Fachhochschule Aachen
Campus Jülich

Comparison of algorithms

for image registration.

Seminar thesis

Astrid Warkentin

Matr. No.: 3568224

Faculty of

Medical Engineering and Applied Mathematics

Study program Angewandte Mathematik und Informatik B.Sc.

Aachen, November 2024

Diese Arbeit ist von mir selbstständig angefertigt und verfasst. Es sind keine anderen als die angegebenen Quellen und Hilfsmittel benutzt worden.

My name ____________________

signature

Diese Arbeit wurde betreut von:

1. Prüfer:	Prof. Dr. Martin Reißel
2. Prüfer:	M.Sc. Nico Blaß

Contents

1 Introduction

As soon as you have two images, it can be useful to combine the information from those images into one. The process with which such a combination is made is called image registration. Be it military, satellite, 3D reconstruction, tracking systems or medical applications - image registration is used in a wide variety of disciplines. It has become an essential part of many image processing applications [1, p. 5–12, p. 2]. Image registration is a crucial step in image analysis tasks in which the final information is gained from the combination of various data sources like in image fusion, change detection and multichannel image restoration. Due to the different image conditions (e.g. different angles, sensors, distortion and time differences) that can arise where image registration is needed, it is usually tailored to a given application and no universal way to register images exists.

> Typically, registration is required in remote sensing (multispectral classification, environmental monitoring, change detection, image mosaicking, weather forecasting, creating super-resolution images, integrating information into geographic information systems (GIS)), in medicine (combining computer tomography (CT) and NMR data to obtain more complete information about the patient, monitoring tumour growth, treatment verification, comparison of the patient's data with anatomical atlases), in cartography (map updating), and in computer vision (target localization, automatic quality control)[...]. [2, p. 977]

More applications of image registration include the tracking of objects, the monitoring of the amount of people in a building such as a shopping mall, or to track traffic in driver assistance systems for vehicles to lower the amount of accidents, as well as in robotics, to allow robots to move and perform operations in unknown buildings and passages. Medical applications include combining medical images, monitoring the growth of tumours, comparing a patient's anatomy with an atlas and to help in epilepsy surgery [1, p. 7-9], as well as the contactless monitoring of newborns.

1.1 Motivation and Scope

> *"[Image registration] is considered one of the most complex and challenging problems in image analysis with no single registration algorithm to be suitable for all the related applications due to the extreme diversity and variety of scenes and scenarios*. [1, p. 1]"

Even though image registration is so widely used, it can be difficult to pick out suited approaches and algorithms for a given application. The amount of algorithms is vast and a lot of them are similar but different still - this seminar thesis aims to explore common algorithms, summarize their strengths and weaknesses and provide the reader with an idea as to which ones might be best suited for their problem and which ones to avoid. This seminar thesis will only explore a small amount of the available methods for image registration and won't cover preprocessing techniques. The algorithms and approaches basic principles and functionality are covered as well as some information on performance and their usual application and suitability.

1.2 Structure

The 2nd chapter will cover basic aspects of image registration and provide insight into the general process as well as provide basic information on why homogenous coordinates are used. The 3rd chapter will explain concrete algorithms, their applications and performance. The 4th chapter aims to provide the reader with an outlook towards the use of other algorithms and the use of neural networks for image registration.

2 Image Registration Fundamentals

> *"Image registration is the process of determining the point-by-point correspondence between two images of a scene."* [3, p. 1]

This chapter provides an overview of the theoretical aspects of image registration. In the process of aligning two images, it is important to consider the information needed for registration, such as the image's content, perspective differences, distortion, noise, scene changes, and modality. There are several approaches to aligning two images, but the underlying principle is minimizing a measure of registration by transforming one image onto the other. This measure can vary, as long as the results are acceptable, and should be considered when balancing accuracy against performance.

2.1 Terminology

The terminology used in this thesis is adapted from the book "2-D and 3-D image registration for medical, remote sensing, and industrial applications" [3].

- **Reference Image / Source Image:** The image used as a reference to fit a target image to.
- **Sensed Image / Target Image:** The image transformed to fit the source image.
- **Transformation Function:** The function that transforms the target image to fit the source image.

