

# Rastreamento de um Micro-robô baseado em imagens

## Tese de

Martin Andrade Jastrzebski [8042499]

Iniciada em 20. Jul 2017 e extendida a partir de 01. Set 2018

Fachgebiet Echtzeitsysteme

PCS3860

Elektrotechnik und  
Informationstechnik (FB18)

Engenharia de Computação-PCS

Zweitmitglied Informatik (FB20)

Prof. Dr. rer.nat. A. Schürr  
Merckstraße 25  
64283 Darmstadt

Prof. Ricardo Nakamura

[www.es.tu-darmstadt.de](http://www.es.tu-darmstadt.de)

Escola Politécnica da Universidade de  
São Paulo



---

## **Ehrenwörtliche Erklärung**

Hiermit versichere ich, Martin Andrade Jastrzebski, die vorliegende Master-Thesis ohne Hilfe Dritter und nur mit den angegebenen Quellen und Hilfsmitteln angefertigt zu haben. Alle Stellen, die Quellen entnommen wurden, sind als solche kenntlich gemacht worden. Diese Arbeit hat in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegen.

Mir ist bekannt, dass im Falle eines Plagiats (§38 Abs.2 APB) ein Täuschungsversuch vorliegt, der dazu führt, dass die Arbeit mit 5,0 bewertet und damit ein Prüfungsversuch verbraucht wird.

Abschlussarbeiten dürfen nur einmal wiederholt werden.

Bei der abgegebenen Thesis stimmen die schriftliche und die zur Archivierung eingereichte elektronische Fassung überein.

Darmstadt, den 20. July 2017

---

(Martin Andrade Jastrzebski)



---

## Contents

|                                       |           |
|---------------------------------------|-----------|
| <b>Introdução</b>                     | <b>8</b>  |
| Motivação                             | 8         |
| Desafios                              | 9         |
| Ambiente/ Robô                        | 9         |
| <b>Definições Básicas</b>             | <b>10</b> |
| <b>Trabalhos Relacionados</b>         | <b>12</b> |
| Estruturas Anatômicas                 | 12        |
| Ferramentas Rígidas                   | 14        |
| Cânulas                               | 16        |
| Estudos não-Médicos Relacionados      | 18        |
| <b>Materiais</b>                      | <b>19</b> |
| Radiografias Digitalmente Construídas | 19        |
| Framework de Registro - ITK           | 21        |
| Elastix                               | 21        |
| Outras bibliotecas de Terceiros       | 21        |
| <b>Methods</b>                        | <b>22</b> |
| The Registration                      | 22        |
| The Metric                            | 28        |
| The Optimizer                         | 28        |
| The API                               | 29        |
| Comparison Implementation             | 30        |
| <b>Tests</b>                          | <b>30</b> |
| Tests Automatization                  | 31        |
| <b>Results</b>                        | <b>32</b> |
| Nois                                  | 37        |
| Metrics                               | 38        |
| Number of Projections                 | 39        |
| <b>Conclusion</b>                     | <b>42</b> |
| Summary                               | 42        |
| Future Works                          | 43        |

---

## Image List

|  |    |
|--|----|
| <b>Image 1:</b> The robot  | 10 |
| <b>Image 2:</b> Classical Registration algorithm loop  | 11 |
| <b>Image 3:</b> Example of an echo probe   | 15 |
| <b>Image 4:</b> Example of Humerus intramedullary nail   | 16 |
| <b>Image 5:</b> Example of use of [14]with surgical instruments  | 18 |
| <b>Image 6:</b> Scheme showing the parameters of projection  | 20 |
| <b>Image 7:</b> Scheme showing the parameters of the Robot   | 20 |
| <b>Image 8:</b> The diagram shows the complete inner workings of the registration as used here           | 23 |
| <b>Image 9:</b> Metric variation walking on X  | 24 |
| <b>Image 10:</b> Metric variation walking on Y   | 24 |
| <b>Image 11:</b> Metric variation walking on Z   | 25 |
| <b>Image 12:</b> An example of the metric variation during registration using Amoeba Optimizer           | 26 |
| <b>Image 13:</b> Comparison between image without (1) and with (2) the Sigmoid Gradient Filter           | 27 |
| <b>Image 14:</b> Final Image with Sigmoid Gradient and Shot Filters                                      | 27 |
| <b>Image 15:</b> Metric Variation using Covariance Matrix Adaptation Evolution Strategy Optimizer        | 29 |
| <b>Image 16:</b> Intensity Difference of images of successful registered robot                           | 32 |
| <b>Image 17:</b> Box Plot comparing robot and screw in pixels( 25db of peak signal-to-noise ratio)       | 33 |
| <b>Image 18:</b> Box Plot-Comparing different noise values in pixels                                     | 38 |
| <b>Image 19:</b> Box Plot-Comparing different noise values in millimeters                                | 38 |
| <b>Image 20:</b> Box Plot-Comparing Coefficient Correlation and Mattes Mutual Information in millimeters | 39 |
| <b>Image 21:</b> Box Plot-Comparing number of projections(35db of noise)                                 | 40 |
| <b>Image 22:</b> Box Plot-comparing robot and screw in millimeters( 25db of peak signal-to-noise ratio)  | 41 |

---

## Table List

|  |    |
|--|----|
| <b>Table 1:</b> Anatomical Structures Papers   | 13 |
| <b>Table 2:</b> Rigid Tools Papers   | 14 |
| <b>Table 3:</b> Cannulas   | 17 |
| <b>Table 4:</b> Robot - Covariance Matrix Adaptation Evolution Strategy Optimizer with noise | 34 |
| <b>Table 5:</b> Screw - Covariance Matrix Adaptation Evolution Strategy Optimizer with noise | 35 |
| <b>Table 6:</b> Robot and Screw- Covariance Matrix Adaptation Evolution Strategy Optimizer   | 37 |

---

# 1 Introdução

---

O Procedimento Minimamente Invasivo é um campo de estudo extenso e multidisciplinar que apresenta diversas vantagens em relação aos procedimentos convencionais devido à redução do trauma e risco de infecção uma vez que uma menor incisão é necessária para o procedimento.

Enquanto a demanda de Procedimentos Minimamente Invasivos(MIPs) vem crescendo, quanto mais avançados e complexos esses procedimentos se tornam, mais assistência é necessária, uma vez que manter o controle da ferramenta durante o procedimento torna-se um desafio por si só devido às limitações na linha de visão e o tamanho reduzido das ferramentas.

A estimativa pose de instrumentos cirúrgicos foi feito de muitas maneiras diferentes por rastreadores auxiliares como Electromagnética e rastreio ópticos. Infelizmente, rastreadores auxiliares podem, quando comparado a outros métodos, acrescentar complexidade extra à sala de operação, as vezes até mudando o fluxo de trabalho da cirurgia, já que novos passos de calibração precisam ser feitos.

E mesmo que os dados pré-operatória sejam usados para o planejamento da cirurgia (como tomografia computadorizada e ressonância magnética) e imagens intra-operatórias (como fluoroscopia e ultra-som) estejam frequentemente disponíveis para orientação ao cirurgião durante o procedimento, muitas vezes é feito o uso de projeções 2D que degradam as informações de profundidade. No final, os cirurgiões têm que confiar principalmente em sua experiência para recriar a cena 3D a partir dessas imagens.

---

## 1.1 Motivação

---

A estimativa de pose é, portanto, uma necessidade que vem aumentando em associação a tais procedimentos, proporcionando grande ajuda aos cirurgiões e um passo fundamental para permitir procedimentos automatizados no futuro.

A estimativa consiste em encontrar a ferramenta em um determinado conjunto de imagens e extrair o máximo de informações possível para recriar a pose da ferramenta. O problema de encontrar a pose ainda não está resolvido e tem sido amplamente estudado nos últimos anos com o surgimento dos Procedimentos Minimamente Invasivos.

Embora a comunidade científica tenha apresentado uma ampla gama de diferentes métodos e resultados [12, 19], cada caso vem com seu conjunto associado de benefícios e desvantagens. Assim, comparações e testes individuais podem ser necessários para encontrar as combinações adequadas de métodos para cada nova aplicação e vários pesquisadores [2, 12] dedicaram seu trabalho a essas comparações e categorizações, a fim de entender melhor os limites de cada abordagem.



---

Com isso em mente, este trabalho tenta descrever a implementação de um procedimento de registro (por exemplo, com base no Algoritmo de Informações Mútuas) e uma comparação com abordagens existentes em um sistema de teste projetado para rastrear um micro robô cirúrgico em imagens de raio-x.

---

## 1.2 Desafios

---

Neste estudo serão utilizadas imagens de fluoroscopia. Apesar de amplamente utilizados no campo da medicina e, muitas vezes presentes em salas de operação, seu uso implica nas seguintes dificuldades e limitações que devem ser tomadas na conta durante a implementação do sistema.

Imagens bidimensionais serão usadas para extrair os dados necessários para recriar a posição 3D e os 2 graus de liberdade de 1 articulação (no meio do robô). Alguns dos principais desafios são encontrar as informações e como usá-las para simular a terceira dimensão. Uma possível solução poderia ser a reconstrução 3D, mas exigiria mais tempo de processamento e várias projeções, o que não é desejável, pois os possíveis riscos de radiação associados são uma preocupação no uso de técnicas de orientação fluoroscópica. Portanto um uso mínimo de projeções poderia evitar exposição à radiação para ambos os médicos e pacientes envolvidos.

Mas isso também significa uma disponibilidade limitada a dados e uma maior sensibilidade a erros aleatórios.

As imagens de fluoroscopia são projeções que levam a uma perda de informação de profundidade e degradação da informação geométrica resultante da distorção radial e efeitos de pincushion inerentes ao processo.

Para compensar essa falta e distorção de informações, serão feitas suposições sobre a geometria da ferramenta. Através do uso de um modelo 3D da ferramenta previamente criado, o sistema deve gerar Radiografias Reconstruídas Digitalmente (DRR) e registrá-las com as imagens da fluoroscopia, mensurando por último o erro entre a posição original da ferramenta e o resultado estimado a partir do sistema.

Como esperado, a imprecisão nas salas de operação pode levar a resultados catastróficos. Isso significa que o principal desafio na implementação do sistema é mantê-lo o mais preciso possível sem a necessidade de um longo tempo computacional e, para isso, a criação de um robusto sistema de teste que pode fornecer as informações necessárias para comparar corretamente os métodos de rastreamento é necessário.

---

### 1.3 Ambiente de uso/ Robô

---

O robô estudado é uma ferramenta para a execução de procedimentos minimamente invasivos (MIPs) para a escavação de estruturas anatômicas dentro do osso petroso, como a cóclea (trado auditivo) ou o conduto auditivo interno.



Fig 1: O robô

Como pode ser observado na Figura 1 o robô apresenta formato cilíndrico e 3 partes principais. A cabeça que pode ou não apresentar uma broca, uma parte central articular que é responsável por dar uma melhor mobilidade ao robô e uma cauda.

---

## 2 Definições Básicas

---

**Radiografia Digitalmente Reconstruída (DRR):** Geração de imagens de raios X sintéticas, através do uso da técnica de Ray-tracing em modelos 3D usando uma geometria de câmera conhecida (ou Volume). Os valores de pixel DRR são simplesmente os somatórios dos valores de CT encontrados ao longo de cada raio de projeção [1].

---

**Processo de Registro (ou Registro de Imagem)** : Processo de encontrar um conjunto de parâmetros que melhor compara um modelo (imagem em movimento) a uma imagem fixa. É comumente feito através de um algoritmo iterativo que tenta otimizar (maximizar ou minimizar) uma determinada métrica de similaridade. Um procedimento de registro sempre pode ser decomposto em três partes principais: a declaração do problema, o paradigma de registro e o procedimento de otimização [2].

**Metric:** The similarity measure used by the Registration process to compare the images. Through the literature several different metrics have been proposed and can be divided as Intensity and Gradient Based.

**Métrica:** A medida de similaridade usada pelo processo de registro para comparar as imagens. Através da literatura, várias métricas diferentes foram propostas e podem ser divididas de acordo com sua base de extração dos dados através da Intensidade ou do Gradiente.

As **métricas baseadas em intensidade** correlacionam a intensidade global das imagens de maneira semelhante aos pixels. Alguns exemplos são Informações Mútuas, Correlação cruzada normalizada e soma de diferenças quadráticas.

**Métricas baseadas em gradientes** por sua vez comparam a intensidade local. Isso significa que duas imagens são consideradas semelhantes se ocorrerem mudanças de intensidade em locais semelhantes. Entre esse tipo de métrica, podemos citar a Correlação de Gradiente (Gradient Correlation) e Informações de Gradiente(Gradient Information) como comumente usadas devido à sua robustez e precisão [11].

Embora vários trabalhos como [16] tenham tentado encontrar métricas “melhores”, todas foram amplamente utilizadas e a opinião diverge sobre qual delas é melhor ou pior.

**Otimizador:** Componente do algoritmo de registro responsável por encontrar o conjunto de parâmetros que otimizam a função de similaridade através da busca de pontos de mínimo ou máximo.

**Target Registration Error (TRE)** : Definido como a distância euclidiana entre a posição registrada calculada e a **posição real**.

**Erro de Distância de Projeção (PDE)** : Distância entre os pontos de referência estimados e verdadeiros **na imagem projetada**.

