Journal of Computer Science

Original Research Paper

A Method for Fast and Robust Intensity Based 2D/3D Registration in Differentiable Framework

[1]Agung Alfiansyah, [2]Izzati Muhimmah, [3]Arif Sasongko, [4]Heru Sulastomo and [4]Nanang Wiyono

[1]*Department of Computer Systems Engineering, Universitas Prasetiya Mulya, Tangerang, Indonesia*
[2]*Department of Informatics, Universitas Islam Indonesia, Yogyakarta, Indonesia*
[3]*School of Electrical Engineering and Informatics, Institut Teknologi Bandung, Bandung, Indonesia*
[4]*Faculty of Medicine, Universitas Sebelas Maret, Surakarta, Indonesia*

Article history
Received: 03-04-2024
Revised: 18-07-2024
Accepted: 28-08-2024

Corresponding Authors:
Agung Alfiansyah
Department of Computer Systems Engineering,
Universitas Prasetiya Mulya,
Tangerang, Indonesia
Email: agung.alfiansyah@prasetiyamulya.ac.id

Abstract: This study introduces an innovative intensity-based registration method designed to align 3D image data volumes, such as Computed Tomography (CT), with 2D data, such as X-ray images, a crucial component of intra-operative navigation in image-guided surgery. Leveraging advanced autodifferentiable reconstruction techniques based on the GPU-accelerated Siddon algorithm, the proposed approach facilitates the generation of differentiable Digitally Reconstructed Radiographs (DRRs), enabling automatic image derivation according to its positional polar coordinates. Prior to the optimization process, an in-depth evaluation of the loss function landscape demonstrates its convex nature, revealing a global minimum aligning with the registered position of both datasets. Subsequently, various optimization algorithms, including Gradient Descent with Momentum, Gradient Descent with Damped Momentum, and Adam, undergo meticulous assessment for their efficacy in minimizing the registration loss function. Among these methods, Adam demonstrates particularly promising outcomes in terms of convergence rate and execution time. Furthermore, robustness evaluation through Monte Carlo simulations underscores the method's remarkable ability to maintain registration accuracy even under perturbations. Despite inherent limitations related to DRR fidelity, such as the absence of reflection and scattering modeling, the proposed registration method achieves alignment. Robustness testing through Monte Carlo simulations, initializing positions at perturbed random locations, and applying the Adam optimization scheme, yields a convergence rate of 82.7%. The study outlines future research aimed at enhancing DRR fidelity through Monte Carlo simulations, refining registration accuracy, and expanding application to real anatomical parts. Overall, the proposed optimization strategies offer a compelling and versatile approach for rapid and precise 2/3D image registration in image-guided surgery, with significant potential for augmenting surgical precision and ultimately improving patient outcomes.

Keywords: 2D/3D Registration, Intensity Based Registration, Autograd, Digitally Reconstructed Radiographs

Introduction

In the hospital setting, it is standard daily practice to obtain preoperative patient images for diagnosis, procedure planning, and creating surgical roadmaps. These preoperative images are typically high-quality 3D scans, often taken with technologies like Magnetic Resonance (MR) or Computed Tomography (CT) scanners.

In contrast, images acquired during medical procedures tend to have lower quality. They have lower signal-to-oise ratios (SNR) and are often 2D slices or projections, lacking the full 3D depth. However, these intra-operative images provide higher spatial resolution due to the capabilities of intra-operative imaging devices. This real-time data is vital for assessing instrument placement and patient anatomy conditions during surgery. While recent advancements have introduced 3D imaging into operating rooms, such as 3D rotational angiography, they are used infrequently due to time constraints and concerns about high X-ray exposure.

When dealing with two sets of data acquired at different times, the registration process is essential to integrate preoperative and intraoperative datasets. This allows the fusion of high-quality preoperative patient anatomy data with real-time surgical data, paving the way for innovative intraoperative roadmapping and navigation methods. This improves treatment efficiency and reduces radiation exposure for both medical professionals and patients.

This system is particularly valuable for assisting surgeons, especially in areas with deformable and noble anatomy, like the thoracic and abdominal regions. Precise alignment in these cases is crucial, as navigating catheters through dynamic structures requires skill, training, and anatomical knowledge. The primary aim of this system is to establish a spatial relationship between a fixed 3D world coordinate system and the coordinate system of a 2D imaging device. In a medical context, this alignment is crucial for matching clinical 3D tomographic images, obtained pre-operatively, with their corresponding 2D projections, often from intra-operative fluoroscopy images. These X-ray images typically come from devices like interventional C-arms or conventional radiography. Importantly, this process does not always have a direct one-to-one correspondence between 3D and 2D data points.

Our paperwork proposes a different approach to the conventional method for solving the 2/3D data registration problem. We use an image intensity-based approach, eliminating the need for feature extraction in both data sets. Additionally, we introduce an optimization scheme based on gradient descent, which includes the reconstruction of simulated DRR images as differentiable data. This differs from the traditional approach, which relies on heuristic methods due to the inability to derive data from images. We also present methods for accelerating DRR reconstruction in this research, achieved through GPU computing platforms that significantly speed up the method.

