RETROSPECTIVE STUDY ON PHANTOM FOR THE APPLICATION OF MEDICAL IMAGE REGISTRATION IN THE OPERATING ROOM SCENARIO

Bogdan Mihai Maris
Altair Robotics Laboratory
Department of Computer Science
University of Verona
Verona, Italy
email: bogdan.maris@univr.it

Paolo Fiorini
Altair Robotics Laboratory
Department of Computer Science
University of Verona
Verona, Italy
email: paolo.fiorini@univr.it

ABSTRACT
This paper presents a phantom study to asses the feasibility of the medical image registration algorithms in the operating room (OR) scenario. The main issues of the registration algorithms in an OR application are, on one hand, the lack of the initial guess of the registration transformation - the images to be registered may be completely independent- and, on the other hand, the multimodality of the data. Other requirements to be addressed by the OR registration algorithms are: real-time execution and the necessity of the validation of the results. This work analyzes how, under these requirements, the current state of the art algorithms in medical image registration may be used and shows which direction should be taken when designing a OR navigation system that includes registration as a component.

KEY WORDS
Medical image registration, image-guided devices and interventions, multimodal imaging, computer aided surgery.

1 Introduction

Image registration is the procedure of aligning two or more images of the same scene taken from different viewpoints, at different time, and/or by different sensors, so that corresponding features can be easily related. Image registration has application to many fields, but the one addressed here is medical imaging and medical applications with a particular regard to OR applications.

In the case of an OR procedure or an image-guided system, the images to be registered are acquired in two steps: the first one takes place before the procedure and we call such dataset the *pre-operative* image, the second one occurs during the procedure and is called *intra-operative* dataset. By registering the two dataset, a spatial relationship between the anatomical structures in the two images and with the body of the patient is established and it is also possible to integrate spatial information about physiological functions and pathologies or other abnormalities.

Deriving the correspondence of spatial information in medical images and equivalent structures in the body is fundamental to image interpretation and analysis, but also to perform navigation in an OR setup.

The correspondence between the image and the physical space of the patient allows the image to provide a map for the navigation with the goal of making the intervention more accurate, safer and less invasive for the patient. The image registration techniques are already present in the clinical set-up for image-guided neurosurgery systems and in orthopedic surgery.

Considering the nature of the problem in medical image registration, that is the deformability of the tissue encountered in most of the medical images, except the images where the main interest is on rigid structures such as bones, most of the current registration algorithms try to find a solution that involves deformation.

Then, the challenge of this paper is to find image registration techniques to be included in an image-guided system that involves structures subject to deformations.

The modern medical image analysis is focused on algorithms that handle more and more complicated transformations needed to model soft tissue deformation. General reviews of the field may be found in [6] and [20].

In the following we present some of the state of the art solutions and we analyze whether these solutions have or have not the desired characteristics so that they can be employed in an image-guided system.

2 Intensity-based medical image registration framework

This class of algorithms is based on three components:

1. A distance measure between images to be registered.
2. A transformation model which could be parametric and nonparametric.
3. An optimization method.

In general, registration can be performed on two or more images. Without loss of generality, we may assume that the registration involve only two images. One of the

images is referred to as target image, denoted in the following by the character $\mathscr{T}$, and the other is referred to as reference image, denoted by $\mathscr{R}$.

Images are considered as mappings from a domain into the real numbers. The domain is denoted by $W \subset \mathbb{R}^d$, where d denotes the spatial dimensionality of the given data.

$$\mathscr{T} : W \to \mathbb{R}, W \subset \mathbb{R}^d \tag{1}$$

Typically d=2 or d=3. To each point in the domain, a gray value is assigned.

The registration process estimates a transformation f that minimizes the following nonlinear functional:

$$\mathscr{D}(\mathscr{T} \circ f, \mathscr{R}) + \mathscr{S}(f) \tag{2}$$

where $\mathscr{T} \circ f$ is the transformed target image, $\mathscr{D}$ measures the distance between images or the image similarity, and $\mathscr{S}$ measures the smoothness of the transform aiming also to favor any specific property in the solution that the user requires.

The transformation f applied to the target image $\mathscr{T}$ is a vector-valued function $f : \mathbb{R}^d \to \mathbb{R}^d$ and

$$(\mathscr{T} \circ f)(x) = \mathscr{T}(f(x)) \tag{3}$$

This approach, called Eulerian, is important from a practical point of view, since when the locations of the pixels/voxels in the target image are mapped to positions that do not correspond to pixels/voxels in the target image, their intensities can be calculated by interpolating the intensity values of the neighboring pixels/voxels. As a result, the corresponding transformation of the target image is counterintuitive: when the grid is rotated counterclockwise, the image is rotated clockwise.

The transformation at every position $x \in W$ may be given in a vector space as the addition of an identity transformation with the displacement field u:

$$f(x) = x + u(x), \tag{4}$$

or as a group structure, where the group operation is the function composition and the identity is given by the identity transformation.

2.1 Distance measures

The objective functional (see equation (2)) is the sum of a measure of distance between the transformed target image and the reference image and the smoothing functional $\mathscr{S}$ over the mapping f.