2.2 Registration Approaches

There are both feature-based and intensity-based registration approaches. In feature-based registration, unique and easily recognizable features are found, and alignment is achieved by matching points and curves extracted from these features. Intensity-based registration is a flexible way to register images, as it utilizes all available information for registration. Instead of using edges or points as in feature-based registration, pixel intensity patterns are used for alignment.

2.3 Summary of a General Image Registration Process

Image registration is not a strictly defined or standardized process; however, there are typical steps usually involved, which may vary based on the application. These steps generally include preprocessing, feature selection and matching or optimization-based alignment, followed by transformation and resampling.

1. **Preprocessing**
   Images are usually preprocessed to prepare them for the chosen registration method. This step is essential for many methods and often includes noise reduction, segmentation, and smoothing. The goal is to optimize the images for the registration method by highlighting relevant information and reducing noise or irrelevant data that can degrade accuracy.

2. **Feature-Based Registration**
   If the registration algorithm relies on features, preprocessing enhances the detectability of these features. The subsequent steps involve:

   - **Feature Selection:** Depending on the images, features such as corners, edges, lines, curves, regions, templates, or patches can be identified. This selection can be done independently for both images or by first extracting features from the source image and then searching for corresponding features in the target image. When features provide significant context or information, like regions and templates, matching features from the source to the target image is preferable. Conversely, when the features contain limited information (e.g., points or simple edges), independent feature extraction from both images is common. [3, p. 4–5]
   - **Correspondence Determination:** For non-point features like regions and curves, pairs or sets of points representing the feature must be identified for both images, as their correspondence is critical for estimating the transformation parameters that align the images.
   - **Transformation Parameter Estimation:** Features with established correspondence are used to calculate the parameters for the transformation function.

3. **Metric-Based Registration (Non-Feature-Based)**
   Algorithms that do not use explicit feature matching optimize a registration metric to align the images. This process involves iteratively transforming one image to maximize a similarity measure, such as mutual information, cross-correlation, or mean squared error.

4. **Transformation and Resampling**
   The final transformation step involves applying the transformation to align the target image with the source image's geometry. Resampling adjusts the pixel values of the target image to fit the transformed grid, ensuring spatial consistency and alignment. This step often employs interpolation techniques to estimate new pixel values for the transformed positions.

2.3.1 Homogeneous Coordinates

Homogeneous coordinates are used in computer graphics as they allow transformations and projections to be represented as matrices by increasing the dimensionality by one. This is useful for image registration because, at the cost of an additional dimension, they can be projected back onto their original dimension. Rotation, scaling, shearing, and translation can all be combined into a single step, whereas in non-homogeneous coordinates, translation, rotation, shearing, and scaling cannot be applied in a single matrix multiplication. For instance, rotating and translating a 2D point requires both a matrix multiplication and addition:

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} * \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} a \\ b \end{bmatrix} \tag{2.1}$$

whereas in homogeneous coordinates, only a single matrix multiplication is needed:

$$\begin{bmatrix} \cos\theta & -\sin\theta & a \\ \sin\theta & \cos\theta & b \\ 0 & 0 & 1 \end{bmatrix} * \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \tag{2.2}$$

Another property of homogeneous points is that scaling a point by a factor does not change its representation as the same point. This is because the extra dimension is used when converting homogeneous coordinates back into Cartesian coordinates:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \xrightarrow{\text{Homogenization}} \begin{pmatrix} \frac{x}{z} \\ \frac{y}{z} \end{pmatrix} \tag{2.3}$$

$$\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} \xrightarrow{\text{Homogenization}} \begin{pmatrix} \frac{1}{1} \\ \frac{2}{1} \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \tag{2.4}$$

$$\begin{pmatrix} 2 \\ 4 \\ 2 \end{pmatrix} \xrightarrow{\text{Homogenization}} \begin{pmatrix} \frac{2}{2} \\ \frac{4}{2} \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \tag{2.5}$$

3 Comparison of Different Algorithms and Their Properties

This chapter focuses on concrete algorithms, their strengths and weaknesses, typical use cases, and performance.

