



## **UNIVERSIDADE FEDERAL DE CAMPINA GRANDE**

**Métodos de calibração de câmeras para sistemas de captura de movimento  
baseados em visão computacional**

**Campina Grande, PB**

**Maio de 2023**

## RESUMO

Este estudo trata dos métodos de calibração de câmeras digitais utilizadas na implementação de sistemas de captura de movimento baseados em visão computacional. O objetivo geral é o desenvolvimento de um procedimento de calibração que seja consistente, acurado, robusto, reproduzível, convergente e genérico. O desenvolvimento deste procedimento de calibração demandará a utilização de modelos matemáticos das câmeras digitais, a formulação matemática do problema de calibração, a realização de múltiplos experimentos práticos para coleta de dados, e a concepção de algoritmos de processamento de dados que permitam a determinação dos parâmetros intrínsecos das câmeras digitais. A coleta dos dados experimentais será efetuada com a plataforma MoCap-Rasp que foi projetada para efetuar a captura e o rastreamento do movimento de mini-drones utilizando quatro câmeras digitais. O procedimento de calibração tem por objetivo aumentar a precisão das medições da posição e da orientação dos mini-drones, de modo a poder considerar que a plataforma de captura de movimento seja tratada como um sensor integrado que fornece os sinais posição tridimensional e de orientação para serem utilizados numa malha de realimentação do sistema de controle de vôo da aeronave.

**Palavras-chave:** câmeras digitais, métodos de calibração, visão computacional, captura de movimento

## 1 INTRODUÇÃO

O avanço contínuo da microeletrônica proporciona o desenvolvimento de sistemas de visão computacional de dimensões físicas reduzidas, de baixo custo, de alta qualidade, e de alta velocidade de aquisição. A disponibilidade destes sistemas enseja a possibilidade de implementação de sistemas de instrumentação baseados em visão computacional. Visão computacional é um campo de estudo que trata de como computadores podem extrair informações úteis do mundo a partir de imagens e vídeos. Imagens e vídeos, adquiridos por intermédio de sensores de imagens, são processadas usando algoritmos que proporcionam a extração dos dados visuais relevantes.

Mediante o uso destes sistemas de instrumentação baseados em visão computacional é possível medir em tempo real a posição e a orientação tridimensionais de um ou mais corpos rígidos que se movam num determinado ambiente [1]. Uma das aplicações de grande relevância é a reconstrução do espaço tridimensional a partir de dados visuais fornecidos por câmeras e outros sensores. Isto permite detectar e monitorar objetos, além de realizar medições do mundo em volta. Esta aplicação é de grande relevância para áreas como a robótica, em que câmeras são usadas para guiar a navegação de robôs [2].

Especificamente, para navegação em ambientes fechados, em que a acurácia das medições de posição do robô necessitam ser milimétricas, usar um sistema de captura de movimento baseado em visão computacional se torna a opção mais viável [3]. Nesse espírito, foi construída a plataforma MoCap-Rasp, um sistema de captura de movimento projetado para realizar o rastreio da trajetória de mini-drones a partir da imagem capturada por quatro câmeras digitais. Para obter a posição e orientação desses corpos, são utilizados marcadores reflexivos conectados ao corpo do mini-drone, que irão ser destacados nas imagens capturadas pela plataforma, de forma a possibilitar a extração dessas informações.

A utilização de câmeras para medição e reconstrução do espaço é extremamente vantajosa, pois é um processo que não requer intervenção manual, é rápido e capaz de gerar medidas precisas. No entanto, para se obter essas medidas com acurácia, é necessário determinar os parâmetros do modelo matemático da câmera utilizada na medição. Uma câmera é geralmente representada com um modelo linear da câmera [4], em que dado um ponto tridimensional no sistema de coordenada da câmera  $\tilde{X}_c$ , a sua projeção na imagem  $\tilde{u}$  é dada por:

$$\tilde{u} = \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = [K|0] \cdot \tilde{X}_c = \begin{bmatrix} f_x & 0 & O_x & 0 \\ 0 & f_y & O_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix}, \quad (1)$$

