VRIJE
UNIVERSITEIT
BRUSSEL

OPTIMIZATION OF REGISTRATION OF THE LIVER TUMOUR RADIOFREQUENCY ABLATION

Master of Biomedical Engineering

Paula Moreno Ruiz

August 23, 2018

Prof. Dr. Ir. Jef Vandemeulebroucke, Ing. Pieter Boonen
Sciences and Bio-Engineering Sciences

Abstract

Liver cancer is one of the most harmful cancers because it affects a vital body organ. There are multiple methods to treat this pathology with surgery being the most effective. However, not all patients can benefit of it due to age or health conditions. Due to its minimal invasiveness, radiofrequency ablation (RFA) is an alternative for some of these patients because the precise image guidance decreases the risk of complications.

It is important to elaborate a procedure that allows to evaluate if either RFA of a liver tumour was successful or an extra treatment like surgery is necessary. Even though there are some studies that assess the effectiveness, very few are reproducible or efficient as they use manual segmentations which are time consuming. Therefore, a registration of pre- and post- RFA images in a group of 9 patients was performed using SimpleITK in python in CT images. Despite registration is a useful method to compare these two images, it is a challenge to achieve a precise registration. It requires to match each voxel of the body of the patient, but not the tumour and ablation region.

The proposed method consists of a rigid transformation followed by a non-rigid transformation. Moreover, the use of different masks were studied to get the most efficient solution. Subsequently, this solution was then improved by tuning the registration parameters such as the grid spacing and the transform bending energy penalty. Finally, an exhaustive evaluation was performed in order to obtain the best parameter tuning basing on the accuracy and the plausibility measures.

For a dataset of 9 patients, it was appreciated how the quality of registration increases by the rise of the grid spacing from 15.0 to 10.0, and from 10.0 to 5.0 according to dice coefficient, false negatives, false positives and jaccard index. However, the weight assigned to the transform bending energy penalty had not the same effect. It could not define a direct effect between the quality of the registration and the increase or decrease of the weight, even having the same grid spacing.

Key words: RFA, CT image registration, non-rigid registration, liver cancer

Acknowledgements

First of all, I would like to thank Ha Manh Luu for sending me the parameter file that used in his article and for responding my doubts during these months. Also, to the UZ Brussel for providing us the images required to the registration, to the Vrije Universiteit of Brussels (VUB) for accepting me in their university and to the Polytechnic University of Valencia for giving me the chance to go to Brussels to do the last semester of the Master.

Secondly, I thank to my promotor Jef Vandemeulebroucke for giving me the opportunity of doing my MSc Thesis at ETRO and to give me his best advices. I am specially grateful to my supervisor Pieter Boonen for his unconditional help and support during the time that he has supervised me that I will never forget. Also, thanks to Lubos Omelina, who helped me with software in my firsts steps in the department always with smiling.

I would like to express a deep appreciation for Jorge, Elena and the group of friends I met in Brussels to cheer me up everyday, my home friends, but also, my office mates that welcomed me in the department.

My heartfelt thanks to my family, especially my mum and dad for listening to me everyday and for their support and attention. Lastly, I could not enjoy this experience without the help, comprehension and affection of Alejandro who encouraged and gave me his love every moment during these months.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Accounting for 8.8 million deaths in 2015, cancer is the leading cause of death worldwide according to OMS. Of these, the second most common cancer death is liver cancer with 788.000 deaths in this period of time [10].

It is one of the most harmful cancers that exists because this organ takes action in digestive procedures, not only breaking down and storing many of the nutrients absorbed from the intestine, but also, secreting bile into it to help to absorb nutrients. Furthermore, it develops refining functions of the body decomposing alcohol, drugs and toxic wastes in the blood [17].

In order to prevent liver cancer, early detection and a detailed diagnosis are important as they increase patient probability of surviving what is fundamental to treat the patient correctly. According to the different types of primary liver cancers, there are some different procedures to treat them as radiotherapy or chemotherapy. However, the most effective treatment is surgery although not all patients can benefit of it. Depending of different factors, there are two possible surgical techniques: the radiofrequency ablation (RFA) and the hepatectomy. [7]

The Interventional Radiologist (IR) commonly chooses RFA as a viable alternative in early metastatic liver cancer in inoperable patients because is a minimally invasive treatment that induces temperature in the tissue using high frequency to desiccate the tumour area with real-time guidance [13, 22]. That allows to remove the tumour with a low risk of complications and in an efficient way.

At the end of this process, it is fundamental to perform contrast material-enhanced Computed Tomography (CME-CT) to the patient in order to, either enable us to assess technical success or reveal any possible procedure-related complications. The treatment is considered successful when the ablation zone completely encom-

passed the tumour observed at CT imaging and treatment failure as incomplete coverage of the tumour seen at CT pre-intervention [1].

Although RFA procedure does not have complications due to follow well standardised protocols, after performing the surgery the IR has to evaluate the follow-up imaging to be sure that the tumour has been removed completely through comparing the post ablation image with the image taken before the ablation. This is a tricky and risky task because they have to take into account changing factors of the liver as the breathing or the weight gain or loss of the patient. In spite of trying to be as accurate as possible, they may make mistakes and because of that, cancerous tissue could still be in the liver, which would cause the recurrence of the tumour being damaging for the patient prognostic. Consequently, IR needs a tool that provides them this accurate and objective evaluation of the RFA intervention.

A potential alternative to evaluate the RFA accuracy objectively, is offered by imaging processing. It would require to select images, pre-process and register them. This registration will consist in comparing the shape, size and position of the original tumour and the treated zone after the intervention as the IR would do it. That procedure can be rigid or non rigid. Performing the rigid one, the image is translated or/and rotated to adapt to the another one, while using non-rigid transformation is provided free movement. The issue will be to find the correct combination of both. Then, it is required the validation of registration in order to assure the reliability of the method.

This could be a solution, although it will be still necessary the IR evaluation in order to assure the correct performance of registration. Therefore, the tool should be always supervised by an expert.

Chapter 2

State of the art

It is important to elaborate a procedure that could evaluate the efficiency of RFA in order to assure the correct deletion of the tumor. Nowadays, this process is still done manually, by comparing visually the image before and after the RFA.

According to this requirement, there are several studies that try to assess the efficiency of RFA reviewing manually CT scans after the procedure [2]. Other researchers try to predict the recurrence comparing the use of different imaging diagnosis as contrast-enhanced ultrasonography (CEUS) [16], multidetector-row computed tomography [4] or positron emission tomography-computed tomography (PET/CT) with fluoride radiolabeled deoxy-glucose [21]. Although none of them uses a tool that could help to assess the post-image obtained from all of these diagnostic imaging procedures objectively, Vandenbroucke *et al.* [3] take profit of an interactive procedure that align the ce-CT-images and, hereafter, verify the tumor coverage 24h after the RFA. Also, Keil *et al.* created a semiautomatic tool to evaluate the initial CT images and the RFA zone [5].

Moreover, there are other authors that try to automatically align both images by different ways. Even though it was some years ago, Meyer *et al.* [12] performed an automatic mutual information-based registration algorithm in phantom studies that requires little pre-processing and minimal user input. Also, Maes *et al.* [9] maximized mutual information with optimization methods and multiresolution strategies to allow a robust and accurate fully automated affine medical image registration tool without using segmentations. Focussing on the liver, Rieder *et al.* performed a rigid transformation avoiding the use of liver masks, but using ablation and tumor masks [14].

Achieving a registration without using masks is an important issue because it would allow doctors to do automatically the comparison of the images. However,

it is also necessary to achieve the best accuracy performing a registration adapted to patient and tissue conditions, as the case of liver tumor. If it is required to compare a region in the images that have different texture in both images, as the ablation zone and the tumor area, it will be necessary to perform different type of transformations. For that reason, Staring *et al.* proposed a local rigidity penalty term which is included in the registration function that penalizes deformation of rigid objects [18].

In 2010, Klein *et al.* [15] developed *elastix*. This toolbox consists in a collection of algorithms used to solve medical image registration problems. It allows to perform medical image registrations while being quickly to configure, test and compare different registration methods. Consequently, Luu *et al.* [8] developed an automatic registration of the images pre- and post- liver tumor ablation using a nonrigid registration with a rigidity penalty as Staring did in a general way using *elastix*. They used as similarity metric $S(T; I_F, I_M)$ normalized cross correlation (NCC) satisfying a linear relation between the intensity values of the fixed and moving images. As a nonrigid transform T they used a B-spline deformation field applying an iterative stochastic gradient descent optimizer.

The method proposed by Luu *et al.* [8] requires a segmentation of the tumor before RFA procedure and after the scanning because they applied a rigidity penalty in that specific region. Despite it was performed in all body, the similarity metric was evaluated only over liver area defined by a liver mask. Regarding evaluation metric, they used one overlap metric as the Dice similarity coefficient, Mean Surface Distance (MSD) and a Local Mean surface distance (LMSD). According to a grid size of 5 mm, $\alpha = 2$ and $\beta = 10$, they obtained a Dice in rigid registration of 87.9%, and a 92.2% in nonrigid with the local rigid deformation.

Chapter 3

The main objective

The initial goal of this project is to develop a registration method without masks and to prove that allows to obtain better results than the one proposed by Luu *et al.* [8] .

To achieve this objective, we examined the optimal combination of masks in order to decrease complexity of the registration method. We used CT images before and after the intervention because it allows us to show the tumor and the ablation zone. Hereafter, different parameters related with the registration process will be tuned in order to find the most efficient and accuracy combination of them.

The validation of this procedure was done using overlap and deformation performance measures that allows to compare results between tests. Thanks to this protocol the interventional radiologist will know if the ablation already done was correct performed to achieve the total deletion of the tumor or, if it was necessary a second intervention.

Chapter 4

Materials and methods

4.1 Data description

Registration was performed on CT follow-up dataset of the thorax of 9 patients having different types of cancers, from primary to metastasis cancer, provided by the UZ Brussel, Belgium. This dataset were from different types of studies and different times. Specifically, the provided dataset consisted of 3 scans per patient: one scan of the day before the surgery, other of the day of the surgery and another of the day after. The images used were CT-PET studies acquired by a Philips CT. These images had slices of 512x512 voxels and a thickness of 2.0 and 5.0 mm.

First of all, patients were selected based on the availability of their images one day before and after the ablation procedure. It was important to obtain these images with the higher resolution that was possible. Because of that, we selected 9 patients from an original dataset of 10, with a thickness of 2.0 mm that was the best available for each patient.

4.2 Image processing

In order to better understand the process that was followed, we are going to describe each step of the flowchart of the figure 4.1.

According to Luu's research, an imaging pre-treatment and a manual segmentation were required before performing registration. Then, registration was composed by two main steps: a rigid ("Initialization" or "Rigid" according to Figure 4.1) and a non-rigid transformation. For each case, a different combination of masks was used in order to get their optimal use. After selecting the best case, a parameter tuning was done and proved by validation metrics.

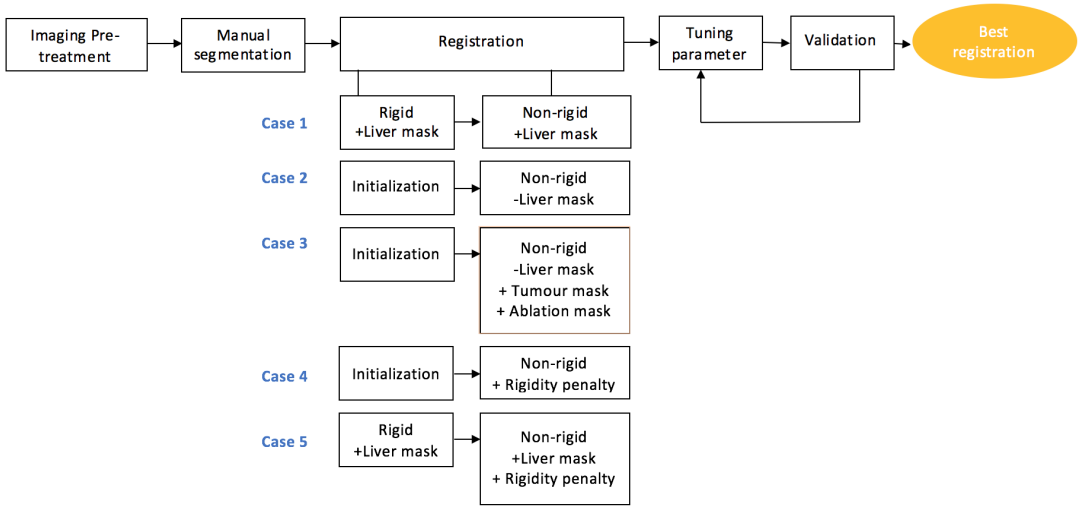
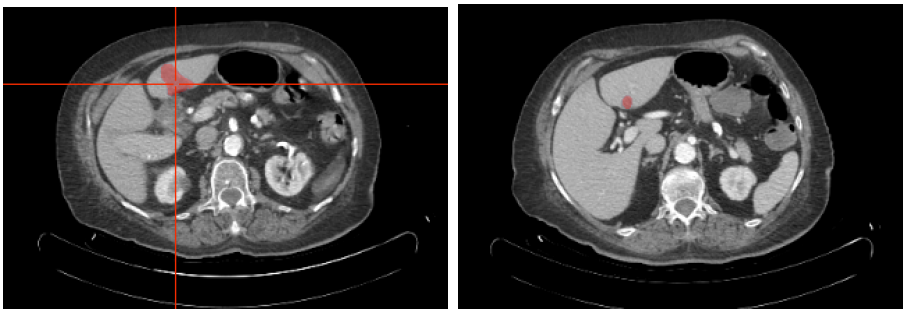


Figure 4.1: Flowchart

4.2.1 Imaging pre-treatment

After the selection of the images before and after the ablation with the highest resolution, it was almost mandatory to crop the images because they had different sizes for each patient in order to standardise the process that it will be followed during the algorithm. Moreover, that allows us to save memory due to the huge space required to store.

For that reason, we decided to fix a lower crop value of 200 and a higher crop value of 200 to decrease the image size, at least a percentage, because in some patients the acquisition has more slices than others.



(a) Ablation zone (b) tumor zone

Figure 4.3: Segmentations in vv

4.2.2 Manual segmentation

A manual segmentation of the liver, tumor and ablation zone were required in order to perform the experiments that will be explained in the following sections. Moreover, they are useful to perform a validation based on overlapping measures. The manual segmentation proved to be challenging due to poor image quality and the presence of multiple tumors. To carry out the segmentation, slicer was used (Figure 4.2). Overlay images were created using vv (Figure 4.3a and 4.3b).

