

UNIVERSIDADE FEDERAL DE MINAS GERAIS
Escola de Engenharia
Programa de Pós Graduação em Engenharia Elétrica

Gabriel Gomes

**Odometria Visual por Fluxo Óptico
Integrada a um Sistema de Localização
Multi-sensor Aplicado a Veículos de
Mineração**

Belo Horizonte

2025

Gabriel Gomes

**Odometria Visual por Fluxo Óptico Integrada a um
Sistema de Localização Multi-sensor Aplicado a Veículos
de Mineração**

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal de Minas Gerais, como requisito parcial para a obtenção do Título de Mestre em Engenharia Elétrica.

Orientador: Prof. Dr. Armando Alves Neto

Coorientador: Prof. Dr. Gustavo Medeiros Freitas

Belo Horizonte

2025

G633o Gomes, Gabriel.
Odometria visual por fluxo óptico integrada a um sistema de localização multi-sensor aplicado a veículos de mineração [recurso eletrônico] / Gabriel Gomes. – 2025.
1 recurso online (84 f. : il., color.) : pdf.

Orientador: Armando Alves Neto.
Coorientador: Gustavo Medeiros Freitas.

Dissertação (mestrado) – Universidade Federal de Minas Gerais, Escola de Engenharia.

Inclui bibliografia.

1. Engenharia elétrica – Teses. 2. Distâncias – Medição – Teses. 3. Sistema de Posicionamento Global – Teses. 4. Veículos a motor – Sistemas automáticos de localização – Teses. 5. Minas e mineração – Teses. I. Alves Neto, Armando. II. Freitas, Gustavo Medeiros. III. Universidade Federal de Minas Gerais. Escola de Engenharia. IV. Título.

CDU: 621.3(043)



UNIVERSIDADE FEDERAL DE MINAS GERAIS

Escola de Engenharia

COLEGIADO DO CURSO DE PÓS-GRADUAÇÃO EM Engenharia Elétrica

FOLHA DE APROVAÇÃO

"Odometria Visual por Fluxo Óptico Integrada A Um Sistema de Localização Multi-sensor Aplicado A Veículos de Mineração"

Gabriel Gomes

Dissertação de Mestrado submetida à Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em Engenharia Elétrica da Escola de Engenharia da Universidade Federal de Minas Gerais, como requisito para obtenção do grau de Mestre em Engenharia Elétrica.

Aprovada em 05 de setembro de 2025.

Por:

Prof. Dr. Armando Alves Neto
DELT (UFMG) - Orientador

Prof. Dr. Gustavo Medeiros Freitas
DEE (UFMG) - Coorientador

Prof. Dr. Gustavo Pessin
ITV (Vale)

Prof. Dr. Paulo Alfredo Frota Rezeck
(Invent Vision)



Documento assinado eletronicamente por **Armando Alves Neto, Professor do Magistério Superior**, em 05/09/2025, às 11:57, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Gustavo Medeiros Freitas, Professor do Magistério Superior**, em 05/09/2025, às 12:43, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Paulo Alfredo Frota Rezeck, Usuário Externo**, em 09/09/2025, às 09:42, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Gustavo Pessin, Usuário Externo**, em 09/09/2025, às 17:26, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site https://sei.ufmg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **4512130** e o código CRC **41AF4091**.

Em honra a quem ajuda os que precisam.

Agradecimentos

Gostaria, primeiramente, de agradecer a Deus por me dar forças para concluir este trabalho, que me acompanhou por tanto tempo e, em muitos momentos, me fez pensar em desistir. Agradeço também aos meus pais, que sempre me apoiaram e deixaram claro que eu teria um lugar para onde retornar.

Agradeço à minha família e aos meus amigos que mesmo distantes mantêm o mesmo carinho sempre que retorno para casa. Agradeço, de forma especial, aos meus amigos Matheus e Marisa que, mesmo compartilhando a distância comigo, me fazem recordar de casa.

Expresso meus agradecimentos aos meus colegas do laboratório CORO que, ao longo dos últimos anos, compartilharam conhecimento, histórias e vivências comigo, aos que estiveram do início até o fim, Arthur Vangasse, João Baião e Gilmar, aos que chegaram e foram embora, Sofia, Gabriel Malaquias, Franz, Natan e Luiza, aos que chegam, somem e depois voltam, Marcos, Elias, Bidas, Leandro e Álvaro, aos que acabaram de chegar, Elan, Thales e Diogo, e aos muitos outros que já estiveram por lá. Deixo também um agradecimento em especial à Vera, que sempre esteve lá cuidando de nossa bagunça.

Gostaria de agradecer à Universidade Federal de Minas Gerais (UFMG), ao Instituto Tecnológico Vale (ITV), à CAPES, ao CNPq e à Fundação Christiano Ottoni pelo apoio tecnológico e financeiro para desenvolvimento desta dissertação. Agradeço também ao meu orientador, Armando Alves Neto, e ao meu coorientador, Gustavo Medeiros Freitas, pelo apoio e ensinamentos durante esses anos de pós-graduação, e pela paciência nesta reta final de conclusão.

*"A vida é assim: esquenta e esfria,
aperta e daí afrouxa,
sossega e depois desinquieta.
O que ela quer da gente é coragem."
João Guimarães Rosa*

Resumo

Esta dissertação apresenta um método de estimação de odometria para veículos em ambientes de mineração, nos quais sistemas de localização baseados em GNSS (Sistema Global de Navegação por Satélite ou *Global Navigation Satellite System*) podem apresentar falhas ou degradação do sinal. Visando mitigar essas falhas, o método proposto recorre à odometria visual a partir de uma câmera direcionada para o solo. As imagens capturadas são usadas para calcular o fluxo óptico, que descreve o movimento das informações visuais em uma sequência de imagens. Esse movimento é combinado com dados de uma IMU (Unidade de Medição Inercial ou *Inertial Measurement Unit*) para determinar o movimento da câmera (Odometria Visual-Inercial). O fluxo óptico é calculado pelo método de Gunnar Farnebäck com processamento em GPU para determinar o movimento da câmera de forma precisa e eficiente. Uma análise de quatro implementações de localização é realizada em simulação: (1) Odometria de Rodas, (2) Odometria Visual-Inercial, (3) fusão da Odometria de Rodas e IMU usando Filtro de Kalman Estendido (EKF), e (4) fusão usando EKF da Odometria Visual-Inercial, Odometria de Rodas e IMU. Os resultados obtidos em simulação ilustram que o sistema de localização multi-sensor proposto apresenta potencial para ser utilizado de forma efetiva para ajudar na localização de veículos de mineração em tarefas de campo.

Palavras-chave: Odometria Visual. Fluxo Óptico. Fusão Sensorial. Filtro de Kalman Estendido.

Abstract

This work presents a method for estimating odometry for vehicles in mining environments, where GNSS-based localization systems may experience failures or signal degradation. To mitigate these faults, the proposed method uses visual odometry from a camera directed towards the ground. Captured images are used to calculate the optical flow, which describes the movement of visual information in a sequence of images. This movement is combined with data from an IMU to determine the camera's movement (Visual-Inertial Odometry). The optical flow is calculated using Gunnar Farneback's GPU processing method to compute the camera motion accurately and efficiently. An analysis of four localization implementations is conducted: (1) Wheel Odometry, (2) Visual-Inertial Odometry, (3) fusion of Wheel Odometry and IMU using Extended Kalman Filter (EKF), and (4) fusion using EKF of Visual-Inertial Odometry, Wheel Odometry, and IMU. Results obtained in the simulation illustrate that the proposed multi-sensor localization system can be effectively used to assist in the localization of mining vehicles in field tasks.

Keywords: Visual Odometry. Optical Flow. Localization. Sensor Fusion. Extended Kalman Filter

Lista de Ilustrações

Figura 1 – Caminhão de mineração e vistas de três minas situadas em diferentes regiões mineradoras.	25
Figura 2 – Exemplos de diferentes plataformas que empregaram odometria visual para estimativa de movimento.	30
Figura 3 – Classificação das abordagens de odometria visual segundo o tipo e o número de câmeras empregados.	30
Figura 4 – Classificação das estratégias de odometria visual de acordo com a posição da câmera.	32
Figura 5 – Classificação dos métodos de odometria visual segundo a informação-chave extraída das imagens.	33
Figura 6 – Composição das rotações em torno dos eixos atuais.	39
Figura 7 – Composição de rotações em torno de eixos fixos.	39
Figura 8 – Representação do Modelo Cinemático Monociclo.	41
Figura 9 – Representação do Modelo Cinemático Diferencial.	42
Figura 10 – Representação do Modelo Cinemático <i>Skid-Steering</i>	43
Figura 11 – Representação do Modelo Cinemático <i>Ackermann</i>	44
Figura 12 – Modelo de projeção central mostrando o plano da imagem e <i>pixels</i> discretos. O raio de luz do ponto P cruza o plano da imagem no ponto p ao passar pelo centro da projeção.	45
Figura 13 – Ilustração do processo de odometria visual.	46
Figura 14 – Ilustração da captura de imagens em diferentes contextos: (a) relação entre distância e campo de visão e (b) sobreposição de quadros consecutivos.	48
Figura 15 – Ilustração de quatro situações distintas do fluxo óptico, cada uma resultante de diferentes movimentos da câmera em relação à cena observada.	49
Figura 16 – Ilustração do Processamento de Imagem usando CPU e GPU.	51
Figura 17 – Fluxo de Operações do Filtro de Kalman.	52
Figura 18 – Diagrama geral do sistema de odometria proposto.	55
Figura 19 – Ilustração das plataformas móveis no ambiente CoppeliaSim.	61
Figura 20 – Diagrama geral do algoritmo de odometria por fluxo óptico.	62
Figura 21 – Diagrama geral do algoritmo de odometria de rodas.	63

Figura 22 – Diagrama geral do algoritmo do Filtro de Kalman Estendido.	63
Figura 23 – EspeleoRobô e o Caminho da Pista <i>Red Bull Ring</i> no Simulador Coppeliasim.	68
Figura 24 – Posições estimadas pelos algoritmos implementados com <i>ground-truth</i> para a simulação do caminho da Pista <i>Red Bull Ring</i>	68
Figura 25 – Histogramas do erro de posição da simulação do EspeleoRobô na pista <i>Red Bull Ring</i> para: (a) Wheel, (b) VO+IMU, (c) EKF W+IMU e (d) EKF VO+W+IMU.	70
Figura 26 – Caminhão Caterpillar CAT793F e o Caminho de Teste no Simulador Coppeliasim.	71
Figura 27 – Posições estimadas pelos algoritmos implementados com <i>ground-truth</i> para a simulação do CAT793F no Caminho de Teste.	71
Figura 28 – Histogramas do erro de posição da simulação do CAT793F no Caminho de Teste para: (a) Wheel, (b) VO+IMU, (c) EKF W+IMU e (d) EKF VO+W+IMU.	73

Lista de Tabelas

Tabela 1 – Tabela de erros de posição referente à simulação do EspeleoRobô na Pista <i>Red Bull Ring</i>	69
Tabela 2 – Tabela dos valores do intervalo de confiança referente à simulação do EspeleoRobô na Pista <i>Red Bull Ring</i>	69
Tabela 3 – Erros de posição da simulação com o caminhão CAT793F no Caminho de Teste.	72
Tabela 4 – Intervalos de confiança para a simulação do CAT793F no Caminho de Teste.	72

Lista de Abreviaturas e Siglas

AKAZE	<i>Accelerated-KAZE</i>
EKF	<i>Extended Kalman Filter</i>
FAST	<i>Features from Accelerated Segment Test</i>
GNSS	<i>Global Navigation Satellite System</i>
GPS	<i>Global Positioning System</i>
GPU	<i>Graphics Processing Unit</i>
IBRAM	<i>Instituto Brasileiro de Mineração</i>
ICR	Centro de Rotação Instantâneo
IMU	<i>Inertial Measurement Unit</i>
ITV	Instituto Tecnológico Vale
LED	<i>Light Emitting Diode</i>
MAV	<i>Micro Aerial Vehicles</i>
ME	<i>Mean Error</i>
MAE	<i>Mean Absolute Error</i>
RGB-D	<i>Red, Green, Blue e Depth</i>
RMSE	<i>Root Mean Squared Error</i>
ROS	<i>Robot Operating System</i>
SUV	<i>Sport Utility Vehicle</i>
ToF	<i>Time of Flight</i>
UFMG	Universidade Federal de Minas Gerais
USV	<i>Unmanned Surface Vehicles</i>

VIO *Visual-Inertial Odometry*

VO *Visual Odometry*

WSD *Water-Shoreline Detection*

Lista de Símbolos

$\{\mathcal{O}\}$	Sistema de coordenadas fixo ao corpo rígido
$\hat{x}_{\mathcal{O}}$	Componente x do sistema de coordenadas $\{\mathcal{O}\}$
$\hat{y}_{\mathcal{O}}$	Componente y do sistema de coordenadas $\{\mathcal{O}\}$
$\hat{z}_{\mathcal{O}}$	Componente z do sistema de coordenadas $\{\mathcal{O}\}$
$\{\mathcal{I}\}$	Sistema de coordenadas inercial
$\hat{x}_{\mathcal{I}}$	Componente x do sistema de coordenadas $\{\mathcal{I}\}$
$\hat{y}_{\mathcal{I}}$	Componente y do sistema de coordenadas $\{\mathcal{I}\}$
$\hat{z}_{\mathcal{I}}$	Componente z do sistema de coordenadas $\{\mathcal{I}\}$
\mathbb{R}^3	Espaço de configuração \mathbb{R} tridimensional
$\mathbf{p}_{\mathcal{O}}^{\mathcal{I}}$	Vetor da posição do sistema de coordenadas $\{\mathcal{O}\}$ com respeito ao sistema inercial $\{\mathcal{I}\}$
$p_{\mathcal{O},x}^{\mathcal{I}}$	Componente x do vetor $\mathbf{p}_{\mathcal{O}}^{\mathcal{I}}$
$p_{\mathcal{O},y}^{\mathcal{I}}$	Componente y do vetor $\mathbf{p}_{\mathcal{O}}^{\mathcal{I}}$
$p_{\mathcal{O},z}^{\mathcal{I}}$	Componente z do vetor $\mathbf{p}_{\mathcal{O}}^{\mathcal{I}}$
$\mathbf{R}_{\mathcal{O}}^{\mathcal{I}}$	Matriz de rotação do sistema de coordenadas $\{\mathcal{O}\}$ para o sistema de coordenadas inercial $\{\mathcal{I}\}$
\mathbb{I}	Matriz identidade
ψ	Ângulo <i>yaw</i>
θ	Ângulo <i>pitch</i>
ϕ	Ângulo de <i>roll</i>
$\mathbf{T}_{\mathcal{O}}^{\mathcal{I}}$	Transformação homogênea do sistema de coordenadas $\{\mathcal{O}\}$ para o sistema inercial $\{\mathcal{I}\}$

$\mathbf{r}_{\mathcal{I}}$	Vetor de posição no sistema de coordenadas $\{\mathcal{I}\}$
$\mathbf{r}_{\mathcal{O}}$	Vetor de posição no sistema de coordenadas $\{\mathcal{O}\}$
$\mathbf{q}_{\mathcal{V}}^{\mathcal{I}}$	Vetor de posição do veículo com respeito ao sistema de coordenadas global
$\{\mathcal{V}\}$	Sistema de coordenadas do veículo
$q_{\mathcal{V}}^{\mathcal{I}}$	Vetor da pose do veículo com respeito ao sistema inercial $\{\mathcal{I}\}$
$x_{\mathcal{V}}^{\mathcal{I}}$	Componente de posição x do vetor $q_{\mathcal{V}}^{\mathcal{I}}$
$y_{\mathcal{V}}^{\mathcal{I}}$	Componente de posição y do vetor $q_{\mathcal{V}}^{\mathcal{I}}$
$\theta_{\mathcal{V}}$	Componente de orientação do vetor $q_{\mathcal{V}}^{\mathcal{I}}$
$R_{\mathcal{V}}^{\mathcal{I}}(\theta_{\mathcal{V}})$	Matriz de rotação ao redor do eixo z do veículo dada pelo ângulo $\theta_{\mathcal{V}}$
$\hat{x}_{\mathcal{V}}$	Componente x do sistema de coordenadas do veículo $\{\mathcal{V}\}$
$\hat{y}_{\mathcal{V}}$	Componente y do sistema de coordenadas do veículo $\{\mathcal{V}\}$
$v_{\mathcal{V}}$	Velocidade linear do veículo na direção x
$\omega_{\mathcal{V}}$	Velocidade angular do veículo em torno do eixo z
$\dot{x}_{\mathcal{V}}^{\mathcal{I}}$	Componente de velocidade x do veículo com respeito ao sistema inercial $\{\mathcal{I}\}$
$\dot{y}_{\mathcal{V}}^{\mathcal{I}}$	Componente de velocidade y do veículo com respeito ao sistema inercial $\{\mathcal{I}\}$
$\dot{\theta}_{\mathcal{V}}$	Componente de velocidade de rotação do veículo em torno do eixo z
t_k	Tempo contínuo amostrado no instante k
k	Índice de tempo discreto
Δt	Intervalo de amostragem
$x_{\mathcal{V},k}^{\mathcal{I}}$	Componente de posição x do veículo no instante k
$y_{\mathcal{V},k}^{\mathcal{I}}$	Componente de posição y do veículo no instante k
$\theta_{\mathcal{V},k}$	Componente de orientação do veículo no instante k
$x_{\mathcal{V},k-1}^{\mathcal{I}}$	Componente de posição x no instante $k - 1$
$y_{\mathcal{V},k-1}^{\mathcal{I}}$	Componente de posição y no instante $k - 1$

$\theta_{\mathcal{V},k-1}$	Componente de orientação no instante $k - 1$
$v_{\mathcal{V},k-1}$	Velocidade linear na direção x no instante $k - 1$
$\omega_{\mathcal{V},k-1}$	Velocidade angular em torno do eixo z no instante $k - 1$
r	Raio da roda do Modelo Diferencial
D	Distância entre as rodas do Modelo Diferencial
$\omega_{\mathcal{V},d}$	Velocidade angular da roda direita do Modelo Diferencial
$\omega_{\mathcal{V},e}$	Velocidade angular da roda esquerda do Modelo Diferencial
$v_{\mathcal{V},d}$	Velocidade das rodas do lado direito do modelo <i>skid-steering</i>
$v_{\mathcal{V},e}$	Velocidade das rodas do lado esquerdo do modelo <i>skid-steering</i>
a	Metade da distância entre os eixos do modelo <i>skid-steering</i>
b	Metade da distância entre as rodas do modelo <i>skid-steering</i>
L	Distância entre rodas dianteiras e traseiras do modelo <i>Ackermann</i>
W	Distância entre rodas no mesmo eixo do modelo <i>Ackermann</i>
δ	Ângulo de viragem central do modelo <i>Ackermann</i>
δ_i	Ângulo de viragem da roda dianteira interna do modelo <i>Ackermann</i>
δ_o	Ângulo de viragem da roda dianteira externa do modelo <i>Ackermann</i>
\mathbf{P}	Vetor de coordenadas de um ponto do mundo
X	Componente x do vetor \mathbf{P}
Y	Componente y do vetor \mathbf{P}
Z	Componente z do vetor \mathbf{P}
f'	Distância focal
$\{\mathcal{P}\}$	Sistema de coordenadas do plano da imagem
$\hat{x}_{\mathcal{P}}$	Componente x do sistema de coordenadas $\{\mathcal{P}\}$
$\hat{y}_{\mathcal{P}}$	Componente y do sistema de coordenadas $\{\mathcal{P}\}$
$\{\mathcal{C}\}$	Sistema de coordenadas da câmera
$\hat{x}_{\mathcal{C}}$	Componente x do sistema de coordenadas $\{\mathcal{C}\}$

\hat{y}_C	Componente y do sistema de coordenadas $\{C\}$
\hat{z}_C	Componente z do sistema de coordenadas $\{C\}$
\mathbf{p}	Vetor de coordenadas de um ponto no plano da imagem
x_P	Componente x do vetor \mathbf{p}
y_P	Componente y do vetor \mathbf{p}
\mathcal{W}	Largura da imagem em pixels
\mathcal{H}	Altura da imagem em pixels
u	Coordenada horizontal do ponto no plano da imagem
v	Coordenada vertical do ponto no plano da imagem
u_0	Coordenada horizontal do ponto principal de uma câmera
v_0	Coordenada vertical do ponto principal de uma câmera
ρ_w	Largura do pixel
ρ_h	Altura do pixel
$i_{0:n}$	Conjunto de imagens do instante de tempo 0 até n
$T_{1:n}$	Conjunto de transformações do instante de tempo 1 até n
$C_{0:n}$	Conjunto de posições da câmera do instante de tempo 0 até n
v_{max}	Velocidade máxima da câmera
α	Campo de visão da câmera
h	Altura da câmera acima do solo
f_s	Frequência de amostragem da câmera
$\frac{\partial I}{\partial x_P}$	Gradiente espacial da intensidade da imagem na direção horizontal
$\frac{\partial I}{\partial y_P}$	Gradiente espacial da intensidade da imagem na direção vertical
$\frac{\partial I}{\partial t}$	Derivada temporal da intensidade
$\frac{dx_P}{dt}$	Velocidade do fluxo óptico na direção horizontal
$\frac{dy_P}{dt}$	Velocidade do fluxo óptico na direção vertical
\dot{u}	Velocidade de fluxo óptico horizontal

\dot{v}	Velocidade de fluxo óptico vertical
\bar{u}	Coordenada normalizada do plano da imagem, em relação ao ponto principal u_0
\bar{v}	Coordenada normalizada do plano da imagem, em relação ao ponto principal v_0
$Z_{\mathcal{P}}$	Profundidade do ponto da imagem
$\mathbf{v}_{\mathcal{C}}$	Vetor de velocidade linear da câmera
$v_{\mathcal{C},x}$	Componente de velocidade linear x do vetor $\mathbf{v}_{\mathcal{C}}$
$v_{\mathcal{C},y}$	Componente de velocidade linear y do vetor $\mathbf{v}_{\mathcal{C}}$
$v_{\mathcal{C},z}$	Componente de velocidade linear z do vetor $\mathbf{v}_{\mathcal{C}}$
$\boldsymbol{\omega}_{\mathcal{C}}$	Vetor de velocidade angular da câmera
$\omega_{\mathcal{C},x}$	Componente de velocidade angular x do vetor $\boldsymbol{\omega}_{\mathcal{C}}$
$\omega_{\mathcal{C},y}$	Componente de velocidade angular y do vetor $\boldsymbol{\omega}_{\mathcal{C}}$
$\omega_{\mathcal{C},z}$	Componente de velocidade angular z do vetor $\boldsymbol{\omega}_{\mathcal{C}}$
$T_{overhead}$	Tempo de sobrecarga da CPU e GPU
T_{upload}	Tempo para enviar dados da CPU para GPU
$T_{download}$	Tempo para receber dados da GPU para CPU
T_{CPU}	Tempo de execução na CPU
T_{GPU}	Tempo de execução na GPU
\mathbf{P}	Posição de referência fornecida pelo simulador (<i>ground-truth</i>)
$\hat{\mathbf{P}}$	Posição estimada pela odometria
\mathbf{Q}	Matriz de covariância do modelo de processo
\mathbf{R}	Matriz de covariância das medições (ruído dos sensores)
e_i	Erro euclidiano no instante i entre a pose estimada e a de referência
s	Desvio padrão amostral dos erros
\bar{e}	Erro médio
σ_k	Intervalo de confiança baseado em k desvios padrão da média

X_0	Estado inicial estimado do Filtro de Kalman
P_0	Matriz de covariância do erro inicial
X_{k-1}	Vetor de estado estimado no instante $k - 1$
P_{k-1}	Matriz de covariância do erro no instante $k - 1$
$\hat{X}_{k k-1}$	Previsão do estado no instante k a partir do instante $k - 1$
$\hat{P}_{k k-1}$	Previsão da covariância do erro no instante k a partir do instante $k - 1$
A	Matriz de transição de estados
B	Matriz de controle
u_k	Vetor de entrada de controle no instante k
w_k	Ruído de processo no instante k
Q_k	Matriz de covariância do ruído de processo
K	Ganho de Kalman
H	Matriz de observação
R	Matriz de covariância do ruído de medição
Y_k	Vetor de medições no instante k
C	Matriz de mapeamento do estado real para a medição
X_{kM}	Estado usado para gerar a medição
Z	Ruído de medição
$g(\cdot)$	Função não linear de transição de estados
$h(\cdot)$	Função não linear de observação
G_k	Jacobiana de $g(\cdot)$
H_k	Jacobiana de $h(\cdot)$

Sumário

1	Introdução	22
1.1	Motivação	23
1.2	Objetivos	25
1.3	Contribuições	26
1.4	Organização do texto	27
2	Revisão Bibliográfica	29
2.1	Odometria Visual	29
2.1.1	Tipo/Número de Câmeras	30
2.1.2	Posição da Câmera	31
2.1.3	Informação-Chave	33
2.2	Fusão Sensorial para Estimação de Movimento	34
3	Fundamentos Teóricos	37
3.1	Cinemática de Corpos Rígidos	37
3.1.1	Posição e Translação	37
3.1.2	Orientação e Rotação	38
3.1.3	Transformação Homogênea	39
3.1.4	Transformação de Velocidades em Corpos Rígidos	40
3.2	Modelos Cinemáticos de Veículos com Rodas	40
3.3	Odometria Visual	44
3.3.1	Formação de Imagem	44
3.3.2	Estimativa Incremental de Poses a Partir de Dados Visuais	47
3.3.3	Fluxo Óptico	48
3.4	Processamento de Imagem usando GPU	50
3.5	Fusão Sensorial	51
3.5.1	Filtro de Kalman	51
3.5.2	Filtro de Kalman Estendido (EKF)	53
4	Metodologia	55
4.1	Localização com Odometria Visual por Fluxo Óptico	56
4.2	Localização com Odometria de Rodas	57
4.3	Localização com Filtro de Kalman Estendido	57
5	Arcabouço Experimental	60
5.1	Plataformas Móveis	60
5.2	Implementações de Algoritmos em Pacotes do ROS	61
5.2.1	Implementação da Odometria Visual por Fluxo Óptico	62
5.2.2	Implementação da Odometria de Rodas	62
5.2.3	Implementação do Filtro de Kalman Estendido	63

5.3	Controle de Navegação e Definição de Velocidades	64
6	Resultados e Discussões	65
6.1	Abordagens de Odometria Consideradas	65
6.2	Técnicas de Avaliação Empregadas	66
6.2.1	Métricas de Erro Euclidiano	66
6.2.2	Análise Estatística dos Erros	67
6.3	Simulação de Ajuste de Covariâncias	67
6.4	Simulação de Validação da Metodologia Proposta	67
6.4.1	Resultados da Simulação do EspeleoRobô	68
6.4.2	Resultados da Simulação do Caminhão Caterpillar CAT793F	71
7	Conclusão	75
7.1	Trabalhos Futuros	76
	Referências	77

Capítulo 1

Introdução

O desenvolvimento de veículos autônomos tem despertado crescente interesse tanto da indústria quanto da comunidade científica, devido aos potenciais benefícios que a navegação autônoma pode proporcionar aos sistemas de transporte. A condução automatizada e a otimização de rotas permitem não apenas a melhoria do fluxo de tráfego, como também a redução do consumo de combustível. Além disso, a segurança nas vias tende a ser aperfeiçoada, uma vez que a automação reduz a ocorrência de falhas humanas e possibilita tempos de reação mais rápidos do que os de um condutor humano [Agostinho et al., 2022]. No entanto, a implementação eficaz dessas tecnologias exige estudos aprofundados, pois diversas variáveis precisam ser controladas para evitar potenciais riscos operacionais.