na qual  $K$  é a matriz de intrínsecos da câmera,  $(f_x, f_y)$  é a distância focal no eixo horizontal e vertical e  $(O_x, O_y)$  são as coordenadas do ponto principal da câmera, esses chamados de parâmetros intrínsecos da câmera. Para deslocar  $\tilde{u}$  do sistema de coordenada da câmera para

o sistema coordenadas tridimensionais de referência  $\tilde{X}_w$  a Equação 1 pode ser reescrita como:

$$\tilde{u} = K \cdot [R|t] \cdot \tilde{X}_w = P \cdot \tilde{X}_w \quad (2)$$

Em que  $(R, t)$  é a matriz de rotação e o vetor de translação da câmera do sistema de coordenadas da câmera para o sistema de referência, os seus parâmetros extrínsecos. Ao unir os parâmetros intrínsecos e os extrínsecos, tem-se a matriz de projeção  $P$ , que rege a transformação dos pontos capturados no sistema de referência para o plano da imagem.

Para câmeras mais atuais, a lente presente provoca distorções na imagem obtida, alterando a posição real dos pontos capturados nas imagens. Isso faz com que o modelo de câmera deixe de ser linear, e uma correção da distorção da lente se torna necessária para encontrar a projeção correta dos pontos pelo modelo matemático da câmera [5]. Os pontos vistos na imagem distorcida  $(u_d, v_d)$  são dados por:

$$u_d = u \cdot f(r, k) \quad (3)$$

$$v_d = v \cdot f(r, k) \quad (4)$$

em que  $f(r, k)$  é a função de distorção, dada por:

$$f(r, k) = \frac{1 + k_1 r^2 + k_2 r^4 + k_3 r^6}{1 + k_4 r^2 + k_5 r^4 + k_6 r^6} \quad (5)$$

onde  $k = [k_1, k_2, k_3, \dots]$  são os coeficientes de distorção.

O processo em que são determinados esses parâmetros é denominado de calibração de câmera. Para determiná-los, primeiro é necessário capturar fotos de um objeto de geometria conhecida. Esses objetos são construídos de forma que localização de cada ponto na imagem já esteja determinada no mundo real.

Figura 1 – Plano de calibração no formato quadricular



Fonte: Preparado pelo autor

Para encontrar a matriz de projeção  $P$ , supõe-se que a distorção seja zero inicialmente.

Em seguida, decompõe-se  $P$  em  $K$ ,  $R$  e  $t$ :

$$P = K \cdot (R|t) \quad (6)$$

$$\begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{bmatrix} = \begin{bmatrix} f_x & 0 & O_x \\ 0 & f_y & O_y \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} = K \cdot R \quad (7)$$

$$\begin{bmatrix} p_{14} \\ p_{24} \\ p_{34} \end{bmatrix} = \begin{bmatrix} f_x & 0 & O_x \\ 0 & f_y & O_y \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix} = K \cdot t \quad (8)$$

Como  $K$  é uma matriz triangular superior e  $R$  uma matriz ortonormal, é possível resolver o produto  $K \cdot R$  utilizando-se fatoração QR [6]. Por fim, encontra-se o vetor de translação  $t$  e determina-se a matriz de projeção  $P$ .

Conhecendo uma estimativa inicial da matriz de projeção, é necessário então considerar a distorção da lente na estimativa final dos parâmetros. Dessa forma, usa-se um algoritmo de otimização não-linear para estimar os coeficientes de distorção e otimizar os parâmetros intrínsecos encontrados anteriormente.

Um fator que influencia na calibração e, consequentemente, na acurácia das medidas calculadas pela câmera é o erro de reprojeção, ou seja, a diferença entre a posição prevista de um objeto e a sua posição real observada na imagem capturada. Na estimativa dos parâmetros da câmera objetiva-se minimizar este erro de reprojeção, de modo a obter valores representativos. Portanto, escolher um modelo de câmera e determinar seus parâmetros é um requisito para obter resultados precisos de medição. Ou seja, um algoritmo de calibração que resulte em parâmetros acurados poderá proporcionar medições mais acuradas e confiáveis para as aplicações que utilizam da câmera como instrumento de medição. No caso específico da plataforma MoCap-Rasp, que foi projetada com componentes de uso geral [7], a maximização da acurácia é fortemente dependente do procedimento de calibração. Um procedimento de calibração adequado almeja permitir que o desempenho da plataforma MoCap-Rasp, em termos de velocidade e precisão, seja comparável ao desempenho de plataformas comerciais tais como a OptiTrack [8].