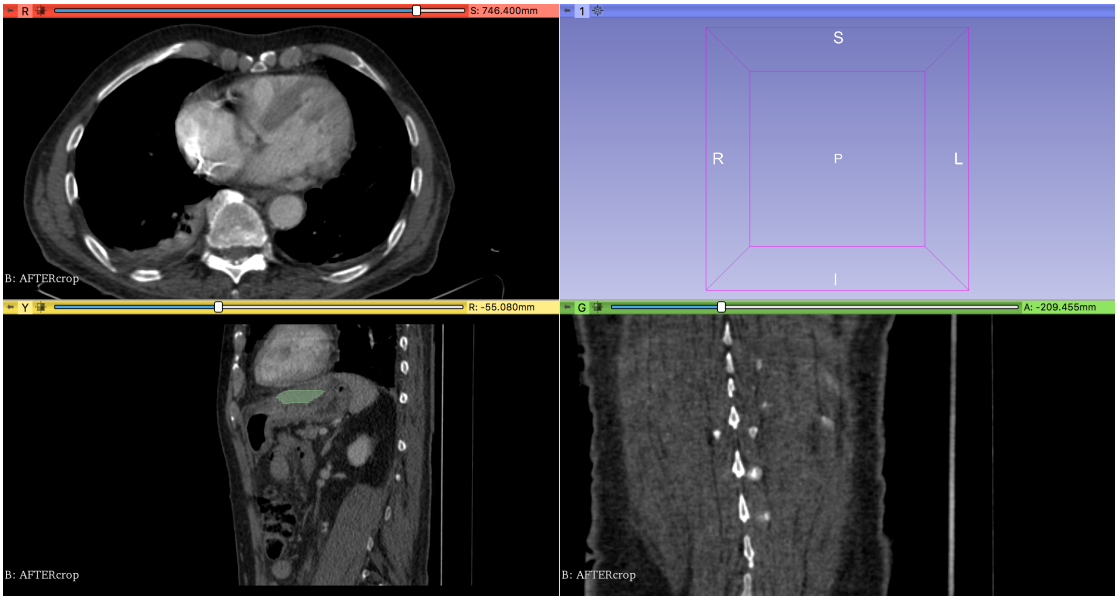


Figure 4.2: Segmentation of the ablation zone by 3D Slice

4.2.3 Registration

Registration is a technique that is used to compare two images though aligning their voxels. In order to do it, it is necessary to define one of them as a fixed image $I_F(\vec{x})$, that would be the reference, and the other as the moving image $I_M(\vec{x})$, which will be modified based on the fixed image. [6]

To achieve this goal, we should apply a function $T(\vec{x}) = \vec{x} + u(\vec{x})$ to transform one of them into the another, applying a transformation field $u(\vec{x})$ to the original space [**?**]. Depending on the degrees of freedom, it exists different types of transformations. In general, they could be rigid and non-rigid. Performing the rigid, we could translate and/or rotate the image to adapt to the another one, while using non-rigid transformation we have free movement. The aim of this field is to find the optimal T(x) for the correct transformation $S(T; I_F, I_M)$.

For the non-rigid registration, a cost function should be minimised depending on parameters:

$$\hat{T} = \arg\min C(T; I_F, I_M)$$

$$C(T; I_F, I_M) = -S(T; I_F, I_M) + \alpha P(T)$$

P(T) is the regularization factor to control the non-rigid transformation and α is the weighting factor that adjusts the relation between $-S(T; I_F, I_M)$ and P(T).

B-Spline deformation is used to model the transformation T that depends on μ. For that reason, it will be the same to find μ to minimise $C(\mu; I_F, I_M)$ in order to find the optimal transformation, given by:

$$\hat{T}_\mu = \arg\min_T C(T_\mu; I_F, I_M)$$

or

$$\mu = \arg\min_T C(\mu; I_F, I_M)$$

Generally, it is required to use a combination of both transformations because of the tissue characteristics: there are tissues more difficult to bend than others, such as bone, which can not deform at all. In our case of study, it is important to align the liver of both images (before and after the ablation procedure) but not tumor and ablation zone of each of them. Therefore, to avoid a locally minimisation of differences between the tumor area and the ablation zone, an Euler transformation as a rigid, followed by a non-rigid B-Spline transformation was used. Performing the registration in that way, we initialised moving and fixed image easily and faster and, later, aligned the deformable parts of the image.

In both cases, an optimal alignment is generally determined by optimizing a similarity metric $-S(T; I_F, I_M)$. In this project, the S of the Mutual Information (MI) between the transformed test image and the reference image was chosen according to Thvenaz and Unser [20]:

$$S(\mu) = \sum_{\iota \in l_T} \sum_{\kappa \in l_R} p(\iota, \kappa; \mu) log_2(\frac{p(\iota, \kappa; \mu)}{p_T(\iota; \mu) p_R(\kappa; \mu)})$$

In order to keep the tumor and the ablation as much fix as possible, it was used penalty terms to restrict the space of T.

- Rigidity penalty. $\mathcal{P}(T)$ was substituted by a rigidity penalty term $\mathcal{P}^{\text{rigid}}(T; I_M)$ that is defined by different parameters: $c(\vec{x})$, affine term $AC_{kij}(x)$, orthonormality term $OC_{kij}(x)$, and a properness term $PC(x)$. $c(\vec{x}) \in [0, 1]$ is a user-predefined coefficient that allows to restrict the rigid movement to the part inside a mask defined as:

$$c(\vec{x}) = \begin{cases} 1 & \text{if } \vec{x} \, \epsilon \, \Omega_R \\ 1 - \dfrac{1}{1 + e^{-\frac{D(\vec{x}) - \beta}{\alpha}}} & \text{else} \end{cases}$$

  where, $D(\vec{x})$ is the Euclidean distance transform of the mask image with zero distance at the tumor boundary. The idea of doing that is to be sure about performing a rigid transformation inside the tumor, almost rigid around and non-rigid outside the tumor area. The difference it is showed in Figure 4.4 and 4.5 [19].

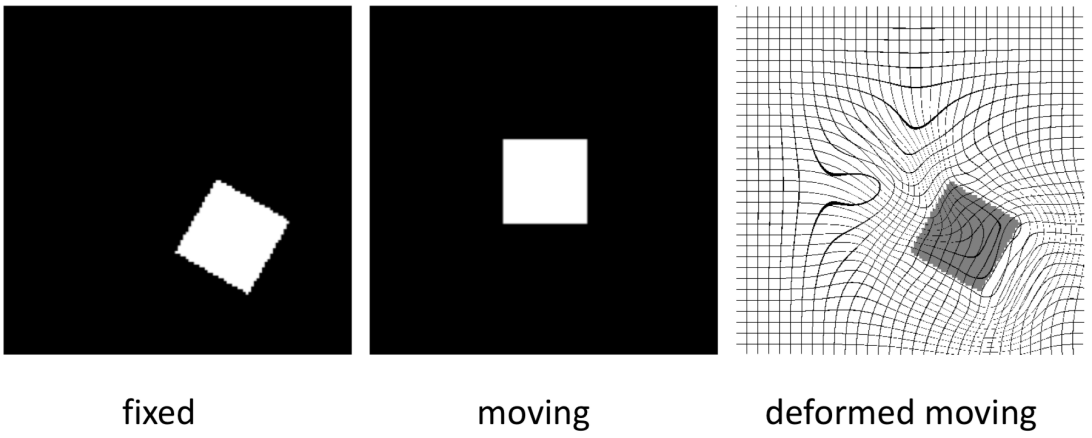


Figure 4.4: Purely nonrigid registration

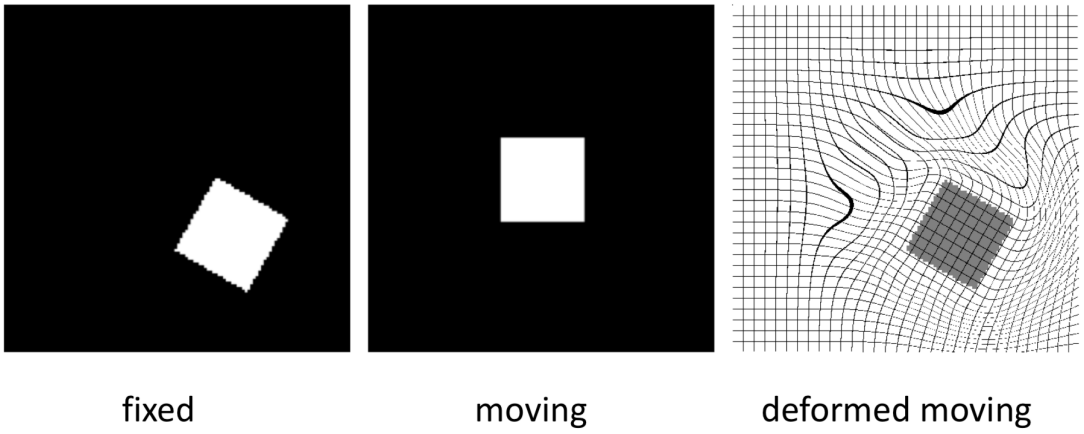


Figure 4.5: Nonrigid registration with rigidity penalty term

- Bending energy penalty. Other option to restrict the deformation is to use the "Transform Bending Energy Penalty"as a second metric. The main difference between both penalty terms is that the rigidity penalty makes rigid deformation around a local area fixed by a mask and the bending energy penalty is applied to the all image equally.

4.2.4 Tuning parameters

Experiments started using a parameter file that has standard parameters. Either in case of rigid and non-rigid, the parameter file has the same general values. Based on the parameter files provided by Luu *et al.* [8] a registration was performed.

Focussing on each parameter, it was defined Pixel type as "short" because saves memory, what is very important when you are using and generating a big amount of data. Regarding the main components, a multiresolution registration with a B-spline Interpolator was carried out. It was applied an adaptive stochastic gradient descent as optimizer and a recursive image pyramid for either fixed and moving. Finally, a random image sampler was used in the 2000 iterations with a B-spline interpolator of order 1.

Whereas the Euler transformation was performed using the same parameter file for all experiments, the B-spline or non-rigid registration used different files depending on the usage of masks or penalties to the standard case. Figure 4.6 shows the standard parameter file used for registration. Specifically, it was important to change the 'ErodeMask' option as false or true, depending on the use of a mask or not. In the case of the application of a rigidity penalty, it was required to add different parameters based on Luu's parameter file [8].

NON RIGID BASE CASE	
Image Types	
Fixed Internal Image Pixel Type	short
MovingInternalImagePixelType	short
UseDirectionCosines	TRUE
Main components:	
Registration	MultiResolutionRegistration
Interpolator	BSplineInterpolator
ResampleInterpolator	FinalBSplineInterpolator
Resampler	DefaultResampler
FixedImagePyramid	FixedRecursiveImagePyramid
MovingImagePyramid	MovingRecursiveImagePyramid
Optimizer	AdaptiveStochasticGradientDescent
Transform	BSplineTransform
Metric	AdvancedMattesMutualInformation
Transformation:	
AutomaticScalesEstimation	TRUE
FinalGridSpacingInPhysicalUnits	15.0
HowToCombineTransforms	Compose
Similarity measures	
NumberOfHistogramBins	32
ErodeMask	FALSE
Multiresolution	
NumberOfResolutions	4
ImagePyramidSchedule	8 8 8 4 4 4 2 2 2 1 1 1
Optimizer	
MaximumNumberOfIterations	2000
Image sampling	
NumberOfSpatialSamples	2048
NewSamplesEveryIteration	TRUE
ImageSampler	Random
Interpolation and Resampling	
BSplineInterpolationOrder	1
FinalBSplineInterpolationOrder	3
DefaultPixelValue	0
WriteResultImage	TRUE
ResultImagePixelType	short
ResultImageFormat	mhd

Figure 4.6: Non-rigid registration standard parameter file

In order to improve the registration tuning parameter of files explained above, we decided to perform different experiments modifying one per one as we are going to explain in the following lines.

The grid spacing of the B-spline transform for each dimension was an important parameter to consider. In order to better understand this concept, it would be useful to show Figure 4.7, where different grid spacings are visualized. As the grid spacing improves, the resolution of the image and thus the registration decreases. Consequently, as lower is the grid spacing, higher is the accuracy of that registration. However, there are inconveniences such as the high time of execution. *Simple Elastix* allows to obtain this argument in millimeters or in "voxel size units", depending on the specific factor name used.

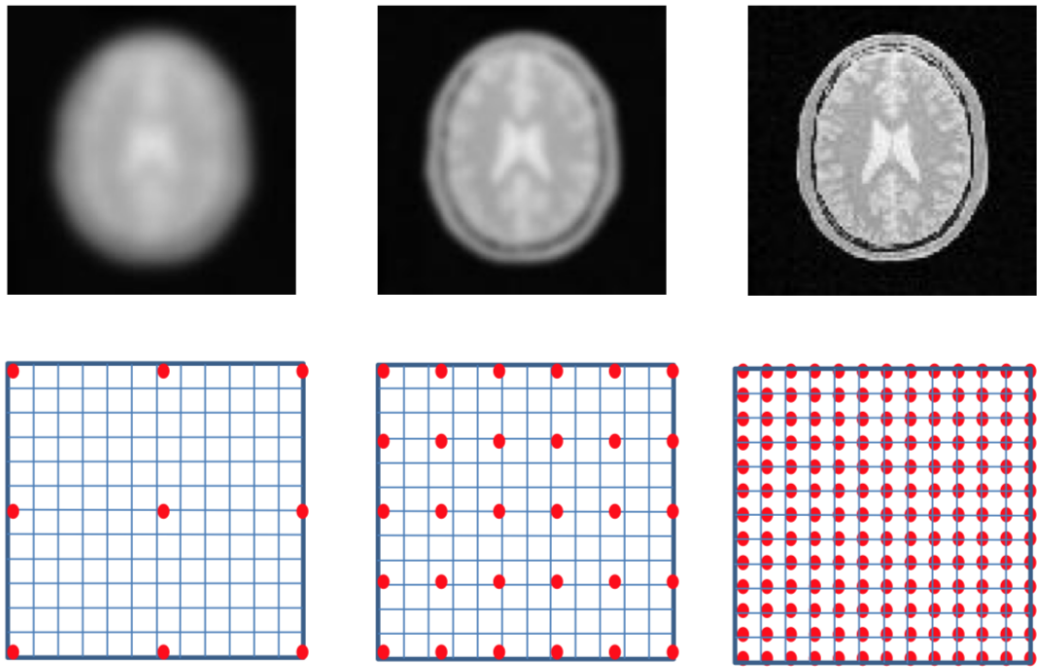


Figure 4.7: Grids of control points

Other alternative parameter to tune in order to improve registration, is the weight of each metric used (Advanced Mattes Mutual Information and Transform Bending Energy Penalty). It is possible to improve the accuracy of registration keeping one of them constant and tuning the other one. In the current project, Transform Bending Energy Penalty weight was tuned keeping the Advanced Mat-

tes Mutual Information as 1 in order to achieve the best results.

