



Registro Automático de Nuvens de Pontos 3D e Refinamento Global de Poses – Contribuições para o Mapeamento e Localização Simultâneo (SLAM)

Automatic Registration of 3D Point Clouds and Global Poses Refinement – Contributions to Simultaneous Localization and Mapping (SLAM)

Rubens Antonio Leite Benevides¹, Daniel Rodrigues dos Santos² e Nadisson Luis Pavan³

¹ Universidade Federal do Paraná (UFPR), Curitiba, Brasil. rubensleite11@gmail.com

ORCID: <https://orcid.org/0000-0003-2605-451X>

² Instituto Militar de Engenharia (IME), Rio de Janeiro, Brasil. daniel.rodrigues@ime.eb.br

ORCID: <https://orcid.org/0000-0001-7977-7426>

³ Universidade Federal do Paraná (UFPR), Curitiba, Brasil. nadissonluisp@gmail.com

ORCID: <https://orcid.org/0000-0002-1917-8576>

Recebido: 07.2022 | Aceito: 01.2024

Resumo: O registro de nuvens de pontos 3D e o refinamento global de poses são dois problemas fundamentais a realizar o Mapeamento e Localização Simultâneos (Simultaneous Localization and Mapping - SLAM) com sensores LIDAR. O registro de nuvens consiste em encontrar transformações de coordenadas que sobreponham localmente pares de nuvens de pontos, chamadas poses relativas. Para referenciar várias nuvens em uma origem global, várias poses relativas precisam ser compostas multiplicativamente em poses absolutas ao longo da trajetória do sensor, como as poses relativas nunca estão isentas de erros, um problema ainda mais geral surge, a deriva (drift) da trajetória do sensor. Para tratar este problema se utilizam Modelos de Refinamento Global (MRG), que refinam simultaneamente todas as poses de uma trajetória. Neste contexto, propõe-se aqui duas contribuições, a primeira, é um método de registro de pares de nuvens de pontos que integra o Fast Global Registration (FGR) e o Generalized Iterative Closest Point (GICP) em uma abordagem multicaminho e em multiescala. Para isto, cada nuvem de um dataset é registrada nas 3 posteriores, criando um grafo de poses, e cada par é sucessivamente registrado em abordagem coarse-to-fine. A segunda contribuição, se trata de um MRG linear e fechado capaz de refinar todas as poses de um circuito, sem necessidade de iterações ou definição de parâmetros. Para isto, as rotações das poses são mapeadas em quatérnios e interpoladas por meio da técnica Spherical Linear Interpolation (SLERP). Em seguida outra otimização linear baseada no modelo LUM-3D é aplicada. A combinação dos modelos foi testada em dois datasets distintos, um com sete nuvens de pontos obtidas por Laser Scanner Terrestre (LST) e outro com 901 nuvens obtidas por Laser Scanner Móvel (LSM). Em ambos, os modelos foram capazes de reconstruir totalmente os datasets e reduzir significativamente os erros de registro e deriva.

Palavras-chave: Laser Scanner. Reconstrução 3D. Otimização Global de Poses. SLAM. SLERP.

Abstract: 3D point cloud registration and global pose refinement are two fundamental problems when performing Simultaneous Localization and Mapping (SLAM) with LIDAR sensors. Cloud registration consists of finding coordinate transformations that locally overlap pairs of point clouds, called relative poses. In order to reference several clouds at a global origin, several relative poses need to be multiplicatively composed into absolute poses along the sensor trajectory, as relative poses are never error-free, an even more general problem arises, the drift of the sensor trajectory. To deal with this, Global Refinement Models (GRM) are used, which simultaneously refine all the poses in a trajectory. In this context, two contributions are proposed here: the first is a method for registering pairs of point clouds that integrates Fast Global Registration (FGR) and Generalized Iterative Closest Point (GICP) in a multipath and multiscale approach. To do this, each cloud in a dataset is registered in the next 3, creating a graph of poses, and each pair is successively registered in a coarse-to-fine approach. The second contribution is a linear, closed MRG capable of refining all the poses in a circuit, without the need for iterations or parameter setting. To do this, the rotations of the poses are mapped onto quaternions and interpolated using the Spherical Linear Interpolation (SLERP) technique. Then another linear optimization based on the LUM-3D model is applied. The combination of models was tested on two different datasets, one with seven point clouds obtained by Laser Scanner Terrestrial (LST) and the other with 901 clouds obtained by Laser Scanner Mobile (LSM). In both cases, the models were able to fully reconstruct the datasets and significantly reduce registration and drift errors.

Keywords: Laser Scanning. 3D reconstruction. Global Pose Optimization. SLAM. SLERP.

1 PROBLEMA E CONTRIBUIÇÃO

A complexidade de mapear objetos em um mundo tridimensional reside na necessidade de alterar a posição do objeto ou do sensor durante o processo. Neste contexto, o desafio do Mapeamento e Localização Simultâneo (*Simultaneous Localization and Mapping – SLAM*) emerge, visando a recuperação precisa da posição do sensor em cada momento do mapeamento, enquanto simultaneamente elabora um mapa detalhado das cenas observadas. A dificuldade intrínseca ao SLAM é a interdependência entre a precisa localização no ambiente, requerendo um mapa acurado, e a construção de um mapa exato, necessitando de uma localização precisa.

A resolução do problema SLAM envolve duas etapas principais, o *back-end* e o *front-end* (CADENA et al., 2016). O *front-end* é responsável pela obtenção dos dados. Para sensores LIDAR (*Ligh Detection and Ranging*), o *front-end* significa registrar sequencialmente múltiplos pares nuvens de pontos 3D. Normalmente, modelos robustos como o RANSAC e suas variações (TORR; ZISSERMAN, 2000), são comuns nesse processo, onde operam integrados com descritores para selecionar pares de correspondências das nuvens de pontos (RUSU et al., 2009) ou entre pares de imagens (LOWE, 2004).

Por sua vez, o *back-end* do SLAM é encarregado de calcular e otimizar trajetórias, baseando-se nas transformações entre os mapas gerados pelo *front-end*. Esta etapa permite a retroalimentação do *front-end*, aprimorando os mapas criados. O *back-end* se beneficia de diversos modelos de otimização, como o framework g2o Grisetti et al. (2010), que integra a Teoria dos Grafos e o método de Levenberg-Marquadt para resolver equações de poses. É importante ressaltar que o *front-end* depende do tipo de sensor utilizado, enquanto o *back-end* é invariante em relação aos dados ou ao sensor.

Neste estudo, apresentamos duas inovações no contexto do LIDAR-SLAM. A primeira, destinada ao *front-end* SLAM, consiste em um método totalmente automatizado para o registro de múltiplas nuvens de pontos 3D, eficaz tanto para dados de Laser Scanner Terrestre (LST) quanto para Laser Scanner Móvel (LSM). A segunda, e principal inovação, destina-se ao *back-end* SLAM. Trata-se de uma otimização global de trajetória que refina todas as poses absolutas de forma linear, eliminando a necessidade de inversões ou decomposições matriciais.

2 REGISTRO DE NUVENS DE PONTOS 3D – ESTADO DA ARTE

Um sensor LIDAR em movimento, que leia continuamente nuvens de pontos 3D, cria a eterna necessidade de registrar múltiplos pares de nuvens de pontos. Nisto o algoritmo *Iterative Closest Point* (ICP) (BESL; MCKAY, 1992) é o mais bem estabelecido na literatura para o registro de pares de nuvens de pontos 3D. O ICP opera solucionando as equações de corpo rígido 3D em cada iteração, para isto usa a Decomposição em Valores Singulares (*Singular Value Decomposition - SVD*) da Matriz Variância-Covariância (MVC) conjunta das nuvens. A solução retorna 6 parâmetros, um vetor $t \in \mathbb{R}^3$ de translação e três parâmetros de rotação em uma matriz de rotação $R \in SO(3)$. $SO(3)$ significa o grupo algébrico das matrizes ortonormais de dimensão 3. Os únicos dois parâmetros de parada do ICP são a quantidade máxima de iterações e a variação mínima de RMSE.

Diversas melhorias foram propostas para o ICP ao longo dos anos, como demonstrado em (RUSINKIEWICZ; LEVOY, 2001). Por exemplo, Chen e Medioni (1992), modificaram a função de minimização do ICP para considerar a aproximação entre pontos e planos. No entanto, essa abordagem é sensível à orientação inicial das nuvens, propensa a convergir para mínimos locais. Para mitigar isso, Segal et al. (2009), propuseram o *Generalized-ICP* (GICP), que incorpora informações das MVCs das vizinhanças de cada ponto para a minimização de distâncias.

Outra abordagem para o registro de nuvens de pontos 3D envolve modelos globais ou 'coarse', que estimam uma transformação inicial independente da orientação das nuvens. A integração desses modelos com o ICP constitui uma estratégia '*coarse-to-fine*', onde um método global inicializa um método local mais refinado. Essa combinação aproveita as vantagens de cada abordagem, garantindo uma inicialização eficaz pelos métodos globais e a precisão do ICP. A vantagem desta combinação é que os ganhos de um método compensam as desvantagens do outro.

Muitos modelos globais utilizam amostragem aleatória e descritores, como o 4-Point Congruent Set (4PCS) de Aiger et al. (2008), que identifica bases de quatro pontos coplanares invariantes às transformações de corpo rígido 3D. Este modelo foi expandido por (BUENO et al., 2018) para o *4-Planes Congruent Set*, uma generalização direta que utiliza bases de 4 planos. O modelo é mais rápido devido ao fato de que planos são mais escassos que pontos, mas não pode ser utilizado, por exemplo, para o registro de nuvens que compreendam ambientes naturais, como cavernas (espeleologia) e florestas, dado que em geral não há feições planas nestes sítios.

Outros modelos descrevem pontos por histogramas invariantes a transformações, sendo um exemplo proeminente o *Fast Point Feature Histogram* (FPFH), de Rusu et al. (2009), autor principal da *Point Cloud Library* (PCL) (RUSU; COUSINS, 2011). O FPFH ganhou destaque em diversos estudos, como em Zhou et al. (2016), que integraram o FPFH no método de registro global *Fast Global Registration* (FGR). Zhou é também o principal autor da biblioteca *Open3D*, amplamente utilizada para processamento de nuvens de pontos 3D em Python (ZHOU et al., 2018).

Combinar diferentes modelos e descritores é uma prática comum, como ilustrado por Yang et al. (2021) que experimentou a integração do 4PCS com descritores como o SIFT-3D (RISTER et al., 2017), ISS (ZHONG, 2009) e HARRIS-3D (SIPIRAN; BUSTOS, 2011). Esses descritores, inspirados em técnicas de registro de imagens, exploram os autovalores da MVC e Diferenças de Gaussianas (*Difference of Gaussians - DoG*) em 3D (HARRIS; STEPHENS, 1988), (LOWE, 2004). Para o registro de nuvens de LST, Theiler et al. (2015) obtiveram bons resultados em ambientes internos e externos combinando o 4PCS com *keypoints* obtidos por DoG (THEILER et al., 2014).

Combinações de descritores e métodos de registro global e local funcionam bem, todavia, por melhor que seja o método de registro adotado, os pares de nuvens nunca estão livres de erros, pois as correspondências nunca são exatas. Logo, a composição multiplicativa de poses relativas, estimadas entre pares subsequentes, para referenciar várias nuvens em uma origem global, por meio de poses absolutas, dá origem ao erro de deriva (*drift*), que sempre cresce com a trajetória do sensor. Para mitigar este erro, são empregados Modelos de Refinamento Global (MRG) no *back-end* de métodos SLAM (CADENA et al., 2016).