3.1 Intensity-Based Registration

Normalized Cross-Correlation

Normalized Cross-Correlation (NCC) compares intensity patterns by normalizing for differences in brightness and contrast. It is based on the Pearson correlation coefficient, and the formula is:

$$NCC(A,B) = \frac{\sum_{x,y}\left((I_A(x,y) - \overline{I}_A) \cdot (I_B(x,y) - \overline{I}_B)\right)}{\sqrt{\sum_{x,y}\left(I_A(x,y) - \overline{I}_A\right)^2 \cdot \sum_{x,y}\left(I_B(x,y) - \overline{I}_B\right)^2}} \tag{3.1}$$

Here, $I_A(x,y)$ and $I_B(x,y)$ are the intensities at position (x,y), and $\overline{I}_A$ and $\overline{I}_B$ are the mean intensities within the overlapping area. Normalized Cross-Correlation is commonly used in single-modality image registration, and its strength lies in registering images under different lighting conditions [4, p. 161].

3.1.1 Mutual Information

Another metric is Mutual Information (MI), where the Shannon entropy $H(A)$, calculated from the images' grayscale values, is utilized. Images with similar grayscale values have low entropy, while more uniformly distributed grayscale values yield higher entropy. The joint histogram of the grayscale values of both images changes as their alignment changes. When images are correctly registered, the joint histogram produces clusters; when misaligned, the histogram is more diffuse. To align images, their mutual information $I(A,B)$ needs to be maximized, which involves minimizing their joint entropy $H(A,B)$ [5, p. 239–240].

$$H(A) = -\sum_{a} p_A(a) \log_2(p_A(a)) \tag{3.2}$$

$$H(A,B) = -\sum_{x\in X}\sum_{y\in Y} p(x,y) \log_2(p(x,y)) \tag{3.3}$$

$$I(A,B) = H(A) + H(B) - H(A,B) \tag{3.4}$$

Mutual Information is sensitive to noise and can be outperformed by NCC in noisy images, but it is powerful for measuring similarity between templates in multimodality images. It is frequently used in multimodality registration, such as CT (Computer Tomography) to MRI (Magnetic Resonance Imaging). The computational complexity of this measure is linked to the accuracy of the registration and is computationally expensive. For images with grayscale values ranging from 0 to 255, the computational complexity is 256^2+n, where n is the number of pixels. This includes n additions to calculate the joint probability distribution and 256^2 multiplications and logarithm evaluations for the joint entropy [4, p. 165], [3, p. 98].

3.1.2 Sum of Squared Differences

A simpler approach is the Sum of Squared Differences (SSD) between individual pixel intensity values. The formula is:

$$SSD(A,B) = \sum_{x,y} (I_A(x,y) - I_B(x,y))^2 \tag{3.5}$$

SSD is often used in single-modality registration, e.g., MRI to MRI. It is sensitive to noise and intensity differences, as it is more affected by large differences than small ones. Consequently, SSD is inferior to normalized cross-correlation when different lighting conditions exist between images [4, p. 168].

3.1.3 Demons Algorithm

Named after Maxwell's demon due to its analogy with the diffusion process, the Demons algorithm is a deformation transformation algorithm that iteratively increases a measure of registration by applying small, smooth displacements. This algorithm requires an initial affine transformation, independent of the Demons algorithm, to maximize the registration measure. It then estimates a deformation field by matching image gradients to fit the images. A sequence of non-parametric transformations is computed, each one increasing the registration measure. These transformations converge iteratively, but excessive iterations may lead to overfitting and susceptibility to noise [1, p. 191–193]. The Demons algorithm is computationally efficient and well-suited for medical imaging tasks where anatomical structures must be aligned. However, it is sensitive to large deformations and noise, often requiring additional preprocessing or regularization to improve accuracy.