A estimativa contínua da própria localização é uma função fundamental para qualquer sistema autônomo que execute tarefas de navegação. O Sistema Global de Navegação por Satélite (GNSS), que engloba constelações como o GPS (Estados Unidos), Galileo (União Europeia), GLONASS (Rússia) e BeiDou (China), é amplamente utilizado para esse fim, principalmente por oferecer uma boa relação entre custo e benefício. No entanto, a precisão proporcionada por essa tecnologia ainda não atende plenamente aos requisitos de diversas aplicações em veículos autônomos [Mohamed et al., 2019]. Por exemplo, receptores GPS convencionais apresentam erros da ordem de metros, o que é inadequado para tarefas que exigem precisão na escala de centímetros, como estacionamento autônomo [Aqel et al., 2016a]. Embora existam versões mais precisas, como o GPS Diferencial (DGPS) e o GPS Cinemático em Tempo Real (RTK), esses sistemas possuem custos mais elevados. Além disso, os sistemas por satélite estão sujeitos a falhas, atrasos e bloqueios de sinal, especialmente em locais com visibilidade limitada do céu, como áreas urbanas densas [Moreau et al., 2017], espaços fechados [Lin and Zhan, 2023] e ambientes de mineração [Szrek et al., 2020].

Uma alternativa para contornar as limitações dos sistemas de posicionamento por satélite é o uso de sensores que operam de forma independente, ou seja, que não dependem

de infraestrutura externa para fornecer estimativas de localização. Esses dispositivos utilizam medições próprias para estimar o deslocamento ao longo do tempo. Contudo, a precisão dessas estimativas depende das características de cada sensor e das condições do ambiente em que o sistema está inserido. Em cenários como os encontrados na mineração, essas limitações se tornam particularmente evidentes. Fatores como extensas áreas abertas, baixa complexidade estrutural e escassez de elementos visuais distintos podem comprometer o desempenho de sensores como radares, LiDARs e câmeras [Abu-Alrub and Rawashdeh, 2024, Domingues et al., 2022].

A fusão sensorial é uma abordagem amplamente utilizada para melhorar a precisão na estimativa de localização por meio da combinação dos dados fornecidos por diferentes sensores. Essa integração permite que os dados sejam utilizados de maneira complementar, possibilitando aproveitar os pontos fortes de cada sensor e minimizar suas limitações individuais. Em aplicações de navegação autônoma, especialmente em ambientes desafiadores comuns na mineração, essa estratégia pode contribuir para aumentar a confiabilidade, a precisão e a capacidade do sistema de lidar com falhas [Alkendi et al., 2021].

A solução proposta utiliza o fluxo óptico, uma técnica de visão computacional que representa o movimento aparente dos pixels entre quadros consecutivos, extraído de imagens capturadas por uma câmera voltada para o solo para estimar o movimento da câmera. Essa estimativa é então combinada com dados de uma unidade de medição inercial (IMU) e da odometria de rodas por meio de um filtro de Kalman dedicado, com o objetivo de estimar a localização de um veículo autônomo mesmo na ausência de sinal de satélite. O processamento do fluxo óptico e o cálculo do movimento da câmera são realizados com o suporte de uma unidade de processamento gráfico (GPU), permitindo processamento paralelo e reduzindo o tempo de execução. A validação do método foi conduzida por meio de simulações em cenários que reproduzem as condições típicas de terrenos de mineração, resultando em dois estudos publicados: um utilizando o EspeleoRobô [Gomes et al., 2024a] e outro com o caminhão Caterpillar CAT793F [Gomes et al., 2024b].

1.1 Motivação

A mineração é uma das bases da economia e tem papel central no fornecimento de recursos minerais que sustentam o desenvolvimento tecnológico e industrial em todo o mundo. Além de sua relevância como fornecedora de matérias-primas, o setor de mineração contribui significativamente para a geração de empregos e para o desenvolvimento socioeconômico. De acordo com dados do Instituto Brasileiro de Mineração (IBRAM), o faturamento da indústria mineral brasileira atingiu R\$ 270,8 bilhões em 2024, representando um aumento de 9,1% em relação ao ano anterior [Instituto Brasileiro de Mineração (IBRAM), 2024], o que reforça sua importância econômica e sua contribuição para o

crescimento do país.

O ambiente de trabalho na mineração pode envolver condições que comprometem a segurança e a saúde dos profissionais. Um exemplo é a instabilidade do terreno, que pode ser agravada pelas constantes atividades de escavação e favorecer a ocorrência de deslizamentos de terra em áreas de operação. Do ponto de vista da saúde, a exposição prolongada à poeira, gases tóxicos e substâncias químicas pode contribuir para o desenvolvimento de doenças respiratórias, dermatológicas e outras complicações de longo prazo [Kaur, 2024]. Nesse contexto, minimizar o tempo de exposição dos trabalhadores em áreas de risco representa uma medida relevante que contribui para a segurança no trabalho.

A substituição da mão de obra humana por sistemas tecnológicos tem se intensificado em diversas áreas, especialmente naquelas que exigem níveis de desempenho superiores às capacidades humanas. Tarefas que impõem grande esforço físico ou cognitivo são propensas à automação, como é o caso do transporte rodoviário de cargas. Nessa atividade, o condutor precisa manter atenção constante por longos períodos, enfrentando rodovias com condições imprevisíveis e, muitas vezes, sob influência de fatores climáticos adversos. Essa exigência prolongada pode levar à fadiga, aumentando o risco de erros na condução e comprometendo não apenas a segurança do motorista, mas também a de outros usuários da via.

A adoção de tecnologias autônomas na mineração pode promover transformações significativas nos processos produtivos. Esses sistemas permitem o planejamento de trajetórias mais seguras e otimizadas, com base em informações geoespaciais, condições do terreno e restrições operacionais, o que contribui para o aumento de produtividade [Tian et al., 2021]. Além disso, a condução automatizada contribui para a eficiência operacional ao garantir padrões de aceleração e frenagem mais suaves e consistentes do que os realizados por motoristas humanos, o que se reflete em menor consumo de combustível e maior eficiência energética ao longo das rotas [Chen et al., 2019].

O uso de caminhões autônomos, como pode ser visto na Figura 1a, já é uma realidade na indústria da mineração, com empresas como a Vale empregando essas tecnologias em operações de larga escala [Global Mining Review, 2021]. No entanto, os caminhões autônomos atuais dependem fortemente de sistemas de posicionamento via satélite, o que representa uma limitação significativa, sobretudo em ambientes onde o sinal pode ser degradado ou interrompido por relevo ou vegetação densa, como mostrado nas Figuras 1b e 1c. Diante desse cenário, torna-se fundamental o desenvolvimento de métodos alternativos ou complementares de localização para garantir que os veículos mantenham operação contínua e segura, mesmo na ausência ou instabilidade do sinal de satélite.



(a) Caminhão de mineração CAT793F da fabricante Caterpillar. [Teles, 2025]



(b) Mina Super Pit em Kalgoorlie. [Jamasmie, 2019]



(c) Mina Gongo Soco em Barão de Cocais. [Rosa, 2025]



(d) Mina de Fortescue Metals em Pilbara. [Bulla, 2021]

Figura 1 – Caminhão de mineração e vistas de três minas situadas em diferentes regiões mineradoras.

1.2 Objetivos

Esta dissertação tem como objetivo propor um método de estimativa de localização para veículos autônomos em ambientes de mineração, capaz de operar sem GNSS quando o sinal estiver ausente ou comprometido. Esses ambientes apresentam desafios relevantes à estimativa de posição com base em visão, principalmente devido à presença de grandes áreas com baixa variação visual e poucas características distintas, o que dificulta a extração de informações úteis a partir das imagens capturadas por câmera, conforme ilustrado na Figura 1d. Dessa forma, os objetivos específicos desta dissertação são:

- Desenvolver um método de odometria visual com uma câmera voltada para o solo;
- Integrar dados de odometria visual, odometria de rodas e IMU por meio de um Filtro de Kalman dedicado;
- Implementar diferentes abordagens de odometria e comparar seus desempenhos com o método proposto;

- Validar o método proposto em modelos virtuais do EspeleoRobô e do caminhão Caterpillar CAT793F;
- Avaliar o desempenho da abordagem proposta por meio de métricas quantitativas, histogramas e intervalos de confiança.

1.3 Contribuições

As contribuições desta pesquisa estão voltadas à estimativa de odometria para veículos autônomos em ambientes de mineração, com foco na fusão sensorial e no uso de câmeras voltadas para o solo.

- Desenvolvimento de um método de odometria visual-inercial para veículos de mineração, capaz de operar com uma câmera rigidamente acoplada ao veículo e voltada para o solo, utilizando a técnica de fluxo óptico;
- Implementação do processamento do fluxo óptico em GPU, por meio do OpenCV 4.9 integrado ao CUDA, resultando em uma redução do tempo de processamento por quadro em cerca de 15 vezes, garantindo maior desempenho computacional e viabilizando aplicações em tempo real;
- Desenvolvimento de um pacote ROS para locomoção do modelo virtual do caminhão Caterpillar CAT793F, implementando a geometria de direção Ackermann;
- Implementação de um Filtro de Kalman Estendido dedicado para a fusão dos dados provenientes da odometria visual por fluxo óptico, da odometria de rodas e da unidade de medição inercial (IMU);
- Simulação do método proposto em ambiente virtual usando os modelos do EspeleoRobô e do caminhão de CAT793F, em cenários que reproduzem as condições de terreno típicas de mineração.

O sistema desenvolvido permite estimar a odometria do veículo quando há falha ou ausência do sinal de GPS. A combinação de sensores e o uso de processamento visual eficiente tornam a solução adequada para navegação autônoma em ambientes desafiadores. A pesquisa teve dois estudos desenvolvidos e publicados em periódicos científicos

- [Gomes et al., 2024a] G. Gomes, G. Cruz, J. Domingues, G. Pessin, A. Neto, and G. M. Freitas. Odometria Visual por Fluxo Óptico Integrada a um Sistema de Localização Multi-sensor para Veículos de Mineração. In Anais do Congresso Brasileiro de Automática 2024, Oct. Neste artigo, o sistema de odometria visual proposto foi

implementado em um modelo virtual do EspeleoRobô, caracterizado por locomoção do tipo *skid-steering*;

- [Gomes et al., 2024b] G. Gomes, G. Cruz, J. Domingues, G. Pessin, A. Neto, and G. M. Freitas. Visual Odometry through Optical Flow Integrated with a Multi-sensor Localization System Applied to Mining Vehicles. In 2024 Latin American Robotics Symposium (LARS), page 1–6. IEEE, Nov. 2024. Neste artigo, o sistema de odometria visual proposto foi implementado em um modelo virtual do caminhão de mineração Caterpillar CAT793F, caracterizado por locomoção do tipo *Ackermann*.

Além das publicações vinculadas a este mestrado, consta ainda um artigo publicado não diretamente associado aos objetivos desta pesquisa

- [Soares et al., 2025] L. G. Soares, N. H. C. Rocha, G. E. Soares, G. P. Cruz Júnior, G. Gomes, G. M. de Oliveira, A. C. Vangasse, H. I. Azpúrua, A. L. M. Cid, and G. M. Freitas. Enhancing the autonomous mapping and navigation unit for service robots. In 2025 Brazilian Conference on Robotics (CROS), volume 1, pages 1–6, 2025.

O mestrado também resultou em contribuições técnicas, documentadas nos seguintes relatórios

- [Gomes et al., 2023] G. Gomes, G. P. Cruz Junior, J. D. Domingues, and G. M. Freitas. Tutorial Sobre a Utilização Das Câmeras Intel Realsense D435i e T265 com o RTAB-Map no ROS. Relatório Técnico, 2023.
- [Freitas et al., 2023] G. M. Freitas, G. P. Cruz Junior, A. C. Vangasse, G. Gomes, S. P. Campos, I. F. S. Amaral, A. L. M. Cid, M. C. Delunardo, N. C. P. S. Thiago Neto, R. S. Pereira Junior, L. G. D. Barros, J. D. Domingues, and G. Pessin. Primeiro Relatório Parcial do Projeto ‘Localização, Mapeamento e Comunicação em Cavas de Minas e Galerias de Barragens de Grande Extensão’. Relatório Técnico, 2023.
- [Freitas et al., 2024] G. M. Freitas, G. P. Cruz Junior, A. C. Vangasse, G. Gomes, S. P. Campos, J. F. R. Baiao, F. A. T. Montano, N. H. C. Rocha, A. L. M. Cid, M. C. Delunardo, N. C. P. S. Thiago Neto, R. S. Pereira Junior, Y. M. Bepaloff, L. G. D. Barros, and G. Pessin. Segundo Relatório Parcial do Projeto ‘Localização, Mapeamento e Comunicação em Cavas de Minas e Galerias de Barragens de Grande Extensão’. Relatório Técnico, 2024.

1.4 Organização do texto

Esta dissertação está organizada em 7 Capítulos. No Capítulo 2 é apresentada uma revisão bibliográfica sobre assuntos relacionados a odometria visual e fusão sensorial.

Os fundamentos teóricos necessários para compreensão de odometria visual por fluxo óptico e fusão sensorial são discutidos no Capítulo 3. O Capítulo 4 descreve a metodologia usada para estimação da odometria visual por fluxo óptico e como é usada para fusão com informações provenientes de outros sensores. O arcabouço experimental e o ambiente usado para simulação são apresentados no Capítulo 5. Os resultados e as comparações com outras estratégias de odometria são mostrados no Capítulo 6. No fim, as conclusões e trabalhos futuros são destacados no Capítulo 7.

Capítulo 2

Revisão Bibliográfica

Este Capítulo apresenta uma revisão bibliográfica dos principais trabalhos relacionados à estimação do movimento de veículos por meio de odometria visual e também à incorporação de sensores auxiliares para aprimorar essa estimativa. Nesse contexto, a primeira parte analisa diferentes formas para a obtenção do movimento do agente a partir da informação visual adquirida por uma ou mais câmeras, considerando as variações nos sistemas de aquisição, a configuração da câmera quanto à sua orientação e os tipos de dados extraídos das imagens. Em seguida, são discutidas técnicas de fusão sensorial que integram dados visuais com informações provenientes de outros sensores, com o objetivo de aumentar a precisão da odometria visual.

2.1 Odometria Visual

A odometria visual (*Visual Odometry* – VO) é uma técnica utilizada para estimar incrementalmente o movimento de um agente a partir da análise de sequências de imagens capturadas por uma ou mais câmeras embarcadas [Scaramuzza and Fraundorfer, 2011]. Essa estimativa é obtida com base nas variações visuais entre quadros consecutivos, permitindo calcular as componentes de translação e rotação do sistema a partir das mudanças observadas na cena [Yousif et al., 2015]. Uma das principais vantagens da VO é sua independência em relação ao tipo de veículo, uma vez que a estimativa do movimento depende exclusivamente da percepção visual do ambiente. Essa característica permite sua aplicação em diferentes tipos de plataformas, como robôs móveis, veículos autônomos e aeronaves, como os mostrados na Figura 2. Conforme discutido em [Mohamed et al., 2019], os métodos de odometria visual podem ser classificados de acordo com diferentes critérios. Entre os principais destacam-se o tipo ou número de câmeras empregadas, o posicionamento da câmera em relação à plataforma e a informação-chave (*key information*) utilizada na estimativa do movimento.



Figura 2 – Exemplos de diferentes plataformas que empregaram odometria visual para estimativa de movimento.

2.1.1 Tipo/Número de Câmeras

O tipo e o número de câmeras empregados para aquisição de imagens constituem um dos critérios para classificar os sistemas de odometria visual, pois influenciam a quantidade e a qualidade das informações disponíveis para a estimativa do movimento. Na literatura, destacam-se as câmeras apresentadas na Figura 3.

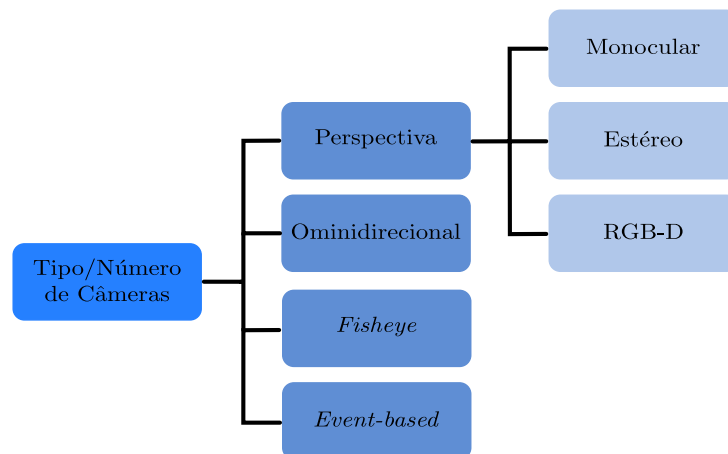


Figura 3 – Classificação das abordagens de odometria visual segundo o tipo e o número de câmeras empregados.

Devido à sua simplicidade e baixo custo, a câmera monocular é frequentemente adotada em sistemas de odometria visual. No trabalho de Engel et al. [2013], os autores utilizaram esse tipo de câmera para determinar em tempo real a pose de uma unidade

operada manualmente. No entanto, em algumas situações, o acesso à informação de distância entre a câmera e os elementos do ambiente pode contribuir para aumentar a precisão da estimativa de movimento. Em [Song et al., 2010], foram usadas as medidas de distância fornecidas por uma câmera estéreo como informação adicional para estimar a velocidade de um robô móvel em terrenos com variação de relevo.

Outra alternativa para obtenção da distância dos componentes da cena é o uso de câmeras RGB-D, que fornecem simultaneamente informações de cor e distância por meio de sensores dedicados, sem depender de correspondência estéreo. A partir das imagens do *dataset TUM RGB-D*, obtidas em ambientes internos por uma câmera *RGB-D Microsoft Kinect*, Steinbrucker et al. [2011] investigaram uma forma de determinar o deslocamento do dispositivo ao longo do tempo.

As limitações do campo de visão de câmeras convencionais têm motivado o uso de sensores visuais com cobertura mais ampla, capazes de registrar uma área maior do ambiente em cada captura. Nesse contexto, os experimentos conduzidos por Valiente García et al. [2012] usaram um robô Pioneer 3-AT equipado com uma câmera omnidirecional DMK 21BF04, cujas imagens panorâmicas foram empregadas para obter a movimentação do robô em ambientes internos. De modo semelhante, Moreau et al. [2017] propuseram um método de localização que utiliza uma câmera olho de peixe direcionada para cima, projetado para lidar com a ausência de sinal GPS em cenários urbanos densos.

Outra linha de VO envolve o uso de câmeras de eventos, que se destacam por sua alta taxa de atualização temporal, baixo consumo de energia e funcionamento eficaz em ambientes com condições adversas de iluminação. Nesse sentido, Hadviger et al. [2021] propuseram uma técnica de odometria visual que utiliza câmeras de eventos DAVIS346 organizadas em um par estéreo. A proposta foi avaliada tanto em um *drone* operando em ambientes internos quanto em um veículo automotivo em cenários urbanos.

Apesar de não serem amplamente adotadas em sistemas de odometria visual, câmeras térmicas são empregadas em aplicações específicas onde sensores ópticos convencionais enfrentam limitações. Em [Khattak et al., 2019], foi proposta uma abordagem que explora dados radiométricos diretos de uma câmera térmica para navegação de robôs aéreos em ambientes de minas subterrâneas e túneis, que não possuem acesso a GPS e são visualmente degradados devido a condições de escuridão e presença de elementos que comprometem a visibilidade, como poeira, neblina e fumaça.

2.1.2 Posição da Câmera

A classificação da odometria visual também pode considerar a posição e a orientação da câmera em relação à plataforma móvel, uma vez que afeta a captura das informações visuais. As principais categorias de posicionamento são apresentadas na Figura 4.

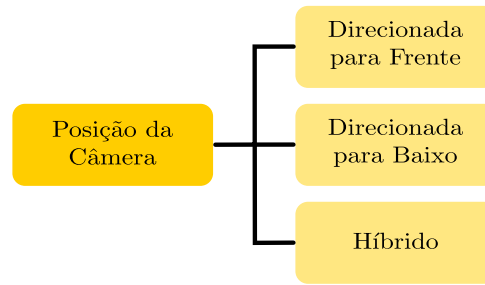


Figura 4 – Classificação das estratégias de odometria visual de acordo com a posição da câmera.

Sistemas de odometria visual convencionais frequentemente adotam câmeras voltadas para frente, uma vez que essa orientação permite capturar um volume maior de informações visuais do ambiente à medida que a plataforma se move. Essa configuração foi adotada por [Sun et al. \[2018\]](#), que implementaram um sistema de odometria visual com câmeras estéreo embarcado em um microveículo aéreo (MAV) denominado FALCON. O sistema demonstrou ser capaz de operar em tempo real, mesmo durante voos rápidos.

Câmeras voltadas para o chão oferecem uma visão mais controlada, com menor influência de variações bruscas de iluminação e profundidade. Essa configuração foi explorada por [Killpack et al. \[2010\]](#), que posicionaram uma câmera na parte inferior de um robô com base omnidirecional, acompanhada de um anel de LED para garantir condições de iluminação estáveis. De forma semelhante, [Dille et al. \[2010\]](#) usaram um sensor óptico composto por sensores de mouse comerciais voltados para o solo na estimação da velocidade do rover lunar Scarab. Com uma proposta alternativa, [Rajesh and Manivannan \[2024\]](#) usaram uma câmera inclinada em 30° em relação ao solo, integrada a um sistema com processamento acelerado por GPU para estimar o deslocamento do veículo.