Uma das formas do MRG operar nas poses de um circuito, é representando por grafos $g = (v, a)$ a trajetória da plataforma. Nestes grafos, cada vértice v representa uma estação do sensor, e cada aresta a que conecte dois vértices/estações, representa uma pose relativa. Grafos, em geral, definem problemas de otimização extremamente não lineares (WILLIAMS; BENNAMOUN, 2001), pois o espaço para otimização simultânea das poses é enorme. Em um circuito com centenas de poses, milhares de parâmetros devem ser estimados simultaneamente, similarmente ao que ocorre nos modelos de *Bundle Adjustment* (TRIGGS et al., 1999).

Primeiramente, um MRG precisa lidar com a topologia do grafo de poses. Em grafos acíclicos, por exemplo, Huber e Hebert (2003), exploram caminhos de menor erro, através da Árvore de Expansão Mínima (*Minimum Spanning Tree - MST*), para então aplicar a solução de Neugebauer (1997) nas equações das poses. Em grafos que formam ciclos, a detecção de lugares revisitados permite distribuir o *drift* propagando-o para poses anteriores ao impor a restrição de fechamento do circuito (*loop-closure*), algo similar ao fechamento de poligonais. Em Theiler et al. (2015), estas duas abordagens são combinadas no grafo completo de estações do LST, onde, posteriormente, as equações são resolvidas pelo MRG de Borrmann et al. (2008). Este MRG ficou conhecido como LUM-3D por ser uma generalização do trabalho de Lu e Milios (1997) em poses tridimensionais.

Para evitar o enorme espaço de busca dos grafos completos, Dong et al. (2018), adotam vetores de descritores na construção do grafo de poses, posteriormente, aplicam o MRG de Williams e Bennamoun (1999) nas equações. Trabalhando com dados RGB-D, Choi et al. (2015), aplicam a estatística robusta de Black e Rangarajan (1996), no grafo de poses, em seguida, executam o *framework graph-2-optimization* (g^2o) de Kümmerle et al. (2011). É válido destacar que o modelo g^2o é adotado em bibliotecas de código aberto como a *Open3D* (ZHOU et al., 2018), e também em *softwares* comerciais, tal qual o MATLAB®.

No contexto do *Kinect*, Weber et al. (2015), refinaram globalmente o grafo de poses ponderando as arestas com base nas correspondências do descritor FPFH, e posteriormente utilizando MST. No contexto de

LSM, Vlaminck et al. (2019), registraram nuvens 3D com o 4PCS e aplicaram um MRG que distribui o erro nas rotações usando a técnica *Spherical Linear Interpolation* (SLERP) (SHOEMAKE, 1985). Em paralelo, Pavan e Dos Santos (2017), utilizaram quatérnios e SVD para refinar as rotações em grafos fechados de poses LST. Em seguida, aplicaram a solução LUM-3D no ajuste das translações.

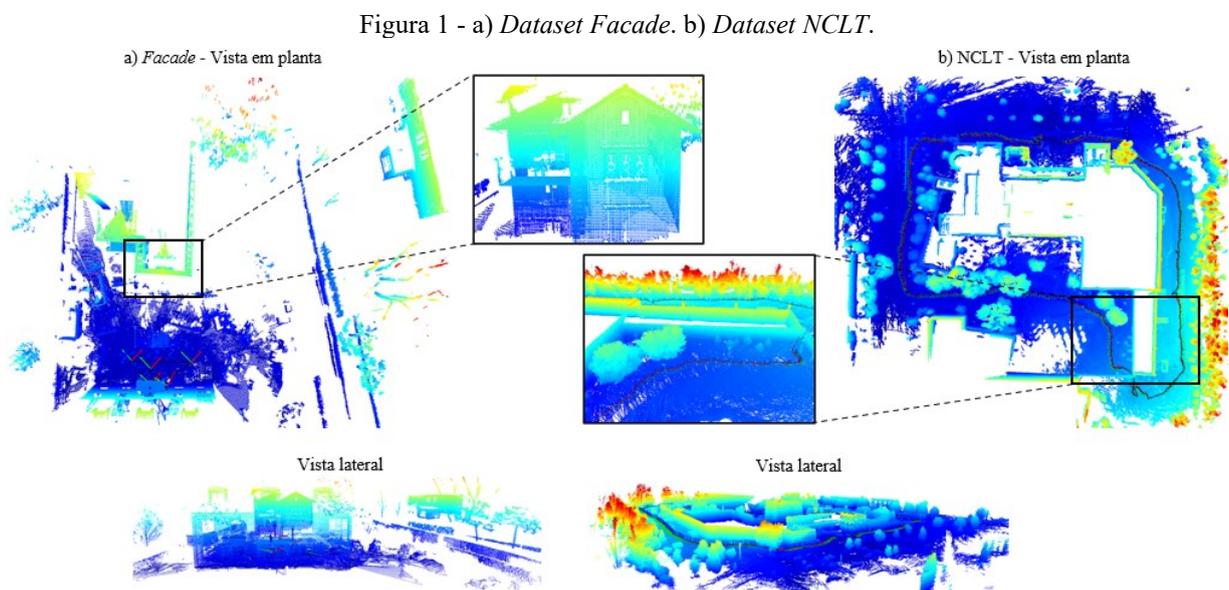
Com exceção do trabalho de Pavan e Dos Santos (2017), que utiliza SVD, todos os trabalhos citados solucionam as equações das poses por meio de métodos iterativos, tais como Gauss-Newton, que em nossa literatura é chamado de Método dos Mínimos Quadrados (MMQ), outros métodos são o Gradiente Descendente e Levenberg-Marquadt, sendo este último uma combinação dos dois anteriores (GAVIN, 2019). Em vista disto, modelos que operam de forma linear são de grande valor na literatura. Neste sentido, propomos um MRG linear, deduzido de forma fechada (*closed-form*), que não necessita inverter ou decompor matrizes. Este modelo aproveita a restrição de fechamento de circuitos de forma análoga aos procedimentos em poligonais (*loop-closure constraint*), e é inspirado em soluções como a de (VLAMINCK et al., 2019).

3 DATASETS

Para a avaliação das propostas de registro em pares e do Modelo de Refinamento Global (MRG), são empregados dois datasets distintos. O primeiro, denominado 'Facade', é composto por sete nuvens de pontos obtidas através de Laser Scanner Terrestre (LST) por (THEILER et al., 2015) em um ambiente externo. Estas nuvens são densas, contendo mais de 20 milhões de pontos cada, com densidade variável devido à natureza angular da aquisição pelo LST. Elas apresentam uma sobreposição média de 60% entre si e retratam um ambiente urbano, incluindo fachadas, vegetação e uma variedade de artefatos móveis, como carros e pessoas em movimento. O propósito principal deste dataset é testar a eficácia do registro de pares de nuvens.

O segundo dataset faz parte do projeto *North Campus Long Term* (NCLT), sendo constituído por um circuito de 645 metros compreendendo 901 nuvens de pontos, todas capturadas por Laser Scanner Móvel (LSM). A sobreposição média entre as nuvens nesse dataset é de 70%, mas reduz-se para até 30% em áreas onde a plataforma executa curvas. Cada nuvem do dataset possui aproximadamente 30 mil pontos, estão presentes elementos como ruas, vegetação, carros, edifícios e artefatos móveis. O objetivo central deste dataset é avaliar a capacidade do MRG em distribuir o erro de deriva (*drift*) ao longo do circuito de poses, realizando o fechamento do circuito.

Todas as nuvens de pontos foram processadas utilizando scripts desenvolvidos em Python 3 com a biblioteca Open3D. Os scripts estão disponíveis em um repositório público no Github. O hardware utilizado neste trabalho consiste de um processador Intel i3-9400KF (4.6 GHz) com 4 núcleos e 8 GB de memória RAM, sem o auxílio de GPU. A Figura 1 ilustra os datasets, em a) e b) estão *Facade* e o circuito NCLT, respectivamente.

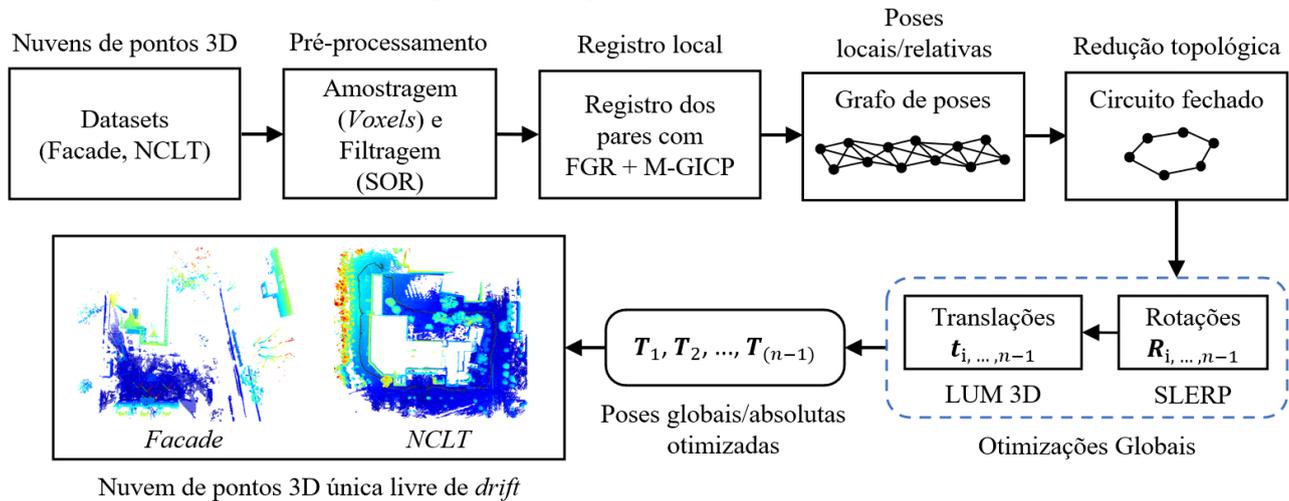


Fonte: Os autores (2022).

4 METODOLOGIA

A Figura 2 apresenta um fluxograma que detalha o processo de registro local e a aplicação do MRG proposto, culminando na geração de modelos 3D a partir das nuvens de pontos. Neste estudo, parte-se do pressuposto de que o circuito de nuvens é fechado, ou seja, a última nuvem de cada dataset possui sobreposição com a primeira. Em cenários onde essa condição não é pré-estabelecida, seria necessário implementar um método de detecção de locais revisitados, o que não é abordado neste trabalho. As etapas do processo são descritas e detalhadas a seguir.

Figura 2 - Fluxograma dos dados do trabalho



Fonte: Os autores (2022).

Conforme ilustrado na Figura 2, o procedimento inicia com o pré-processamento dos datasets. Posteriormente, realiza-se o registro par-a-par de nuvens com o método *coarse-to-fine* FGR+M-GICP proposto, resultando na construção de um grafo de poses. Este grafo é então simplificado topologicamente para formar um circuito fechado, sobre o qual o MRG proposto é aplicado. O MRG divide-se em duas fases: inicialmente, aplica-se a interpolação linear esférica (SLERP) nas matrizes de rotação R_i das poses; em seguida, emprega-se o modelo LUM-3D para as translações t_i . O resultado desse processo são poses globais otimizadas T_i , onde $i = 1, 2, \dots, n$, sendo n a quantidade de nuvens de pontos dentro do *dataset*. Essas poses globais otimizadas são então aplicadas nas nuvens dos *datasets* para obter uma reconstrução 3D automática e otimizada.

