

# Pré-alinhamento Baseado na Distância de Wasserstein para o Registro de Nuvens de Pontos

Patricia J. O. Martins<sup>1</sup>, Jefferson C. Figueiredo<sup>2</sup>, José M. Soares<sup>3</sup> e George A. P. Thé<sup>4</sup>  
UFC, Fortaleza, CE

**Resumo.** O registro de nuvens de pontos é uma atividade desafiadora devido às transformações geométricas complexas e à dificuldade de encontrar correspondências entre os pontos. Quanto maior o desalinhamento entre as nuvens, maior a dificuldade imposta aos algoritmos destinados ao registro. Para viabilizar o uso de algoritmos disponíveis em bibliotecas que apresentam essa limitação, propõe-se neste trabalho uma técnica baseada na distância de Wasserstein em uma etapa prévia ao registro, permitindo resultados satisfatórios em casos de desalinhamentos severos. Além de aumentar significativamente a qualidade do alinhamento feito exclusivamente por algoritmos de registro, a técnica é simples e possui potencial para ser otimizada para diferentes cenários e necessidades.

**Palavras-chave.** Nuvens de Pontos, Registro de Nuvens de Pontos, Distância de Wasserstein

## 1 Introdução

Nuvens de pontos 3D são conjuntos de pontos distribuídos no espaço. Quando duas ou mais nuvens são capturadas de um ambiente ou objeto, escaneadas em diferentes perspectivas, com o objetivo de representá-las em ambiente digital, o registro de nuvem de pontos se faz necessário. O processo de registro de nuvens se dá pelo alinhamento dessas nuvens em um mesmo sistema de coordenadas. Frequentemente, esse processo está incluso nos *pipelines* de processamento de nuvens de pontos e, por isso, diversas técnicas e estratégias de registro foram criadas. Algumas abordagens de registro podem ser vistas em [4], [10] e [11].

Tipicamente, o registro de nuvem de pontos consiste em determinar uma nuvem como referência, fixa durante todo o processo (denominada *target*), e outra nuvem a sofrer uma transformação de modo a alinhá-la com a primeira (denominada *source*). A técnica de registro busca a transformação que melhor alinha as nuvens.

As técnicas de registro podem ser divididas em Globais e Locais quanto ao método de alinhamento [6] e frequentemente são utilizadas em conjunto para um alinhamento grosseiro (registro global) seguido de um alinhamento fino (registro local) [4]. Isso ocorre porque as técnicas costumam ter limitações, necessitando uma combinação delas no processo de registro.

Um exemplo dessas limitações pode ser visto em [10], onde é mostrado que uma técnica clássica como o ICP (*Iterative Closest Point*) [2] pode falhar em registrar adequadamente quando há uma grande diferença entre as nuvens de pontos, uma vez que o algoritmo é suscetível a cair em mínimo local que o impede de seguir se aproximando do alinhamento adequado. De fato, o desalinhamento severo entre duas nuvens pode tornar técnicas de escopo local como ICP menos efetivas. Já as técnicas globais são tipicamente menos precisas no alinhamento e utilizam alguma forma de extração

---

<sup>1</sup>patricia.martins@ifce.edu.br

<sup>2</sup>jefferson.figueiredo@ifce.edu.br

<sup>3</sup>marques@ufc.br

<sup>4</sup>george.the@ufc.br

de características, o que é um processo lento e de alto custo computacional. Não raramente, essas técnicas se baseiam em *Deep Learning*, enfatizando o nível de complexidade do problema [4].

Tendo isso em mente, neste trabalho é proposta uma estratégia de registro que considera um passo inicial de alinhamento baseado em distância de Wasserstein que aumenta a eficiência de algoritmos como o ICP em condições de severo desalinhamento. Além disso, a técnica proposta não implica em aumento significativo de custo computacional, especialmente quando o registro é feito com nuvens de pontos muito grandes. Isso ocorre porque realiza um pré-alinhamento através de um processamento indireto, cujo custo independe no número de pontos da nuvem, não necessitando de técnicas complexas de extração de características ou de operações sucessivas sobre toda a nuvem de pontos *source*. Para isso, basta extrair informações básicas da distribuição das nuvens a serem registradas: o baricentro e a matriz de covariância.

## 1.1 Distância de Wasserstein

A distância de Wasserstein pode ser entendida como uma medida do menor custo de transporte de uma distribuição de massas para uma nova configuração e posição. Assim, pode-se enxergar essa distância como uma medida de dissimilaridade entre distribuições, pois quanto mais similares, menor será o custo de transporte. Em [3] e [9], é possível encontrar mais detalhes sobre essa métrica e suas propriedades matemáticas.

