Preprint of a Chapter to appear in *Multi-Sensor Image Fusion and Its Applications*, R. S. Blum and Zh. Liu editors. Marcel Dekker Inc, 2005.

Non-Rigid MR/US Registration for Tracking Brain Deformations

Xavier Pennec[1], Alexis Roche[2], Pascal Cathier[3], Nicholas Ayache[1]

[1] EPIDAURE, INRIA Sophia-Antipolis, BP 93, F-06902 Sophia Antipolis Cedex, France
{Xavier.Pennec, Nicholas.Ayache}@sophia.inria.fr

[2] Service Hospitalier Frdric Joliot, CEA, F-91401 Orsay, France
roche@shfj.cea.fr

[3] CAD Solutions Group, Siemens Medical Solutions, Malvern, PA, USA
Pascal.Cathier@siemens.com

February 6, 2004

Abstract

During a neurosurgical intervention, the brain tissues shift and warp. In order to keep an accurate positioning of the surgical instruments, one has to estimate this deformation from intra-operative images. 3D ultrasound (US) imaging is an innovative and low-cost modality which appears to be suited for such computer-assisted surgery tools. In this paper, we present an image-based technique to register intra-operative 3D US with pre-operative Magnetic Resonance (MR) data. A first automatic rigid registration is achieved by maximizing a similarity measure that generalizes the correlation ratio. Then, brain deformations are tracked in the 3D US time-sequence using a "demon's" like algorithm. Experiments show that a registration accuracy of the MR voxel size is achieved for the rigid part, and a qualitative accuracy of a few millimeters could be obtained for the complete tracking system.

1 Introduction

The use of stereotactic systems is now a quite standard procedure for neurosurgery. However, these systems assume that the brain is in fixed relation to the skull during surgery. In practice, relative motion of the brain with respect to the skull (also called brain shift) occurs, mainly due to tumor resection, cerebrospinal fluid drainage, hemorrhage or even the use of diuretics. Furthermore, this motion is likely to increase with the size of the skull opening and the duration of the operation.

Over the last years, the development of real-time 3D ultrasound (US) imaging has revealed a number of potential applications in image-guided surgery as an alternative approach to open MR and intra-interventional CT. The major advantages of 3D US over existing intra-operative imaging techniques are its comparatively low cost and simplicity of use. However, the automatic processing of US images has not gained the same degree of development as other medical imaging modalities, probably due to the low signal-to-noise ratio of US images.

1.1 Context

We present in this chapter a feasibility study of a tracking tool for brain deformations based on intra-operative 3D ultrasound (US) image sequences. This work was performed within the framework of the European project ROBOSCOPE, a collaboration between The Fraunhofer Institute (Germany), Fokker Control System (Netherlands), Imperial College (UK), INRIA (France), ISM-Salzburg and Kretz Technik (Austria). The goal of the whole project is to assist neurosurgical operations using real-time 3D ultrasound images and a robotic manipulator arm (fig. 1).

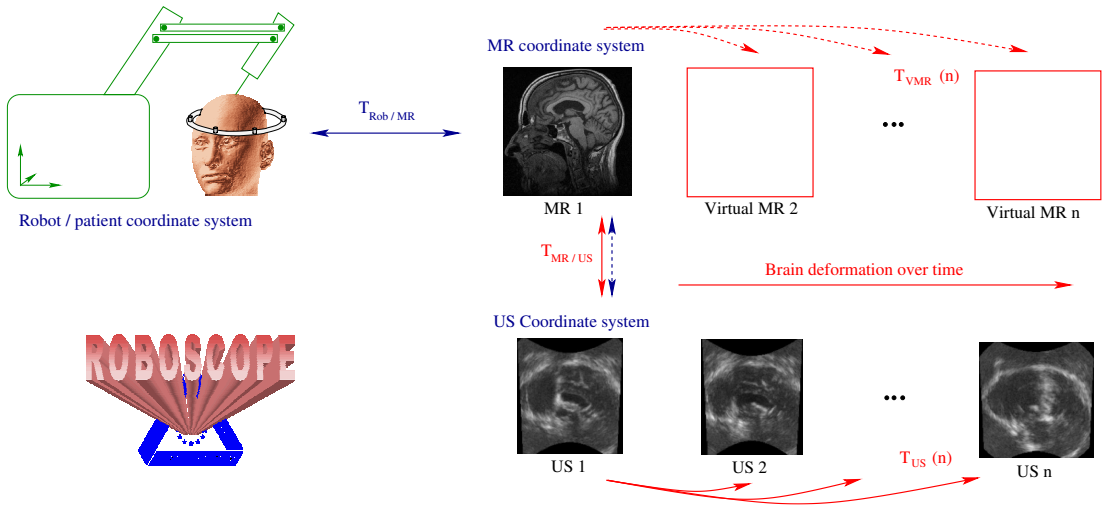

Figure 1: Overview of the image analysis part of the Roboscope project.

The operation being planned on a pre-operative MRI (MR1), the idea is to track in real time the deformation of anatomical structures using 3D US images acquired during surgery. To calibrate the system (i.e. to relate the MR and the US coordinate systems) and possibly to correct for the distortions of the US acquisition device, a first US image (US1) has to be acquired with dura mater still closed in order to perform a rigid registration with the preoperative MR. Then, per-operative 3D US images are continuously acquired during surgery to track the brain deformations. From these deformations, one can update the preoperative plan and synthesize a virtual MR image that matches the current brain anatomy.

1.2 MR/US registration

The idea of MR/US registration is already present in [1, 2] where the US probe is calibrated (i.e. registered to the surgical space) and then tracked using an optical device. Using standard stereotactic neurosurgical procedures, they also register the MR image to the surgical space, thus relating the MR and US coordinate systems. [3] improved this system by designing a real-time low-cost US imaging system based on a PCI bus, while [4] used a DC magnetic position sensor to track the US probe. [5] interactively delineates corresponding surfaces in both MR and US images and performs the registration by visually fitting the surfaces using a 6D space-mouse. In [6], the outlines of the 2D US image is registered to the MR surface using a Chamfer matching technique. All these techniques only perform a rigid registration of the MR and the US image (i.e. the calibration of their coordinate systems) using external apparatus or interactive techniques.

For a non rigid registration (i.e. a brain shift estimation), we have to turn to [7, 8, 9], where the 2D US probe is still optically and rigidly tracked but the corresponding MR slice is displayed to

the user who marks corresponding points on MR and US slices. Then, a thin plate spline warp is computed to determine the brain shift. This method was also developed in [10] with the possibility of using 3D US images and a deformation computed using a spring model instead of splines. More recently, Ionescu *et al* [11] registered US with Computed Tomography (CT) data after automatically extracting contours from the US image using a watershed segmentation. In these studies, there is no processing of a full time sequence of US images: the brain shift estimation is limited to a few samples from times to times as the user interaction is required at least to define the landmarks.

Up to our knowledge, only [12] deals with an automatic non-rigid MR/US registration: the idea is to register a surface extracted from the MR image to the 3D US image using a combination of the US intensity and the norm of its gradient in a Bayesian framework. The registration is quite fast (about 5mn), even if the compounding of the 3D US and the computation of its gradient takes much longer. However, experiments are presented only on phantom data and our experience (see section 4) is that real US images may lead to quite different results.

Be what it may, the first phase of all MR/US registrations is always a calibration step, i.e. a rigid registration, in order to relate the US probe position with the MR image coordinate system. Thus, we chose to split it into two subproblems: first a rigid registration is performed between the MR image and a first US image acquired before the opening of the dura matter (there is no brain shift yet, so the transformation may safely be considered as rigid) ; then we look for the non-rigid motion within the US time-sequence.

1.3 Tracking methods in sequences of US images

There are few articles on the registration of 3D US images. [13] use a maximum-likelihood approach to deduce a similarity measure for ultrasound images corrupted by a Rayleigh noise and a block-matching strategy to recover the rigid motion. In [14], the correlation of the norm of the image gradient is used as the similarity measure to rigidly register two US images in replacement of the landmark-based RANSAC registration of [15]. However, these methods only deal with rigid motion and consider only two images, eluding the tracking problem.

One has to move to cardiac application to find some real tracking of non-rigid motion in US images. In [16], the endo- and epi- cardial surfaces are interactively segmented on each 2D image plane. Then, a shape-memory deformable model determines the correspondences between the points of the 3D surfaces of successive images. These correspondences are used to update an anisotropic linear elastic model (finite element mesh). The approach is appealing but relies once again on an interactive segmentation. In [17], a combination of feature point extraction (phase-based boundaries) and a multi-scale fuzzy clustering algorithm (classifying the very low intensities of intra-ventricular pixels) is used to segment the surface of the left ventricular cavity. This process is done in 2D+T and then reconstructed in 3D. Thus it exploits the whole sequence before tracking the motion itself, which is not possible for our application. These two methods are well suited for the shape of the cardiac ventricle using dedicated surface models. If they could be adapted to the brain ventricles, it seems difficult to extend them to the tracking of the volumetric deformations of the whole brain.

1.4 Intensity based non-rigid registration algorithms

Since feature or surface extraction is especially difficult in US images, we believe that an intensity-based method can more easily yield an automatic algorithm. Over recent years, several non-rigid registration techniques have been proposed. [18] differentiated the linear correlation criterion and used a fixed fraction of its gradient as an external force to interact with a linear elasticity model.

Christensen shows in [19] that the linear elasticity, valid for small displacements, cannot guarantee the conservation of the topology of the objects as the displacements become larger: the Jacobian of the transformation can become negative. Thus, he proposed a viscous fluid model of transformations as it can handle larger displacement. This model is also linearized in practice. Bro Nielsen started from the fluid model of Christensen and used the linearity of partial derivative equations to establish a regularization filter, several order of magnitude faster than the previous finite element method [20]. He also justified his forces as the differential of the sum of square intensity differences criterion, but he still used a fixed fraction of this gradient, and shows that Gaussian smoothing is an approximation of the linear elastic model.

Some authors [21] tried to apply to non-rigid registration some criteria developed for rigid or affine matching using block-matching techniques. However, these criteria require a minimal window size, thus limiting the resolution of the result. Moreover, the regularization of the displacement field is usually implicit, i.e. only due to the integration of the criterion over the window, which means that it is difficult to explicitly control the regularity of the sought transformation.

Thirion proposed to consider non rigid registration as a diffusion process [22]. He introduced in the images entities (demons) that push according to local characteristics of the images in a similar way Maxwell did for solving the Gibbs paradox in thermodynamics. The forces he proposed were inspired from the optical flow equations. This algorithm is increasingly used in several teams as reported by [23, 24, 25, 26]. In [27], we investigated the non-rigid registration using gradient descent techniques. Differentiating the sum of square intensity differences criterion (SSD), we showed that the demons forces are an approximation of a second order gradient descent on this criterion. The same gradient descent techniques were applied to a more complex similarity measure in [28]: the sum of Gaussian-windowed local correlation coefficients (LCC).

1.5 Overview of the article organization

The first part of this article deals with the rigid registration (calibration) of the pre-operative MR image with the first US image. The proposed method expands on the correlation ratio (CR) registration method [29] in order to deal with the specificities of the US acquisition. In essence, we have improved the method in following three distinct axes [30, 31]: using the gradient information from the MR image, reducing the number of intensity parameters to be estimated, and using a robust intensity distance.