## 2 OBJETIVOS

O objetivo geral é avaliar os métodos já existentes para a determinação dos parâmetros da câmera e os seus impactos na acurácia do rastreio de objetos em sistemas de capturas baseados em visão computacional e propor um procedimento de calibração geral e de fácil uso, o qual permita estimar acuradamente os valores dos parâmetros intrínsecos da câmera e aumentar a precisão das medições realizadas em sistemas de captura.

## 2.1 OBJETIVOS ESPECÍFICOS

- Compreender a modelagem matemática de uma câmera digital, bem como o processo de determinação dos parâmetros intrínsecos.
- Estudar as diversas estratégias e algoritmos de calibração já existentes e selecionar aqueles que, a depender do equipamento disponível, permitam produzir os resultados mais acurados.
- Entender o processo de otimização de parâmetros utilizando algoritmos não-lineares e propor um estudo da otimização dos parâmetros de câmera utilizando o erro de reprojeção como forma de atingir resultados mais acurados.

## 3 JUSTIFICATIVA E RELEVÂNCIA

É perceptível o aumento no interesse da área de robótica e visão computacional, uma vez que, devido a disseminação dos princípios da indústria 4.0, sistemas ópticos de captura de movimento estão sendo largamente utilizados em aplicações de controle realimentado e de automação, e.g., vision-based robot control [9]. Com isso, para suprir a necessidade dessas aplicações, surgem novos algoritmos de reconstrução espacial, e com eles, a necessidade de estratégias práticas de calibração.

Em meio a diversidade de procedimentos existentes para calcular os parâmetros intrínsecos, é uma tarefa difícil escolher a estratégia que melhor se adequa à câmera que será utilizada, bem como o conjunto de práticas necessárias no momento da captura das imagens que serão utilizadas na calibração, a fim de minimizar o erro de reprojeção final do processo, algo que impacta diretamente na precisão das medidas obtidas por esses sistemas de captura.

Assim, o desenvolvimento de um guia prático, contendo o procedimento necessário desde a coleta dos dados, bem como um estudo sobre as estratégias existentes para a estimativa dos parâmetros, é de extrema relevância para as aplicações que utilizam a reconstrução 3D e captura de movimento.

## 4 ESTADO DA ARTE

Em diversos campos da robótica e da indústria, câmeras estão sendo largamente utilizadas atualmente para extrair informações importantes do ambiente. No entanto, a modelagem e calibração de câmeras é foco de estudo desde o surgimento dos primeiros algoritmos de visão computacional.

Uma câmera captura os pontos do ambiente no formato de uma imagem, ou seja, transforma pontos do mundo real no plano 2D da imagem. Essa transformação de coordenadas

é chamada de projeção e é regida pela matriz de projeção da câmera. Saber modelar a câmera é saber descrever essa transformação a partir da geometria da câmera.

O primeiro modelo proposto para a câmera foi o “modelo pinhole”, um modelo baseado apenas na geometria da câmera que desconsidera a distorção da imagem pelas lentes da câmera. Ele é uma aproximação da projeção feita pela câmera dos pontos 3D do mundo real para o plano 2D da imagem capturada. A partir deste modelo, foram desenvolvidas técnicas de calibração e estimativa dos parâmetros intrínsecos.

Uma dessas técnicas é a utilização da Transformação Linear Direta [10], em que a partir do método de mínimos quadrados, são relacionados pontos 3D com suas projeções 2D. Uma desvantagem desse método é que ela desconsidera a distorção em certos pontos devido a lente da câmera. Para sistemas de aquisição com muita distorção essa técnica é inadequada. Para superar esta limitação, é necessário incluir a distorção das lentes nos pontos da imagem no modelo. Essa distorção pode ser tanto radial como tangencial e irá proporcionar um deslocamento da posição real do ponto 3D para a sua projeção [5]. O modelo usual para a modelagem é por uma função polinomial na qual seus coeficientes descrevem a distorção das lentes.