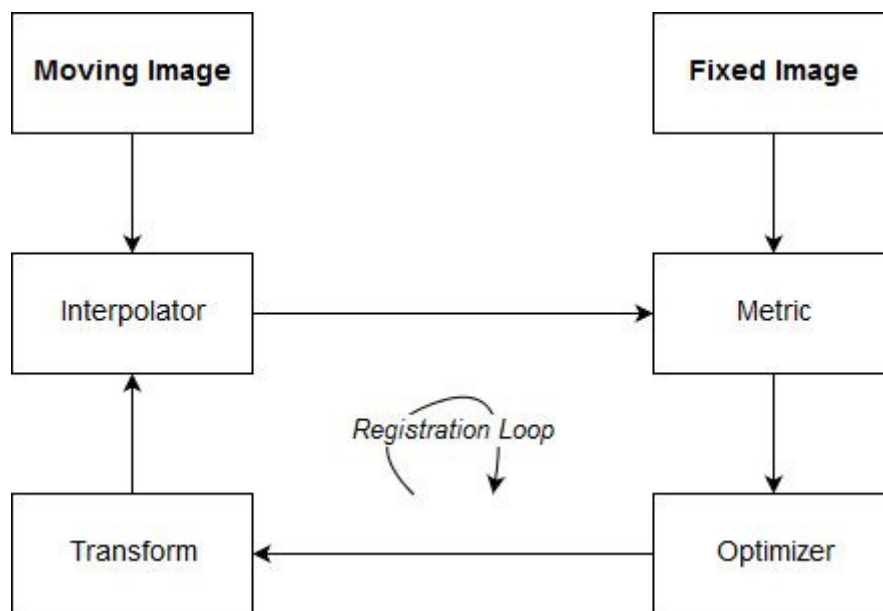


Figura 2: Loop do algoritmo de registro clássico

### 3 Trabalhos relacionados

Como mencionado anteriormente, vários trabalhos foram publicados sobre a extração da pose 3D de um modelo fazendo uso de uma ou mais projeções. Seus métodos, resultados e aplicações diferem uns dos outros, mas uma breve explicação sobre como alguns deles podem ser comparados com a contribuição desta tese poderia ajudar a entender tanto a contribuição quanto o ambiente atual do estado da arte do registro.

Este projeto implementará um sistema para estimar a pose de um micro robô usando um modelo 3D e imagens de raios-X.

A estimativa da pose em 3D a partir de projeções é um problema que vem sendo estudado há muitos anos na área de imagens médicas [3,6-11,14] e possui diversas aplicações conforme explicado anteriormente.

Abordagens semelhantes às apresentadas aqui já foram utilizadas, principalmente na detecção de ferramentas rígidas em toda a literatura [7,8,9], mas as particularidades no comportamento do robô a ser rastreado aqui se faz necessário um tratamento especial, uma vez que a flexão e extensão da articulação podem estar presentes devido ao centro (parte do meio) do robô.

Com isso em mente, para uma melhor compreensão e definição do problema, é necessário entender e levar em conta pesquisas relacionadas a registros não rígidos como [3,6].

Além disso, a comparação de métricas a serem utilizadas apresenta grande peso no desenvolvimento do sistema. Enquanto alguns grupos [16] compararam métricas existentes para entender quais são superiores em quais situações, outros grupos como [13,15] tentam aprimorar ou criar novas métricas para estender a capacidade dos métodos de registro atuais.

Para entender melhor as diferentes áreas estudadas e relacionadas ao problema, as próximas seções irão apresentar uma explicação sobre os trabalhos mencionados acima e contribuições na literatura e como seus métodos, resultados ou aplicações podem ser relacionados a esta tese.

Como forma de organização da apresentação desses trabalhos, haverá uma divisão quanto ao objetivo do estudo.

---

### 3.1 Estruturas Anatômicas

---

A detecção e alinhamento de estruturas anatômicas em imagens médicas tem recebido foco dos pesquisadores com o crescimento dos MIP e se mostraram importantes tanto para o planejamento da operação quanto para auxiliar na visualização durante o procedimento.

Estudos sobre a estimagem da posição de ossos, órgãos e vasos de imagens médicas e não médicas abrangem várias metodologias de detecção uma vez que deformações e tecidos moles podem estar presentes e atrapalhar o processo, razão pela qual essa área pode ser vista como mais desafiadora do ponto de vista técnico.

Uma forma recorrente de registro 2D-3D é a comparação entre tomografias e raios-x. Infelizmente, o uso de dados 3D geralmente pode gerar custos extras para diagnósticos e cirurgias. Como consequência, o uso de tais informações em situações de rotina geralmente não está disponível [17].

| Número | Anatomia        | Método  | Resultado   |
|--------|-----------------|---|---|
| [17]   | Ossos Genéricos | 2D-to-3D Generic Bone Model to Fluoroscopy using Distance as metric | taxa de sucesso de registro é de 93,9% e erro médio de 3 pixels entre os pontos de contorno do modelo e os pontos de contorno da imagem |
| [10]   | Vertebras       | 2D-to-3D Gradient Information                                       | Sucesso em mais de 99,99% dos testes e exigiu menos de 5 segundos para computar   |
| [11]   | Vertebras       | 2D-to-3D Gradient Information                                       | A PDE < 5mm foi   |

|      |         |                         |  |
|------|---------|-------------------------|--|
|      |         |                         | alcançada em > 99,99% dos ensaios                    |
| [13] | Crânios | “Nova” Informação Mútua | mTRE 23,92% menor do que a informação mútua clássica |

Em [17] foi proposto um método para recuperar poses 3D de imagens de raios X e com o uso de modelos ósseos genéricos. O registro das imagens utilizadas por eles é baseado principalmente na comparação entre os contornos da projeção do modelo ósseo e o raio X. Pontos correspondentes são aproximados a partir de um “chute” inicial e refinado através das iterações do sistema. O uso da distância como métrica foi justificado com a tentativa de evitar a necessidade de uma boa inicialização. Parte do desafio aqui foi o registro de duas imagens que não provinham do mesmo espécime (as DRRs eram provenientes de um osso genérico) e as escolhas de design feitas para evitar a necessidade de grandes algoritmos de segmentação.

[10,11] propuseram métodos para identificar vértebras de pacientes usando uma tomografia computadorizada pré-operatória que recebe a enumeração da vértebra manualmente. Depois disso, as DRRs das vértebras seriam criadas e comparadas às imagens de raios X para rotular os números das vértebras e tentar reduzir os casos de cirurgias no local errado (comum em cirurgias em vértebras e costelas).

O registro de imagem de dados que contém padrões repetitivos, como vértebras ou nervuras, pode representar um desafio para as métricas em pixels, pois vários máximos e mínimos locais podem fazer com que o otimizador estagne em extremos locais, em vez de globais. Diversos trabalhos têm utilizado métricas baseadas em Gradiente nos casos em que há risco de falsa convergência [10,11].

Devido a esses problemas [13] realizou-se um estudo sobre a implementação de dados espaciais na métrica da informação mútua através da combinação de intensidades e suas coordenadas espaciais. Os testes com crânios apresentaram melhores resultados quando comparados com a métrica clássica de informação mútua regular com média de mTRE e iteração inferior a 23,92% e 52,06% que a de informação mútua respectivamente, com 14,8 vezes maior taxa de sucesso.

Esses tipos de esforços, mesmo que inicialmente destinados a diferentes aplicações, mostram as limitações e possíveis melhorias que as soluções atuais podem sofrer quando submetidas aos novos desafios de diferentes ambientes.

---

## 3.2 Ferramentas Rígidas

---

Embora o registro não-rígido (que poderia incluir deformações) tenha sido amplamente utilizado para melhorar a detecção de órgãos, vasos e ferramentas semelhantes a cânulas, o

registro rígido (somente translação e rotação) foi fortemente estudado para detectar ossos e algumas ferramentas, apresentando-se normalmente como um problema menos complexo devido à falta de parâmetros de deformação que também teriam que ser resolvidos.

Entre os vários métodos propostos para encontrar a pose 3D de ferramentas, é possível destacar alguns deles, já que suas abordagens utilizam os dados extraídos de imagens em um fluxo de trabalho similar, como apresentado aqui:

| Número | Ferramenta                                 | Método   | Resultado   |
|--------|--|--|---|
| [7]    | Sonda transesofágica 3D                    | Diferença de gradiente e intensidade de padrão (Gradient Difference and Pattern Intensity) | Erros de registro médio de 1.5–4.2 mm, intervalo de captura de 95% de 8.7–11.4 mm em termos de TRE. |
| [8]    | Pinos Intramedulares                       | Recriar a pose 3D através da localização de buracos presentes no pino                      | Tradução <0.5mm Pitch and Roll <2 ° e yaw <2 °  |
| [9]    | Parafuso de pedículo de cirurgia de coluna | Correlação de gradiente (Gradient correlation)   | TRE 1–4 mm e < 5 ° usando modelos paramétricos simples.   |
| [5]    | Instrumentos Laparoscópicos                | Landmarks  | precisão de posição de 0,438 ± 0,231 mm.  |

[7], por exemplo, fizeram uso de um fluxo de trabalho muito semelhante para rastrear a cabeça de uma sonda de eco transesofágica 3D. Um modelo 3D da cabeça foi criado e um DRR foi gerado para ser comparado às imagens de raios-x. Seu objetivo principal era poder fazer uso dos dados provenientes da sonda e da radiografia juntos. A posição da cabeça da sonda precisava ser encontrada para executar o procedimento de calibração necessário. Durante a etapa de registro, foi feito o uso de diferentes métricas (Gradient Difference e Pattern Intensity) e não foram consideradas transformações não rígidas.

O uso de Gradient Difference e Pattern Intensity foi baseado na vantagem de que eles tendem a ser menos afetados pela presença de tecido mole e, portanto, mais precisos neste caso e seus resultados de registro mostraram erros de registro médios de 1.5–4.2 mm e 95% alcance de captura de 8,7 a 11,4 mm em termos de TRE.



Figura 3: Exemplo do tipo de sonda usada por [7]

[source : <http://www.axisimagingnews.com/2015/07/axess-ultrasound-expands-probe-repair-capabilities/>]

Em [8] fez-se o rastreamento da posição de pinos intramedulares com o menor número possível de imagens de raios-x tiradas para reduzir a quantidade de exposição à radiação. Imagens de raios-X forneceriam projeção do pino e, usando um modelo matemático da projeção dos furos presentes no mesmo, a posição do pino poderia ser reconstruída a partir de duas projeções. Esta solução usou as particularidades geométricas dos parafusos para resolver o problema da pose 3D sem a necessidade da geração de DRRs.

Embora, neste caso, a extração de características não tenha sido usada em um algoritmo de registro, é comum reduzir a quantidade de dados a serem comparados por meio da detecção e da correspondência de pontos de referência.



Figura 4: Exemplo de pino intramedular do úmero

[source: <http://www.medicaexpo.com/prod/aesculap/product-70641-764345.html>]

[9] Registrou Imagens 3D pré operatórias e imagens 2D intraoperatórias da anatomia do paciente para encontrar os parâmetros de pose. Em seguida, utilizaram a conhecida geometria dos



---

parafusos pediculares da cirurgia da coluna para gerar diversos modelos 3D que simplificam seus detalhes e os registram nas imagens intraoperatórias. Aqui, a correlação de Gradiente foi usada devido à vantagem de filtrar as diferenças de baixa frequência espacial, concentrando-se nos limites dos componentes.

Este trabalho levanta uma interessante discussão sobre a precisão do modelo 3D dos parafusos (o corpo a ser registrado) e como isso afeta a precisão do posicionamento do parafuso. Seus resultados apontam, conforme o esperado, que quanto mais preciso for o modelo, melhor será o desempenho de seu sistema.

Como uma forma alternativa de rastreamento, os pontos de referência que poderiam gerar padrões únicos de cada ponto de vista também foram usados. Um bom exemplo de uso desses marcos foi no estudo sobre o sistema de rastreamento de instrumentos laparoscópicos feitos em [5,4] onde as propriedades da colocação das referências na ferramenta foram utilizadas para resolver o problema da pose como um sistema matemático que usa sua posição como parâmetros.

---

### 3.3 Cânulas

---

Ainda dentro da área de rastreamento de ferramentas cirúrgicas, as ferramentas semelhantes a cânulas precisam de uma abordagem especial devido ao seu comportamento único quando dentro do espécime observado.

Sensores cinéticos estão comumente presentes através da maioria das ferramentas de cânulas como cateteres, mas como cada junta tem uma certa imprecisão, a previsão final da ponta da ferramenta é passiva para sofrer um acúmulo de erros, resultando em uma grande disparidade entre o valor final estimado da ferramenta e sua posição real. Isso, juntamente com a atuação ruidosa resultante da miniaturização, e as incertezas resultantes da interação dispositivo-tecido criam um grande desafio no rastreamento dessas ferramentas.

Devido a isso, uma abordagem baseada em rastreamento de imagens pode ser vista como uma solução razoável.

| Número | Método   | Resultado   |
|--------|--|---|
| [3]    | Rastreamento de fio-guia que combina recursos semelhantes a segmentos (GSEG) (Explicação detalhada disponível em [18]) | Erro médio de reconstrução da forma de 1,22 mm e erro de ponta de 4,23 mm                   |
| [6]    | Registro de componente conhecido deformável (dKC-Reg)  | Erro de translação $1,51 \pm 0,9$ mm e erro de angulação da trajetória de $0,58 \pm 0,22$ ° |

---

Como forma de superar a incerteza dos sensores cinéticos, a equipe responsável por [3] criou um sistema para integrar informações desses sensores e imagens de raios X para recriar um modelo 3D.

Os métodos que dependem da reconstrução do modelo precisam de mais de uma imagem obtida de diferentes pontos de vista. Essa metodologia, quando aplicada a esse tipo de ferramenta, baseia-se na busca de pontos-chave que são usados como base para a produção do modelo. Uma posição inicial pode ser tomada a partir das informações provenientes dos sensores e ocorre um processo iterativo de refinamento, fazendo uso do registro de imagens para calcular a similaridade entre as imagens.

[6] fizeram uso de uma descrição paramétrica das ferramentas da cânula com curvas B-spline para reconstruir a pose delas em imagens obtidas durante o procedimento. O trabalho deles recriou o modelo à medida que os pontos-chave da curva eram alterados para se adequar melhor à imagem e a Correlação de Gradiente usada como métrica para o registro. O uso de B-splines para descrever as possíveis formas que a ferramenta pode assumir é um exemplo da modelagem matemática do problema para reduzir a liberdade dos valores dos parâmetros, ajudando a garantir a convergência e um menor tempo de processamento.