Materials and Methods

The following exposition of our proposed methods will begin with a discussion of the techniques used in DRR reconstruction, reformulation, and vectorization to accelerate DRR. We will explore how auto grad can be applied to DRR, making the data auto-differentiable on open-access computing platforms. The paper will then delve into the gradient descent-based 2/3D data registration technique using DRR images obtained from our proposed method. In the end, we conduct an intensive experiment to evaluate our method and look for a feasibility study to apply the method in computer-assisted heart surgery.

Digitally Reconstructed Radiographs

Digital Reconstructed Radiographs (DRRs) are computer-simulated radiographic images that result from projecting a Three-Dimensional (3D) volume onto a Two-Dimensional (2D) image plane. Typically, DRRs are created by simulating X-ray images from Computed Tomography (CT) volumes. DRRs offer several advantages over traditional X-rays, including the ability to generate images from various angles in real time and manipulate parameters like the emitter distance, which would be impractical with conventional X-rays. In the medical field, DRRs are increasingly being used to enable continuous monitoring of a patient's position during medical procedures using 3D imaging technology. For instance, a pre-operative CT scan can generate DRRs, which can then be matched with X-rays taken during the actual procedure. This registration of real and simulated X-rays provides valuable real-time information, accounting for minor changes in the patient's position. This process is both fast and accurate, making it suitable for online clinical applications.

To simulate two-dimensional images from CT scan volumetric data, we conduct mathematical radiographic projection as illustrated in Fig. (1). Let $s \in R \neq$ represent the X-ray source point in the fluoroscopy imaging device, and $p \in R \neq$ denote the estimated pixel intensity on the imaging sensor plane. A ray originating from the point $s = \alpha = 0$propagates through the volume to be projected and ultimately strikes the sensors on the detector plane at the endpoint $p = \alpha = 1$. The cumulative energy attenuation of the X-ray upon reaching pixel p can be expressed mathematically as follows:

$$E(R) = \|x+y\|_2 \int_0^1 V(s+\alpha(p-s))\, d\alpha \quad (1)$$

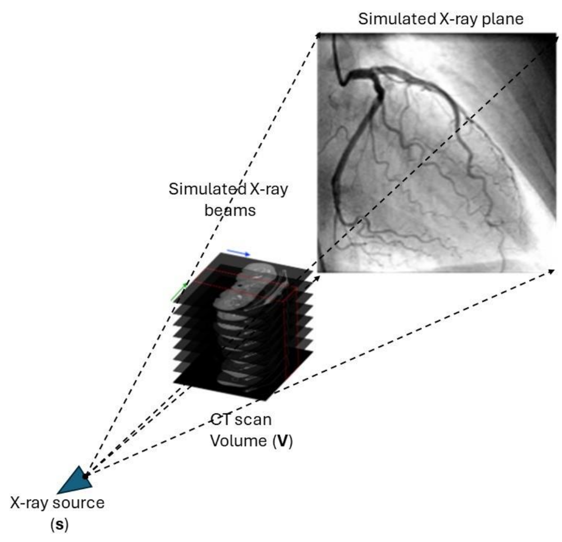

Fig. 1: DRR using the Siddon method

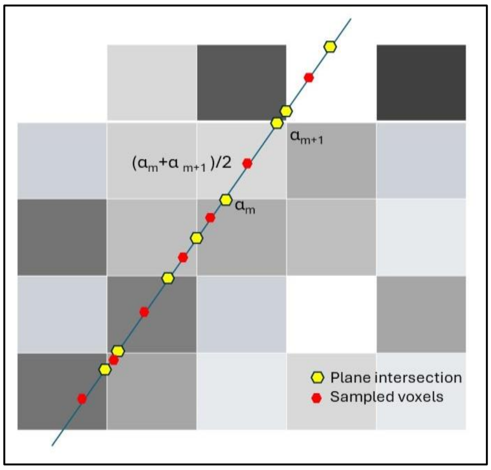

Fig. 2: Discrete approximation of the DRR volume

where, $V \rightarrow R$ denotes the data volume used in the registration process. The expression $|x + y|$ 2 represents the physical unit of length for the dimensionless domain $d\alpha$. A graphical illustration of the discrete domain approximation of the DRR is presented in Fig. (2). For DRR synthesis, Vis approximated in the three-dimensional discrete domain of CT Scan using (Siddon, 1985) as follows:

$$E(R) = \|x+y\|_2 \sum_{m=0}^{M} (\alpha_{m+1} - \alpha_m) V \left[s + \frac{\alpha_m + \alpha_{m+1}}{2} (p - s) \right] \quad (2)$$

where, αm defines the parameterization of locations where the ray R intersects the orthogonal planes comprising the CT volume, and M denotes the total number of these intersections (Van der Bom *et al*., 2011). It is important to note that this approximation does not consider the reflection and scattering patterns that usually occur in actual particle interaction in X-ray systems.

Siddon with GPU Acceleration

The Siddon method is an efficient and precise algorithm for computing the X-ray attenuation along multiple rays passing through a 3D volume to create a DRR. The method employs a ray-casting approach, where rays are projected through the 3D volume in a way that simulates the path of X-rays. The 3D volume is divided into tiny cubic elements called voxels. The algorithm calculates the contribution of each voxel along the path of each ray. This contribution represents the attenuation of X-rays as they pass through the voxel. The algorithm accumulates these contributions along the entire path of each ray. As the rays pass through different voxels, the Siddon method keeps track of the accumulated attenuation values. The accumulated values are then used to create a 2D projection image, which simulates what an X-ray image would look like if it were captured from a particular viewpoint.