The second part of the article develops an automatic intensity-based non-rigid tracking algorithm suited for real-time US images sequences, based on encouraging preliminary results reported in [27, 32]. We first present the registration method for two US images and how the method is turned into a tracking algorithm [33].

In section 4, we present some results of the rigid MR/US registration on clinical data (a baby and an actual surgical case), along with the results of an original evaluation of the registration accuracy. Then, we present qualitative results of the tracking algorithm on a sequence of 3D US animal images and a qualitative evaluation of the complete tracking system on a sequence of images of an MR and US compatible phantom.

2 Rigid MR/US Registration

2.1 Classical intensity-based similarity metric

Given two images I and J (considered as intensity functions of the position is space), the basic principle of intensity-based registration methods is to search for a spatial transformation T such

that the intensities of the transformed image $J \circ T$ be as similar as possible to the intensities of image I. The whole question is how to quantify the similarity between the two images intensities.

2.1.1 Sum of Squared Differences (SSD)

If we assume that there are no intensity changes between the two image (except a Gaussian additive noise), then a good similarity metric (which may be justified by a maximum likelihood approach) is the Sum of Squared Differences (SSD) criterion:

$$SSD(T) = \|I - J \circ T\|^2 = \int [I(x) - J(T(x))]^2 .dx \tag{1}$$

2.1.2 Correlation Coefficient (CC)

Now, let us assume that there is moreover a linear bias, so that the the intensity $I(x)$ of image I at any point x may be modeled as $a.J(T(x)) + b$ plus some Gaussian noise, where a and b are some constant multiplicative and additive biases. A maximum likelihood approach leads to minimize the criterion

$$C(T, a, b) = \|I - a.J \circ T - b\|^2 = \int [I(x) - a.J(T(x)) - b]^2 .dx$$

Let us denote by $\bar{I}$ the mean intensity value of image I, $\mathrm{Var}(I) = \|I - \bar{I}\|^2$ its variance, $\mathrm{Cov}(I, J) = \int (I - \bar{I}).(J - \bar{J})$ the covariance of image I and J, and finally $\rho(I, J) = \mathrm{Cov}(I, J)/\sqrt{\mathrm{Var}(I).\mathrm{Var}(J)}$ the *correlation coefficient* (CC) between the two images. Computing the optimal values of the constants a and b and eliminating them, we get a criterion that only depends on the transformation:

$$C(T) = \mathrm{Var}(I) - \frac{\mathrm{Cov}(I, J \circ T)^2}{\mathrm{Var}(J \circ T)} = \mathrm{Var}(I).\left(1 - \rho(I, J \circ T)^2\right)$$

One problem is that we can compute this criterion only on the overlapping part of the transformed images. In order to avoid a minimum when the image overlap is small, we need to renormalise the criterion: a good choice, justified in [34], is to look for a large variance of I in the overlapping region (we are trying to register informative parts), so that the criterion becomes:

$$C(T) = \frac{\|I - a.J \circ T - b\|^2}{\mathrm{Var}(I)} = 1 - \rho(I, J \circ T)^2 \tag{2}$$

where the integrals are computed over the intersection of the domains of I and $J \circ T$. Thus, minimizing $C(T)$ amounts to find the transformation that maximizes the squared or absolute value of the *correlation coefficient* (CC) $\rho(I, J \circ T)$ between the two images.

2.1.3 Correlation ratio (CR)

Now, let us assume that the intensity mapping between image J and image I is an arbitrary functional relationship, instead of being just linear. With the same arguments as above, the renormalized maximum likelihood criterion is now:

$$C(T, f) = \frac{\|I - f(J \circ T)\|^2}{\mathrm{Var}(I)} \tag{3}$$

This formulation is asymmetric in the sense that the cost function changes when permuting the roles of I and J. Since the positions and intensities of J actually serve to predict those of I, we

will call J the "template image". In the context of US/MR image registration, we always choose the MR image as the template.

If no constraint is imposed to the intensity mapping f, an important result is that the optimal f at fixed T enjoys an explicit form that is very fast to compute [29]. The minimization of eq (3) may then be performed by traveling through the minima of $C(T, f)$ at fixed T. This yields the correlation ratio, $\eta^2_{I|J}(T) = 1 - \min_f C(T, f)$, a measure that reaches its maximum when $C(T, f)$ is minimal. In practice, the maximization of η^2 is performed using Powell's method [35]. Another important point is the discretisation scheme used to compute the criterion, leading to the choice of an interpolation scheme [34]. In this paper, we use the Partial Volume interpolation [36], that avoids many artifacts and local minima [37, 38].

2.2 Bivariate correlation ratio

Ultrasound images are commonly said to be "gradient images" as they enhance the interfaces between anatomical structures. The physical reason is that the amplitudes of the US echos are proportional to the squared *difference* of acoustical impedance caused by successive tissue layers. Ideally, the US signal should be high at the interfaces, and low within homogeneous tissues.

Thus, assuming that the MR intensities describe homogeneous classes of tissues amounts to consider the acoustic impedance Z as an unknown function of the MR intensities: $Z(x) = g(J(x))$. Now, when the ultrasound signal emitted from the probe encounters an interface (i.e. a high gradient of Z), the proportion of the reflected energy is $R = \|\nabla Z\|^2/Z^2$. Adding a very simple model of the log-compression scheme used to visualize the US images, we obtain the following US image acquisition model: $I(x) = a. \log\left(\|\nabla Z\|^2/Z^2\right) + b + \epsilon(x)$. Using $Z(x) = g(J(x))$ finally gives an unknown bivariate function: $I(x) = f\left(J(x), \|\nabla J(x)\|\right) + \epsilon(x)$. Our new correlation ratio criterion is then:

$$C(T, f) = \frac{\|I - f(J \circ T, \|\nabla J \circ T\|)\|^2}{\mathrm{Var}(I)} \tag{4}$$

The MR gradient is practically computed by convolution with a Gaussian kernel.

In this rather simple modeling of the relationship between the two images, we ignored the incidence angle between the scan direction u of the US probe and the "surface normal" (the gradient): we should ideally use the cross product $< \nabla J \circ T\,,\, u >$ instead of just the gradient norm $\|\nabla J \circ T\|$. However, recomputing the dot product image at each transformation trial entails a massive increase of the computation time in a straightforward implementation. Moreover, the received echo is actually much less anisotropic than would be the case with a perfectly specular reflection because of the diffraction of the ultrasound beam at the interfaces. Thus, we believe that ignoring the gradient direction is a good first order approximation on the theoretical side and a good tradeoff from the computational point of view.

Another feature of the US images that has been occulted in our functional model is the presence of speckle. Reflection of the US waves do not only occur at acoustic impedance boundaries, but also within tissues due to small inhomogeneities (compared to the US wavelength) that are almost invisible in the MR. As a consequence, homogeneous tissue regions generally appear in the US image with a non-zero mean intensity and a strong texture. Also, since our model does not take attenuation into account, unrealistic intensity values are predicted in some areas, especially outside the skull. These limitations of our model strongly suggest to robustify our criterion, which will be done in Section 2.4.

2.3 Parametric intensity fit

We are now looking for a bivariate intensity mapping f with floating values for the MR gradient component, instead of a simple function of already discretized intensity values in the standard Correlation Ratio (CR) registration method. The problem is that there is no natural way to discretize the MR gradient values and even if we found one that respects the dynamic of the gradient values, the number of parameters of the (discretized) function f grows as the square (e.g. 256^2 instead of 256 only for the classical CR). Thus, instead of the close-form solution of the discrete version, we chose to use a continuous but regularized description of the mapping f: we restrain our search to a polynomial function f of degree d. The number of parameters describing f then reduces to $(d+1)(d+2)/2$. In this paper, the degree was set to $d = 3$, implying that 10 coefficients were estimated. Finding the coefficient of the polynomial minimizing eq (4) amounts to solve a weighted least square linear regression problem which is solved using the singular value decomposition (SVD).

However, this polynomial fitting procedure adds a significant extra computational cost and cannot be done for each transformation trial of a Brent/Powell minimization as before. Instead, we chose to perform an alternated optimization of the criterion along the transformation T and the intensity mapping f. In the first step, we find the best polynomial f (at a fixed T) and remap J and $\|\nabla J\|$ accordingly. In the second step, we minimize $C(T, f)$ with respect to T using Powell's method given the intensity corrected image $f(J, \|\nabla J\|)$. This iterative process is stopped when T and f do not evolve any more. This alternate minimization strategy saves us a lot of computation time (speed-up factors are in the range of 2 to 10 when setting the polynomial degree to $d = 3$). This is guaranteed to converge at least to a local maximum of the registration criterion. In practice, we did not observe any alteration of the performances with respect to the original technique.

2.4 Robust intensity distance

Our method is based on the assumption that the intensities of the US may be well predicted from the information available in the MR. Due to several ultrasound artifacts, we do not expect this assumption to be perfectly true. As discussed above, shadowing, duplication or interference artifacts (speckle) may cause large variations of the US intensity from its predicted value, even when the images are perfectly registered. To reduce the sensitivity of the registration criterion to these outliers, we propose to use a robust estimation of the intensities differences using a one-step S-estimator [39]: the quadratic error function $\int \left[I(x) - f(J(T(x)))\right]^2 .dx$ is then replaced with

$$S^2(T, f) = \frac{S_0^2}{K} . \int \Phi\left(\frac{[I(x) - f(J(T(x)))]}{S_0}\right) .dx,$$

where K is a normalization constant that ensures consistency with the normal distribution, and S_0 is some initial guess of the scale. In our implementation, we have opted for the Geman-McClure redescending function $\Phi(x) = \frac{1}{2}x^2/(1 + x^2/c^2)$, for its computational efficiency and good robustness properties, to which we always set a cut-off distance $c = 3.648$ corresponding to 95% Gaussian efficiency.

The new registration criterion requires very few modifications in our alternate optimization scheme: the polynomial function f is now estimated using a simple iterative WLS procedure and the transformation is still found using Powell's method. Initially, the intensity mapping f is estimated in a non-robust fashion. The starting value S_0 is then computed as the median of absolute intensities deviations. Due to the initial misalignment, it tends to be overestimated and may not efficiently reject outliers. For that reason, it is re-estimated after each alternated minimization step.

3 Tracking deformations in US images

Now that we have the method to register the MR image to the first US image, the goal is to track the brain deformations in 3D US time-sequences. When analyzing the problem, we made the following observations. First, deformations are small between successive images in a real-time sequence, but they are possibly large deformations around the surgical tools with respect to the pre-operative image. Thus, the transformation space should allow large deformations, but only small deformations have to be retrieved between successive images. Second, there is a poor signal to noise ratio in US images and the absence of information in some areas. However, the speckle (inducing localized high intensities) is usually persistent in time and may produce reliable landmarks for successive images [40]. As a consequence, the transformation space should be able to interpolate in areas with few information while relying on high intensity voxels for successive images registration. Last but not least, the algorithm is designed in view of a real-time registration during surgery, which means that, at equal performances, one should prefer the fastest method.

Following the encouraging results obtained in [27] for the intensity based free-form deformations registration of two 3D US images, we have adapted in [33] the method according to the previous observations. In the sequel of this section, we first detail the parameterization of our non-rigid transformations. Then we investigate the similarity criterion and the optimization strategy. Finally, we show how to turn this registration algorithm into a tracking tool suited for time sequences [32].