3.2 Feature-Based Algorithms

3.2.1 Harris Corner Detector

The Harris Corner Detector is an intuitive way to detect corners and edges. Within a small window of the image, there can be either a flat area, an edge, or a corner. By calculating derivatives over pixel intensities in the x and y directions, edges produce high derivative values. For each window (a small part of the image), the structure tensor is computed from the gradients in the window, and its eigenvalues are calculated. Together with a threshold,

(a) *Flat Region* (b) *Edge* (c) *Corner*

Figure 1: *Possible image window regions*

these eigenvalues determine whether the region is flat, an edge, or a corner. The Harris Corner Detector is effective for detecting corners and edges in feature-based registration but is sensitive to noise. Preprocessing steps, such as smoothing, are often necessary to ensure robust performance [4, p. 60]. It is widely used in mono-modal image registration, where feature stability is crucial. For an $n \times n$ pixel window in an $M \times N$ image, the computational complexity involves MNn^2 multiplications, and a square matrix must be calculated at each pixel [4, p. 62].

3.2.2 Laplacian of Gaussian

The Laplacian of Gaussian (LoG) detector is a common blob detector used to find the centers of dark and bright spots. The algorithm calculates the Laplacian, $I_{xx}(x,y) + I_{yy}(x,y)$, of a grayscale image $I(x,y)$. The Laplacian, representing the sum of second-order derivatives, highlights edges and blobs. Applying the Laplacian to a Gaussian-blurred image reduces sensitivity to small changes and noise, highlighting blobs of different sizes based on the Gaussian blur radius [4, p. 51–53]. LoG is effective for identifying regions of interest, making it suitable for applications requiring precise feature detection. It is computationally expensive, with complexity on the order of MNn^2 multiplications for an $M \times N$ image [4, p. 51–53].

3.2.3 Iterative Closest Point

The Iterative Closest Point (ICP) algorithm is a versatile algorithm widely used for aligning n-dimensional point clouds, including 2D-2D image registration. Given two point sets, S_1 and S_2, where S_2 is transformed to align with S_1, the steps are:

1. Apply an initial transformation as an estimated registration.

2. For each point in S_2, find the closest point in S_1.

3. Compute a rigid transformation to best align each point in S_2 with its corresponding closest point in S_1.
4. Transform S_2 using the computed transformation.
5. Repeat until a predefined accuracy threshold or the maximum number of iterations is reached.

While ICP is adaptable, it requires an initial transformation estimate to converge to the correct solution. Without it, the algorithm may converge to a local minimum [1, p. 89]. The computational cost depends on the number of points and the closeness measure used [1, p. 164].

4 Outlook

The algorithms and principles introduced in this seminar thesis represent only a small subset of the vast and diverse methods available for image registration. Numerous other algorithms exist, including those that rely on patterns and templates for registering images, which are widely utilized across various applications.
A steadily growing and highly dynamic area within this field is the application of neural networks. These approaches have demonstrated impressive results in image registration, and the number of research papers published on this topic continues to increase each year.

References

[1] ARGYRIOU, Vasileios: *Image, video & 3D data registration: Medical, satellite & video processing applications with quality metrics*. Chichester West Sussex United Kingdom : Wiley, 2015. – ISBN 9781118702468

[2] ZITOVÁ, Barbara ; FLUSSER, Jan: Image registration methods: a survey. In: *Image and Vision Computing* 21 (2003), No. 11, p. 977–1000. `http://dx.doi.org/10.1016/S0262-8856(03)00137-9`. – DOI 10.1016/S0262–8856(03)00137–9. – ISSN 0262–8856

[3] GOSHTASBY, Ardeshir: *2-D and 3-D image registration for medical, remote sensing, and industrial applications*. Hoboken NJ : J. Wiley & Sons, 2005. – ISBN 0471649546

[4] GOSHTASBY, Ardeshir: *Theory and applications of image registration*. Hoboken NJ : John Wiley & Sons, 2017. – ISBN 9781119171720

[5] LEE, Myung-Eun ; KIM, Soo-Hyung ; SEO, In-Hye: Intensity-based registration of medical images. In: *2009 International Conference on Test and Measurement*, IEEE, 2009. – ISBN 978–1–4244–4699–5, p. 239–242