Alguns sistemas de odometria visual adotam configurações híbridas, combinando câmeras voltadas para diferentes direções a fim de ampliar o campo de visão e melhorar a estimativa de movimento. [Piyathilaka and Munasinghe \[2010\]](#) exploraram essa abordagem em um veículo terrestre com condução por *skid-steering*, aplicando simultaneamente câmeras orientadas para frente e para o solo. Outras estratégias recorrem a diferentes combinações de orientação. Um exemplo é o sistema proposto por [Liu et al. \[2021\]](#), que faz uso de imagens adquiridas nas direções frontal, esquerda e direita.

[Zou et al. \[2022\]](#) propuseram o Water-VO, uma solução de odometria visual projetada para USVs navegarem em ambientes fluviais com ausência GPS. Nessa estratégia, a câmera é instalada lateralmente, com o campo de visão voltado para a margem, e apenas os elementos visuais das regiões não-aquáticas são utilizados na estimação do movimento. O sistema baseia-se em um algoritmo prévio de detecção da linha d'água (WSD), desenvolvido em um trabalho anterior [[Zou et al., 2020](#)], capaz de segmentar imagens em áreas aquáticas e não-aquáticas.

2.1.3 Informação-Chave

No contexto da informação-chave para estimativa do movimento, as técnicas de odometria visual podem ser agrupadas em três categorias principais: métodos diretos, métodos indiretos e métodos híbridos. Os métodos diretos, também conhecidos como baseados em aparência (*appearance-based*), utilizam como principal fonte de informação as intensidades brutas dos *pixels* [He et al., 2019, Scaramuzza and Fraundorfer, 2011]. Em contraste, os métodos indiretos, ou baseados em características (*feature-based*), extraem e rastreiam características visuais entre quadros consecutivos, como pontos ou linhas [Aquel et al., 2016b, Scaramuzza and Fraundorfer, 2011]. Já os métodos híbridos combinam o uso das intensidades de *pixels* e das características visuais, buscando integrar as vantagens de ambas as estratégias. A Figura 5 resume as categorias, incluindo as subdivisões mais comuns.

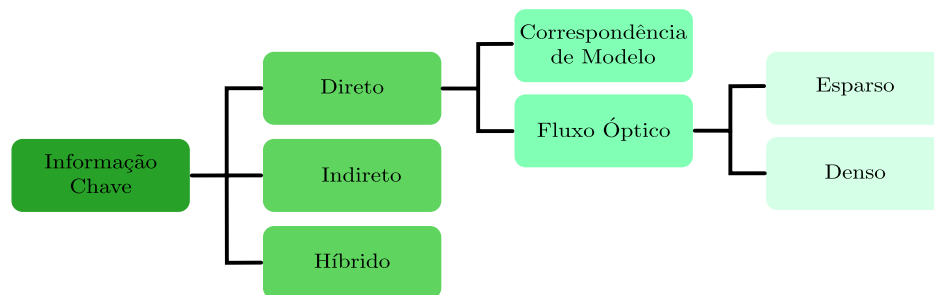


Figura 5 – Classificação dos métodos de odometria visual segundo a informação-chave extraída das imagens.

A correspondência de modelo (*template matching*) é um método direto que busca em um novo quadro de imagem a região que melhor corresponde a um padrão previamente selecionado no quadro anterior. Essa correspondência é baseada em medidas de similaridade entre pequenas janelas de *pixels*. No trabalho de Nourani-Vatani et al. [2009], foi empregada para calcular o deslocamento de uma empilhadeira industrial e de um Toyota Prado SUV. Esse trabalho foi posteriormente estendido em [Nourani-Vatani and Borges, 2011], com o uso de múltiplas janelas e um filtro preditivo para restringir a área de busca. De forma semelhante, Yu et al. [2011] usaram uma correspondência rotacionada de modelo, na qual um trecho da imagem é comparado em diferentes ângulos para determinar tanto a translação quanto a orientação, sendo esse processo validado em um robô *Pioneer 3-AT*.

O fluxo óptico (*optical flow*) é uma abordagem direta que estima o movimento da câmera com base nas variações de intensidade dos *pixels* entre quadros consecutivos, por meio da minimização do erro fotométrico. Essa técnica pode ser implementada de forma esparsa ou densa. Na forma esparsa, o cálculo do deslocamento é realizado apenas em pontos selecionados da imagem. Um exemplo de VO com fluxo óptico esparsa é apresentado em [Campbell et al., 2005], que usa o método de *Lucas-Kanade* em um sistema embarcado

em pequenos robôs móveis. A estratégia adotada consistiu em dividir os vetores de fluxo em regiões de solo e céu, sendo a rotação obtida com base nos vetores do céu e a translação a partir dos vetores do solo.

Em contraste com a estimativa esparsa, o fluxo óptico denso explora as variações de intensidade de todos os *pixels* da imagem, ou de uma parte significativa dela. Essa formulação é adequada para regiões com baixa textura, mas exige maior capacidade computacional e apresenta maior sensibilidade a variações de iluminação. Em [Domingues et al., 2022], foi utilizado o algoritmo de fluxo óptico denso de *Farnebäck* com o objetivo de determinar o deslocamento de um robô móvel em ambientes de mineração simulados e reais.

Nos métodos indiretos, o movimento da câmera é calculado pela correspondência entre características visuais extraídas das imagens, como pontos de interesse, cantos ou linhas. Essa abordagem identifica elementos distintos da cena que podem ser reconhecidos em quadros seguintes, permitindo medir a variação de pose. Lee et al. [2024] apresentaram um método baseado em características extraídas do solo com o detector *Accelerated-KAZE* (AKAZE), usando uma câmera monocular voltada para o chão, um anel de LED para iluminação e um sensor ToF para obter a distância da câmera ao chão.

Apesar das particularidades de cada estratégia, métodos híbridos têm sido investigados para integrar os dados brutos de intensidade das imagens dos métodos diretos com as correspondências de pontos-chave dos métodos indiretos. Nesse sentido, Al-Hadithi et al. [2024] propõe uma técnica que combina uma etapa indireta para extração e rastreamento de características visuais entre quadros consecutivos, com uma etapa direta de refinamento por *template matching*. Outro exemplo é o método proposto por Naixin et al. [2019], que combina a extração de cantos pelo algoritmo FAST, típica de métodos indiretos, com o alinhamento de imagem por erro fotométrico, característico dos processos diretos.

2.2 Fusão Sensorial para Estimação de Movimento

A fusão sensorial é uma estratégia utilizada para integrar informações de diferentes sensores, com o objetivo de melhorar a precisão e a confiabilidade dos resultados. No contexto da estimativa de pose, uma das abordagens mais comuns é a odometria visual-inercial, que combina imagens de câmera com medições de IMU para calcular a posição e orientação de um sistema. De acordo com Mohamed et al. [2019], essa integração pode ser classificada de acordo com o momento da combinação dos dados, como *loosely* ou *tightly coupled*, e pelo método de fusão adotado, seja por filtros ou por otimização.

A combinação entre câmera e IMU pode ocorrer de forma *loosely-coupled*, com os dados dos sensores processados separadamente e integrados posteriormente, ou *tightly-coupled*, com a integração direta dos dados brutos de ambos os sensores durante a estimação.

Zeng et al. [2021] adotam uma abordagem *loosely-coupled*, na qual a translação é estimada visualmente e combinada ao ângulo de guinada fornecido pela IMU. Já em [Zhang and Singh, 2015], os autores propuseram um método para aeronaves no qual a orientação é fornecida por um INS e também utilizada diretamente no processo visual de estimação da translação.

A fusão entre dados visuais e inerciais pode ser realizada por diferentes estratégias, sendo as mais comuns as que utilizam filtros e as que se baseiam em otimização. Em [Bloesch et al., 2015], os autores apresentaram um algoritmo de odometria visual-inercial para UAV baseado em filtro de Kalman estendido, no qual os erros de intensidade de *patches* de imagem são usados como termo de inovação na atualização do filtro. Por outro lado, Usenko et al. [2016] propuseram um método visual-inercial para câmeras estéreo, baseado em mapas de profundidade semi-densos obtidos por estereoscopia estática e temporal. A fusão com dados da IMU é feita por otimização conjunta de erros fotométricos e inerciais. Essa abordagem foi validada em experimentos com *drones* e veículos terrestres. Além da VIO, diferentes abordagens exploram a fusão da odometria visual com outros sensores adicionais. Yuan et al. [2023] desenvolveram o SDV-LOAM, um sistema de odometria visual-LiDAR baseado em câmera monocular e LiDAR 3D. Por sua vez, Cheng et al. [2024] apresentaram o MS-VRO, um sistema de odometria visual baseado na fusão de câmera monocular e radar mmWave. E Zhu et al. [2021] elaboraram um sistema de odometria visual-inercial com encoders de roda que emprega fusão adaptativa por filtro de Kalman.

Uma abordagem distinta das técnicas convencionais de odometria visual é o uso de redes neurais profundas treinadas para inferir diretamente o movimento da câmera. Em vez de depender de correspondência de características, cálculo explícito de fluxo óptico ou otimização de poses, essas redes aprendem a mapear imagens ou sequências de imagens diretamente para os parâmetros de movimento. [Gilles and Ibrahimovic, 2021] propuseram dois modelos de redes convolucionais treinados de forma não supervisionada para estimar incrementalmente o movimento de um robô com a câmera voltada para o chão.

Fusão da Odometria Visual-Inercial com Odometria de Rodas e IMU Realizada Nesta Dissertação

Essa dissertação propõe um sistema de odometria baseado na fusão de dados provenientes de uma câmera orientada para o solo, de uma unidade inercial (IMU) e da odometria de rodas, utilizando um filtro de Kalman como estrutura de integração sensorial. A câmera foi utilizada para estimar o movimento visual por meio do cálculo de fluxo óptico denso utilizando o método de Farneback, uma abordagem baseada em aparência que permite explorar a totalidade da imagem. Essa técnica apresenta menor complexidade computacional em comparação com métodos baseados em características e é

mais adequada a superfícies de baixa textura, como terra, areia compactada e cascalho fino, comumente encontradas em ambientes de mineração. Considerando que o processamento de imagens impõe elevada demanda computacional, a implementação foi acelerada em GPU, de modo a garantir viabilidade prática para aplicações embarcadas em tempo real.

Em comparação ao método apresentado por [Domingues et al. \[2022\]](#), que estima a odometria a partir da fusão de fluxo óptico, IMU e odometria de rodas, estabelecendo uma relação geométrica entre a projeção em perspectiva da câmera e o veículo, esta dissertação introduz melhorias como o processamento do fluxo óptico em GPU e o desenvolvimento de um Filtro de Kalman dedicado para a integração das três fontes de dados, aumentando a eficiência computacional e a precisão das estimativas.

Capítulo 3

Fundamentos Teóricos

Este capítulo apresenta os fundamentos teóricos necessários para o desenvolvimento e a compreensão do sistema de odometria proposto. Inicialmente, são abordados os conceitos de cinemática de corpos rígidos, incluindo a representação da posição e orientação, transformações homogêneas e a propagação de velocidades entre referenciais. Em seguida, são descritos os modelos cinemáticos comumente utilizados para veículos com rodas, abrangendo as configurações de unicycle, diferencial, *skid-steer* e Ackermann. Na sequência, os princípios da odometria visual são introduzidos, desde a formação da imagem até os métodos de estimação incremental de movimento baseados em dados visuais, com destaque para o uso de fluxo óptico e aceleração computacional por GPU. Por fim, são discutidas as técnicas de fusão sensorial, com ênfase nos Filtros de Kalman e Kalman Estendido.

3.1 Cinemática de Corpos Rígidos

A cinemática de corpos rígidos estuda o movimento de objetos sem considerar as forças ou torques que o causam. Um corpo rígido é definido como um objeto cuja forma e dimensões não se alteram durante o movimento. Em robótica, a cinemática é usada para descrever como esses corpos se movem ao longo do tempo. O movimento de um corpo rígido é caracterizado por sua *pose*, isto é, sua posição e orientação no espaço, bem como pelas derivadas temporais associadas, como velocidade e aceleração.

3.1.1 Posição e Translação

A posição pode ser representada pela localização da origem de um sistema de coordenadas fixo ao corpo rígido, denotado por $\{\mathcal{O}\}$ com eixos $[\hat{x}_{\mathcal{O}}, \hat{y}_{\mathcal{O}}, \hat{z}_{\mathcal{O}}]^T$, em relação a um sistema de coordenadas inercial $\{\mathcal{I}\}$ com eixos $[\hat{x}_{\mathcal{I}}, \hat{y}_{\mathcal{I}}, \hat{z}_{\mathcal{I}}]^T$. Essa posição é representada por um vetor coluna $\mathbf{p}_{\mathcal{O}}^{\mathcal{I}} \in \mathbb{R}^3$, que define a posição da origem de $\{\mathcal{O}\}$ expressa no sistema $\{\mathcal{I}\}$. Uma translação é um tipo de deslocamento no qual nenhum ponto do corpo permanece em sua posição inicial. Dessa forma, a posição de um corpo rígido pode ser interpretada

como uma translação que leva o sistema $\{\mathcal{O}\}$, inicialmente coincidente com $\{\mathcal{I}\}$, até sua posição atual, onde as origens estão separadas pelo vetor $\mathbf{p}_{\mathcal{O}}^{\mathcal{I}}$:

$$\mathbf{p}_{\mathcal{O}}^{\mathcal{I}} = \begin{bmatrix} p_{\mathcal{O},x}^{\mathcal{I}} \\ p_{\mathcal{O},y}^{\mathcal{I}} \\ p_{\mathcal{O},z}^{\mathcal{I}} \end{bmatrix}, \quad (3.1)$$

sendo $p_{\mathcal{O},x}^{\mathcal{I}}$, $p_{\mathcal{O},y}^{\mathcal{I}}$ e $p_{\mathcal{O},z}^{\mathcal{I}}$ as coordenadas cartesianas x , y e z de $\{\mathcal{O}\}$ no referencial $\{\mathcal{I}\}$.

3.1.2 Orientação e Rotação

A orientação de um corpo rígido descreve como ele está rotacionado em relação a um sistema de referência. Diferente da posição, que pode ser representada por um vetor em \mathbb{R}^3 , a orientação possui mais representações. Uma rotação é um tipo de deslocamento no qual pelo menos um ponto do corpo rígido permanece fixo, enquanto outras partes mudam de orientação em relação a um referencial inercial. Por exemplo, um corpo em movimento circular pode girar em torno de um eixo, mantendo seus pontos nesse eixo em posição constante. Assim como a posição está associada à translação, a orientação está relacionada diretamente à rotação, e uma pode ser usada para descrever a outra.

A orientação de um corpo rígido pode ser descrita pela rotação de um sistema de coordenadas $\{\mathcal{O}\}$, fixo ao corpo, em relação a um sistema de referência $\{\mathcal{I}\}$. Essa rotação é representada por uma matriz $\mathbf{R}_{\mathcal{O}}^{\mathcal{I}} \in \mathbb{R}^{3 \times 3}$, cujas colunas correspondem aos vetores de base de $\{\mathcal{O}\}$ expressos em $\{\mathcal{I}\}$. Como esses vetores são unitários e mutuamente ortogonais, a matriz $\mathbf{R}_{\mathcal{O}}^{\mathcal{I}}$ é ortogonal, ou seja, sua transposta é igual à sua inversa, $(\mathbf{R}_{\mathcal{O}}^{\mathcal{I}})^{\top} \mathbf{R}_{\mathcal{O}}^{\mathcal{I}} = \mathbb{I}$, onde \mathbb{I} é a matriz identidade. Além disso, seu determinante é igual a $+1$, o que classifica essa matriz como pertencente ao grupo ortogonal especial $SO(3)$. A matriz $\mathbf{R}_{\mathcal{O}}^{\mathcal{I}}$ permite transformar vetores expressos no sistema $\{\mathcal{O}\}$ para o sistema $\{\mathcal{I}\}$ e pode ser interpretada como uma rotação que leva $\{\mathcal{O}\}$ da orientação inicial coincidente com $\{\mathcal{I}\}$ até sua orientação atual.

Além das matrizes de rotação, outras formas utilizadas para representar a orientação de corpos rígidos são os ângulos de Euler e os ângulos *roll*, *pitch* e *yaw* (*RPY*). Essas representações descrevem a orientação como uma sequência de rotações elementares em torno de eixos específicos. Embora ambos os métodos utilizem três ângulos para definir uma orientação no espaço tridimensional, eles se baseiam em convenções distintas de ordem e eixos de rotação.

Os ângulos de Euler representam a orientação de um sistema de coordenadas $\{\mathcal{O}\}$ em relação a um sistema de referência $\{\mathcal{I}\}$ como uma sequência de três rotações sucessivas em torno dos eixos do próprio sistema em movimento $\{\mathcal{O}\}$. Cada um dos ângulos corresponde a uma rotação em torno de um dos eixos, sendo que cada rotação ocorre em um referencial já afetado pela rotação anterior, como exemplificado na Figura 6. Por essa

razão, a ordem das rotações deve ser especificada junto aos ângulos para que a orientação seja corretamente determinada.

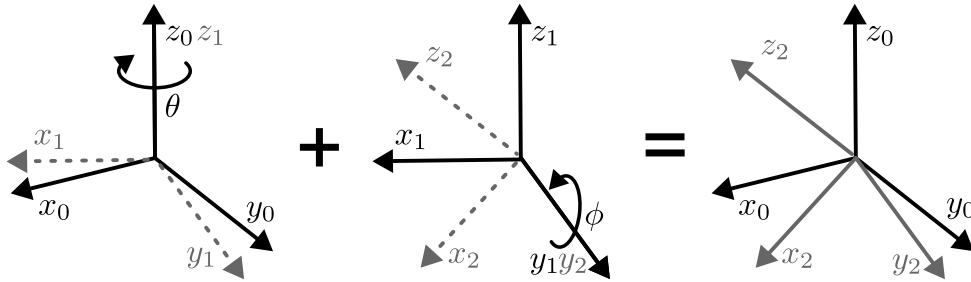


Figura 6 – Composição das rotações em torno dos eixos atuais.

A orientação de um sistema de coordenadas $\{\mathcal{O}\}$ em relação a um sistema de referência $\{\mathcal{I}\}$ pode ser expressa como uma sequência de três rotações sucessivas em torno dos eixos do sistema de coordenadas fixo $\{\mathcal{I}\}$. Os ângulos *RPY*, denotados por $[\psi, \theta, \phi]^T$, correspondem a rotações em torno dos eixos do referencial inercial, como ilustrado na Figura 7.

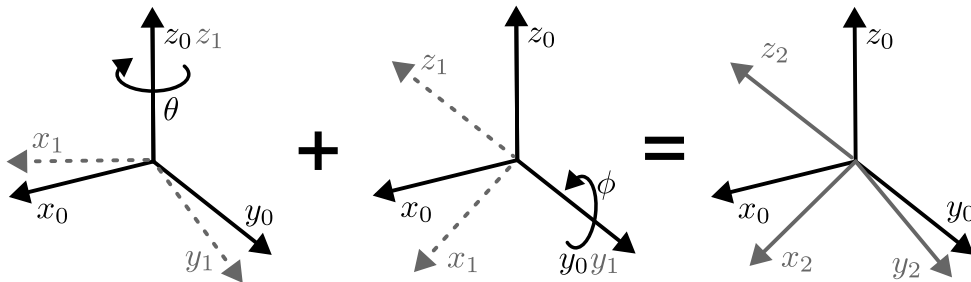


Figura 7 – Composição de rotações em torno de eixos fixos.

3.1.3 Transformação Homogênea

Uma vez definidos os conceitos de posição e orientação de um sistema de coordenadas $\{\mathcal{O}\}$ em relação a um sistema de referência $\{\mathcal{I}\}$, é possível combiná-los por meio de uma matriz de transformação homogênea. Essa matriz representa simultaneamente a translação e a rotação entre os dois sistemas, sendo expressa como uma matriz 4×4 que atua sobre representações homogêneas de vetores. A transformação homogênea de $\{\mathcal{O}\}$ em relação a $\{\mathcal{I}\}$ é definida como:

$$\mathbf{T}_{\mathcal{O}}^{\mathcal{I}} = \begin{bmatrix} \mathbf{R}_{\mathcal{O}}^{\mathcal{I}} & \mathbf{p}_{\mathcal{O}}^{\mathcal{I}} \\ \mathbf{0}^T & 1 \end{bmatrix}, \quad (3.2)$$

onde $\mathbf{R}_{\mathcal{O}}^{\mathcal{I}} \in \mathbb{R}^{3 \times 3}$ representa a orientação e $\mathbf{p}_{\mathcal{O}}^{\mathcal{I}} \in \mathbb{R}^3$ representa a posição do sistema $\{\mathcal{O}\}$ em relação a $\{\mathcal{I}\}$. Essa matriz permite transformar pontos e vetores expressos em $\{\mathcal{O}\}$ para o sistema $\{\mathcal{I}\}$:

$$\begin{bmatrix} \mathbf{r}_{\mathcal{I}} \\ 1 \end{bmatrix} = \underbrace{\begin{bmatrix} \mathbf{R}_{\mathcal{O}}^{\mathcal{I}} & \mathbf{p}_{\mathcal{O}}^{\mathcal{I}} \\ \mathbf{0}^T & 1 \end{bmatrix}}_{\mathbf{T}_{\mathcal{O}}^{\mathcal{I}}} \begin{bmatrix} \mathbf{r}_{\mathcal{O}} \\ 1 \end{bmatrix}, \quad (3.3)$$

em que $\mathbf{r}_{\mathcal{I}}$ é a posição em $\{\mathcal{I}\}$ e $\mathbf{r}_{\mathcal{O}}$ é a posição em $\{\mathcal{O}\}$.

3.1.4 Transformação de Velocidades em Corpos Rígidos

Em um corpo rígido, diferentes pontos ou referenciais fixados à estrutura compartilham o mesmo movimento global. Quando as velocidades linear e angular são conhecidas em um ponto do corpo, é possível propagá-las para outro ponto rigidamente conectado por meio de uma transformação adjunta [Spong et al., 2006]. As velocidades linear $\mathbf{v}_{\mathcal{A}} = [v_{\mathcal{A},x} \ v_{\mathcal{A},y} \ v_{\mathcal{A},z}]^T$ e angular $\boldsymbol{\omega}_{\mathcal{A}} = [\omega_{\mathcal{A},x} \ \omega_{\mathcal{A},y} \ \omega_{\mathcal{A},z}]^T$ expressas no referencial $\{\mathcal{A}\}$ são propagadas para o referencial $\{\mathcal{B}\}$ pela Equação (3.4):

$$\begin{bmatrix} \mathbf{v}_{\mathcal{B}} \\ \boldsymbol{\omega}_{\mathcal{B}} \end{bmatrix} = \begin{bmatrix} \mathbf{R}_{\mathcal{A}}^{\mathcal{B}} & \mathbf{S}(d)\mathbf{R}_{\mathcal{A}}^{\mathcal{B}} \\ 0 & \mathbf{R}_{\mathcal{A}}^{\mathcal{B}} \end{bmatrix} \begin{bmatrix} \mathbf{v}_{\mathcal{A}} \\ \boldsymbol{\omega}_{\mathcal{A}} \end{bmatrix}, \quad (3.4)$$

sendo $\mathbf{S}(d)$ a matriz anti-simétrica da distância entre os dois referenciais, $\mathbf{R}_{\mathcal{A}}^{\mathcal{B}}$ é a matriz de rotação de $\{\mathcal{A}\}$ para $\{\mathcal{B}\}$.

3.2 Modelos Cinemáticos de Veículos com Rodas

Os modelos cinemáticos descrevem o movimento do sistema sem considerar suas causas, focando na relação entre os parâmetros de entrada e o comportamento do sistema. Esses modelos são representados por um conjunto de equações diferenciais de primeira ordem. A modelagem cinemática leva em conta as contribuições individuais de cada roda para o deslocamento do veículo, bem como as restrições que essas rodas impõem ao seu movimento [Klančar et al., 2017].