Nuvens de pontos são distribuições de dados num ambiente de coordenadas virtual. Sendo assim, podem ser utilizadas com a distância de Wasserstein nas diversas aplicações. Contudo, as nuvens de pontos podem apresentar uma distribuição de alta complexidade, devido ao grande número de pontos e às inúmeras formas que um objeto pode tomar. Essa complexidade pode levar a modelos desnecessariamente complexos e custosos. Em casos assim, dados podem ser modelados com base em distribuições conhecidas, a mais comum sendo a gaussiana. Dessa forma, opta-se por achar uma gaussiana que modele segmentos da nuvem original (denominados fatias ou *slices*) a fim de achar a menor distância de Wasserstein para gaussianas, simplificando o processo de pré-alinhamento a ser investigado.

A expressão usada como critério de similaridade é dada por:

$$W_2(\mu_1, \mu_2)^2 = |m_1 - m_2|_2^2 + \text{tr}(C_1 + C_2 - 2(C_2^{1/2}C_1C_2^{1/2})^{1/2}), \quad (1)$$

em que  $\mu_i$ ,  $m_i$  e  $C_i$  representam, respectivamente, uma distribuição de probabilidade (neste trabalho, uma nuvem de pontos), seu vetor de médias e sua matriz de covariância;  $\text{tr}(\cdot)$  representa a operação matricial traço.

Uma propriedade importante dessa métrica para este experimento é o fato de que a distância de Wasserstein é variante a rotação. Isso significa que a distância entre duas nuvens iguais será zero somente quando houver sobreposição total entre elas, ou seja, translação e rotação zero entre ambas. Logo, em distribuições iguais ou aproximadamente iguais, é possível buscar uma posição ótima em que a distância de Wasserstein é a menor possível, de modo que tais nuvens irão se sobrepor de modo a encontrar a melhor superposição. O registro entre duas nuvens iguais, portanto, pode ser obtido usando essa métrica como parâmetro de qualidade. Assim, o registro de nuvens pode se beneficiar significativamente da minimização dessa distância.

Contudo, é importante ressaltar que, raramente, as nuvens a serem alinhadas são idênticas, embora costumeiramente tenham algum grau de sobreposição para possibilitar o registro. Assim, para os experimentos apresentados, buscou-se extrair *slices* parecidos entre as duas imagens, cujo grau de similaridade morfológica aparente fosse o maior possível. O objetivo é investigar se segmentos de imagens diferentes, mas suficientemente similares, podem ser usados em conjunto com a distância de Wasserstein para realizar o passo inicial do alinhamento. Nesse momento, entende-se como similaridade suficiente aquela apresentada por dois segmentos de nuvem cuja aparência visual

seja, entre si, percebida como idênticos a olho nu. A primeira coluna da Figura 2 mostra segmentos usados nos experimentos. A olho nu, quase todas parecem idênticas, mesmo nos casos em que as nuvens são efetivamente diferentes.

## 1.2 Rotação Indireta

Durante o estudo da distância de Wasserstein como critério de registro para imagens similares, notou-se que é possível determinar a distância entre *source* e *target* sem a necessidade de realizar transformações na *source* previamente. Dessa forma, a nuvem *source* só sofre modificações reais quando a transformação que resulta na menor distância de Wasserstein é encontrada.

Para entender essa estratégia, é necessário observar como é feita a transformação na nuvem de pontos que causa a translação e rotação:

$$\bar{X} = R.X + t, \quad (2)$$

em que  $\bar{X}$  é a nuvem  $X$  transformada com uma matriz de rotação  $R$  e um vetor de translação  $t$ .

Substituindo os valores de  $\bar{X}$  transformados em função de  $X$ ,  $R$  e  $t$  no cálculo da média e matriz de covariância da nuvem transformada, obtém-se o seguinte resultado:

$$\begin{aligned} \bar{m}_i &= R.m_i + t, \\ \bar{C}_i &= R.C_i.R^T, \end{aligned} \quad (3)$$

Reduz-se, dessa maneira, o custo computacional àquele de produtos de matrizes  $[3 \times 3]$  ou menor, sendo mais simples do que realizar operações diretamente sobre a nuvem, cujo tamanho pode ser da ordem de milhares ou até mesmo milhões de pontos. Esses parâmetros, uma vez calculados, podem ser substituídos diretamente na Equação 1, fornecendo-nos assim o resultado de uma transformação na nuvem, sem o custo de transformá-la. Assim, o custo computacional desse processo é invariante ao tamanho da nuvem, uma vez que sua matriz de covariância e seu baricentro vão sempre ter a mesma dimensionalidade.