The Siddon method stands out for its computational efficiency, primarily attributed to its selective computation approach, which calculates interactions between rays and voxels only at their intersections. This targeted computation significantly reduces the computational load and accelerates the process of generating Digitally Reconstructed Radiographs (DRRs). Particularly beneficial for real-time or near-real-time applications like intraoperative imaging and image-guided surgery, this method ensures swift access to DRRs, which is critical for guiding surgical procedures effectively.

A more efficient variant of the Siddon method enhances performance by employing an iterative approach to identify successive intersecting planes (Alvarez-Gomez *et al*., 2021). This algorithm iteratively adjusts the alpha parameter until it reaches the boundaries of the CT volume, thereby improving memory efficiency by significantly reducing the necessity to store intermediate values during processing. The modified approach of the Siddon-Jacobs method is implemented in accelerated environments using Compute Unified Device Architecture (CUDA) and mid-level programming languages, enabling the leverage of the advantages associated with multi-threaded, GPU-accelerated DRR image generators. These implementations effectively exploit data parallelism by assigning distinct threads to independently trace rays that intersect with the detector plane (De Greef *et al*., 2009; Mori *et al*., 2009; Ruijters *et al*., 2008).

The enhancement of the traditional Siddon method is achieved by converting the computations into a comprehensive series of tensor operations. This vectorized approach utilizes advanced GPU compilers and memory allocators that are specifically designed for optimizing large-scale deep learning architectures and computationally intensive tasks, leading to improved overall performance and efficiency during image generation.

Auto Differentiable Image

In our objective to create differentiable DRRs, we undertook a redefinition of the positional parameters within the C-arm system to conform them with the requirements of image registration. Our approach is to envision the C-arms acquisition space as a spherical volume, where both the X-ray source's position (s) and the center of the detector plane are assumed to be perfectly aligned diametrically, with the detector plane tangent of the sphere.

Within this novel framework, the 3D object's position became a subject of six degrees of freedom, as can be seen in Fig. (3). These encompassed a scaling factor ρ that effectively represented the sphere's radius, three rotational degrees of freedom (θ, φ, γ), and three translational degrees of freedom (bx, by, bz). To provide more clarity, in spherical coordinates, the position of the

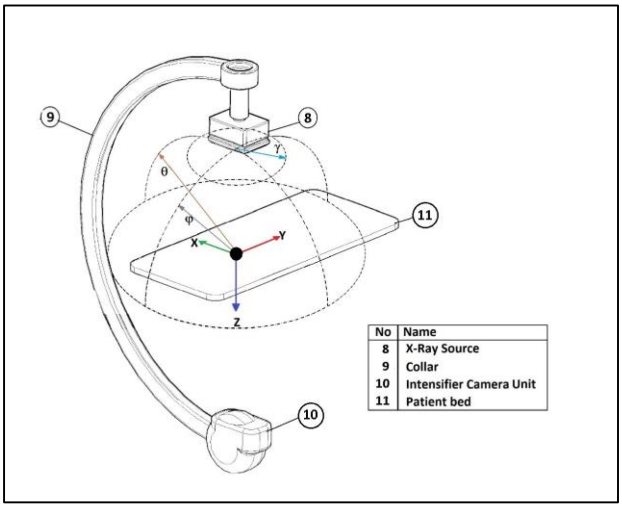

Fig. 3: The parameter diagram illustrates the six Degrees of Freedom (6 DoF) parameters utilized in our registration method for reconstructing auto-differentiable DRRs. The large dot at the center represents the radiographic center of the imaged volume. Our goal is to calculate the displacements in the x, y, and z axes between the volumetric center and the fluoroscopic isocenter (cm), along with the cranial/caudal (θ), RAO/LAO (φ), and receiver's rotational (γ) angles.

X-ray source (s) was expressed as (θ, φ, γ). Here, ρ denoted half the distance between the source and detector, while φ and γ referred to the azimuthal and polar angles, respectively.

In practice, we carried out each stage of the process, which encompassed generating pixels on the detector plane and calculating pixel intensities, within the PyTorch tensor framework. This decision was driven by PyTorch's strong capabilities in automatic differentiation (autograd), enabling us to compute the gradient of our loss function in relation to the DRR. This essential feature allowed us to tackle the primary challenge of image registration, optimizing our registration approach for any differentiable loss function and set of parameters. Since every aspect of our pipeline-from generating pixels on the detector plane to computing pixel intensities-is practically carried out within the PyTorch tensor framework, the resulting DRRs images are then differentiable with respect to the parameters of CT scan position in the spaces (Margossian, 2019; Baydin *et al*., 2018).

Registration

In Computer Assisted Medical Intervention, intra-operative execution time is crucial because of its association with potential patient infections, cumulative X-ray exposure, surgical bleeding risks, and the amount of contrast agents to be injected. Considering these critical factors, the entire pipeline involved in image-guided surgery must be conducted swiftly while minimizing manual interventions in the operating room. With these considerations in mind, this study introduces an intensity-based registration technique, eliminating the need for feature extraction in both pre-operative and intra-operative data. Some methods for enhancing anatomical features with specific geometries may still be acceptable if performed in a pre-operative manner.

