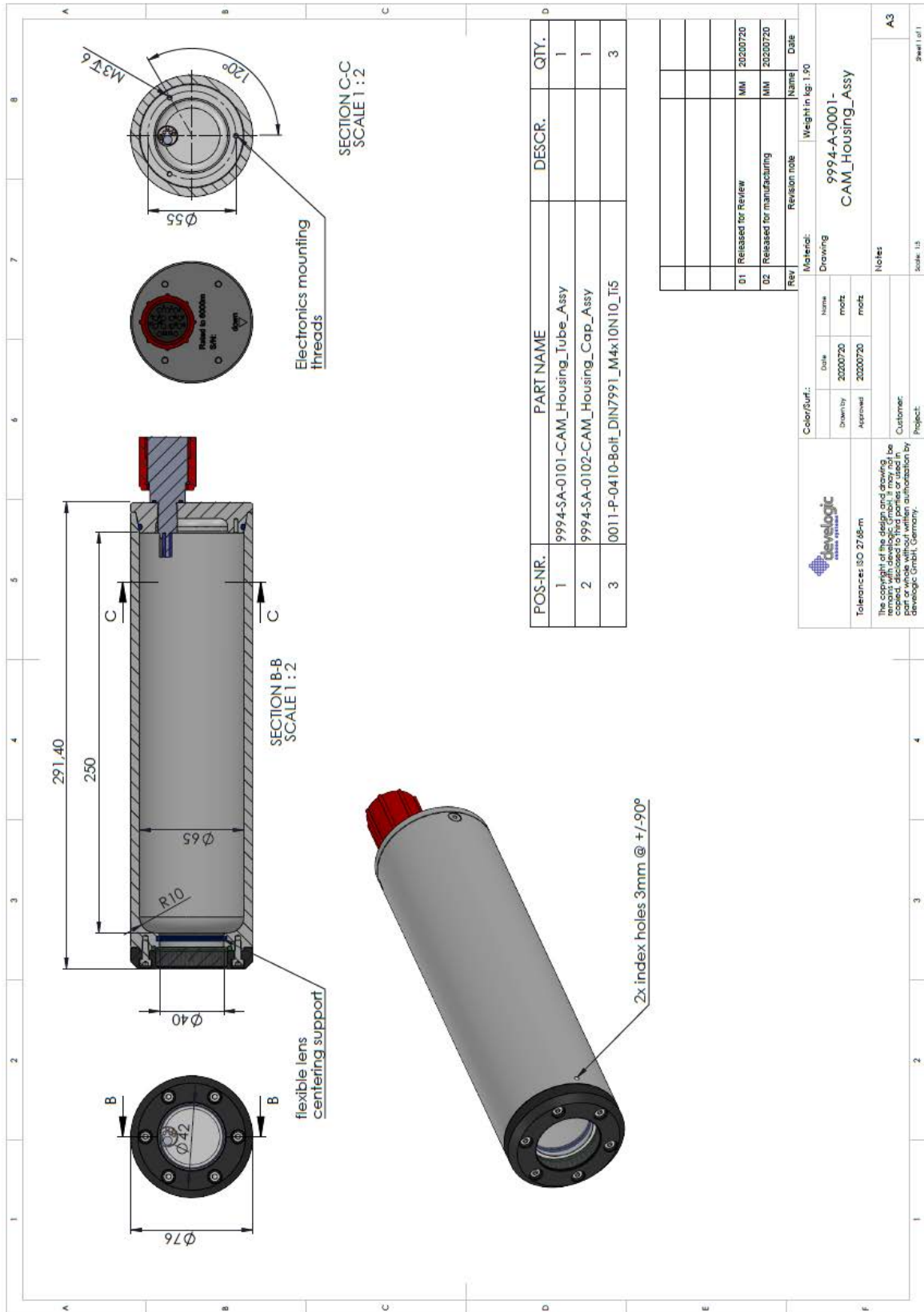


Figura A.11: Desenho técnico de construção do encapsulamento da câmera.

TRABALHOS PUBLICADOS

Artigos publicados:

Nardy, L., Pinheiro, O., Lepikson, H. . (2021). COMPUTER SYSTEM INTEGRATED WITH DIGITAL MODELS FOR RECONSTRUCTION OF UNDERWATER STRUCTURES WITH HIGH DEFINITION. *IEEE Latin America Transactions*, 20(2), 283–290. Disponível em: <https://latamt.ieeer9.org/index.php/transactions/article/view/5634>.

TECCHIO, PEDRO P. V.; VIEIRA, MIGUEL; BATISTA, LUCAS; CERQUEIRA, JORSIELE; LIMA, REBECA; PINHEIRO, OBERDAN; NARDY, LEONARDO; DOS REIS, MARCO A.; LEPIKSON, HERMAN; LOPES, LUIZ F. M. Overview of Robotic Applications on Offshore OGI IMR: Splash Zone In: OCEANS 2021: San Diego Porto, 2021, San Diego. OCEANS 2021: San Diego - Porto. Porto, Portugal: IEEE OCEANS 2021, 2021.

Nardy, L., Pinheiro, O., Lepikson, H. . INTELLMANIPULATOR: DESENVOLVIMENTO DE MANIPULADORES SUBAQUÁTICOS INTELIGENTES BASEADO EM SISTEMAS MULTIAGENTE, *International Symposium on Innovation and Technology (SIINTEC)* 2019.

ZAGATTI, R. et al. Flatfish resident AUV: Leading the autonomy era for subsea oil and gas operations. *Offshore Technology Conference*, Abril 2018. Disponível em: <https://doi.org/10.4043/28881-MS>.

Artigo submetido (aguardando análise do periódico):

Nardy, L., Pinheiro, O., Lepikson, H. . (2022). HIGH DEFINITION 3D SYSTEM FOR REMOTE UNDERWATER GEOMETRIC PRECISION MAPPING AND MEASUREMENT. *International Journal of Computer Vision*. Número de registro da submissão: VISI-D-22-00671.

Patente requerida:

DIGITALIZADOR 3D SUBMARINO. Tipo: Patente de Invenção (PI). Protocolo: BR 10 2021 012710 4.

Sistema computacional integrado a modelos digitais para reconstrução de estruturas subaquáticas com alta definição.

Leonardo de Paula Nardy

Salvador, 2022.