4.2.5 Validation

Validation is required to obtain an qualitative assessment of the registration. It is possible to use a segmentation approach to evaluate image registration through comparing the segmentation before and after the registration quantitatively. In this project, Dice Similarity Coefficient (DSC), False negatives (FN), False positives (FP) and Jaccard Similarity were used. Considering Figure 4.8 these metrics are going to be explained in detail according to Crum [11].

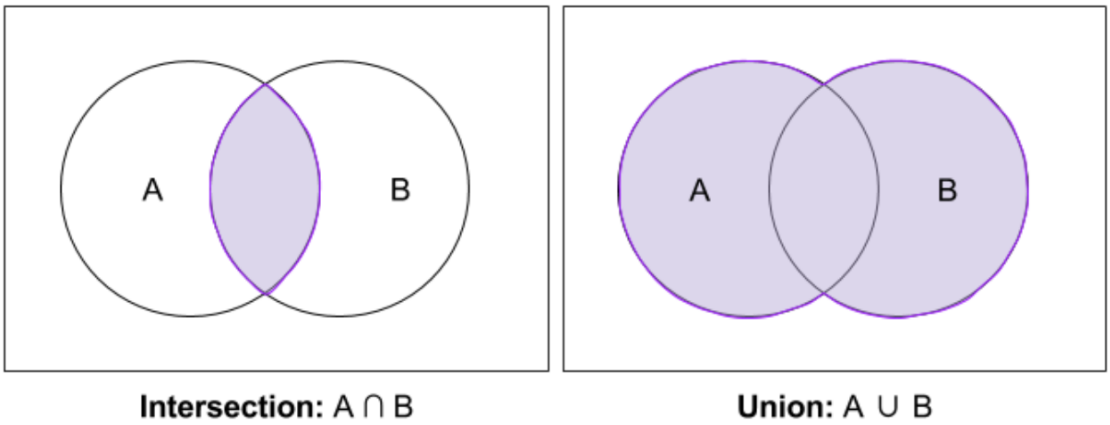


Figure 4.8: Intersection and union

- Dice Similarity Coefficient. For A and B as an overlapping regions this metric represents the ratio of the intersection to the mean label volume. It is defined as followed:

$$DSC = \frac{2(A \cap B)}{A + B}$$

- Jaccard Similarity or Tannimoto Coefficient is the ratio of the number of voxels of the intersection and the union.

$$Jaccard = \frac{A \cap B}{A \cup B}$$

  The value of that index should be between $[0 : 1]$ being $Jaccard(A, B) = 1$ when $A \cap B = A \cup B$.

These two metrics are related by the following expression. They are equal at the extrema 0,1 and between those limits DSC >Jaccard.

$$DSC = \frac{2Jaccard}{Jaccard + 1}$$

- False negatives (FN) and false positives (FP). False negatives and false positives measures are easier to understand looking at the Figure 4.9. In both cases, the optimal values are close to 0, what means a perfect overlapping between the segmentation after the registration and the ground truth that, in the current project would be the segmentation before registration.

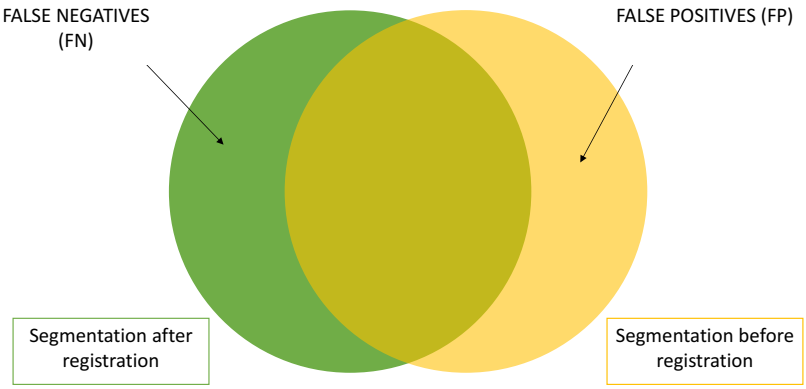


Figure 4.9: False negatives and false positives

In addition, it is useful to obtain texture measures through the Jacobian determinant of the registration results. In a 3D image, this element is a matrix composed of the determinant of the Jacobian matrix of each voxel that describes how the transformation looks like when you zoom near a specific point in the image. This transformation is represented by the following equation in 2D case:

$$Jacobianmatrix = \begin{bmatrix} \partial f_1/\partial x & \partial f_1/\partial y \\ \partial f_2/\partial x & \partial f_2/\partial y \end{bmatrix}$$

From this concept the Smoothness was obtained to have an idea about how was distributed the deformation in the image. That parameter is calculated obtaining

the standard deviation (ST) of the determinant of the Jacobian matrix as it is showed below [6].

$$Smoothness = ST(JacobianDeterminant)$$

For that reason, the closer to 0 the Smoothness is, the more smooth will be the deformation of this part of the image. Depending on the part of the body that is being evaluated, this value should be lower or higher. For example, in the case of bones, the smoothness must be 0 or almost 0. If not, it can be assumed that registration was performed incorrect. Regarding the liver, it is expected to obtain higher values but not up to 0.8. This value will use to fix a threshold that allows to obtain an extra metric that will give the number of voxels (NOV) that exceed this value in the liver region.

4.3 Software

The code was developed with Python 3.6.3 using Anaconda. To perform the registration, Simple ITK and Simple Elastix were used because of the application of useful functions.
Apart of the code, other programs were used in different parts of the project:

- VV was used to visualize images, masks and registration results. That allows to follow each step of the algorithm checking its correct development.
- Horos allowed to organize the initial database obtaining images with higher resolution and also gave easy access to their metadata.
- Slicer is the program used to segmentate liver, tumor and ablation region of each patient. It provides useful tools to read, draw, visualize and write images.

Chapter 5

Experiments and results

The following chapter is structured in two main parts. Firstly, an study of the use of masks is described to find the best registration method using different combination of masks. Secondly, one of the methods was selected in order to improve results tuning specific parameters. The evaluation of them is going to be explained at the end of each part with box-and-whisker plots.

At the end of the chapter some examples are going to be visualized in order to understand better results obtained.

5.1 Study of the use of masks

In order to investigate the effect of using masks on the registration, we performed 5 methods, applying different mask combinations. The registrations were done using liver-, tumor- and ablation zone masks. The following table is an overview of the registrations methods.

Table 5.1: Registration methods used

	Base case		
Case 1	INITIALIZATION	RIGID + Liver mask	NON RIGID + Liver mask
Case 2	INITIALIZATION	NON RIGID - Liver mask	
Case 3	INITIALIZATION	NON RIGID - Liver mask + Ablation mask + Tumor mask	
Case 4	INITIALIZATION	NON RIGID - Liver mask + Rigidity penalty (tumor)	
Case 5	INITIALIZATION	RIGID + Liver mask	NON RIGID + Liver mask + Rigidity penalty

It is important to define some concepts shown in the table:

- Initialization is the part of the process that allows us to situate the fixed and moving image in the same origin using a rigid transformation without applying a mask. After the initialization, we performed the additional registrations, applying masks. It is required to perform a second rigid transformation after the initialization when it is going to use a liver mask in the non rigid one.
- Rigid is the second Euler transformation required to use a liver mask.

5.1.1 Case 1

In this case, an initialization step was performed to center the fixed and the moving image. Consequently, we did another Euler transformation but using the result of the initialization step as the moving image of this second rigid registration. Finally, a non-rigid registration was performed with a liver mask applied to the fixed image. As the used mask corresponds to the fixed image, it was not required to perform an initialization of it.

We obtained dice coefficient, false negatives, false positives, jaccard index and smoothness in order to compare each case. We created tables to store results for each case as the following Table 5.2 that is for case 1. The per patient validation

results for the remaining cases will be provided in the appendix A since we will make a decision based on the boxplots provided in 5.2

Table 5.2: Results general approach case 1

	Dice	FN	FP	Jaccard	Smoothness
Patient 1	0,8868	0,0569	0,1632	0,7966	0,0728
Patient 2	0,8970	0,0632	0,1395	0,8133	0,0666
Patient 3	0,8326	0,1232	0,2073	0,7132	0,0705
Patient 4	0,8275	0,1256	0,2145	0,7058	0,0721
Patient 5	0,9092	0,0351	0,1404	0,8336	0,0463
Patient 6	0,8801	0,0735	0,1404	0,7859	0,0430
Patient 7	0,9113	0,0466	0,1272	0,8371	0,0739
Patient 8	0,6395	0,3428	0,3772	0,4701	
Patient 9	0,8201	0,1417	0,2149	0,6950	0,0681

5.1.2 Case 2

One of the objectives was to remove masks of the algorithm in order to do the registration as much automatic as possible. For that reason, this case was performed just doing the initialization of the fixed and moving image followed by a non rigid registration, without using any mask.

5.1.3 Case 3

Even though we tried to remove all masks, we could not ignore that the use of them is relevant to the improvement of the registration. Therefore, we decided to perform a non rigid registration using two different masks: on the one hand, an ablation zone mask applied to the fixed image, that is the image acquired after the ablation performed and, on the other hand, a tumor mask applied to the moving image that was the image before the ablation procedure.

Due to the use of a moving mask, it was required to apply an initialization step to the moving mask as well in order to locate the moving mask in the same position as the moving image.

5.1.4 Case 4

Following the same idea, we performed a non rigid registration without using any mask, but, in that case, applying a rigidity penalty [18].

As we explained in the Material and methods section, in order to apply a penalty to the registration it is required to add specific lines to the parameter file with details of the mask where the rigidity penalty will be applied. In addition, it is a key to specify the path of the mask as Figure 5.1 showed.

```
(RigidityPenaltyWeight 0.1 0.1 0.1 4.0)
(LinearityConditionWeight 100.0)
(OrthonormalityConditionWeight 1.0)
(PropernessConditionWeight 2.0)
(UseLinearityCondition "true")
(UseOrthonormalityCondition "true")
(UsePropernessCondition "true")
(CalculateLinearityCondition "true")
(CalculateOrthonormalityCondition "true")
(CalculatePropernessCondition "true")
(DilateRigidityImages "false" "false" "false" "true")
(DilationRadiusMultiplier 2.0) // used to be 2
//(UseFixedRigidityImage "false")
//(FixedRigidityImageName "dummy")
(UseMovingRigidityImage "true")
(MovingRigidityImageName "/Users/paula/Virtual_environment/ExamplePieter/
Registration/Registration_case4_1/0013/NonRigid/Tumour.mhd")
```

Figure 5.1: Parameters added to the non rigid parameter file to apply a rigidity penalty

5.1.5 Case 5

Lastly, a liver mask to the rigid registration and non-rigid registration, as for case 1, but a rigidity penalty (Figure 5.1) was applied to the non-rigid registration. The procedure was the same as case 1, but adding the lines of Figure 5.1.

5.1.6 Overview of results

In order to compare the different cases explained above, boxplots are used to visualised dice coefficient, false negatives, false positives, jaccard index and smoothness. For each case, the average of all patients is taken. For every validation measurement, the initalization is included as well.

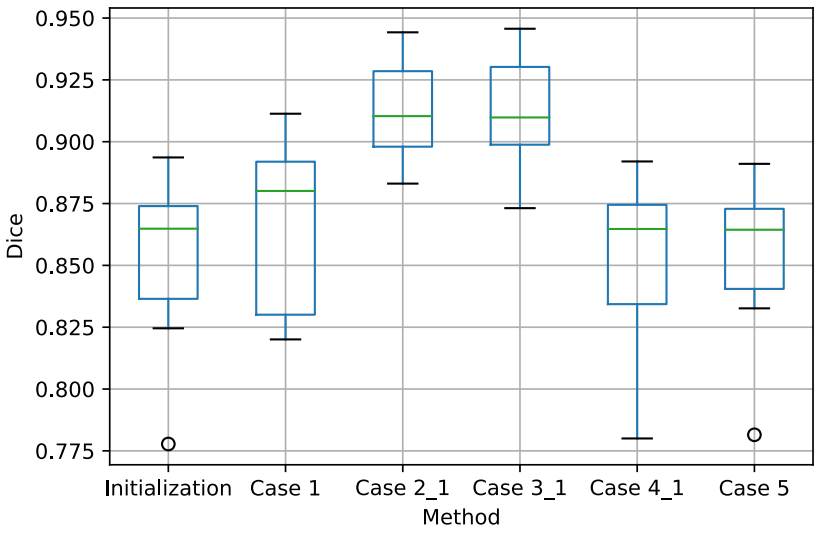
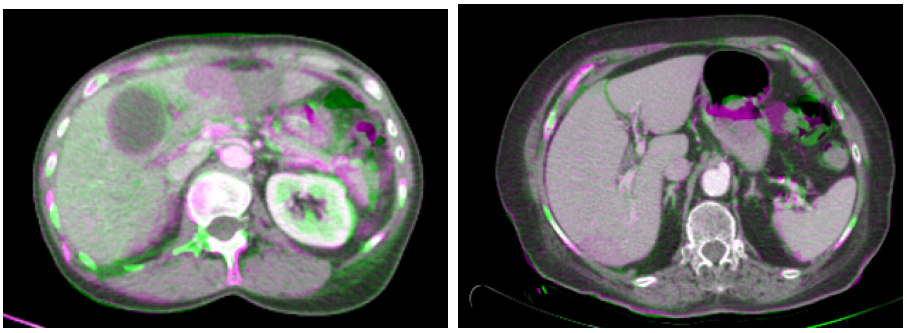


Figure 5.2: Dice coefficient. Best case would be high values

As it is showed in Figure 5.2, cases with best dice results are case 2 and 3. Any of them used liver mask, but case 3 applied an ablation- and a tumor mask to do the registration. In both cases, the distribution of patients is regular (between a similar 75th and 25th percentile), whereas in case 1, 4 and 5, the interquartile range is wider. Also, it is possible to appreciate a higher standard deviation in case 4 and some outliers in case 5. That means that registration is performing correctly depending on the patient, what gives uncertainty.