To address the 2/3D registration problem, we assume that a 3D dataset will register well with intra-operative 2D data when the projections of the 3D data obtained through DRR are closely similar to the actual intra-operative 2D data. The precise position of the 3D data needs to be iteratively determined by means of the procedure depicted in Fig. (4). For each unregistered position, the method reconstructs a 2D image using DRR and compares it to the intra-operative 2D image to measure a similarity score. Subsequently, an optimization process is performed to estimate a new position with a better similarity score. At this new position, the process is repeated by generating a DRR image and comparing it again to the intra-operative one. This iterative process results in progressively smaller similarity scores, and the process terminates when convergence is achieved.

Similarity Metrics

A major hurdle in intensity-based registration is the calculation of derivatives for the registration loss, which renders gradient-based optimization impractical in numerous applications (Van der Bom *et al*., 2011). Although deep learning methods have demonstrated the capability to achieve high-accuracy registration between X-ray and CT images (Hou *et al*., 2017), they typically require a large amount of training data, which is often unfeasible for certain interventional contexts. Consequently, many applications turn to iterative gradient-free approaches, such as the Nelder-Mead method (Nelder and Mead, 1965) and the Powell-Brent.

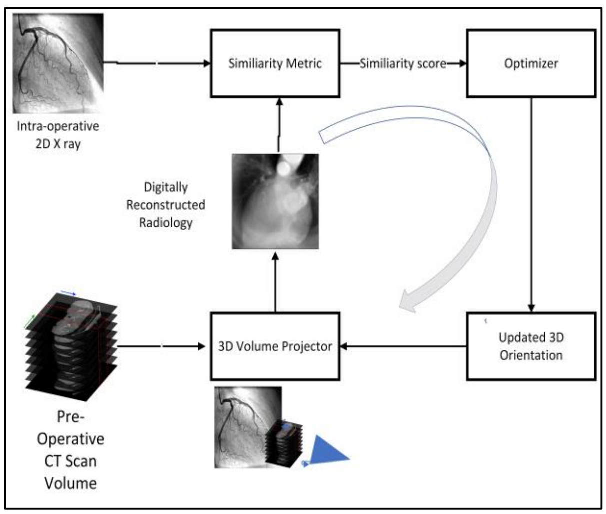

Fig. 4: Registration cycle process

method (Powell, 1964), to optimize image similarity metrics while taking DRR generator parameters into account. These techniques are particularly effective at optimizing highly nonlinear loss functions. However, our research indicates that the loss landscapes, especially in 2/3D registration, tend to be convex in a significant area surrounding the optimal solution. This convexity renders gradient-based optimization methods less appropriate.

To address this challenge, the choice of the loss function is critical to achieving a concave loss landscape with a global optimum at the perfectly registered position. In this study, we evaluated a statistical-based loss function, specifically the Normalized Cross-Correlation between pre-operative and intra-operative data. The landscape of this loss function is visualized by computing the normalized cross-correlation (Mohammadi and Keyvanpour, 2021) between CT projection data at the initial position and other images that are shifted and rotated along all axes. These values are then plotted as a three-dimensional curve as shown in Fig. (5).

In the loss function graph, it's clear that the landscape curve forms elongated valley-shaped regions, especially when rotation is applied around the Y-axis. This phenomenon is attributed to the cranio-caudal orientation of the skull, which resembles a cylindrical shape, leading to the presence of multiple local minima during rotations around that axis.

With this plot as our basis, we have a strong conviction that Gradient Descent is a viable approach for optimizing the registration problem we are tackling.

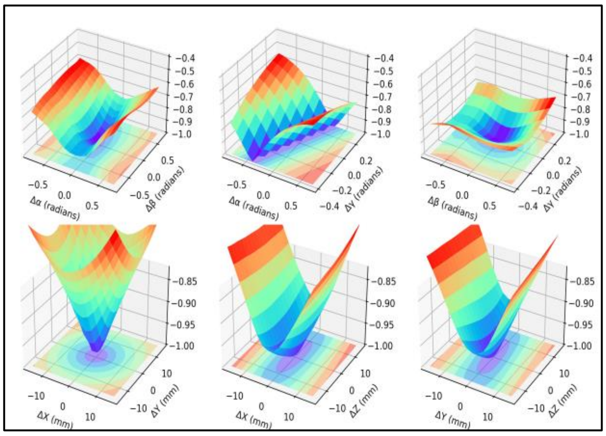

Fig. 5: Loss function landscape

Optimization

Gradient descent is a widely used optimization algorithm in image registration for finding the minimum of a loss function that reflects dissimilarity between the two data sets. It iteratively updates the CT Scan position parameters in the direction that reduces the loss, making it a fundamental building block for image registration applications. We are starting our discussion with the basic gradient descent and then discuss its available extensions: Gradient descent with momentum, gradient descent with damped momentum, and the adam optimizer.

Gradient descent adjusts model parameters by subtracting the gradient of the loss function relative to those parameters. This procedure is carried out repeatedly until convergence is reached or a specified number of iterations is completed. During each iteration, the gradient of the loss with respect to the model parameters is computed, and the parameters are updated by subtracting the scaled gradient, controlled by a learning rate (α). Mathematically, the parameter update step θ at iteration t can be expressed as:

$$\theta^{(t+1)} = \theta^{(t)} - \alpha \nabla J(\theta^{(t)}) \tag{3}$$

where, θ (t) represents the parameters at iteration t. α is the learning rate, a hyperparameter that controls the step size.