When the same anatomical structures are assumed to correspond to similar intensity values, correlation based matching produces dense depth maps by calculating the disparity at each pixel/voxel within a neighborhood.

The simplest distance measures in this case are given by the L_p-norms of the intensity differences. The most used norms in image registration are the L_1 norm, or *sum of absolute differences* SAD, and the L_2 norm, or *sum of squared differences* SSD. SAD measure works better when the number of pixels/voxels is small and the intensities differences between images are large. Both measures can handle Gaussian noise.

When the intensity values of one of the images is linearly shifted by different settings on the image acquisition scanner, or is affected by non-Gaussian noise distributions the *normalized cross correlation coefficient* (NCC) and the *zero-mean NCC* (ZNCC) are introduced.

From a geometric point of view, SAD and SSD may be understood as the norm of the vector that has as components the differences between images intensities at corresponding pixels/voxels locations, therefore is a measure to be minimized, meanwhile NCC and ZNCC can be seen as the cosine of the angle between the two normalized vectors that have as components the intensities of each image, therefore is a measure to be maximized in order to achieve the perfect alignment.

Applications of these measures of distance in medical image registration may be found in [10] or [1].

The choice of an appropriate distance measure is a harder task in the case of multi-modal imaging. In some works the multi-modal problem is reduced to a single-modal problem, by deriving one modality form another [18], [24]. The most used distance measure for multi-modal registration called *mutual information* (MI) and derived from the information theory, was first introduced by [25] and [5]. The MI measures how much information two images to be registered share together and reaches its maximum when the two images are aligned.

A compromise between SSD and mutual information, based on normalized image intensity gradients was introduced by Haber and Modersitzki in [8]. This distance measure is based on the observation that even for images of different modalities, intensity changes appear at corresponding positions. However, the gradient also measures the strength of the change which is an unwanted information for multi-modal information, therefore the gradient is normalized by its norm.

2.2 Parametric and non parametric solutions

The solution of the registration process, or the registration transformation, may be generated from a physical model that constrains the registration by the smoothness term $\mathscr{S}$ in (2), or by a parametrization of the transformation.

The type of the mapping is of paramount importance for the registration, as it reflects the class of transformations that are desirable or acceptable, and therefore limits the solution to a large extent. The registration parameters estimated through the optimization step correspond to the degrees of freedom of the transformation involved. Their number varies greatly, from six in the case of global rigid transformations in the three dimensional space, to a number equal to the number of pixels/voxels of the image in the

case of a dense transformation.

In the following, we call the transformations constrained to belong to a certain class of functions such as rigid, linear or affine, polynomial, radial basis functions, free form deformations, B-splines, thin plate spline, that have a relatively low number of parameters, *parametric transformations* while the transformations given as the discretized numerical solution of the equation (2), constrained by the chosen regularizer, *non parametric transformation*.

In the case of parametric transformations the models are derived from linear or nonlinear interpolation or approximation theories. The nonlinear methods range from polynomial to spline-based transformations that are piecewise polynomial functions with a predefined degree of regularity (see [22] for details).

When the transformation is derived from physical models, the displacement is given as the reaction of the model to a force. The force is generated by the similarity between the images. In the linear models or the *elastic body deformation*, first introduced by Broit [3] in medical image registration, the deformation is described by the Navier-Caucy partial differential equation, while the image grid is considered as an elastic membrane.

The *optical flow* techniques model the image as a function in space and time and assume that a particle located at $x(t)$ at time t does not change intensity. The optical flow leads the deformable target toward the reference. Example of algorithms for medical image registration based on the optical flow are in [21] and [4].

2.3 Optimization

The last step of the classical approach to image registration is the optimization. The aim of this step is to derive the optimal transformation that best aligns the two images according to the objective function given by the equation (2).

In this work we shall focus on optimization based on continuous data, since most of the registration algorithms use this approach. The reasons for using a continuous model are because it is more practical, since the transformed object does not always align with the pixel/grid and because of its computational efficiency. To derive a continuous model from the discrete data, interpolation has to be used.

The major difficulties in image registration, from the optimization point of view, is the handling of a variety of local and even global minimum. Multilevel methods are thus essential and it is not recommended to solve a problem using one fixed level.

Except *Powell's conjugate direction method* [17], all the commonly used methods are based on the computation of the gradient: *gradient descent*, *conjugate gradient*, *quasi Newton*, *Gauss-Newton*, *stochastic gradient descent*. A more detailed description of these and other methods of optimization may be found in [16].

When using an MI-based distance measure, one must pay particular attention to the optimization schemes that require the computation of the Hessian matrix due to its high computational cost. One of the most successful approach is the Broyden-Fletcher-Goldfarb-Shanno (BFGS) method which estimates the Hessian by an update from an initial approximation and a sequence of differences of search directions and gradients.

A comparative study of optimization strategies in image registration using mutual information is reported by Klein et al. in [11].

3 Landmark-based registration framework

In the case of landmark registration, the input of the algorithm is represented by two sets of landmarks. The first set of landmarks include points belonging to the target image and the second set is compose by landmarks from the reference image.

The main limitation for the extraction of landmarks from medical images is that they are not as rich in details as for instance the digital photographs. The extraction of landmarks has been studied more in the case of 2D images and less in the case of 3D images.