Quando considerada essa distorção das lentes, as técnicas de calibração tornaram-se não-lineares, com os parâmetros obtidos a partir de algoritmos de iteração e otimização dos parâmetros intrínsecos. Estratégias mais simples e robustas, como proposto em [4] surgiram, em que com apenas poucos pontos de um plano de calibração é possível concluir a calibração. Planos de calibração são objetos de geometria conhecida e de fácil detecção dos pontos, como tabuleiros de xadrez, grades circulares e marcadores fiduciais. O formato escolhido pode afetar diretamente a acurácia da calibração, como visto em [11].

Com câmeras com maior campo de visão, como câmeras “olho de peixe”(fisheye), o modelo de câmera pinhole não é mais suficiente para modelar toda as características. Dessa forma, outros modelos de câmeras tornam-se necessários, como [12] e [13], propondo uma nova maneira de olhar para a câmera com modelos genéricos, capazes de se adequar a câmera e suas características.

Outras técnicas de calibração não-lineares surgiram, com foco em algoritmos que proporcionam uma calibração mais acurada. Utilizando a proposta de Zhang [4], é possível refinar a estimativa inicial e reduzir a influência do erro de reprojeção, como proposto em [14]. Com foco na redução da influência do erro de reprojeção, diversos algoritmos de calibração também foram propostos [15]. É possível detectar, selecionar e extrair os pontos com maior erro de reprojeção nas imagens capturadas para a calibração, de forma a determinar os parâmetros utilizando somente os pontos mais acurados. Estratégias como essa foram propostas por [16] e [17].

## 5 METODOLOGIA E CRONOGRAMA

Este projeto propõe um estudo sobre as técnicas de calibração disponíveis a partir de uma análise minuciosa da acurácia do processo. Para isso, inicialmente, devem ser estudados os modelos matemáticos de câmeras digitais e o processo de estimativa dos parâmetros intrínsecos da câmera, bem como a sua importância para o rastreio de corpos rígidos. Além disso, também deve ser alvo de estudo as diferentes técnicas de calibração existentes, a fim de entender a formulação matemática dessas técnicas e o passo a passo a ser seguido nas etapas pré-calibração. Isto será feito a partir de uma atualização da revisão bibliográfica e uma busca detalhada na literatura, a fim de pesquisar e compreender os diferentes modelos de câmeras e as técnicas de calibração já existentes [4, 12, 15].

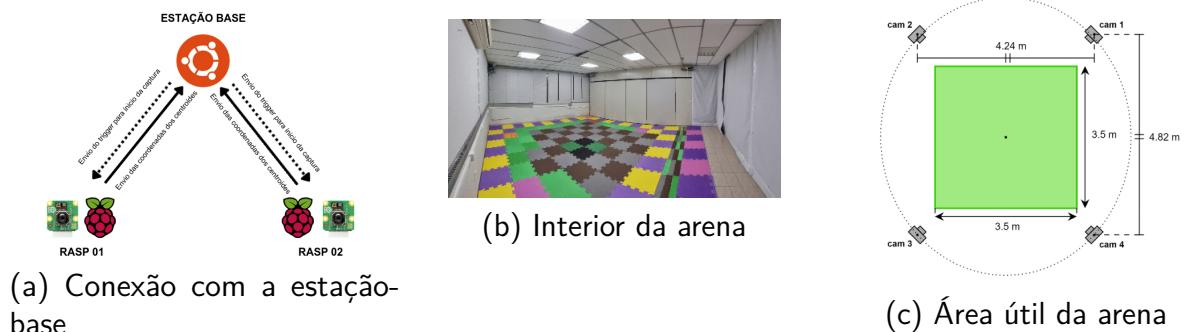
A partir disso, deve ser selecionadas as estratégias que maximizem a acurácia da calibração. Essa escolha é baseada numa série de experimentos práticos, utilizando os dados colhidos do sistema de captura MoCap-Rasp, em que as técnicas de calibração serão testadas e avaliadas com base na exatidão do processo. Para a investigação e seleção das estratégias, deve ser utilizadas ferramentas computacionais para a realização das análises do erro de reprojeção final de cada técnica de calibração. Após selecionadas, deve ser determinado qual conjunto de práticas necessárias para a captura dos dados utilizadas na calibração e de que modo essas estratégias devem ser empregadas, a fim de minimizar o erro de reprojeção na calibração.