4.1 Pré-processamento

O pré-processamento é um procedimento crítico que envolve a amostragem e filtragem das nuvens de pontos, sendo essencial para a diminuição do volume de dados e correção das variações de densidade. Estas ações são fundamentais, pois conferem um balanceamento na distribuição dos pontos, mitigando potenciais enviesamentos locais durante o registro. Realiza-se a amostragem por meio do cálculo dos centroides dos voxels, termo que deriva de 'volume-elements', originados da divisão espacial da nuvem em octantes. A escolha de voxels com dimensão de 10 cm é baseada no equilíbrio estabelecido entre a fidelidade geométrica e a eficiência da amostragem para os processos de registro, conforme sugerido por Theiler et al. (2015) e Dong et al. (2018). Para a filtragem, emprega-se o método Statistical Outlier Removal (SOR), detalhado por Rusu e Cousins (2011).

4.2 Proposta de Registro Local – FGR + GICP em Multiescala

No processo de registro local dos pares de nuvens, emprega-se uma abordagem *coarse-to-fine* que integra o Fast Global Registration (FGR) e o Generalized-ICP (GICP). Aproveita-se da ampla sobreposição existente entre nuvens adjacentes ao tentar o registro de cada nuvem com as três subsequentes no dataset.

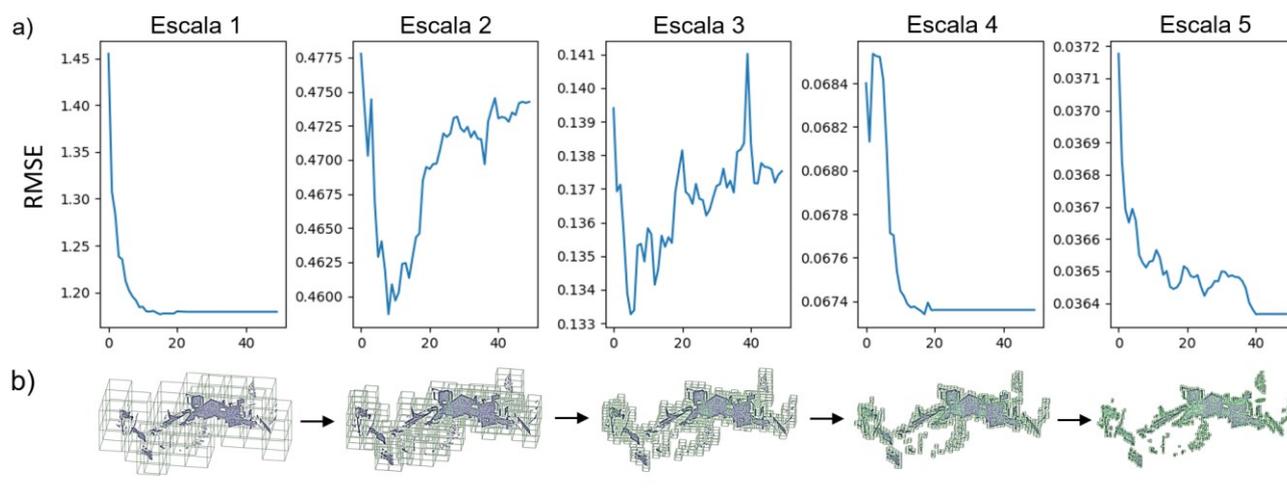
Essa estratégia estabelece um grafo de poses, no qual cada vértice pode se conectar com até seis outros vértices, representando as arestas do grafo. Estas arestas representam as poses relativas e carregam informações obtidas durante o registro dos pares de nuvens, tais como o número de correspondências encontradas utilizando o descritor Fast Point Feature Histogram (FPFH), expressas em termos de sobreposição percentual, e o Root Mean Square Error (RMSE) das correspondências. A sobreposição é calculada pela razão entre o número de pontos correspondentes e o total de pontos na nuvem alvo, expressa em porcentagem; valores mais altos indicam resultados melhores. O RMSE é mensurado em metros, com valores inferiores indicando uma precisão superior.

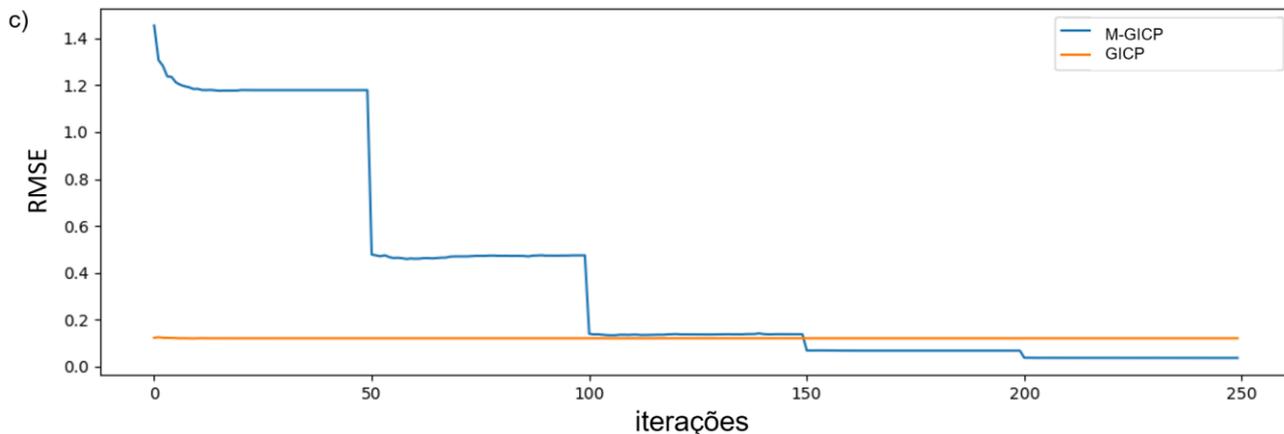
Para garantir mais robustez na etapa *fine*, o GICP é implementado em multiescala por amostragens com *voxels* decrescentes, definindo o que chamaremos de M-GICP. Conforme (PARK et al., 2017), isto torna qualquer modalidade do ICP mais rápida sem comprometer a precisão. Além disso, uma função de ponderação com métrica *L1*, como sugere Babin et al. (2019), é adotada entre as correspondências. Esta função é livre de parâmetros, pois os pares são ponderados apenas pelo inverso da distância.

O método M-GICP inicia com a transformação estimada pelo FGR, aplicando o registro do par de nuvens em uma sequência de escalas progressivamente mais detalhadas a partir do voxel de amostragem inicial. Em escalas maiores, o voxel ampliado resulta em uma amostragem significativa da nuvem, o que contribui para suavizar mínimos locais e previne a redução prematura do RMSE. Posteriormente, em escalas mais finas, procede-se à amostragem com *voxels* progressivamente menores, utilizando os resultados da escala anterior para inicializar o GICP. Este procedimento permite que a convergência do M-GICP seja acelerada pelo uso reduzido de pontos em escalas mais abrangentes, assegurando, simultaneamente, elevada precisão em escalas mais detalhadas, evitando a convergência para mínimos locais. Assim, alcança-se um equilíbrio entre velocidade e precisão, aumentando a robustez do GICP em relação à orientação inicial das nuvens de pontos. O tamanho do voxel de amostragem inicial *v* é o único parâmetro definido para todo o processo de registro par-a-par FGR+M-GICP.

A Figura 3 ilustra o funcionamento do M-GICP. Na subfigura a), observa-se a evolução do RMSE por iteração e por escala para o par de nuvens 0-1 do dataset *Facade*. Na subfigura b), visualizam-se os *voxels* de tamanho decrescente que são empregados na amostragem. A subfigura c) faz a concatenação de todos os gráficos da subfigura a) e compara com o RMSE obtido no GICP convencional. Por fim, a subfigura d) exibe o resultado entre o GICP convencional e o M-GICP em um mesmo par de nuvens.

Figura 3 – a) RMSE vs Iteração vs Escala. b) Amostragem sucessiva da nuvem por *voxels* decrescentes. c) RMSE vs Iteração para GICP e M-GICP.

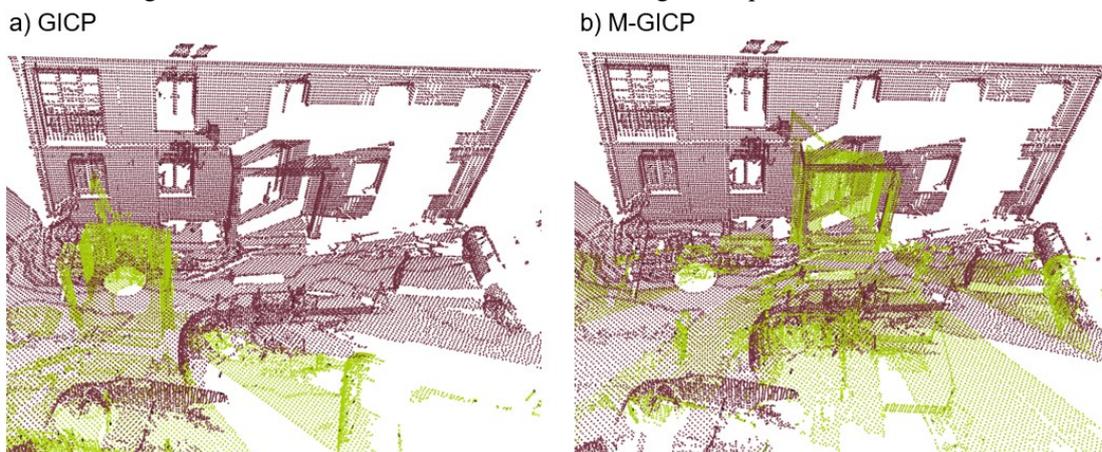




Fonte: Os autores (2022)

Na Figura 3 a) está o RMSE de 5 escalas calculadas com o *voxel* inicial como segue: $5v$, $2.5v$, v , $v/2$ e $v/4$. Os octantes definidos pelos voxels de amostragem podem ser vistos na Fig. 3 b). Na Figura 3 c) está a comparação dos gráficos de RMSE do GICP e do M-GICP. Ambos os algoritmos foram inicializados com o mesmo par de nuvens, na mesma orientação e utilizando o mesmo *voxel* de amostragem: 10 cm. Os critérios de parada também são iguais, $\Delta RMSE = 10^{-6}$ m e 250 iterações para cada. O resultado visual de ambas as formas de registro é apresentado na Figura 4 a) e b), respectivamente.

Figura 4 – Par de nuvens 0-1 do dataset Facade registrado por GICP e M-GICP.



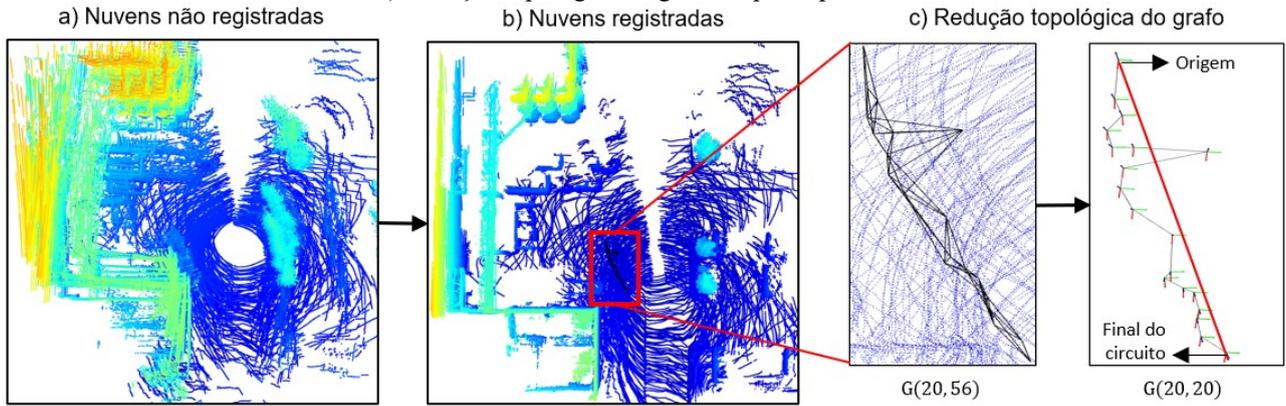
Fonte: autores (2022)

Na Figura 4, se percebe que o M-GICP (Figura 4b) sobrepôs as nuvens de pontos corretamente, por outro lado, o GICP tradicional (Figura 4a), convergiu em uma posição praticamente igual a configuração inicial do par. Voltando à Figura 3 c), nela é possível entender por que o GICP, mesmo após todas as iterações, praticamente não mudou a nuvem móvel (amarela) do lugar. Isto acontece porque, no *voxel* inicial de 10 cm, rapidamente o GICP converge para o mínimo local em função de suas correspondências próximas, o que impede a nuvem de se movimentar. Esse mínimo local é evidenciado na invariância do RMSE do GICP tradicional, que praticamente não muda ao longo das 250 iterações.