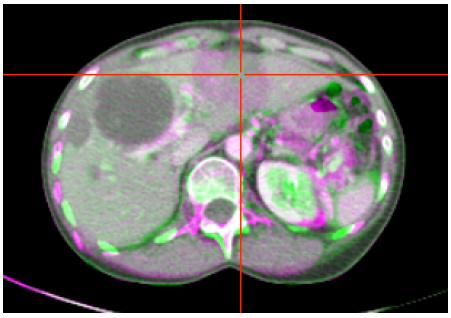
To better understand what disparities could appear between registrations, some patients are going to be visualised. As the dataset is composed of different patients, it is difficult to create a uniform registration. For example, for patient 4 there was more movement during the scan compared to other patients. Consequently, applying the same case to this patient (Figure 5.3a) and to other easier one, as number 7 (Figure 5.3b), it is showed that results are better in the second one due to conditions of the acquisition (Figure 5.3).



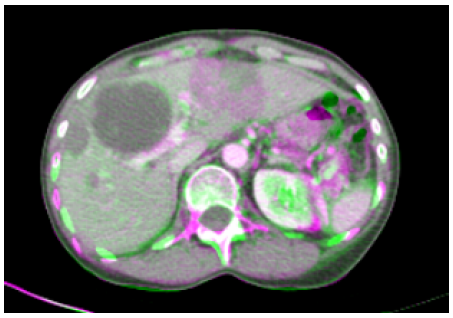
(a) Patient 4 (b) Patient 7

Figure 5.3: Case 2. Initialization followed by non rigid registration with no masks

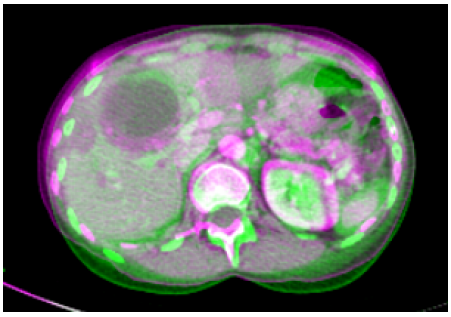
For that reason, it is notable that registration in patient 7 is better done than in patient 4. However, if the comparison is done between cases applied to the same patient, the behaviour is almost the same in both of them. It is appreciable how case 4 and 5 present problems matching liver edges either in patient 4 (Figure 5.4) and 7 (Figure 5.5). Furthermore, it is clearly showed, that in patient 4 these matching problems are more significant than in patient 7 (Figure 5.4 and 5.5)



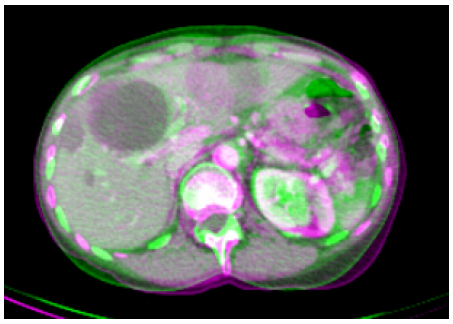
(a) Case 2 in patient 4



(b) Case 3 in patient 4

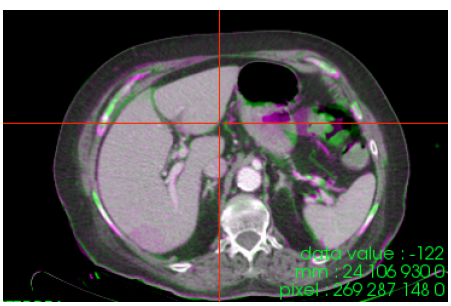


(c) Case 4 in patient 4

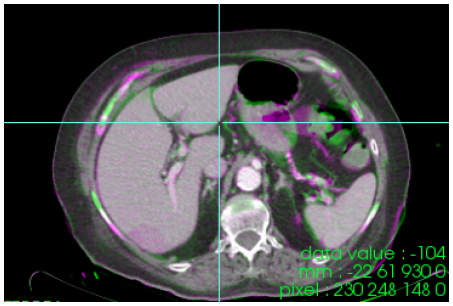


(d) Case 5 in patient 4

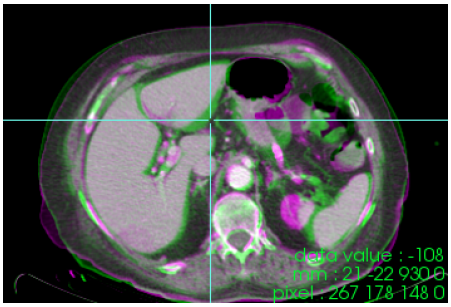
Figure 5.4: Comparison between cases in patient number 4



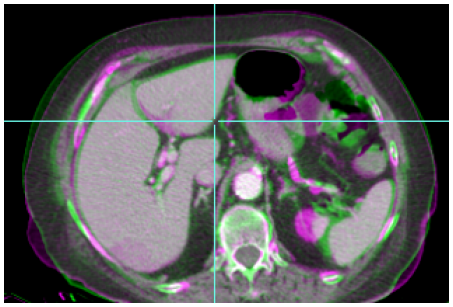
(a) Case 2 in patient 7



(b) Case 3 in patient 7



(c) Case 4 in patient 7



(d) Case 5 in patient 7

Figure 5.5: Comparison between cases in patient number 7

The Initialisation step is equivalent to not doing any registration at all, just an alignment of the images origin. Applying the transformation of case 4 & 5 did not improve the dice coefficient compared to the initialization.

Figure 5.6 and 5.7 show the same tendencies for the false negatives and false positives as the dice coefficient does. Case 1,4 and 5 offer similar results compared to the baseline and case 2 & 3 show an improvement.

Focussing on False negatives (Figure 5.6), best results are obtained for case 3, as the average of patients has a lower value than case 2. However, in the false positives plot (Figure 5.7) both cases have apparently the same average.

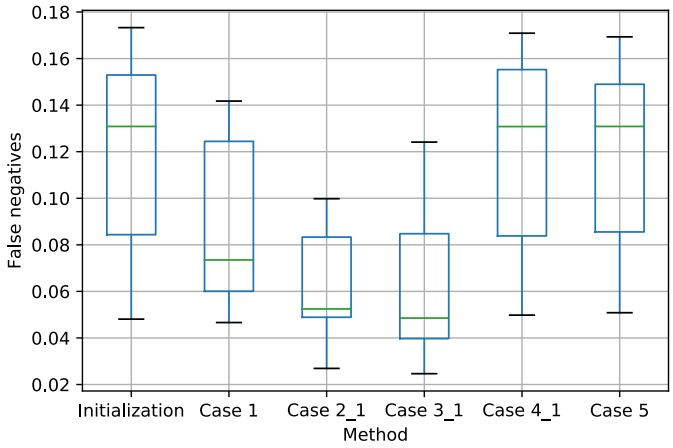


Figure 5.6: False negatives. Best case corresponds to lower values

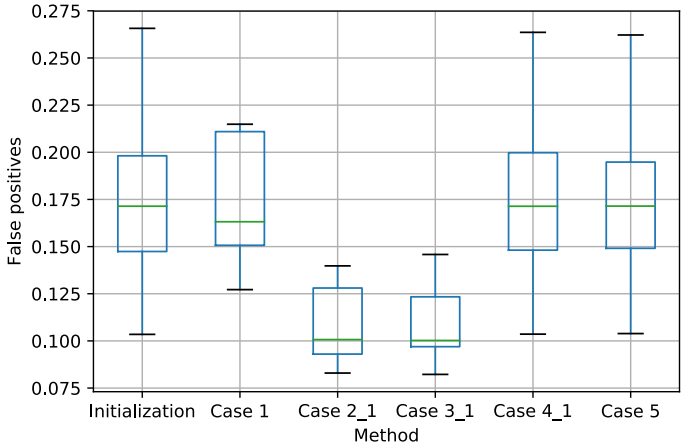


Figure 5.7: False positives. Best case corresponds to lower values

As the definition of Jaccard index explains, best results are the higher, as the dice coefficient. For that reason, case 2 & 3 are the best ones comparing to initialization step (Figure 5.8).

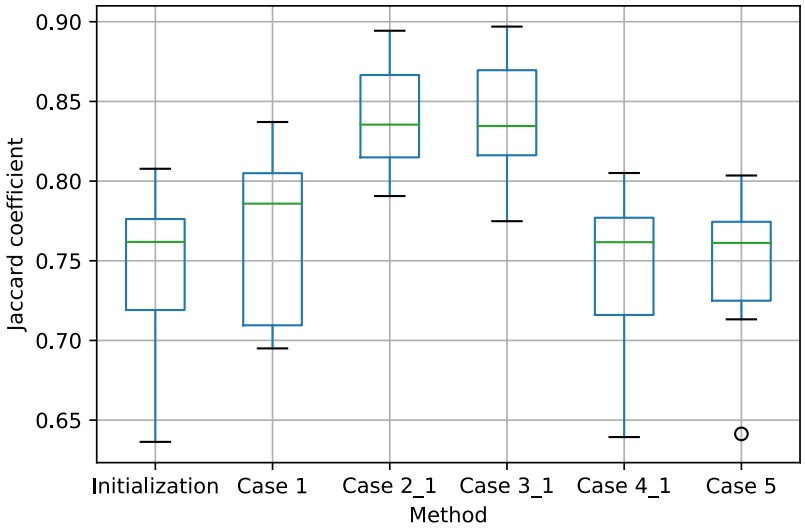


Figure 5.8: Jaccard index. Best case corresponds to higher values

According to the visualisation of each metric, the best results were obtained in case 2 and 3 as it proves all metrics. Despite case 3 had best values in false negatives metric, case 2 was chosen for the tuning of parameters because no mask was used.

In the following section, tuning of some parameters used to perform case 2 is going to be developed in order to achieve better results. Also, the execution time will be taken into account because registration should be a routine procedure in a clinical environment.

5.2 Tuning registration parameters

As it was explained in the section before, the registration method of case 2 (no mask, non-rigid registration) was selected to improve the registration via parameter tuning. That section is divided in two parts: the first one is related with parameters that were tuned, and the second one is about values obtained by the application of tuned parameter file.

5.2.1 Parameter file setting

After the study of the optimal use of mask done in the section before, next step is to evaluate which parameters are sensitive to tune in order to obtain a better registration.

For that reason, as it was explained in subsection 4.2.4, two parameters were chosen to achieve an improvement of registration: metric, and the final grid spacing, which are remarked in Figure 5.9.

Initially, advanced mattes mutual information was used as a single metric, and then, a transform bending energy penalty was added as a second one. When more than one metric appears, a new parameter (weight of the metric) should be created in order to control the influence of each one in the registration. Consequently, it was assigned a weight value for each metric.

Figure 5.9: Parameter file. Parameters tuned remarked in blue colour

NON RIGID	BASE CASE	EX 1	EX 2	EX 3
Image Types				
Fixed Internal Image Pixel Type	short	short	short	short
MovingInternalImagePixelType	short	short	short	short
UseDirectionCosines	TRUE	TRUE	TRUE	TRUE
Main components:				
Registration	MultiResolutionRegistration	MultiMetricMultiResolutionRegistration	MultiMetricMultiResolutionRegistration	MultiMetricMultiResolutionRegistration
Interpolator	BSplineInterpolator	BSplineInterpolator	BSplineInterpolator	BSplineInterpolator
ResampleInterpolator	FinalBSplineInterpolator	FinalBSplineInterpolator	FinalBSplineInterpolator	FinalBSplineInterpolator
Resampler	DefaultResampler	DefaultResampler	DefaultResampler	DefaultResampler
FixedImagePyramid	FixedRecursiveImagePyramid	FixedRecursiveImagePyramid	FixedRecursiveImagePyramid	FixedRecursiveImagePyramid
MovingImagePyramid	MovingRecursiveImagePyramid	MovingRecursiveImagePyramid	MovingRecursiveImagePyramid	MovingRecursiveImagePyramid
Optimizer	AdaptiveStochasticGradientDescent	AdaptiveStochasticGradientDescent	AdaptiveStochasticGradientDescent	AdaptiveStochasticGradientDescent
Transform	BSplineTransform	BSplineTransform	BSplineTransform	BSplineTransform
Metric	AdvancedMattesMutualInformation	AdvancedMattesMutualInformation	AdvancedMattesMutualInformation	AdvancedMattesMutualInformation
		TransformBendingEnergyPenalty	**TransformBendingEnergyPenalty**	**TransformBendingEnergyPenalty**
		Weight0 1	**Weight0 1**	**Weight0 1**
		Weight1 x	**Weight1 x**	**Weight1 x**
Transformation:				
AutomaticScalesEstimation	TRUE	TRUE	TRUE	TRUE
FinalGridSpacingInPhysicalUnits	15.0	**15.0**	**10.0**	**5.0**
HowToCombineTransforms	Compose	Compose	Compose	Compose
Similarity measures				
NumberOfHistogramBins	32	32	32	32
ErodeMask	FALSE	FALSE	FALSE	FALSE
Multiresolution				
NumberOfResolutions	4	4	4	4
ImagePyramidSchedule	8 8 8 4 4 4 2 2 2 1 1 1	8 8 8 4 4 4 2 2 2 1 1 1	8 8 8 4 4 4 2 2 2 1 1 1	8 8 8 4 4 4 2 2 2 1 1 1
Optimizer				
MaximumNumberOfIterations	2000	2000	2000	2000
Image sampling				
NumberOfSpatialSamples	2048	2048	2048	2048
NewSamplesEveryIteration	TRUE	TRUE	TRUE	TRUE
ImageSampler	Random	Random	Random	Random
Interpolation and Resampling				
BSplineInterpolationOrder	1	1	1	1
FinalBSplineInterpolationOrder	3	3	3	3
DefaultPixelValue	0	0	0	0
WriteResultImage	TRUE	TRUE	TRUE	TRUE
ResultImagePixelType	short	short	short	short
ResultImageFormat	mhd	mhd	mhd	mhd

As Figure 5.9 shows, a constant value was assigned as a weight for the advance mattes mutual information (AMMI). During experiments, AMMI was kept as 1, whereas transform bending energy penalty was tuned from 0 to 1 in a first step, and from 1 to 10, in a second step. AMMI could have been fixed in other value, but this one was used based on parameter files of Staring *et al.* [18]

5.2.2 Quantitative results

Experiment 1

In Experiment 1 (EX 1 in Figure 5.9), a constant value of 15.0 mm was chosen as final grid spacing. That is the most rough approximation, equivalent to first picture of Figure 4.7. For that reason, in terms of time, it was the fastest execution because the algorithm had to take less points per grid than in following

experiments.