3.1 Correspondences and transformation

The transformation can be estimated using interpolation strategies, when the correspondences are known (see section 2.2).

The best known method that infers both the correspondences and the transformation is the iterative closest point (ICP) method, proposed by Besl and McKay [2].

In the following we assume that the correspondences are known *a priori* and the landmark-based registration techniques we shall test are linear, polynomial and quadratic.

If we denote $\{T_i\}$ the set of landmarks in the target image and $\{R_i\}$ the set of landmarks in the reference image, the goal is to find a transformation f such that

$$f(T_i) = R_i, \forall i. \tag{5}$$

Observe that this approach, called Lagrangian, is the opposite of the Eulerian approach presented in the previous section, in the equation (3). The same framework, given by the equation 3, may be implemented by simply substitute f with f^{-1}, when feasible.

More generally, the interpolation conditions (5) are replaced with approximation conditions:

$$\sum_i \| f(T_i) - R_i \|^2 = min. \tag{6}$$

If we consider the input data given as a binary image, where the foreground is given by the landmarks, this is nothing more than the SSD measure.

The linear solution, when the image domain is a subset of $\mathbb{R}^d$, looks for a $d \times d$ matrix A and a d-dimensional vector v such that $f(T_i) = AT_i + v$ and the minimization in the equation (6) takes place. Given n landmarks, the number of conditions is nd, while the number of parameters is $d(d+1)$. Considering the low dimensionality of the data and a relatively high number of landmarks, the match is not perfect, except in the case when $nd = d(d+1)$ (see section 6).

Enlarging the transformation space, for example by choosing polynomial transformation, the number of parameters increases with the polynomial degree. Writing the transformation f in vectorial form $f = (f_1, ..., f_d)$, $T_i = (T_i^1, ...T_i^d)$ and considering the quadratic transformation, then we have:

$$f_j(T_i) = a^j + \sum_{k=1}^{d} b_k^j T_i^k + \sum_{m \neq n} c_{m,n}^j T_i^m T_i^n + \sum_{k=1}^{d} (T_i^k)^2 \quad (7)$$

From equation (7) it follows that the number of parameters of the quadratic transformation is $\frac{d^2}{2} + \frac{3}{2}d + 1$.

If we want to solve equation (5) and find a function whose bending energy is minimum, we obtain the thin-plate-spline (TPS) transformation [23]. The bending energy is the integral of the square of the second derivative:

$$\mathscr{S}(f) = \sum_{j=1}^{d} \int \sum_{m,n=1}^{d} \left(\frac{\partial^2 f_j}{\partial x_m \partial x_n} \right)^2 dx_1 ... dx_d, \quad (8)$$

where $f = (f_1, ..., f_d) : \mathbb{R}^d \to \mathbb{R}^d$.

Therefore, the equation we want to solve is

$$\mathscr{S}(f) = min, \quad (9)$$

subject to (5).

The solution of this equation belongs to a space that is spanned by shifts of a known radial basis function ρ by a linear term:

$$f_j(x) = \sum_i c_j^i \rho(\|x - T_i\|) + Ax, \quad (10)$$

where A is a $d \times d$ matrix representing the affine transformation, $c = \{c_j^i\}$ is a $n \times d$ warping coefficient matrix representing the non-affine deformation and $\rho(r) = r^2 log(r)$ represent the kernel function.

The interpolation condition (5) can be relaxed and replaced by an approximation condition if we solve for both equations (6) and (9):

$$\sum_i \| f(T_i) - R_i \|^2 + \alpha \mathscr{S}(f) = min, \quad (11)$$

where $\alpha \geq 0$ is a parameter that balances the interpolation and the approximation scheme. This equation is nothing more than the registration formula we have already introduced in the equation (2). The solution is again given by equation (10) [14].

A nice property of the TPS is that it can always be decomposed into a global affine and a local non-affine component. Consequently, the TPS smoothness term in equation (11) is solely dependent on the non-affine components. This is a desirable property, especially when compared to other splines, since the global pose parameters included in the affine transformation are not penalized.

4 Experimental set-up for testing

To test some of the registration algorithms presented so far, we have used medical images obtained from a triple modality 3D abdominal phantom, developed by Computerized Imaging Reference Systems Inc (CIRS, Norwalk, VA). The model 057 interventional 3D abdominal phantom, is designed to address minimally invasive procedures and to be used in different abdominal scan techniques such as CT, MRI and US, developing imaging protocol and system testing and validation .

The reason for the choice of this set-up, beside the multi-modal capability, is due also to the complete control over the physical position of the scanned sections. We wanted to obtain two completely aligned 2D slices in two different modalities, therefore to create the ideal framework for the registration algorithms, considering the nature of our problem which is the registration of real-time acquired images with accurate pre-operative scans.

The difficulties encountered by the registration algorithms are due on one hand to the different physical process that generates the images and, on the other hand, by the deformation of the phantom generated by the pressure of the US probe and by the different covering of the US convex probe with respect to the CT slice which is a condition often encountered in practice (see figure 1, the deformation takes place in the upper part of the US image).