A pose de um veículo em um plano pode ser representada por um vetor de estado $\mathbf{q}_{\mathcal{V}}^{\mathcal{I}}$, que expressa a posição e orientação do sistema de coordenadas do veículo $\{\mathcal{V}\}$ em relação ao referencial inercial $\{\mathcal{I}\}$:

$$\mathbf{q}_{\mathcal{V}}^{\mathcal{I}} = \begin{bmatrix} x_{\mathcal{V}}^{\mathcal{I}} \\ y_{\mathcal{V}}^{\mathcal{I}} \\ \theta_{\mathcal{V}} \end{bmatrix}. \quad (3.5)$$

Nesse vetor, os dois primeiros termos indicam as coordenadas do veículo no plano global, enquanto o último representa sua orientação angular em relação ao eixo z do referencial. A relação entre o sistema de coordenadas móvel, fixado ao veículo, e o

referencial inercial é definida por essa translação $[x_V^I \ y_V^I]^T$ e por uma matriz de rotação $R_V^I(\theta_V)$, dada por:

$$R_V^I(\theta_V) = \begin{bmatrix} \cos \theta_V & -\sin \theta_V & 0 \\ \sin \theta_V & \cos \theta_V & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (3.6)$$

• Modelo Cinemático Monociclo

Um dos modelos cinemáticos mais simples para representar o movimento de veículos no plano é o modelo do monociclo, como pode ser visto na Figura 8. Nesse modelo, o veículo é idealizado como uma roda que pode se deslocar em linha reta com velocidade linear v_V e girar em torno de seu próprio centro com velocidade angular ω_V , sem ocorrência de deslizamento tanto na direção longitudinal quanto lateral. Sua estrutura pode ser diretamente associada à representação do estado do veículo, dada na Equação (3.5), sendo este controlado pelas variáveis v_V e ω_V , como mostra a Equação (3.7):

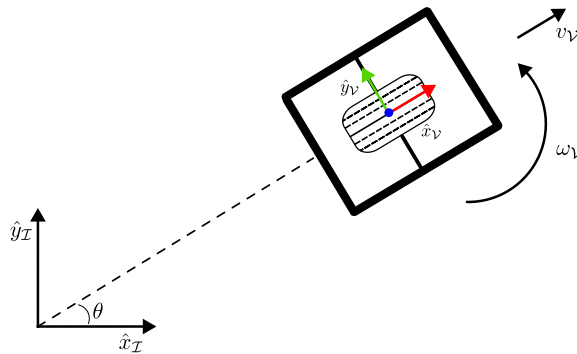


Figura 8 – Representação do Modelo Cinemático Monociclo.

$$\begin{bmatrix} \dot{x}_V^I \\ \dot{y}_V^I \\ \dot{\theta}_V \end{bmatrix} = \begin{bmatrix} \cos \theta_V & 0 \\ \sin \theta_V & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} v_V \\ \omega_V \end{bmatrix}. \quad (3.7)$$

Apesar de ser uma abstração, este modelo é empregado como aproximação para outros modelos de veículos, desde que as velocidades v_V e ω_V possam ser controladas.

• Modelo Cinemático Diferencial

O acionamento diferencial, como ilustrado na Figura 9, é um mecanismo utilizado especialmente em plataformas de menor porte, devido à sua simplicidade construtiva e facilidade de controle. Nesse tipo de configuração, o veículo possui duas rodas principais alinhadas em um mesmo eixo, cada uma acionada por um motor independente, e geralmente

conta com uma ou mais rodas livres para apoio e estabilidade. O modelo é descrito pela Equação 3.8.

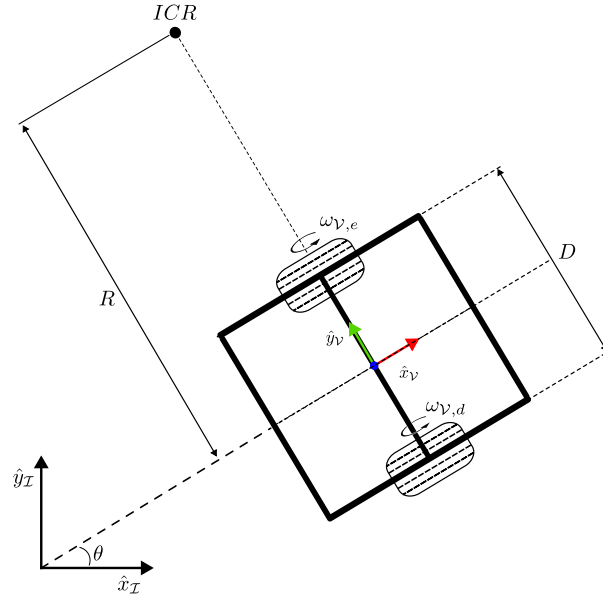


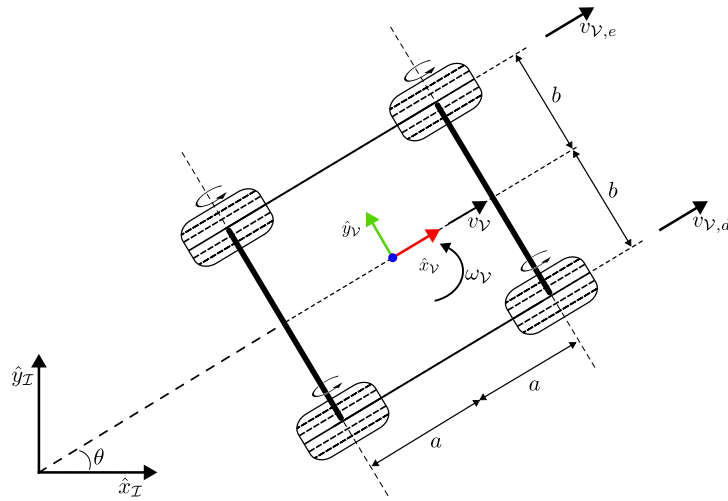
Figura 9 – Representação do Modelo Cinemático Diferencial.

$$\begin{bmatrix} \dot{\hat{x}}_v^I \\ \dot{\hat{y}}_v^I \\ \dot{\theta}_v \end{bmatrix} = \begin{bmatrix} \frac{r}{2} \cos \theta_v & \frac{r}{2} \cos \theta_v \\ \frac{r}{2} \sin \theta_v & \frac{r}{2} \sin \theta_v \\ \frac{r}{D} & -\frac{r}{D} \end{bmatrix} \begin{bmatrix} \omega_{v,d} \\ \omega_{v,e} \end{bmatrix}, \quad (3.8)$$

onde $\omega_{v,d}$ e $\omega_{v,e}$ são as velocidades angulares das rodas direita e esquerda, respectivamente, D é a distância entre as rodas, r é o raio da roda. Esse modelo também serve como base para a descrição do movimento de veículos com tração do tipo *skid-steering*. Nessa configuração, a direção do movimento é obtida pela diferença de velocidade entre os lados esquerdo e direito do veículo, tal como no diferencial, mas o giro ocorre pelo deslizamento lateral controlado das rodas ou esteiras.

• Modelo Cinemático *Skid-Steering*

Plataformas com locomoção do tipo *skid-steer*, como mostrado na Figura 10, utilizam rodas fixas ou esteiras dispostas paralelamente à estrutura do veículo, sendo que as rodas de cada lado giram à mesma velocidade. Os movimentos da plataforma são determinados pelas velocidades impostas em cada lado, de tal forma que essas velocidades resultam em deslocamentos lineares, rotações ou curvas. Esse tipo de locomoção é empregada em veículos terrestres, especialmente em ambientes irregulares, como operações em mineração e agricultura [Mandow et al., 2007]. O modelo cinemático é descrito pela Equação (3.9), conforme apresentado por Dogru and Marques [2021]:


 Figura 10 – Representação do Modelo Cinemático *Skid-Steering*.

$$\begin{bmatrix} v_V \\ \omega_V \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{b}{2(a^2+b^2)} & -\frac{b}{2(a^2+b^2)} \end{bmatrix} \begin{bmatrix} v_{V,d} \\ v_{V,e} \end{bmatrix}, \quad (3.9)$$

onde $v_{V,d}$ e $v_{V,e}$ são as velocidades lineares do veículo nas rodas direita e esquerda, respectivamente, a é a metade da distância entre os eixos e b é a metade da distância entre as rodas. Esse modelo pode ser usado para calcular o deslocamento do veículo no referencial inercial usando o modelo monociclo da Equação (3.7).

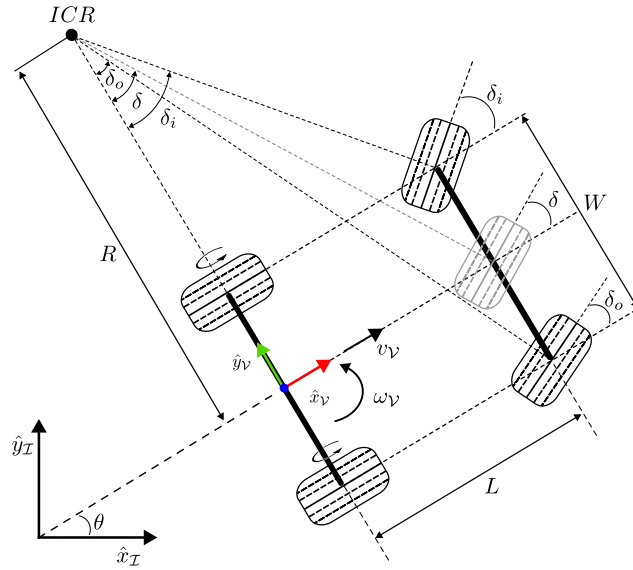
• Modelo Cinemático *Ackermann*

Veículos com geometria de direção do tipo *Ackermann*, como os automóveis convencionais, possuem um eixo traseiro motorizado e um par de rodas dianteiras direcionáveis, como mostrado na Figura 11. Tradicionalmente, um mecanismo de *Ackermann* é empregado no sistema de direção para garantir que ambas as rodas dianteiras se alinhem em torno de um Centro de Rotação Instantâneo (ICR) comum durante curvas. Para isso, os ângulos de viragem das rodas dianteiras, δ_i e δ_o , são diferentes, pois descrevem trajetórias com curvaturas distintas. Esses ângulos são calculados a partir de um ângulo de viragem central virtual, denotado por δ , conforme definido nas Equações (3.10) e (3.11):

$$\delta_i = \frac{L \tan(\delta)}{L - \frac{W}{2} \tan(\delta)}, \quad (3.10)$$

$$\delta_o = \frac{L \tan(\delta)}{L + \frac{W}{2} \tan(\delta)}. \quad (3.11)$$

O movimento do veículo, por sua vez, é descrito por esse ângulo e pela velocidade v_V no ponto central do eixo traseiro, que pode ser aproximada pela média das velocidades das rodas traseiras.


 Figura 11 – Representação do Modelo Cinemático *Ackermann*.

A partir dessa simplificação, o modelo de *Ackermann* pode ser reformulado considerando o modelo monociclo, no qual o estado do veículo é descrito diretamente por v_V e δ , conforme mostrado na Equação (3.12):

$$\begin{bmatrix} \dot{x}_V^I \\ \dot{y}_V^I \\ \dot{\theta}_V \end{bmatrix} = \begin{bmatrix} \cos \theta_V \\ \sin \theta_V \\ \frac{\tan \delta}{L} \end{bmatrix} v_V. \quad (3.12)$$

3.3 Odometria Visual

A odometria visual é uma técnica que permite estimar o movimento de um agente a partir da análise sequencial de imagens. Esta seção aborda os principais conceitos que sustentam essa abordagem, começando pela formação da imagem e pela forma como as informações visuais são capturadas e representadas. Em seguida, são discutidos os métodos de estimação incremental de pose com base em dados visuais, que permitem reconstruir a trajetória do agente ao longo do tempo. Também são apresentados os princípios do fluxo óptico, amplamente utilizado na VO para inferir deslocamentos entre quadros consecutivos. Por fim, destaca-se o papel do processamento paralelo por GPU, que viabiliza a execução eficiente de algoritmos de visão computacional em tempo real.

3.3.1 Formação de Imagem

A Figura 12 ilustra o modelo de formação de imagem em uma câmera de perspectiva. Dentre os modelos existentes, o modelo pinhole é o mais conhecido e empregado para descrever o processo de formação de imagens em sistemas de visão computacional. Nesse modelo, assume-se que a cena tridimensional é projetada em uma superfície bidimensional

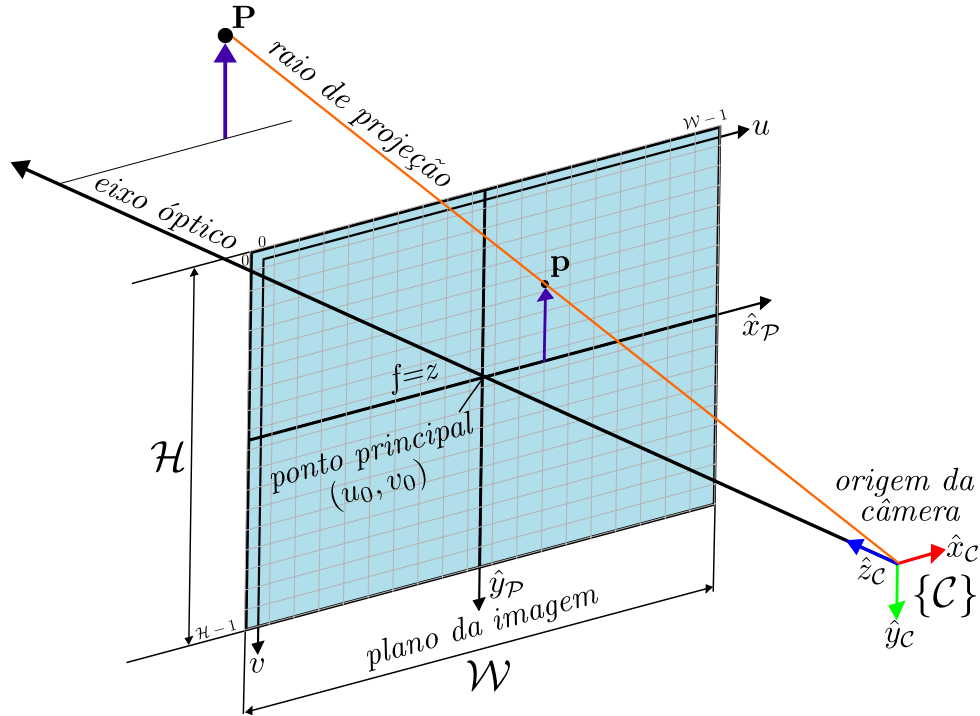


Figura 12 – Modelo de projeção central mostrando o plano da imagem e *pixels* discretos. O raio de luz do ponto P cruza o plano da imagem no ponto p ao passar pelo centro da projeção.

chamada plano da imagem. Por meio de semelhança de triângulos, um ponto do mundo com vetor de coordenadas $\mathbf{P} = [X, Y, Z]^T$, por exemplo, é mapeado para um ponto sobre esse plano. O plano da imagem possui um sistema de coordenadas 2D denotado por $\{\mathcal{P}\}$, e está posicionado a uma distância f do centro óptico ao longo do eixo \hat{z}_c no sistema de coordenadas da câmera $\{\mathcal{C}\}$, onde f representa a distância focal da lente. A projeção ocorre pela interseção do raio que conecta o ponto \mathbf{P} à origem de $\{\mathcal{C}\}$ com o plano da imagem. O ponto resultante da projeção no plano da imagem é $\mathbf{p} = [x_p, y_p]^T$, determinado pelas Equações (3.13) e (3.14):

$$x_p = f \frac{X}{Z}, \quad (3.13)$$

$$y_p = f \frac{Y}{Z}. \quad (3.14)$$

Essa transformação é conhecida como projeção em perspectiva e, por reduzir a dimensionalidade da cena de três para duas dimensões, elimina a informação de profundidade. Como consequência, diferentes pontos ao longo da mesma linha de visão, mas em distâncias distintas, podem ser projetados nas mesmas coordenadas (x_p, y_p) . Isso torna impossível, apenas com base na imagem formada, distinguir se um objeto é pequeno e próximo ou grande e distante.

Na prática, o plano da imagem é discretizado em uma grade bidimensional de resolução $\mathcal{W} \times \mathcal{H}$, onde \mathcal{W} representa a largura da imagem em *pixels* e \mathcal{H} sua altura. Essa grade é composta por elementos sensores sensíveis à luz, que convertem a radiação luminosa incidente em sinais digitais. Cada elemento sensor corresponde diretamente a um pixel da imagem digital gerada, formando uma malha de amostragem uniforme.

As coordenadas do plano da imagem $\{\mathcal{P}\}$ são então convertidas para coordenadas de pixel, representadas pelo vetor de duas dimensões $[u, v]^T$, que representam índices na grade digital, e, por convenção, sua origem fica no canto superior esquerdo da imagem. A discretização do plano da imagem para coordenadas em pixel é realizada por meio da relação:

$$u = \frac{x_{\mathcal{P}}}{\rho_w} + u_0, \quad (3.15)$$

$$v = \frac{y_{\mathcal{P}}}{\rho_h} + v_0, \quad (3.16)$$

em que ρ_w e ρ_h correspondem ao tamanho do pixel nas direções horizontal e vertical, respectivamente. O termo (u_0, v_0) representa o ponto principal da imagem, ou seja, a coordenada de pixel correspondente à interseção do eixo óptico com o plano da imagem. Essa transformação permite mapear pontos contínuos da projeção para uma representação discreta que pode ser processada computacionalmente.

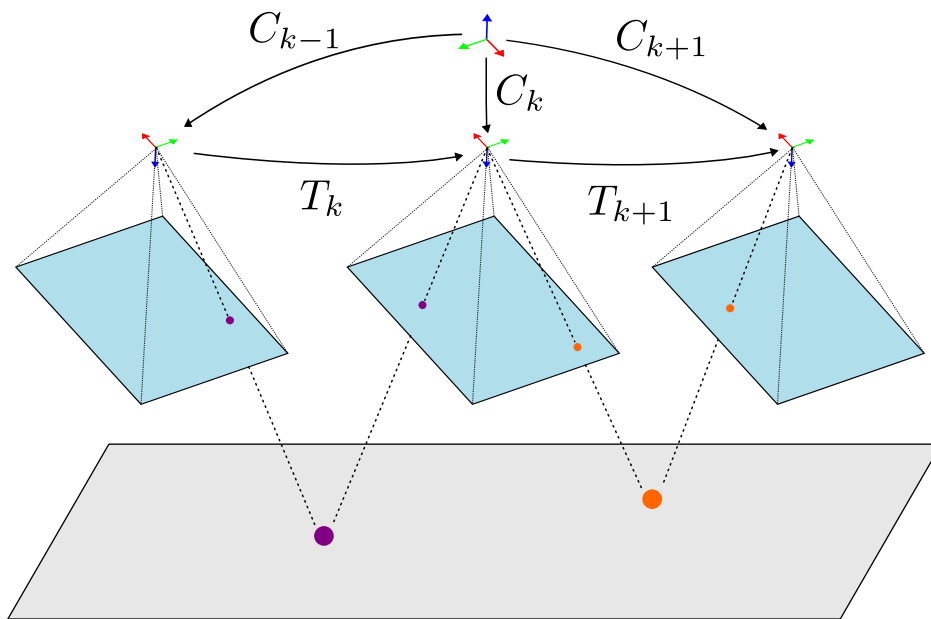


Figura 13 – Ilustração do processo de odometria visual.

3.3.2 Estimativa Incremental de Poses a Partir de Dados Visuais

A odometria visual permite estimar o deslocamento de um sistema com base na análise sequencial de imagens adquiridas por uma câmera ao longo do tempo. Como ilustrado na Figura 13, à medida que a câmera se move, uma mesma região do ambiente é registrada em diferentes posições e instantes, possibilitando a identificação de correspondências visuais entre quadros consecutivos. A análise dessas variações espaciais e temporais fornece a base para inferir a trajetória percorrida pela câmera a partir da própria percepção visual.

Durante o deslocamento em um ambiente, um agente registra imagens em instantes de tempo discretos k por meio de uma câmera rigidamente acoplada à sua estrutura. No caso monocular, o conjunto de imagens adquiridas no instante k é representado por $i_{0:n} = i_0, \dots, i_n$, sendo cada imagem indexada por sua respectiva captura. As posições da câmera nos instantes $k - 1$ e k são relacionadas por uma transformação rígida, e o conjunto dessas transformações sucessivas é denotado por $T_{1:n} = T_{1,0}, \dots, T_{n,n-1}$. Já o conjunto de poses absolutas da câmera, em relação à posição inicial $k = 0$, é dado por $C_{0:n} = C_0, \dots, C_n$, onde cada pose pode ser obtida por meio da concatenação das transformações incrementais, ou seja, $C_n = C_{n-1}T_n$, sendo C_0 a posição da câmera no instante $k = 0$. Uma vez que as poses são estimadas no referencial da câmera, a obtenção da pose do agente exige a consideração da transformação rígida entre os dois sistemas de coordenadas.

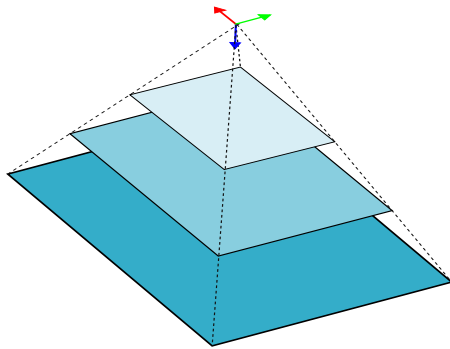
Para computar o movimento relativo T_k entre pares de imagens, podem ser adotadas duas abordagens principais, conhecidas como métodos baseados em aparência e métodos baseados em características. Os primeiros operam diretamente sobre os níveis de intensidade dos *pixels* de toda a imagem, buscando minimizar discrepâncias globais entre os quadros sucessivos. Já os métodos baseados em características funcionam extraindo pontos salientes e repetíveis, como cantos ou bordas, e associando-os entre imagens consecutivas. Neste trabalho, optou-se por utilizar uma abordagem baseada em aparência, que apresenta vantagens em ambientes com textura contínua e em aplicações onde o padrão visual do solo pode ser explorado de forma densa.

Um requisito para o funcionamento da odometria visual é que haja sobreposição entre imagens consecutivas, de modo que os pontos de interesse detectados em uma imagem possam ser localizados na imagem seguinte. Em câmeras voltadas para frente, isso geralmente não é um problema por causa do amplo campo de visão em expansão. No entanto, em câmeras voltadas para o solo, esse requisito depende fortemente da velocidade de deslocamento do agente e da altura da câmera em relação ao solo. A sobreposição entre quadros consecutivos diminui à medida que a velocidade aumenta, já que a distância percorrida entre capturas é maior. Por outro lado, o aumento da altura da câmera amplia o campo de visão sobre o solo, o que contribui para aumentar a área comum entre imagens

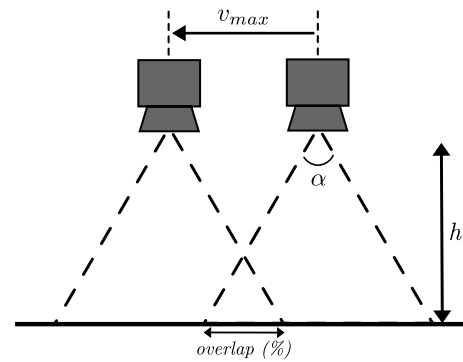
sucessivas. A relação entre esses fatores foi descrita em [Ericson and Åstrand, 2018], e pode ser expressa pela Equação (3.17), que define a sobreposição em função da velocidade v_{\max} em metros por segundo, da altura da câmera acima do solo h em metros, da frequência de amostragem f_s em hertz e do campo de visão α :

$$\text{overlap} = 1 - \frac{v_{\max}}{2hf_s \tan\left(\frac{\alpha}{2}\right)}. \quad (3.17)$$

A Figura 14a ilustra como a área capturada varia em função da distância até a câmera, enquanto a Figura 14b mostra a relação entre sobreposição, altura da câmera e velocidade máxima admissível.



(a) Variação da Área Capturada em Função da Distância da Câmera.



(b) Sobreposição máxima para câmera voltada para baixo.

Figura 14 – Ilustração da captura de imagens em diferentes contextos: (a) relação entre distância e campo de visão e (b) sobreposição de quadros consecutivos.

3.3.3 Fluxo Óptico

O fluxo óptico é um conceito utilizado para descrever o movimento dos pontos de uma cena projetados no plano da imagem, resultante do deslocamento relativo entre a câmera e o ambiente tridimensional, seja por movimento da própria câmera, da cena ou de ambos. Ele é representado por um campo vetorial que indica a velocidade de deslocamento da informação visual entre imagens sucessivas, como ilustrado na Figura 15. A estimativa do fluxo óptico pode ser realizada por diferentes abordagens, que se distinguem pelas suposições adotadas sobre o comportamento da imagem e o movimento da câmera. Neste trabalho, serão exploradas duas dessas abordagens, sendo a primeira baseada na variação de intensidade dos *pixels* ao longo do tempo e a segunda formulada diretamente a partir dos parâmetros de movimento da câmera.