Keeping this parameter constant, the weight of the transform bending energy penalty was altered between 0 and 1. These values were chosen in order to arrive to the same range of values than used in the weight of AMMI.

In order to have a control of which values are assigned to each metric, and to clear understand how change according to the final grid spacing, some detailed tables were elaborated.

This document was composed by different parts: a description of the experiments, plots and detailed tables. The experiment table was divided in three sections, one per experiment. Table 5.3 visualizes the results of the dice coefficients using different weights. Each value is compared with the base case in order to know if it is improving or even decreasing the registration efficiency with each experiment. The experiment document had a table for each metric (false negatives, false positives, jaccard and smothness) that are included in appendix B.

In general, it is possible to show that dice obtained using a metric of 0, 0.3 or 0.9 are better than the rest. However, best value depends on the patient. For example, for patient 21 best value is 0.9 whereas for patient 22 is better the base case.

Table 5.3: Dice metric. weight tuning for a grid spacing of 15.0 mm

weight 1-weight X / FinalGridSpacingInPhysicalUnits 15.0							
DICE	BASE CASE	0	0.1	0.3	0.7	0.9	1
Patient 1	0,9018	0,9043	0,9077	0,9073	0,9019	0,9109	0,9064
Patient 2	0,9257	0,9268	0,9273	0,9275	0,9269	0,9268	0,9240
Patient 3	0,9313	0,9252	0,9277	0,9309	0,9269	0,9287	0,9273
Patient 4	0,8942	0,8962	0,8975	0,8971	0,8909	0,8945	0,8917
Patient 5	0,9210	0,9224	0,9215	0,9226	0,9223	0,9228	0,9208
Patient 6	0,9103	0,9102	0,9100	0,9133	0,9096	0,9103	0,9126
Patient 7	0,9442	0,9452	0,9459	0,9454	0,9449	0,9457	0,9455
Patient 8	0,9024	0,9000	0,9020	0,9011	0,8994	0,9060	0,8963
Patient 9	0,8831	0,8804	0,8777	0,8787	0,8806	0,8787	0,8791

Experiment 2

It is not possible to know which is the best result just with one grid spacing test because registration depends on the combination of the weight and the grid spacing value. For that reason, a second experiment (EX 2 in Figure 5.9) was done using a final grid spacing of 10.0 mm. That change induced a significant execution time rise due to the decrease of the grid spacing. However, that also carries an increment in the registration quality what would be supported by the different metrics.

In that case, more weights were used in order to find the best combination. As has been done in the previous section, a table with values is added as an example of the rest metrics (Table 5.4)

Table 5.4: Dice metric. weight tuning for grid spacing of 10.0 mm

		weight 1-weight X / FinalGridSpacingInPhysicalUnits 10.0								
DICE	BASE CASE	0	0.1	0.3	0.7	0.9	1	2.5	5	7.5
Patient 1	0,9018	0,9116	0,9105	0,9104	0,9106	0,9104	0,9105	0,9123	0,9125	0,9110
Patient 2	0,9257	0,9341	0,9304	0,9316	0,9316	0,9323	0,9297	0,9311	0,9320	0,9305
Patient 3	0,9313	0,9340	0,9315	0,9351	0,9328	0,9330	0,9337	0,9347	0,9352	0,9344
Patient 4	0,8942	0,9015	0,9009	0,9029	0,9030	0,9045	0,9029	0,9044	0,9040	0,9039
Patient 5	0,9210	0,9238	0,9230	0,9240	0,9242	0,9230	0,9235	0,9227	0,9224	0,9237
Patient 6	0,9103	0,9112	0,9094	0,9105	0,9111	0,9153	0,9116	0,9137	0,9088	0,9107
Patient 7	0,9442	0,9466	0,9481	0,9474	0,9475	0,9474	0,9484	0,9484	0,9469	0,9475
Patient 8	0,9024	0,8972	0,8988	0,8957	0,8953	0,8967	0,8945	0,8945	0,8938	0,9010
Patient 9	0,8831	0,8837	0,8799	0,8794	0,8818	0,8800	0,8806	0,8806	0,8835	0,8857

Experiment 3

Experiment 3 (EX 3 in Figure 5.9) was carried out with a final grid spacing of 5.0 mm, what should result in the best results compared to the previous experiments. The execution time, however, is increased exponentially. For that reason, experiment 3 was done just in patient 1 and 2 as Table 5.5 shows.

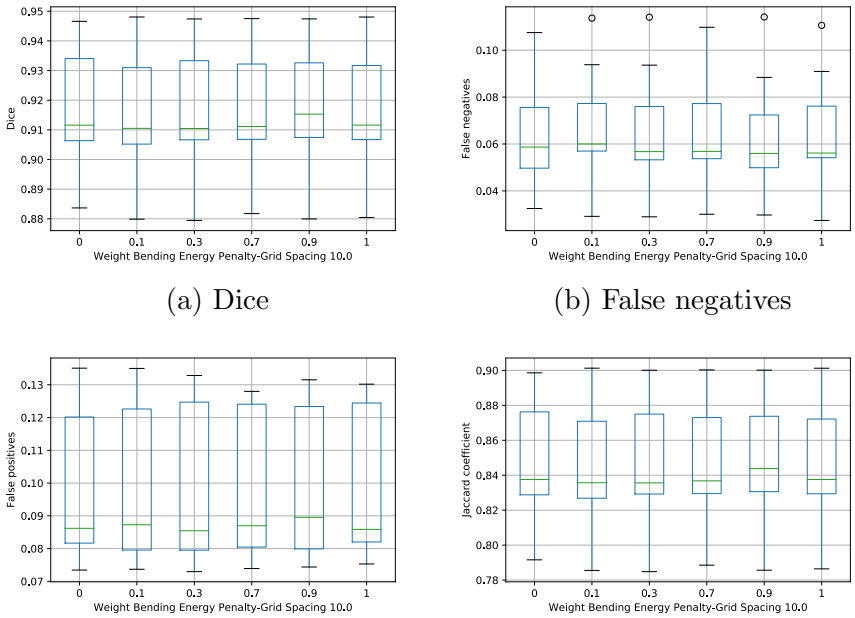
Table 5.5: Dice metric. weight tuning for a grid spacing of 5.0 mm

		weight 1-weight X / FinalGridSpacingInPhysicalUnits 5.0								
DICE	BASE CASE	0	0.1	0.3	0.7	0.9	1	2.5	5	7.5
Patient 1	0,9018	0,9136	0,9140	0,9158	0,9140	0,9125	0,9142	0,9127	0,9147	0,9147
Patient 2	0,9257	0,9362	0,9346	0,9351	0,9343	0,9349	0,9369	0,9348	0,9348	0,9332

As demonstrated, a decreasing grid size result in an improved accuracy in most of patients, but not in case of weight values. That means that an increase of weight given to the transform bending energy is not directly related to an improvement of the registration according to dice and jaccard coefficient.

In case of false positives and false negatives, the behaviour is completely opposite because a decrease of values corresponds to an increase of the quality of the registration, but the conclusion is the same.

The same tendencies are shown when we average the results over all patients, that allows to get a general statement. For that reason, a boxplot for each metric was created in order to visualise the general behaviour by different measures.



(a) Dice

(b) False negatives

(c) False positives

(d) Jaccard

Figure 5.10: Box-and-whisker plots case 2 of grid spacing of 10.0 mm for each weight

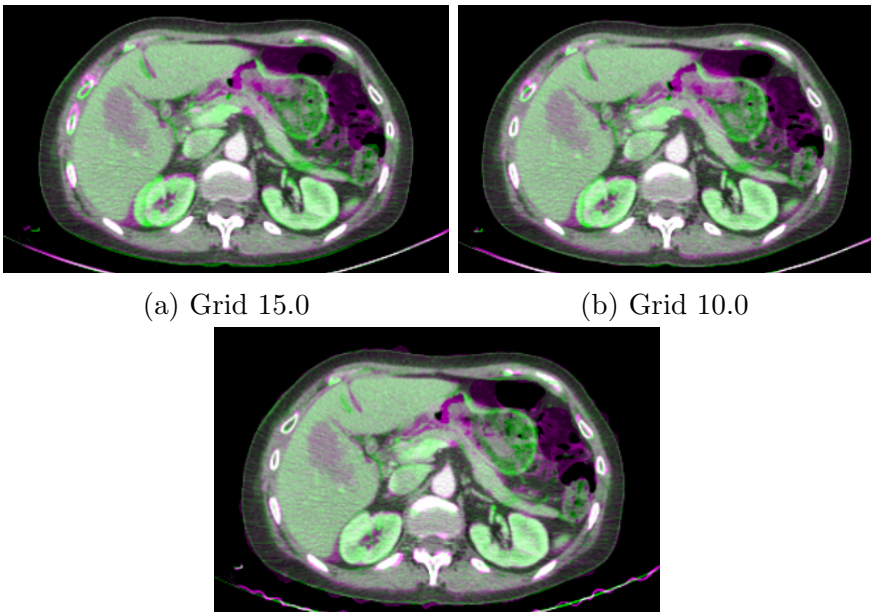
None of these graphics (Figure 5.10) can emphasise an specific weight that improves registration in a general way. It is appreciable how the distribution values for each weight is too large due to the dependence of characters of each patient for that level of registration improvement.

5.2.3 Qualitative results

Grid spacing evaluation

Unfortunately, results depend too much on the patient. For that reason, some of them are going to be analysed. This analysis is going to be done in a transverse, but also in a coronal and sagittal plane. In the following images green colour part corresponds to the moving image (image before the ablation) and the pink colour to the fixed image (image after the ablation).

As patient 2 is one of the most convenient to perform a registration correctly, each patient is going to be compared with it to proof how dependable the registration process is on the patient. In a first visualisation, comparison between grid spacing is going to be done for the same weight. For that reason, patient 1 and 2 are going to be showed because they have results of grid 5.0 as well.

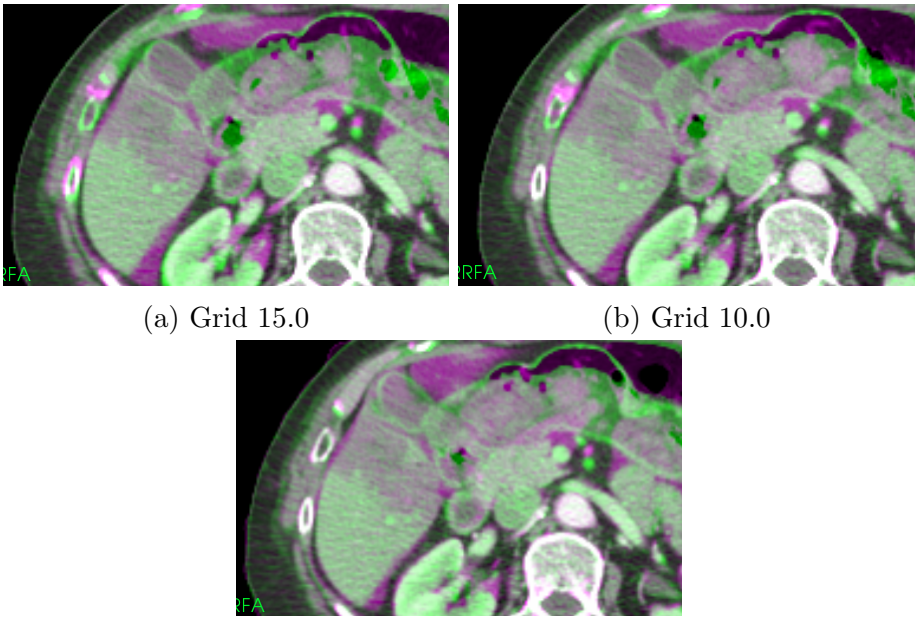


(a) Grid 15.0 (b) Grid 10.0

(c) Grid 5.0

Figure 5.11: Grid 15.0, 10.0 and 5.0 mm with a weight 0 in patient 2. Focussing in ribs.

Figure 5.11 shows that a decreasing grid space will optimize the placement of the ribs. Figure 5.13b, that corresponds to a grid spacing of 10.0 mm, is not completely matched whereas 5.13c that is the grid spacing of 5.0 mm is almost completely registered.



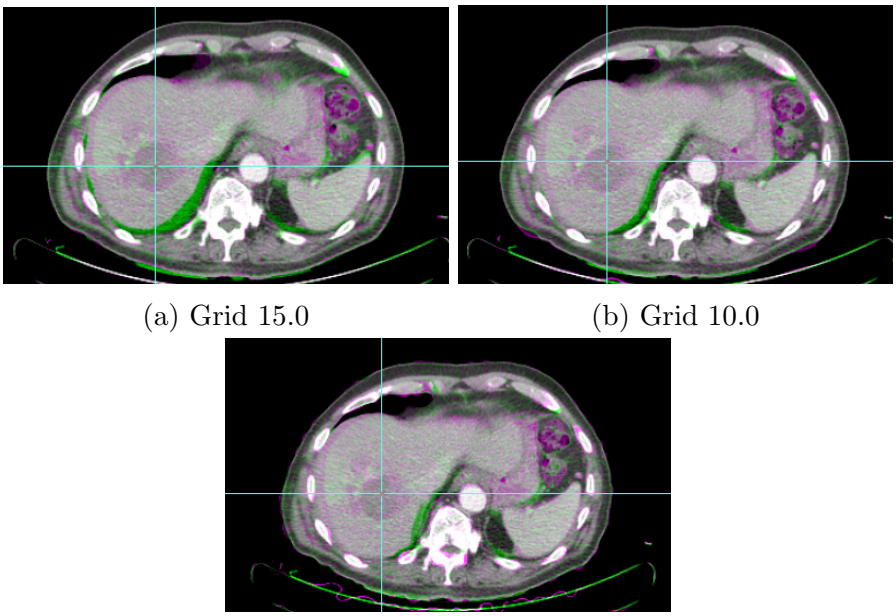
(a) Grid 15.0 (b) Grid 10.0

(c) Grid 5.0

Figure 5.12: Grid 15.0, 10.0 and 5.0 with a weight 0 in Patient 2. Visualising tumor region.

In both, Figure 5.11 and 5.12, there is a clearly difference in ribs matching, but in Figure 5.12 it is possible to see the tumor (green region) inside the ablation zone (purple region) without deformation of any of them.

Regarding patient 1, Figure 5.13 shows how the matching of the liver edges improves according to the decrease of the grid spacing value.