O sistema de captura MoCap-Rasp opera com quatro câmeras digitais conectadas a quatro Raspberry-Pi, com uma área útil de  $12,25m^2$ . Os principais elementos da configuração atual da plataforma MoCap-Rasp são apresentados na Figura 2. Cada câmera se comunica com uma estação-base, onde se encontra o computador central que realiza o processamento dos dados da plataforma. A captura inicia-se com o trigger enviado pelo computador central. As câmeras gravam a trajetória do corpo rígido e as Raspberry-Pi extraem as informações importantes das imagens, como a localização dos marcadores. Essas informações são enviadas a estação-base, que irá realizar o processo de triangulação dos marcadores e reconstrução visual. A acurácia da reconstrução visual está relacionada com um bom processo de calibração câmera, uma vez que estimar corretamente os parâmetros intrínsecos da câmera torna a projeção dos pontos na triangulação menos suscetível a erros.

Após determinado o procedimento de calibração, deve ser estudado a atuação do erro de reprojeção na estimativa dos parâmetros da câmera, com o objetivo de entender de que forma ele influencia no processo de calibração e procurar minimizá-lo a partir da concepção de um novo algoritmo de otimização de parâmetros.

Esse algoritmo utilizará de uma otimização não linear, em que os dados com maior erro de reprojeção devem ser menos relevantes na estimativa dos parâmetros. Concebido o algoritmo, este será testado utilizando os dados capturados do sistema de captura MoCap-Rasp e sua eficácia e precisão comparada com os algoritmos de otimização já existentes [14].

Figura 2 – Plataforma MoCap-Rasp: a) conexão lógica entre duas das estações de coleta de imagem e a estação-base (a conexão das estações 03 e 04 é semelhante), b) fotografia do ambiente, e c) posicionamento das câmeras e dimensões físicas.



Fonte: Preparado pelo autor

## 5.1 CRONOGRAMA

- Etapa 1 Estudar os diferentes tipos de modelagem de câmera e compreender a descrição matemática do processo de determinação e otimização dos parâmetros intrínsecos da câmera.
- Etapa 2 Analisar as diferentes técnicas de calibração disponíveis em níveis de precisão e acurácia dos valores encontrados.
- Etapa 3 Estudar e encontrar uma forma de minimizar a influência do erro de reprojeção na estimativa dos parâmetros de câmera.
- Etapa 4 Documentar os resultados obtidos em relatórios e artigos. Apresentar os resultados da pesquisa nos congressos de iniciação científica da Universidade Federal de Campina Grande.

	2023				2024							
	SET	OUT	NOV	DEZ	JAN	FEV	MAR	ABR	MAI	JUN	JUL	AGO
Etapa 1												
Etapa 2												
Etapa 3												
Etapa 4												