- **Modelagem do Fluxo Óptico a Partir da Restrição de Intensidade**

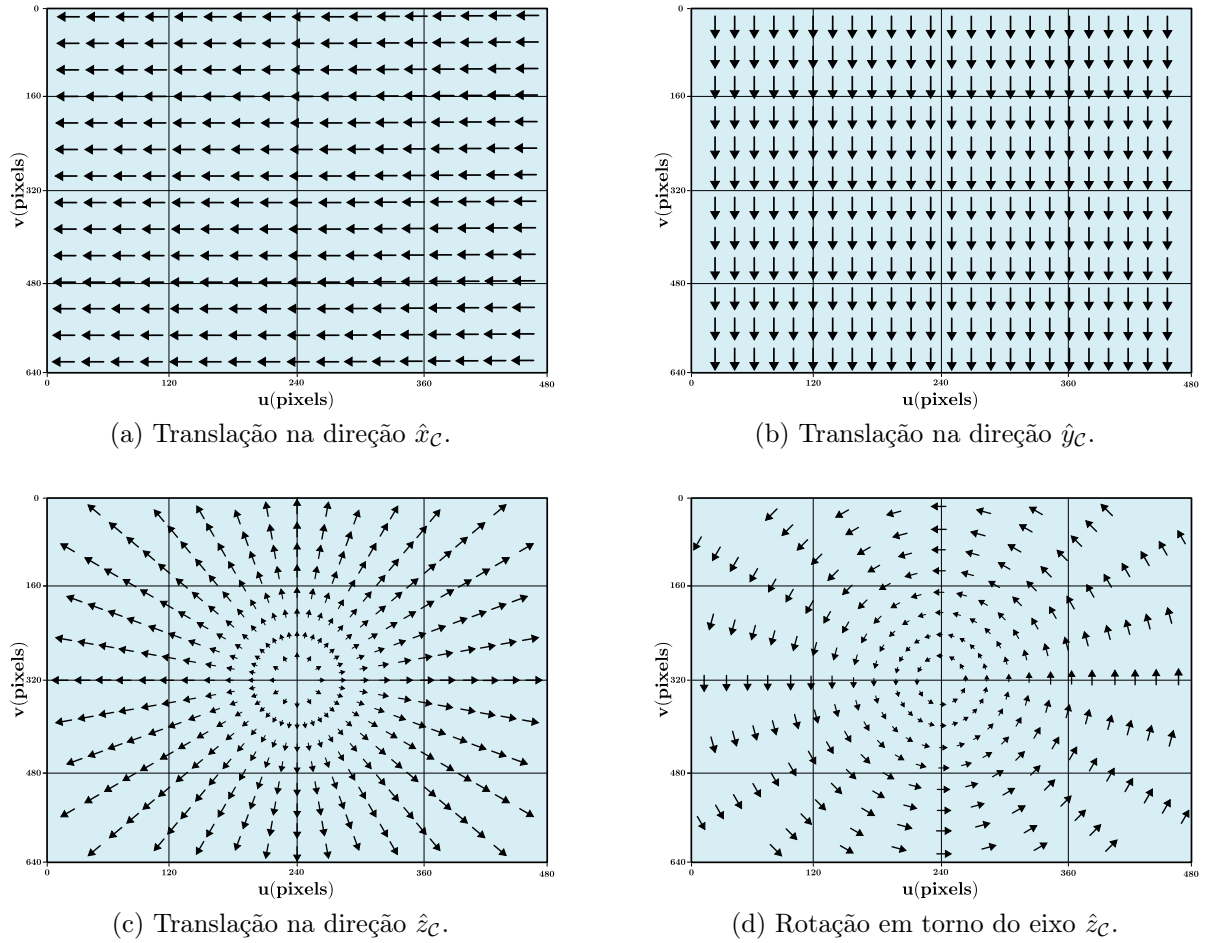


Figura 15 – Ilustração de quatro situações distintas do fluxo óptico, cada uma resultante de diferentes movimentos da câmera em relação à cena observada.

O cálculo do fluxo óptico parte do princípio de que o brilho de um ponto da imagem não se altera significativamente entre quadros consecutivos. Com base nessa suposição, é possível derivar a chamada equação de restrição do fluxo óptico, que expressa uma relação entre os gradientes espaciais e temporais da intensidade da imagem:

$$\frac{\partial I}{\partial x} \frac{dx_{\mathcal{P}}}{dt} + \frac{\partial I}{\partial y} \frac{dy_{\mathcal{P}}}{dt} + \frac{\partial I}{\partial t} = 0, \quad (3.18)$$

onde $\frac{\partial I}{\partial x_{\mathcal{P}}}$ e $\frac{\partial I}{\partial y_{\mathcal{P}}}$ correspondem aos gradientes espaciais da intensidade da imagem nas direções horizontal e vertical, respectivamente, enquanto $\frac{\partial I}{\partial t}$ representa a derivada temporal da intensidade. Os termos $\frac{dx_{\mathcal{P}}}{dt}$ e $\frac{dy_{\mathcal{P}}}{dt}$ denotam as componentes do vetor de fluxo óptico, ou seja, as velocidades do ponto da imagem nas direções x e y . Para resolver essa equação e estimar os vetores de fluxo, são empregados algoritmos específicos, como o método de Gunnar-Farneback [Farneback, 2003], que utiliza uma abordagem densa baseada em modelos polinomiais locais para capturar os deslocamentos entre os quadros consecutivos.

• Modelagem do Fluxo Óptico a Partir do Movimento da Câmera

O fluxo óptico também pode ser descrito como resultado direto do movimento tridimensional da câmera e da forma como os pontos da cena são projetados no plano da imagem. A magnitude do fluxo óptico está relacionada à velocidade da câmera e à distância dos pontos em relação a ela, o que leva a uma ambiguidade de escala, pois diferentes combinações de velocidade e profundidade podem produzir padrões de fluxo indistinguíveis na imagem. Para lidar com essa limitação, é comum integrar informações adicionais, como dados de profundidade ou medições inerciais. A relação entre a velocidade dos pontos da imagem e o movimento da câmera é formalizada pela Equação (3.19), frequentemente referida como a equação do fluxo óptico [Corke, 2023]:

$$\begin{bmatrix} \dot{u} \\ \dot{v} \end{bmatrix} = \begin{bmatrix} -\frac{f'}{Z_p} & 0 & \frac{\bar{u}}{Z_p} & \frac{\bar{u}\bar{v}}{f'} & -\frac{f'^2+\bar{u}}{f'} & \bar{v} \\ 0 & -\frac{f'}{Z_p} & \frac{\bar{v}}{Z_p} & \frac{f'^2+\bar{v}}{f'} & -\frac{\bar{u}\bar{v}}{f'} & -\bar{u} \end{bmatrix} \begin{bmatrix} \mathbf{v}_C \\ \boldsymbol{\omega}_C \end{bmatrix}, \quad (3.19)$$

no qual \dot{u} e \dot{v} representam as velocidades de fluxo em *pixels/s*, f' é a distância focal da câmera em *pixels*, Z_p é a profundidade do *pixel* em metros, $\bar{u} = u - u_0$ é a coordenada horizontal do *pixel* com respeito ao centro da imagem (u_0), $\bar{v} = v - v_0$ é a coordenada vertical do *pixel* com respeito ao centro da imagem (v_0), $\mathbf{v}_C = [v_{C,x} \ v_{C,y} \ v_{C,z}]^T$ são as velocidades lineares da câmera e $\boldsymbol{\omega}_C = [\omega_{C,x} \ \omega_{C,y} \ \omega_{C,z}]^T$ as velocidades angulares da câmera.

3.4 Processamento de Imagem usando GPU

O processamento de imagens constitui uma parte significativa da visão computacional, envolvendo operações que se beneficiam da execução paralela. Nesse contexto, as unidades de processamento gráfico (GPUs), compostas por milhares de núcleos de execução, oferecem uma arquitetura paralela capaz de acelerar significativamente algoritmos de processamento de imagem. A plataforma CUDA, desenvolvida pela NVIDIA, permite a utilização desses núcleos para cálculos de uso geral, viabilizando a implementação de rotinas otimizadas em bibliotecas como o OpenCV, uma biblioteca amplamente utilizada para processamento de imagens e visão computacional. Apesar do suporte ao módulo CUDA no OpenCV desde 2010, seu uso ainda não é amplamente difundido [Cervera, 2020].

No processamento realizado exclusivamente pela CPU, os dados são armazenados na memória principal e todas as operações são executadas pela própria CPU, como ilustrado na Figura 16a. Já quando se utiliza uma GPU, o fluxo envolve a transferência dos dados da memória da CPU para a memória da GPU, a execução paralela dos algoritmos nos núcleos da GPU e, por fim, a cópia dos resultados de volta para a memória principal, conforme representado na Figura 16b.

Esse processo introduz uma sobrecarga própria ao uso da GPU, decorrente da movimentação de dados entre os diferentes espaços de memória. O tempo total de transferência é dado por:

$$T_{overhead} = T_{upload} + T_{download}, \quad (3.20)$$

onde T_{upload} representa o tempo necessário para enviar os dados da CPU para a GPU, e $T_{download}$ corresponde ao tempo gasto para retornar os dados processados à memória principal. Para que o uso da GPU proporcione ganho de desempenho, é necessário que o tempo de execução na CPU seja superior ao tempo de transferência mais o tempo de execução na GPU, como definido por:

$$T_{CPU} > T_{overhead} + T_{GPU}. \quad (3.21)$$

A efetividade desse ganho depende tanto da arquitetura de hardware utilizada quanto do grau de paralelização permitido pelo algoritmo de processamento.

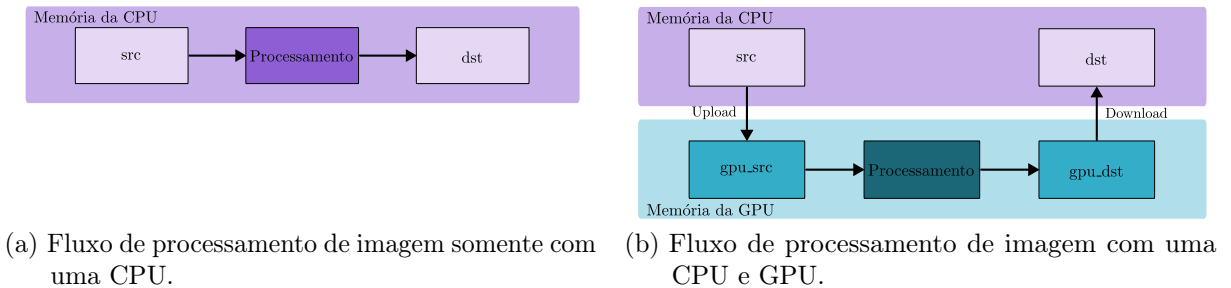


Figura 16 – Ilustração do Processamento de Imagem usando CPU e GPU.

3.5 Fusão Sensorial

O filtro de Kalman é um algoritmo recursivo que utiliza um conjunto de equações e entradas de dados sucessivas para estimar o estado de sistemas dinâmicos modelados por distribuições gaussianas, mesmo quando as medições contêm erros imprevisíveis, ruído aleatório, incertezas ou variações. Devido à sua capacidade de combinar medições ruidosas com modelos preditivos, o filtro é amplamente adotado em áreas como navegação, rastreamento e controle.

3.5.1 Filtro de Kalman

Neste contexto, a crença sobre o estado do sistema em cada instante de tempo é representada por uma distribuição gaussiana caracterizada por sua média X_t e sua covariância P_t . O filtro de Kalman opera periodicamente em duas etapas principais, predição e correção, conforme ilustrado na Figura 17.

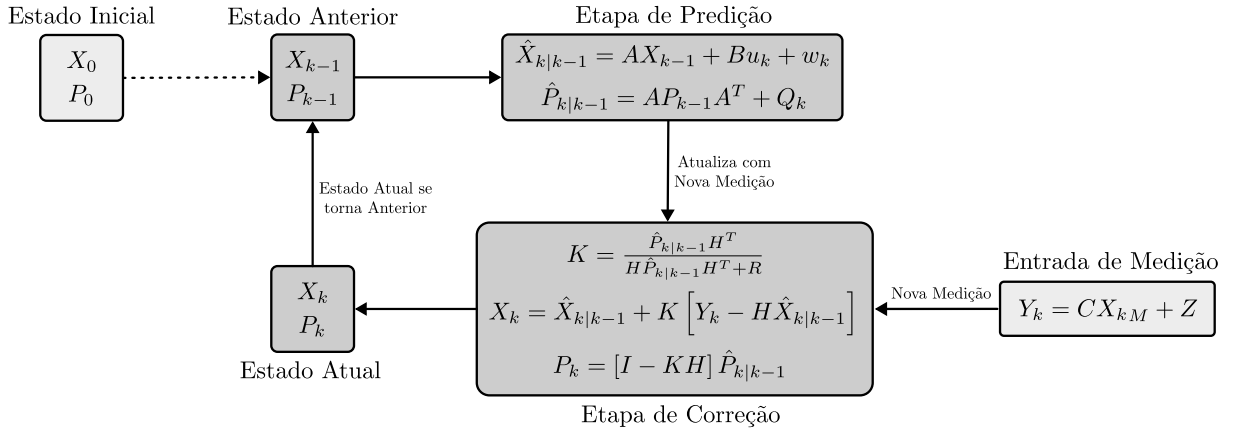


Figura 17 – Fluxo de Operações do Filtro de Kalman.

O funcionamento do filtro de Kalman pode ser descrito por meio das seguintes etapas:

1. Estado Inicial

O filtro é iniciado com uma estimativa inicial do estado X_0 e sua incerteza associada P_0 . Essas estimativas podem ser obtidas a partir de conhecimento prévio ou definidas com base em suposições iniciais.

2. Estado Anterior

A partir da etapa inicial, assume-se que o sistema possui uma estimativa do estado em um instante anterior X_{k-1} e sua covariância P_{k-1} , que servirão como entrada para a próxima iteração do filtro.

3. Etapa de Predição

O filtro projeta a estimativa do estado futuro com base no modelo dinâmico do sistema e na entrada de controle u_k :

$$\hat{X}_{k|k-1} = AX_{k-1} + Bu_k + w_k, \quad (3.22)$$

$$\hat{P}_{k|k-1} = AP_{k-1}A^T + Q_k, \quad (3.23)$$

onde w_k representa o ruído do processo, assumido gaussiano com média zero e covariância Q_k .

4. Entrada de Medição

A medição Y_k é obtida conforme o modelo de observação linear:

$$Y_k = CX_{kM} + Z, \quad (3.24)$$

sendo C a matriz de observação e Z o ruído de medição, com distribuição gaussiana de média zero e covariância R .

5. Etapa de Correção

Nesta etapa, o filtro ajusta a estimativa predita com base na nova medição Y_k , conforme as seguintes operações:

- *Ganho de Kalman:*

$$K = \hat{P}_{k|k-1} H^\top (H \hat{P}_{k|k-1} H^\top + R)^{-1}. \quad (3.25)$$

- *Inovação (diferença entre medição e previsão):*

$$Y_k - H \hat{X}_{k|k-1}. \quad (3.26)$$

- *Atualização do estado estimado:*

$$X_k = \hat{X}_{k|k-1} + K(Y_k - H \hat{X}_{k|k-1}). \quad (3.27)$$

- *Atualização da covariância do erro:*

$$P_k = (I - KH) \hat{P}_{k|k-1}. \quad (3.28)$$

O ganho de Kalman K determina o quanto a medição deve influenciar a estimativa final. A inovação representa a discrepância entre a medição observada e a prevista pelo modelo. Ambas são fundamentais para ajustar a crença sobre o estado do sistema.

6. Iteração do Processo

Ao final da etapa de correção, o estado atualizado X_k e a covariância P_k passam a ser considerados como os novos valores anteriores, reiniciando o processo para o próximo instante de tempo $k + 1$.

3.5.2 Filtro de Kalman Estendido (EKF)

Embora o filtro de Kalman clássico seja adequado para sistemas lineares com ruído gaussiano, muitos sistemas reais apresentam dinâmicas não lineares. O Filtro de Kalman Estendido é uma generalização que permite tratar essas não linearidades por meio da linearização local das funções que descrevem o sistema.

No EKF, as equações lineares da etapa de predição do estado e da medição são substituídas por funções não lineares $g(\cdot)$ e $h(\cdot)$:

- A predição do estado é dada por: $\hat{X}_{k|k-1} = g(u_k, X_{k-1})$
- A predição da medição é dada por: $\hat{Y}_k = h(\tilde{X}_{k|k-1})$

As matrizes A , B e C utilizadas no filtro linear são substituídas pelas jacobianas das funções g e h avaliadas nos pontos de operação atuais:

- G_k : Jacobiana de $g(u_k, X_{k-1})$ com relação ao estado
- H_k : Jacobiana de $h(\hat{X}_{k|k-1})$ com relação ao estado

O restante da estrutura do filtro é mantido, incluindo o cálculo do ganho de Kalman, da inovação e da atualização da covariância. A seguir, apresenta-se o algoritmo do EKF.

Algorithm 1 Filtro de Kalman Estendido (EKF)

- 1: **Entrada:** $X_{k-1}, P_{k-1}, u_k, Y_k$
 - 2: **Predição:**
 - 3: $\hat{X}_{k|k-1} = g(u_k, X_{k-1})$
 - 4: $\hat{P}_{k|k-1} = G_k P_{k-1} G_k^\top + Q_k$
 - 5: **Correção:**
 - 6: $K = \hat{P}_{k|k-1} H_k^\top (H_k \hat{P}_{k|k-1} H_k^\top + R)^{-1}$
 - 7: $X_k = \hat{X}_{k|k-1} + K[Y_k - h(\hat{X}_{k|k-1})]$
 - 8: $P_k = (I - KH_k)\hat{P}_{k|k-1}$
 - 9: **Saída:** X_k, P_k
-

Capítulo 4

Metodologia

Neste capítulo, é apresentada a metodologia desenvolvida para a estimação da pose de veículos com rodas em ambientes onde o sinal de GNSS pode ser limitado ou ausente. A Figura 18 apresenta o diagrama da estratégia de fusão sensorial adotada. Nessa abordagem, uma câmera monocular orientada para o solo e rigidamente acoplada à estrutura do veículo, captura imagens que são utilizadas para calcular o fluxo óptico e estimar a odometria visual. Essas informações são combinadas aos dados de uma unidade de medição inercial (IMU) e à odometria de rodas por meio de um filtro de Kalman, com o objetivo de aprimorar a estimação da pose. O sistema fornece como saída a odometria resultante da fusão sensorial.

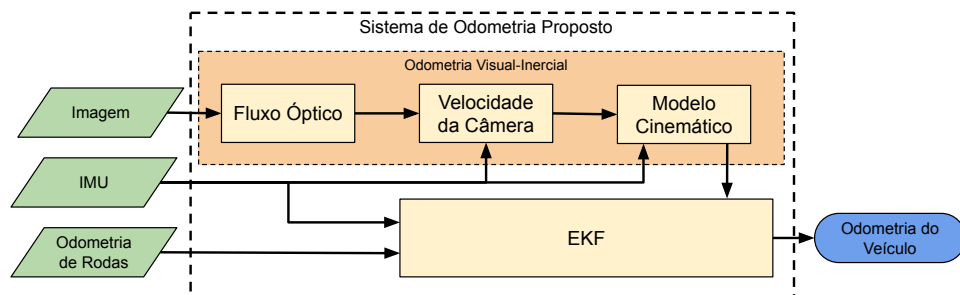


Figura 18 – Diagrama geral do sistema de odometria proposto.

As seções seguintes detalham as etapas do método proposto, organizadas conforme o fluxograma apresentado anteriormente. Inicialmente, é descrita a etapa de odometria visual baseada em fluxo óptico. Em seguida, é abordada a odometria de rodas. Por fim, apresenta-se a fusão sensorial por meio de um Filtro de Kalman Estendido (EKF), responsável por integrar as estimativas provenientes da odometria visual e da odometria de rodas com as medições inerciais da IMU.

4.1 Localização com Odometria Visual por Fluxo Óptico

Considerando que o veículo se desloca sobre uma superfície plana, admite-se que sua movimentação ocorre apenas no plano horizontal, o que permite restringir os graus de liberdade do movimento da câmera a translações nos eixos x e y e rotação em torno do eixo z , no referencial da câmera $\{\mathcal{C}\}$. A partir dessa premissa, é possível simplificar o modelo que relaciona o movimento da câmera com o fluxo óptico observado, dado pela Equação (3.19).

Esse modelo baseia-se na equação do fluxo óptico derivada da projeção de pontos 3D no plano da imagem, e descreve como as velocidades lineares e angulares da câmera afetam o deslocamento dos pixels entre quadros consecutivos. Quando o movimento é restrito ao plano, como no caso de uma câmera rigidamente fixada a um veículo terrestre voltada para o solo, a equação geral pode ser reduzida à forma planar, conforme mostrado na Equação (4.1):

$$\begin{bmatrix} \dot{u} \\ \dot{v} \end{bmatrix} = \begin{bmatrix} 0 & \omega_{\mathcal{C},z} & -\frac{f'}{Z_p} & 0 \\ -\omega_{\mathcal{C},z} & 0 & 0 & -\frac{f'}{Z_p} \end{bmatrix} \begin{bmatrix} \boldsymbol{\rho} \\ \bar{\mathbf{v}}_{\mathcal{C}} \end{bmatrix}, \quad (4.1)$$

em que $\boldsymbol{\rho} = [\bar{u} \ \bar{v}]^T$ representa a posição do ponto em relação ao centro óptico da imagem, $\bar{\mathbf{v}}_{\mathcal{C}} = [v_{\mathcal{C},x} \ v_{\mathcal{C},y}]^T$ corresponde às velocidades de translação no plano e $\omega_{\mathcal{C},z}$ à rotação em torno do eixo vertical da câmera.

A Equação (4.1) pode ser reorganizada para isolar as velocidades de translação da câmera, conforme mostrado na Equação (4.2):

$$\begin{bmatrix} v_{\mathcal{C},x} \\ v_{\mathcal{C},y} \end{bmatrix} = \begin{bmatrix} 0 & \frac{Z_p \omega_{\mathcal{C},z}}{f'} & -\frac{Z_p}{f'} & 0 \\ -\frac{Z_p \omega_{\mathcal{C},z}}{f'} & 0 & 0 & -\frac{Z_p}{f'} \end{bmatrix} \begin{bmatrix} \boldsymbol{\rho} \\ \dot{\boldsymbol{\rho}} \end{bmatrix}, \quad (4.2)$$

onde $\dot{\boldsymbol{\rho}} = [\dot{u} \ \dot{v}]^T$ são às componentes do fluxo óptico estimado no ponto.

Nessa dissertação, o fluxo óptico estimado por Farnebäck é utilizado como entrada na Equação (4.2), permitindo a estimação das velocidades de translação da câmera no plano. Esse método calcula um fluxo óptico denso por meio do ajuste de polinômios quadráticos em pequenas vizinhanças da imagem e da comparação dos coeficientes entre quadros consecutivos, o que proporciona estimativas mais suaves de deslocamento. Embora o método de Farnebäck seja aplicado a toda a imagem, apenas uma região central é considerada no cálculo da odometria visual, com o objetivo de reduzir os efeitos de distorções nas bordas, que são comuns em métodos densos, e concentrar a análise na área de maior consistência do movimento. A Equação (4.2) é aplicada ponto a ponto para estimar o vetor de velocidade da câmera correspondente a cada pixel da região selecionada.

Em seguida, as velocidades $v_{C,x}$ e $v_{C,y}$ são obtidas por meio da média das estimativas individuais do fluxo. As velocidades da câmera com respeito ao sistema de coordenadas inercial de referência são dadas pela seguinte equação:

$$\begin{bmatrix} \dot{x}_{C,k}^{\mathcal{I}} \\ \dot{y}_{C,k}^{\mathcal{I}} \\ \dot{\psi}_{C,k}^{\mathcal{I}} \end{bmatrix} = \begin{bmatrix} \cos(\theta_{C,k-1}) & -\sin(\theta_{C,k-1}) & 0 \\ \sin(\theta_{C,k-1}) & \cos(\theta_{C,k-1}) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} v_{C,x;k-1} \\ v_{C,y;k-1} \\ \omega_{C,z;k-1} \end{bmatrix}. \quad (4.3)$$

4.2 Localização com Odometria de Rodas

A odometria das rodas é uma das técnicas mais simples para estimar a distância percorrida e a pose atual de um veículo ou de um robô, a partir das velocidades das rodas [Liao et al., 2019]. Essa estratégia usa modelos cinemáticos e medições fornecidas por sensores, geralmente *encoders* acoplados aos eixos das rodas, para calcular o deslocamento do sistema.

O modelo cinemático do unicyclo, representado na Equação (3.7), descreve de forma simplificada o movimento planar de veículos com rodas, dadas as velocidades linear $v_{\mathcal{V}}$ e angular $\omega_{\mathcal{V}}$ como entradas. Esse modelo é frequentemente empregado como aproximação para outros tipos de plataformas móveis, como veículos com locomoção por *Skid-steering* ou com geometria de direção do tipo *Ackermann*, desde que seja possível controlar diretamente as velocidades $v_{\mathcal{V}}$ e $\omega_{\mathcal{V}}$.