Por não exigir a iteração sobre os pontos, este processo é aqui denominado Rotação Indireta. A ação preliminar do processo de registro para nuvens com desalinhamentos severos é centralizar as nuvens na origem, sendo esta uma aproximação inicial suficiente para a translação. Em seguida, busca-se indiretamente a rotação necessária para minimizar a distância de Wasserstein entre as duas nuvens.

## 2 Metodologia

O processo de registro de nuvens proposto é feito em dois passos: um alinhamento preliminar, usando a distância de Wasserstein como guia, e o alinhamento final, que fará um alinhamento fino no processo de registro [7]. Esse segundo passo é realizado usando qualquer algoritmo de registro de nuvem de pontos conhecido. Neste trabalho, usamos o ICP (*Iterative Closest Point*) Ponto ao Plano<sup>5</sup>.

Para avaliar os efeitos do alinhamento preliminar, diversos cenários com desalinhamento severo foram elaborados, sobretudo no que se refere à rotação da imagem. Para explorar os limites da estratégia proposta, foi realizada uma rotação aleatória (com os ângulos dos 3 eixos principais podendo variar de 1° a 360°) na nuvem *source*. Alguns casos em que a severidade do desalinhamento é visível foram catalogados e repetidos a fim de verificar a reprodutividade do experimento.

<sup>5</sup>Configuração do algoritmo: máximo de 30 iterações; e tolerância de 0.01 para o translado e de 0.5° para a rotação, dessa forma, o algoritmo pára quando a média da diferença em translação e rotação das 3 últimas iterações é menor que esses parâmetros.

Dentre os casos catalogados (determinados como conjuntos de ângulos para os eixos  $x$ ,  $y$  e  $z$ , respectivamente), estão:  $[252^\circ, 321^\circ, 346^\circ]$ ,  $[93^\circ, 303^\circ, 92^\circ]$ ,  $[126^\circ, 71^\circ, 91^\circ]$  e  $[137^\circ, 205^\circ, 28^\circ]$ . A Figura 2 mostra exemplos dos experimentos realizados. A segunda coluna apresenta o estado de desalinhamento severo e a terceira coluna, uma tentativa de registro usando exclusivamente o ICP ponto ao plano, resultando em um alinhamento insatisfatório. Para essa figura, foram selecionados exemplos de desalinhamentos severos em que o ICP falharia, a fim de verificar se o procedimento proposto ampliaria as situações de sucesso do algoritmo, se usado como pré-alinhamento.

## 2.1 Estrutura do Experimento

A estratégia proposta consiste em 6 etapas, que podem ser observadas na Figura 1. O experimento utiliza dados das bases de dados [8], [1] e [5]. As duas primeiras etapas do *pipeline* foram realizadas para condicionar a base de dados para o experimento e, portanto, não fazem parte do escopo deste estudo. As etapas são discutidas a seguir.

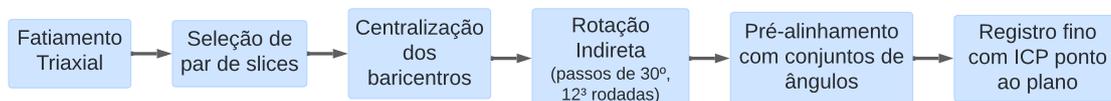


Figura 1: Pipeline da Estratégia de Registro Proposta. Fonte: Os autores.

**Fatiamento triaxial:** todas as imagens foram divididas em três fatias (*slices*) igualmente espaçadas para ambas as imagens *source* e *target*. Esse fatiamento é feito nos 3 eixos ( $x$ ,  $y$  e  $z$ ), sendo assim gerados 9 *slices* por nuvem. O método de segmentação no eixo é o mesmo visto em [7].

**Seleção de par de *slices*:** para cada experimento de registro, é selecionado um par de *slices* (um da nuvem *source* e uma da *target*). O critério de seleção é a similaridade visual, sendo a seleção feita manualmente. O entendimento é que, uma vez verificado que duas partes similares das nuvens podem gerar um bom registro com a estratégia proposta, é justificada a busca por técnicas para detectar automaticamente *slices* similares entre as nuvens *source* e *target*. Assim, a automatização desse processo encontra-se fora do escopo deste estudo.