O micro-robô aqui rastreado deve ter seus parâmetros de posição e ângulos extraídos de imagens de raios-x. Entre as características deste robô, podemos destacar o fato de que a articulação que conecta as duas partes do robô cria dois novos graus de liberdade, implicando um aumento no número de parâmetros a serem resolvidos. As ferramentas tipo cânula têm uma grande liberdade de deformação, o que força sua pesquisa a focar tanto na resolução desses parâmetros quanto na criação de modelos matemáticos que restrinjam esses valores e facilitem a análise das informações disponíveis.

---

### **3.4 Estudos não médicos relacionados**

---

O registro de imagem tem raízes na visão computacional. Ambas as áreas crescem relativamente juntas, uma vez que os resultados obtidos em pesquisas em uma área podem beneficiar pesquisas em outra.

Como o robô a ser detectado apresenta áreas rígidas cuja geometria pode ser aproximada a cilindros, estudos sobre a solução destes podem apresentar resultados úteis como modelos matemáticos para acelerar a convergência de valores no registro ou simplificar a parametrização das peças.

O estudo em [14] apresenta um algoritmo de dois estágios para a solução de pose a partir de uma única visão, dado seu contorno aparente. Seu método inclui o uso do contorno do cilindro como pontos para resolver um sistema que determina a posição do centro e o raio do cilindro. Este sistema é um modelo geométrico com uma matriz singular composta com as coordenadas de

---

Plücker do eixo de simetria e o raio do cilindro

Esse resultado tem aplicação direta na cirurgia assistida por computador, pois pode ser usado como base para encontrar e rastrear todo tipo de ferramentas com peças cilíndricas.

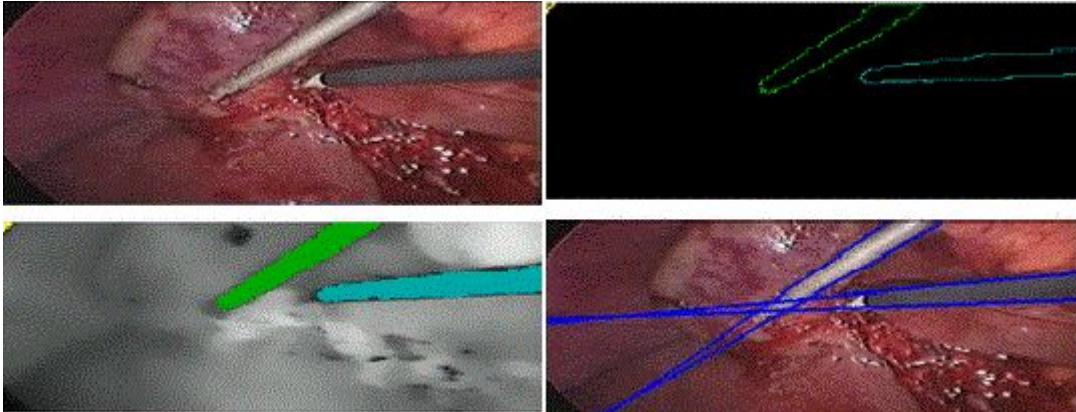


Figura 5:Exemplo de uso de [14] com instrumentos cirúrgicos

Vários estudos também vêm aprimorando as métricas existentes tentando superar suas limitações.

De maneira similar a [13], [15] define uma nova forma de informação mútua que combina informação mútua com uma função de ponderação baseada na diferença absoluta de valores de pixel correspondentes. Eles pretendiam incorporar na informação mútua informação espacial (a função de peso acima mencionada) considerando os vizinhos de cada pixel. Os testes feitos pela equipe mostraram que ela é mais robusta ao ruído e mais precisa do que os métodos padrão usados.

---

## 4 Materiais

---

O sistema a ser implementado nesta tese é baseado no trabalho de vários autores e tem como ponto de partida alguns resultados previamente prontos para avançar.

Como afirmado anteriormente, o sistema de reconstrução de raios-x digital, que desempenha um papel fundamental no registro de imagens 2D-3D, foi fornecido como o ponto de partida do projeto.

Utilizando este sistema de DRR como base, utilizou-se o Insight Segmentation and Registration Toolkit como principal estrutura de registro das imagens, fornecendo o loop interno do registro.

Além destes dois componentes principais, é possível enfatizar o uso do kit de ferramentas VTK, RTK e CUDA como ferramentas usadas para otimizar e depurar tanto o código quanto a

operação lógica do sistema.

## 4.1 Digitally Reconstructed Radiography

O gerador de imagens provido para executar este projeto pode ser dividido em um leitor de arquivos JSON, um módulo para manipulação de malhas e volumes e um projetor.

O leitor de arquivos JSON é responsável por ler arquivos e transferir as informações para os objetos corretos no código. Também foi de fundamental importância para a geração de testes, uma vez que os arquivos a serem lidos poderiam apresentar distribuições estatísticas ou combinatórias nos parâmetros a serem passados.

O segundo módulo do gerador de imagens recebe malhas (conjunto de pontos e índices) e gera volumes (imagens 3d com informação volumétrica). Também disponibiliza um conjunto de ferramentas para manipular essas malhas, transformação no espaço e geração a partir dos parâmetros dados.

Finalmente, o projetor é responsável por gerar imagens assim que o volume é projetado. Para isso, ele aceita um vetor que contém as propriedades de cada projeção necessária.

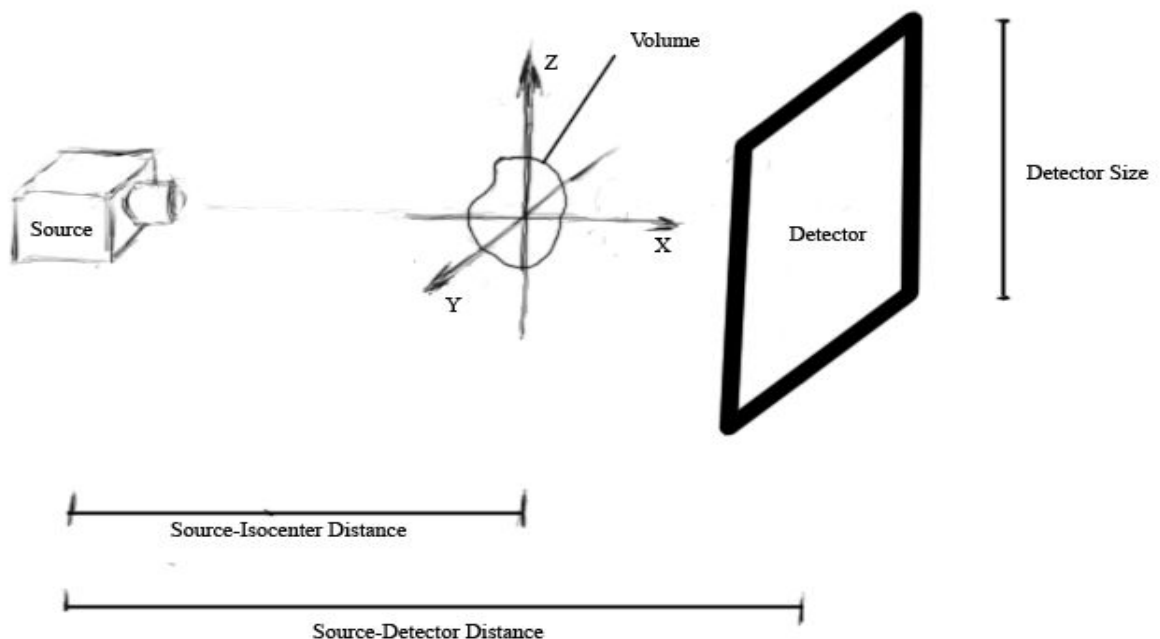


Figura 6: Esquema mostrando os parâmetros de projeção

A Projeção foi definida pelo conjunto de parâmetros mostrados acima somado a um vetor 3D de rotação ao redor do centro e os valores de deslocamento necessários que ajudariam a definir a posição do centro relacionado ao volume.

A malha do robô foi gerada em cada etapa, o que exigiria uma mudança nos parâmetros que definem sua forma. A geração seria feita por meio de um programa de terceiros, o que significa que o sistema teria que ler e carregar a nova malha do disco, levando a um aumento na carga de trabalho e, portanto, a um desempenho pior.

---

Devido a isso a maioria dos testes foram feitos utilizando malhas fixas que poderiam ser mantidas na memória durante todo o processo e acelerar a geração de imagens.

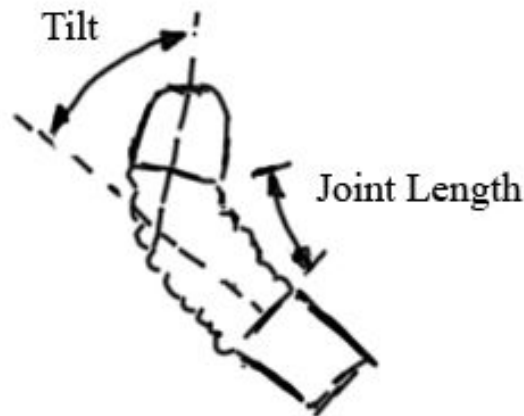


Figura 7: Esquema mostrando os parâmetros do robô

---

## 4.2 Framework de registro - The ITK

---

O ITK (Insight Segmentation and Registration Toolkit) é um sistema de plataforma cruzada de código aberto que fornece ao usuário uma ampla variedade de ferramentas para trabalhar com imagens.

O ITK tem como principal vantagem a flexibilidade, dada pelo design extremamente modular e o emprego de "Templates" na maioria das classes, deixando a escolha de como e onde usar a ferramenta nas mãos do desenvolvedor.

A possibilidade de trabalhar com uma quantidade indefinida de dimensões fornecida por este projeto tornou possível executar o registro do projeto sem grandes mudanças na implementação durante a fase de testes, onde os parâmetros a serem encontrados seriam frequentemente alterados.

O ITK começou sua história em 1999, quando a Biblioteca Nacional de Medicina dos EUA do National Institutes of Health recebeu um contrato de 3 anos para desenvolver um kit de ferramentas para registro e segmentação de imagens que mais tarde viriam a ser chamadas de ITK.

O ITK foi utilizado neste projeto como base de registro de imagem. O loop de registro básico do mesmo foi usado com métricas e interfaces geradas aqui para executar o registro.

---

## 4.3 Elastix

---

O Elastix é um conjunto de ferramentas para registro de imagens rígidas e não rígidas. Ele é baseado no Insight Segmentation and Registration Toolkit e tem uma compatibilidade direta com ele, o que facilita o acréscimo ao nosso projeto e a rápida configuração e teste.

Esse kit de ferramentas forneceu o otimizador da Covariance Matrix Adaptation Evolution

---

Strategy usado para alguns de nossos testes, mas o mesmo apresenta muito mais e está disponível para ser usado com C ++, Python, Java, R, Ruby, C # e Lua. O artigo completo sobre sua criação e uso está disponível em [20,21].

---

#### 4.4 Outras bibliotecas de terceiros

---

Além das ferramentas mencionadas anteriormente, algumas outras desempenharam um papel menor neste trabalho enquanto ajudavam com partes menos críticas do sistema.

O Visualization Toolkit (VTK) é um software de código aberto que foi usado para depurar o sistema, fornecendo visualização em tempo real das imagens à medida que eram criadas.

O Kit de Ferramentas de Reconstrução (RTK) é um software de código aberto e de plataforma cruzada baseado no ITK usado pelo DRR para implementar a aceleração CUDA da projeção de volume.

O kit de ferramentas CUDA é o kit de ferramentas fornecido pela Nvidia para o desenvolvimento de softwares capazes de usar a GPU. Foi utilizado neste projeto em uma otimização da DRR para melhoria de desempenho para o trabalho com grandes volumes.

---

### 5 Métodos

---

Este capítulo apresentará o sistema implementado nesta tese, bem como explicará as opções de design e entrará em detalhes de sua operação. O sistema desenvolvido pode ser dividido em três partes:

- A interface que comunica o sistema de registro ao sistema de geração de imagens.
- A métrica que implementa internamente quatro tipos diferentes de métricas (Informações Mútuas, Informações Mútuas Mútuas, Diferenças de Gradiente e Correlação de Gradiente) e adapta o registro do ITK que conecta o número de dimensões da imagem a ser avaliada ao número de parâmetros a serem variados.
- O sistema de teste automatizado.

---

#### 5.1 O Registro

---

O registro de imagem, como explicado anteriormente, tenta maximizar a métrica de similaridade de imagem variando os parâmetros fornecidos.

Isso é feito dentro de um loop que consiste em um otimizador, uma métrica e o gerador de imagens. O Otimizador passa para a métrica os parâmetros do robô, a métrica por meio da API sua vez passa para o gerador de imagem que finalmente devolve as imagens para serem comparadas pela métrica. Um diagrama operacional completo é mostrado na próxima imagem.

O ITK tem em seu loop base uma limitação entre o número de parâmetros a serem otimizados e a dimensionalidade das imagens a serem comparadas. Isso criou a necessidade de modificar a sua implementação para se adaptar a nossa realidade (o loop clássico de registro

---

como mencionado no primeiro capítulo), fazendo com que o loop base otimize a quantidade de parâmetros necessários para o projeto, enquanto a métrica modificada continha métricas secundárias que comparam apenas imagens bidimensionais.

Tentativas iniciais produziram resultados não muito robustos. Após testes e análise, concluiu-se que muito disso se devia à quantidade de ruído que as imagens como um todo criavam sob o valor da métrica quando comparados com a variação proveniente da mudança de posição e do formato do robô.

Como forma de mitigar esse efeito foi criado um sistema de corte de imagem que focaliza o registro dentro de uma determinada área ao redor do robô na imagem fixa. Isso pode ser alcançado em um cenário real, sem grandes problemas com um algoritmo de segmentação inicial ou entrada do usuário, por tanto não invalidando os resultados aqui apresentados.

Como resultado, nossa ferramenta começou a apresentar resultados mais consistentes e trabalhou de acordo com o esperado.