Como não há uma forma direta de medir instantaneamente a posição de um veículo, a estimativa da pose é realizada por meio da integração das velocidades fornecidas pelo modelo. A equação abaixo pode ser utilizada para esse fim, aplicando o método de integração de Euler, com avaliação em instantes de tempo discretos $t_k = k\Delta t$, $k = 1, 2, \dots$, sendo Δt o intervalo de amostragem. No entanto, essa estratégia está sujeita a erros numéricos de integração, que tendem a se acumular ao longo do tempo, comprometendo a precisão da odometria estimada:

$$\begin{bmatrix} x_{\mathcal{V},k}^{\mathcal{I}} \\ y_{\mathcal{V},k}^{\mathcal{I}} \\ \theta_{\mathcal{V},k} \end{bmatrix} = \begin{bmatrix} x_{\mathcal{V},k-1}^{\mathcal{I}} \\ y_{\mathcal{V},k-1}^{\mathcal{I}} \\ \theta_{\mathcal{V},k-1} \end{bmatrix} + \begin{bmatrix} \cos(\theta_{\mathcal{V},k-1}) & 0 \\ \sin(\theta_{\mathcal{V},k-1}) & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} v_{\mathcal{V},k-1} \\ \omega_{\mathcal{V},k-1} \end{bmatrix} \Delta t. \quad (4.4)$$

4.3 Localização com Filtro de Kalman Estendido

Para fundir os dados da odometria visual com a odometria de rodas e IMU, foi implementado um Filtro de Kalman Estendido com um modelo com 12 DoF. Os estados do filtro são definidos por $\mathbf{x} = [\mathbf{r}^{\mathbf{T}} \ \boldsymbol{\vartheta}^{\mathbf{T}} \ \mathbf{v}^{\mathbf{T}} \ \boldsymbol{\omega}^{\mathbf{T}}]^{\mathbf{T}}$, no qual $\mathbf{r} = [x \ y \ z]^{\mathbf{T}}$ é o vetor de posição, $\boldsymbol{\vartheta} = [\phi \ \theta \ \psi]^{\mathbf{T}}$ é o vetor da representação mínima dada pelos ângulos *roll*, *pitch*, *yaw*,

$\mathbf{v} = [v_x \ v_y \ v_z]^T$ é o vetor de velocidades lineares e $\boldsymbol{\omega} = [\omega_x \ \omega_y \ \omega_z]^T$ é o vetor de velocidades angulares. Sendo assim, a predição do estado a priori $\mathbf{X}_{k|k-1}$ é definida como um modelo de um bloco deslizante dado por:

$$\hat{\mathbf{X}}_{k|k-1} = f(\hat{\mathbf{X}}_{k-1}, \Delta t) = \begin{bmatrix} \mathbf{r}_{k-1} + \mathbf{R}(\boldsymbol{\vartheta}_{k-1})\mathbf{v}_{k-1}\Delta t \\ \boldsymbol{\vartheta}_{k-1} + \mathbf{J}(\boldsymbol{\vartheta}_{k-1})\boldsymbol{\omega}_{k-1}\Delta t \\ \mathbf{v}_{k-1} \\ \boldsymbol{\omega}_{k-1} \end{bmatrix}, \quad (4.5)$$

na qual $\mathbf{R}(\boldsymbol{\vartheta}_{k-1})$ é a matriz de rotação definida pela componente $\boldsymbol{\vartheta}$, $\mathbf{J}(\boldsymbol{\vartheta})$ é a matriz jacobiana da orientação $\boldsymbol{\vartheta}$ que mapeia as velocidades angulares para a derivada dos ângulos de Euler e Δt é o passo de tempo amostral. Já a matriz de covariâncias da transição de estados é definida como:

$$\mathbf{Q}_f = \sigma_f \mathbf{I}_{12}, \quad (4.6)$$

onde $\mathbf{I}_{12} \in \mathbb{R}^{12 \times 12}$ é uma matriz identidade e $\sigma_f = 0.001$, definido empiricamente.

A covariância a priori $\hat{\mathbf{P}}_{k|k-1}$ calculada na etapa de predição é dada por:

$$\hat{\mathbf{P}}_{k|k-1} \leftarrow \mathbf{G}\hat{\mathbf{P}}_{k-1}\mathbf{G}^T + \mathbf{Q}_f, \quad (4.7)$$

onde \mathbf{G} é a matriz jacobiana da função $g(\hat{\mathbf{X}}, \Delta t)$ definida no estado $\hat{\mathbf{X}}_{k-1}$ estimada na última iteração do filtro.

As etapas de correções são realizadas conforme as medidas da odometria das rodas e IMU são recebidas. Assim sendo, o estado a posteriori $\hat{\mathbf{X}}_{k|k}$ é determinado por:

$$\hat{\mathbf{X}}_{k|k} = \hat{\mathbf{X}}_{k|k-1} + \mathbf{K}_i (\mathbf{Y}_i - h_i(\mathbf{X}_{k|k-1})), \quad (4.8)$$

onde $h_i(\mathbf{X}_{k|k-1})$ é o modelo da medida i dado o estado $\mathbf{X}_{k|k-1}$ e \mathbf{K}_i é o ganho de Kalman:

$$\mathbf{K}_i = \hat{\mathbf{P}}_{k-1}\mathbf{H}^T \left(\mathbf{H}\hat{\mathbf{P}}_{k-1}\mathbf{H}^T + \mathbf{Q}_i \right)^{-1}, \quad (4.9)$$

no qual \mathbf{H} é a matriz Jacobiana da função $h(\mathbf{X}_{k|k-1})$.

Assim sendo, a covariância a posteriori $\hat{\mathbf{P}}_{k|k}$ do estado é dada por:

$$\hat{\mathbf{P}}_{k|k} = (\mathbf{I}_{12} - \mathbf{K}_i\mathbf{H})\hat{\mathbf{P}}_{k|k-1}. \quad (4.10)$$

Dessa forma, na etapa de correção da odometria de rodas a medida é definida como $\mathbf{Y}_W = [v_{w,x} \ \omega_{w,z}]^T$, e o modelo da medida é dado por:

$$h_W(\hat{\mathbf{X}}) = \begin{bmatrix} v_x \\ \omega_z \end{bmatrix}. \quad (4.11)$$

Semelhante ao apresentado em [Libby and Kantor, 2011], duas matrizes de covariância foram consideradas, uma para movimentos lineares e outra para movimentos

angulares. Portanto, a matriz de covariância da medição da odometria da roda é definida como:

$$\mathbf{Q}_W = \text{diag} \left(\sigma_{v_{w,x}}^2, \sigma_{\omega_{w,z}}^2 \right), \quad (4.12)$$

no qual $\sigma_{v_{w,x}}^2$ é a variância da velocidade linear $v_{w,x}$ e $\sigma_{\omega_{w,z}}^2$ é a variância da velocidade angular $\omega_{w,z}$.

Na etapa de correção da IMU a medida é definida como $\mathbf{Y}_I = [\omega_{I,x} \ \omega_{I,y} \ \omega_{I,z}]^T$. Dessa forma, o modelo da medida é dado por:

$$h_I(\hat{\mathbf{X}}) = \boldsymbol{\vartheta}, \quad (4.13)$$

e a matriz de covariância é:

$$\mathbf{Q}_I = \text{diag} \left(\sigma_{\phi_I}^2, \sigma_{\theta_I}^2, \sigma_{\psi_I}^2 \right), \quad (4.14)$$

na qual $\sigma_{\phi_I}^2$, $\sigma_{\theta_I}^2$ e $\sigma_{\psi_I}^2$ são as variâncias dos ângulos de Euler ϕ_I , θ_I e ψ_I .

Por fim, na etapa de correção da odometria visual, a medida é definida por $\mathbf{Y}_V = [v_{V,x} \ v_{V,y} \ \omega_{V,z}]^T$ e o modelo da medida é dado por:

$$h_V(\hat{\mathbf{X}}) = \begin{bmatrix} v_x \\ v_y \\ \omega_z \end{bmatrix}. \quad (4.15)$$

A matriz de covariâncias da medida da odometria visual é definida como:

$$\mathbf{Q}_V = \text{diag} \left(\sigma_{v_{V,x}}^2, \sigma_{v_{V,y}}^2, \sigma_{\omega_{V,z}}^2 \right), \quad (4.16)$$

onde $\sigma_{v_{V,x}}^2$, $\sigma_{v_{V,y}}^2$ e $\sigma_{\omega_{V,z}}^2$ são as variâncias das medidas $v_{V,x}$, $v_{V,y}$ e $\omega_{V,z}$ da odometria visual.

Capítulo 5

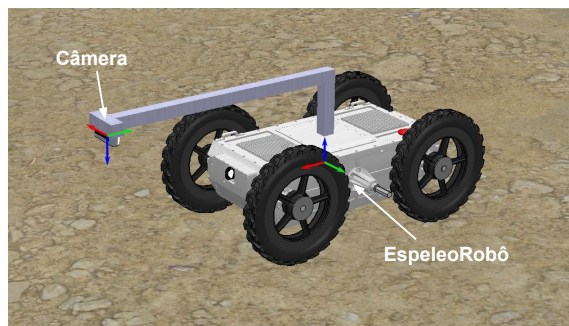
Arcabouço Experimental

Neste capítulo, são descritas as ferramentas, configurações e procedimentos adotados para a validação do sistema proposto. A validação foi conduzida por meio de simulações no ambiente *CoppeliaSim*, no qual dois modelos de veículos virtuais previamente disponíveis foram utilizados. O simulador foi integrado ao *framework Robot Operating System* (ROS), na versão *Noetic Ninjemys*, executada sobre o sistema operacional Ubuntu 20.04 LTS. As rotinas de processamento de imagem foram desenvolvidas com a biblioteca OpenCV 4.9, com suporte à aceleração por GPU via CUDA 12.0. As simulações foram executadas em uma estação de trabalho equipada com um processador Intel i9 14900K e uma GPU Nvidia RTX 4090.

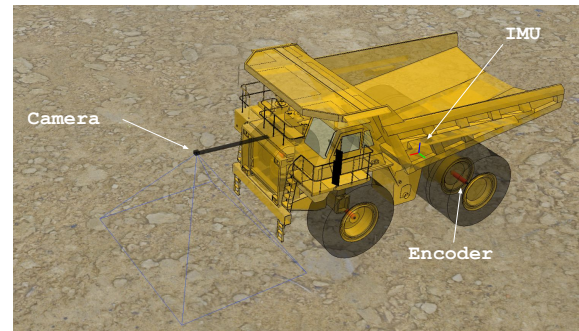
5.1 Plataformas Móveis

As simulações realizadas no ambiente CoppeliaSim utilizaram dois modelos de plataformas móveis. O primeiro é o EspeleoRobô, uma plataforma terrestre de menor porte disponível para experimentos físicos. Embora não represente o veículo final de aplicação, ele será empregado em etapas posteriores para validação prática dos algoritmos desenvolvidos, servindo como intermediário entre a simulação e os testes em campo. O segundo modelo é uma representação virtual do caminhão de mineração CAT793F, veículo de grande porte utilizado em operações a céu aberto. Este modelo foi adotado como base para o desenvolvimento do sistema de odometria, por refletir de forma mais fiel as características operacionais e o contexto real de aplicação.

Os modelos virtuais do EspeleoRobô e do caminhão CAT793F foram configurados de forma semelhante no ambiente CoppeliaSim, permitindo a aplicação de um mesmo sistema de odometria com pequenas adaptações. Ambos os veículos possuem uma câmera rigidamente acoplada e orientada para o solo, simulada com base no modelo *Intel RealSense D435i*, que possui uma frequência de amostragem de 30 fps e inclui uma unidade inercial integrada. Contudo, neste estudo foi empregado apenas o modo monocular, sem emprego



(a) Configuração do EspeleoRobô com câmera direcionada para o chão.



(b) Configuração do caminhão CAT793F com câmera direcionada para o chão.

Figura 19 – Ilustração das plataformas móveis no ambiente CoppeliaSim.

dos dados de profundidade do sistema RGB-D, utilizando imagens com resolução de 640×480 pixels. As informações de calibração intrínseca e extrínseca da câmera foram obtidas por meio de uma simulação dedicada para calibração, assegurando que os parâmetros empregados refletissem o modelo virtual adotado. No modelo do CAT793F, ela foi posicionada a uma altura de 1,87m em relação ao solo, enquanto no EspeleoRobô a altura adotada foi de 0,54m. Além disso, cada plataforma conta com uma IMU adicional posicionada em seu centro de rotação. As informações de sensores e atuadores são disponibilizadas por meio de tópicos ROS, viabilizando a integração com os algoritmos desenvolvidos. Apesar das similaridades nos sensores e na estrutura da simulação, os dois modelos diferem em sua arquitetura de locomoção. O EspeleoRobô adota um modelo *skid-steering*, enquanto o CAT793F opera com locomoção do tipo *Ackermann*. Essa distinção exige tratamentos diferentes nos modelos cinemáticos empregados em cada caso. Para compor o ambiente de simulação, foi utilizado um terreno gerado a partir de texturas obtidas no repositório PolyHaven¹, com o objetivo de representar os ambientes de mineração.

É importante destacar que, nas simulações, não foram considerados os efeitos de vibração, ruídos dos sensores e demais perturbações inerentes ao ambiente real, como irregularidades do terreno, variações de iluminação ou interferências eletromagnéticas. Assim, os resultados obtidos representam um cenário idealizado, cujo desempenho poderá sofrer variações quando o sistema for testado em condições reais de operação.

5.2 Implementações de Algoritmos em Pacotes do ROS

A arquitetura do sistema foi organizada em três pacotes distintos, desenvolvidos no ambiente ROS. O primeiro trata da odometria visual, baseada em fluxo óptico calculado a partir de imagens monoculares voltadas para o solo. O segundo implementa a odometria de rodas, por meio de dados provenientes dos *encoders* acoplados ao sistema de tração para

¹ <https://polyhaven.com/textures/terrain>

estimar o deslocamento do veículo. O terceiro pacote corresponde ao Filtro de Kalman Estendido, responsável pela fusão sensorial entre as estimativas visuais, os dados da odometria de rodas e as medições inerciais.

5.2.1 Implementação da Odometria Visual por Fluxo Óptico

O pacote do ROS de odometria visual por fluxo óptico foi implementado em C++ e possui basicamente 3 módulos: Odometria Visual por Fluxo Óptico, Cálculo das Velocidades da Câmera e Aplicação do Modelo Cinemático, conforme Figura 20. As entradas do algoritmo são as imagens da câmera RGB e os dados da IMU. Já as saídas são a odometria da câmera composta pela pose e velocidades linear e angular e as respectivas covariâncias.

O fluxo óptico é calculado usando o método de Gunnar Farneback do OpenCV, integrado ao CUDA para processamento em GPU. Os parâmetros do algoritmo foram ajustados manualmente para este trabalho, sem a adoção de um padrão pré-estabelecido, de modo a atender às características específicas do cenário analisado. Devido ao processamento paralelo proporcionado pelas GPUs, os tempos de execução obtidos são significativamente menores que aqueles verificados em CPUs [Cervera, 2020]. Para aproveitar essa capacidade de processamento, o cálculo das velocidades da câmera, descrito na Seção 4.1, também foi implementado em GPU. Em seguida, o modelo cinemático da Equação (4.3) é utilizado para calcular as velocidades da câmera.

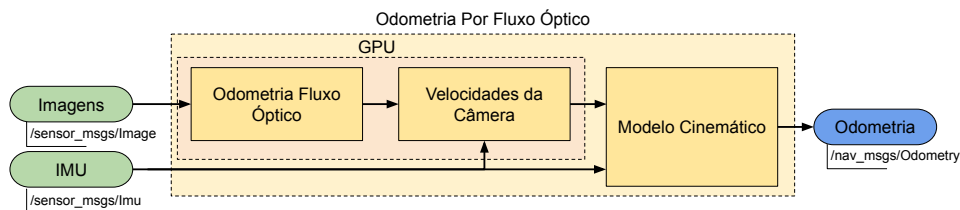


Figura 20 – Diagrama geral do algoritmo de odometria por fluxo óptico.

5.2.2 Implementação da Odometria de Rodas

O pacote de odometria de rodas foi implementado na linguagem de programação *Python*, com adaptações específicas para cada uma das arquiteturas cinemáticas utilizadas nos modelos simulados. No caso do EspeleoRobô, a abordagem segue uma estrutura compatível com sua configuração *skid-steering*. As entradas consistem nas leituras dos *encoders* das rodas do flanco esquerdo e direito, utilizadas para o cálculo das velocidades correspondentes. A partir dessas velocidades, o algoritmo estima a pose do robô em relação a um referencial inercial por meio do modelo cinemático *skid-steering*. A saída do pacote inclui a pose, as velocidades linear e angular, bem como as covariâncias associadas a essas estimativas.

Para o caminhão CAT793F, as entradas do algoritmo são as leituras dos *encoders* empregadas para o cálculo das velocidades das rodas traseiras e da direção das rodas dianteiras. Com base nessas informações, determinam-se as velocidades linear e angular do veículo e, por meio do modelo cinemático Ackermann, estima-se a pose em relação a um referencial inercial. A saída do algoritmo compreende a pose atual do veículo, suas velocidades linear no eixo x e angular no eixo z , além das covariâncias correspondentes.

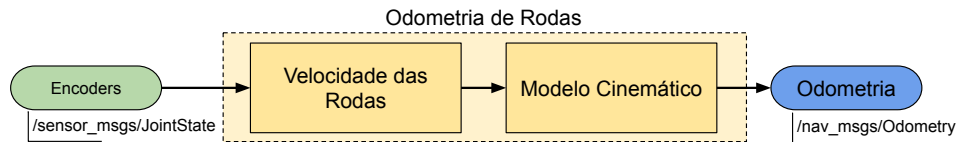


Figura 21 – Diagrama geral do algoritmo de odometria de rodas.

5.2.3 Implementação do Filtro de Kalman Estendido

O Filtro de Kalman Estendido foi implementado como um pacote do ROS utilizando a linguagem de programação C++. O pacote recebe como entradas os dados provenientes da odometria visual, da odometria de rodas e da unidade inercial (IMU). Como a odometria visual é expressa no referencial inercial da câmera, é necessário transformá-la para o referencial inercial do veículo antes de ser utilizada no filtro. Essa conversão é realizada por meio de uma transformação homogênea entre os sistemas de coordenadas da câmera e do corpo do veículo, dada pela Equação (3.4).

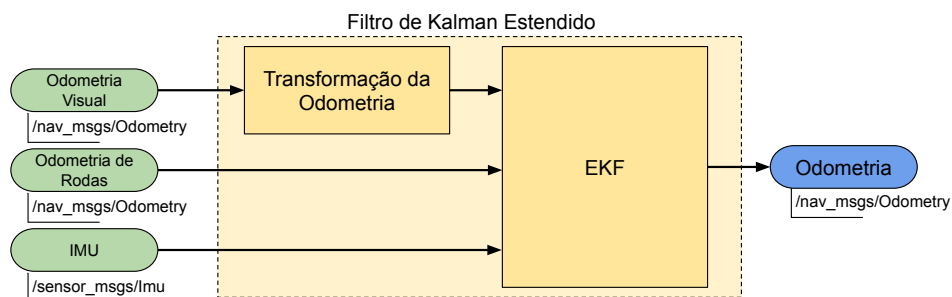


Figura 22 – Diagrama geral do algoritmo do Filtro de Kalman Estendido.

A Figura 22 apresenta o diagrama geral de funcionamento do pacote. A saída do filtro é a odometria combinada, composta pela estimativa da pose e das velocidades do veículo, com base na fusão das diferentes fontes de informação disponíveis. O pacote inclui um arquivo de configurações que permite habilitar ou desabilitar entradas específicas e ajustar parâmetros do filtro, como a frequência de publicação das mensagens. Essa estrutura possibilita a aplicação do mesmo algoritmo para diferentes configurações de entrada, como a fusão entre odometria de rodas e IMU ou a fusão completa com odometria visual-inercial. A frequência de predição foi definida em 200 Hz para ambos os casos, enquanto a frequência de correção variou conforme as limitações impostas pela simulação:

20 Hz para o modelo do EspeleoRobô e 60 Hz para o modelo do CAT793F. Observou-se que essa diferença de frequências foi influenciada pelo modelo utilizado, embora as causas exatas dessa variação não tenham sido determinadas no presente estudo.

5.3 Controle de Navegação e Definição de Velocidades

A navegação dos veículos ao longo dos percursos definidos nas simulações foi realizada por meio de um pacote ROS baseado no algoritmo de campos vetoriais artificiais, conforme proposto por [Rezende et al. \[2021\]](#). Para garantir o funcionamento adequado da odometria visual, as velocidades máximas dos veículos foram determinadas com base na equação de sobreposição entre quadros consecutivos, apresentada na Seção 3.3.2. Essa equação leva em consideração a velocidade de deslocamento v_{\max} , a altura da câmera em relação ao solo h , a frequência de aquisição das imagens f_s e o campo de visão α . Adotou-se como critério mínimo uma sobreposição de aproximadamente 70% entre imagens consecutivas, condição necessária para preservar a correspondência de pontos visuais usada no cálculo do fluxo óptico. Com base nessa restrição, a velocidade máxima foi definida como 0,3 m/s para o EspeleoRobô, cuja câmera está posicionada a 0,54 m de altura, e 10 km/h para o caminhão CAT793F, com câmera a 1,87 m do solo.

Capítulo 6

Resultados e Discussões

Este capítulo apresenta os resultados obtidos com a aplicação da metodologia proposta, além da análise dos experimentos conduzidos. Inicialmente, são descritas as abordagens de odometria consideradas e os critérios adotados para avaliação da precisão das estimativas. Em seguida, é explicada a simulação utilizada para o ajuste das covariâncias no Filtro de Kalman Estendido. Por fim, são apresentados os resultados das simulações de validação.

6.1 Abordagens de Odometria Consideradas

Para possibilitar a análise comparativa do sistema proposto, foram definidas quatro abordagens distintas de estimação de odometria, cada uma explorando diferentes combinações de sensores e estratégias de fusão para estimar a pose do veículo:

- **W**: Estimativa baseada unicamente na odometria de rodas, a partir do modelo cinemático e das velocidades medidas pelos atuadores do veículo;
- **VO+IMU**: Combinação direta entre a odometria visual, obtida a partir do fluxo óptico, e os dados da unidade de medição inercial (IMU), por meio de um modelo acoplado;
- **EKF W+IMU**: Fusão sensorial por filtro de Kalman estendido entre a odometria de rodas e a IMU;
- **EKF VO+W+IMU**: Abordagem completa que realiza a fusão da odometria visual-inercial, da odometria de rodas e da IMU em um Filtro de Kalman Estendido.

6.2 Técnicas de Avaliação Empregadas

A avaliação das abordagens de odometria implementadas foi conduzida com base em métricas quantitativas e análises estatísticas que permitiram comparar a acurácia e a consistência das trajetórias estimadas em relação ao *ground-truth* do simulador. As técnicas adotadas estão organizadas nas subseções a seguir.

6.2.1 Métricas de Erro Euclidiano

A avaliação da acurácia das estimativas de pose geradas pelas diferentes abordagens de odometria foi realizada com base na comparação direta com o *ground-truth* fornecido pelo simulador. Para quantificar os erros cometidos ao longo da trajetória, foram utilizadas três métricas adotadas na literatura: o erro médio (ME), o erro médio absoluto (MAE) e a raiz quadrada do erro médio (RMSE).

Considere N instantes de tempo discretos, nos quais a posição estimada do veículo é representada por $\hat{\mathbf{p}}_i = [\hat{x}_i, \hat{y}_i]^T$ e a posição de referência pelo simulador (ou *ground-truth*) é dada por $\mathbf{p}_i = [x_i, y_i]^T$. A distância euclidiana entre os pares de posições em cada instante é calculada por:

$$e_i = \|\mathbf{p}_i - \hat{\mathbf{p}}_i\| = \sqrt{(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2}. \quad (6.1)$$

Com base nessa medida de erro ponto a ponto, são definidas as seguintes métricas globais:

- **Erro médio (ME)**: representa a média aritmética dos erros.

$$\text{ME} = \frac{1}{N} \sum_{i=1}^N e_i. \quad (6.2)$$

- **Erro médio absoluto (MAE)**: mede a magnitude média dos erros, desconsiderando o sinal.

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |e_i|. \quad (6.3)$$

- **Raiz quadrada do erro médio (RMSE)**: atribui maior penalização a erros maiores.

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N e_i^2}. \quad (6.4)$$

Essas métricas foram aplicadas às trajetórias estimadas em relação ao *ground-truth* para cada uma das configurações testadas.

6.2.2 Análise Estatística dos Erros

Como complemento à avaliação numérica dos erros médios, foi realizada uma análise estatística das distribuições dos erros euclidianos ao longo do tempo. Para isso, foram utilizados histogramas e intervalos de confiança. Os histogramas foram construídos a partir dos valores de e_i definidos na Equação (6.1). Essa representação gráfica fornece uma compreensão intuitiva sobre o comportamento global de cada abordagem.