However, vanilla gradient descent has some issues, such as slow convergence and oscillations in the shallow attracting basin during the parameter updates.

Gradient descent with momentum. momentum is an extension of gradient descent that addresses the slow convergence problem. It introduces a momentum term (usually denoted as β) to the update rule by adding the β factor to the previous velocity and subtracting the gradient scaled by the learning rate. It also accumulates a fraction of the previous gradients to smooth out the parameter updates:

$$\begin{aligned} v(t+1) &= \beta \cdot v(t) + (1-\beta)\nabla J(\theta(t)) \\ \theta(t+1) &= \theta(t) - \alpha \cdot v(t+1) \end{aligned} \tag{4}$$

where, $v(t)$ represents the velocity of the parameters at iteration t. β is the momentum coefficient. $\theta(t)$ represents the parameters at iteration t. α is the learning rate, a hyperparameter that controls the step size.

The introduction of momentum terms helps dampen oscillations and accelerates convergence along flat or shallow regions of the loss landscape.

Gradient descent with damped momentum is an improvement over standard momentum, where the velocity is damped to avoid overshooting the minimum. It introduces an additional parameter γ to control the damping. The update rule becomes:

$$\begin{aligned} v(t+1) &= \gamma \cdot v(t) + (1-\gamma)\nabla J(\theta(t)) \\ \theta(t+1) &= \theta(t) - \alpha \cdot v(t+1) \end{aligned} \tag{5}$$

where, $v(t)$ represents the velocity of the parameters at iteration t. γ is the damping coefficient. $\theta(t)$ represents the parameters at iteration t. α is the learning rate, a hyperparameter that controls the step size.

Damping the velocity helps stabilize the optimization process, especially when dealing with noisy or ill-conditioned gradients.

Adam (Adaptive Moment Estimation) adam (Kingma and Ba, 2017) is an adaptive optimization algorithm that combines ideas from both Momentum and Root Mean Square Propagation. It maintains two moving averages, the first moment (mean) of the gradients (similar to momentum) and the second moment (uncentered variance) of the gradients (similar to RMS prop).

Adam has two hyperparameters, $\beta 1$, and $\beta 2$, which control the decay rates of the first and second moments. It also has a small ε term added to prevent division by zero.

Mathematically, the update step in Adam is:

$$m(t+1) = \beta_1 \cdot m(t) + (1-\beta_1)\nabla J(\theta(t))$$
$$v(t+1) = \beta_2 \cdot v(t) + (1-\beta_2)(\nabla J(\theta(t))) \quad (6)$$
$$\hat{m} = \frac{m(t+1)}{1-\beta_1^{t+1}} \hat{v} = \frac{v(t+1)}{1-\beta_2^{t+1}} \theta(t+1) = \theta(t) - \alpha \frac{\hat{m}}{\sqrt{\hat{v}+\varepsilon}}$$

where, $m(t)$ represents the first moment (mean) of the gradients at iteration t. $v(t)$ represents the second moment (uncentered variance) of the gradients at iteration t $\beta 1$ and $\beta 2$ are the exponential decay rates for the first and second moments, respectively. $\hat{}m$ and $\hat{}v$ are bias-corrected estimates of the first and second moments and ε is a small constant added to prevent division by zero.

Adam customizes the learning rates for each parameter on an individual basis, which makes it highly effective for various optimization tasks.

In summary, gradient descent and its extensions are the foundation of many optimization algorithms to overcome its limitations by introducing concepts of velocity, damping, and adaptive learning rates to improve convergence and stability during optimization. These algorithms are widely applied in the training of deep neural networks and optimizing complex loss functions.

Results and Discussion

We evaluated the implementation of the proposed algorithm using CT scans of a plastic skull model as objects. Our assessment primarily centered on the algorithm's performance, quantifying the time needed to generate Digitally Reconstructed Radiography (DRR) images from the dataset. Furthermore, we illustrated the practical utility of these DRR reconstructions for image registration by employing the method in various gradient-based registration techniques. Finally, we assessed the robustness of our registration method using Monte Carlo simulations.

To ensure that our work is accessible and reproducible by others, we performed this evaluation in the Google Colab environment, utilizing a machine equipped with a T4 GPU.

DRR Generation Evaluation

The reconstructed DRR data set is presented in Fig. (6), we need to provide parameters of image size, and resolution as well as CTscan position and orientation. For source-to-sensor distance, this parameter is determined by the C arm calibration process.

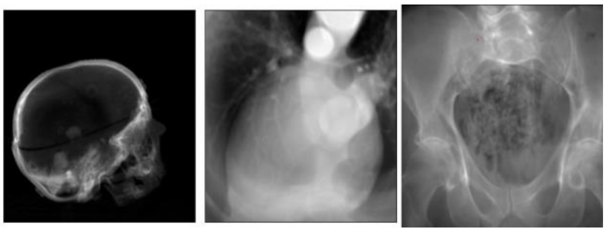

Fig. 6: DDR results in different objects, plastic phantoms, heart, and pelvis

Table 1: DRR execution time for different image sizes and patches. Each metric is averaged over 15 runs

DRR Size in pixel	Time (in ms)	Patch size
100×100	21.2±2.86	
200×200	70.1±3.87	
300×300	150.0±0.94	
400×400	265.0±0.92	
500×500	393.0±55.3	
600×600	600.0±125	
600×600	320.0±1.01	150×150
900×900	601.0±6.12	150×150
1200×1200	1140.0±23.5	150×150
1500×1500	1700.0±24.1	150×150
1800×1800	1490.0±26.4	150×150
1024×1024	7450.0±647	16×16
1024×1024	1970.0±177	32×32
1024×1024	1080.0±52.5	64×64
1024×1024	859.0±218	128×128
1024×1024	881.0±396	256×256
1024×1024	1230.0±112	512×512

We observe the execution time to reconstruct DRR images for different sizes as presented in Table (1).