## Referências

- 1 SHIRMOHAMMADI, S.; FERRERO, A. Camera as the instrument: the rising trend of vision based measurement. *IEEE Instrumentation & Measurement Magazine*, v. 17, n. 3, p. 41–47, jun. 2014. ISSN 1094-6969, 1941-0123. Disponível em: <<https://ieeexplore.ieee.org/document/6825388/>>. Citado na página 2.
- 2 OLIVEIRA, D. N. et al. Void detection for uav based on optical flow and vanishing points. In: *Simpósio Brasileiro de Automação Inteligente-SBAI*. Rio Grande: SBA, 2021. p. 1–6. Citado na página 2.
- 3 OLIVEIRA, D. N. P. *Heterogeneous and asynchronous optical tracking system for mobile vehicles surveillance*. Dissertação (Mestrado) — Universidade Federal de Campina Grande, Campina Grande, 2022. Citado na página 2.
- 4 ZHANG, Z. A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 22, n. 11, p. 1330–1334, nov. 2000. ISSN 01628828. Disponível em: <<http://ieeexplore.ieee.org/document/888718/>>. Citado 3 vezes nas páginas 2, 6 e 7.
- 5 MA, L.; CHEN, Y.; MOORE, K. L. *A Family of Simplified Geometric Distortion Models for Camera Calibration*. arXiv, 2003. ArXiv:cs/0308003. Disponível em: <<http://arxiv.org/abs/cs/0308003>>. Citado 2 vezes nas páginas 3 e 6.
- 6 ANDERSON, E.; BAI, Z.; DONGARRA, J. Generalized QR factorization and its applications. *Linear Algebra and its Applications*, Elsevier, v. 162, p. 243–271, 1992. Citado na página 4.
- 7 OLIVEIRA, D. N.; MORAIS, M. R.; LIMA, A. M. Optical tracking system based on cots components. In: *5th International Conference on Image Processing Applications and Systems (IPAS)*. Genova: IEEE, 2022. p. 1–6. Citado na página 4.
- 8 FURTADO, J. S. et al. Comparative analysis of optitrack motion capture systems. In: *Advances in Motion Sensing and Control for Robotic Applications: Selected Papers from the Symposium on Mechatronics, Robotics, and Control - CSME*. Toronto: Springer, 2019. p. 15–31. Citado na página 4.
- 9 SHAHRIA, M. T. et al. A comprehensive review of vision-based robotic applications: Current state, components, approaches, barriers, and potential solutions. *Robotics*, MDPI, v. 11, n. 6, p. 139, 2022. Citado na página 5.
- 10 ABDEL-AZIZ, Y.; KARARA, H. Direct Linear Transformation from Comparator Coordinates into Object Space Coordinates in Close-Range Photogrammetry. *Photogrammetric Engineering & Remote Sensing*, v. 81, n. 2, p. 103–107, fev. 2015. ISSN 00991112. Disponível em: <<http://openurl.ingenta.com/content/xref?genre=article&issn=0099-1112&volume=81&issue=2&spage=103>>. Citado na página 6.
- 11 MALLON, J.; WHELAN, P. F. Which pattern? Biasing aspects of planar calibration patterns and detection methods. *Pattern Recognition Letters*, v. 28, n. 8, p. 921–930, jun. 2007. ISSN 01678655. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0167865506003114>>. Citado na página 6.

- 12 KANNALA, J.; BRANDT, S. A generic camera model and calibration method for conventional, wide-angle, and fish-eye lenses. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 28, n. 8, p. 1335–1340, ago. 2006. ISSN 0162-8828, 2160-9292. Disponível em: <<http://ieeexplore.ieee.org/document/1642666>>. Citado 2 vezes nas páginas 6 e 7.
- 13 SCARAMUZZA, D.; MARTINELLI, A.; SIEGWART, R. A Flexible Technique for Accurate Omnidirectional Camera Calibration and Structure from Motion. In: *Fourth IEEE International Conference on Computer Vision Systems (ICVS'06)*. New York, NY, USA: IEEE, 2006. p. 45–45. ISBN 978-0-7695-2506-8. Disponível em: <<http://ieeexplore.ieee.org/document/1578733>>. Citado na página 6.
- 14 HU, G. et al. Non-linear calibration optimisation based on the levenberg–marquardt algorithm. *IET Image Processing*, Wiley Online Library, v. 14, n. 7, p. 1402–1414, 2020. Citado 2 vezes nas páginas 6 e 7.
- 15 SALVI, J.; ARMANGUÉ, X.; BATLLE, J. A comparative review of camera calibrating methods with accuracy evaluation. *Pattern Recognition*, v. 35, n. 7, p. 1617–1635, jul. 2002. ISSN 00313203. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0031320301001261>>. Citado 2 vezes nas páginas 6 e 7.
- 16 ZHOU, F. et al. Accurate and robust estimation of camera parameters using RANSAC. *Optics and Lasers in Engineering*, v. 51, n. 3, p. 197–212, mar. 2013. ISSN 01438166. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0143816612003016>>. Citado na página 6.
- 17 JU, H.; YUNHUI, L.; MING, Y. Multi-camera calibration method based on minimizing the difference of reprojection error vectors. *Journal of Systems Engineering and Electronics*, BIAI, v. 29, n. 4, p. 844–853, 2018. Citado na página 6.