Além disso, foram aplicados intervalos estatísticos baseados em múltiplos do desvio padrão:

- σ_1 : intervalo em torno da média que abrange aproximadamente 68,3% dos dados;
- σ_2 : intervalo que cobre cerca de 95,4% dos erros em torno da média;
- σ_3 : intervalo que inclui aproximadamente 99,7% dos valores.

Esses limites foram calculados a partir da média \bar{e} e do desvio padrão s da amostra de erros, conforme:

$$\sigma_k = \bar{e} \pm k \cdot s \quad \text{com } k \in 1, 2, 3. \quad (6.5)$$

6.3 Simulação de Ajuste de Covariâncias

Com o objetivo de calibrar as matrizes de covariância utilizadas no Filtro de Kalman Estendido (EKF), foi conduzida uma simulação preliminar baseada em uma trajetória com formato de Lemniscata de Bernoulli. Essa curva foi escolhida por permitir a avaliação do sistema em diferentes condições de movimento, incluindo variações de curvatura e velocidade angular, além de possibilitar o retorno periódico ao mesmo ponto da trajetória, o que facilita a comparação entre estimativas e *ground-truth*. Durante essa simulação, o veículo percorreu duas voltas no caminho definido.

6.4 Simulação de Validação da Metodologia Proposta

Nesta seção, são apresentados os resultados obtidos com a aplicação das abordagens de odometria consideradas. As análises foram realizadas em dois cenários distintos. Para o EspeleoRobô, foi adotado um caminho derivado da pista *Red Bull Ring*. No caso do caminhão Caterpillar CAT793F, o trajeto seguiu um Caminho de Teste. Em ambas as simulações, os resultados das odometrias foram comparados com o *ground-truth* do simulador, por meio de métricas quantitativas e análises estatísticas.

6.4.1 Resultados da Simulação do EspeleoRobô

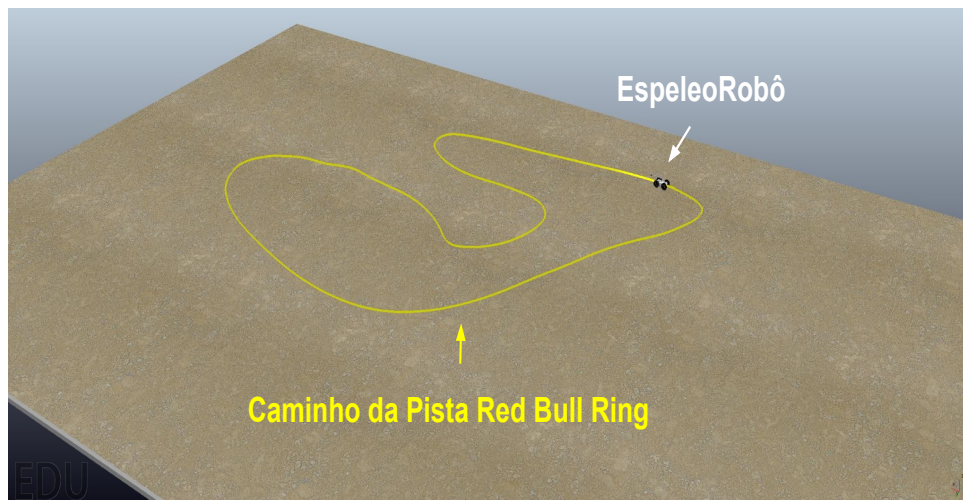


Figura 23 – EspeleoRobô e o Caminho da Pista *Red Bull Ring* no Simulador CoppeliaSim.

A simulação do EspeleoRobô foi realizada seguindo um caminho gerado a partir da pista *Red Bull Ring*, Figura 23, no qual o robô executou uma volta. Na Figura 24 são apresentadas as posições estimadas por cada uma das odometrias implementadas com o *ground-truth*.

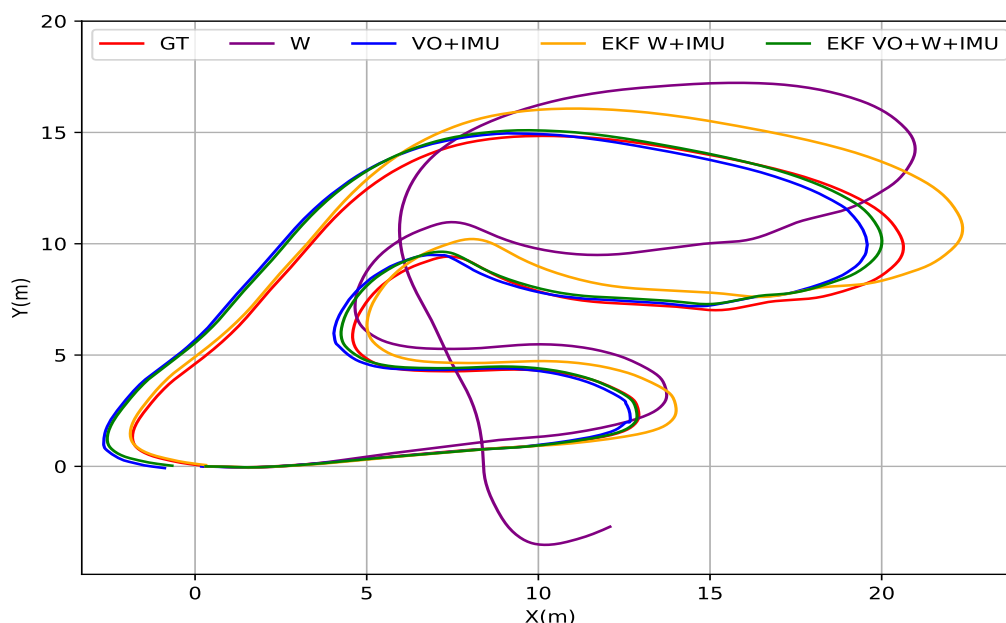


Figura 24 – Posições estimadas pelos algoritmos implementados com *ground-truth* para a simulação do caminho da Pista *Red Bull Ring*.

A Tabela 1 mostra os valores das métricas ME, MAE e RMSE. A metodologia proposta (EKF VO+W+IMU) mais uma vez apresentou os melhores resultados, com um

ME de 1,19 m, MAE de 0,56 m e RMSE de 0,64 m. Novamente seguida pela odometria visual e IMU com um ME de 1,237 m, MAE de 0,67 m e RMSE de 0,73 m. Já a odometria de rodas e IMU acumulou mais erros que resultaram em um ME de 1,93 m, MAE de 1,03 m e RMSE de 1,15 m. A odometria de rodas obteve os piores resultados assim como antes.

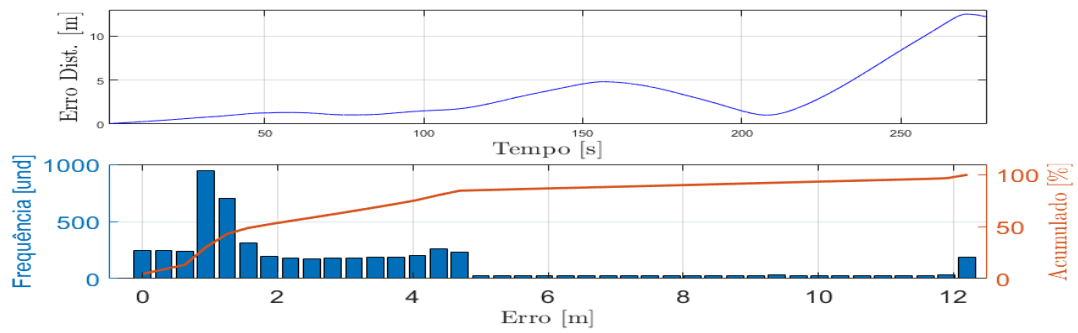
Os histogramas dos erros estão representados na Figura 25. Na Tabela 2 são apresentados os valores σ referente aos intervalos de confiança dos modelos das medidas. A metodologia proposta (EKF VO+W+IMU) possui um erro de posição inferior a 0,80 m para 68% dos dados (σ_1), inferior a 1,08 m para 95% dos dados (σ_2) e um erro inferior a 1,17 m para 99% dos dados (σ_3). Os resultados foram muito próximos da odometria visual e IMU que apresentou um erro de 0,87 m para σ_1 , 1,12 m para σ_2 e 1,21 m para σ_3 . Apesar da odometria de rodas possuir resultados ruins, quando combinado com a IMU usando o EKF apresentou um erro de 1.29 m para σ_1 , 1,86 m para σ_2 e 1,91 m para σ_3 .

Tabela 1 – Tabela de erros de posição referente à simulação do EspeleoRobô na Pista *Red Bull Ring*.

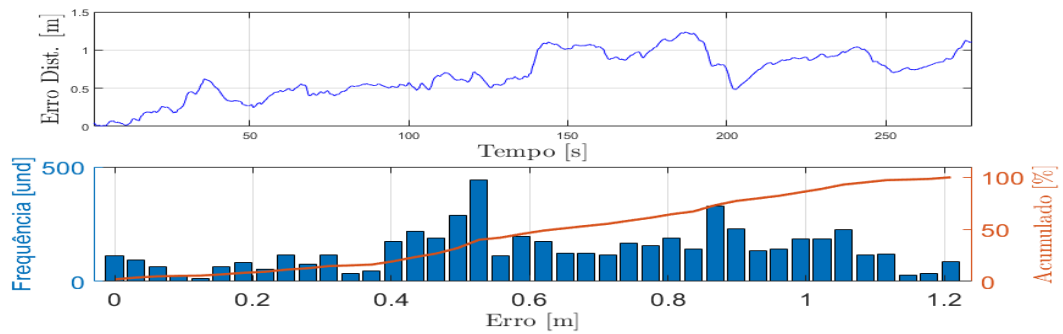
Método	EP(m)		
	ME	MAE	RMSE
W	12,515	3,240	4,483
VO+IMU	1,237	0,674	0,737
EKF W+IMU	1,931	1,032	1,155
EKF VO+W+IMU	1,190	0,562	0,638

Tabela 2 – Tabela dos valores do intervalo de confiança referente à simulação do EspeleoRobô na Pista *Red Bull Ring*.

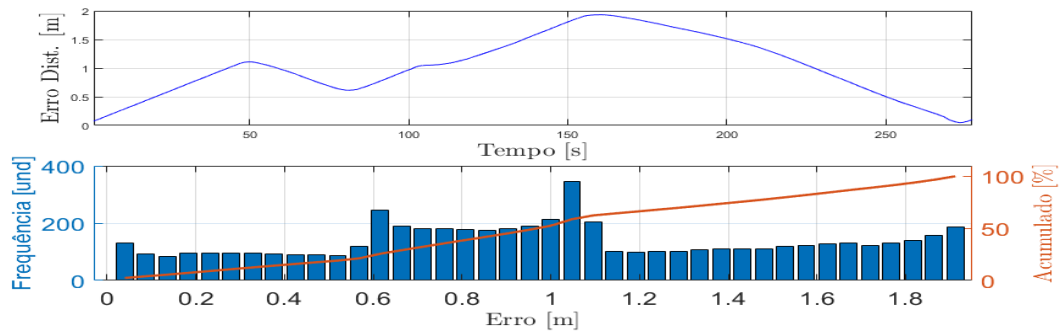
Método	σ_1	σ_2	σ_3
W	3,443	11,581	12,207
VO+IMU	0,868	1,116	1,209
EKF W+IMU	1,288	1,864	1,912
EKF VO+W+IMU	0,745	1,073	1,162



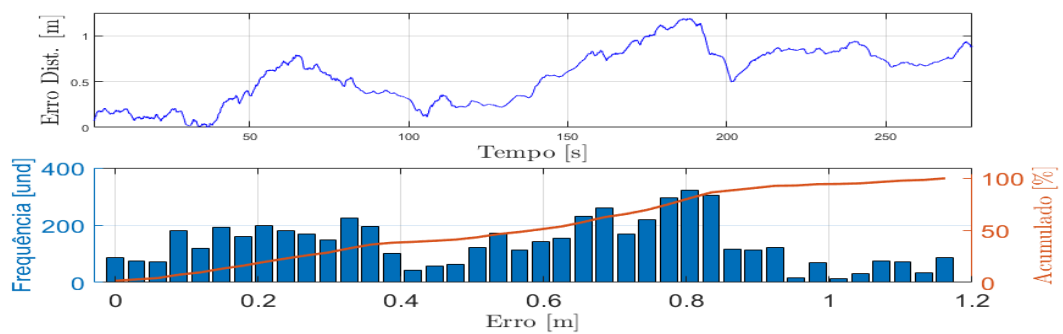
(a)



(b)



(c)



(d)

Figura 25 – Histogramas do erro de posição da simulação do EspeleoRobô na pista *Red Bull Ring* para: (a) Wheel, (b) VO+IMU, (c) EKF W+IMU e (d) EKF VO+W+IMU.

6.4.2 Resultados da Simulação do Caminhão Caterpillar CAT793F



Figura 26 – Caminhão Caterpillar CAT793F e o Caminho de Teste no Simulador CoppeliaSim.

A Figura 26 ilustra o caminhão Caterpillar CAT793F no ambiente simulado, juntamente com o caminho de teste definido para a validação. A trajetória consiste em segmentos de reta e curvas para ambos os lados, permitindo avaliar o desempenho das abordagens de odometria em diferentes condições de deslocamento. O veículo foi programado para realizar uma volta.

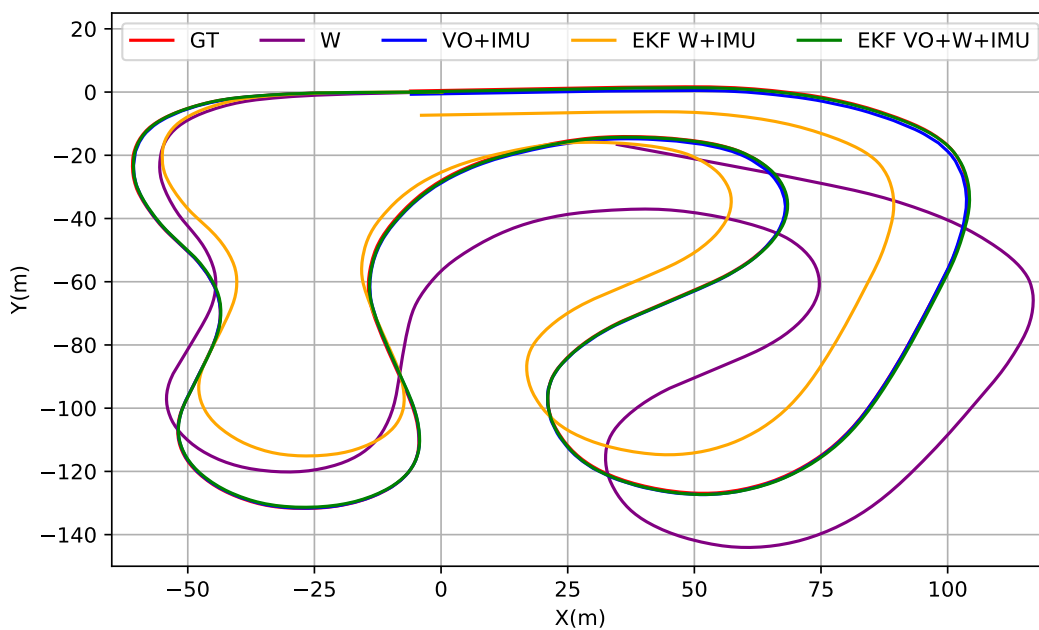


Figura 27 – Posições estimadas pelos algoritmos implementados com *ground-truth* para a simulação do CAT793F no Caminho de Teste.

A Figura 27 apresenta a comparação visual das trajetórias estimadas por cada abordagem em relação ao *ground-truth*. As estimativas de pose foram comparadas com a trajetória de referência utilizando a distância euclidiana como métrica de erro. Os resultados são apresentados na Tabela 3, que evidencia o melhor desempenho da abordagem proposta (EKF VO+W+IMU), com um erro médio (ME) de 1,09 m, erro médio absoluto (MAE) de 0,49 m e raiz quadrada do erro médio (RMSE) de 0,54 m. A odometria visual e IMU também apresentou bom desempenho, enquanto a combinação entre odometria de rodas e IMU resultou em erros mais elevados. A odometria de rodas isolada apresentou os piores resultados.

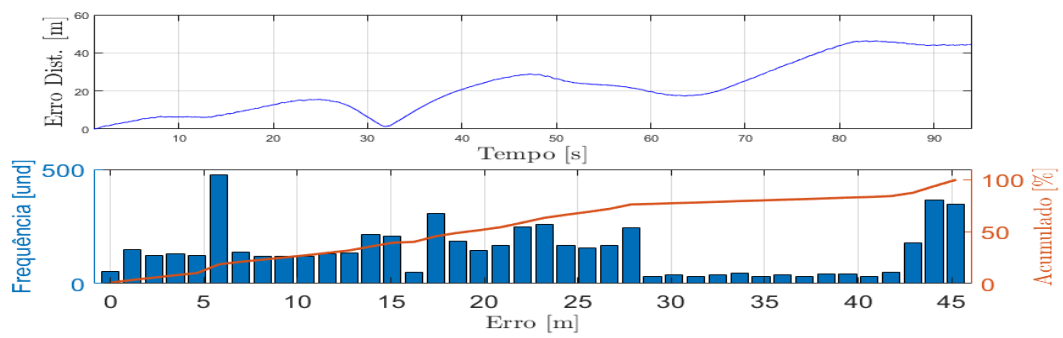
Os histogramas dos erros de posição são mostrados na Figura 28, assim como os gráficos que mostram a distância da posição estimada em relação ao *ground-truth* no decorrer do tempo. Os valores dos intervalos de confiança σ_1 , σ_2 e σ_3 são apresentados na Tabela 4. A metodologia proposta apresentou um erro inferior a 0,57 m em 68% das amostras, inferior a 0,82 m em 95% e inferior a 1,01 m em 99%, demonstrando alta precisão e consistência estatística.

Tabela 3 – Erros de posição da simulação com o caminhão CAT793F no Caminho de Teste.

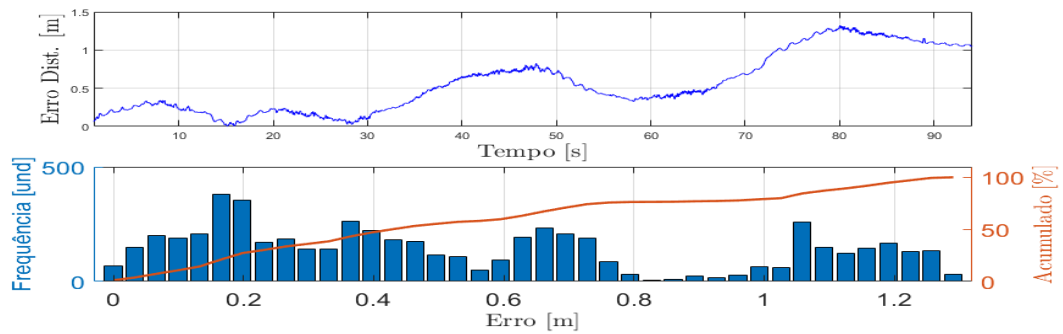
Método	EP(m)		
	ME	MAE	RMSE
W	46,308	21,764	25,647
VO+IMU	1,324	0,559	0,676
EKF W+IMU	16,994	10,544	11,247
EKF VO+W+IMU	1,089	0,494	0,537

Tabela 4 – Intervalos de confiança para a simulação do CAT793F no Caminho de Teste

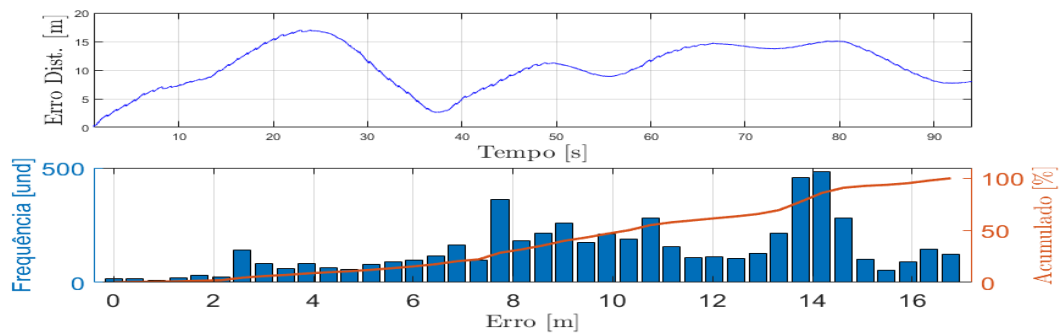
Método	σ_1	σ_2	σ_3
W	25,520	45,240	45,240
VO+IMU	0,695	1,225	1,291
EKF W+IMU	13,330	16,340	16,770
EKF VO+W+IMU	0,573	0,819	1,010



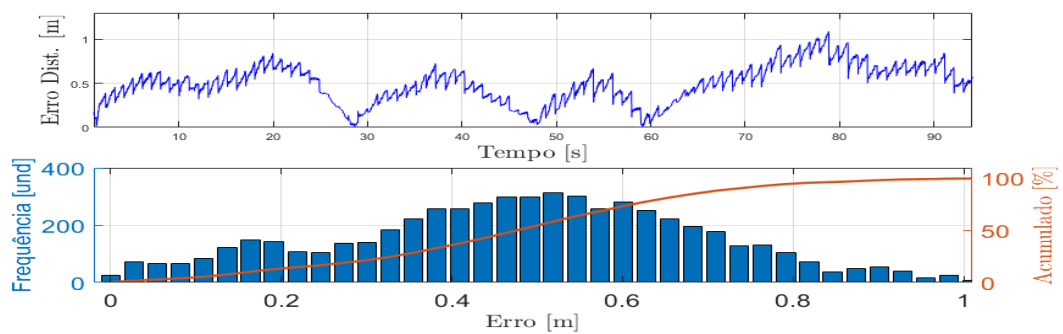
(a)



(b)



(c)



(d)

Figura 28 – Histogramas do erro de posição da simulação do CAT793F no Caminho de Teste para: (a) Wheel, (b) VO+IMU, (c) EKF W+IMU e (d) EKF VO+W+IMU.

Os resultados obtidos nas simulações demonstram que a metodologia proposta contribuiu significativamente para a melhoria na estimação da pose dos veículos. A odometria visual, que já apresentava bom desempenho de forma isolada, foi ainda mais aprimorada com a incorporação dos dados da odometria de rodas e da IMU por meio da fusão sensorial. Além disso, a utilização da GPU para o cálculo do fluxo óptico resultou em uma expressiva redução no tempo de processamento. Enquanto a execução na CPU demandava aproximadamente 0,03 segundos por quadro, o processamento paralelo na GPU reduziu esse tempo para cerca de 0,002 segundos, representando uma melhoria de aproximadamente 15 vezes na taxa de processamento.

Capítulo 7

Conclusão

Essa dissertação apresentou uma metodologia de estimação de pose voltada à navegação de veículos em ambientes de mineração, combinando odometria visual baseada em fluxo óptico, dados inerciais de uma IMU e informações de odometria de rodas por meio de um filtro de Kalman estendido (EKF). A proposta demonstrou desempenho superior em relação às abordagens individuais ao ser avaliada em diferentes trajetórias simuladas no ambiente CoppeliaSim.

A odometria visual-inercial com uma câmera voltada para o solo mostrou-se uma alternativa promissora aos sensores tradicionais de odometria, como *encoders* nas rodas, que frequentemente enfrentam limitações em veículos industriais, como dificuldades de instalação, imprecisões causadas por variações de pressão nos pneus ou pelo tipo de sistema de locomoção empregado. Nesse contexto, a câmera oferece uma solução mais versátil e de fácil adaptação, independente do modelo do veículo.

Além disso, o uso de aceleração por GPU no processamento do fluxo óptico resultou em ganhos expressivos de desempenho computacional. Isso representa um passo importante para aplicações em um veículo real. A integração sensorial permitiu aproveitar as vantagens de cada fonte de informação, compensando suas limitações individuais e melhorando a qualidade das estimativas de pose.

Durante o desenvolvimento, também foi observada a influência de fatores práticos que devem ser considerados em aplicações reais. A metodologia assumiu um terreno plano, mas em cenários reais, a irregularidade do solo afeta a precisão da estimativa de velocidade. A altura da câmera em relação ao solo influencia diretamente as informações visuais. Além disso, a presença de sombras, poeira e vibrações da plataforma pode comprometer o desempenho da odometria visual, sendo aspectos que demandam estratégias adicionais de tratamento ou o uso de sensores complementares.

De forma geral, os resultados obtidos comprovam a eficácia do método proposto como alternativa de navegação quando o GNSS se torna falho ou indisponível, reforçando

o potencial da fusão de sensores visuais e inerciais com otimização computacional para sistemas embarcados em mineração.