3.1 Parameterization of the transformation

Simple transformations, like rigid or affine ones, can be represented by a small number of parameters (resp. 6 and 12 in 3D). When it comes to Free-Form Deformations (FFD), we need to specify the coordinates $T(x)$ of each point x of the image after the transformation. Such a non-parametric transformation is usually represented by its displacement field $U(x) = T(x) - x$ (or $U = T - \mathrm{Id}$), sampled at each voxel, and a regularization term has to be added to the criterion in order to reduce the effective number of degrees of freedom. This strategy proved to be successful in textured enough regions but induces some convergence problems in large uniform areas (as it is the case in the phantom sequence of section 4.1.3).

We found that a re-parameterization of the transformation was necessary to promote a better conditioning of the problem. In a standard free form deformation (FFD), we have a displacement t_i for each voxel position x_i. Let now t_i be a parameter of a smooth transformation defined by:

$$T(t_1, ...t_n)(x) = \sum_i t_i . G_\sigma(x - x_i) \tag{5}$$

Notice that when σ goes to 0, the parameterization tends toward the standard FFD description. The transformation being described a sum of Gaussians, rather than a sum of Diracs, the gradient descent algorithm uses the derivatives of the similarity with respect to the displacement of an entire group of voxel, which is more robust to noise and may propagate the motion more rapidly in uniform areas. This new parameterization can be seen as a regularization of the similarity energy (see Sec. 3.4 below) and not as a regularization of the transformation as we can still interpolate any displacement value at each site x_i.

3.2 Similarity energy

Even if there is a poor signal to noise ratio in US images, the speckle is usually persistent in time and may produce reliable landmarks within the time-sequence [40]. Hence, it is desirable to use a

similarity measure which favors the correspondence of similar high intensities for the registration of successive images in the time-sequence. First experiments presented in [27] indicated that the simplest one, the sum of square differences (SSD), could be suited. Let I be the reference image and $J \circ T$ the transformed image to register; the criterion to minimize is:

$$SSD(T) = \int (I - J \circ T)^2 \,,$$

In [28], we developed a more complex similarity measure: the sum of Gaussian-windowed local correlation coefficients (LCC). Let $G \star f$ be the convolution of f by the Gaussian, $\bar{I} = (G \star I)$ the *local* mean, $\mathrm{Var}(I) = G \star (I - \bar{I})^2$ the *local* variance, $\mathrm{COV}(I, J \circ T) = G \star \left[(I - \bar{I})(J \circ T - \overline{J \circ T})\right]$ the *local* covariance between image I and image $J \circ T$, and finally $\rho(I, J \circ T) = \mathrm{COV}(I, J \circ T)/\sqrt{\mathrm{Var}(I).\mathrm{Var}(J \circ T)}$ the *local correlation coefficient*. Contrarily to Section 2.1, where the mean, variance, covariance and correlation coefficient where global (i.e. constants computed for the whole image), these values are here *functions* of the position in space, just like the images, because of the locality of their computation. In particular, we cannot optimize the local correlation coefficient directly: we have to integrate it to obtain the sum of the local correlation coefficients:

$$LCC(T) = \int \rho(I, J \circ T)(x).dx = \int \frac{\mathrm{COV}(I, J \circ T)}{\sqrt{\mathrm{Var}(I).\mathrm{Var}(J \circ T)}}$$

We have shown in [27] and [28] how the SSD and LCC criteria can be optimized using first and second order gradient descent techniques with a general free-form deformation field by computing the gradient and the Hessian of the criteria. We summarize the main ideas in the following sections.

3.3 Minimizing the SSD for a free-form deformation

Let us consider first the SSD criterion and the standard FFD description of the transformation. Let T be the current estimation of the transformation and $(\nabla_J \circ T)(x)$ (resp. $(\mathcal{H}_J \circ T)(x)$) be the transformed gradient (resp. Hessian) of the image J. The gradient and the Hessian of the SSD criterion are:

$$\begin{aligned} \nabla_{SSD}(T) &= 2(J \circ T - I).(\nabla_J \circ T) \\ \mathcal{H}_{SSD}(T) &= 2(\nabla_J \circ T).(\nabla_J \circ T)^{\mathrm{T}} + 2(J \circ T - I).(\mathcal{H}_J \circ T) \end{aligned}$$

Let us now approximate the criterion by its tangential quadratic form at the current transformation T, or equivalently approximate the criterion gradient at the first order for a small perturbation by a displacement field $u(x)$. We have: $\nabla_{SSD}(T + u) \simeq \nabla_{SSD}(T) + \mathcal{H}_{SSD}(T).u$ Assuming that the Hessian matrix of the criterion is positive definite, the minimum is obtained for a null gradient, i.e. for $u = -\mathcal{H}^{(-1)}_{SSD}(T).\nabla_{SSD}(T)$. This formula require to invert the Hessian matrix $\mathcal{H}_{SSD}(T)$ at each point x of the image. To speed up the process, we approximate this matrix by the closest scalar matrix. Using this approximation, we get the following adjustment vector field:

$$u \simeq \frac{-3.(J \circ T - I).(\nabla_J \circ T)}{\|\nabla_J \circ T\|^2 + (J \circ T - I).(\Delta_J \circ T)}$$

In fact, when minimizing the reverse SSD criterion $\int (I \circ T^{(-1)} - J)^2$, one finds that the optimal adjustment is given by [27]:

$$\hat{T} = T \circ (\mathrm{Id} + u') \qquad \text{with} \qquad u' = \frac{3.(I - J \circ T).\nabla_I}{\|\nabla_I\|^2 + (I - J \circ T).\Delta_I}$$

which justifies the empirical force used by Thirion's demons:

$$v = \frac{(I - J \circ T).\nabla_I}{\|\nabla_I\|^2 + \alpha.(I - J \circ T)^2}$$

In practice, we have modified the Newton optimization scheme described above into a Levenberg-Marquardt method where the adjustment vector field is given at each step by $u = -(\lambda.\mathrm{Id} + \mathcal{H}_{SSD})^{(-1)}.\nabla_{SSD}$. Dropping the (possibly negative) second order terms in the Hessian, we are left with:

$$u = \frac{-3.(J \circ T - I) \cdot (\nabla_J \circ T)}{\|\nabla_J \circ T\|^2 + \lambda^2} \tag{6}$$

The parameter λ performs a tradeoff between a first order gradient descent ($\lambda \gg 1$ means that we don't trust the approximated Hessian matrix and we simply go along the gradient with a small time-step) and a second order gradient descent ($\lambda \ll 1$ means that we use our simplified Hessian matrix). At each step, λ is divided by a fixed value α (typically 5) if the similarity criterion decreased, and the criterion is re-estimated with λ multiplied by α otherwise until the criterion decreases.

3.4 Minimizing the similarity energy for the new parameterization

We now detail the differences induced by our new parameterization of the free-form transformation on the SSD criterion. Using the Gaussian parameterization of the transformation (eq. 5), t_i is now a parameter of the transformation. Deriving the SSD w.r.t. this parameter gives:

$$\nabla_{SSD}(T) \quad = \quad 2\, G_\sigma \star \Big((J \circ T - I).(\nabla J \circ T) \Big)$$

Thus, the Gaussian parameterization acts as a smoothing on the gradient of the energy. Therefore, it will be more robust and may escape from previous local minima. The minimization is performed as above with a Levenberg-Marquardt method using these regularized version of the energy derivatives.

We now tun to the minimization of the LCC criterion. We have detailed in [28] how to compute efficiently this criterion and its gradient using convolutions. Our conclusion was that the following approximation of the gradient was performing an ideal tradeoff w.r.t. computation time:

$$\nabla_{LCC} \simeq \left((I - \bar{I}) - (J \circ T - \overline{J \circ T}).\frac{(\overline{I.J \circ T} - \bar{I}.\overline{J \circ T})}{\sqrt{\mathrm{Var}(J \circ T)}} \right) \frac{\nabla_J \circ T}{\sqrt{\mathrm{Var}(I)} \cdot \sqrt{\mathrm{Var}(J \circ T)}}$$

Like above for the SSD, using our Gaussian parameterization of the displacement field amounts to convolve this gradient with G_σ.

However, as LCC is not any more a least-square problem, it seems difficult to derive a Gauss-Newton minimization scheme as we have done for the SSD. Instead, we remarked that Eq. 6 may be rewritten:

$$u(x) = -3\frac{2.E(x).\nabla E(x)}{\|\nabla E(x)\|^2 + 4\lambda^2.E(x)}$$

where $E(x) = (I(x) - (J \circ T)(x))^2$ is our local similarity energy and $\nabla E(x) = 2.(I(x) - (J \circ T)(x)).\nabla J \circ T(x)$ is its gradient w.r.t. the transformation $T(x)$. With this formulation, it is now easy to replace local similarity energy by our local correlation coefficient LCC (and its gradient). In practice, we found that this adaptation performed very well despite its weak theoretical background.

3.5 Regularization energy

As said before, there are too many degrees of freedom with free-form deformations, and we have to regularize them. Thus, there is a trade-off to find between the similarity energy, reflected by the visual quality of the registration, and the smoothing energy, reflected by the regularity of the transformation.

In the regularization theory framework, one minimizes the weighted sum of the energies: $E_{sim} + \lambda.E_{reg}$. This formulation has proven to be successful for data approximation, and has been used for various approaches of non-rigid registration algorithms [41]. However, there is an important difference between data approximation and image registration. In data approximation, both energies measure different properties of the same object (the similarity and the smoothness of the data), while the two energies relate to different objects in image registration (the intensities of the images for the matching energy and the transformation for the regularization energy). Thus, one has to find a non linear tradeoff between the two energies.

Another widely spread method attempts to separate the image measure from the transformation measure, and could be compared with the approach of game theory. It consists in alternatively decreasing the similarity energy and the smoothing energy. This approach is chosen in many block-matching algorithms [42] and in some optical-flow-based techniques [22]. In view of a real-time system, this is particularly well suited for the stretch energy (or membrane model) $E_{reg} = \|\nabla T|\|^2 = \int \mathrm{Tr}(\nabla T.\nabla T^{\mathrm{T}})$ as the associated Euler-Lagrange evolution equation corresponds to the heat propagation in a homogeneous material. Thus, one step of gradient descent corresponds to convolution of the transformation by a Gaussian with a standard deviation linked to the time step of the gradient descent [43]. This way, we get a simple regularization by a Gaussian smoothing of the transformation parameters t_i with a smoothing parameter (the σ_T of this Gaussian) that has a physical meaning. The final algorithm consists in alternatively performing one step of gradient descent on the similarity energy E_{sim} and one step of transformation smoothing by Gaussian filtering of standard deviation σ_T.

3.6 From registration to tracking

In the previous sections, we studied how to register two US images together. We now have to estimate the deformation of the brain between the first image (which is assumed to correspond to the MR image of the preoperative brain) and the current image of the sequence. One could think of registering directly US_1 (taken at time t_1) and US_n (at time t_n) but the deformations could be quite large and the intensity changes important. To constrain the problem, we need to exploit the temporal continuity of the deformation.

First, assuming that we already have the deformation $T_{US}(n)$ from image US_1 to US_n, we register US_n with the current image US_{n+1}, obtaining the transformation $dT_{US}(n)$. If the time step between two images is short with respect to the deformation rate (which should be the case in real-time sequences at a rate ranging from 1 to 5 images per second), this registration should be easy. Moreover, the intensity changes should be small. For this step, we believe that the SSD criterion is well adapted.