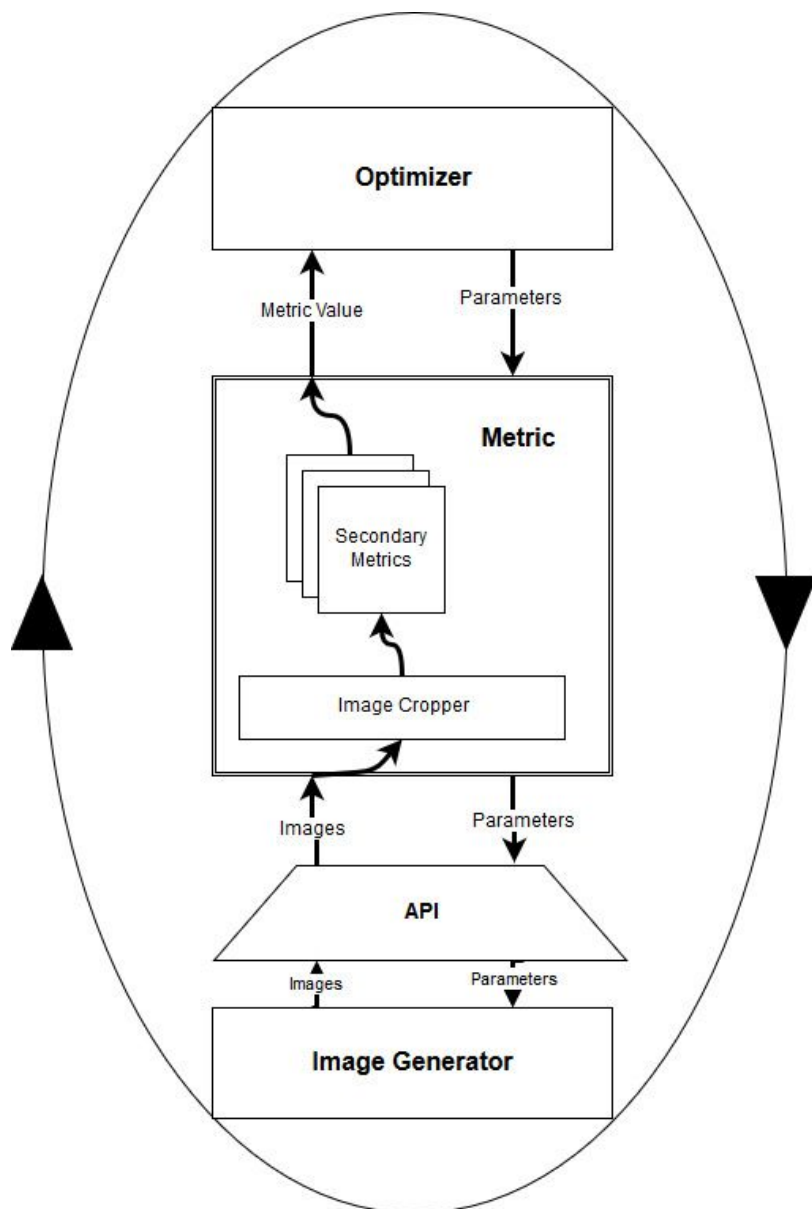


Figura 8: O diagrama mostra o funcionamento interno completo do registro como usado aqui.

Uma vez que resultados consistentes foram obtidos, um estudo do comportamento da variação métrica em função da dimensão variada foi realizado para melhor compreender que tipo de limitações e dificuldades são encontradas pelo otimizador durante um processo de registro de imagem usando o método descrito acima.

O principal problema e limitação encontrada veio do fato de que cada parâmetro tem seus próprios valores máximos em diferentes locais.

Para visualizar melhor que os gráficos seguintes mostram a variação da informação mútua durante uma busca exaustiva (todos os valores em uma grade da busca em cada eixo (e um somente)).



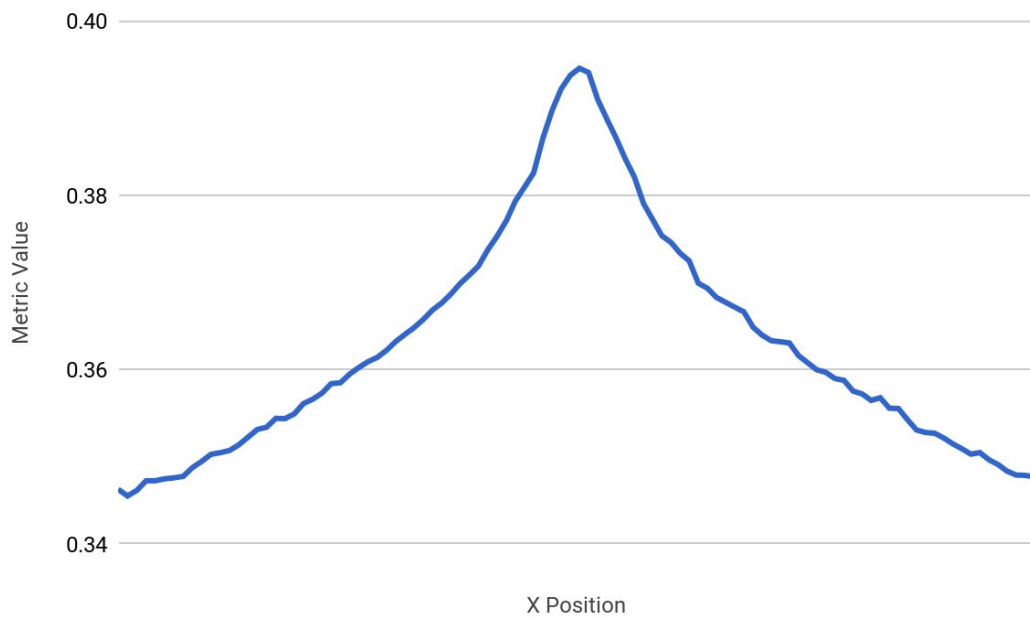


Figura 9: O gráfico mostra a variação do valor da métrica à medida que o robô caminha ao longo do eixo X e passa seu valor real com um passo de 0,05 milímetros por amostra.



Figura 10: O gráfico mostra a variação do valor da métrica à medida que o robô caminha ao longo do eixo Y e passa seu valor real com um passo de 0,05 milímetros por amostra.

Acima pode-se observar que o ponto de valor máximo não se apresentou no lugar certo.

Isso significa que uma possível informação de profundidade perdida ocorreu na projeção ou devido ao ruído adicionado às imagens. A posição verdadeira era [41.1307], mas o programa a encontrou em [ 41.28071] (diferença de 0.15 mm, aproximadamente de 0.5 pixels)

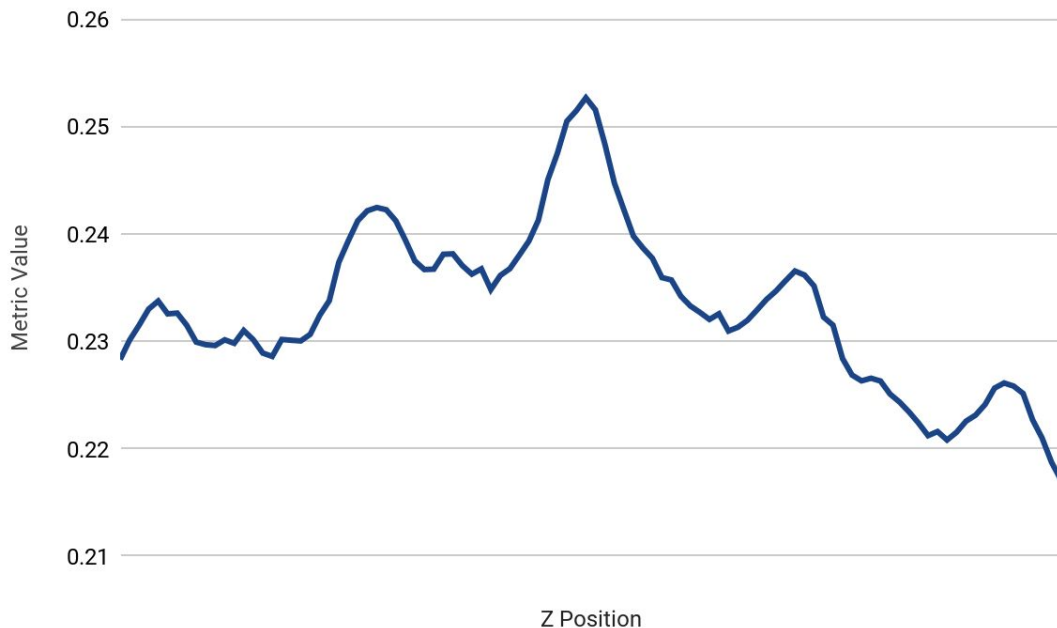


Figure 11: O gráfico mostra a variação do valor da métrica à medida que o robô caminha ao longo do eixo Z e passa seu valor real com um passo de 0,05 milímetros por amostra.

Como se pode observar com os 3 gráficos acima, o valor da métrica apresenta vários valores máximos locais que poderiam impedir o otimizador de alcançar o local correto de convergência.

O aumento no número de dimensões aumenta ainda mais esse problema.

A fim de aumentar o alcance da pesquisa bem-sucedida do registro, os seguintes métodos foram usados durante a fase de teste:

- Um filtro de desfoque foi aplicado para remover ruídos de baixa frequência e tentar aumentar a faixa de detecção da ferramenta durante as iterações iniciais do registro. Como o filtro de desfoque alterou as imagens, toda a métrica a ser calculada estava sob influência da intensidade decrescente do filtro, que resultou principalmente em erro no valor da métrica. Para evitar isso, os registros foram divididos em várias partes e o filtro de desfoque foi mantido constante durante cada parte e nulo durante a última parte. Este método não foi usado durante o período de avaliação, pois os testes foram feitos com distâncias pequenas (cerca de 2 milímetros) do ponto correto, mas alguns testes realizados mostraram resultados promissores, ajudando a aumentar o intervalo de registro nos casos em que o ponto de partida foi maior.
- As imagens geradas tinham 1024 x 1024 pixels. Enquanto o robô, o maior modelo a ser detectado, poderia facilmente caber dentro de 200 x 200 pixels nos casos estudados. Isso significa que o background teria uma influência muito alta no valor da métrica, muitas

vezes maior que as alterações devido a variações nos parâmetros do registro. Como forma de reduzir esse ruído na métrica, as imagens avaliadas foram recortadas em torno da posição do modelo a ser encontrado na imagem fixa. Isso apresentou uma boa mudança na precisão dos registros e foi utilizado em todos os testes durante a avaliação.

O uso de diferentes tipos de otimizadores também demonstrou ter uma grande influência na suscetibilidade do registro a valores extremos locais.

Embora otimizadores operacionais mais simples, como Gradient Descendent ou Amoeba, tivessem fortes tendências para estabilizar seus valores em extremos locais dependentes da localização inicial do registro, outros, como o One-Plus-One e o otimizador da Covariance Matrix Adaptation Evolution Strategy, eram mais robustos ( embora o otimizador One-Plus-One mostre uma maior variabilidade de resultados).

### Mutual Information Value per iteration

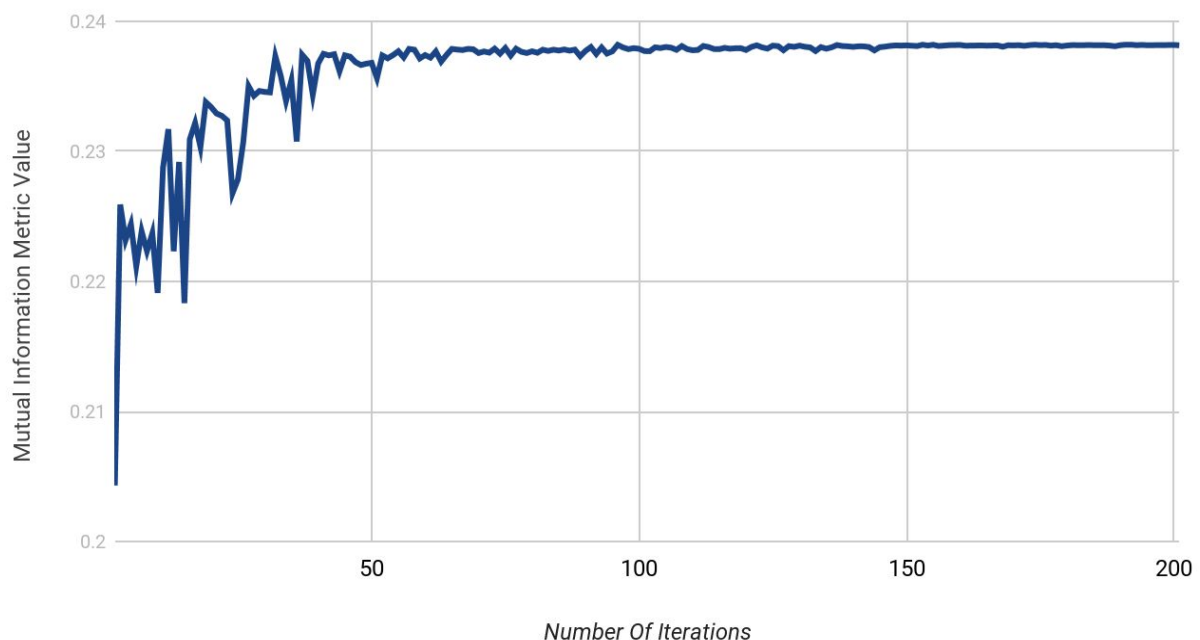


Figura 12: Um exemplo da variação da métrica durante o registro usando o Amoeba Optimizer

A fim de se aproximar do nível de complexidade apresentado na realidade, algumas modificações foram feitas nas imagens a serem analisadas(as fixas). Um filtro sigmóide foi aplicado em pixels, resultando em uma transformação linear na intensidade do pixel. A transformação total resultante é dada por:

$$f(x) = (Max - Min) * 1/(1 + e^{-(x-\beta)/\alpha}) + Min$$

Onde Max e Min são os valores máximos e mínimos de intensidade de pixel presentes na imagem e  $\alpha$  e  $\beta$  são parâmetros do formato da sigmóide.