A Figura 3 a) mostra que não se tem o máximo de vantagem do M-GICP com os atuais critérios de parada, que estão iguais por escala. Percebe-se, na Figura 3 a), que apenas 10 iterações bastaram para a primeira escala convergir, e, nesta escala, ocorreu a maior variação do RMSE. Para fazer um *fine-tuning* apropriado, é razoável supor, que, nas escalas iniciais são necessárias mais iterações para o M-GICP convergir, pois há maior liberdade de movimentação das nuvens nessas escalas. Por outro lado, nas escalas mais finais, supõe-se que o par estará mais próximo do mínimo global, necessitando de menos iterações. Isto também é válido para a variação mínima de RMSE ($\Delta RMSE$), que deve ser menos restritiva no início e mais restritiva no final. Em face dessas observações e dos testes conduzidos, adotou-se os seguintes critérios de parada por escala: $\Delta RMSE = 10^{-3}$, 10^{-3} , 10^{-4} , 10^{-5} , 10^{-6} , valores em metros; *iterações* = 50, 40, 30, 30, 20.

Após a aplicação do método FGR+M-GICP nos pares de nuvens, obtém-se um grafo de poses. Para transformá-lo em um circuito fechado, realiza-se uma redução topológica, escolhendo as arestas mais adequadas do grafo com base em critérios específicos. A Figura 5 demonstra este procedimento usando as 20 primeiras nuvens do dataset NCLT. Na Figura 5 a), são exibidas as nuvens antes do registro; a Figura 5 b) mostra as nuvens após a reconstrução pelo método de registro proposto; e a Figura 5 c) ilustra a redução topológica do grafo de poses, resultando em um número igual de vértices e arestas. Esta redução é feita selecionando arestas que apresentam simultaneamente o menor RMSE e a maior sobreposição, garantindo assim a qualidade e a precisão do circuito fechado.

Figura 5 – a) 20 primeiras nuvens do *dataset* NCLT não registradas. b) O mesmo conjunto de 20 nuvens após o registro FGR+M-GICP. c) Redução topológica do grafo de poses para um circuito fechado.



Fonte: Os autores (2022).

4.3 Proposta de MRG - Refinamento global das rotações por SLERP

Definido um circuito fechado de poses relativas, as matrizes de rotação das poses são mapeadas para quatérnios unitários $|\mathbf{q}_i| = 1$, onde \mathbf{q}_i é o quatérnio que representa a rotação relativa do vértice X_i referenciado no vértice X_{i-1} , com $i = 1, 2, \dots, n$; sendo n a quantidade de nuvens do *dataset*. Os quatérnios são então compostos para a origem global, seguindo a sequência de multiplicações conforme a Eq. (1):

$$\begin{aligned}
 \mathbf{r}_0 &= \mathbf{q}_I \\
 \mathbf{r}_1 &= \mathbf{q}_1 \mathbf{r}_0 \\
 \mathbf{r}_2 &= \mathbf{q}_2 \mathbf{r}_1 \\
 &\vdots \\
 \mathbf{r}_n &= \mathbf{q}_n \mathbf{r}_{n-1} \\
 \mathbf{r}_{n+1} &= \mathbf{q}_{n+1} \mathbf{r}_n
 \end{aligned} \tag{1}$$

onde, \mathbf{q}_I é a identidade dos quatérnios; \mathbf{r}_i o quatérnio que representa a rotação global do vértice X_i referenciado no vértice X_0 , com $i = 1, 2, \dots, n$. Se as estimativas das rotações não apresentassem erros, o quatérnio \mathbf{r}_{n+1} seria igual à identidade. Na prática, esse quatérnio apresenta uma aproximação com a identidade, pois os erros se acumulam. A composição de rotações pode ser feita no sentido contrário da trajetória do sensor, então, as rotações dos vértices X_1, X_2, \dots, X_n referenciadas no vértice X_0 , podem ser obtidas conforme a Eq. (2):

$$\begin{aligned}
 \mathbf{p}_1 &= (\mathbf{q}_2, \mathbf{q}_3, \dots, \mathbf{q}_{n+1})^{-1} \\
 \mathbf{p}_2 &= (\mathbf{q}_3, \mathbf{q}_4, \dots, \mathbf{q}_{n+1})^{-1} \\
 &\vdots \\
 \mathbf{p}_n &= (\mathbf{q}_{n+1})^{-1}
 \end{aligned} \tag{2}$$

onde, \mathbf{p}_i é o quatérnio que representa a rotação global do vértice X_i referenciado no vértice X_0 , com $i = 1, 2, \dots, n$. Para otimizar as rotações, os quatérnios são interpolados com a SLERP, conforme a Eq. (3):

$$\begin{aligned} \hat{q}_1 &= r_1(r_1^{-1}p_1)^{\tau_1} \\ \hat{q}_2 &= r_2(r_2^{-1}p_2)^{\tau_2} \\ &\vdots \\ \hat{q}_n &= r_n(r_n^{-1}p_n)^{\tau_n} \end{aligned} \tag{3}$$

onde, \hat{q}_i é o quatérnio interpolado entre r_i e p_i , que representa a rotação ótima do vértice X_i referenciado ao vértice X_0 , com $i = 1, 2, \dots, n$. Os valores $\tau_i \in [0,1]$ são os intervalos de interpolação, com $i = 1, 2, \dots, n$. As propriedades de multiplicação dos quatérnios permitem escrever p_i como:

$$\begin{aligned} p_1 &= r_1 r_{n+1}^{-1} \\ p_2 &= r_2 r_{n+1}^{-1} \\ &\vdots \\ p_n &= r_n r_{n+1}^{-1} \end{aligned} \tag{4}$$

substituindo a Eq. (4) na Eq. (3) se obtém:

$$\begin{aligned} \hat{q}_1 &= r_1(r_1^{-1}r_1 r_{n+1}^{-1})^{\tau_1} = r_1 r_{n+1}^{-\tau_1} \\ \hat{q}_2 &= r_2(r_2^{-1}r_2 r_{n+1}^{-1})^{\tau_2} = r_2 r_{n+1}^{-\tau_2} \\ &\vdots \\ \hat{q}_n &= r_n(r_n^{-1}r_n r_{n+1}^{-1})^{\tau_n} = r_n r_{n+1}^{-\tau_n} \end{aligned} \tag{5}$$

para otimizar os intervalos τ_i , toma-se vantagem da restrição de que o inverso de \hat{q}_i , ao ser multiplicado por $q_i \hat{q}_{i-1}$, deve se aproximar o máximo possível do quatérnio identidade, como segue:

$$\begin{aligned} \hat{q}_1^{-1}q_1 &= (r_1 r_{n+1}^{-\tau_1})^{-1}q_1 \approx q_I \\ \hat{q}_2^{-1}q_2 \hat{q}_1 &= (r_2 r_{n+1}^{-\tau_2})^{-1}q_2 r_1 r_{n+1}^{-\tau_1} \approx q_I \\ &\vdots \\ \hat{q}_n^{-1}q_n \hat{q}_{n-1} &= (r_n r_{n+1}^{-\tau_n})^{-1}q_n r_{n-1} r_{n+1}^{-\tau_{n-1}} \approx q_I \\ q_{n+1} \hat{q}_n &= q_{n+1} r_n r_{n+1}^{-\tau_n} \approx q_I \end{aligned} \tag{6}$$

com as devidas inversões de quatérnios e a substituição da Eq. (1) na Eq. (6), a sequência acima se torna:

$$\begin{aligned} \hat{q}_1^{-1}q_1 &= r_{n+1}^{\tau_1} r_1^{-1} r_1 = r_{n+1}^{\tau_1} \approx q_I \\ \hat{q}_2^{-1}q_2 \hat{q}_1 &= r_{n+1}^{\tau_2} r_2^{-1} r_2 r_{n+1}^{-\tau_1} = r_{n+1}^{\tau_2} r_{n+1}^{-\tau_1} \approx q_I \\ &\vdots \\ \hat{q}_n^{-1}q_n \hat{q}_{n-1} &= r_{n+1}^{\tau_n} r_n^{-1} r_n r_{n+1}^{-\tau_{n-1}} = r_{n+1}^{\tau_n} r_{n+1}^{-\tau_{n-1}} \approx q_I \\ q_{n+1} \hat{q}_n &= r_{n+1} r_{n+1}^{-\tau_n} \approx q_I \end{aligned} \tag{7}$$

agora, utilizando a propriedade da multiplicação de potência de quatérnios iguais $q^{\tau_1} q^{\tau_2} = q^{\tau_1 + \tau_2}$, tem-se a Eq. (7) reescrita como a Eq. (8):

$$\begin{aligned} \hat{q}_1^{-1}q_1 &= r_{n+1}^{\tau_1} \approx q_I \\ \hat{q}_2^{-1}q_2 \hat{q}_1 &= r_{n+1}^{\tau_2} r_{n+1}^{-\tau_1} = r_{n+1}^{\tau_2 - \tau_1} \approx q_I \\ &\vdots \\ \hat{q}_n^{-1}q_n \hat{q}_{n-1} &= r_{n+1}^{\tau_n} r_{n+1}^{-\tau_{n-1}} = r_{n+1}^{\tau_n - \tau_{n-1}} \approx q_I \\ q_{n+1} \hat{q}_n &= r_{n+1} r_{n+1}^{-\tau_n} = r_{n+1}^{1 - \tau_n} \approx q_I \end{aligned} \tag{8}$$

assim como os números reais, todo quatérnio elevado a 0 deve ser igual à identidade. Então, para obter intervalos τ_i ótimos, é necessário minimizar o sistema de equações formado nos expoentes da Eq. (8). Este é

um sistema de equações lineares simples, como mostra a Eq. (9):

$$\begin{aligned} \tau_1 &= \theta_1 \\ \tau_2 - \tau_1 &= \theta_2 \\ &\vdots \\ \tau_n - \tau_{n-1} &= \theta_{n-1} \\ 1 - \tau_n &= \theta_n \end{aligned} \tag{9}$$

a solução que minimiza esse sistema, que tem apenas 1 grau de liberdade, é obtida por MMQ, conforme a Eq. (10):

$$\hat{\tau} = (J^T J)^{-1} J^T \mathbf{b} \tag{10}$$

onde J é a matriz jacobiana da Eq. (9) em relação às variáveis $\tau_1, \tau_2, \dots, \tau_n$; \mathbf{b} é o vetor $(\theta_1, \theta_2, \dots, \theta_n)$. A solução da Eq. (10) fornece o vetor de intervalos da SLERP, $\hat{\tau} = (\tau_1, \tau_2, \dots, \tau_n)$. Observe que o formato da Eq. (9) é função apenas da quantidade de vértices, para um circuito fechado de n vértices, tem-se $n - 1$ equações, portanto, o vetor $\hat{\tau}$ é previamente conhecido, pois J e \mathbf{b} seguem um padrão simples, conforme a Eq. (11):