**Centralização de baricentros:** todas as imagens são transladadas de modo que seus baricentros estejam localizados na origem. Os *slices* selecionados são transladados segundo o mesmo princípio, ou seja, obedecem o baricentro da imagem completa e não seus próprios em separado. Esse passo é utilizado como uma estratégia simples para mitigar a severidade do desalinhamento via translação.

**Rotação indireta:** corresponde ao passo de pré-alinhamento da técnica proposta, no qual é realizada uma busca em todas as combinações de ângulos possíveis entre  $0^\circ$  e  $360^\circ$  nos eixos  $x$ ,  $y$  e  $z$ , em passos de  $30^\circ$ . Essas combinações de ângulos são usadas como possíveis rotações para o *slice* da imagem *source*, porém, a transformação de rotação não é realizada diretamente a fim de reduzir o custo computacional do processo. O conjunto de ângulos é apenas usado para calcular indiretamente a distância de Wasserstein que seria obtida entre os *slices* caso a rotação fosse executada. Para dar uma volta completa a passos de  $30^\circ$ , são necessárias 12 iterações. Assim, para obter todas as combinações possíveis nos três eixos, são realizados  $12^3$  cálculos da distância de Wasserstein. Um total de 1728 iterações.

**Pré-alinhamento com conjunto de ângulos:** aqui é feita a seleção do conjunto de ângulos que gerou a menor distância de Wasserstein entre os *slices* dos modelos *source* e *target*. O conjunto de ângulos selecionado é então usado como matriz de rotação para transformar o *slice* da nuvem

*source*. Essa nova configuração será fornecida como entrada na próxima etapa, iniciando assim o registro final.

**Registro final com ICP ponto ao plano:** aqui são fornecidos ambos os *slices* para um algoritmo de registro conhecido. Neste experimento, foi usado o ICP (*Iterative Closest Point*) ponto ao plano. A matriz de transformação obtida é, então, aplicada à imagem *source* completa, obtendo-se assim o resultado final do registro.

## 2.2 Passo de Rotação

Durante os testes exploratórios, o comportamento da Distância de Wasserstein com relação à rotação relativa entre *source* e *target* foi avaliado com diversos passos de rotação: 5°, 15°, 20°, 30°, 60°, 90° e 120°. Em se tratando de um processo de busca exaustiva (testando todas as combinações possíveis), o custo computacional de tal operação é da ordem de  $O(n^3)$ . Quanto menor esse ângulo, mais iterações serão necessárias. Porém, dado que essa etapa tem por objetivo eliminar a alta severidade do desalinhamento, é desnecessário tentar uma quantidade muito grande de combinações. É, portanto, desejável um ângulo suficientemente grande que resulte no menor custo computacional possível.

Contudo, um ângulo grande demais pode inviabilizar o ajuste adequado da rotação. Um passo de 120° resulta em  $3^3 = 27$  iterações, o que é um custo bastante reduzido, porém, com resultados insatisfatórios. Já um passo de 5° resulta em  $72^3 = 373.248$  iterações, o que é um custo proibitivo. Porém, os resultados são satisfatórios.

Após uma avaliação empírica desses passos, observou-se ser suficiente rotações de 30° a cada iteração, gerando-se  $12^3 = 1.728$  possíveis combinações, o que resulta em menos que 0.5% do número de iterações com passo de 5°. O desempenho dessa configuração é de aproximadamente 0.5 segundos<sup>6</sup> para os testes realizados. A título de comparação, o ICP levou, nesses experimentos, cerca de 1 segundo a 0.1 segundo, dependendo do tamanho das nuvens de pontos.

## 3 Resultados

Nesta seção são exibidos e discutidos os resultados do experimento. Na Figura 2 são mostradas várias nuvens passando pelas etapas do experimento, onde cada linha corresponde a um modelo 3D diferente. Da esquerda para direita, as colunas representam: (I) os *slices* escolhidos; (II) as nuvens completas severamente desalinhadas com os *slices* destacados para perspectiva; (III) o resultado do registro com apenas o ICP Ponto ao Plano a título de comparação; (IV) o resultado das nuvens após o pré-alinhamento e, por último; (V) o resultado do alinhamento final proposto, combinando o pré-alinhamento da coluna anterior com o ICP Ponto ao Plano.

