8DC00 Medical Image Analysis

Project 1 - Registration

Report

1-10-2021
Group 3
Aiik Biermans – 1241616
Brigitte van der Geest – 1464027
Pauline Haulez – 1462245
Willem Schellekens - 1636308

Index

1. Introduction

Image registration, the process of aligning two images, is a vital practice used throughout the field of medical imaging. It is used in the comparison of medical images of different moments in time, different modalities or different patients, necessary for clinical evaluation E.g., images from before and after treatment of a patient can be used to evaluate the effectiveness of the treatment. Computed tomography (CT) images can be aligned with Positron Emission Tomography (PET) images to further improve tumor detection and segmentation of structures in an image can be made easier when aligned with an atlas. For good and relevant comparisons between images, accurate alignment is needed (Kostelec & Periaswamy, 2003). The structures in the images must overlap in the area of interest, with as little error as possible to achieve this. Image registration can be done automatically to eliminate the human factor and to make the registration objective. Various algorithms which optimize the alignment of the images can be used, each with its own characteristics and features.

In this report, different methods of registration are elaborated and applied on a provided dataset of MR brain scans. A transformation matrix is calculated either semi-automatically or automatically and then applied to the moving image to align it with a fixed image. The methods are evaluated with the use of different metrics to assess the quality of the registration. The goal of this project is to compare and evaluate the different inter-and intra-modal registration methods. Experiments will be performed to demonstrate the advantages and disadvantages of each.

2. Methods

The used dataset, transformations, methods of registration and how the evaluation of the registration is done are elaborated below. Then, the approach which uses these methods is elaborated.

2.1 Dataset

The dataset provided consists of transverse MR brain scan slices. For a total of three patients, three different slices are used, which are all scanned both T1-weighted and by T2-FLAIR (Mendrik, et al., 2015). For the T1-weighed scans, both the original and a randomly transformed image are used. Throughout this report, the following format is used: {Patient ID}_{Slice ID}_{Sequence}, where patient ID is one of the three patients, slice ID is one of the three slices and sequence is the modality used. For the randomly transformed image, a ‘d’ is appended after the sequence number.
For intra-modal registration the images 1_1_t1 versus 1_1_t1_d and 3_3_t1 versus 3_3_t1_d are used. For inter-modal registration the image 1_1_t1 versus 1_1_t2 and 3_3_t1 versus 3_3_t2 are used

2.2 Transformations

The geometrical transformations are the transformations applied to the moving image, $x' = T(x)$. These include:
Translation: $\boldsymbol{x}' = \boldsymbol{x} + \boldsymbol{t}$, with $\boldsymbol{x}$ the original position vector and $\boldsymbol{t}$ the translation.
Rotation: $\boldsymbol{x}' = \boldsymbol{R}\boldsymbol{x}$, with $\boldsymbol{R} = \begin{bmatrix} \cos(\phi) & -\sin(\phi) \\ \sin(\phi) & \cos(\phi) \end{bmatrix}$, with ϕ the rotation angle
Scaling: $\boldsymbol{x}' = \boldsymbol{S}\boldsymbol{x}$, with $\boldsymbol{S} = \begin{bmatrix} s_x & 0 \\ 0 & s_y \end{bmatrix}$
Shearing: $\boldsymbol{x}' = \boldsymbol{H}\boldsymbol{x}$, with $\boldsymbol{H} = \begin{bmatrix} 1 & h_x \\ h_y & 1 \end{bmatrix}$.

Rigid transformations consist of only translation and rotation, i.e., the dimensions of the moving image do not change. Affine transformations can be combinations of translation, rotation, scaling and shearing, which does change the dimensions of the moving image.

2.3 Point-based and intensity-based registration

Point-based registration is performed semi-automatically based on user input in the form of reference points picked by the user. These points that are selected in the fixed image that are considered reliable, based on distinguishable features, are the fiducials (Fitzpatrick, J.M., et al.). The same points are selected in the moving image, after which a transformation matrix is calculated, using the difference in x and y coordinates, giving t_x and t_y. These values are then used in an affine transformation on the moving image.