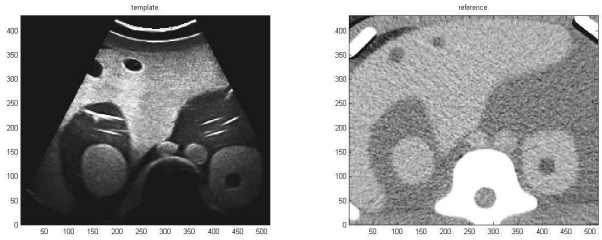

Figure 1. An example of images used in the registration process. On the left hand side the US target image; on the right hand side the CT reference (fixed) image. Some of the structures visible in both images are: vertebra, liver and kidneys with simulated lesions, abdominal aorta, ribs.

The phantom was equipped with 4 markers in order to register it with the CT dataset (figure 2). The global coordinate system is given by an optical tracking system composed by infrared light emitting cameras.

The US images were acquired with a 2D probe equipped with markers in order to map its position to the

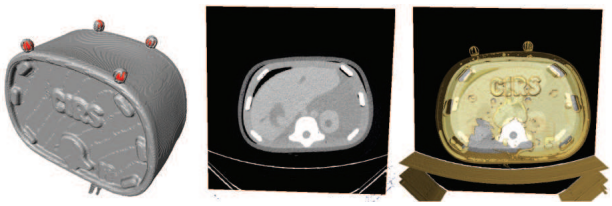

Figure 2. From left to right: the 3D image of the phantom with 4 markers that identify its position in space; a CT slice of the phantom; the same CT slice with the outline of the phantom in transparency.

global coordinate system. The calibration process of the US probe converts the point (u_i, v_i) in image plane coordinates (pixels) in 3D homogeneous coordinates defined with respect to the global reference system by the following formula:

$$(x_i\ y_i\ z_i\ 1)^\top = T_p T_{pi} (s_u u_i\ s_v v_i\ 0\ 1)^\top, \qquad (12)$$

where T_p is the pose matrix that encodes the pose of the markers and is given by the tracking system, while T_{pi} is the transform estimated by the calibration procedure together with the scale factors s_u and s_v that we assume isotropic, therefore $s_u = s_v = s_{uv}$.

We have used methods for multi-modal registration with transformation ranging from parametric (affine, spline) to non-parametric based on curvature or elastic. In the case of non-parametric registration the multi-level approach was necessary.

The tests were performed using MATLAB environment and FAIR toolbox [14].

5 Intensity-based registration results

5.1 Parametric registration

We report here (figure 3) the results we have obtained using two types of parametric registration (affine and spline) and two types of distance measures (mutual information and normalized gradient field).

The results obtained using MI as distance measure are not satisfactory in both affine and spline cases. The optimization process tends very quick to a local minimum, and the transformation leaves the image almost unchanged, in the case of affine transformation or applies a very small local deformation, in the case of the spline deformation.

The trend of the MI measure is to quickly reach a stable position (figures 4 and 5) at around 66% in the affine case, and 87% in the spline case, of the of the initial value.

The computation time is very long.

The minimization scheme used was Gauss-Newton.

The NGF measure is more efficient in terms of computation time but gives bad results in qualitative terms. Visually, figure 3 shows that in both affine and spline cases

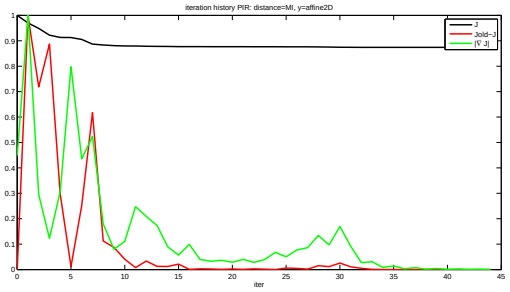

Figure 4. Affine registration using MI. The graphical trend of the objective function J during 44 iterations (the black line). The green line gives the graph of $|\nabla J|$, while the red shows line the graph of the difference between the previous J value and the current value.

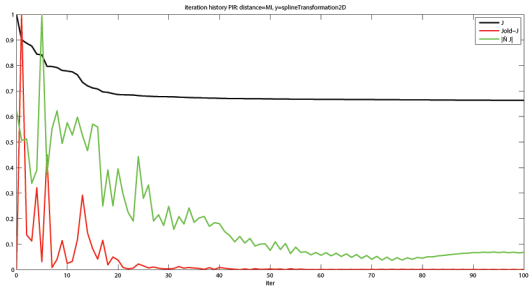

Figure 5. Spline registration using MI. The graphical trend of the objective function J during 100 iterations (the black line).

the results are meaningless and the deformation applied is arbitrary.

As it can be seen in figures 6 and 7, the values of the objective function do not change over the iterations, they are indistinguishable in the graphic. In fact, they remain always very near to the initial value 1.

The only advantage over the previous approach was the computation time.

Also in this case we have used Gauss-Newton for the minimization.

5.2 Multilevel parametric and non parametric registration

The multilevel representation of the input data is required first of all in order to reduce the risk of being trapped by local minimum. At the same time, a solution of a coarse representation of the problem serves as a starting point for a representation with more details. Starting with a very coarse representation, the procedure is repeated on each level, until all details provided by the initial data are resolved.