As the size of the Digital Reconstructed Radiograph (DRR) increases, the reconstruction time also experiences a linear increase. This correlation comes from the fact that larger images demand more computational resources and time. However, due to the limited memory of GPUs, it becomes necessary to reconstruct images using a patching approach. While it's feasible to compute all the rays in a DRR simultaneously on the GPU, as DRRs grow in size, they can quickly exceed the available memory, resulting in CUD A memory errors. Consequently, patching becomes crucial, ensuring efficient utilization of GPU memory.

Leveraging GPU capabilities, the reconstruction of CT projection images into 2D images is an expedited process, particularly for relatively small image sizes (greater than 600×600 pixels). However, when dealing with higher-resolution reconstructions, the GPU's memory constraints necessitate a patching process.

Throughout the development of this method, we actively sought feedback from radiologists and medical professionals. They expressed concerns about the image quality, as the reconstructed images did not closely resemble the original radiographic images. This issue arises because the reconstruction method, based on DRR, simplifies physical phenomena by not accounting for effects like scattering and reflection. These simplifications, although efficient, cannot replicate realistic X-rays. Nevertheless, the model outlined in Eq. (2) has a proven track record in slice-to-volume registration (Van der Bom *et al*., 2011), as demonstrated also in this study for image registration purposes.

Registration Evaluation

In our registration test, we implemented our approach to register 2/3D data. This involved generating fixed DRR data at a specified initial position, denoted as $\eta0(\theta, \varphi, \gamma, bx$, by, $bz)$, and moving DRR images from random positions. The registration was achieved through the optimization of a loss function, as described in section 2.2. To minimize the cost function between the pre-determined DRR position and those in the varied position, we employed various gradient descent-based methods, each aimed at finding an optimal rigid transformation denoted as $\eta*$.

The table reveals that vanilla gradient descent struggled to converge for our phantom. This lack of convergence was primarily attributed to oscillations observed when employing the Gradient Descent optimizer, particularly when the solution was in close proximity, as depicted in Fig. (7). Nevertheless, in general, Gradient Descent proves effective in optimizing the cost function when the starting position is sufficiently close to the optimal value, and the attraction basin is not overly shallow around the optimum position. The enhancements introduced in the gradient descent method, as presented in section 2.2, improved the ability to find the optimal position while significantly reducing the execution time of the optimization process, as demonstrated in Table (2).

Given the successful execution of the method based on autograd reconstruction in DRR-based registration combined with gradient-based registration, we are confident in the applicability of this registration pipeline in the operating room.

A study assessing the robustness of the registration involved initializing positions at perturbed random locations (45-degree rotation and 60 mm translation along the axes) in 1000 instances. Subsequently, the registration method with the Adam optimization scheme was applied. This choice was made due to its higher convergence rate, even though its execution time is slightly longer compared to other optimization schemes. The convergence graph, as illustrated in Fig. (8), demonstrates that the proposed optimization successfully brings the majority of the initial positions to convergence points.

In the context of registration, achieving desired results is more feasible when initial positions are already reasonably close to the global optimum loss value. Estimating such initial positions can be accomplished by selecting pairs of anatomical features in both 2 and 3D data and performing an initial registration, similar to methods described in (Arun *et al*., 1987). This approach has the potential to reduce the unconvergence rate of optimization time and simultaneously avoid local optimum positions.

Table 2: Performance of optimization methods in registering CT scans from arbitrary positions to their projections. Experimentation involved 300 random positions and the method is considered successfully optimizing when the loss function value is below 10-6. GD is for Gradient Descent, GDM is for GD with momentum, GDDM is for GD with dampened momentum

Optimization	Convergence rate	Time (in ms)
GD	Not-convergence	-
GDM	66.8%	15.56
GDDM	75.6%	6.96
Adam	85.7%	7.00

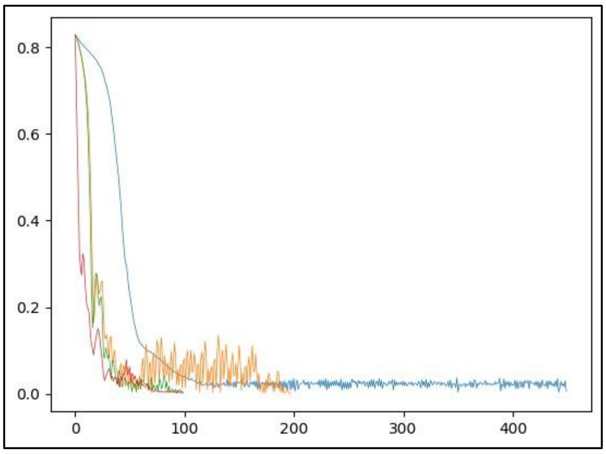

Fig. 7: Convergence of various optimization methods. Gradient descent often fails to find the optimal solution due to oscillations (blue line) around the minimum position, but the proposed improvements can effectively overcome this issue