Intensity-based registration is performed automatically. The intensity values of the pixels in the images are used as the basis for the registration. The alignment of the fixed and moving images are iteratively improved. The amount of movement per iteration is determined by the learning rate. An optimal registration is reached when the values used to evaluate the alignment reach an optimum for normalized cross-correlation or mutual information, as explained in the following paragraph.

2.4 Methods of registration evaluation

The different intensity-based similarity measures used in the evaluation of the alignment are:

Normalized cross-correlation (NCC)

Normalized cross-correlation assumes there is a linear relationship between the pixel intensities in the two images. However, for inter-modality registration, there is a lack of linearity. Normalized cross correlation is calculated with formula 2.1.

$$NCC(I,J) = \frac{\sum_{i=1}^{n}(I(i)-\bar{I})(J(i)-\bar{J})}{\sqrt{\sum_{i=1}^{n}(I(i)-\bar{I})^2 \sum_{i=1}^{n}(J(i)-\bar{J})^2}} \tag{2.1}$$

In which, I and J are the two images and i is the pixel number. The NCC value is normalized, meaning it will have a value between $0-1$. The higher the value returned, the better the registration. It is possible the value of 1 is not reached when more transformations than translation are done because of inverse mapping.

Mutual information (MI)

The pixels of the images have a finite set of possible values; thus, the probability mass function (PMF) can be defined which maps each possible value to a probability. Mutual information uses this probability mass function (PMF) of the two discrete images, defined with normalized image histograms. The formula for mutual information is stated below.

$$MI(I,J) = \sum_{i=1}^{n}\sum_{j=1}^{n} p_{I,J}(i,j)\log\left(\frac{p_{I,J}(i,j)}{p_I(i)p_J(j)}\right) \tag{2.2}$$

In which I and J are again the images, $p_I(i) = P(I = i)$ and $p_{I,J}(i,j) = P(I = i, J = j)$. The logarithm has base 2, therefore, the MI is expressed in units of bits. Again, the higher the value returned, the better the registration.

Target registration error (TRE)
The evaluation of the quality of the alignment is determined by calculating the error. This is done by taking the L^2-norm for every chosen pixel in the fixed and moving image. The pixels chosen are in relevant locations, and are different than the fiducial points chosen for the image registration. The formula for the target registration error is stated below.

$$TRE = \|\boldsymbol{TX}' - \boldsymbol{X}\|_2^2 \tag{2.3}$$

Then the mean distance error is calculated by dividing this error by the total amount of pixels.

Gradient ascent
The gradient is used to approach the maximal values for the NCC and MI. The optimum of these values is where the gradient is zero. The gradient can be approached as the following formula for discrete functions:

$$f(x,y) = \begin{bmatrix} \frac{\partial}{\partial x} f(x,y) \\ \frac{\partial}{\partial y} f(x,y) \end{bmatrix} \approx \begin{bmatrix} \frac{f\left(x+\frac{h}{2},y\right) - f\left(x-\frac{h}{2},y\right)}{h} \\ \frac{f\left(x,y+\frac{h}{2}\right) - f\left(x,y-\frac{h}{2}\right)}{h} \end{bmatrix} \tag{2.4}$$

In which h is a very small number, which is used to increase and decrease the values for x and y in very small steps in every iteration. The parameters will then be updated in the direction of the gradient until the gradient reaches zero.
The size of the step in the ascent is determined by the learning rate mu (μ). For maximizing the function $f(w)$, the function will thus be:

$$w \leftarrow w + \mu \nabla_w f(x,y) \tag{2.5}$$

2.5 Approach
The following image registrations and evaluations will be performed:

- Intra-modal rigid intensity-based registration, evaluated through the method of normalized cross-correlation, using slices 1_1_t1 and 1_1_t1_d. These slices are identical.
- Intra-modal affine intensity-based registration, evaluated through the method of normalized cross-correlation, also using slices 1_1_t1 and 1_1_t1_d.
- Inter-modal affine intensity-based registration, evaluated through the method of normalized cross-correlation, using slices 1_1_t1 vs 1_1_t2.
- Intra-modal affine intensity-based registration, evaluated through the method of mutual information, using slices 1_1_t1 vs 1_1_t1_d again.
- Inter-modal affine intensity-based registration, evaluated through the method of mutual information, using slices 1_1_t1 vs 1_1_t2 again.