From an optimization point of view, the multilevel representation yield a smoother representation of the ob-

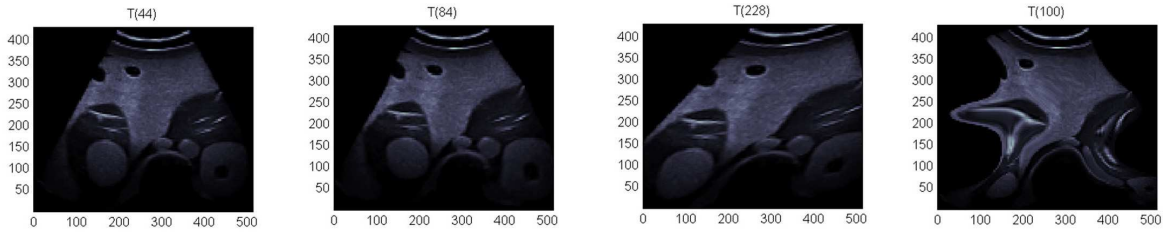

Figure 3. The target image after the parametric multi-modal registration process. From left to right: affine registration using MI after 44 iterations, spline registration using MI after 84 iterations, affine registration using NGF after 228 iterations, spline registration using NGF after 100 iterations.

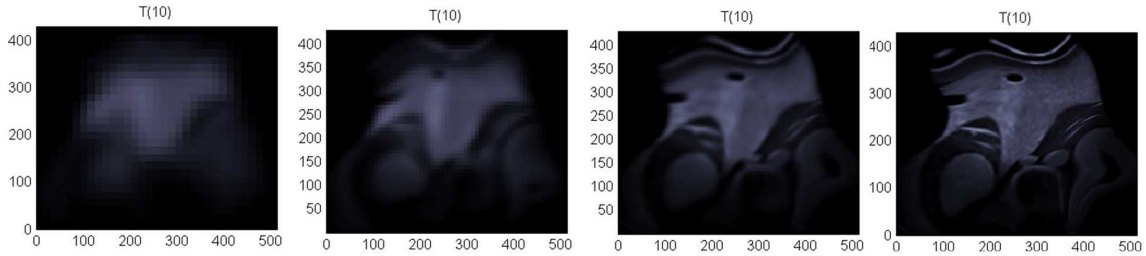

Figure 8. Multilevel spline registration

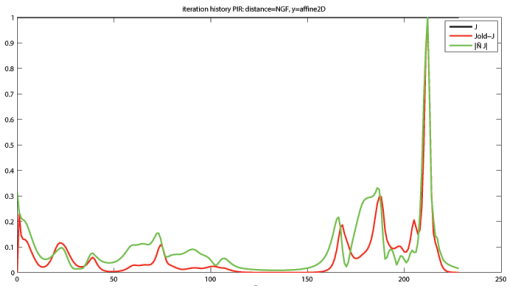

Figure 6. Affine registration using NGF. The graphical trend of the objective function J during 228 iterations (the black line).

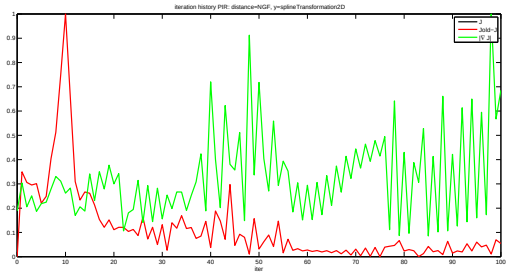

Figure 7. Spline registration using NGF. The graphical trend of the objective function J during 100 iterations (the black line).

jective function. A smooth problem may be easier to resolve and, based on a good starting point, the more detailed problem can be solved quicker.

The multilevel representation is also useful to reduce the computational time. The complexity of the registration algorithms increases exponentially with the dimension of the input data.

5.2.1 Parametric multilevel spline registration

Figure 8 reports the results of the multilevel spline registration. After 10 iterations on each level, the result is passed to the first iteration of the next level.

The transformation also in these cases is not natural, after a strong bending obtained at the first level, the other levels tend to bend less the image.

The computation time is very high and an ulterior refinement at each level through the other iterations does not improve the results as it can be seen in the figure 8.

The tests using the affine transformation are not reported here but the results are similar to the non-level approach, and the transformation blocks quickly into a local minimum.

5.2.2 Non parametric multilevel registration

Figure 9 shows the results of the non-parametric elastic multi-level registration. The algorithm starts at the coarser level using an affine transformation, then the registration at each level starts using the previous obtained result. After the initialization, there is not much interaction of the algo-

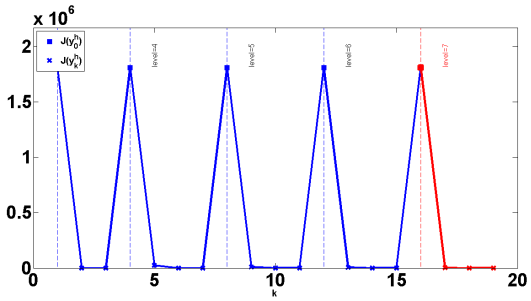

Figure 9. Iteration history of multilevel non-parametric elastic registration: vertical lines separate different levels. The squares represent the initial value of MI, while the crosses show the value on each iteration.

rithm with the data, the images remain almost unchanged and the number of iterations is very low on each level. Even so, the computation time is very high.