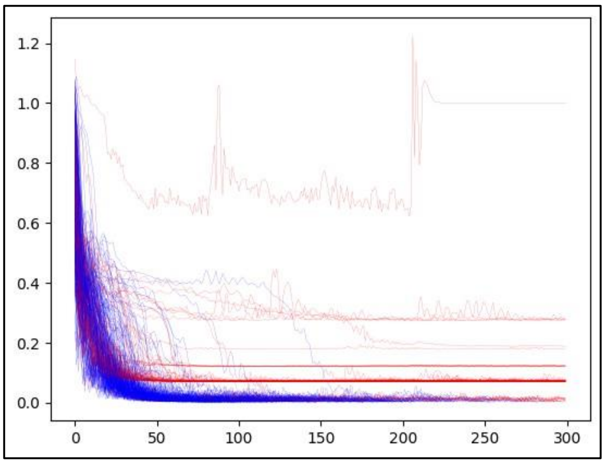

Fig. 8: The convergence graph displays results from diverse initial positions, with successful registrations represented by the blue line and unsuccessful ones by the red line

DRR-Based Gradient Descent 2/3D Registration

Our method utilizing the auto-differentiable DRR generator encompasses a comprehensive 2-3D image registration technique facilitated by synthetic DRRs. Initially, we generate a fixed DRR by employing a specific set of ground truth parameters, denoted as $\eta* = (\theta, \phi, \gamma, bx, by, bz)$, which represent key attributes such as rotational angles and translational displacements. Subsequently, we create a moving DRR utilizing a separate set of randomly initialized parameters $\eta 0$, allowing for a diverse range of positional and orientational adjustments.

To effectively address the registration challenge, we optimize the loss function by employing gradient descent methods. These methods aim to minimize the negative Zero-Normalized Cross-Correlation (ZNCC) between the fixed DRR image and the moving images obtained from real-time CT scan projections. By focusing on the correlation between these images, we enhance the registration accuracy, allowing for improved alignment of the synthetic DRRs with the actual scanned data. This systematic approach not only facilitates precise image registration but also leverages the advantages of auto-differentiable techniques to ensure efficient computations throughout the optimization process.

Employing gradient-based optimization necessitates a convex loss function in the vicinity of the optimal position to facilitate effective convergence. To explore this essential property, we conduct a comprehensive simulation study that examines and illustrates the convexity of the loss landscape surrounding the ground truth parameters $\eta*$. In our simulations, moving DRRs are generated by systematically sampling both rotational and translational displacements, with parameters uniformly drawn from specified ranges that encompass the ground truth parameters η∗. While our generator can produce gradients for additional model parameters such as the source-to-detector distance (2ρ) and the dimensions and pixel spacing of the DRR (H, W, Δx, Δy) we specifically concentrate our analysis on the challenge posed by these six Degrees of Freedom (DoF) registration problem.

Our findings reveal that the negative Zero-Normalized Cross-Correlation (ZNCC) exhibits local convexity, as demonstrated in Fig. (5). This characteristic strongly supports the suitability of negative ZNCC as an effective loss function for optimization through gradient descent methodologies. Furthermore, similar convex loss landscapes are observed when evaluating performance with the L2 norm, underscoring the robustness of our approach and its potential applicability in various imaging registration contexts, thereby enhancing the accuracy and reliability of the registration process.

Differentiable DRR Registration Converges Quickly

The optimization procedure involves adjusting the parameters of the moving DRR using gradient descent. We applied distinct update rates for the rotational and translational parameters due to the variations in their units, in addition to incorporating a momentum factor. To investigate potential local minima, we broadened the range of initializations beyond a convex neighborhood. Each DRR underwent 250 iterations of gradient descent, with convergence defined as achieving a negative ZNCC value below a certain threshold. Out of 1,000 random initializations, 745 succeeded in converging, resolving the registration issue in roughly 65 iterations. Visualizations

of the optimization steps indicate successful recovery of the true pose parameters from challenging initial configurations, although some instances did not converge due to local minima. Our primary goal is to develop an efficient DRR synthesis method to facilitate a variety of downstream optimization tasks, including registration.

The three-dimensional CT scan data collected for this study was obtained using a plastic skull phantom, with x-ray acquisition parameters set to 146 mA and 120 kV. Preprocessing is not required during the described 2D/3D registration process.

We have tested this registration method and found that it is quite robust to the initial positions of both CT and X-ray data. This method is also sufficiently fast to be feasible for clinical application in the operating room. This is further supported by the fact that preprocessing is not required during the procedure, which does not disrupt the workflow in the operating room. We received feedback from radiologists regarding the quality of the generated DRRs, which, although fairly good, often lack realism. This might be improved by integrating refraction and scattering phenomena during the DRR reconstruction process. We will apply this registration method to cardiac surgery applications, particularly in percutaneous coronary intervention (PCI).

Conclusion

In this study, we introduce a fast automatic intensity-based registration method utilizing auto-differentiable DRR image generation in parallel computation. These images are then integrated into registration process with a pure gradient scheme to optimize the image loss function. Our approach addresses previous challenges in 2D/3D registration, which typically relied on finite difference methods or other heuristic optimization strategies.

This method is sufficiently fast to be feasible for clinical application in the operating room. In future research, we will apply this registration method to cardiac surgery applications, particularly in percutaneous coronary intervention (PCI).