---

Este filtro foi aplicado no volume a ser projetado e posteriormente na própria imagem.

Depois disso, um filtro de ruído gaussiano foi adicionado no final do pipeline.

A diferença entre as imagens geradas com o filtro de gradiente e a ser comparada pode ser observada a seguir.

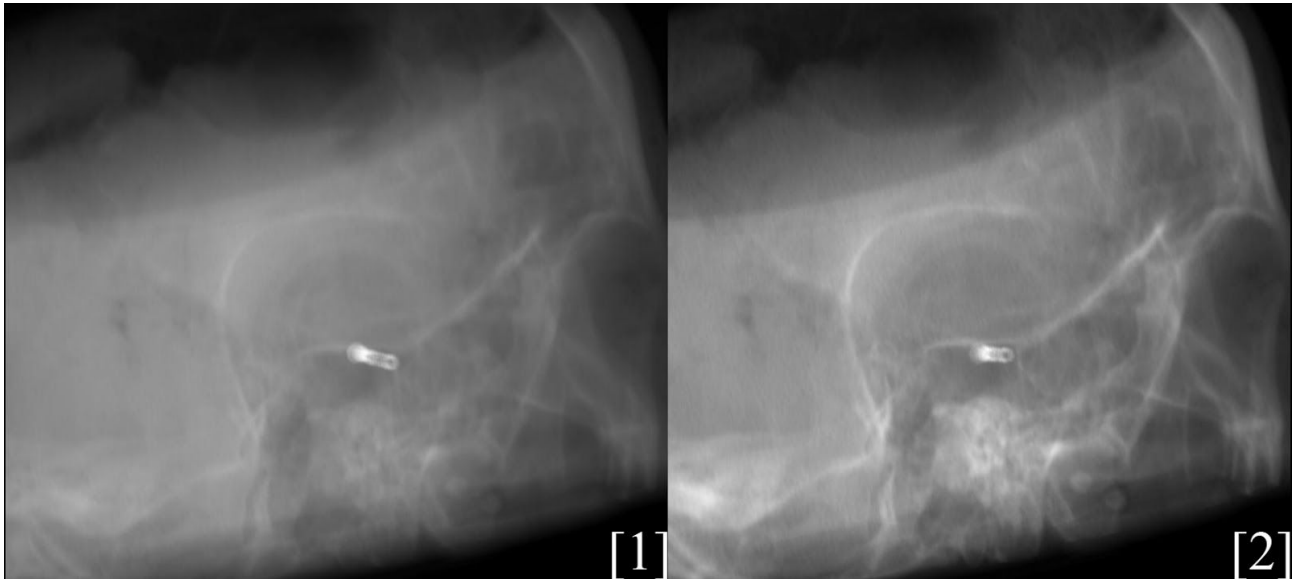
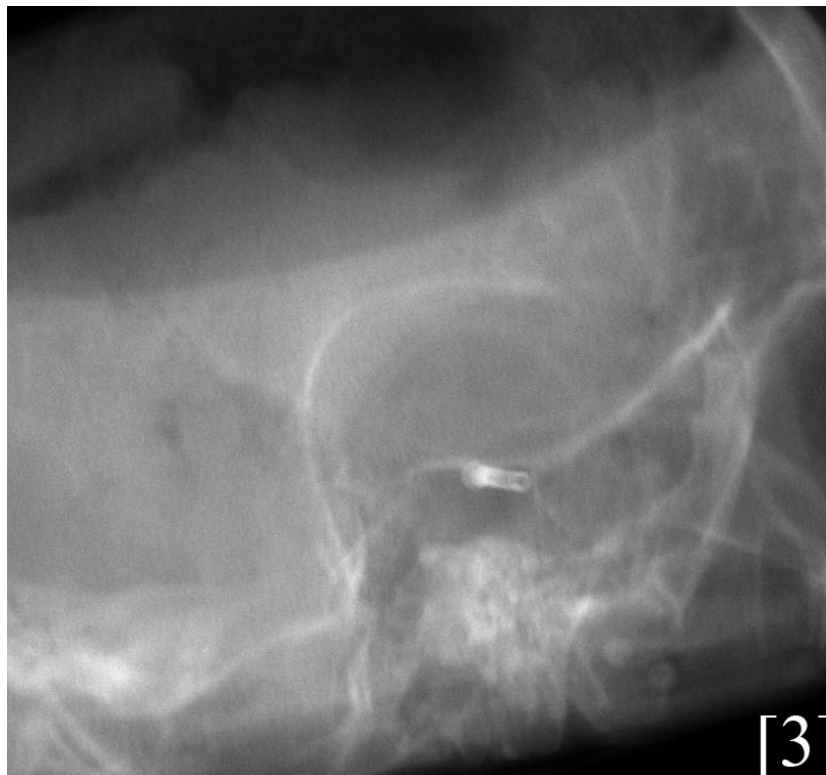


Figura 13: Comparação entre a imagem sem (1) e com (2) o filtro de gradiente sigmóide

A imagem final com adição de ruído pode ser vista abaixo.



---

Figura 1 4 : Imagem Final com Gradiente Sigmóide e Filtros de Tiro

---

## 5.2 A Métrica

---

Como mencionado anteriormente, nossa métrica foi adaptada para operação com um número de parâmetros a serem otimizados diferindo em relação à dimensionalidade da imagem a ser analisada.

Isso foi possível devido à mudança na forma como o cálculo da similaridade ocorreu. A nova métrica chama a API responsável pela comunicação com o gerador de imagens para passar os parâmetros a serem otimizados.

O fato de termos controle sobre a aquisição das imagens e o cálculo das métricas permitiu a implementação de métodos capazes de realizar esse cálculo para um número arbitrário de imagens com ou sem o corte do mesmo.

Uma vez com as imagens geradas, a métrica cria objetos clássicos de métricas do ITK para comparar imagens de dois a dois, a soma dessa métrica é então passada para o otimizador para o cálculo da próxima etapa.

Quatro métricas de similaridade foram implementadas em uma fase inicial de teste da ferramenta. Mas um número indefinido deles pode ser facilmente implementado mais tarde, de acordo com as necessidades.

Essas métricas são: Informações Mútuas, Diferença de Gradiente, Informações Mútuas de Mattes e Coeficiente de Correlação.

A métrica também é responsável por recortar as imagens de acordo com a necessidade. Os dois tipos de cortes usados até agora são feitos ao redor do modelo na imagem a ser analisada ou um corte duplo da cabeça e cauda do robô para comparação de posição usando os pontos dados como referência pela API geradora de imagem.

Em um caso real, esses pontos de referência poderiam ser dados como uma entrada do usuário ou como um processo de segmentação inicial (não relevantes no estudo apresentado aqui).

---

## 5.3 O Otimizador

---

Os otimizadores calculam a direção da variação dos parâmetros a serem testados em cada etapa da iteração.

Vários tipos de otimizadores existem e cada um tem um ponto de vantagem em alguma função em comparação a outros, mas nosso trabalho de registro 2D envolve a geração de imagens e isso cria um gargalo na possibilidade de escolher otimizadores que precisam de muitas iterações até alcançar uma convergência aceitável. .

Devido a isso, ficamos restritos a testes envolvendo apenas otimizadores de convergência não-derivativos e relativamente rápidos.

Além disso, damos preferência a otimizadores determinísticos sobre não determinísticos para evitar variações adicionais nos resultados.

Para a avaliação da ferramenta, utilizamos o Otimizador Covariance Matrix Adaptation Evolution Strategy como bases de operação.

O Otimizador Covariance Matrix Adaptation Evolution Strategy (CMA) retira uma quantidade de amostras com uma distribuição normal ao redor do ponto inicial e dependendo do valor da métrica em cada amostra, aumenta ou diminui a probabilidade dos próximos pontos serem amostrados mais perto desse ponto .

O desempenho dos otimizadores é geralmente altamente dependente do ajuste fino dos parâmetros operacionais de cada um. Como forma de evitar uma grande variação entre as configurações do otimizador e evitar a perda de tempo nessa configuração que não se encontra como objetivo desse estudo, todos foram usados com sua parametrização padrão.

Em contraste com a variação de valores de cada iteração do Otimizador Amoeba anteriormente apresentada, a imagem abaixo mostra a variação ao longo de um processo de registro que usa o Otimizador de estratégias Covariance Matrix Adaptation Evolution.

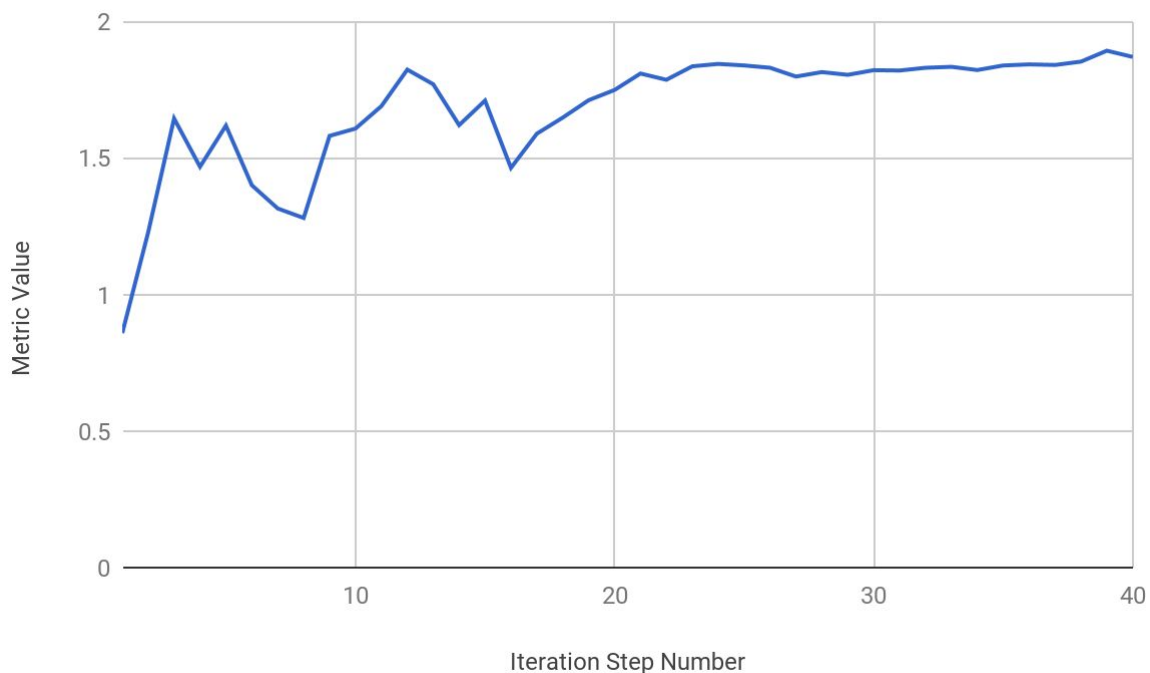


Figura 15:Variação de métrica durante um registro genérico usando o Covariance Matrix Adaptation Evolution Strategy Optimizer.

O sistema é compatível com qualquer otimizador que aceite um valor único, não derivativo. Embora uma implementação para a obtenção do valor derivativo tenha sido feita durante a criação do sistema, foi deixada de fora dos testes finais, uma vez que requer a geração de uma imagem para cada derivada possível (uma por parâmetro) e, portanto, precisaria também muito tempo para ser avaliado.

A implementação ainda está disponível no código-fonte do sistema para um exemplo de 3 parâmetros. Pode ser estendido facilmente para qualquer número necessário.

---

## 5.4 The API

---

O sistema responsável pela geração de imagens foi dado como material. Como consequência do fato de ainda estar em produção, diversas versões foram implementadas ao longo da execução deste projeto e, conseqüentemente, foi necessário criar uma interface de conexão entre o gerador de imagens e o sistema de registro de imagens sem dependência direta entre a implementação do gerador e do registro.

Nosso objeto de interface também foi responsável pelo carregamento de malhas, geração de volume e manipulação de malhas utilizadas durante a geração de imagens (utilizando as ferramentas fornecidas pelo gerador de imagens).

Tal interface é responsável por chamar gerador de imagem para gerar um volume a partir da anatomia dada e da matriz de transformação calculada pelo sistema. O gerador de imagens então combina ambos os volumes (utilizando o sistema previamente pronto ou a função de combinação de volume GPU otimizado feita durante esta tese) e no final alimenta o gerador de imagem com esse volume final de projeção.

As propriedades da projeção são derivadas da leitura de um arquivo JSON que descrevem de forma determinística ou estatística a distribuição dos ângulos de projeção.

---

## 5.5 Implementação da Comparação

---

Como forma de entender melhor onde o nosso trabalho se localiza no ambiente de pesquisa. Uma implementação de comparação foi feita.

O artigo [7] foi usado como modelo devido a proximidade dos resultados entre a implementação de nosso sistema e do paper, os resultados foram posteriormente avaliados (podem ser encontrados no próximo capítulo).

Como dito no paper, aproximadamente 200 imagens de raio-X de um corpo de ensaio foram tomadas e Coeficiente de Correlação foi usada para avaliar tais imagens 3 a 3 extraídas com um ângulo  $< 20$  graus. Como resultado, obtiveram um erro médio de registro de 1.5–4.2 mm, intervalo de captura de 95% de 8.7–11.4 mm em termos de TRE.

Nossa implementação de referência tentou chegar a tais resultados usando o mesmo tipo de teste (três projeções e abertura  $< 20$  graus) e sem fazer uso do método de corte de imagem, uma vez que não foi mencionado no artigo.

A ferramenta a ser detectada no referido papel tinha dimensões maiores que o nosso modelo, o que, conforme esperado, alterou o intervalo de captura.

---

## 6 Tests

---

Os testes tiveram como objetivo submeter a ferramenta ao maior número de casos possível dentro do período de tempo delimitado.