Then, composing with the previous deformation, we obtain a first estimation of $T_{US}(n+1) \simeq dT_{US}(n) \circ T_{US}(n)$. However, the composition of deformation fields involves interpolations and just keeping this estimation would finally lead to a disastrous cumulation of interpolation errors:

$$T_{US}(n+1) = dT_{US}(n) \circ dT_{US}(n-1) \ldots dT_{US}(2) \circ dT_{US}(1)$$

Moreover, a small systematic error in the computation of $dT_{US}(n)$ leads to a huge drift in $T_{US}(n)$ as we go along the sequence.

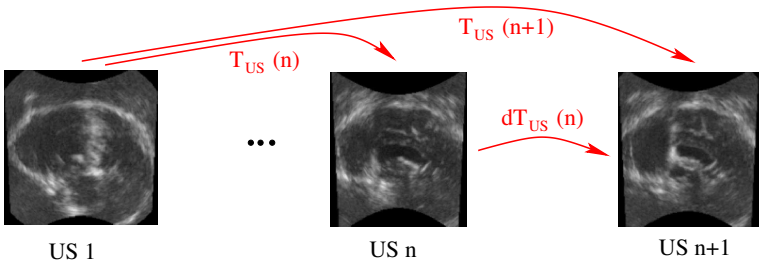

Figure 2: The deformations computed in the tracking algorithm.

Thus, we only use $dT_{US}(n) \circ T_{US}(n)$ as an initialization for the registration of US_1 to US_n. Starting from this position, the residual deformation should be small (it corresponds to the correction of interpolation and systematic error effects) but the difference between homologous point intensities might remain important. In this case, the LCC criterion might be better than the SSD one despite its worse computational efficiency. In practice, we run most of the experiments presented in section 4 with the SSD and LCC criterion without finding significant differences. Since the LCC is still around 2 times slower than the SSD and since the computation time of the US-US non-rigid registration is a key issue for real-time motion tracking, we preferred to keep the SSD criterion. We believe that this choice could be reconsidered if the sequence was to present some important intensity changes along time.

One of the main consequences of our tracking method is that the first US image has to be of very high quality since it will be the only reference for tracking deformations along the whole sequence. One possibility consists in acquiring several images of the still brain in order to compute a mean image of better quality. Another possibility consists in performing some anisotropic diffusion on US_1 to improve its quality.

4 Experiments

In this section, we present quantitative results of the rigid MR/US registration algorithm on real brain images, and qualitative results of the tracking algorithm and its combination with the MR/US registration on animal and phantom sequences. The location of the US probe being linked to the pathology and its orientation being arbitrary (the rotation may be superior to 90 degrees), it was necessary to provide a rough initial estimate of the MR/US transformation. This was done using an interactive interface that allows to draw lines in the images and match them. This procedure was carried out by a non-expert, generally taking less than 2 minutes (see Fig. 3). However this user interaction could be alleviated using a calibration system such as the one described in [4]. After initialization, we observed that the algorithm found residual displacements up to 10 mm and 10 degrees.

4.1 Data

All 3D-US images were acquired using a commercial 3D-US volume scanner Voluson 530 D from Kretz Technology (4-9 MHz, 90 degrees aperture). It should be emphasized that all the US images provided to us in this project were already resampled in a Cartesian format. As a consequence,

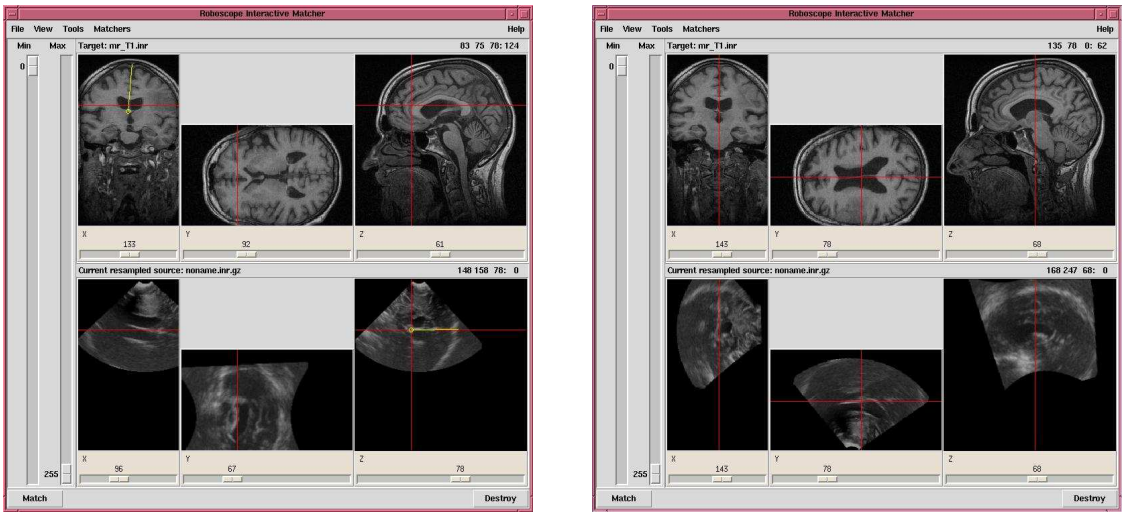

Figure 3: Interactive interface to perform the MR/US initial registration. Left: the users specifies a corresponding point and direction in one of the slice of each image. Right: the software computes the 3D rigid transformation that superimposes the two slices and the user features. The process is usually iterated twice with the remaining slices in order to better adjust the transformation. The whole process generally takes less than 2 minutes.

the images suffer from interpolation (blurring) artifacts in areas far from the probe, and the details close to the probe are averaged out. Moreover, border parts of the images were often clipped to fit into a cubic shape, which implied some important loss of information. In the future, we will have to deal directly with the original geometry of the US images.

4.1.1 Baby dataset

This clinical dataset was acquired to simulate the degradation of the US images quality with respect to the number of converters used in the probe. Here, we have one MR T1 image of a baby's head and 5 transfontanel US images with different percentages of converters used (40%, 60%, 70%, 90% and 100%). The MR image has 256x256x124 voxels of size 0.9 mm^3. The Cartesian US images have 184x184x184 voxels of size 0.29x0.29x0.29 mm^3.

As we have no deformations within the images, we can only rigidly register all the US images onto our single MR. An example result is presented in Fig.4. The visual quality of the registration is very good. We did not notice a significant change of the registration result with respect to the number of converters used in the US probe. This emphasizes the robustness of the information used in the US image for the registration (basically the mid-sagittal plane and the ventricular surface). However, other experiments with neonates images showed that the quality of the MR was a crucial parameter for the registration to succeed, since it is difficult to keep a neonate motionless in a MRI scanner all along the acquisition time.

4.1.2 Patient images during tumor resection

This dataset is an actual surgical case: two MR T1 images with and without a contrast agent were acquired before surgery. After craniotomy (dura mater still closed), a set of 3D US images was acquired to precisely locate the tumor to resect. The MR images have 256x256x124 voxels of size

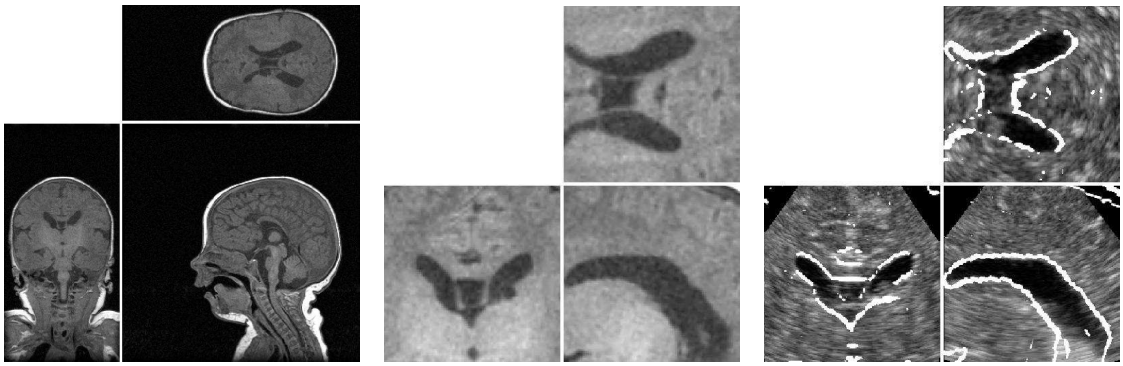

Figure 4: Example registration of MR and US images of the baby. From left to right: original MR T1 image, closeup on the ventricle area, and registered US image with MR contours superimposed.

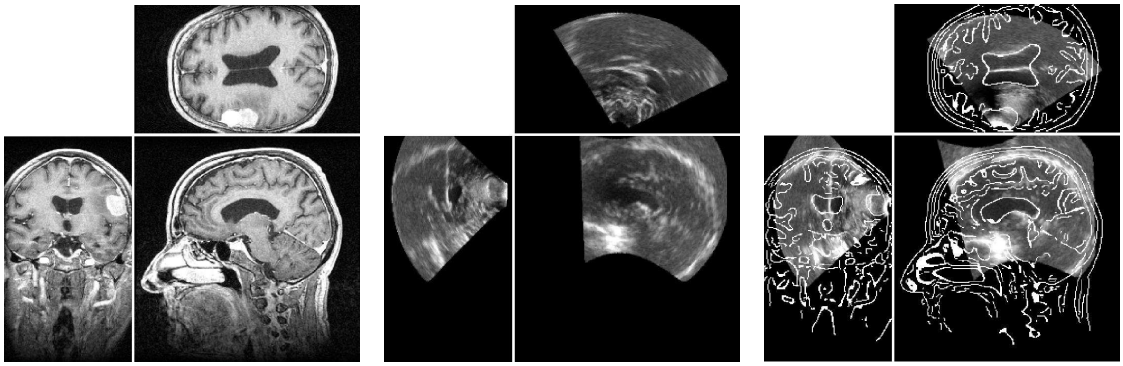

Figure 5: Example registration of MR and US images of the patient. From left to right: MR T1 image with a contrast agent, manual initialization of the US image registration, and result of the automatic registration of the US image with the MR contours superimposed.

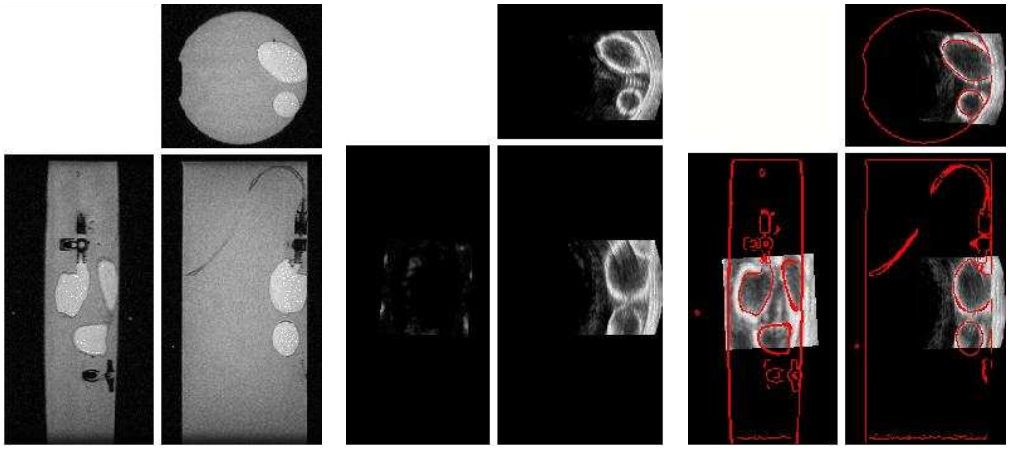

Figure 6: Example registration of MR and US images of the phantom. From left to right: MR image, manual initialization of the US image registration, and result of the automatic registration of the US image with the MR contours superimposed.