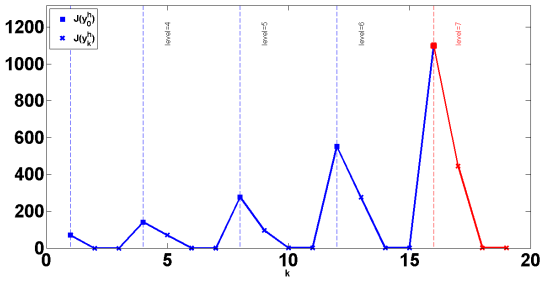

Figure 10. Iteration history of multilevel non-parametric curvature registration: vertical lines separate different levels. The squares represents the initial value of MI, while the crosses show the value on each iteration.

The results obtained using the curvature registration are very similar with those obtained using the elastic registration (figure 10). On each level the algorithm stops after 2-3 iterations.

5.3 Computation time

We report in table 1 the time taken for each of the test we have presented in the previous sections.

Registration type	Transformation type	Time (sec.)
Parametric MI	Affine	15
	Spline	87
Parametric NGF	Affine	6
	Spline	18
Multilevel parametric	Affine	29
	Spline	143
Multilevel non-parametric	Elastic	214
	Curvature	168

Table 1. Computation time of the principal registration algorithm tested.

6 Landmark-based registration results

6.1 Linear registration

The linear registration in 2D involves the computation of 5 parameters, 2 for translation, 1 for rotation and 2 for scaling, therefore a minimum number of 3 landmarks is required.

The solution with 3 landmarks align perfectly only the markers, but the overall result is not qualitatively acceptable (figure 11, top row).

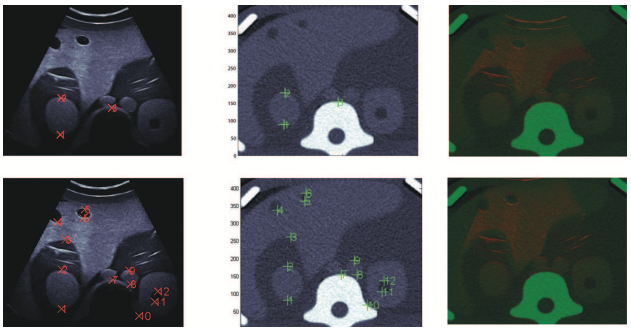

Figure 11. Linear registration. Top row, first two images from the left: 3 selected landmarks in the target and reference image, respectively. On the right: the two images registered with the target on the red channel and the reference on the green channel. Bottom row, first two images from the left: 12 selected landmarks in the target and reference image, respectively. On the right: the two images registered.

Increasing the number of landmarks, the registration problem becomes overdetermined and the solution is given by minimizing the sum of the distances between every pair of corresponding landmarks.

Even though not all the landmarks will be aligned , the result improves a lot compared to the 3 landmarks solution (figure 11,bottom row).

6.2 Non-linear registration

The results of the linear registration may be improved by using a non-linear approach. The quadratic solution in 2D has a number of 12 parameters (see equation (7)) and is completely solved by 6 landmarks. In this case the error in the alignment of the landmarks is negligible because depends only on the computational precision (figure 12, row (a)).

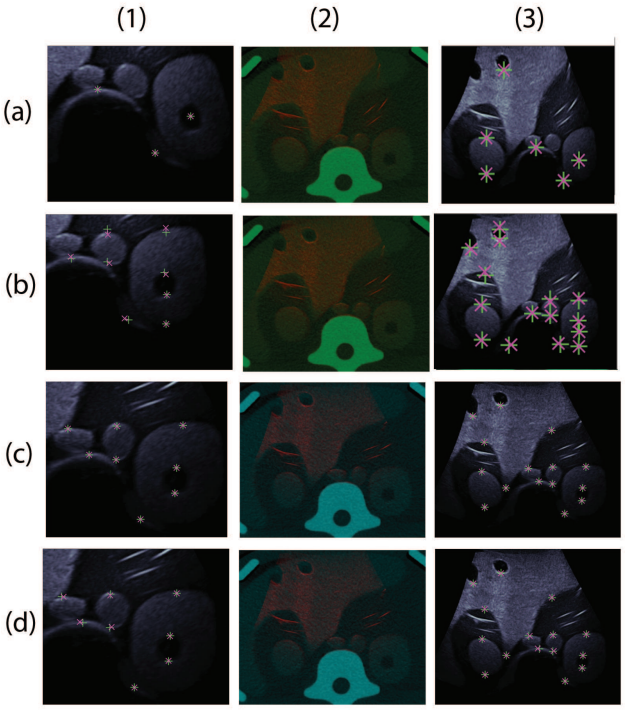

Figure 12. Non-linear registration. Columns: (1) A close-up of the landmarks after the registration (green for the reference, red for the target); (2) The two images registered overlapped on red channel (target image) and on the green channel (reference image); (3) The target image registered. Rows: (a) Quadratic registration results with 6 landmarks; (b) Quadratic registration with 15 landmarks; (c) Spline registration with 15 landmarks and $\alpha = 0$; (d) Spline registration with 15 landmarks and $\alpha = 1000$

The registration result applied to the entire image, even in the 6 landmarks case, is acceptable. As in the linear case, increasing the number of landmarks, the solution is overdetermined but the overall result is better. By using the quadratic term, the alignment of 15 landmarks gives a minor error compared with the use of 12 landmarks in the linear case.