Para melhorar o desempenho dos testes foram feitas frequentemente reduções no número de parâmetros a serem otimizados.

Além disso, nossa validação é baseada na suposição de que nosso modelo está sempre dentro de 2 mm do início da pesquisa. Tal restrição foi dada como parte das características originais dos requisitos do sistema, uma vez que um sistema de segmentação se mostra complexo o suficiente para ser desenvolvido como um sistema a parte.

A avaliação da ferramenta foi utilizada para melhor entender quais fatores afetam o sucesso do registro do robô. Como consequência, manteve-se em mente ao decidir os cenários avaliados sempre o objetivo de melhor entender a influência de cada elemento.

Os elementos estudados foram:

- Número de projeções usadas na otimização da métrica de similaridade.
- Qual métrica foi usada.
- O tamanho do objeto a ser encontrado.
- A presença ou não de ruídos (gerados por computador) que variam entre 25 e 45 db (relação sinal-ruído de pico), abaixo desse valor a qualidade das imagens torna-se muito fraca para qualquer experimentação significativa.

---

### 6.1 Automatização de Testes

---

Cada caso de estudo geraria 2 posições aleatórias, dentro de 2 mm ao redor do valor real para o robô. Depois disso, o registro foi alimentado com as imagens fixas e os parâmetros iniciais da pesquisa.

Os testes foram feitos duas vezes para cada modelo e seguiram o padrão abaixo:

#### **Com o ruído ao redor de 35 db + CMA Optimizer:**

25 casos usando 1 projeção e Métrica de Correlation Coefficient Histogram.

25 casos usando 1 projeção e Métricas de Mattes Mutual Information.

25 casos usando 2 projeções e Métrica de Correlation Coefficient Histogram.

25 casos usando 2 projeções e Métricas de Mattes Mutual Information.

25 casos usando 4 projeções e Métrica de Correlation Coefficient Histogram.

25 casos usando 4 projeções e Métricas de Mattes Mutual Information.



---

**Com ruído ao redor de 25 db + CMA Optimizer:**

25 casos usando 1 projeção e Métrica de Correlation Coefficient Histogram.

25 casos usando 1 projeção e Métrica de Mattes Mutual Information.

25 casos usando 2 projeções e Métrica de Correlation Coefficient Histogram.

25 casos usando 2 projeções e Métrica de Mattes Mutual Information.

25 casos usando 4 projeções e Métrica de Correlation Coefficient Histogram.

25 casos usando 4 projeções e Métricas de Mattes Mutual Information.

**Com o ruído ao redor de 45 db + CMA Optimizer:**

(Parafuso) 25 casos usando 1 projeção e Métrica de Correlation Coefficient Histogram.

(Robot) 25 casos usando 1 projeção e Métrica de Mattes Mutual Information.

(Parafuso) 25 casos usando 2 projeções e Métrica de Correlation Coefficient Histogram.

(Robot) 25 casos usando 2 projeções e Métrica de Mattes Mutual Information.

(Parafuso) 25 casos usando 4 projeções e Métrica de Correlation Coefficient Histogram.

(Robot) 25 casos usando 4 projeções e Métrica de Mattes Mutual Information.

**Sem ruído + CMA Optimizer:**

25 casos usando 2 projeções e Métrica de Correlation Coefficient Histogram.

25 casos usando 2 projeções e Métrica de Mattes Mutual Information.

---

## 7 Resultados

---

Usando os padrões de teste apresentados acima, tentamos responder se nosso sistema pode ser considerado válido tanto como forma de avaliação dos métodos de registro quanto para uso futuro, dependendo da exatidão encontrada.

Esta geração e avaliação de testes foi destinada a ajudar a encontrar os principais pontos que afetam o registro com as referidas métricas, otimizadores e modelos.

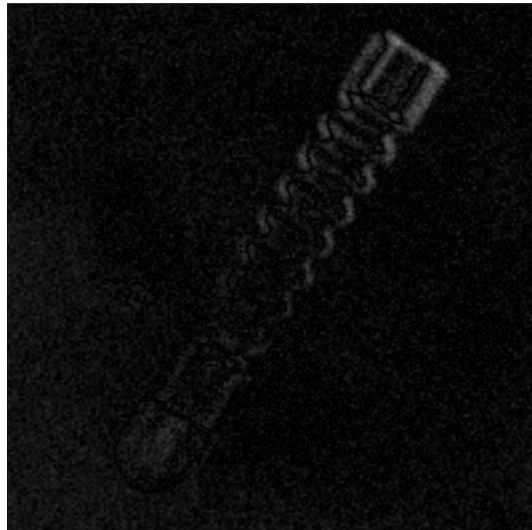


Figura 16: Diferença de intensidade das imagens do robô registrado com sucesso.

A figura acima mostra o resultado da diferença entre a imagem resultante e a imagem fixa em um caso típico de um registro bem-sucedido do robô em que a cabeça do robô está quase perfeitamente alinhada com a da imagem fixa, mostrando um bom desempenho resultante.

Avaliamos um registro como bem sucedido se a transformação resultante reduzir o distância euclidiana de ambos os parâmetros de posição ( $x$ ,  $y$ ,  $z$ ) e a orientação. Como nossos modelos tinham simetria radial, a rotação ao redor do eixo de simetria foi ignorada na medida.

Nos casos de 8 DOFs, essa simetria desapareceria, mas a rotação no eixo central poderia ser compensada por diferentes posições das juntas, por isso ainda não foi usado para decidir o sucesso do registro.

Também monitoramos as distâncias nas imagens resultantes. Mas ao otimizar os parâmetros para várias imagens, era comum haverem algumas imagens que tivessem um impacto maior nos valores da métrica, o que significa que eles apresentariam uma melhor aproximação à verdade básica. Com isso em mente, usamos apenas a melhor imagem resultante em cada registro para avaliar a distância em pixels entre as imagens fixas e em movimento.

Nossos testes apresentaram os seguintes resultados:

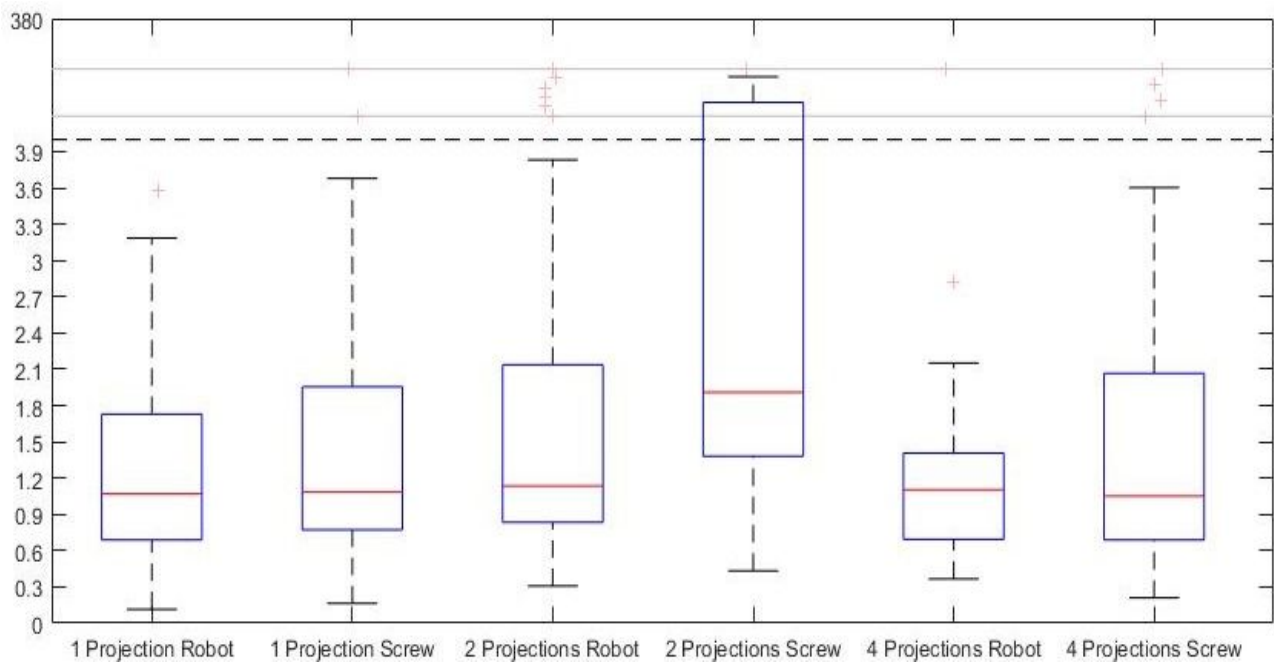


Figura 17: Gráfico Box Plot mostrando as distâncias em pixels entre as imagens registradas e de fundo para ambos os modelos e com 25db de pico relação sinal-ruído

| <b>Robô - Covariance Matrix Adaptation Evolution Estratégia Optimizer (+ ruído)</b> |                     |                          |                                     |                                 |                                    |
|---|---------------------|--------------------------|-------------------------------------|---------------------------------|------------------------------------|
| Métrica   | Número de projeções | Relação pico sinal-ruído | Distância Final Mediana (em pixels) | Distância final mediana (em mm) | Taxa de Sucesso (em 50 tentativas) |
| Mattes MI   | 1                   | 25                       | 1.69                                | 1.02                            | 68%                                |
| Mattes MI   | 2                   | 25                       | 1.47                                | 0.35                            | 100%                               |
| Mattes MI   | 4                   | 25                       | 1.39                                | 0.32                            | 100%                               |
| Correlation Coefficient Histogram Metric  | 1                   | 25                       | 1.07                                | 1.57                            | 84%                                |
| Correlation   | 2                   | 25                       | 1.13                                | 0.35                            | 94%                                |

|  |   |    |      |      |      |
|--|---|----|------|------|------|
| Coefficient Histogram Metric             |   |    |      |      |      |
| Correlation Coefficient Histogram Metric | 4 | 25 | 1.10 | 0.25 | 100% |
| Mattes MI                                | 1 | 35 | 1.25 | 1.57 | 92%  |
| Mattes MI                                | 2 | 35 | 1.21 | 0.27 | 98%  |
| Mattes MI                                | 4 | 35 | 1.08 | 0.24 | 98%  |
| Correlation Coefficient Histogram Metric | 1 | 35 | 0.96 | 1.12 | 88%  |
| Correlation Coefficient Histogram Metric | 2 | 35 | 1.29 | 0.32 | 94%  |
| Correlation Coefficient Histogram Metric | 4 | 35 | 1.01 | 0.23 | 100% |
| Mattes MI                                | 1 | 45 | 1.08 | 1.13 | 84%  |
| Mattes MI                                | 2 | 45 | 1.11 | 0.29 | 96%  |
| Mattes MI                                | 4 | 45 | 1.14 | 0.24 | 98%  |

| <b>Parafuso - Covariance Matrix Adaptation Evolution Strategy Optimizer(+ ruído)</b> |                     |                          |  |                                 |                 |
|--|---------------------|--------------------------|--|---------------------------------|-----------------|
| Métrica  | Número de projeções | Relação pico sinal-ruído | Distância final mediana (em pixels na melhor projeção) | Distância final mediana (em mm) | Taxa de sucesso |
| Mattes MI  | 1                   | 25                       | 3.43   | 1.45                            | 40%             |
| Mattes MI  | 2                   | 25                       | 5.58   | 0.77                            | 68%             |
| Mattes MI  | 4                   | 25                       | 3.09   | 0.49                            | 62%             |
| Correlation Coefficient Histogram Metric   | 1                   | 25                       | 1.08   | 0.97                            | 72%             |
| Correlation Coefficient Histogram Metric   | 2                   | 25                       | 1.91   | 0.30                            | 86%             |
| Correlation Coefficient Histogram Metric   | 4                   | 25                       | 1.09   | 0.23                            | 92%             |
| Mattes MI  | 1                   | 35                       | 2.89   | 1.23                            | 60%             |
| Mattes MI  | 2                   | 35                       | 3.25   | 0.80                            | 68%             |
| Mattes MI  | 4                   | 35                       | 1.70   | 0.32                            | 72%             |
| Correlation Coefficient Histogram Metric   | 1                   | 35                       | 0.87   | 0.72                            | 86%             |
| Correlation Coefficient Histogram Metric   | 2                   | 35                       | 1.35   | 0.30                            | 86%             |
| Correlation  | 4                   | 35                       | 1.28   | 0.20                            | 84%             |

| Coefficient Histogram Metric             |   |    |      |      |     |
|--|---|----|------|------|-----|
| Correlation Coefficient Histogram Metric | 1 | 45 | 0.74 | 0.81 | 68% |
| Correlation Coefficient Histogram Metric | 2 | 45 | 1.26 | 0.28 | 80% |
| Correlation Coefficient Histogram Metric | 4 | 45 | 1.19 | 0.21 | 82% |

| <b>Robô - Covariance Matrix Adaptation Evolution Strategy Optimizer (sem ruído)</b>     |                     |                             |                      |                             |
|---|---------------------|-----------------------------|----------------------|-----------------------------|
| Métrica   | Número de Projeções | Distância mediana em Pixels | Distância Mediana mm | Taxa de sucesso no registro |
| Mattes MI   | 2                   | 0.76                        | 0.22                 | 88%                         |
| Correlation Coefficient Histogram Metric  | 2                   | 0.62                        | 0.17                 | 96%                         |
| <b>Parafuso - Covariance Matrix Adaptation Evolution Strategy Optimizer (sem ruído)</b> |                     |                             |                      |                             |
| Mattes MI   | 2                   | 1.01                        | 0.22                 | 40%                         |
| Correlation Coefficient Histogram Metric  | 2                   | 0.34                        | 0.087                | 94%                         |

---

## 7.1 Resultados relacionados a ruídos

---

Como esperado, a quantidade de ruído e a métrica usada tiveram um grande efeito nas distâncias resultantes em ambas as imagens.