0.9x0.9x1.1 mm^3, while the US images have various dimensions with a cubic voxel size ranging from 0.17 mm to 0.95 mm.

In this experiment, we use the three US images that are large enough to contain the ventricles. Unfortunately, we could only test for the rigid MR/US registration as we have no US images during surgery. An example of the registration results is presented in Fig.5. One can see that the ventricles and some of the sulci visible in the US image are very well matched. When looking at the image super-imposition more closely (using a dynamic fusion visualization), it appears that a small residual deformation (about 1 to 2mm maximum) is not corrected. This deformation could be due to the sound speed that vary depending on the brain tissues (it is assumed to be constant for the US image reconstruction), or to a real brain deformation during the opening of the skull. More experiments would be need to determine the right hypothesis. Nevertheless, the validation scheme of Section 4.2.3 shows that a registration accuracy of 0.6 mm at the center and 1.6 mm in the whole brain area is achieved. This example is interesting at it demonstrate that the MR/US registration could be performed in clinical conditions with a sufficient accuracy.

4.1.3 A Phantom study

For the evaluation of the US images acquisitions, Prof. Auer and his colleagues at ISM (Austria) developed an MR and US compatible phantom to simulate brain deformations. It is made of two balloons, one ellipsoid and one ellipsoid with a "nose", that can be inflated with known volumes. Each acquisition consists in one 3D MR and one 3D US image. Both balloons were initially filled with 40 ml of fluid. During a first phase, the small ellipsoidal balloon only was filled by steps of 10 ml (acquisitions 1 to 4 in Fig 7). Then, this balloon was deflated, while the "nose" balloon was filled (still by steps of 10 ml), thus creating much larger deformations (acquisitions 5 and 6 in Fig 7). The MR images have 256 x 256 x 124 voxels with a voxel size of 0.9 x 0.9 x 1 mm^3. The Cartesian US images are cubic with 184 voxels of size 0.41 mm on each side.

The first MR image was rigidly registered to the first US image with our bivariate correlation ratio method. An example result is provided in Fig. 6. Then, deformations are estimated using the tracking algorithm on the US sequence, and the corresponding virtual MR image is computed. Since the US probe had to be removed from the phantom for each MR scans, we had to rigidly re-register all the US images before tracking in order to minimize the amount of global motion between US frames. Due to this resampling and to the too focused conversion from polar to Cartesian coordinates by the US machine, one can observe some missing information in the borders (first row of Fig.7). Last but not least, the remaining MR images can be used to assess the quality of the tracking thanks to a rigid MR/US registration for each acquisition (third and fourth row of Fig.7).

Even if there are very few salient landmarks (all the information is located in the thick and smooth balloons boundaries, and thus the tracking problem is loosely constrained), results are globally good all along the sequence. This shows that the SSD criterion correctly captures the information at edges and that our parameterized deformation interpolates reasonably well in uniform areas. When looking at the virtual MR in more details, one can however find some places where the motion is less accurately recovered: the contact between the balloons and borders of the US images. Indeed, the parameterization of the transformation and especially its smoothing are designed to approximate the behavior of a uniform elastic like body. If this assumption can be justified for the shift of brain tissues, it is less obvious for our phantom where balloons are placed into a viscous fluid. In particular, the fluid motions between the two balloons cannot be recovered. On the borders of the US images, there is often a lack of intensity information (due to the inadequate

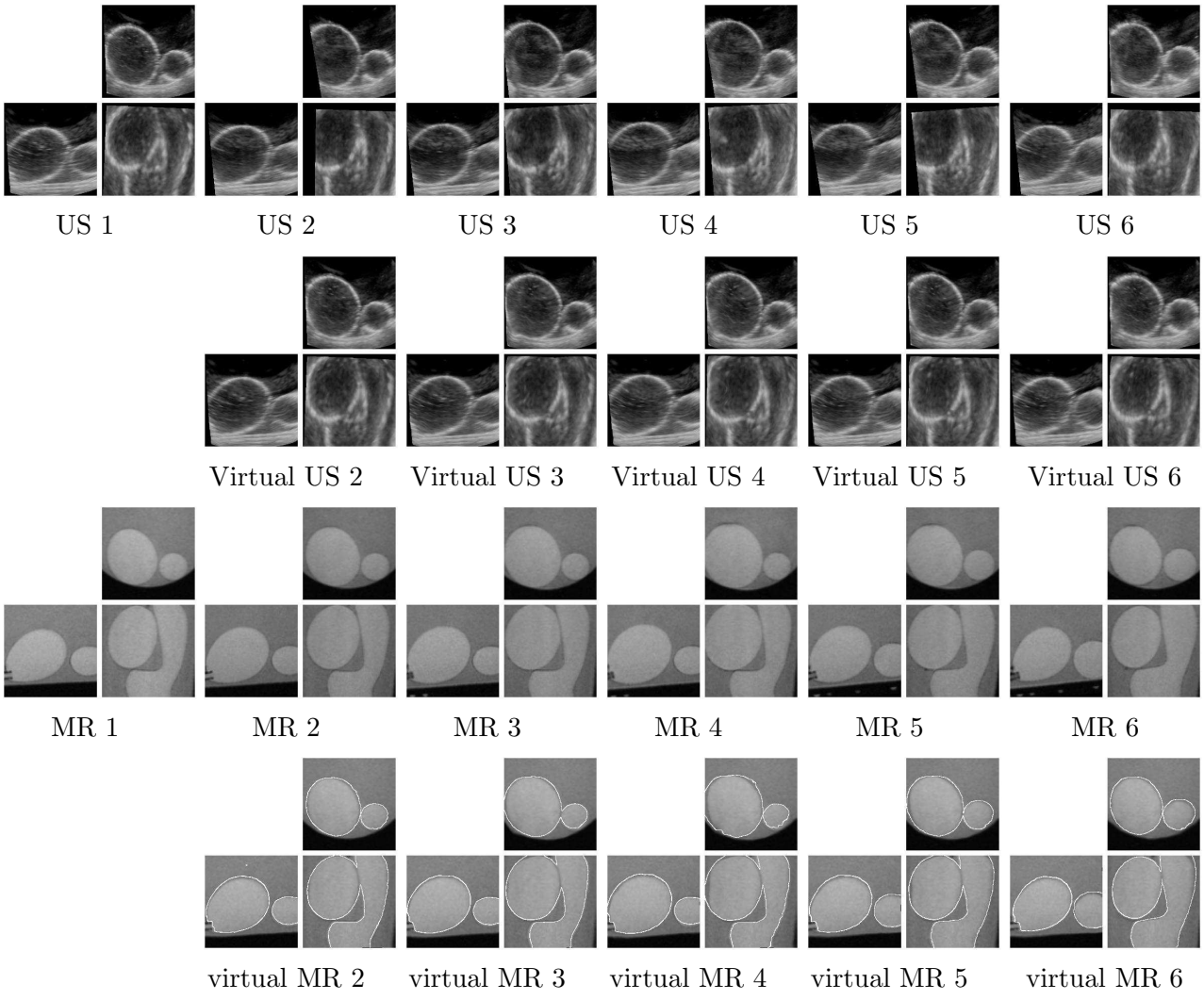

Figure 7: **Top:** 6 US images of the sequence after a rigid registration to compensate for the motion of the probe and the "virtual" US images (US 1 deformed to match the current US image) resulting from the tracking. **Bottom:** The "original" MR images (rigidly registered to the corresponding US images to correct for the probe motion and the phantom motion between MR acquisitions) and the virtual MR image synthesized using the deformation field computed on the US images. To assess the quality of the tracking, we superimposed the contours of the "original" MR images. The volume of the balloons ranges from 60 to 90 ml for the ellipsoid one and 40 to 60 ml for the more complex one. Reprinted, with permission, from *Pattern Recognition Letters 24 (4-5), X. Pennec, P. Cachier, and N. Ayache: Tracking Brain Deformations in Time-Sequences of 3D US Images; pages 801–813, Copyright Elsevier February 2003.*

conversion from polar to Cartesian coordinates by the US machine) and the deformation can only be extrapolated from the smoothing of neighboring displacements. Since we are not using a precise geometrical and physical model of the observed structures, one cannot expect this extrapolation to be very accurate.

4.1.4 US images of a balloon inflated in a pig brain

This dataset was obtained by Dr. Ing. V. Paul at IBMT, Fraunhofer Institute (Germany) from a pig brain at a post-lethal status. A cyst drainage has been simulated by deflating a balloon catheter with a complete volume scan at three steps. All US images have a Cartesian dimension of 184x184x184 voxels of size 0.26x0.26x0.26 mm^3. Unfortunately, we have no MR image in this dataset, so that we could only run the tracking algorithm.

We present the results in Fig. 8. Since we have no corresponding MR image, we present on the two last lines the deformation of a grid (a virtual "MR" image...), to emphasize the regularity of the estimated deformation, and the deformation of a segmentation of the balloon. The registration of each image of the sequence takes between 10 and 15 minutes on a Pentium II 450 Mhz running Linux. To visually assess the quality of the registration, we segmented the balloon on the first image. Then, this segmentation is deformed using the transformation found and superimposed to the corresponding original US image (last row of figure 8).

The correspondence between the original and the virtual (i.e. deformed US 1) images is qualitatively good. In fact, if the edges are less salient than in the phantom images (see next section), we have globally a better distribution of intensity features over the field of view due to the speckle in these real brain images. One should also note on the deformed grid images that the deformation found is very smooth.

4.2 MR/US rigid registration consistency evaluation

The usual way of measuring a registration accuracy is to provide a Target registration Error (TRE), i.e. the mean RMS error in millimeters that the registration did w.r.t. the ground truth on a specified set of target points. Using the methods developed in [44], we can also measure this error directly on the transformation in order to provide a full covariance matrix on the transformation parameters. As this matrix is rather difficult to interpret, we can simplify it and provide the mean rotation error σ_{rot} in degrees and the mean translation error σ_{trans} in mm. We may also propagate this covariance matrix onto a set of pre-specified target or test points to obtain a TRE σ_{test}.

4.2.1 Registration loops

It is very difficult to obtain a gold standard registration between MR and US images since external markers cannot be seen in US images (US waves do not properly propagate in the air) and even internal markers detection is prone to inaccuracy. Moreover, if such a calibration setup was possible for phantom data in a controlled environment, it is not really feasible for clinical data. To get around the problem, we adapted in [31] the registration loop protocol introduced in [45, 30, 46, 47].