Acknowledgment

Thank you to the publisher for their support in the publication of this research article. We are grateful for the resources and platform provided by the publisher, which have enabled us to share our findings with a wider audience. We appreciate the efforts of the editorial team in reviewing and editing our work, and we are thankful for the opportunity to contribute to the field of research through this publication.

Funding Information

We would like to thank National Research and Innovation Agency Lembaga Pengelola Dana Pendidikan (LPDP) Ministry of Finance of the Republic of Indonesia based on the Decree of the main director of the education fund management institution number: KEP-6/LPDP/LPDP.4/2023 dated 13 June 2023 which has provided support for this research through Riset Dan Inovasi untuk Indonesia Maju Batch III with the agreement/contract number 70/ IV/KS/05/2023 and number 1/2/04.04/1114/05/2023.

Author's Contributions

Agung Alfiansyah: Conception, design, implementation, data analysis, data interpretation, and writing.

Izzati Muhimmah: Conception, design, analysis, data interpretation, and writing.

Arif Sasongko: Design, analysis, data interpretation, and writing.

Heru Sulastomo: Data acquisition, analysis, and writing.

Nanang Wiyono: Data acquisition, data interpretation, and writing.

Ethics

The authors hereby declare that this manuscript is original and has not been submitted for publication in any other journal. Co-authors have reviewed and approved the final version of the manuscript for submission. The research presented does not involve any ethical concerns. The authors declare no conflicts of interest that could potentially influence the work presented in this study.

References

Alvarez-Gomez, J. C., Jimenez Palavicini, G., Roth, H., & Wahrburg, J. (2021). An Ultra-Fast Digitally Reconstructed Radiograph (DRR) Implementation of the Siddon-Jacobs Algorithm Using Parallel Computing: Runtime Improvement of an Intensity-Based 2D/3D Registration. *Current Directions in Biomedical Engineering*, *7*(2), 25–28. https://doi.org/10.1515/cdbme-2021-2007

Arun, K. S., Huang, T. S., & Blostein, S. D. (1987). Least-Squares Fitting of Two 3-D Point Sets. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *PAMI-9*(5), 698–700. https://doi.org/10.1109/tpami.1987.4767965

Baydin, A. G., Pearlmutter, B. A., Radul, A. A., & Siskind, J. M. (2018). Automatic Differentiation in Machine Learning: A Survey. *Journal of Machine Learning Research*, *18*(153), 1–43.

de Greef, M., Crezee, J., van Eijk, J. C., Pool, R., & Bel, A. (2009). Accelerated Ray Tracing for Radiotherapy Dose Calculations on a GPU. *Medical Physics*, *36*(9Part1), 4095–4102. https://doi.org/10.1118/1.3190156

Hou, B., Alansary, A., McDonagh, S., Davidson, A., Rutherford, M., Hajnal, J. V., Rueckert, D., Glocker, B., & Kainz, B. (2017). Predicting Slice-to-Volume Transformation in Presence of Arbitrary Subject Motion. *Springer International Publishing*, 296–304. https://doi.org/10.1007/978-3-319-66185-8_34

Kingma, D. P., & Jimmy, L. B. (2017). Adam: A Method for Stochastic Optimization. *ArXiv*, arXiv:1412.6980. https://doi.org/10.48550/arXiv.1412.6980

Margossian, C. C. (2019). A Review of Automatic Differentiation and Its Efficient Implementation. *WIREs Data Mining and Knowledge Discovery*, *9*(4), e1305. https://doi.org/10.1002/widm.1305

Mohammadi, Z., & Keyvanpour, M. R. (2021). Similarity Measures in Medical Image Registration a Review Article. *2021 12th International Conference on Information and Knowledge Technology (IKT)*, 89–95. https://doi.org/10.1109/ikt54664.2021.9685453

Mori, S., Kobayashi, M., Kumagai, M., & Minohara, S. (2009). Development of a GPU-Based Multithreaded Software Application to Calculate Digitally Reconstructed Radiographs for Radiotherapy. *Radiological Physics and Technology*, *2*(1), 40–45. https://doi.org/10.1007/s12194-008-0040-3

Nelder, J. A., & Mead, R. (1965). A Simplex Method for Function Minimization. *The Computer Journal*, *7*(4), 308–313. https://doi.org/10.1093/comjnl/7.4.308

Powell, M. J. D. (1964). An Efficient Method for Finding the Minimum of a Function of Several Variables Without Calculating Derivatives. *The Computer Journal*, *7*(2), 155–162. https://doi.org/10.1093/comjnl/7.2.155

Ruijters, D., ter Haar Romeny, B. M., & Suetens, P. (2008). Gpu-Accelerated Digitally Reconstructed Radiographs. *BioMED*, *8*, 431–435.

Siddon, R. L. (1985). Fast Calculation of the Exact Radiological Path for a Three-Dimensional CT Array. *Medical Physics*, *12*(2), 252–255. https://doi.org/10.1118/1.595715

Van Der Bom, I. M. J., Klein, S., Staring, M., Homan, R., Bartels, L. W., & Pluim, J. P. W. (2011). Evaluation of Optimization Methods for Intensity-Based 2D-3D Registration in X-Ray Guided Interventions. *Image Processing*, 796223. https://doi.org/10.1117/12.877655