The polynomial solution may produce arbitrary deformations especially when the number of landmarks is low. The approach based on the thin-plate-spline (TPS) functions further refines the previous results. The spline functions may yield the perfect matching of the landmarks (interpolation) or an approximation, when the landmarks are not completely overlapped. These two conditions are realized by varying a smoothing parameter α (see equation (11)). The solutions range from a low degree of bending when α is large to a high degree of bending in the case of small α. However, considering a number of landmarks from 6 to 15 and α ranging from 0 to 1000, this phenomenon is barely visible and the distance between the two sets of registered landmarks increases visibly only in the case of $\alpha = 1000$ (figure 12, rows (c) and (d)).

7 Conclusions

Why the intensity-based medical image registration framework cannot work for OR applications?

The major issues we have identified that makes this approach unfeasible in OR applications are:

- The distance measure MI: this measure is very sensible at noise, incomplete data, no completely overlapping domains.
- The distance measure NGF: in our tests this measure has not produced the results we expected.
- The optimization: since the objective function is highly non convex, all the optimization methods fails in finding the global minimum or maximum.
- The multilevel approach: even though the multilevel approach is vital when using the non-parametric approach, in the case of multi-modal images the low level approximation of the registration transformation is not accurate at all and generates wrong results at higher levels.
- The computation time is very high even when only a couple of 2D images is involved.

Hadamard [9] defined a problem well-posed if it has a solution, this solution is unique and depends continuously on the data. In this sense, the registration problem is ill-posed since for every spatial location $x \in W \subset \mathbb{R}^d$, we search a vector $f(x) \in \mathbb{R}^d$, but usually only a scalar information $\mathscr{T}(y(x))$ is given.

[19] gives a simple example of how ill-posedness may give arbitrary results that seem very good from an optimization point of view but are completely useless in practice. The authors called this approach CURT (completely useless registration tool), which is a very simple registration algorithm based on correspondence of pixels sorted by increasing intensities, and showed that this method outperform other registration methods in the following cases: SSD difference, NCC image correlation and NMI (normalized MI) image similarity.

Except the well established and accepted solutions for image registration, some newer solution were introduced and most of them address new distance measures such as a locally evaluation of MI in combination with standard

global MI [26], the residual complexity to account for complex spatially varying intensity distortions [15], learning based multi-modal registration using Kullback-Leibler divergence for non-rigid registration [7] and rigid registration using learning based Jensen-Shannon divergence [12].

These new methods tend to ulteriorly complicate the already complex registration framework and most of them follow the three steps approach.

It would be difficult to test all of them but we intuit that our case study cannot be solved by the classical approach of medical registration framework and we shall need additional information to solve the registration in the OR.

Why the landmark-based registration works better than the intensity-based registration for an OR application?

The first observation is that, in both cases, the transformations we are interested in have no physical meaning. Except in the rigid case, they are just mathematical tools that help to solve the registration. Following this observation, the intensity-based completely automatic algorithms fail and give arbitrary results if no prior information is given. In the case of landmark registration, even with a small number of landmarks and the rigid/linear constraint the results are much better. Extracting corresponding landmarks in multi-modality images excludes the computation of a multi-modal distance measure. As we have seen, the computation of MI and NGF generates misleading results and the computational time is very high, even in the 2D case. On the other hand, the landmark registration will always yield a decent result and the computation time is very low.

What are the limitations of the landmark-based registration for an OR application?

Even if the registration results in the landmark case are preferable, the bottleneck of this method is not only the detection of landmarks, but also the computation of the correspondences. Having the correspondences, we can choose a mathematic model and, based on the number of landmarks, the registration is solved as some form of interpolation (section 3.1). This solution is simple and intuitive and it may always isolate a linear transformation, even with the TPS solution, which is an advantage since in the OR scenario a rigid motion is always required.

What is the direction for the future works?

Another research area in the field of medical image processing is the segmentation, that is the extraction of contours or boundaries from the interest area. This type of processing may generate binary images, where the foreground is given by the extracted contours and the background is given by the remaining domain.

The paradigm of landmark registration may be extended to this type of images but, in this case, the point to point correspondence is lost. Most of the radiologic information will disappear but, as we have seen, we don't need all the radiologic information. In fact, the radiologic information should be used only to detect the relevant features during the segmentation process.

A rigid transformation should be first identified for the global alignment and, when required, a non-rigid transformation should be derived as a local refinement.

A new approach to the registration of dense sets of landmarks based on the computation of correspondences, derived from the inter-point distances in each dataset, followed by the computation of the registration was recently published in [13].

This paper has not talked much about the validation of the registration algorithms. Most of the solutions we have presented here were estimated only visually. This was due to the fact that in an image-guided procedure the validation depends on the task we are addressing. For instance, in a targeted procedure for the ablation or the biopsy of a tumor, the parameter to be estimated for the validation is the overlapping volume of the tumor in the two image datasets: pre-operative and intra-operative. In a simplified scenario, when the insertion of an instrument toward a target point is required, the registration may be validated by computing the distance between the virtual point in the pre-operative image identified once the registration took place and the real position of the physical point we want to reach.