The principle is to acquire several MR and US images of the same subject (or object), and to compose a series registration transformations in a loop that lead ideally to the identity transformation. Typical loops in the case of the phantom data, sketched in Fig. 9, are sequences like $MR_i \rightarrow US_i \rightarrow US_j \rightarrow MR_j \rightarrow MR_i$. If we were given perfectly registered images within each modality, applying the transformations of the loop in turn to a test point in MR_i would lead to a

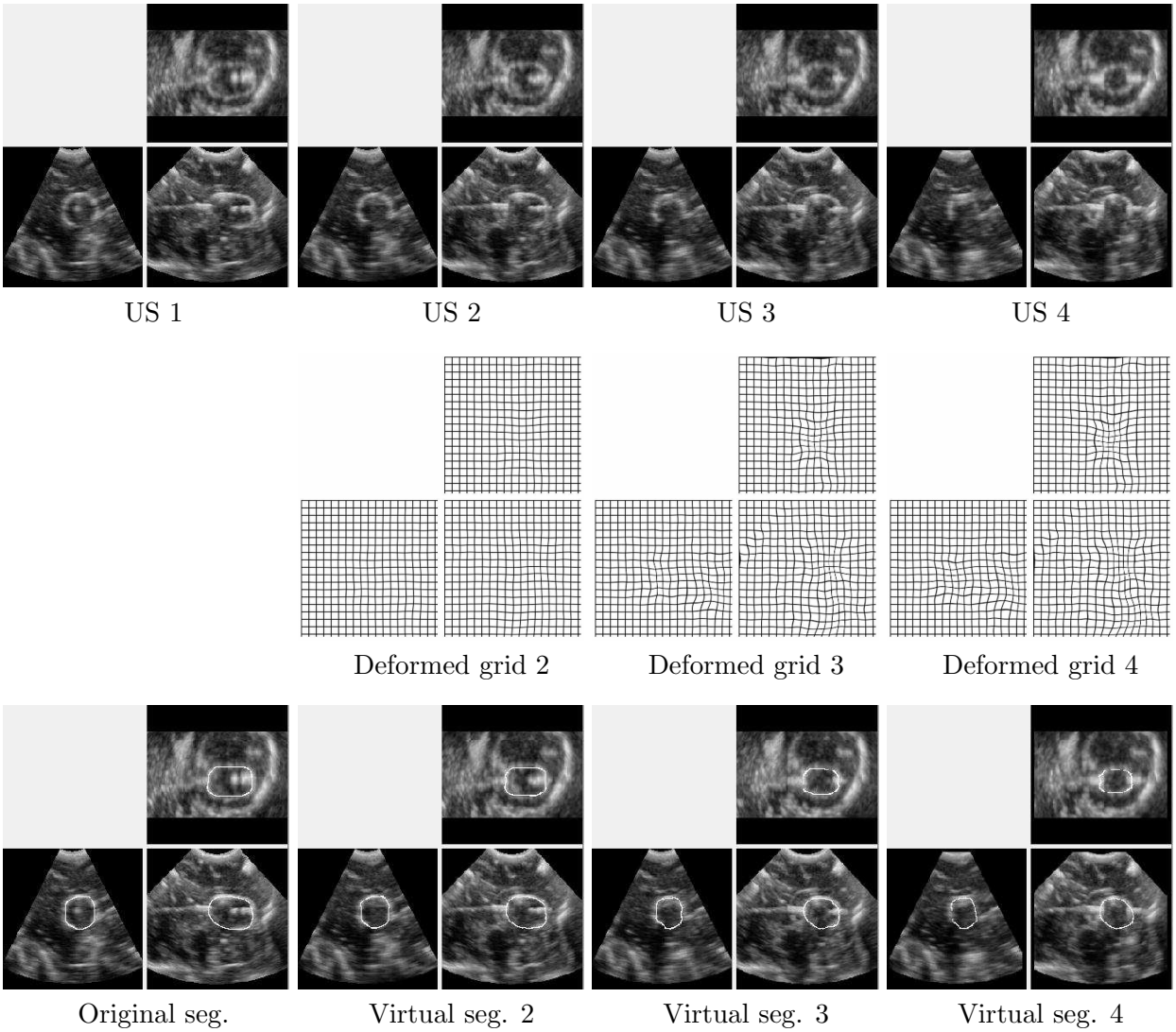

Figure 8: **Tracking deformations on a pig brain**. In this figure, each triplet of 2D images represents 3 orthogonal views resliced from the 3D image. **Top:** The 4 images of the pig brain with a deflating balloon simulating a cyst drainage. **Middle:** deformation of a grid to visualize more precisely the location of the deformations found. These images correspond to the deformation of an image of a 3D grid (a "virtual MR" image) with strips orthogonal to each 2D resliced plane: they allow to visualize the in-plane deformation for each 2D slice. **Bottom:** We segmented the balloon on the first image. Then, this segmentation is deformed using the transformation found and superimposed to the corresponding original US image.

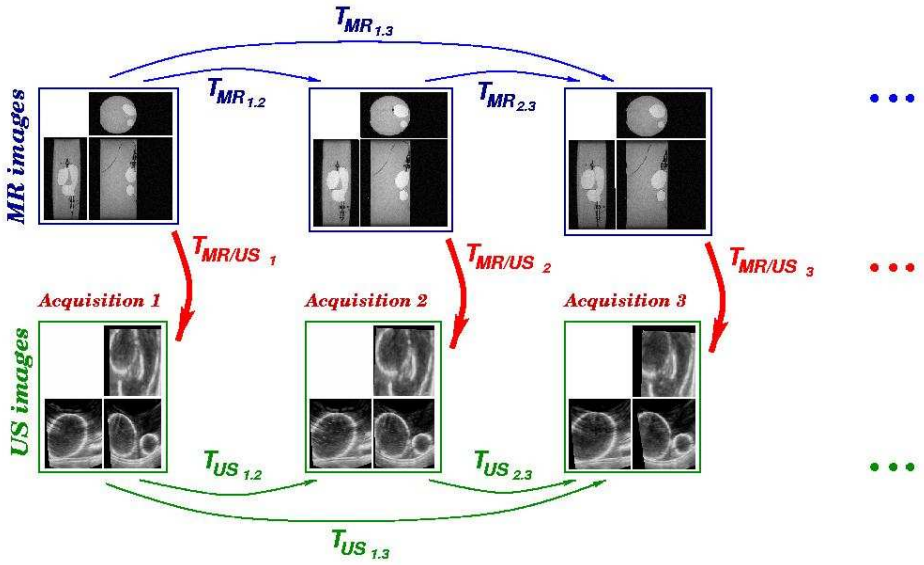

Figure 9: Registration loops used to estimate the registration consistency. Adapted, with permission, from *IEEE Transactions on Medical Imaging, 20(10), A. Roche, X. Pennec, G. Malandain, and N. Ayache, Rigid Registration of 3D Ultrasound with MR Images: a New Approach Combining Intensity and Gradient Information; pages 1038–1049, Copyright IEEE October 2001.*

displacement of this test point that it due to the errors of the two MR/US registrations. Since variance are additives, the observed TRE should be $\sigma^2_{loop} = 2\sigma^2_{MR/US}$. In most cases, the intramodality transformations are not perfectly known and need to be estimated. Thus, we have to take into account their variability and the error on the loop becomes: $\sigma^2_{loop} = 2\sigma^2_{MR/US} + \sigma^2_{MR/MR} + \sigma^2_{US/US}$. Finally, we can estimate the *expected accuracy* or consistency using

$$\sigma_{MR/US} = \sqrt{\frac{1}{2}(\sigma^2_{loop} - \sigma^2_{MR/MR} - \sigma^2_{US/US})}$$

Notice that neglecting the intramodality errors and setting $\sigma_{MR/US} = \sigma_{loop}/\sqrt{2}$ amounts to take a more conservative error estimation.

Using different MR and US images in the loop allows to de-correlate the MR/US transformation (none of the data is common to the registrations), but this is not always possible. For instance, we have only one MR image in the Baby dataset. Moreover, if the intramodality registration is also done using the images and not using external informations (e.g. markers), a special feature of one image may similarly affect the different registrations involved in the loop and consequently hides the bias due to the variability of this feature. This is why the registration loop protocol only leads to a measure of consistency which is a lower bound on the accuracy.

4.2.2 Bronze standard registration

Our goal now is to compute the $n-1$ most reliable transformations $\bar{T}_{i,i+1}$ that relate the n (successive) MR_i (or US_i) images. Estimations of these transformations are readily available by computing all the possible registrations $T_{i,j}$ between the MR images using m different methods. Then, the transformations $\bar{T}_{i,i+1}$ that best explain these measurements are computed by minimizing the sum

of the squared distance between the observed transformations $T_{i,j}$ and the corresponding combination of the sought transformation $\bar{T}_{i,i+1} \circ \bar{T}_{i+1,i+2} \ldots \bar{T}_{j-1,j}$. The distance between transformations is chosen as a robust variant of the left invariant distance on rigid transformation developed in [45].

In this process, we do not only estimate the optimal transformation, but also the rotational and translational variance of the "transformation measurements", which are propagated through the criterion to give an estimated of the variance of the optimal transformations.

The estimation $\bar{T}_{i,i+1}$ of the perfect registration $T_{i,i+1}$ is called *bronze standard* because the result converges toward the true transformation $T_{i,i+1}$ if there is a sufficiently high number of images (to decrease the noise level) up to the intrinsic bias (if there is any) introduced by the method. Now, using different registration procedures based on different methods, the intrinsic bias of each method also becomes a random variable, which is hopefully centered around zero and averaged out in the minimization procedure. The different bias of the methods are now integrated into the transformation variability. To fully reach this goal, it is important to use as many independent registration methods as possible.

In our setup, we used as many images as possible within each dataset, and up to three different methods with different geometric features or intensity measures. Two of these methods are intensity-based and were used for both MR/MR and US/US registrations: the algorithm `aladin` [42] has a block matching strategy where matches are determined using the coefficient of correlation (CC), and the transformation is robustly estimated using a least-trimmed-squares; the algorithm `yasmina` uses the Powell algorithm to optimize the SSD or a robust variant of the correlation ratio (CR) metrics between the images [31]. For the MR/MR registration, we added a feature-based methods, the `crest lines` registration described and thoroughly evaluated in [48].

4.2.3 Consistency results

We ran the MR/US registration on the phantom, baby and patient datasets. Results are summarized in Table 1. The optimization of the bivariate correlation ratio was realized using Powell's method and took 5 to 10 min on a Pentium II 450 MHz running Linux.

For the Phantom data, we used the 54 loops $US_i \rightarrow MR_i \rightarrow MR_j \rightarrow US_j \rightarrow US_i$. Eight test points were placed at the corners of the Cartesian US image in order to fully enclose the balloons area. On can see that the measured consistency is of the order of the MR voxel size. One could probably reduce slightly the uncertainty of the US/US registration by incorporating rigid markers in the phantom. This would strengthen our belief in this result, but we believe that it would not drastically change it.

In the case of the baby dataset, we used 5 loops $US_i \rightarrow MR \rightarrow US_j \rightarrow US_i$ since only one MR image was available. In this case, the loop may hide a MR induced bias as the two MR/US registrations share the same MR image. We indeed observe a consistency which is slightly less than the MR voxel size. This better result may also be due to US images that are much more focused on the ventricular area (we experimentally observed that the ventricles were the main intensity features practically used by the registration algorithm).