$$J = \begin{bmatrix} 1 & 0 & \dots & 0 & 0 \\ -1 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & -1 & 1 \\ 0 & 0 & \dots & 0 & -1 \end{bmatrix}; \mathbf{b} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \tag{11}$$

em um circuito fechado com $n = 5$ vértices, a solução de $\hat{\tau}$ é dada pela Eq. (12):

$$\hat{\tau} = [\tau_1 \quad \tau_2 \quad \tau_3 \quad \tau_4] = \left[\frac{1}{5} \quad \frac{2}{5} \quad \frac{3}{5} \quad \frac{4}{5} \right] \tag{12}$$

isto é, a solução segue a série da Eq. (13):

$$\frac{1}{n}, \dots, \frac{n-2}{n}, \frac{(n-1)}{n} \tag{13}$$

A dedução por MMQ mostra que o intervalo $[0,1]$ é dividido em n partes lineares. A interpretação geométrica deste resultado é simples, a interpolação ótima entre \mathbf{r}_i e \mathbf{p}_i , é mais próxima de \mathbf{r}_i na primeira metade do circuito, e mais próxima de \mathbf{p}_i na segunda metade. Quanto mais rotações forem acumuladas no quatérnio \mathbf{r}_i , mais próximo do quatérnio \mathbf{p}_i a interpolação ocorrerá, e vice-versa, pois \mathbf{p}_i é obtido no sentido inverso do circuito. Em um circuito com uma quantidade de vértices pares, um dos pares de quatérnios será interpolado exatamente em $\frac{1}{2}$, pois ir pelo sentido direto ou inverso compõe a mesma quantidade de rotações em ambos os quatérnios.

4.4 Proposta de MRG – Refinamento global das translações por LUM-3D

O refinamento global das translações, chamado de LUM-3D, é explicado de forma suficiente para fins de implementação, mas não é deduzido. Para mais detalhes sugere-se a consulta do artigo original de Lu e Milios, (1997) e sua adaptação para 3D de Borrmann et al. (2008). Seguindo o exemplo anterior, em circuito fechado de 5 vértices e 5 arestas, primeiro, constrói-se a hipermatriz jacobiana \mathbf{H} , conforme a Eq. (14):

$$H = \begin{bmatrix} I & 0 & 0 & 0 \\ -I & I & 0 & 0 \\ 0 & -I & I & 0 \\ 0 & 0 & -I & I \\ 0 & 0 & 0 & -I \end{bmatrix}; \text{ onde } I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (14)$$

segundo, monta-se o vetor de observações \mathbf{d} , como mostra a Eq. (15):

$$\mathbf{d} = \begin{bmatrix} \mathbf{R}_I \mathbf{t}_{10} \\ \widehat{\mathbf{R}}_{10} \mathbf{t}_{21} \\ \widehat{\mathbf{R}}_{20} \mathbf{t}_{32} \\ \widehat{\mathbf{R}}_{30} \mathbf{t}_{43} \\ \widehat{\mathbf{R}}_{40} \mathbf{t}_{04} \end{bmatrix} \quad (15)$$

onde, \mathbf{R}_I é a rotação identidade; $\widehat{\mathbf{R}}_{i,0}$ são as matrizes de rotação obtidas dos quatérnios interpolados $\widehat{\mathbf{q}}_i$; e $\mathbf{t}_{i,i-1}$ são as translações relativas dos pares do circuito. A solução que minimiza o erro das translações é dada pelo MMQ conforme a Eq. (16):

$$\hat{\mathbf{t}} = (\mathbf{H}^T \mathbf{H})^{-1} (\mathbf{H}^T \mathbf{d}) \quad (16)$$

sendo $\hat{\mathbf{t}}$ o vetor das translações absolutas ($\mathbf{t}_{40}, \mathbf{t}_{30}, \dots, \mathbf{t}_{10}$) refinadas. Novamente, como a matriz \mathbf{H} cresce seguindo um padrão simples, $(\mathbf{H}^T \mathbf{H})^{-1}$ é previamente conhecida. Logo, toda a otimização global das poses é feita sem a necessidade de inversão ou decomposição de matrizes. Resumidamente, as rotações são refinadas com a SLERP e as translações são refinadas por LUM-3D, definindo nossa proposta de MGR. Esta solução para a otimização global de poses é linear, livre de parâmetros e sem necessidade de iterações, algo que os autores desconhecem na literatura de refinamento global de poses.

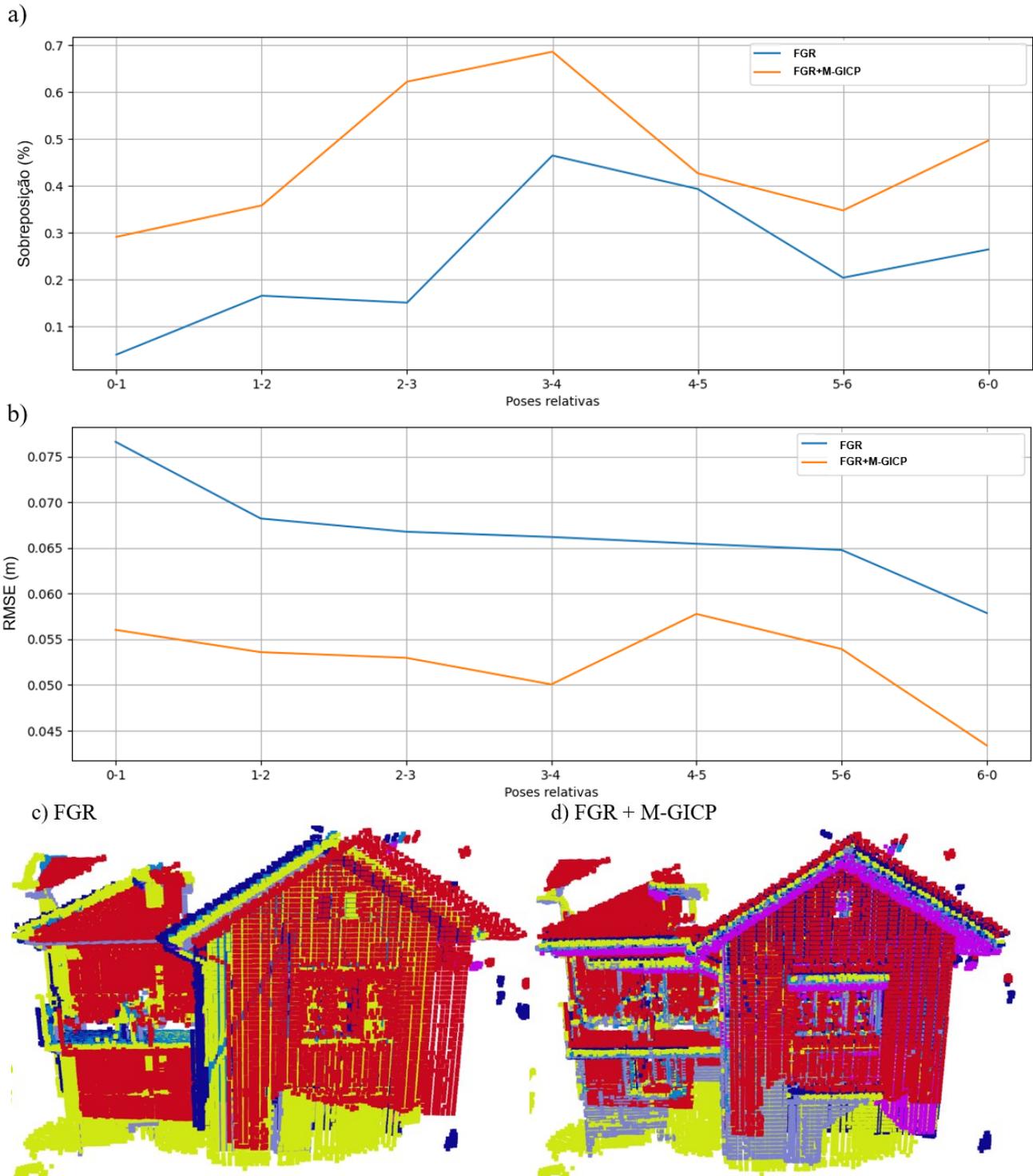
5 RESULTADOS

A seguir, o registro *coarse-to-fine* FGR+M-GICP e o MRG proposto SLERP+LUM-3D é analisado em cada *dataset*. Para a análise do modelo FGR+M-GICP, observam-se as diferenças que as etapas do modelo fazem no RMSE das correspondências e na sobreposição dos pares de nuvens (poses relativas). Para a análise do modelo SLERP+LUM-3D, observa-se a diferença na translação $\varepsilon t_i = \|\mathbf{t}_i^g - \mathbf{t}_i\|$ e na rotação $\varepsilon R_i = \|\mathbf{R}_i^g - \mathbf{R}_i\|$ de cada pose estimada em relação ao *groundtruth*. Onde $\varepsilon R_i \in [0, 2\sqrt{2}]$. O superíndice g refere-se às poses do *groundtruth* do *dataset* em questão, com $i = 1, 2, \dots, n$. O erro na translação é uma simples distância euclidiana. O erro na rotação é a distância de Frobenios ou norma-2 do grupo $SO(3)$, uma métrica adimensional para rotações (HUYNH, 2009). Em ambas, quanto menor, melhor.

5.1 Dataset *Facade* (LST)

A Figura 6 apresenta o resultado da proposta de registro local nos 7 pares do circuito de *Facade*, as nuvens estão coloridas aleatoriamente para facilitar a visualização. Em a) e b) tem-se os gráficos da sobreposição e do RMSE, respectivamente. Em c) e d), estão os resultados visuais da reconstrução 3D. Apresenta-se um recorte da mesma cena.

Figura 6 – FGR vs FGR+M-GICP. a) Sobreposição entre os pares de nuvens. b) RMSE entre os pares de nuvens. c) Reconstrução 3D com FGR. d) Reconstrução 3D com FGR+M-GICP.