(a) Grid 15.0 (b) Grid 10.0

(c) Grid 5.0

Figure 5.13: Grid 15.0, 10.0 and 5.0 with a weight 0 in Patient 1.

Transform bending energy weight evaluation

In a second analysis, the weight significance is going to be showed per patient, in order to appreciate how changes are exclusive for each patient. A table with best weight per patient is going to be compared between grid spacing 15.0 and 10.0:

- Grid spacing 15.0.
  Surprisingly, there are some patients that obtained better results according to dice using just the base case than using the transform bending energy penalty for a grid spacing of 15.0 as is the case of patient 3 and 9 (Table 5.6).

Table 5.6: Patients best results for grid spacing of 15.0

Grid 15	**Pat 1**	**Pat 2**	**Pat 3**	**Pat 4**	**Pat 5**	**Pat 6**	**Pat 7**	**Pat 8**	**Pat 9**
weight X	0.9	0.3	Base case	0.1	0.9	0.3	0.1	0.9	Base case

- Grid spacing 10.0.
  Table 5.7 shows that, even in some patients best weight for a grid spacing of 10.0 match, as patient 4 and 6 or 8 and 9, it is not possible to gain a general weight suitable for all dataset.

Table 5.7: Patients best results for grid spacing of 10.0

Grid 10	**Pat 1**	**Pat 2**	**Pat 3**	**Pat 4**	**Pat 5**	**Pat 6**	**Pat 7**	**Pat 8**	**Pat 9**
weight X	5	0	5	0.9	0.7	0.9	2.5	7.5	7.5

This comparison allows to understand how best weight values change for each patient even using the same grid spacing. Also, it shows the challenge of choosing one for all of them.

In addition, the best result for each patient is going to be analysed based on dice coefficient because results are similar to jaccard index. In case of the first two patients, grid 5 is included in this analysis. It is not possible to focus on all metrics because, even having similar results, are not exactly the same. For example, the best case according to dice for patient 1 is using grid spacing of 5 and weight of 0.3, but according to false positives, is 0.1. However, the difference between false positives value of weight 0.1 and 0.3 is small as it is showed in Table 5.8.

Table 5.8: Comparison between two better results of false positives in patient 1

weight 1-weight X / FinalGridSpacingInPhysicalUnits 5.0			
FP	BASE CASE	0.1	0.3
Patient 1	0,1398	0,1099	0,1091

In case of patient 1, best registration results according to dice, corresponds to a grid spacing of 5.0 and a weight of 0.3 (Figure 5.14).

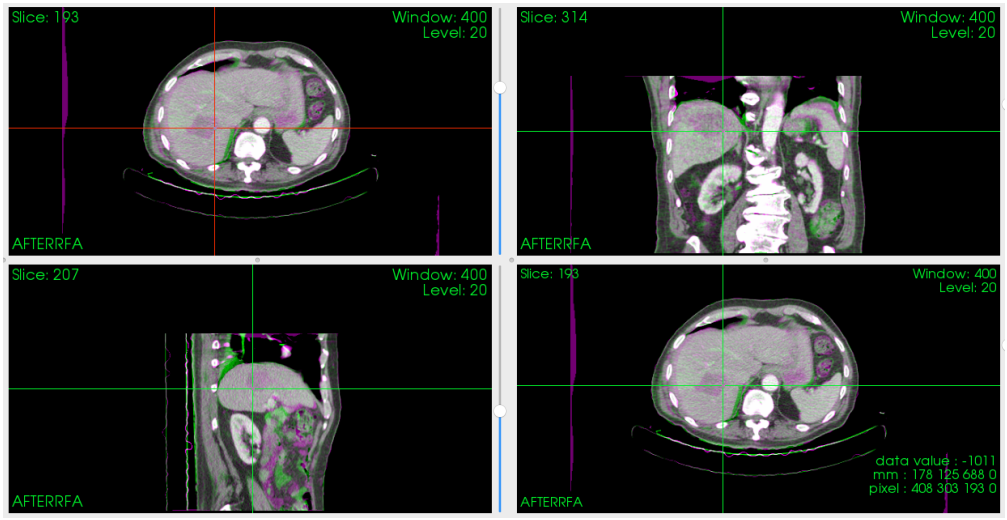


Figure 5.14: Best registration of patient 1

However, for patient 2, best registration results according, corresponds to a grid spacing of 5.0 and a weight of 1, as it is showed in Figure 5.15.

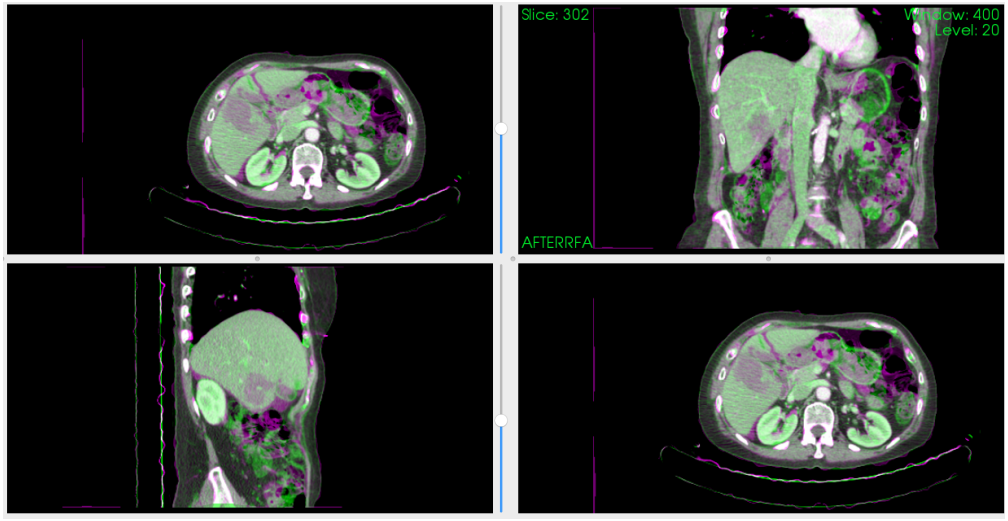


Figure 5.15: Best registration of patient 2

In patient 3 (Figure 5.16), it is possible to show a correct registration for a grid spacing of 10.0 and a weight of 5, even it was difficult to match some voxels because of the movement in the y direction.

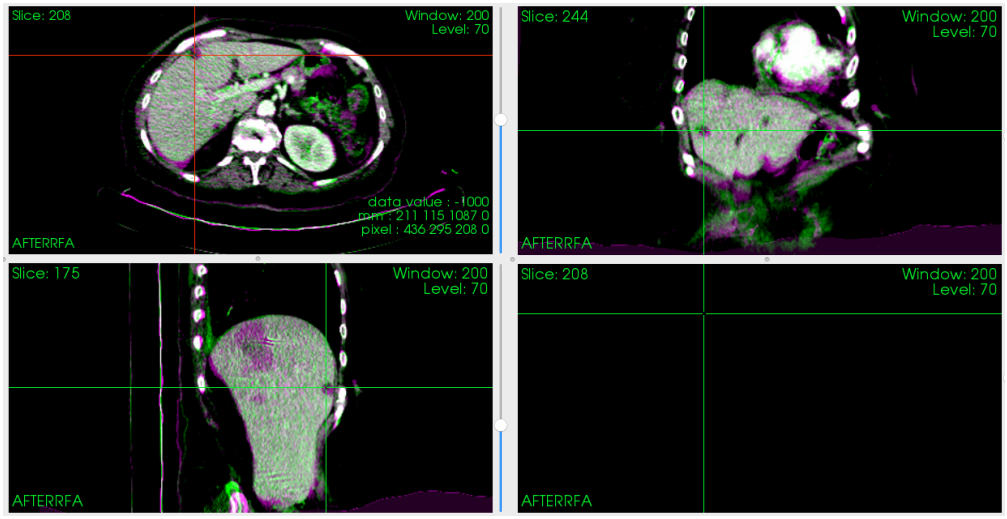


Figure 5.16: Best registration of patient 3

Patient 4 was a good case to prove the registration method due to the difficulty added by the dark space that is showed in Figure 5.17 that may correspond to an previous RFA.

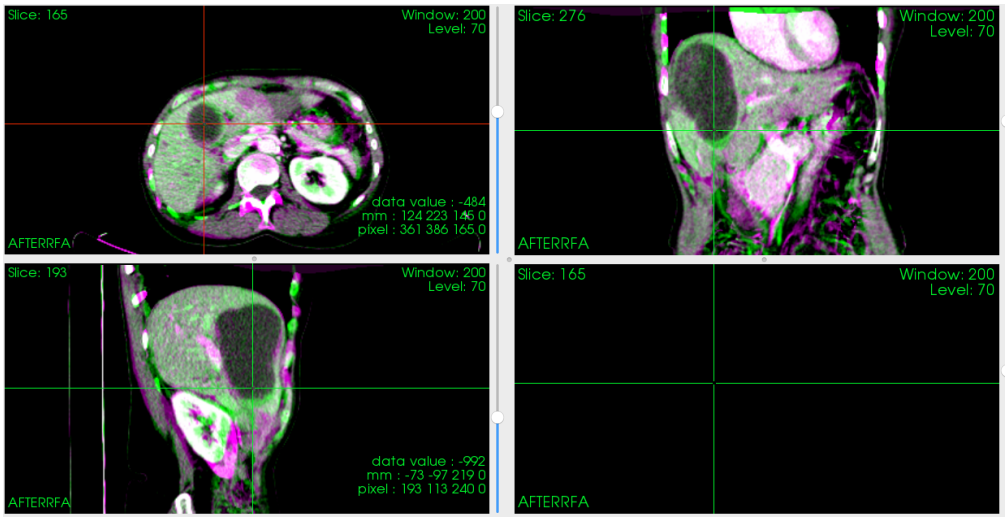


Figure 5.17: Best registration of patient 4

Registration process was almost completely successful for patient 5, due to the easily matching from the beginning of the rigid transformation (Figure 5.18).

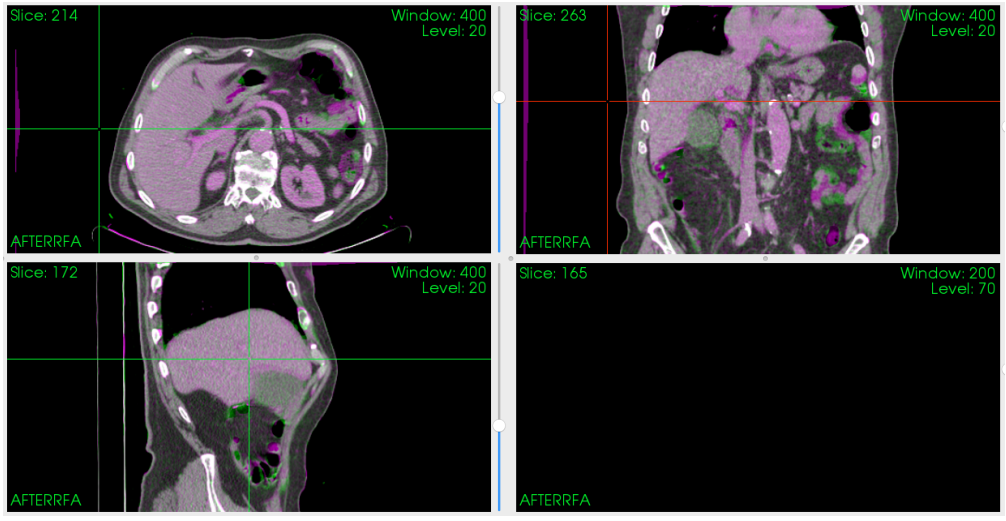


Figure 5.18: Best registration of patient 5

Liver of patient 6 was almost aligned just with the initialization process but not the column. For that reason, it was a hard case (Figure 5.19).

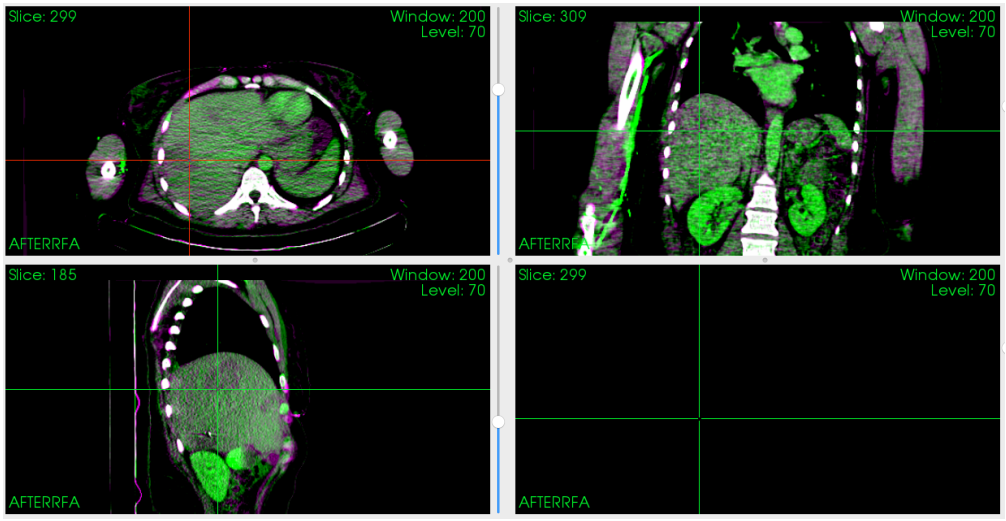


Figure 5.19: Best registration of patient 6

This patient was used in previous sections because of its perfect alignment that made successful registration (Figure 5.20).