For the per-operative ("patient") data, the three US images cannot be easily registered into the same coordinate system as they were acquired at different times of the operation. Thus, we were left with only 3 registration loop $US_i \rightarrow MR_1 \rightarrow MR_2 \rightarrow US_i$. Our validation scheme exhibit a registration accuracy of 0.6 mm at the center and 1.6 mm in the whole brain area (the test points are taken here at the corners of a 80 x 80 x 80 x 80 mm^3 cube centered in the Cartesian US image so that they lie in the acquisition cone). However, when we look more carefully at the results, we find that the loop involving the smallest US image (real Cartesian size $150 \times 85 \times 100$ mm, voxel

Table 1: Estimated registration errors: standard deviations of the residual rotation (σ_{rot}), of the translation (σ_{trans}) - which corresponds to the error of a test point at the center of the image -, and on the test points that enclose the area of interest (σ_{test}). Reprinted, with permission, from *IEEE Transactions on Medical Imaging, 20(10), A. Roche, X. Pennec, G. Malandain, and N. Ayache, Rigid Registration of 3D Ultrasound with MR Images: a New Approach Combining Intensity and Gradient Information; pages 1038–1049, Copyright IEEE October 2001.*

		σ_{rot} (deg)	σ_{trans} (mm)	σ_{test} (mm)
Phantom dataset	Multiple MR/MR	0.06	0.10	0.13
MR: $0.9 \times 0.9 \times 1$ mm^3	Multiple US/US	0.60	0.40	0.71
US: $0.41 \times 0.41 \times 0.41$ mm^3	Loop	1.62	1.43	2.07
	Expected MR/US	1.06	0.97	1.37
Baby dataset	Multiple US/US	0.10	0.06	0.12
MR: $0.9 \times 0.9 \times 0.9$ mm^3	Loop	1.71	0.51	1.27
US: $0.3 \times 0.3 \times 0.3$ mm^3	Expected MR/US	1.21	0.36	0.89
Patient dataset	MR/MR	0.06	0.06	0.10
MR: $0.9 \times 0.9 \times 1.1$ mm^3	Loop	2.22	0.82	2.33
US: 0.63^3 and 0.95^3 mm^3	Expected MR/US	1.57	0.58	1.65

size 0.63^3 mm^3) is responsible for a test point error of 2.6 mm (0.85 mm at the center) while the loops involving the two larger US images (real size $170 \times 130 \times 180$, voxels size 0.95^3 mm^3) do have a much smaller test point error of about 0.84 mm (0.4 mm at the center). We suspect that a non-rigidity in the smallest US could account for the registration inaccuracy. Another explanation could be a mis-estimation of the sound speed for this small US acquisition leading to a false voxel size and once again the violation of the rigidity assumption.

4.3 3D US tracking performances

When it comes to 3D deformations (and even worse in our case: 3D non-rigid tracking), it becomes much more difficult to asses the algorithm performances since they depend on the location in space.

Qualitatively, we expect to obtain good matches were we have some time and space consistent information (edges, speckle), but we need to regularize in other places. This means that the registration quality highly depends on the fit between the regularization model and the mechanical properties of the observed tissues. On a quantitative point of view, we would need to know the displacement of each point in the images in order to evaluate the accuracy of the estimated deformation. If it is sometimes possible to determine some fuzzy corresponding landmarks in our US images, it seems difficult to generalize that approach to the whole image.

However, we can usually determine corresponding surfaces, and measure the distance between the surfaces or the difference in the volume they enclose. In the Phantom case, we segmented the surface of the balloons in the MR images. The maximal distance between the deformed surface of the first MR and the surface of the current MR was 2 to 3 mm all along the sequence, while the mean distance was below one millimeter. This shows that the tracking is quite accurate at the places where we have information (here the balloon edges). With these data, we have no mean to assess the accuracy of the "interpolated" motion within and outside the balloons.

In the pig brain images, we have no MR and it is much more difficult to segment faithfully the edges of the inflated balloon in each US images. However, we know the theoretical volume of the

balloon in each image. Thus, we (approximately) segmented the balloon in the first US image, and deformed its volume after deformation (bottom line of Fig. 8). Since our segmentation tends to overestimate the balloon volume, we believe that it is more interesting to compare the ratio of the volumes than the volumes themselves.

Image number	1	2	3	4
Original balloon volume (cm^3)	1.25	1.00	0.75	0.5
Relative volume ratio		0.8	0.6	0.4
Measured balloon volume	1.28	1.10	0.80	0.67
Measured volume ratio		0.86	0.62	0.53

The measurements indicates that we are overestimating the volume (underestimating the deformation) by 7.5% for image 2, by 3.3% for image 3, and by 30% for image 4. However, one should note that volume measurements are very sensitive as they relate to the cube of the balloon dimension: this corresponds to an error of less than one millimeter on the balloon diameter. This could be explained by an occlusion of the lower part of the balloon probably due to an air bubble trapped inside the balloon during the experience: on US 4, almost the entire lower half of the balloon is shadowed by the air bubble. In these conditions, one cannot expect a perfect retrieval. The estimated deformation at the occlusion being computed thanks to the regularization of the deformation field from neighboring structures, it is expected to be less than the real deformations (maximal at the balloon boundaries).

In both pig brain and phantom experiments, reducing the smoothing of the transformation could allow the algorithm to find a closer fit. However, this could allow some unwanted high frequency deformations due to the noise in the US images. We believe that it is better to recover the most important deformations and miss some smaller parts than trying to match exactly the images and have the possibility to create some possibly large deformations.

5 Discussion

We presented in the first part a new automated method to rigidly register 3D US with MR images. It is based on a multivariate and robust generalization of the correlation ratio (CR) measure that allows to better take into account the physics of the US images acquisition. The assumption is that the US signal may be approximated by a function of both the MR intensity and its gradient magnitude. This model does not take into account the speckle and attenuation effects, which are important aspects of the US physics. However, as we introduced a robust intensity distance measure in our bivariate correlation ratio criterion, the functional relationship assumption does not need to hold throughout the whole image.

Our implementation using Powell's optimization method was successful in registering more than 20 MR/US volume pairs from phantom and clinical data. To evaluate the registration performances, we designed an original approach to establish a "bronze standard" in the absence of ground truth. We found the worst registration errors (maximum error in the region of interest defined by the US cone) to be of the order of the MR voxel size (1 to 1.5 mm). A robustness study performed in [31] also showed that the bivariate CR significantly outperforms the conventional correlation ration (CR) and Mutual Information (MI) similarity metrics. Incidentally, we believe that the generalized CR could be considered in other multi-modal registration problems where conventional similarity measures experiment a lack of robustness, such as CT or PET to MR registration [49].

We believe that our algorithm may be improved in several ways to better take into account the specific nature of the US images. The first improvement would be use sampling techniques adapted

to the polar geometry of the US image and the speckle size [50]. The second improvement would be to use the gradient orientation w.r.t. the the US scan-line in addition to its magnitude. Another interesting development would be to take into account (and possibly estimate during registration) the spatial variation of the speed of the sound within tissues. This last development would allow for the correction of the main US image distortions without going for a full free-form deformation.

In the second part of this chapter, we developed a tracking algorithm adapted to time sequences of US images and not only to the registration of two images. The algorithm is able to recover an important part of the deformations and issues a smooth deformation, despite the noisy nature of the US images. Experiments on animal and phantom data show that this allows to simulate virtual MR images qualitatively close to the real ones. Quantitative measurement remains to be done, but it seems that an accuracy of 1-2 mm is achievable in the areas where there is an elastic deformation. This is encouraging since the accuracy of the clinicians without per-operative imaging is estimated to be around 3-5 mm. In our experiments, we observed that the SSD criterion was well adapted to the registration of successive US images of the time sequence, an also performed well for the update of the global transformation. However, the appearance of some tissues in US images is known to change along time, and we believe that the LCC criterion will be more adapted to times sequences of brain US images in a clinical setup.

The computation times are still far from real time for a continuous tracking of deformations during surgery, but this implementation was focused on generic components in order to test different criteria and gradient descent approaches. A dedicated re-implementation of the method may gain a factor 4 to 8, leading to a clinically useful tool for brain shift estimation (one estimation every minute or 2). To be further accelerated and reach real-time (video rate for instance), the algorithm needs to be parallelized. This has been partly realized in [51], with an acceleration of a factor 11 on 15 processors. However, the limiting factor becomes then the latency of the network to transfer the input images ($184^3 = 6$ Mb in the case of our US images) and more importantly the resulting deformation field (72 Mb at full resolution). This last transfer could be avoided by only providing the current resampled MR image.

The type of transformation is a very sensitive choice for such a tracking algorithm. In this work, we made the assumption of a "uniform elastic" like material. If this is adapted for the brain tissues (white and gray matter), some improvements of the algorithm will be needed to cope with the non-elastic deformations that occur with the CSF leakage (particularly in the ventricles). Likewise, the introduction of surgical instruments will create shadows and artifacts that will hide important information in the US image, and may mislead the registration algorithm. Thus, we would need to work with dense transformation with a space-varying regularization, depending on the underlying brain tissue type, and also a space-varying trade-off between the similarity and regularization criteria, in order to use image information where it is reliable and regularization around an behind instruments. First steps toward these goals have been done recently with the introduction of an anatomically informed regularization of the deformation [52] with an adaptive similarity/regularization tradeoff, implemented in a GRID compatible framework. Results are already very good on the registration of MR images and we believe that the adaptation to time series of US images will lead to excellent results.

Acknowledgments

This work was partially supported by the EC-funded ROBOSCOPE project HC 4018, a collaboration between The Fraunhofer Institute (Germany), Fokker Control System (Netherlands), Imperial College (UK), INRIA (France), ISM-Salzburg and Kretz Technik (Austria). The authors address

special thanks to Dr. Ing. V. Paul at IBMT, Fraunhofer Institute for the acquisition of the pig brain images, and to Prof. Auer and his colleagues at ISM for the acquisitions of all other images.

This chapter synthesizes and partly reprints material, with permission, from *IEEE Transactions on Medical Imaging, 20(10), A. Roche, X. Pennec, G. Malandain, and N. Ayache, Rigid Registration of 3D Ultrasound with MR Images: a New Approach Combining Intensity and Gradient Information; pages 1038–1049, Copyright IEEE October 2001* for the MR/UR registration part, and from *Pattern Recognition Letters 24 (4-5), X. Pennec, P. Cachier, and N. Ayache: Tracking Brain Deformations in Time-Sequences of 3D US Images; pages 801–813, Copyright Elsevier February 2003* for the tracking part.

References

[1] J.W. Trobaugh, W.D. Richard, K.R. Smith, and R.D. Bucholz. Frameless stereotactic ultrasonography - method and applications. *Computerized Medical Imaging and Graphics*, 18(4):235–246, 1994.

[2] J.W. Trobaugh, D.T. Trobaugh, and W.D. Richard. Three-dimensional imaging with stereotactic ultrasonography. *Computerized Medical Imaging and Graphics*, 18(5):315–323, 1994.

[3] W.D. Richard, D.M. Zar, E.L LaPresto, and C.P. Steiner. A low-cost PCI-bus-based ultrasound system for use in image-guided neurosurgery. *Computerized Medical Imaging and Graphics*, 23(5):267–276, 1999.

[4] N. Pagoulatos, W.S. Edwards, D.R. Haynor, and Y. Kim. Interactive 3-D Registration of Ultrasound and Magnetic Resonance Images Based on a Magnetic Position Sensor. *IEEE Transactions on Information Technology In Biomedicine*, 3(4):278–288, December 1999.

[5] H. Erbe, A. Kriete, A. Jödicke, W. Deinsberger, and D.-K. Böker. 3D-Ultrasonography and Image Matching for Detection of Brain Shift During Intracranial Surgery. *Computer Assisted Radiology*, pages 225–230, 1996.