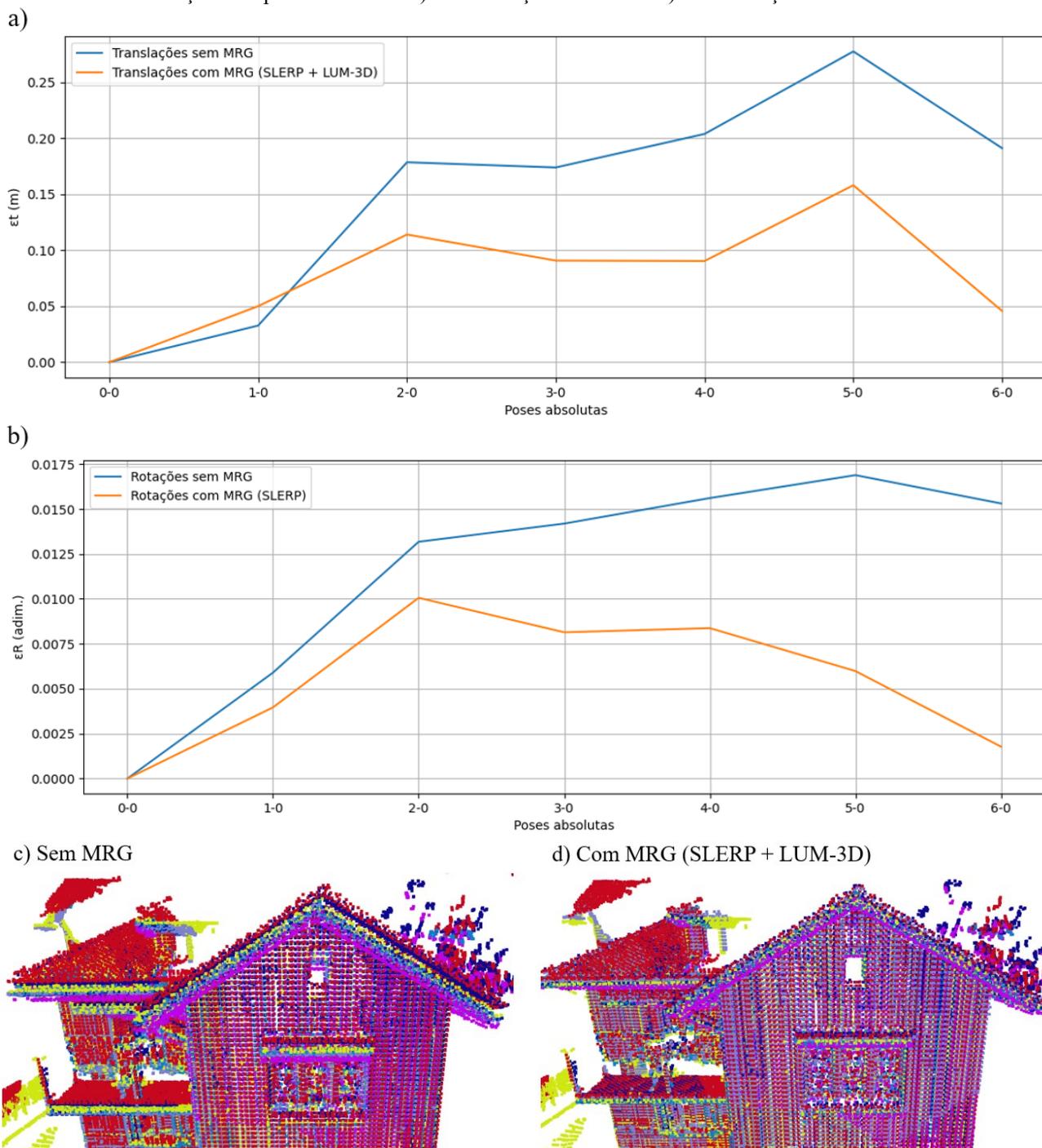
Fonte: Os autores (2022).

Na Figura 6 a), constata-se um aumento na sobreposição entre todos os pares de nuvens após o refinamento com o M-GICP, atingindo um incremento médio de 48%. Na Figura 6 b), observa-se que o RMSE das correspondências diminuiu em média 21% em todos os pares, posicionando-se abaixo dos 6 cm. Este resultado é considerado eficaz, especialmente porque as nuvens foram amostradas utilizando voxels de 10 cm. A Figura 6 c) ilustra a cena da fachada do dataset *Facade*, reconstruída apenas com o registro FGR (etapa coarse), enquanto a Figura 6 d) mostra a mesma cena após aplicação da estratégia coarse-to-fine FGR+M-GICP. O resultado do FGR isoladamente não é satisfatório, mas é adequado para posicionar o M-GICP de maneira aproximadamente correta, levando a uma melhor sobreposição das nuvens em d). Contudo, o resultado na Figura 6 d) ainda apresenta limitações para uso prático, com espaços visíveis entre as nuvens, notadamente no telhado e nas janelas da fachada.

Em relação às sete nuvens do dataset *Facade*, a estratégia de registrar cada nuvem nas três subsequentes culminaram em um total de 21 poses relativas (arestas) para compor o grafo de poses. O registro par-a-par demandou, em média, 25 segundos por par, totalizando aproximadamente 190 segundos para o grafo completo. Considerando a rapidez no processo de conversão das poses relativas em absolutas, este tempo representa a duração total para a reconstrução do dataset *Facade*. Theiler et al. (2015) reportaram um período similar de 200 segundos para este dataset com uma estratégia de registro coarse-to-fine implementada em C++.

A Figura 7 exibe os resultados antes e depois da aplicação do MRG nas poses absolutas do dataset *Facade*. Na Figura 7 a), é exibido o erro na translação de cada pose absoluta; na 7 b), o erro na rotação; na 7 c), a reconstrução sem o MRG; e na 7 d), com o MRG. As subfiguras 7 c) e 7 d), embora idênticas em termos de recorte e cor das nuvens, diferem por um leve zoom no telhado. O resultado na Figura 7 c) é o mesmo da Figura 6 d).

Figura 7 – Aplicação do MRG na reconstrução de *Facade*. a) Erro nas translações das poses absolutas. b) Erro nas rotações das poses absolutas. c) Reconstrução sem MRG. c) Reconstrução com MRG.



Fonte: Os autores (2022).

Nas Figuras 7 a) e b), observa-se que os erros de translação e rotação, antes e após a aplicação do MRG, tendem a se intensificar com a composição das poses, um resultado esperado devido ao acúmulo multiplicativo dos erros. Comparando com as poses do groundtruth, houve uma diminuição média de 46% no erro de translação e de 52% nas rotações. A pose 0, que é a identidade, não sofre alterações, permanecendo fixa para estabelecer a origem na primeira nuvem. É relevante destacar que ajustes nas rotações influenciam as translações, mas o inverso não ocorre. Esta particularidade justifica a escolha de alguns autores, como Theiler et al. (2015), de focar somente no impacto do MRG nas translações. No entanto, devido à nossa proposta original de otimizar as rotações por SLERP, optou-se por analisar ambos os tipos de erro.

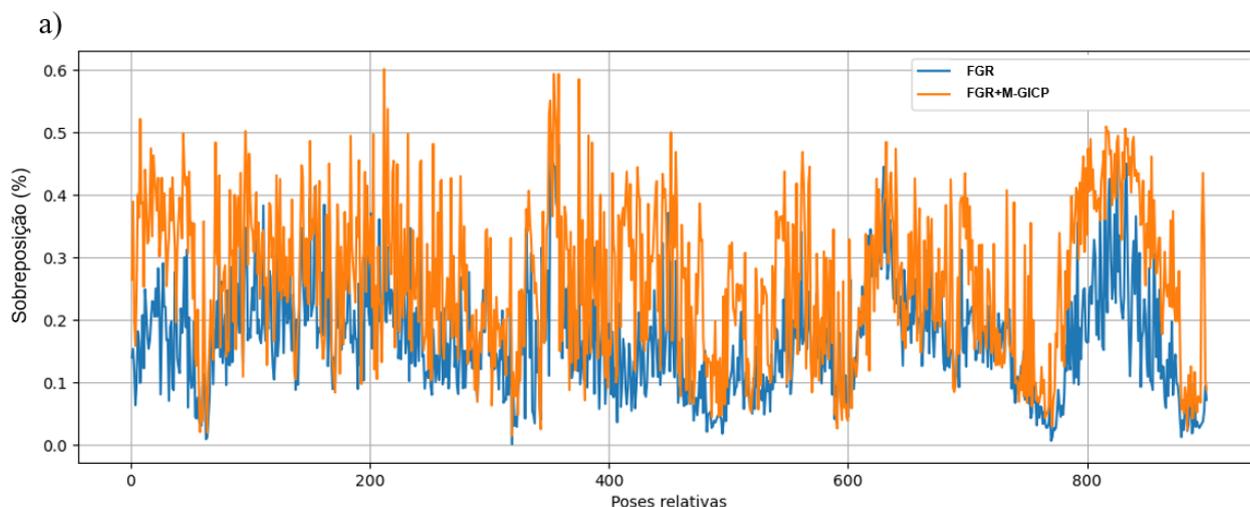
Nem todas as poses se beneficiaram igualmente do MRG proposto. Especificamente, a primeira pose absoluta, 1-0, teve um leve aumento no seu erro de translação. Isso ocorre porque o MRG obtém vantagem do fechamento do circuito de poses, retropropagando o erro da pose final para as anteriores. Assim, nas primeiras poses, onde os erros são originalmente menores, observa-se um ligeiro aumento, o que permite uma redução drástica do erro nas poses finais, mais propensas a acumular imprecisões.

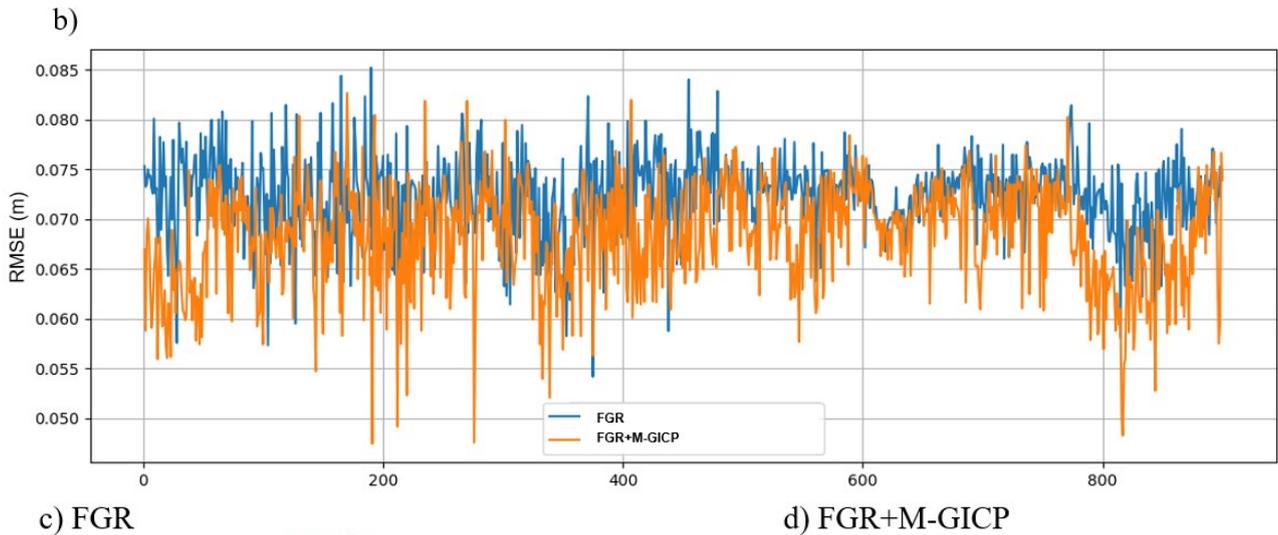
Na Figura 7 c), é apresentada a reconstrução 3D realizada com o método de registro local dos pares. Este resultado é o mesmo exibido anteriormente na Figura 6 c). Mesmo após a operação do GICP, persiste um erro significativo, sugerindo a necessidade de um maior número de iterações. A Figura 7 c) mostra uma melhoria sutil na sobreposição das nuvens após a aplicação do MRG, especialmente visível no vão da janela central da fachada. Embora o principal objetivo do MRG seja corrigir o drift em datasets com muitas poses, a limitação do número de nuvens, neste caso específico, restringe a observação clara de seus efeitos. Com nuvens de LST, encontrar as transformações adequadas entre os pares é o desafio mais crítico, eficientemente superado pela abordagem FGR+M-GICP.