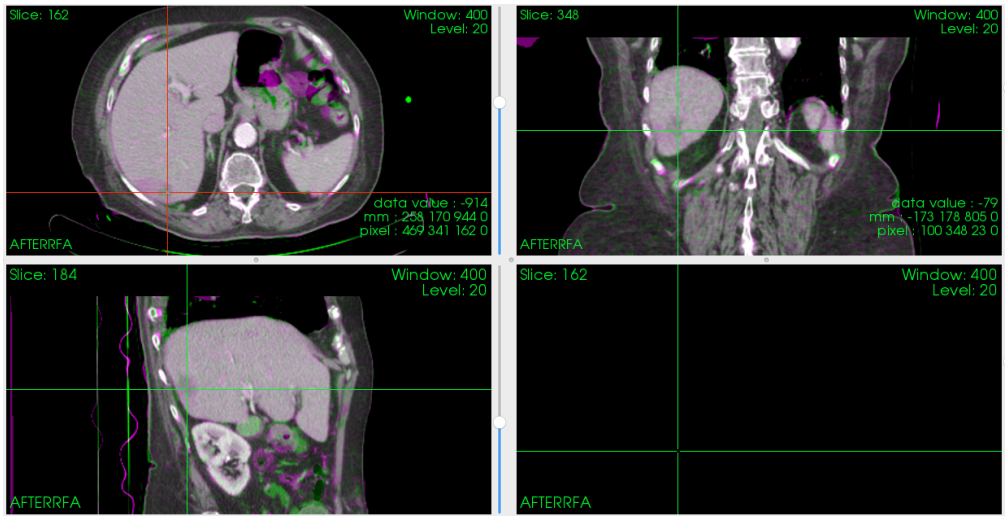


Figure 5.20: Best registration of patient 7

Patient 8 (Figure 5.21) required more execution time due to the different positions of the patient in images before and after the ablation procedure. However, registration was performed almost correctly.

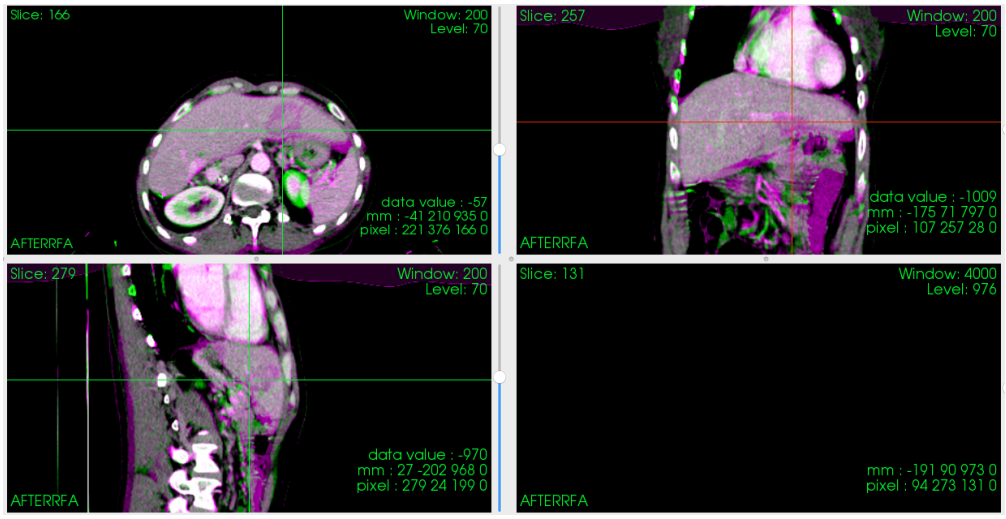


Figure 5.21: Best registration of patient 8

This case was almost impossible to align because patient 22 seemed to have liquid in one lung in the image after the ablation, but not in the image before. This means that, during registration, some element was in the fixed image that was not possible to match to any element in the moving image because it does not exist in it. However, some good results were obtained close to the liver (Figure 5.22).

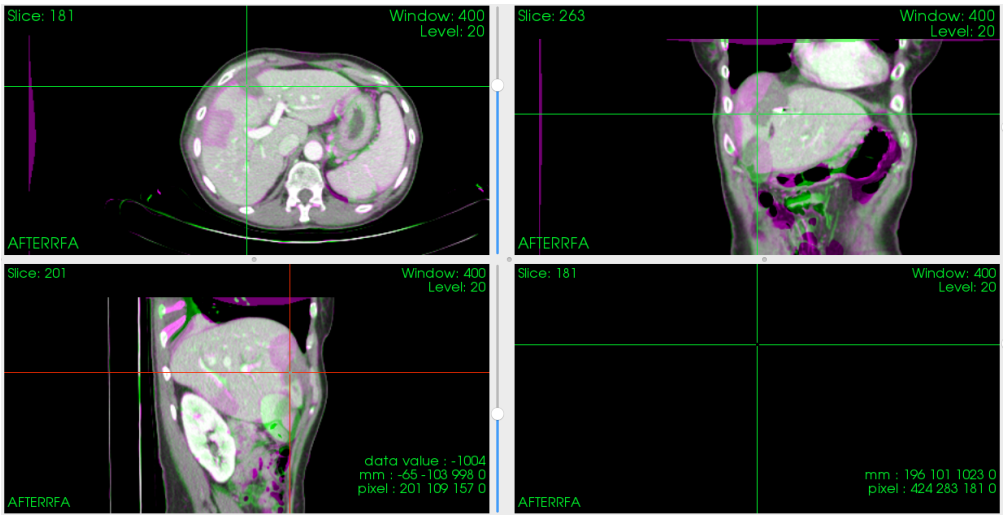


Figure 5.22: Best registration of patient 9

Chapter 6

Discussion

A registration process composed by a rigid and a non-rigid transformation was performed on 9 patients. The influence of masks was studied in order to determine which ones were necessary for the registration process.

Masks can be used in many different ways as the literature shows. For that reason, in the first step of this project, some papers were analysed, in order to find different methods using masks to apply and evaluate them. 5 cases were developed, some of them based on Luu *et al.* [8] and Staring *et al.* [18] papers, as case 4 and 5.

After a quantitatively evaluation of each of them, case 2 was chosen as the best case for several reasons. It was performed without any mask using a non-rigid transformation after the initialization. It had good results according to Dice coefficient, False negatives, False positives and Jaccard coefficient as it shows Figure 5.2, 5.6, 5.7 and 5.8. It was not possible to compare these results to other research studies because no other was found without using any mask. Even case 3 had competent results as well, it was required to use an ablation and a tumor mask, what makes registration dependent on the quality of the segmentation. Moreover, the creation of these masks required manual segmentation what supposed a lot of time.

Using a non-rigid registration with a local rigid deformation equivalent to case 4, and a B-spline grid size of 5 mm, Luu *et al.* [8] obtained a Dice coefficient of 92.2% . In that project, a Dice of 94.4% was obtained for the best patient (Patient 7) using method of case 2 with a grid spacing of 15.0 mm. Luu *et al.* achieved better differences between tumor and ablation zone due to the use of a rigidity penalty around the tumor. It was also relevant the quality of images that they used.

In addition, the use of a higher amount of patients supposes a difference between their research and this current project. It makes more difficult the decision-making

according to all of them, but also, it gives the advantage to probe better the method because of taking into account more cases.

The evaluation of the use of masks in the different registration methods leads to the study of other relevant parameters. Due to the importance of the grid spacing value and its dependence on the weight assigned to the transform bending energy penalty, a second study was focussed on the analysis of these parameters.

In a first step, grid spacing was kept constant to 15.0 mm in order to study the influence of the weight assigned to the transform bending energy penalty in registration for each patient. After tuning the value of the weight from 0 to 10 it was realised that the variation of this value has not influence in all patients in the same way. Consequently, grid spacing was increased to 10.0 and then to 5.0 in order to probe what was obtained from the first evaluation of the weight. Whereas an increment of the quality of registrations were appreciated according to the decrease of the grid spacing value, a non defined behaviour of registration was observed according to the weight variation.

Parameters were evaluated using four different metrics: Dice, false negatives, false positives, jaccard coefficient and smoothness. Last of them was used to have an overview of the deformation in each voxel. In addition, a qualitative evaluation was performed focussing on each patient.

The examination of different mask combinations and according to the results, the proposed registration method is composed by an initialization, followed by a non-rigid B-spline registration without using any mask.

For a grid spacing of 15.0 mm and a weight of bending energy penalty from 0 to 1, patients 1, 5 and 8 obtained the best result using a weight of 0.9, patients 2 and 6 using 0.3, and patient 4 and 7 using 0.1, according to dice. However, for a grid spacing of 10.0 mm and a weight from 0 to 10, patient 1 and 3 obtained best results using a weight of 5, patient 4 and 6 using a weight of 0.9, patient 8 and 9 using 7.5, patient 2 using 0, patient 5 using 0.7 and patient 7, 2.5.

After the evaluation of the grid spacing in all patients (5.0 mm just in patient 1 and 2) and the weight of the penalty, it was found that each patient has different optimal weight according to the relation with each grid. However, the best result was obtained decreasing grid spacing in almost all patients.

6.1 Limitations of the study

There are some limitations to comment according to the project. Firstly, the dataset images had not the competent quality to perform correctly segmentation and registration. For that reason, it was so difficult to distinguish tumor and ablation zone in some images. Secondly, segmentations should have been provided by an specialist to assure the correct election of the method. This will allow to perform a more accurate evaluation of each case based on medical knowledge. Thirdly and more important, the memory capacity of the computers used was insufficient to execute registration. In some of the cases, days were spent to execute an specific case for all patients.

6.2 Suggestions for future work

The method proposed in this project could be improved in some aspects. Despite the fact that, it is possible to avoid the use of masks in the registration process, it is not possible in the validation. Masks are required in order to have the ground truth to obtain an overlapping metric. For that reason, the selection of other metrics to evaluate the registration method would be very interesting and useful to achieve the complete deletion of the use of masks.

Regarding the rough approach experiments of the project, an interesting point would be to introduce a new case using an ablation zone rigidity penalty instead tumor penalty. That improvement will allow to fix the ablation zone, what hopefully includes the tumor region.

Finally, I suspect significant differences using a grid spacing of 5.0 mm to all patients and if it is possible, to continue decreasing this value in order to obtain the best case possible according to this parameter.

Chapter 7

Conclusion

Liver cancer is one of the most harmful cancer that exists. Despite there are different options to treat it, radiofrequency ablation (RFA) is the most efficient and it is available to almost all patients. In order to know if the tumor has been removed, it is required to visualise images before and after the ablation procedure, and to compare them. Nowadays, visual evaluation is done manually by the interventional radiologist (IR) what induces an imprecision in the results due to the subjectivity added.

Until now, there are several studies that investigated different ways to automate this process. Some authors try to automatically align both images using different methods [5], [12], [9], or tools [15], and also trying to remove the use of some masks of the registration process [14].

This project was inspired by Luu *et al.* who developed an automatic registration of the images pre- and post- liver tumor ablation using *elastix* [15], [8].

A registration approach consisting of multiple transformations is proposed in this thesis using different combinations of masks. According to the evaluation performed by dice coefficient, false negatives, false positives and jaccard, best method was case 2, which performs an initialization, followed by a non-rigid B-spline registration without using liver, tumor or ablation mask. This method was chosen to be improved it in the next step.

An specific approach was done in order to study the influence of some parameters using the registration method of case 2. The balance between the grid spacing value and the weight assigned to the transform bending energy penalty was evaluated keeping one of them constant and tuning the other one. A grid spacing of 15.0, 10.0 and 5.0 mm was tuned for a weight from 0 to 10.

After a qualitative and quantitative evaluation, a significant influence of grid spacing on the improvement of registration was achieved by decreasing the value. However, non general relation was showed tuning weight value in the registration quality. That means that for each patient the relation between weight value and the registration quality was different. Consequently, a significant improvement of the registration has been achieved avoiding the use of masks and, the best case was obtained decreasing the grid spacing value until 5.0 mm.

Appendices

Appendix is divided in two sections: Appendix A contains results of general approach for case 2, 3, 4 and 5. Appendix B contains registration results of case 2 for each weight and final grid spacing combination according to false negatives, false positives and jaccard index.

Appendix A

Appendix A

General approach for case 2, 3, 4 and 5.

Table A.1: Results general approach case 2

	Dice	FN	FP	Jaccard	Smoothness
Patient 1	0,9018	0,0524	0,1398	0,8211	0,1145
Patient 2	0,9257	0,0500	0,0973	0,8617	0,1153
Patient 3	0,9313	0,0478	0,0887	0,8715	0,1651
Patient 4	0,8942	0,0883	0,1226	0,8086	0,1118
Patient 5	0,9210	0,0375	0,1170	0,8536	0,0815
Patient 6	0,9103	0,0783	0,1007	0,8354	0,1918
Patient 7	0,9442	0,0269	0,0830	0,8944	0,1474
Patient 8	0,9024	0,1393	0,0516	0,8222	
Patient 9	0,8831	0,0998	0,1334	0,7906	0,1086

Table A.2: Results general approach case 3

	Dice	FN	FP	Jaccard	Smoothness
Patient 1	0,9002	0,0485	0,1458	0,8185	0,1138
Patient 2	0,9282	0,0467	0,0956	0,8660	0,1238
Patient 3	0,9323	0,0327	0,1002	0,8732	0,2815
Patient 4	0,8974	0,0876	0,1172	0,8139	0,1170
Patient 5					
Patient 6	0,9098	0,0820	0,0983	0,8345	0,1970
Patient 7	0,9457	0,0246	0,0822	0,8970	0,1328
Patient 8					
Patient 9	0,8731	0,1241	0,1296	0,7748	0,1278

Table A.3: Results general approach case 4

	Dice	FN	FP	Jaccard
Patient 1	0,8690	0,0651	0,1883	0,7683
Patient 2	0,8800	0,1025	0,1368	0,7857
Patient 3	0,8202	0,1457	0,2112	0,6952
Patient 4	0,7800	0,1709	0,2636	0,6394
Patient 5	0,9082	0,0696	0,1131	0,8318
Patient 6	0,8647	0,1648	0,1036	0,7617
Patient 7	0,8920	0,0498	0,1594	0,8051
Patient 8				
Patient 9	0,8484	0,1308	0,1714	0,7368

Table A.4: Results general approach case 5

	Dice	FN	FP	Jaccard
Patient 1	0,8668	0,0675	0,1903	0,7649
Patient 2	0,8790	0,1036	0,1378	0,7840
Patient 3	0,8326	0,1328	0,1993	0,7132
Patient 4	0,7815	0,1693	0,2622	0,6414
Patient 5	0,9088	0,0689	0,1125	0,8328
Patient 6	0,8644	0,1651	0,1039	0,7612
Patient 7	0,8911	0,0508	0,1604	0,8035
Patient 8				
Patient 9	0,8484	0,1308	0,1715	0,7366

Appendix B

Appendix B

Registration results of case 2 for each weight and final grid spacing combination according to false negatives, false positives, jaccard index and smoothess.