With these registrations, the difference in TRE between rigid and affine transformations can be determined, with the transformation yielding the lowest error being the preferred transformation. Also, the difference in NCC, MI and TRE values for inter- and intra-modal transformations will be shown, which will illustrate how the registration of images with different intensity values affects the outcomes, when compared to images with similar intensity values. Lastly, an intra- and intermodal case is tested and compared with variable learning rate and compared to a fixed learning rate. With a variable learning rate (mu), the learning rate will decrease when the similarity becomes larger.

3. Results

In the following table the target registration error (TRE) is given for affine transformation of point-based registration.

Table 1. Point-based affine image registration

Intra-modal registration	**TRE**
1_1_t1 vs 1_1_t1_d	22.80
3_3_t1 vs 3_3_t1_d	12.27
Inter-modal registration	
1_1_t1 vs 1_1_t2	10.65
3_3_t1 vs 3_3_t2	8.76

In table 2 the values of the normalized cross-correlation (NCC) or mutual information (MI) are mentioned of affine intensity-based registration. The values are mentioned based on the standard parameters of the rigid transformation ([0, 0, 0]) or affine transformation ([0, 1, 1, 0, 0, 0, 0]). Also, the values of NCC or MI based on the optimal transformation calculated after the gradient ascent are mentioned. The TRE is also mentioned in table 2. A complete overview with the values of the learning rate and the resulting optimal transformation matrix is shown in table 4 in appendix A.

Table 2. Intensity-based image registration (NCC = normalized cross-correlation, MI = mutual information)

Registration method	**Used images**	**Result NCC or MI Before optimalization**	**Result NCC or MI After optimalization**	**TRE**
Rigid Intra-modal	1_1_t1 vs 1_1_t1_d 3_3_t1 vs 3_3_t1_d	NCC = 1 NCC = 0.60	NCC = 1 NCC = 0.63	5.10 444.88
Affine Intra-modal	1_1_t1 vs 1_1_t1_d 3_3_t1 vs 3_3_t1_d	NCC = 1 NCC = 0.60	NCC = 1 NCC = 0.74	18.21 193.38
Affine Inter-modal	1_1_t1 vs 1_1_t2 3_3_t1 vs 3_3_t2	NCC = 0.53 NCC = 0.56	NCC = 0.54 NCC = 0.60	213.02 118.94
Affine Intra-modal	1_1_t1 vs 1_1_t1_d 3_3_t1 vs 3_3_t1_d	MI = 3.02 MI = 0.63	MI = 1.92 MI = 0.83	2.20 122.35
Affine Inter-modal	1_1_t1 vs 1_1_t2 3_3_t1 vs 3_3_t2	MI = 0.72 MI = 0.71	MI = 0.85 MI = 0.82	157.95 97.27

In the following table 3 the results of MI are shown for fixed and variable learning rates (mu).

Table 1. Intensity based image registration with fixed or variable mu

Registration method	**Used images**	**Learning rate fixed/variable**	**Result MI**	**TRE**
Affine Intra-modal	3_3_t1 vs 3_3_t1_d	Fixed	MI = 0.63	444.88
Affine Intra-modal	3_3_t1 vs 3_3_t1_d	Variable	MI = 1.0	0.96
Affine Inter-modal	1_1_t1 vs 1_1_t2	Fixed	MI = 0.72	157.95
Affine Inter-modal	1_1_t1 vs 1_1_t2	Variable	MI = 0.84	1.22

4. Discussion

Point-based affine image registration is done to compute transformation matrices for both intra- and inter-modal cases. To evaluate the accuracy of this registration, the target registration error is calculated. The distance between a chosen set of points for both intra- and inter-modal cases ranges from 8.76 to 22.80 and is on average 13.6 squared pixels. This error is mainly caused due to human imprecision. Point-based registration requires a manual input for picking coordinates in two different images. It is unlikely that those points are picked precisely on the same spot in both images. Besides, the same method is used for picking new points to calculate the target registration error.
In order to avoid human errors during registration, an automated form of intensity-based registration is also evaluated. Several cases consisting of combinations of both rigid or affine, normalized cross-correlation or mutual information and inter- or intra-modal situations are evaluated. The learning rates for those cases differ significantly, and were hand-chosen to obtain the most precise results. The target registration error ranges from 2.20 to 444.88, indicating a large variation and a larger average of 137.33 squared pixels for intensity-based registration.