5.2 Dataset NCLT (LSM)

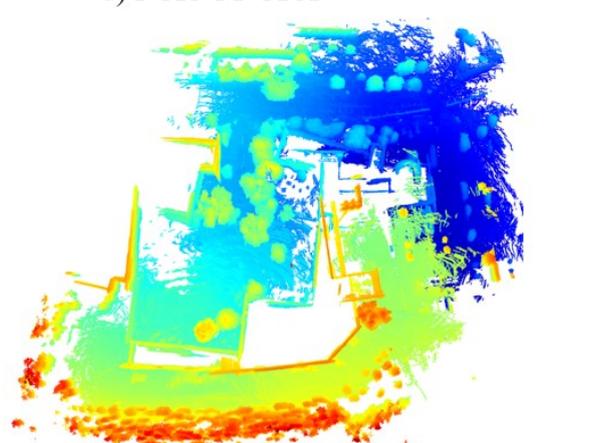
A avaliação do dataset NCLT segue o mesmo procedimento aplicado ao *Facade*. Inicialmente, as poses relativas são analisadas com base na sobreposição, conforme ilustrado na Figura 8 a). Posteriormente, avalia-se o RMSE das correspondências entre os pares, apresentado na Figura 8 b). As reconstruções resultantes das duas fases do registro FGR+M-GICP são exibidas nas Figuras 8 c) e d). Devido à grande quantidade de poses relativas do dataset NCLT, os índices são simplificados: o índice n refere-se ao valor obtido para o par de nuvens n e $n + 1$. Para facilitar uma análise mais abrangente dos resultados, as Figuras 8 c) e d) proporcionam vistas em planta e lateral das reconstruções.

Figura 8 – FGR vs FGR+M-GICP. a) Sobreposição entre pares de nuvens. b) RMSE entre pares de nuvens. c) Reconstrução com FGR. d) Reconstrução com FGR+M-GICP.

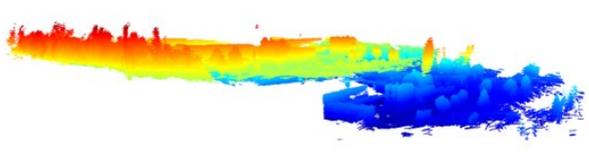
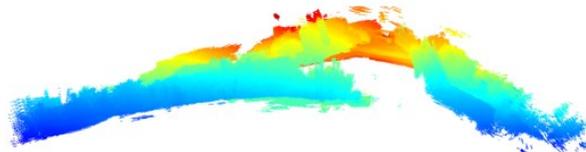




Vista lateral



Vista lateral

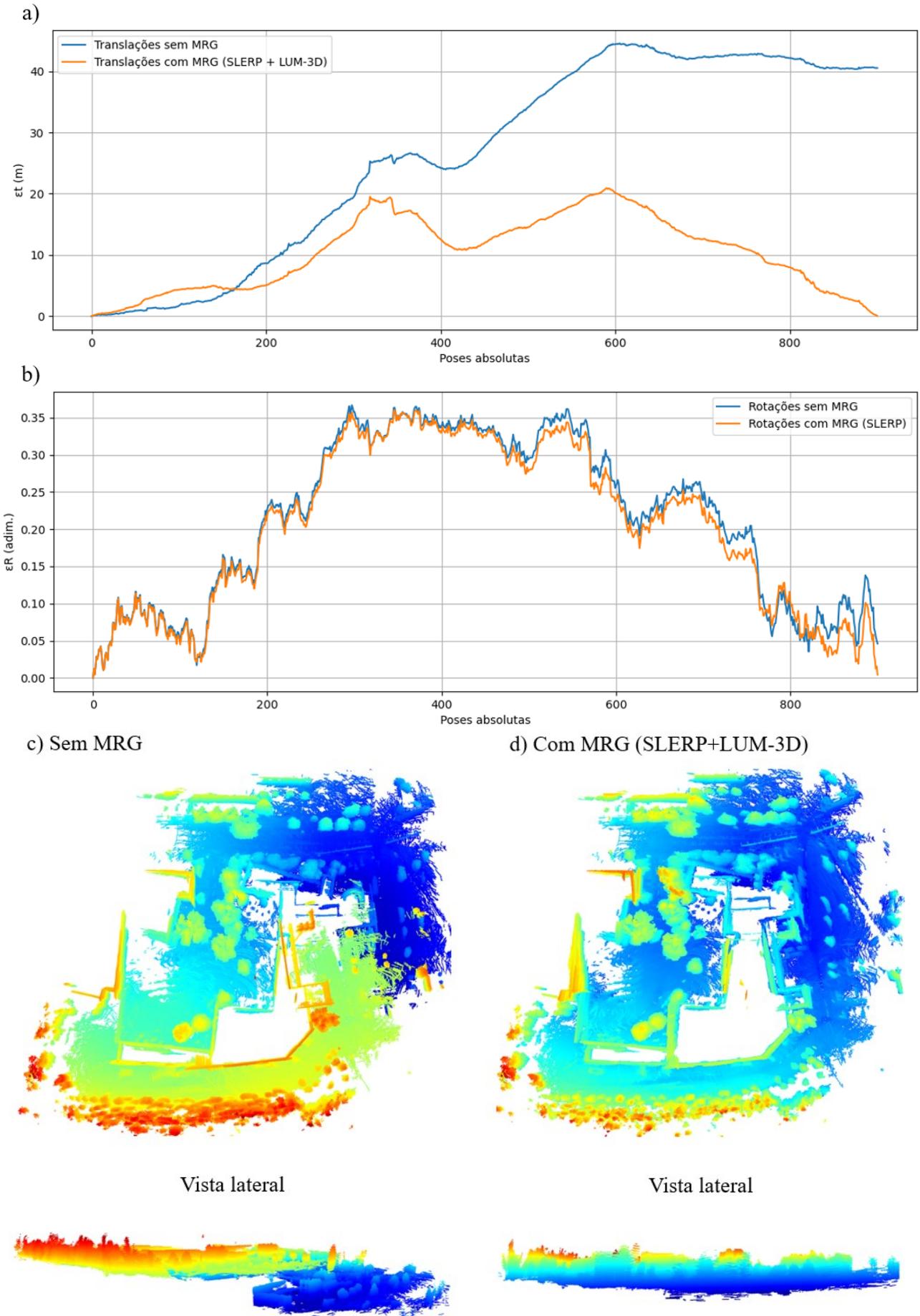


Fonte: Os autores (2022).

Na Figura 8 a), observa-se um aumento significativo na sobreposição entre os pares após a aplicação do M-GICP, com um incremento médio de 36%. Já na Figura 8 b), o RMSE entre as correspondências apresentou uma redução média de 5,5%, caindo de 7,2 cm para 6,8 cm em termos absolutos. A Figura 8 c) revela que essas métricas não refletem plenamente o resultado da reconstrução. É perceptível que o circuito do dataset NCLT, reconstruído pelo FGR, está distorcido, com o acúmulo de erros das 900 nuvens de pontos comprometendo a distinção entre o início e o final do percurso devido a várias distorções. A vista lateral na Figura 8 c) mostra distorções não apenas no plano XY, mas também em planos perpendiculares, evidenciadas por abaulamentos. A Figura 8 d) exhibe o resultado após a aplicação do FGR+M-GICP, destacando a diferença significativa que o refinamento das poses locais trouxe, tornando o circuito agora reconhecível e mais alinhado ao *groundtruth*. No entanto, o início e o final do circuito ainda não se alinham, indicando que o M-GICP por si só não é suficiente para eliminar completamente o drift. No dataset NCLT, foram geradas 2703 arestas no grafo de poses, consumindo um total de 4153 segundos para o FGR e 4690 segundos para o FGR+M-GICP, resultando em um tempo médio de 1,7 segundos por registro de par.

Prosseguindo, a Figura 9 ilustra os resultados após a aplicação do MRG nas poses absolutas, seguindo o mesmo formato de apresentação usado para o dataset *Facade*. Na Figura 9 a), é apresentado o impacto do MRG nas translações das poses absolutas, enquanto a Figura 9 b) mostra o efeito nas rotações. A Figura 9 c) exhibe a reconstrução utilizando o registro FGR+M-GICP, correspondendo ao resultado da Figura 8 c). Finalmente, a Figura 9 d) revela o resultado com a implementação do MRG proposto.

Figura 9 – Aplicação do MRG na reconstrução do NCLT. a) Erro nas translações das poses absolutas. b) Erro nas rotações das poses absolutas. c) Reconstrução sem MRG. d) Reconstrução com o MRG.



Fonte: Os autores (2022).

Na Figura 9 a), inicialmente, identificam-se grandes variações nas translações, enquanto nas rotações, conforme a Figura 9 b), as diferenças são menos acentuadas. Essas variações tendem a aumentar com a composição multiplicativa das poses, mas são neutralizadas no final pelo MRG, que impõe uma condição de fechamento (loop-closure). Isso é evidente no extremo direito dos dois gráficos, na última pose, após a aplicação do MRG. O erro médio de rotação diminuiu relativamente em 4,8%, enquanto nas translações, a melhoria foi de 61%. Em termos de impacto absoluto do MRG, as rotações foram mais beneficiadas: das 900 poses absolutas, 840 (93%) apresentaram reduções no erro de rotação, enquanto 739 poses tiveram melhorias na translação.

Analisando o gráfico do erro de translação na Figura 9 a), observa-se um aumento do erro nas primeiras 161 poses. Em contraste, no gráfico de rotações da Figura 9 b) mostra que, em geral, após a aplicação do MRG, o erro de rotação sempre permanece abaixo do gráfico sem MRG, mas próximo do final (por volta da pose 800), volta a intersectá-lo. Esse fenômeno é resultado do comportamento linear das rotações, que influenciam os ajustes nas translações, mas não o inverso. Isso justifica a escolha de alguns autores, como Vlaminc et al. (2019), de focar apenas no erro de translação.

A Figura 9 c) exibe a reconstrução sem o uso do MRG, enquanto a Figura 9 d) mostra o resultado após sua aplicação. A mudança mais notável é a correção do *drift* pelo fechamento do circuito. Uma métrica comum para medir o drift é a distância entre as extremidades do circuito em relação ao seu comprimento total. No circuito do dataset NCLT, a distância entre as origens era de 41,5 metros antes do MRG, equivalente a um drift de 6,4%. Após a aplicação do MRG, essa distância reduziu-se para apenas 1,1 cm, o que representa um drift de apenas 0,0018%. Embora o circuito final ainda apresente discrepâncias, com algumas poses desviando quase 20 metros do *groundtruth*, o resultado é claramente superior ao obtido sem o MRG.

6 CONCLUSÕES

Este estudo introduziu duas inovações significativas para a área de SLAM. A primeira é uma estratégia de registro de pares de nuvens de pontos 3D, operando em modo coarse-to-fine e combinando os modelos FGR e GICP em uma abordagem multiescala e multicaminho através de um grafo de poses. A segunda inovação é um MRG linear e deduzido de forma fechada, aplicável a qualquer circuito fechado de poses, com a vantagem de não requerer iterações, definição de parâmetros ou inversão de matrizes. Testou-se ambos os modelos em dois datasets distintos de nuvens de pontos 3D, um obtido por LST e outro por MLS em plataforma robótica.

O modelo de registro par-a-par FGR+M-GICP demonstrou ser eficaz na reconstrução de ambos os datasets, exibindo robustez e capacidade de generalização, apesar de não estar completamente isento de erros. Esta eficácia deve-se, em parte, à implementação de um grafo de poses que explora a sobreposição múltipla entre nuvens próximas e ao uso inovador do GICP em multiescala, conforme recomendado por Park et al. (2017). O único parâmetro necessário em todo o processo de reconstrução 3D é o tamanho inicial do voxel de amostragem, escolhido com base em Dong et al. (2018) e Theiler et al. (2015).

O MRG proposto conseguiu aprimorar todas as poses absolutas dos datasets, corrigindo eficientemente o drift resultante da composição das poses relativas. Comparado-se ao *groundtruth*, observou-se uma redução de mais de 40% no erro no dataset *Facade* (LST) e de mais de 60% no dataset NCLT (MLS). Esta melhoria foi corroborada por análises visuais em ambos os datasets, que mostraram melhorias na sobreposição entre os pares de nuvens de pontos 3D.