Table B.1: False negatives metric. Weight tuning for a grid spacing of 15.0 mm

Weight 1-Weight X / FinalGridSpacingInPhysicalUnits 15.0

FN	BASE CASE	0	0.1	0.3	0.7	0.9	1
Patient 1	0,0524	0,0522	0,0553	0,0534	0,0529	0,0503	0,0504
Patient 2	0,0500	0,0514	0,0490	0,0454	0,0487	0,0490	0,0501
Patient 3	0,0478	0,0573	0,0521	0,0546	0,0586	0,0536	0,0551
Patient 4	0,0883	0,0381	0,0393	0,0391	0,0367	0,0359	0,0390
Patient 5	0,0375	0,0381	0,0393	0,0391	0,0367	0,0359	0,0390
Patient 6	0,0783	0,0875	0,0812	0,0735	0,0864	0,0825	0,0779
Patient 7	0,0269	0,0262	0,0275	0,0282	0,0296	0,0304	0,0273
Patient 8	0,1393	0,1281	0,1326	0,1327	0,1383	0,1232	0,1449
Patient 9	0,0998	0,1064	0,1118	0,1119	0,1074	0,1129	0,1117

Table B.2: False positives metric. Weight tuning for a grid spacing of 15.0 mm

Weight 1-Weight X / FinalGridSpacingInPhysicalUnits 15.0							
FP	BASE CASE	0	0.1	0.3	0.7	0.9	1
Patient 1	0,1398	0,1354	0,1265	0,1288	0,1392	0,1248	0,1331
Patient 2	0,0973	0,0941	0,0953	0,0981	0,0963	0,0961	0,1005
Patient 3	0,0887	0,0917	0,0916	0,0831	0,0871	0,0883	0,0897
Patient 4	0,1226	0,1197	0,1163	0,1219	0,1235	0,1189	0,1228
Patient 5	0,1170	0,1140	0,1145	0,1128	0,1153	0,1151	0,1162
Patient 6	0,1007	0,0921	0,0986	0,0995	0,0944	0,0968	0,0966
Patient 7	0,0830	0,0818	0,0794	0,0797	0,0793	0,0771	0,0803
Patient 8	0,0516	0,0700	0,0604	0,0623	0,0594	0,0628	0,0583
Patient 9	0,1334	0,1324	0,1325	0,1304	0,1310	0,1296	0,1299

Table B.3: Jaccard index. Weight tuning for a grid spacing of 15.0 mm

Weight 1-Weight X / FinalGridSpacingInPhysicalUnits 15.0							
Jaccard	BASE CASE	0	0.1	0.3	0.7	0.9	1
Patient 1	0,8211	0,8253	0,8310	0,8304	0,8213	0,8364	0,8288
Patient 2	0,8617	0,8635	0,8644	0,8648	0,8638	0,8637	0,8587
Patient 3	0,8715	0,8608	0,8652	0,8708	0,8638	0,8669	0,8645
Patient 4	0,8086	0,8120	0,8140	0,8135	0,8033	0,8092	0,8045
Patient 5	0,8536	0,8560	0,8545	0,8563	0,8558	0,8567	0,8533
Patient 6	0,8354	0,8352	0,8349	0,8405	0,8342	0,8353	0,8393
Patient 7	0,8944	0,8961	0,8973	0,8964	0,8955	0,8969	0,8966
Patient 8	0,8222	0,8182	0,8215	0,8201	0,8172	0,8281	0,8121
Patient 9	0,7906	0,7863	0,7821	0,7837	0,7867	0,7836	0,7843

Table B.4: False negatives metric. Weight tuning for grid spacing of 10.0 mm

		Weight 1-Weight X / FinalGridSpacingInPhysicalUnits 10.0								
FN	BASE CASE	0	0.1	0.3	0.7	0.9	1	2.5	5	7.5
Patient 1	0,0524	0,0598	0,0582	0,0541	0,0547	0,0560	0,0561	0,0546	0,0577	0,0549
Patient 2	0,0500	0,0468	0,0557	0,0524	0,0527	0,0448	0,0545	0,0499	0,0517	0,0530
Patient 3	0,0478	0,0526	0,0607	0,0567	0,0568	0,0549	0,0538	0,0540	0,0538	0,0529
Patient 4	0,0883	0,0587	0,0600	0,0584	0,0637	0,0564	0,0613	0,0598	0,0605	0,0635
Patient 5	0,0375	0,0399	0,0388	0,0384	0,0393	0,0413	0,0384	0,0393	0,0388	0,0392
Patient 6	0,0783	0,0914	0,0938	0,0936	0,0907	0,0884	0,0909	0,0872	0,0938	0,0926
Patient 7	0,0269	0,0325	0,0291	0,0289	0,0300	0,0297	0,0274	0,0294	0,0294	0,0288
Patient 8	0,1393	0,1385	0,1362	0,1408	0,1409	0,1404	0,1335	0,1448	0,1407	0,1297
Patient 9	0,0998	0,1075	0,1137	0,1141	0,1098	0,1142	0,1106	0,1118	0,1050	0,1023

Table B.5: False positives metric. Weight tuning for grid spacing of 10.0 mm

		Weight 1-Weight X / FinalGridSpacingInPhysicalUnits 10.0								
FP	BASE CASE	0	0.1	0.3	0.7	0.9	1	2.5	5	7.5
Patient 1	0,1398	0,1154	0,1188	0,1225	0,1216	0,1209	0,1205	0,1185	0,1155	0,1206
Patient 2	0,0973	0,0843	0,0830	0,0839	0,0836	0,0896	0,0855	0,0872	0,0838	0,0854
Patient 3	0,0887	0,0791	0,0761	0,0730	0,0773	0,0789	0,0785	0,0763	0,0755	0,0780
Patient 4	0,1226	0,1351	0,1350	0,1328	0,1280	0,1315	0,1302	0,1288	0,1290	0,1265
Patient 5	0,1170	0,1098	0,1123	0,1108	0,1097	0,1101	0,1116	0,1123	0,1133	0,1106
Patient 6	0,1007	0,0862	0,0873	0,0854	0,0870	0,0810	0,0859	0,0853	0,0885	0,0859
Patient 7	0,0830	0,0734	0,0737	0,0751	0,0739	0,0744	0,0753	0,0728	0,0757	0,0750
Patient 8	0,0516	0,0639	0,0632	0,0646	0,0653	0,0630	0,0640	0,0625	0,0688	0,0660
Patient 9	0,1334	0,1250	0,1265	0,1269	0,1265	0,1258	0,1284	0,1268	0,1277	0,1259

Table B.6: Jaccard index. Weight tuning for grid spacing of 10.0 mm

		Weight 1-Weight X / FinalGridSpacingInPhysicalUnits 10.0								
Jaccard	BASE CASE	0	0.1	0.3	0.7	0.9	1	2.5	5	7.5
Patient 1	0,8211	0,8375	0,8357	0,8355	0,8359	0,8355	0,8357	0,8388	0,8391	0,8366
Patient 2	0,8617	0,8763	0,8699	0,8719	0,8719	0,8731	0,8687	0,8710	0,8726	0,8701
Patient 3	0,8715	0,8762	0,8718	0,8781	0,8741	0,8743	0,8756	0,8774	0,8783	0,8769
Patient 4	0,8086	0,8207	0,8197	0,8229	0,8231	0,8256	0,8230	0,8255	0,8248	0,8247
Patient 5	0,8536	0,8584	0,8570	0,8587	0,8590	0,8570	0,8580	0,8565	0,8560	0,8582
Patient 6	0,8354	0,8369	0,8339	0,8356	0,8368	0,8438	0,8376	0,8412	0,8329	0,8361
Patient 7	0,8944	0,8986	0,9013	0,9001	0,9003	0,9001	0,9012	0,9019	0,8991	0,9002
Patient 8	0,8222	0,8136	0,8163	0,8110	0,8105	0,8127	0,8181	0,8091	0,8080	0,8199
Patient 9	0,7906	0,7916	0,7855	0,7848	0,7885	0,7857	0,7864	0,7867	0,7913	0,7949

Table B.7: False negatives metric. Weight tuning for a grid spacing of 5.0 mm

		Weight 1-Weight X / FinalGridSpacingInPhysicalUnits 5.0								
FN	BASE CASE	0	0.1	0.3	0.7	0.9	1	2.5	5	7.5
Patient 1	0,0524	0,0625	0,0608	0,0579	0,0581	0,0606	0,0561	0,0596	0,0579	0,0579
Patient 2	0,0500	0,0465	0,0492	0,0494	0,0496	0,0508	0,0488	0,0492	0,0477	0,0498

Table B.8: False positives metric. Weight tuning for a grid spacing of 5.0 mm

		Weight 1-Weight X / FinalGridSpacingInPhysicalUnits 5.0								
FP	BASE CASE	0	0.1	0.3	0.7	0.9	1	2.5	5	7.5
Patient 1	0,1398	0,1091	0,1099	0,1091	0,1123	0,1129	0,1136	0,1134	0,1111	0,1112
Patient 2	0,0973	0,0806	0,0811	0,0799	0,0812	0,0791	0,0770	0,0806	0,0821	0,0831

Table B.9: Jaccard index. Weight tuning for a grid spacing of 5.0 mm

		Weight 1-Weight X / FinalGridSpacingInPhysicalUnits 5.0								
Jaccard	BASE CASE	0	0.1	0.3	0.7	0.9	1	2.5	5	7.5
Patient 1	0,8211	0,8409	0,8416	0,8446	0,8416	0,8391	0,8420	0,8395	0,8428	0,8428
Patient 2	0,8617	0,8800	0,8772	0,8781	0,8768	0,8777	0,8812	0,8776	0,8776	0,8748

Table B.10: Smoothness. Weight tuning for a grid spacing of 5.0 mm

		Weight 1-Weight X / FinalGridSpacingInPhysicalUnits 5.0								
Smoothness	BASE CASE	0	0.1	0.3	0.7	0.9	1	2.5	5	7.
Patient 1	0,1145	0,4051	0,3730	0,3932	0,3941	0,3903	0,3885	0,3685	0,3737	0,33
Patient 2	0,1153	0,2552	0,2398	0,2337	0,2333	0,2294	0,2233	0,2152	0,2139	0,20

Bibliography

[1] Lee J. M. A, E. H. H. C. C. Clinical outcomes of radiofrequency ablation for early hypovascular hcc: A multicenter retrospective study. *Vascular and interventional radiology*, 286(1):338 349, 2018.

[2] 3rd Chintapalli KN Leyendecker JR Karahan OI Rhim H Chopra S, Dodd GD. Tumor recurrence after radiofrequency thermal ablation of hepatic tumors: spectrum of findings on dual-phase contrast-enhanced ct. *AJR Am J Roentgenol*, 177:381–387, 2001.

[3] N. Buls P. R. Ros F. Vandenbroucke, J. Vandemeulebroucke and J. de Mey. Can tumor coverage evaluated 24h post radiofrequency ablation predict local tumor progression of liver metastases? 2013.

[4] Iezzi R. Di Fabio F. Cianci R. Grassedonio E. & Storto M. L. Filippone, A. Multidetector-row computed tomography of focal liver lesions treated by radiofrequency ablation: spectrum of findings at long-term follow-up. *J Comput Assist Tomogr*, 31(1):42?52, 2007.

[5] Bruners P. Schiffl K. Sedlmair M. Mhlenbruch G. Gnther R. W. et al Keil, S. Radiofrequency ablation of liver metastases-software-assisted evaluation of the ablation zone in mdct: Tumor-free follow-up versus local recurrent disease. *CardioVascular and Interventional Radiology*, 33(2):297?306, 2010.

[6] I. Lara Riera. *Optimizing Lung Registration for CT Ventilation Estimation*. Brussels, Belgium, 2017.

[7] Heo J. S. Cho Y. B. Yun S. H. Kim H. C. Lee W. et al Lee, H. Hepatectomy vs radiofrequency ablation for colorectal liver metastasis: A propensity score analysis. *World Journal of Gastroenterology*, 21(11):3300?3307, 2015.

[8] Niessen W. Van Walsum T. M. Luu, H. An automatic registration method for pre- and post- interventional ct images for assessing treatment success in liver rfa treatment. 42(9), 2015.

[9] Vandermeulen D. Suetens P. Maes, F. Comparative evaluation of multiresolution optimization strategies for multimodality image registration by maximization of mutual information. 3.

[10] World Health Organisation. Cancer: The problem. pursue prevention strategies. early detection. early diagnosis. treatment.

[11] Camara R. Crum, W. and D. O., L. G. Hill. Generalized overlap measures for evaluation and validation in medical image analysis. *IEEE transactions on medical imaging*, 25(11):1451–1461.

[12] L. Boes J. Kim B. H. Bland P. R. Zasadny K. V. Kison P. Kiral K. A. Frey K.-L.Wahl R. R. Meyer, C. Demonstration of accuracy and clinical versatility of mutual information for automatic multimodality image fusion using affine and thin-plate spline warped geometric deformations. 1.

[13] Brandmaier P. Seider D. Kolesnik M. Jenniskens S. Sequeiros et al" title= "A prospective development study of software-guided radiofrequency ablation of primary Reinhardt, M., planning secondary liver tumors: Clinical intervention modelling, and pages="25?32" year="2017" DOI="http://doi.org/10.1016/j.conctc.2017.08.004" proof for ablation cancer treatment", journal= "Contemporary Clinical Trials Communications".

[14] Wirtz S. Strehlow J. Zidowitz S. et al. Rieder, C. Automatic alignment of pre- and post-interventional liver ct images for assessment of radiofrequency ablation. *Medical Imaging 2012: Image-guided procedures, robotic interventions, and modeling*, 2012.

[15] K. Murphy M.A. Viergever J.P.W. Pluim S. Klein, M. Staring. elastix: a toolbox for intensity based medical image registration. 29.

[16] Gervais D. A. Mueller P. R. & Arellano R. S. Sainani, N. I. Imaging after percutaneous radiofrequency ablation of hepatic tumors: Part 2, abnormal findings. *American Journal of Roentgenology*, 200(1):194?204, 2013.

[17] American Cancer Society. About liver cancer.

[18] Klein S. P.W. Pluim J. Staring, M. A rigidity penalty term for nonrigid registration. 2007.

[19] Marius Staring. Regularization in deformable registration.

[20] P. Thvenaz and M. Unser. Optimization of mutual information for multiresolution image registration. 9(12), 2000.

[21] Trifir G. Ravasi L. Monfardini L. Della Vigna P. Bonomo G. et al Travaini, L. L. Role of fdg-pet/ct after radiofrequency ablation of liver metastases: Preliminary results. *European Journal of Nuclear Medicine and Molecular Imaging*, 35(7):1316?1322, 2008.

[22] Kudo M. Zheng R. Q. Minami Y. Chung H. Suetomi Y. Wen, Y. L. Radiofrequency ablation of hepatocellular carcinoma. 181(July):57?63, 2003.