From these results the conclusion can be made that under these conditions point-based registration gives a lower target registration error than intensity based registration, therefore making point-based registration the more accurate method.

The large errors are caused by the learning rate (mu). A smaller mu results in the more precise results, but might take many more iterations if the initial displacement is large. Less iterations are often needed with a larger learning rate. However, generally, a larger mu results in a smaller final correlation. In almost half of the cases the similarity (NCC or MI) increases significantly after optimization compared to the similarity before optimization. This means that the current setting of 200 iterations should be increased. A significant difference between inter- and intra-modal cases is expected in cases using cross correlation, due to the large differences in intensities in intra-modal image registration. This, however, is also not shown in the results, probably cause by the issues with the learning rate.

The best way to counter the issues with the learning rate is to apply a variable mu. This is obtained by making the learning rate decrease when the similarity increases. The results show that a variable learning rate increases the accuracy significantly. Less iterations are needed to obtain similar results compared to a method with a fixed learning rate. The difference between inter- and intra-modal registration now also becomes more clear. The mutual information in the inter-modal case is relatively low, while the target registration error is very favorable. With the use of a variable learning rate, intensity based registration becomes a more accurate method than point based registration.

Another way to combat this issue could be by combining point-based registration with intensity-based registration. By using point-based registration an initial transformation vector can be obtained to bring the moving image closer to the fixed image. A small learning rate can then easily be used without taking too many iterations. Further research should also imply larger sample sizes in order to make more correct conclusions in the future.

References

Kostelec, P., Periaswamy, S. 2010/09/30. Image registration for MRI. Cambridge University press, Cambridge: MSRI publications

Mendrik, Adriënne M., e.a. 'MRBrainS Challenge: Online Evaluation Framework for Brain Image Segmentation in 3T MRI Scans'. Computational Intelligence and Neuroscience, vol. 2015, 2015, pp. 1–16. DOI.org (Crossref), https://doi.org/10.1155/2015/813696.

Appendix A – A complete overview of the values used for the optimization of the intensity-based registration

Table 4 – The used learning rate (mu) and resulting initial parameter vector for the optimization of the intensity-based registration

Registration method	**Used images**	**Learning rate (mu)**	**Best initial parameter vector**
Rigid Intra-modal	1_1_t1 vs 1_1_t1_d	0.0001	[1.55e-06, 2.24e-06, -2.19e-06]
Affine Intra-modal	1_1_t1 vs 1_1_t1_d	0.0001	[0, 1, 1, -9.31e-08, -8.34e-08, 3.19e-07, 5.03e-09]
Affine Inter-modal	1_1_t1 vs 1_1_t2	0.00001	[0, 1, 1, 0.0024, -0.0025, 0.0019, -0.0019]
Affine Inter-modal	1_1_t1 vs 1_1_t1_d	0.00001	[0, 1, 1, 2.32e-03, 8.89e-04, 1.16e-03, 1.22e-04]
Affine Inter-modal	1_1_t1 vs 1_1_t2	0.001	[0, 1, 1, 0.058, -0.050, -0.30, -0.078]
Rigid Intra-modal	3_3_t1 vs 3_3_t1_d	0.0015	[-0.031, -0.088, -0.015]
Affine Intra-modal	3_3_t1 vs 3_3_t1_d	0.0015	[0, 1, 1, 0.0033, -0.081, -0.021, -0.094]
Affine Inter-modal	3_3_t1 vs 3_3_t2	0.001	[0, 1, 1, 0.047, -0.044, -0.053, -0.079]
Affine Intra-modal	3_3_t1 vs 3_3_t1_d	0.001	[0, 1, 1, 0.0030, -0.080, -0.042, -0.10]
Affine Inter-modal	3_3_t1 vs 3_3_t2	0.001	[0, 1, 1, 0.051, -0.061, -0.069, -0.066]