Para pesquisas futuras, sugere-se a inclusão de pesos nas etapas do MRG, uma implementação facilitada pela dedução de Gauss-Newton, que permite a inserção de uma matriz de pesos entre as jacobianas J do modelo SLERP e H do LUM-3D. Outra extensão possível, mais complexa, seria a integração da etapa SLERP com o framework g2o de Kümmerle et al. (2011), o que implicaria perda da linearidade em favor de possíveis benefícios compensatórios. Também se recomenda explorar a integração com modelos de detecção de loop-closure para automatizar ainda mais a reconstrução 3D. A própria estratégia FGR+M-GICP poderia ser adaptada para este fim, embora mais testes sejam necessários.

Agradecimentos

Os autores agradecem o apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pela concessão da bolsa de estudo, e à Universidade Federal do Paraná (UFPR) por fornecer a infraestrutura necessária à condução da pesquisa em seus laboratórios.

Contribuição dos Autores

Conceptualização: R. A. L. B., D. R. S. e N.L.P.; Curadoria dos dados: R. A. L. B. e D. R. S.; Análise formal: R. A. L. B. e D. R. S.; Aquisição de financiamento: não houve; Investigação: R. A. L. B.; Metodologia: R. A. L. B. e D. R. S.; Administração do projeto: R. A. L. B. e D. R. S.; Recursos: R. A. L. B. e D. R. S.; Software: R. A. L. B.; Supervisão: R. A. L. B. e D. R. S.; Validação: R. A. L. B.; Visualização: R. A. L. B.; Redação – minuta inicial: R. A. L. B. e D. R. S.; Redação – revisão e edição: R. A. L. B. e D. R. S..

Conflitos de Interesse

Os autores declaram ausência de quaisquer conflitos de interesse.

Referências

- AIGER, D.; MITRA, N. J.; COHEN-OR, D. 4-points congruent sets for robust pairwise surface registration. **ACM SIGGRAPH 2008 papers**. p.1–10, 2008.
- BABIN, P.; GIGUERE, P.; POMERLEAU, F. Analysis of robust functions for registration algorithms. 2019 International Conference on Robotics and Automation (ICRA). **Anais...** . p.1451–1457, 2019. IEEE.
- BESL, P. J.; MCKAY, N. D. Method for registration of 3-D shapes. Sensor fusion IV: control paradigms and data structures. **Anais...** . v. 1611, p.586–606, 1992. Spie.
- BLACK, M. J.; RANGARAJAN, A. On the unification of line processes, outlier rejection, and robust statistics with applications in early vision. **International journal of computer vision**, v. 19, n. 1, p. 57–91, 1996. Springer.
- BORRMANN, D.; ELSEBERG, J.; LINGEMANN, K.; NÜCHTER, A.; HERTZBERG, J. Globally consistent 3D mapping with scan matching. **Robotics and Autonomous Systems**, v. 56, n. 2, p. 130–142, 2008. Elsevier.
- BUENO, M.; BOSCHÉ, F.; GONZÁLEZ-JORGE, H.; MARTÍNEZ-SÁNCHEZ, J.; ARIAS, P. 4-Plane congruent sets for automatic registration of as-is 3D point clouds with 3D BIM models. **Automation in Construction**, v. 89, p. 120–134, 2018. Elsevier.
- CADENA, C.; CARLONE, L.; CARRILLO, H.; et al. Past, present, and future of simultaneous localization and mapping: Toward the robust-perception age. **IEEE Transactions on robotics**, v. 32, n. 6, p. 1309–1332, 2016. IEEE.
- CHEN, Y.; MEDIONI, G. G. Object modeling by registration of multiple range images. **Image Vision Comput.**, v. 10, n. 3, p. 145–155, 1992.
- CHOI, S.; ZHOU, Q.-Y.; KOLTUN, V. Robust reconstruction of indoor scenes. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. **Anais...** . p.5556–5565, 2015.
- DONG, Z.; YANG, B.; LIANG, F.; HUANG, R.; SCHERER, S. Hierarchical registration of unordered TLS point clouds based on binary shape context descriptor. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 144, p. 61–79, 2018. Elsevier.
- GAVIN, H. P. The Levenberg-Marquardt algorithm for nonlinear least squares curve-fitting problems. **Department of Civil and Environmental Engineering, Duke University**, Academic Report, 2019.
- GRISSETTI, G.; KUMMERLE, R.; STACHNISS, C.; BURGARD, W. A tutorial on graph-based SLAM.

- IEEE Intelligent Transportation Systems Magazine**, v. 2, n. 4, p. 31–43, 2010. IEEE.
- HARRIS, C.; STEPHENS, M. A combined corner and edge detector. Alvey vision conference. **Anais...** . v. 15, p.10–5244, 1988. Citeseer.
- HUBER, D. F.; HEBERT, M. Fully automatic registration of multiple 3D data sets. **Image and Vision Computing**, v. 21, n. 7, p. 637–650, 2003. Elsevier.
- HUYNH, D. Q. Metrics for 3D rotations: Comparison and analysis. **Journal of Mathematical Imaging and Vision**, v. 35, n. 2, p. 155–164, 2009. Springer.
- KÜMMERLE, R.; GRISETTI, G.; STRASDAT, H.; KONOLIGE, K.; BURGARD, W. g 2 o: A general framework for graph optimization. 2011 IEEE International Conference on Robotics and Automation. **Anais...** . p.3607–3613, 2011. IEEE.
- LOWE, G. Sift-the scale invariant feature transform. **Intelligence Journal**, v. 2, n. 91–110, p. 2, 2004.
- LU, F.; MILIOS, E. Globally consistent range scan alignment for environment mapping. **Autonomous robots**, v. 4, n. 4, p. 333–349, 1997. Springer.
- NEUGEBAUER, P. J. Reconstruction of real-world objects via simultaneous registration and robust combination of multiple range images. **International journal of shape modeling**, v. 3, n. 01n02, p. 71–90, 1997. World Scientific.
- PARK, J.; ZHOU, Q.-Y.; KOLTUN, V. Colored point cloud registration revisited. Proceedings of the IEEE international conference on computer vision. **Anais...** . p.143–152, 2017.
- PAVAN, N. L.; DOS SANTOS, D. R. A global closed-form refinement for consistent TLS data registration. **IEEE Geoscience and Remote Sensing Letters**, v. 14, n. 7, p. 1131–1135, 2017. IEEE.
- RISTER, B.; HOROWITZ, M. A.; RUBIN, D. L. Volumetric image registration from invariant keypoints. **IEEE Transactions on Image Processing**, v. 26, n. 10, p. 4900–4910, 2017. IEEE.
- RUSINKIEWICZ, S.; LEVOY, M. Efficient variants of the ICP algorithm. Proceedings third international conference on 3-D digital imaging and modeling. **Anais...** . p.145–152, 2001. IEEE.
- RUSU, R. B.; BLODOW, N.; BEETZ, M. Fast point feature histograms (FPFH) for 3D registration. 2009 IEEE international conference on robotics and automation. **Anais...** . p.3212–3217, 2009. IEEE.
- RUSU, R. B.; COUSINS, S. 3d is here: Point cloud library (pcl). 2011 IEEE international conference on robotics and automation. **Anais...** . p.1–4, 2011. IEEE.
- SEGAL, A.; HAEHNEL, D.; THRUN, S. Generalized-icp. **Robotics: science and systems**. **Anais...** . v. 2, p.435, 2009. Seattle, WA.
- SHOEMAKE, K. Animating rotation with quaternion curves. Proceedings of the 12th annual conference on Computer graphics and interactive techniques. **Anais...** . p.245–254, 1985.
- SIPIRAN, I.; BUSTOS, B. Harris 3D: a robust extension of the Harris operator for interest point detection on 3D meshes. **The Visual Computer**, v. 27, n. 11, p. 963–976, 2011. Springer.
- THEILER, P. W.; WEGNER, J. D.; SCHINDLER, K. Keypoint-based 4-points congruent sets—automated marker-less registration of laser scans. **ISPRS journal of photogrammetry and remote sensing**, v. 96, p. 149–163, 2014. Elsevier.
- THEILER, P. W.; WEGNER, J. D.; SCHINDLER, K. Globally consistent registration of terrestrial laser scans via graph optimization. **ISPRS journal of photogrammetry and remote sensing**, v. 109, p. 126–138, 2015. Elsevier.
- TORR, P. H. S.; ZISSERMAN, A. MLESAC: A new robust estimator with application to estimating image geometry. **Computer vision and image understanding**, v. 78, n. 1, p. 138–156, 2000. Elsevier.
- TRIGGS, B.; MCLAUHLAN, P. F.; HARTLEY, R. I.; FITZGIBBON, A. W. Bundle adjustment—a modern synthesis. International workshop on vision algorithms. **Anais...** . p.298–372, 1999. Springer.
- VLAMINCK, M.; LUONG, H.; PHILIPS, W. Have I seen this place before? A fast and robust loop detection and correction method for 3D lidar SLAM. **Sensors**, v. 19, n. 1, p. 23, 2019. Multidisciplinary Digital Publishing Institute.

- WEBER, T.; HÄNSCH, R.; HELLWICH, O. Automatic registration of unordered point clouds acquired by Kinect sensors using an overlap heuristic. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 102, p. 96–109, 2015. Elsevier.
- WILLIAMS, J.; BENNAMOUN, M. Multiple view 3D registration using statistical error models. *Vision modeling and visualization*. **Anais...** . v. 99, 1999.
- WILLIAMS, J.; BENNAMOUN, M. Simultaneous registration of multiple corresponding point sets. **Computer Vision and Image Understanding**, v. 81, n. 1, p. 117–142, 2001. Elsevier.
- YANG, Z.; WANG, X.; HOU, J. A 4PCS Coarse Registration Algorithm Based on ISS Feature Points. 2021 40th Chinese Control Conference (CCC). **Anais...** . p.7371–7375, 2021. IEEE.
- ZHONG, Y. Intrinsic shape signatures: A shape descriptor for 3d object recognition. 2009 IEEE 12th International Conference on Computer Vision Workshops, ICCV Workshops. **Anais...** . p.689–696, 2009. IEEE.
- ZHOU, Q.-Y.; PARK, J.; KOLTUN, V. Fast global registration. European conference on computer vision. **Anais...** . p.766–782, 2016. Springer.
- ZHOU, Q.-Y.; PARK, J.; KOLTUN, V. Open3D: A modern library for 3D data processing. **arXiv preprint arXiv:1801.09847**, 2018.

Biografia do autor principal



Rubens Antonio Leite Benevides, nascido em Teresina-PI, 1995. Formado em Eng. Cartográfica e de Agrimensura na Universidade Federal do Piauí (UFPI). Mestre em Ciências Geodésicas pela Universidade Federal do Paraná (UFPR), na área de Fotogrametria e Sensoriamento Remoto. Atualmente é professor de Geoprocessamento no Instituto Federal do Piauí (IFPI) e faz doutorado na UFPR, onde desenvolve pesquisas envolvendo o registro de nuvens de pontos 3D e correção de trajetória/drift no Mapeamento e Localização Simultâneos (SLAM). Tem interesse no processamento geral de nuvens de pontos 3D, classificação, registro, e modelagem matemática por MMQ e RANSAC em Python e MATLAB.



Esta obra está licenciada com uma Licença [Creative Commons Atribuição 4.0 Internacional](https://creativecommons.org/licenses/by/4.0/) – CC BY. Esta licença permite que outros distribuam, remixem, adaptem e criem a partir do seu trabalho, mesmo para fins comerciais, desde que lhe atribuam o devido crédito pela criação original.