O impacto do ruído pode ser visto principalmente no registro do parafuso (um modelo menor). A quantidade de projeções a serem medidas ajudou a reduzir o número de registros fracassados, mas ambas as métricas mostraram uma baixa taxa de sucesso com altos níveis de ruído, isso se mostra como um resultado possível da relação entre o tamanho do modelo e a quantidade de ruído. Com o robô, a quantidade de ruído não causaria grandes mudanças para a parte do histograma que o robô se encontrava. Com o parafuso, por outro lado, alteraria grande parte dos pixels correspondentes a projeção do modelo.

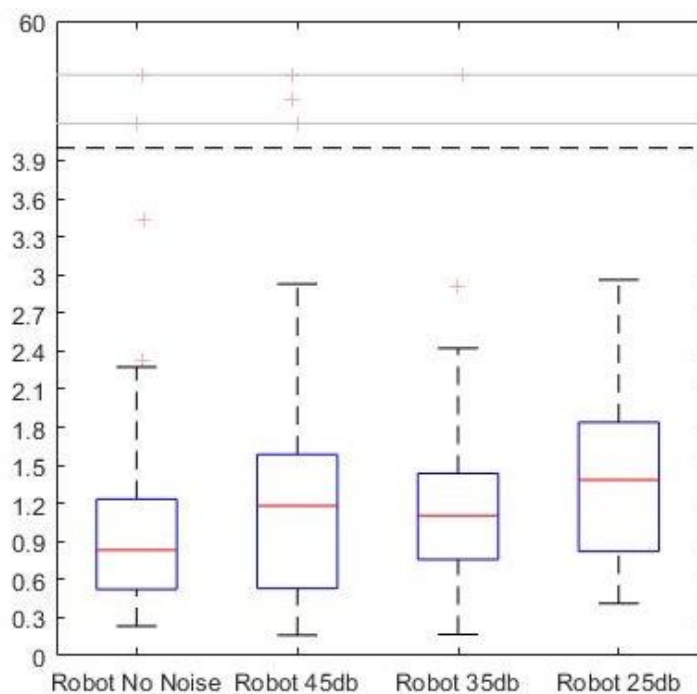


Figure 18: Gráfico Box Plot mostrando as distâncias em pixels entre as imagens reais registradas e terrestres ao comparar diferentes valores de ruído

---

Um resultado inesperado mostrou que a precisão da estimativa de pose, quando fazendo uso de várias projeções, tende a variar muito pouco com o aumento do ruído, enquanto a taxa de sucesso do registro sofreu uma variação maior

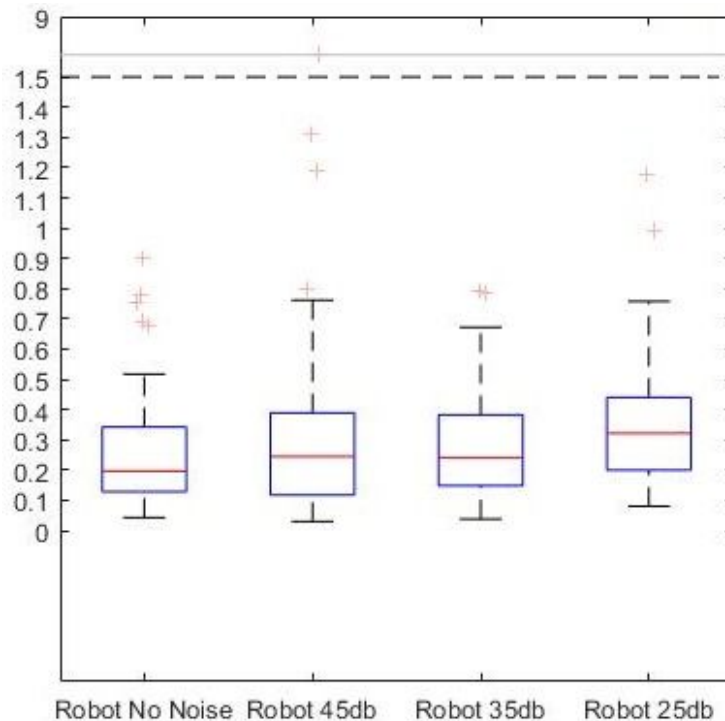


Figure 19: Gráfico Box Plot mostrando as distâncias em milímetros entre as imagens verdadeiras registradas e a realidade com o uso de 4 projeções

---

## 7.2 Resultados relacionados a Métrica

---

A mudança das métricas usadas teve impacto tanto na precisão quanto na robustez. Coefficient Correlation mostrou-se um pouco mais precisa na maioria das vezes, enquanto Mattes Mutual Information com mais frequência convergiu para os extremos locais, mas funcionou bem com grandes quantidades de ruído. Isso era esperado uma vez que [10] mencionou que as métricas baseadas em correspondência na intensidade local eram mais robustas e precisas para o registro 3D-2D.

Coefficient Correlation apresentou uma maior sensibilidade para a mudança no número de imagens. É possível observar que, ao sofrer com uma quantidade elevada de ruído, a mudança de 1 para 2 projeções teria, muitas vezes, pouco impacto no registro quando comparada ao Mattes, ao passo que, ao usar 4, alcançamos melhores resultados.



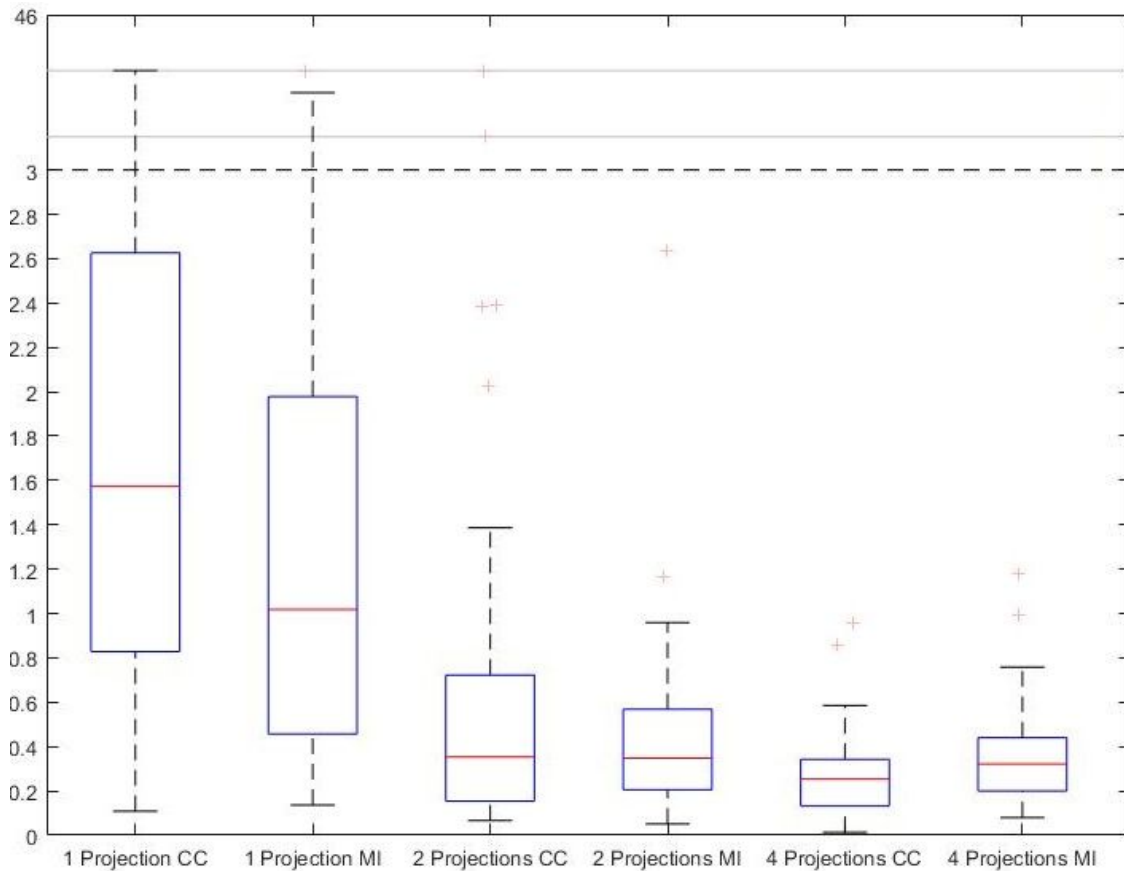


Figura 20: Gráfico Box Plot mostrando as distâncias em milímetros entre as imagens verdadeiras registradas e terrestres comparando a Coefficient Correlation e a Mattes Mutual Information

### 7.3 Número de projeções

Os resultados mostraram que o uso de uma projeção muitas vezes resultaria em um melhor posicionamento quando comparamos as distâncias nas imagens. Embora inesperado, isso pode ser explicado, por exemplo, pelo fato de que mudanças nas métricas originadas por uma ou mais imagens poderiam mascarar mudanças nas métricas originadas de outras. Como consequência disso, é possível ver que muitas vezes a configuração das duas projeções tinha uma distância média de imagem pior entre o modelo e a verdade do solo para o caso do parafuso. Isso pode significar que há um número ideal de projeções a serem usadas ao tentar minimizar as distâncias nas imagens.

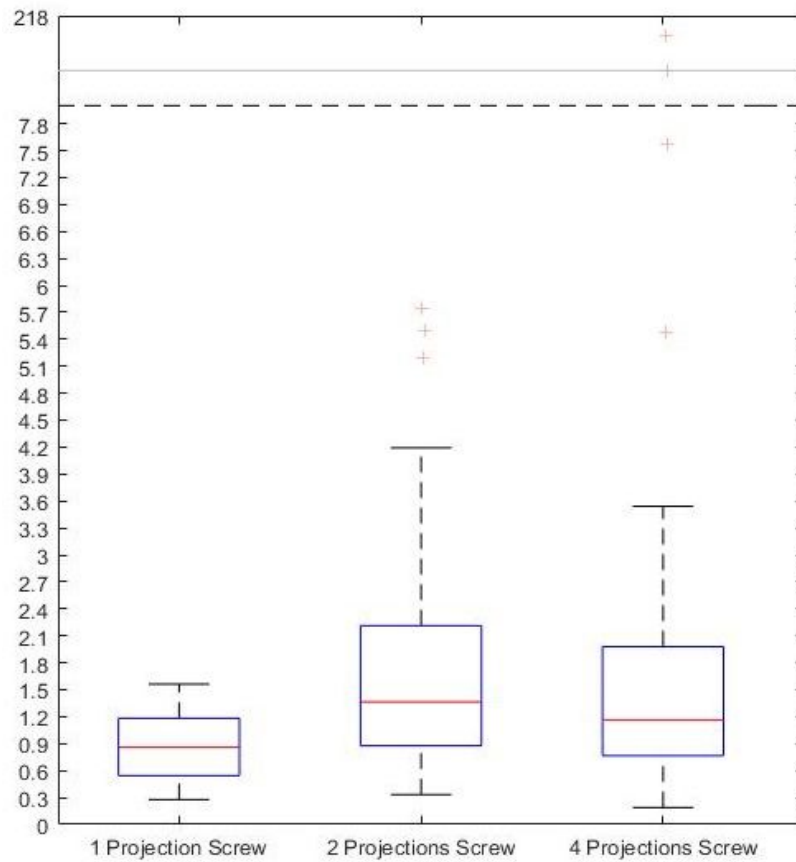


Figura 21: Gráfico Box Plot mostrando as distâncias médias em pixels entre as imagens geradas e as de referência, comparando os resultados do registro com o número crescente de projeções (35db de ruído)

Como mencionado anteriormente, pode-se também observar uma forte tendência à robustez à medida que o número de projeções aumenta. Ela ajuda tanto a aumentar a taxa de sucesso dos registros quanto a estimar a estimativa de grandes variações devido ao ruído.

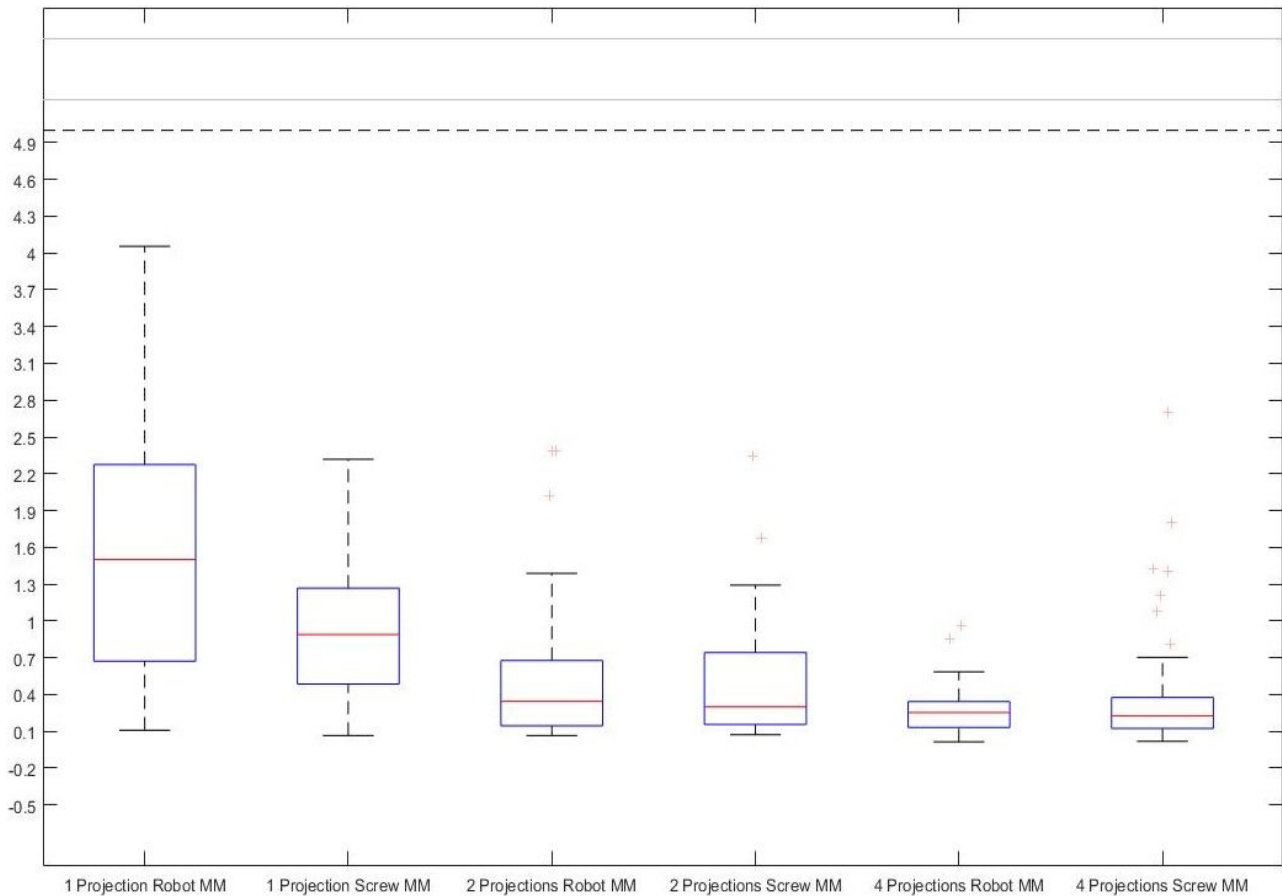


Figura 22: Gráfico Box Plot mostrando as distâncias em milímetros entre as imagens reais registradas e realidade (25db de pico relação sinal-ruído)

A distância média nas imagens melhoraria quanto menos projeções fossem utilizadas, mas, embora a precisão da estimativa de pose também mostrasse ser altamente dependente da quantidade de projeções, melhoraria quanto mais projeções fossem usadas em todos os casos de teste. Isso é esperado, uma vez que, como mencionado anteriormente, a projeção de imagens tende a degradar uma certa quantidade das informações de profundidade. A variação dos ângulos de projeção ajudou a dissipar este erro entre as várias dimensões.