7.1 Trabalhos Futuros

Considerando os próximos passos, espera-se realizar a validação da metodologia proposta em veículos reais. Como etapa inicial, propõe-se a utilização do EspeleoRobô, empregado em fases anteriores da pesquisa e adequado para transições do ambiente simulado para o real. Os testes em campo terão como objetivo estimar as covariâncias e calibrar com maior precisão os parâmetros do EKF, preparando o sistema para aplicação em veículos de maior porte, como caminhões de mineração. Essa aplicação prática também permitirá avaliar o impacto de variáveis como iluminação, textura do solo e vibrações na performance da odometria visual-inercial.

Adicionalmente, é necessário expandir a metodologia para cenários tridimensionais, em que variações de altura, desníveis e inclinações do terreno sejam devidamente modelados. Isso inclui, por exemplo, a inclusão de uma etapa de compensação da altura da câmera e a adaptação do modelo de movimento a superfícies inclinadas.

Outra perspectiva está na integração de sensores adicionais, como GPS-RTK, sensores ultrassônicos, LiDAR ou sistemas de radiofrequência, com o objetivo de melhorar a confiabilidade da localização em condições adversas, especialmente em ambientes com baixa visibilidade ou presença de poeira.

Por fim, a odometria visual também pode ser aprimorada para operar com taxas de quadros mais altas, aproveitando o potencial de câmeras de maior frequência (FPS), desde que o sistema mantenha capacidade de processamento em tempo real. Ainda nesse sentido, é desejável que o algoritmo seja adaptado para operar em velocidades compatíveis com veículos reais de mineração, que podem alcançar até 40 km/h.

Referências

- N. J. Abu-Alrub and N. A. Rawashdeh. Radar Odometry for Autonomous Ground Vehicles: A Survey of Methods and Datasets. *IEEE Transactions on Intelligent Vehicles*, 9(3):4275–4291, Mar. 2024. ISSN 2379-8858. doi: 10.1109/tiv.2023.3340513. URL <http://dx.doi.org/10.1109/TIV.2023.3340513>.
- L. R. Agostinho, N. M. Ricardo, M. I. Pereira, A. Hiolle, and A. M. Pinto. A Practical Survey on Visual Odometry for Autonomous Driving in Challenging Scenarios and Conditions. *IEEE Access*, 10:72182–72205, 2022. ISSN 2169-3536. doi: 10.1109/access.2022.3188990. URL <http://dx.doi.org/10.1109/ACCESS.2022.3188990>.
- B. M. Al-Hadithi, D. Thomas, and C. Pastor. Hybrid Visual Odometry Algorithm Using a Downward-Facing Monocular Camera. *Applied Sciences*, 14(17):7732, Sept. 2024. ISSN 2076-3417. doi: 10.3390/app14177732. URL <http://dx.doi.org/10.3390/app14177732>.
- Y. Alkendi, L. Seneviratne, and Y. Zweiri. State of the Art in Vision-Based Localization Techniques for Autonomous Navigation Systems. *IEEE Access*, 9:76847–76874, 2021. ISSN 2169-3536. doi: 10.1109/access.2021.3082778. URL <http://dx.doi.org/10.1109/ACCESS.2021.3082778>.
- M. O. A. Aqel, M. H. Marhaban, M. Iqbal Saripan, N. B. Ismail, and A. Khmag. Optimal Configuration of a Downward-Facing Monocular Camera for Visual Odometry. *Indian Journal of Science and Technology*, 8(32), May 2016a. ISSN 0974-6846. doi: 10.17485/ijst/2015/v8i32/92101. URL <http://dx.doi.org/10.17485/ijst/2015/v8i32/92101>.
- M. O. A. Aqel, M. H. Marhaban, M. I. Saripan, and N. B. Ismail. Review of visual odometry: types, approaches, challenges, and applications. *SpringerPlus*, 5(1), Oct. 2016b. ISSN 2193-1801. doi: 10.1186/s40064-016-3573-7. URL <http://dx.doi.org/10.1186/s40064-016-3573-7>.
- M. Bloesch, S. Omari, M. Hutter, and R. Siegwart. Robust visual inertial odometry using a direct EKF-based approach. In *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, page 298–304. IEEE, Sept. 2015. doi: 10.1109/iros.2015.7353389. URL <http://dx.doi.org/10.1109/IROS.2015.7353389>.

- O. Bulla. Na máxima em 2 anos, minério de ferro reanima vale e desbanca petróleo, 2021. URL <https://investnews.com.br/financas/na-maxima-em-2-anos-minerio-de-ferro-reanima-vale-e-desbanca-petroleo/>. Acessado em: 25 set. 2025.
- J. Campbell, R. Sukthankar, I. Nourbakhsh, and A. Pahwa. A Robust Visual Odometry and Precipice Detection System Using Consumer-grade Monocular Vision. In *Proceedings of the 2005 IEEE International Conference on Robotics and Automation*, page 3421–3427. IEEE, 2005. doi: 10.1109/robot.2005.1570639. URL <http://dx.doi.org/10.1109/ROBOT.2005.1570639>.
- E. Cervera. GPU-Accelerated Vision for Robots: Improving System Throughput Using OpenCV and CUDA. *IEEE Robotics & Automation Magazine*, 27(2):151–158, June 2020.
- Y. Chen, J. Gonder, S. Young, and E. Wood. Quantifying autonomous vehicles national fuel consumption impacts: A data-rich approach. *Transportation Research Part A: Policy and Practice*, 122:134–145, Apr. 2019. ISSN 0965-8564. doi: 10.1016/j.tra.2017.10.012. URL <http://dx.doi.org/10.1016/j.tra.2017.10.012>.
- Y. Cheng, M. Jiang, and Y. Liu. MS-VRO: A Multistage Visual-Millimeter Wave Radar Fusion Odometry. *IEEE Transactions on Robotics*, 40:3004–3023, 2024. ISSN 1941-0468. doi: 10.1109/tro.2024.3400941. URL <http://dx.doi.org/10.1109/TR0.2024.3400941>.
- P. Corke. *Robotics, Vision and Control: Fundamental Algorithms in Python*. Springer International Publishing, 2023.
- M. Dille, B. Grocholsky, and S. Singh. *Outdoor Downward-Facing Optical Flow Odometry with Commodity Sensors*, page 183–193. Springer Berlin Heidelberg, 2010. ISBN 9783642134081. doi: 10.1007/978-3-642-13408-1_17. URL http://dx.doi.org/10.1007/978-3-642-13408-1_17.
- S. Dogru and L. Marques. An improved kinematic model for skid-steered wheeled platforms. *Autonomous Robots*, 45(2):229–243, Jan. 2021. ISSN 1573-7527. doi: 10.1007/s10514-020-09959-0. URL <http://dx.doi.org/10.1007/s10514-020-09959-0>.
- J. Domingues, H. Azpurua, G. Freitas, and G. Pessin. Localization of mobile robots through optical flow and sensor fusion in mining environments. In *Latin American Robotics Symposium (LARS), 2022 Brazilian Symposium on Robotics (SBR), and Workshop on Robotics in Education (WRE)*. IEEE, Oct. 2022.

- J. Engel, J. Sturm, and D. Cremers. Semi-dense visual odometry for a monocular camera. In *2013 IEEE International Conference on Computer Vision*, page 1449–1456. IEEE, Dec. 2013. doi: 10.1109/iccv.2013.183. URL <http://dx.doi.org/10.1109/ICCV.2013.183>.
- S. K. Ericson and B. S. Åstrand. Analysis of two visual odometry systems for use in an agricultural field environment. *Biosystems Engineering*, 166:116–125, Feb. 2018. ISSN 1537-5110. doi: 10.1016/j.biosystemseng.2017.11.009. URL <http://dx.doi.org/10.1016/j.biosystemseng.2017.11.009>.
- G. Farneback. *Two-Frame Motion Estimation Based on Polynomial Expansion*, page 363–370. Springer Berlin Heidelberg, 2003. ISBN 9783540451037. doi: 10.1007/3-540-45103-x_50. URL http://dx.doi.org/10.1007/3-540-45103-X_50.
- G. M. Freitas, G. P. Cruz Junior, A. C. Vangasse, G. Gomes, S. P. Campos, I. F. S. Amaral, A. L. M. Cid, M. C. Delunardo, N. C. P. S. Thiago Neto, R. S. Pereira Junior, L. G. D. Barros, J. D. Domingues, and G. Pessin. Primeiro Relatório Parcial do Projeto ‘Localização, Mapeamento e Comunicação em Cavas de Minas e Galerias de Barragens de Grande Extensão’. Relatório Técnico, 2023.
- G. M. Freitas, G. P. Cruz Junior, A. C. Vangasse, G. Gomes, S. P. Campos, J. F. R. Baiao, F. A. T. Montano, N. H. C. Rocha, A. L. M. Cid, M. C. Delunardo, N. C. P. S. Thiago Neto, R. S. Pereira Junior, Y. M. Bepaloff, L. G. D. Barros, and G. Pessin. Segundo Relatório Parcial do Projeto ‘Localização, Mapeamento e Comunicação em Cavas de Minas e Galerias de Barragens de Grande Extensão’. Relatório Técnico, 2024.
- M. Gilles and S. Ibrahimasic. Unsupervised deep learning based ego motion estimation with a downward facing camera. *The Visual Computer*, 39(3):785–798, Nov. 2021. ISSN 1432-2315. doi: 10.1007/s00371-021-02345-6. URL <http://dx.doi.org/10.1007/s00371-021-02345-6>.
- Global Mining Review. Vale deploys autonomous trucks at Carajás iron ore complex, 2021. URL <https://www.globalminingreview.com/mining/06092021/vale-deploys-autonomous-trucks-at-carajas-iron-ore-complex/>.
- G. Gomes, G. P. Cruz Junior, J. D. Domingues, and G. M. Freitas. Tutorial Sobre a Utilização Das Câmeras Intel Realsense D435i e T265 com o RTAB-Map no ROS. Relatório Técnico, 2023.
- G. Gomes, G. Cruz, J. Domingues, G. Pessin, A. Neto, and G. M. Freitas. Odometria Visual por Fluxo Óptico Integrada a um Sistema de Localização Multi-sensor para Veículos de Mineração. In *Anais do Congresso Brasileiro de Automática 2024*, Oct. 2024a.

- G. Gomes, G. Cruz, J. Domingues, G. Pessin, A. Neto, and G. M. Freitas. Visual Odometry through Optical Flow Integrated with a Multi-sensor Localization System Applied to Mining Vehicles. In *2024 Latin American Robotics Symposium (LARS)*, page 1–6. IEEE, Nov. 2024b. doi: 10.1109/lars64411.2024.10786421. URL <http://dx.doi.org/10.1109/LARS64411.2024.10786421>.
- A. Hadviger, I. Cvisic, I. Markovic, S. Vrazic, and I. Petrovic. Feature-based Event Stereo Visual Odometry. In *2021 European Conference on Mobile Robots (ECMR)*, page 1–6. IEEE, Aug. 2021. doi: 10.1109/ecmr50962.2021.9568811. URL <http://dx.doi.org/10.1109/ECMR50962.2021.9568811>.
- M. He, C. Zhu, Q. Huang, B. Ren, and J. Liu. A review of monocular visual odometry. *The Visual Computer*, 36(5):1053–1065, June 2019. ISSN 1432-2315. doi: 10.1007/s00371-019-01714-6. URL <http://dx.doi.org/10.1007/s00371-019-01714-6>.
- Instituto Brasileiro de Mineração (IBRAM). Dados IBRAM - Setor Mineral, 2024. URL <https://ibram.org.br/publicacoes/dados-ibram/>. Instituto Brasileiro de Mineração.
- C. Jamasmie. Northern star goes after barrick gold’s super pit, 2019. URL <https://www.mining.com/northern-star-goes-after-barrick-golds-super-pit/>. Acessado em: 25 set. 2025.
- D. Kaur. The impact of autonomous vehicles on mining operations: Enhancing safety and productivity through technological advancements. *Scholarly Review Journal*, SR Online: Showcase(Equinox 2024), Oct. 2024. ISSN 2996-8380. doi: 10.70121/001c.124875. URL <http://dx.doi.org/10.70121/001c.124875>.
- S. Khattak, C. Papachristos, and K. Alexis. Keyframe-based Direct Thermal–Inertial Odometry. In *2019 International Conference on Robotics and Automation (ICRA)*. IEEE, May 2019. doi: 10.1109/icra.2019.8793927. URL <http://dx.doi.org/10.1109/ICRA.2019.8793927>.
- M. Killpack, T. Deyle, C. Anderson, and C. C. Kemp. Visual odometry and control for an omnidirectional mobile robot with a downward-facing camera. In *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*, page 139–146. IEEE, Oct. 2010. doi: 10.1109/iros.2010.5649749. URL <http://dx.doi.org/10.1109/IROS.2010.5649749>.
- G. Klančar, A. Zdešar, S. Blažič, and I. Škrjanc. *Introduction to Mobile Robotics*, page 1–11. Elsevier, 2017. ISBN 9780128042045. doi: 10.1016/b978-0-12-804204-5.00001-9. URL <http://dx.doi.org/10.1016/B978-0-12-804204-5.00001-9>.

- H. Lee, H. Lee, I. Kwak, C. Sung, and S. Han. Effective Feature-Based Downward-Facing Monocular Visual Odometry. *IEEE Transactions on Control Systems Technology*, 32(1):266–273, Jan. 2024. ISSN 2374-0159. doi: 10.1109/tcst.2023.3294843. URL <http://dx.doi.org/10.1109/TCST.2023.3294843>.
- J. Liao, Z. Chen, and B. Yao. Model-Based Coordinated Control of Four-Wheel Independently Driven Skid Steer Mobile Robot with Wheel–Ground Interaction and Wheel Dynamics. *IEEE Transactions on Industrial Informatics*, 15(3):1742–1752, Mar. 2019.
- J. Libby and G. Kantor. Deployment of a point and line feature localization system for an outdoor agriculture vehicle. In *ICRA*, pages 1565–1570. IEEE, 2011.
- H.-Y. Lin and J.-R. Zhan. GNSS-denied UAV indoor navigation with UWB incorporated visual inertial odometry. *Measurement*, 206:112256, Jan. 2023.
- D. Liu, Y. Cui, X. Guo, W. Ding, B. Yang, and Y. Chen. Visual Localization for Autonomous Driving: Mapping the Accurate Location in the City Maze. In *25th International Conference on Pattern Recognition (ICPR)*. IEEE, Jan. 2021.
- A. Mandow, J. L. Martinez, J. Morales, J. L. Blanco, A. Garcia-Cerezo, and J. Gonzalez. Experimental kinematics for wheeled skid-steer mobile robots. In *2007 IEEE/RSJ International Conference on Intelligent Robots and Systems*, page 1222–1227. IEEE, Oct. 2007. doi: 10.1109/iros.2007.4399139. URL <http://dx.doi.org/10.1109/IROS.2007.4399139>.
- S. A. S. Mohamed, M.-H. Haghbayan, T. Westerlund, J. Heikkonen, H. Tenhunen, and J. Plosila. A Survey on Odometry for Autonomous Navigation Systems. *IEEE Access*, 7:97466–97486, 2019. ISSN 2169-3536. doi: 10.1109/access.2019.2929133. URL <http://dx.doi.org/10.1109/ACCESS.2019.2929133>.
- J. Moreau, S. Ambellouis, and Y. Ruichek. Fisheye-Based Method for GPS Localization Improvement in Unknown Semi-Obstructed Areas. *Sensors*, 17(12):119, Jan. 2017.
- Q. Naixin, Y. Xiaogang, L. Chuanxiang, L. Xiaofeng, Z. Shengxiu, and C. Lijia. Monocular Semidirect Visual Odometry for Large-Scale Outdoor Localization. *IEEE Access*, 7:57927–57942, 2019. ISSN 2169-3536. doi: 10.1109/access.2019.2914033. URL <http://dx.doi.org/10.1109/ACCESS.2019.2914033>.
- N. Nourani-Vatani, J. Roberts, and M. Srinivasan. Practical visual odometry for car-like vehicles. In *2009 IEEE International Conference on Robotics and Automation*, page 3551–3557. IEEE, May 2009. doi: 10.1109/robot.2009.5152403. URL <http://dx.doi.org/10.1109/ROBOT.2009.5152403>.

- N. Nourani-Vatani and P. V. K. Borges. Correlation-based visual odometry for ground vehicles. *Journal of Field Robotics*, 28(5):742–768, Aug. 2011. ISSN 1556-4967. doi: 10.1002/rob.20407. URL <http://dx.doi.org/10.1002/rob.20407>.
- L. Piyathilaka and R. Munasinghe. Multi-camera visual odometry for skid steered field robot. In *2010 Fifth International Conference on Information and Automation for Sustainability*, page 189–194. IEEE, Dec. 2010. doi: 10.1109/iciafs.2010.5715658. URL <http://dx.doi.org/10.1109/ICIAFS.2010.5715658>.
- R. Rajesh and P. V. Manivannan. Enhanced Monocular Visual Odometry: A Robust Appearance-Based Method for Accurate Vehicle Pose Estimation. *IEEE Access*, 12: 106176–106192, 2024. ISSN 2169-3536. doi: 10.1109/access.2024.3437658. URL <http://dx.doi.org/10.1109/ACCESS.2024.3437658>.
- A. M. C. Rezende, V. M. Goncalves, and L. C. A. Pimenta. Constructive Time-Varying Vector Fields for Robot Navigation. *IEEE Transactions on Robotics*, 38(2):852–867, 2021. doi: 10.1109/TRO.2021.3093674.
- G. Rosa. A vale em barão de cocais, são gonçalo do rio abaixo, santa bárbara e rio piracicaba (mg), 2025. URL <https://vale.com/pt/baraodecocais>. Acessado em: 25 set. 2025.
- D. Scaramuzza and F. Fraundorfer. Visual Odometry [Tutorial]. *IEEE Robotics & Automation Magazine*, 18(4):80–92, Dec. 2011. ISSN 1558-223X. doi: 10.1109/mra.2011.943233. URL <http://dx.doi.org/10.1109/MRA.2011.943233>.
- L. G. Soares, N. H. C. Rocha, G. E. Soares, G. P. Cruz Júnior, G. Gomes, G. M. de Oliveira, A. C. Vangasse, H. I. Azpúrua, A. L. M. Cid, and G. M. Freitas. Enhancing the autonomous mapping and navigation unit for service robots. In *2025 Brazilian Conference on Robotics (CROS)*, volume 1, pages 1–6, 2025. doi: 10.1109/CROS66186.2025.11066085.
- X. Song, K. Althoefer, and L. Seneviratne. A robust downward-looking camera based velocity estimation with height compensation for mobile robots. In *2010 11th International Conference on Control Automation Robotics & Vision*, page 378–383. IEEE, Dec. 2010. doi: 10.1109/icarcv.2010.5707269. URL <http://dx.doi.org/10.1109/ICARCV.2010.5707269>.
- M. W. Spong, S. Hutchinson, and M. Vidyasagar. *Robot Modeling and Control*, volume 3. Wiley New York, 2006.
- F. Steinbrucker, J. Sturm, and D. Cremers. Real-time visual odometry from dense RGB-D images. In *2011 IEEE International Conference on Computer Vision Workshops (ICCV Workshops)*. IEEE, Nov. 2011. doi: 10.1109/iccvw.2011.6130321. URL <http://dx.doi.org/10.1109/ICCVW.2011.6130321>.

- K. Sun, K. Mohta, B. Pfrommer, M. Watterson, S. Liu, Y. Mulgaonkar, C. J. Taylor, and V. Kumar. Robust Stereo Visual Inertial Odometry for Fast Autonomous Flight. *IEEE Robotics and Automation Letters*, 3(2):965–972, Apr. 2018. ISSN 2377-3774. doi: 10.1109/lra.2018.2793349. URL <http://dx.doi.org/10.1109/LRA.2018.2793349>.
- J. Szrek, P. Trybała, M. Góralczyk, A. Michalak, B. Ziętek, and R. Zimroz. Accuracy Evaluation of Selected Mobile Inspection Robot Localization Techniques in a GNSS-Denied Environment. *Sensors*, 21(1):141, Dec. 2020.
- R. Teles. Vale e caterpillar reforçam colaboração com foco em produtividade, inovação e descarbonização, 2025. URL <https://vale.com/pt/w/vale-e-caterpillar-reforcamos-colaboracao-com-foco-em-productividade-inovacao-e-des>. Acessado em: 25 set. 2025.
- F. Tian, R. Zhou, Z. Li, L. Li, Y. Gao, D. Cao, and L. Chen. Trajectory Planning for Autonomous Mining Trucks Considering Terrain Constraints. *IEEE Transactions on Intelligent Vehicles*, 6(4):772–786, Dec. 2021.
- V. Usenko, J. Engel, J. Stuckler, and D. Cremers. Direct visual-inertial odometry with stereo cameras. In *2016 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, May 2016. doi: 10.1109/icra.2016.7487335. URL <http://dx.doi.org/10.1109/ICRA.2016.7487335>.
- D. Valiente García, L. Fernández Rojo, A. Gil Aparicio, L. Payá Castelló, and O. Reinoso García. Visual Odometry through Appearance- and Feature-Based Method with Omnidirectional Images. *Journal of Robotics*, 2012:1–13, 2012. ISSN 1687-9619. doi: 10.1155/2012/797063. URL <http://dx.doi.org/10.1155/2012/797063>.
- K. Yousif, A. Bab-Hadiashar, and R. Hoseinnezhad. An Overview to Visual Odometry and Visual SLAM: Applications to Mobile Robotics. *Intelligent Industrial Systems*, 1(4):289–311, Nov. 2015. ISSN 2199-854X. doi: 10.1007/s40903-015-0032-7. URL <http://dx.doi.org/10.1007/s40903-015-0032-7>.
- Y. Yu, C. Pradalier, and G. Zong. Appearance-based monocular visual odometry for ground vehicles. In *2011 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM)*. IEEE, July 2011. doi: 10.1109/aim.2011.6027050. URL <http://dx.doi.org/10.1109/AIM.2011.6027050>.
- Z. Yuan, Q. Wang, K. Cheng, T. Hao, and X. Yang. SDV-LOAM: Semi-Direct Visual-LiDAR Odometry and Mapping. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(9):11203–11220, Sept. 2023. ISSN 1939-3539. doi: 10.1109/tpami.2023.3262817. URL <http://dx.doi.org/10.1109/TPAMI.2023.3262817>.

- Q. Zeng, B. Ou, C. Lv, S. Scherer, and Y. Kan. Monocular Visual Odometry Using Template Matching and IMU. *IEEE Sensors Journal*, 21(15):17207–17218, Aug. 2021. ISSN 2379-9153. doi: 10.1109/jsen.2021.3078847. URL <http://dx.doi.org/10.1109/JSEN.2021.3078847>.
- J. Zhang and S. Singh. Visual–Inertial Combined Odometry System for Aerial Vehicles. *Journal of Field Robotics*, 32(8):1043–1055, July 2015. ISSN 1556-4967. doi: 10.1002/rob.21599. URL <http://dx.doi.org/10.1002/rob.21599>.
- L. Zhang, D. Wisth, M. Camurri, and M. Fallon. Balancing the Budget: Feature Selection and Tracking for Multi-Camera Visual-Inertial Odometry. *IEEE Robotics and Automation Letters*, 7(2):1182–1189, Apr. 2022. ISSN 2377-3774. doi: 10.1109/lra.2021.3137910. URL <http://dx.doi.org/10.1109/LRA.2021.3137910>.
- J. Zhu, Y. Tang, X. Shao, and Y. Xie. Multisensor Fusion Using Fuzzy Inference System for a Visual-IMU-Wheel Odometry. *IEEE Transactions on Instrumentation and Measurement*, 70:1–16, 2021. ISSN 1557-9662. doi: 10.1109/tim.2021.3051999. URL <http://dx.doi.org/10.1109/TIM.2021.3051999>.
- X. Zou, C. Xiao, W. Zhan, C. Zhou, S. Xiu, and H. Yuan. A Novel Water-Shore-Line Detection Method for USV Autonomous Navigation. *Sensors*, 20(6):1682, Mar. 2020. ISSN 1424-8220. doi: 10.3390/s20061682. URL <http://dx.doi.org/10.3390/s20061682>.
- X. Zou, J. Gao, H. Li, Y. Zhang, F. Liu, H. Wu, and R. Yao. Novel Visual Odometry Method for Water Autonomous Navigation. In *2022 2nd International Conference on Computation, Communication and Engineering (ICCCE)*, page 13–17. IEEE, Nov. 2022. doi: 10.1109/iccce55785.2022.10036173. URL <http://dx.doi.org/10.1109/ICCCE55785.2022.10036173>.