Nos casos em que as nuvens são iguais, é possível ver que o alinhamento resulta numa superposição total e somente uma cor pode ser vista. Esse resultado nos apresenta distância de Wasserstein zero, ou um valor muito próximo de zero. Isso fornece subsídio para justificar a busca por regiões equivalentes em um caso de registro de nuvens distintas, pois pode ser possível realizar o alinhamento apenas minimizando a distância de Wasserstein em qualquer contexto.

É possível observar que, de fato, a distância de Wasserstein contribuiu para um registro satisfatório, mesmo no caso de nuvens diferentes, desde que os *slices* utilizados sejam suficientemente similares. Isso ocorre porque, quanto mais similares são as distribuições, mais próximo da distância zero se pode chegar. Atendendo a esse requisito, a técnica torna o registro robusto a desalinhamento severo, mas com um procedimento de custo computacional invariante ao tamanho da nuvem.

<sup>6</sup>Notebook com Processador Intel(R) Core(TM) i7-3632QM CPU @ 2.20GHz, 2.20 GHz, 8,00 GB de memória RAM e placa de vídeo integrada Intel(R) HD Graphics 4000.

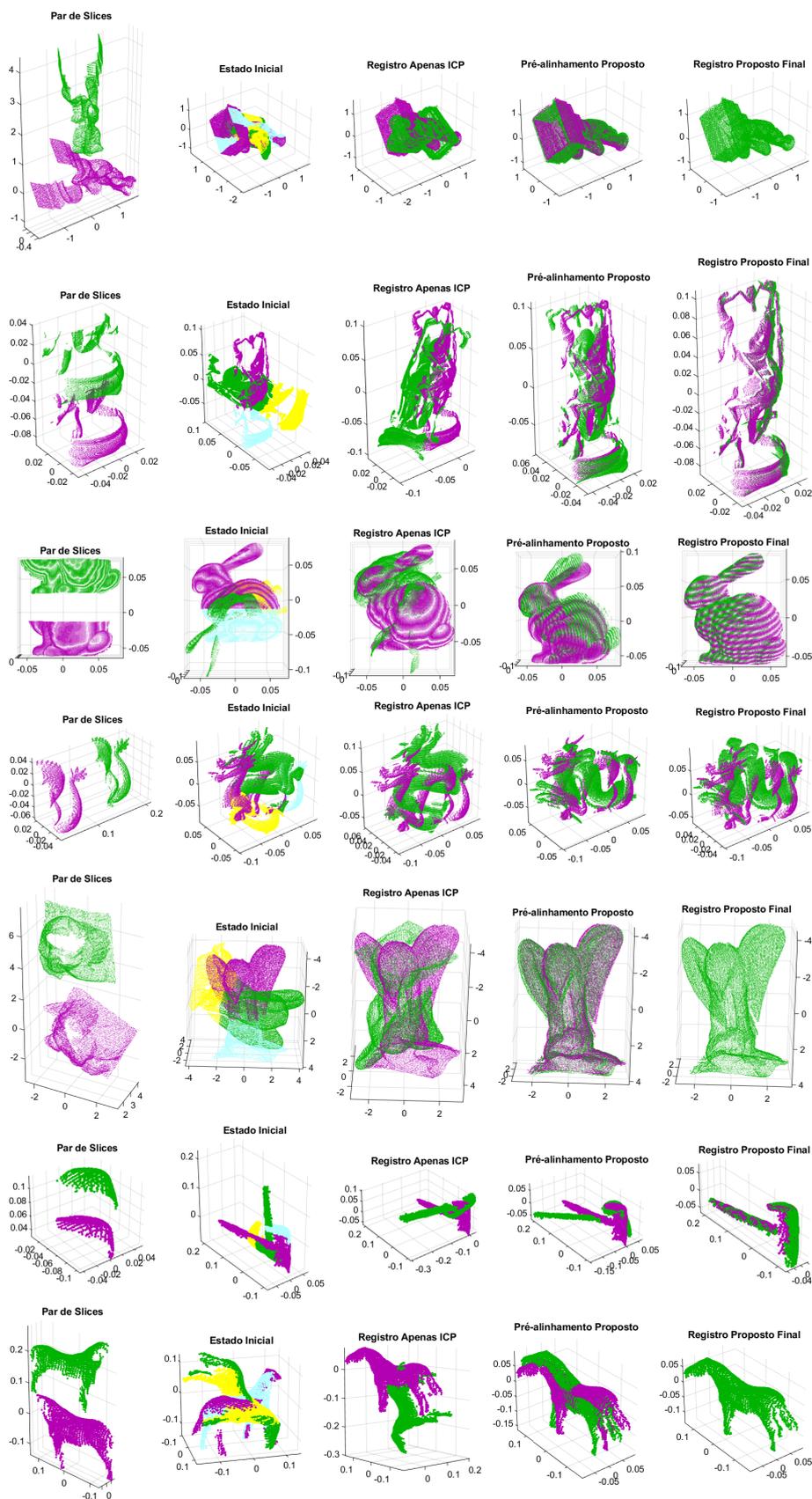


Figura 2: Nuvens *Aquarius*, *Buddha*, *Bunny*, *Dragon*, *Eagle*, *Hammer* e *Horse* em diversas situações de desalinhamento passando pelo experimento de registro. Fonte: Os autores.

## 4 Considerações Finais

A técnica foi validada, mostrando ser possível utilizar a distância de Wasserstein para mitigar os efeitos de um desalinhamento severo no processo de registro de nuvem de pontos. Tendo isso em vista, neste trabalho abre-se caminho para novas investigações. O passo de busca de  $30^\circ$  gera uma quantidade de combinações viável para a técnica realizar o registro com relativa velocidade, porém é perceptível que muitas combinações não são boas candidatas e não precisariam ser testadas e calculadas se um algoritmo utilizasse o critério da distância de Wasserstein como guia direto da busca, não apenas como critério de seleção. Técnicas de otimização como o gradiente descendente são candidatas naturais para essa análise.

## Referências

- [1] J. Aleotti, D. L. Rizzini e S. Caselli. “Perception and Grasping of Object Parts from Active Robot Exploration”. Em: **Journal of Intelligent Robotic Systems** 76.3–4 (mar. de 2014), pp. 401–425. DOI: 10.1007/s10846-014-0045-6.