Além desses três fatos mais evidentes, podemos ver vários resultados bem conhecidos, como o fato da quantidade de ruído reduzir a robustez e a precisão do registro, que o otimizador utilizado teve um grande impacto sobre a precisão final e que um corpo maior foi mais facilmente registrado e menos afetado pela quantidade de ruído.

Os resultados já conhecidos, juntamente com os inesperados que poderiam ser explicados, mostram que o sistema implementado apresentou o comportamento esperado.

---

## 8 Conclusion

---

No trabalho aqui evidenciado, apresentamos uma implementação de um sistema de registro 2D / 3D focado na avaliação de diferentes configurações de métricas, otimizadores, métodos e modelos.

Além da implementação de base que usa a métrica de Mutual Information e o corte de imagens para medir a semelhança, uma implementação secundária baseada em [7] foi feita no mesmo seguindo o mesmo sistema como base de comparação.

Cerca de 1400 testes que variaram o número de projeções, métricas e modelos foram utilizados como avaliação comparando imagens geradas fixas, modificadas para tentar simular possíveis complexidades quando comparadas com casos reais, com imagens geradas puras.

Nossos resultados se apresentaram na sua maior parte conforme o esperado, com uma precisão altamente dependente da quantidade de ruído e da razão entre a quantidade de ruído e o tamanho do modelo a ser encontrado.

A variação dos otimizadores também teve forte influência na robustez dos registros, com a CMA apresentando-se como a melhor.

Pode-se dizer que o sistema alcançou o que era exigido dele e foi apresentado com melhorias quando comparado com o propósito inicial.

---

### 8.1 Resumo

---

Foi apresentado aqui um sistema de estimativa de pose tanto de instrumentos cirúrgicos rígidos quanto não-rígidos que modela o problema como um registro de imagens. Utilizando informações da anatomia a ser estudada e o modelo 3D conhecido do instrumento.

A geração de imagens mostrou-se como o principal gargalo para o registro, pois requer o uso de programas externos na geração da malha e diversas etapas para o carregamento, geração e projeção dos volumes necessários para o sistema.

Para alcançar um nível de complexidade mais alto, um pipeline de geração de ruído foi adicionado à geração das imagens fixas. Este foi dividido em 2 partes, o primeiro mudando a intensidade dos pixels para uma distribuição não linear tanto do volume como da imagem. O segundo adicionando ruído gaussiano à imagem final.

As imagens também passaram por um método de corte que isolaria a área ao redor do alvo. Isso foi feito para reduzir os ruídos originados das imagens como um todo.

Como as imagens foram geradas de acordo com o número de projeções requeridas, elas seriam comparadas com a respectiva imagem fixa e o valor total seria enviado ao otimizador como valores da iteração. Isso mostrou resultados muito bons na estimativa de pose ao usar mais de uma projeção.

Como forma de validar nossa implementação, uma extensa série de testes foi realizada. Os resultados foram comparados com conhecimento bem estabelecido e mostraram que o sistema apresentou um comportamento correto.

---

Uma vez que o sistema foi considerado funcionando corretamente, foi feito uso do mesmo para o estudo das influências do uso de várias projeções no registro e foi encontrado que a variação da geração de ruído produzia um impacto tanto na robustez quanto na precisão dos resultados obtidos.

Utilizando o modelo do robô, com 35db de razão Peak Signal-to-Noise, usando Coeficiente de Correlação e o Otimizador de Covariance Matrix Adaptation Evolution Strategy obtivemos um erro linear médio de 0,25 mm e erro angular médio de 1,30 graus na estimativa de pose, comparável a resultados obtidos em estudos recente publicados.

---

## 8.2 Trabalhos Futuros

---

Como continuações, a melhoria adicional do método poderia concentrar-se no problema dos extremos locais que, embora reduzido usando o desfoque e o recorte das imagens, permanece como um grande obstáculo no registro.

Como os testes mostraram que o principal gargalo do desempenho foi a geração das imagens, principalmente no trabalho relacionado com a manipulação de volumes.

Uma boa melhoria no desempenho poderia ser alcançada pela mudança no sistema de trabalho com volumes, onde poderia-se manter o volume da anatomia na memória da GPU. Dessa forma, poderíamos evitar ter que enviá-lo para a GPU em cada iteração, o que ajudaria a evitar o gargalo inerente à limitação no volume e velocidade da troca de dados com a GPU. Outro ponto de melhoria poderia concentrar-se na melhoria da posição encontrada no caso de menos projeções.

A maior parte dos testes mostraram que, ao trabalhar com uma única projeção, mesmo que a precisão do registro pareça ser realmente boa (0,5 pixels ou menos), a posição 3D encontrada e a verdadeira não estavam tão próximas quanto se poderia esperar.

Embora fosse esperado desde o início, dado o fato de que a projeção perde informações principalmente em relação à profundidade, às vezes a posição resultante levaria o robô ainda mais longe da posição real quando comparado com o ponto inicial.

---

## 9 References

---

[1]-Rohlfing, Torsten; Russakoff, Daniel B.; Denzler, Joachim; Mori, Kensaku; Maurer, Calvin R. (2005): Progressive attenuation fields: fast 2D-3D image registration without precomputation. In *Medical physics* 32 (9), pp. 2870–2880. DOI: 10.1118/1.1997367.

[2]-Maintz, J.B., Viergever, M.A., 1998. A survey of medical image registration. *Med. Image Anal.* 2 (1), 1–36]

[3]-Vandini, Alessandro; Bergeles, Christos; Glocker, Ben; Giataganas, Petros; Yang, Guang-Zhong (2017): Unified Tracking and Shape Estimation for Concentric Tube Robots. In *IEEE Trans. Robot.*, pp. 1–15. DOI: 10.1109/TRO.2017.2690977.]

- 
- [4]-Jain, Ameet Kumar; Mustafa, Tabish; Zhou, Yu; Burdette, Clif; Chirikjian, Gregory S.; Fichtinger, Gabor (2005): FTRAC--a robust fluoroscope tracking fiducial. In *Medical physics* 32 (10), pp. 3185–3198. DOI: 10.1118/1.2047782.]
- [5]-Shin, Sangkyun; Kim, Youngjun; Cho, Hyunchul; Lee, Deukhee; Park, Sehyung; Kim, Gerard Jounghyun; Kim, Laehyun (2014): A single camera tracking system for 3D position, grasper angle, and rolling angle of laparoscopic instruments. In *Int. J. Precis. Eng. Manuf.* 15 (10), pp. 2155–2160. DOI: 10.1007/s12541-014-0576-6.]
- [6]-Uneri A. et al. (2016) Deformable 3D-2D Registration of Known Components for Image Guidance in Spine Surgery. In: Ourselin S., Joskowicz L., Sabuncu M., Unal G., Wells W. (eds) *Medical Image Computing and Computer-Assisted Intervention -- MICCAI 2016*. MICCAI 2016. Lecture Notes in Computer Science, vol 9902. Springer, Cham
- [7]-Gao, Gang; Penney, Graeme; Ma, Yingliang; Gogin, Nicolas; Cathier, Pascal; Arujuna, Aruna et al. (2012): Registration of 3D trans-esophageal echocardiography to X-ray fluoroscopy using image-based probe tracking. In *Medical image analysis* 16 (1), pp. 38–49. DOI: 10.1016/j.media.2011.05.003.
- [8]-Esfandiari, H.; Amiri, S.; Lichti, D. D.; Anglin, C. (2016): A fast, accurate and closed-form method for pose recognition of an intramedullary nail using a tracked C-arm. In *International journal of computer assisted radiology and surgery* 11 (4), pp. 621–633. DOI: 10.1007/s11548-015-1294-y.
- [9]-Uneri, A.; Silva, T. de; Stayman, J. W.; Kleinszig, G.; Vogt, S.; Khanna, A. J. et al. (2015): Known-component 3D-2D registration for quality assurance of spine surgery pedicle screw placement. In *Physics in medicine and biology* 60 (20), pp. 8007–8024. DOI: 10.1088/0031-9155/60/20/8007.
- [10]-Otake, Y.; Schafer, S.; Stayman, J. W.; Zbijewski, W.; Kleinszig, G.; Graumann, R. et al. (2012): Automatic localization of vertebral levels in x-ray fluoroscopy using 3D-2D registration: a tool to reduce wrong-site surgery. In *Physics in medicine and biology* 57 (17), pp. 5485–5508. DOI: 10.1088/0031-9155/57/17/5485.
- [11]-Otake, Yoshito; Wang, Adam S.; Stayman, J. Webster; Uneri, Ali; Kleinszig, Gerhard; Vogt, Sebastian et al. (2013): Robust 3D–2D image registration. Application to spine interventions and vertebral labeling in the presence of anatomical deformation. In *Physics in medicine and biology* 58 (23), pp. 8535–8553. DOI: 10.1088/0031-9155/58/23/8535.
- [12]- Markelj, P.; Tomažević, D.; Likar, B.; Pernuš, F. (2012): A review of 3D/2D registration methods for image-guided interventions. In *Medical image analysis* 16 (3), pp. 642–661. DOI: 10.1016/j.media.2010.03.005.
- [13]-Wang, Lei; Gao, Xin; Fang, Qiang (2013): A novel mutual information-based similarity measure for 2D/3D registration in image guided intervention. In : 2013 International Conference on Orange Technologies (ICOT). 12-16 March 2013, Tainan, Taiwan. 2013 1st International Conference on Orange Technologies (ICOT 2013). Tainan. International Conference on Orange Technologies; ICOT. Piscataway, NJ: IEEE, pp. 135–138.
- [14]-Doignon, Christophe; Mathelin, Michel de (2007): A Degenerate Conic-Based Method for a Direct Fitting and 3-D Pose of Cylinders with a Single Perspective View. In : *Proceedings 2007*

---

IEEE International Conference on Robotics and Automation. 2007 IEEE International Conference on Robotics and Automation. Rome, Italy: IEEE, pp. 4220–4225.

[15]-Amankwah, Anthony (2013): Image registration by automatic subimage selection and maximization of combined mutual information and spatial information. In : 2013 IEEE International Geoscience and Remote Sensing Symposium - IGARSS. IGARSS 2013 - 2013 IEEE International Geoscience and Remote Sensing Symposium. Melbourne, Australia: IEEE, pp. 4379–4382.

[16]-Russakoff D *et al* 2003a Evaluation of intensity-based 2D-3D spine image registration using clinical gold-standard data *Biomedical Image Registration* ed J Gee, J Maintz and M Vannier (Berlin/Heidelberg: Springer) pp 151–160

[17]-Bhunre, Piyush; Leow, Wee; Howe, Tet (2007): Recovery of 3D Pose of Bones in Single 2D X-ray Images. In : 2007 IEEE Workshop on Applications of Computer Vision (WACV '07). 2007 IEEE Workshop on Applications of Computer Vision (WACV '07). Austin, TX, USA: IEEE, p. 48.

[18]- P. Dupont, J. Lock et al., “Design and control of concentric-tube robots,”*IEEE Trans. Robot.*, vol. 26, no. 2, pp. 209–225, Apr. 2010.

[19]-Penney, G. P.; Weese, J.; Little, J. A.; Desmedt, P.; Hill, D. L.; Hawkes, D. J. (1998): A comparison of similarity measures for use in 2-D-3-D medical image registration. In *IEEE transactions on medical imaging* 17 (4), pp. 586–595. DOI: 10.1109/42.730403.

[20]-S. Klein, M. Staring, K. Murphy, M.A. Viergever, J.P.W. Pluim, "elastix: a toolbox for intensity based medical image registration," *IEEE Transactions on Medical Imaging*, vol. 29, no. 1, pp. 196 - 205, January 2010.

[21]-D.P. Shamonin, E.E. Bron, B.P.F. Lelieveldt, M. Smits, S. Klein and M. Staring, "Fast Parallel Image Registration on CPU and GPU for Diagnostic Classification of Alzheimer's Disease", *Frontiers in Neuroinformatics*, vol. 7, no. 50, pp. 1-15, January 2014.