8 Acknowledgements

The work described in this paper has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement no. 688188 as part of the MURAB (MRI and Ultrasound Robotic Assisted Biopsy) project.

References

[1] Avants, B.B., Epstein, C.L., Grossman, M., Gee, J.C.: Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain. Medical Image Analysis (1), 26–41

[2] Besl, P.J., McKay, N.D.: A method for registration of 3-D shapes. IEEE Transactions on Pattern Analysis and Machine Intelligence **14**(2), 239–256 (1992). DOI 10.1109/34.121791

[3] Broit, C.: Optimal registration of deformed images. Ph.D. dissertation, Univ. Pennsylvania, Philadelphia (1981)

[4] Christensen, G.E., Johnson, H.J.: Consistent image registration. IEEE Trans. Med. Imaging **20**(7), 568–582 (2001)

[5] Collignon, A., Maes, F., Delaere, D., Vandermeulen, D., Suetens, P., Marchal, G.: Automated multimodality medical image registration using information theory. In: Y. Bizais, C. Barillot (eds.) Information Pro-

cessing in Medical Imaging, pp. 263–274. Kluwer Academic Publishers, Dordrecht (1995)

[6] Goshtasby, A.A.: Image Registration: Principles, Tools and Methods. Springer (2012)

[7] Guetter, C., Xu, C., Sauer, F., Hornegger, J.: Learning based non-rigid multi-modal image registration using kullback-leibler divergence. In: MICCAI (2), pp. 255–262 (2005)

[8] Haber, E., Modersitzki, J.: Intensity gradient based registration and fusion of multi-modal images. In: MICCAI (2), pp. 726–733 (2006)

[9] Hadamard, J.: Sur les problèmes aux dérivés partielles et leur signification physique. Princeton University Bulletin **13**, 49–52 (1902)

[10] Kim, J., Fessler, J.A.: Intensity-based image registration using robust correlation coefficients. IEEE Trans. Med. Imaging (11), 1430–1444

[11] Klein, S., Staring, M., Pluim, J.P.W.: Evaluation of optimization methods for nonrigid medical image registration using mutual information and b-splines. IEEE Transactions on Image Processing **16**(12), 2879–2890 (2007)

[12] Liao, R., Guetter, C., Xu, C., Sun, Y., Khamene, A., Sauer, F.: Learning-based 2d/3d rigid registration using jensen-shannon divergence for image-guided surgery. In: MIAR, pp. 228–235 (2006)

[13] Maris, B.M., Fiorini, P.: Generalized shapes and point sets correspondence and registration. Journal of Mathematical Imaging and Vision **52**(2), 218–233

[14] Modersitzki, J.: FAIR: Flexible Algorithms for Image Registration. SIAM, Philadelphia (2009)

[15] Myronenko, A., Song, X.B.: Intensity-based image registration by minimizing residual complexity. IEEE Trans. Med. Imaging **29**(11), 1882–1891 (2010)

[16] Nocedal, J., Wright, S.: Numerical Optimization. Springer Series in Operations Research. Springer (1999). URL http://books.google.de/books?id=epc5fX0lqRIC

[17] Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P.: Numerical Recipes 3rd Edition: The Art of Scientific Computing, 3 edn. Cambridge University Press, New York, NY, USA (2007)

[18] Roche, A., Pennec, X., Malandain, G., Ayache, N.: Rigid registration of 3d ultrasound with mr images: a new approach combining intensity and gradient information. IEEE Transactions on Medical Imaging **20**, 1038–1049 (2001)

[19] Rohlfing, T.: Image similarity and tissue overlaps as surrogates for image registration accuracy: Widely used but unreliable. IEEE Trans. Med. Imaging **31**(2), 153–163 (2012)

[20] Sotiras, A., Davatzikos, C., Paragios, N.: Deformable medical image registration: A survey. Medical Imaging, IEEE Transactions on **32**(7), 1153–1190 (2013)

[21] Thirion, J.P.: Non-rigid matching using demons. In: CVPR, pp. 245–251 (1996)

[22] Unser, M.: Splines: A perfect fit for signal/image processing. IEEE Signal Processing Magazine **16**, 22–38 (1999)

[23] Wahba, G.: Spline Models for Observational Data, *CBMS-NSF Regional Conference Series in Applied Mathematics*, vol. 59. SIAM (1990). URL http://www.ams.org/mathscinet-getitem?mr=1045442

[24] Wein, W., Brunke, S., Khamene, A., Callstrom, M.R., Navab, N.: Automatic ct-ultrasound registration for diagnostic imaging and image-guided intervention. Medical Image Analysis (5), 577–585

[25] Wells, W.M., Viola, P., Atsumi, H., Nakajima, S., Kikinis, R.: Multi-modal volume registration by maximization of mutual information. (1996)

[26] Zhuang, X., Arridge, S.R., Hawkes, D.J., Ourselin, S.: A nonrigid registration framework using spatially encoded mutual information and free-form deformations. IEEE Trans. Med. Imaging **30**(10), 1819–1828 (2011)