- [2] P.J. Besl e N. D. McKay. “A method for registration of 3-D shapes”. Em: **IEEE Transactions on Pattern Analysis and Machine Intelligence** 14.2 (fev. de 1992), pp. 239–256. DOI: 10.1109/34.121791.
- [3] R. Bhatia, T. Jain e Y. Lim. “On the bures–wasserstein distance between positive definite matrices”. Em: **Expositiones Mathematicae** 37.2 (jun. de 2019), pp. 165–191. DOI: 10.1016/j.exmath.2018.01.002.
- [4] N. Brightman, L. Fan e Y. Zhao. “Point cloud registration: a mini-review of current state, challenging issues and future directions”. Em: **AIMS Geosciences** 9.1 (2023), pp. 68–85. ISSN: 2471-2132. DOI: 10.3934/geosci.2023005.
- [5] EPFL. **Statues Model Repository**. Online. Acessado em 11/03/2024, [https://lgg.epfl.ch/statues\\_dataset.php](https://lgg.epfl.ch/statues_dataset.php).
- [6] X. Huang, G. Mei e J. Zhang. “Feature-Metric Registration: A Fast Semi-Supervised Approach for Robust Point Cloud Registration Without Correspondences”. Em: **2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)**. IEEE, jun. de 2020. DOI: 10.1109/cvpr42600.2020.01138.
- [7] P. S. Neto, J. M. Soares e G. A. P. Thé. “Uniaxial partitioning strategy for Efficient Point Cloud Registration”. Em: **Sensors** 22.8 (abr. de 2022), p. 2887. DOI: 10.3390/s22082887.
- [8] Stanford. **The Stanford 3D Scanning Repository**. Online. Acessado em 11/03/2024, <https://graphics.stanford.edu/data/3Dscanrep/>.
- [9] A. Takatsu. “On wasserstein geometry of gaussian measures”. Em: **Advanced Studies in Pure Mathematics** 57 (jan. de 2010). DOI: 10.2969/aspm/05710463.
- [10] D. Xie, W. Zhu, F. Rong, X. Xia e H. Shang. “Registration of Point Clouds: A Survey”. Em: **2021 International Conference on Networking Systems of AI (INSAI)**. 2021, pp. 136–142. DOI: 10.1109/INSAI54028.2021.00034.
- [11] N. Xu, R. Qin e S. Song. “Point cloud registration for LiDAR and photogrammetric data: A critical synthesis and performance analysis on classic and deep learning algorithms”. Em: **ISPRS Open Journal of Photogrammetry and Remote Sensing** 8 (2023), p. 100032. ISSN: 2667-3932. DOI: <https://doi.org/10.1016/j.ophoto.2023.100032>.