[6] N. Hata, M. Suzuki, T. Dohi, H. Iseki, K. Takakura, and D. Hashimoto. Registration of Ultrasound echography for Intraoperative Use: A Newly Developed Multiproperty Method. In *Proc. of VBC'94*, SPIE 2359, pages 251–259, Rochester, MN, USA, October 1994.

[7] D.G. Gobbi, R.M. Comeau, and T.M. Peters. Ultrasound Probe Tracking for Real-Time Ultrasound/MRI Overlay and Visualization of Brain Shift. In *Proc of MICCAI'99*, LNCS 1679, pages 920–927, Cambridge, UK, September 1999.

[8] D.G. Gobbi, R.M. Comeau, and T.M. Peters. Ultrasound/MRI Overlay with Image Warping for Neurosurgery. In *Proc of MICCAI'00*, LNCS 1935, pages 106–114, Pittsburgh, October 2000.

[9] R.M. Comeau, A.F. Sadikot, A. Fenster, and T.M. Peters. Intraoperative ultrasound for guidance and tissue shift correction in image-guided neurosurgery. *Med. Phys.*, 27(4):787–800, April 2000.

[10] R.D. Bucholz, D.D. Yeh, B.S. Trobaugh, L.L. McDurmont, C.D. Sturm, Baumann C., Henderson J.M., Levy A., and Kessman P. The correction of stereotactic inaccuracy caused by brain shift using an intraoperative ultrasound device. In *Proc of CVRMed-MRCAS'97*, LNCS 1205, pages 459–466, 1997.

[11] G. Ionescu, S. Lavallée, and J. Demongeot. Automated Registration of Ultrasound with CT Images: Application to Computer Assisted Prostate Radiotherapy and Orthopedics. In *Proc. MICCAI'99*, LNCS 1679, pages 768–777, Cambridge (UK), October 1999.

[12] A.P. King, J.M. Blackall, G.P. Penney, P.J. Edwards, D.L.G. Hill, and D.J. Hawkes. Baysian Estimation of Intra-operative Deformation for Image-Guided Surgery Using 3-D Ultrasound. In *Proc of MICCAI'00*, LNCS 1935, pages 588–597, October 2000.

[13] M. G. Strintzis and I. Kokkinidis. Maximum Likelihood Motion Estimation in Ultrasound Image Sequences. *IEEE Signal Processing Letters*, 4(6), June 1997.

[14] R. N. Rohling, A. H. Gee, and L. Berman. Automatic registration of 3-D ultrasound images. *Medicine and Biology*, 24(6):841–854, July 1998.

[15] R. N. Rohling, A. H. Gee, and L. Berman. Three-Dimensional Spatial Compounding of Ultrasound Images. *Medical Image Analysis*, 1(3):177–193, 1997.

[16] X. Papademetris, A.J. Sinusas, D.P. Dione, and J.S. Duncan. 3D Cardiac Deformation from Ultrasound Images. In *Proc. of MICCAI'99*, LNCS 1679, pages 421–429, Cambridge, UK, September 1999.

[17] G.I. Sanchez-Ortiz, J. Declerck, M. Mulet-Parada, and J.A. Noble. Automatic 3D echocardiographic image analysis. In *Proc. of MICCAI'00*, LNCS 1935, pages 687–696, Pittsburgh, USA, October 2000.

[18] R. Bajcsy and S. Kovačič. Multiresolution Elastic Matching. *Computer Vision, Graphics and Image Processing*, 46:1–21, 1989.

[19] G. E. Christensen, S. C. Joshi, and M. I. Miller. Volumetric Transformation of Brain Anatomy. *IEEE Trans. on Medical Imaging*, 16(6):864–877, December 1997.

[20] M. Bro-Nielsen. *Medical Image Registration and Surgery Simulation.* PhD thesis, Institut for Matematisk Modellering, Danmarks Tekniske Universitet, Lyngby, Denmark, 1996.

[21] J. B. A. Maintz, E. H. W. Meijering, and M. A. Viergever. General Multimodal Elastic Registration based on Mutual Information. *Image Processing*, 1998.

[22] J.-P. Thirion. Image matching as a diffusion process: an analogy with Maxwell's demons. *Medical Image Analysis*, 2(3), 1998.

[23] B.M. Dawant, S.L. Hartmann, and Gadamsetty S. Brain atlas deformation in the presence of large space-occupying tumors. In *Proc. of MICCAI'99*, LNCS 1679, pages 589–596, Cambridge, UK, September 1999.

[24] I. Bricault, G. Ferretti, and P. Cinquin. Registration of Real and CT-Derived Virtual Bronchoscopic Imag es to Assist Transbronchial Biopsy. *Transactions in Medical Imaging*, 17(5):703–714, 1998.

[25] J. Webb, A. Guimond, N. Roberts, P. Eldridge an D. Chadwick, J. Meunier, and J.-P. Thirion. Automatic Detection of Hippocampal Atrophy on Magnetic Resonnance Images. *Magnetic Resonance Imaging*, 17(8):1149–1161, April 1999.

[26] S. Prima, J.-P. Thirion, G. Subsol, and N. Roberts. Automatic Analysis of Normal Brain Dissymmetry of Males and Females in MR Images. In *Proc. of MICCAI'98*, volume 1496 of *Lecture Notes in Computer Science*, pages 770–779, 1998.

[27] X. Pennec, P. Cachier, and N. Ayache. Understanding the "demon's algorithm": 3D non-rigid registration by gradient descent. In *Proc. of MICCAI'99*, LNCS 1679, pages 597–605, Cambridge, UK, September 1999.

[28] P. Cachier and X. Pennec. 3D non-rigid registration by gradient descent on a gaussian-windowed similarity measure using convolutions. In *Proc. of MMBIA'00*, pages 182–189, Hilton Head Island, USA, June 2000. IEEE Comput. society.

[29] A. Roche, G. Malandain, X. Pennec, and N. Ayache. The correlation ratio as a new similarity measure for multimodal image registration. In *Proc. of MICCAI'98*, LNCS 1496, pages 1115–1124, Cambridge, USA, October 1998.

[30] A. Roche, X. Pennec, M. Rudolph, D. P. Auer, G. Malandain, S. Ourselin, L. M. Auer, and N. Ayache. Generalized Correlation Ratio for Rigid Registration of 3D Ultrasound with MR Images. In *Proc. of MICCAI'00*, LNCS 1935, pages 567–577, October 2000. Submitted to IEEE TMI.

[31] A. Roche, X. Pennec, G. Malandain, and N. Ayache. Rigid registration of 3D ultrasound with MR images: a new approach combining intensity and gradient information. *IEEE Transactions on Medical Imaging*, 20(10):1038–1049, October 2001.

[32] X. Pennec, P. Cachier, and P. Ayache. Tracking brain deformations in time sequences of 3D US images. In M.I. Insana and R.M. Leahy, editors, *Proc. of IPMI'01*, volume 2082 of *LNCS*, pages 169–175, Davis, CA, USA, June 2001. Springer Verlag.

[33] X. Pennec, P. Cachier, and N. Ayache. Tracking brain deformations in time-sequences of 3D US images. *Pattern Recognition Letters - Special Issue on Ultrasonic Image Processing and Analysis*, 24(4-5):801–813, February 2003.

[34] A. Roche, G. Malandain, and N. Ayache. Unifying Maximum Likelihood Approaches in Medical Image Registration. *International Journal of Imaging Systems and Technology: Special Issue on 3D Imaging*, 11:71–80, 2000.

[35] W.H. Press, B.P. Flannery, S.A Teukolsky, and W.T. Vetterling. *Numerical Recipices in C*. Cambridge Univ. Press, 1991.

[36] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens. Multimodality image registration by maximization of mutual information. *IEEE Trans. Med. Imaging*, 16:187–198, 1997.

[37] F. Maes, D. Vandermeulen, and P. Suetens. Comparative evaluation of multiresolution optimization strategies for multimodality image registration by maximization of mutual information. *Medical Image Analysis*, 3(4):373–386, 1999.

[38] J.P.W. Pluim, J.B.A. Maintz, and M.A. Viergever. Mutual information matching and interpolation artefacts. In *Proc. SPIE 3661*, pages 56–65, 1999.

[39] Peter J. Rousseeuw and Annick M. Leroy. *Robust Regression and Outlier Detection.* Wiley, 1987.

[40] J. Meunier and M. Bertrand. Ultrasonic Texture Motion Analysis: Theory and Simulation. *IEEE Transactions on Medical Imaging*, 14(2), June 1995.

[41] M. Ferrant, S. K. Warfield, C. R. G. Guttmann, R. V. Mulkern, F. A. Jolesz, and R. Kikinis. 3D Image Matching using a Finite Element Based Elastic Deformation Model. In *Proc. of MICCAI'99*, LNCS 1679, pages 202 – 209, Cambridge, UK, September 1999.

[42] S. Ourselin, A. Roche, S. Prima, and N. Ayache. Block Matching: A General Framework to Improve Robustness of Rigid Registration of Medical Images. In *Proc. of MICCAI'2000*, volume LNCS 1935, pages 557–566. Springer, 2000.

[43] J.-M. Morel and S. Solimini. *Variational Methods in Image Segmentation.* Progress in Nonlinear Differential Equations and Their Applications. Birkhuser, 1995.

[44] X. Pennec and J.-P. Thirion. A framework for uncertainty and validation of 3D registration methods based on points and frames. *Int. Journal of Computer Vision*, 25(3):203–229, 1997.

[45] X. Pennec, C.R.G. Guttmann, and J.-P. Thirion. Feature-based registration of medical images: Estimation and validation of the pose accuracy. In *Proc. of First Int. Conf. on Medical Image Computing and Computer-Assisted Intervention (MICCAI'98)*, volume 1496 of *LNCS*, pages 1107–1114, Cambridge, USA, October 1998. Springer Verlag.

[46] M. Holden, D.L.G. Hill, E.R.E. Denton, J.M. Jarosz, T.C.S. Cox, T. Rohlfing, J. Goodey, and D.J. Hawkes. Voxel Similarity Measures for 3D Serial MR Brain Image Registration. *IEEE Trams. Med. Imaging*, 19(2), 2000.

[47] G.P. Penney, J.M. Blackall, D. Hayashi, T. Sabharwal, A. Adam, and D.J. Hawkes. Overview of an Ultrasound to CT or MR Registration System for use in Thermal Ablation of Liver Metastases. In *Proc of Medical Image Understanding and Analysis (MIUA'01)*, July 2001.

[48] X. Pennec, N. Ayache, and J.-P. Thirion. Landmark-based registration using features identified through differential geometry. In I. Bankman, editor, *Handbook of Medical Imaging*, chapter 31, pages 499–513. Academic Press, September 2000.

[49] J.P.W. Pluim, J.B.A. Maintz, and M.A. Viergever. Image registration by maximization of combined mutual information and gradient information. *IEEE Transactions on Medical Imaging*, 19:809–814, 2000.

[50] W.L. Smith and A. Fenster. Optimum scan spacing for three dimensionnal ultrasound by speckle statistics. *Ultrasound Med. Biol.*, 26(4):551–562, 2000.

[51] Radu Stefanescu, Xavier Pennec, and Nicholas Ayache. Parallel non-rigid registration on a cluster of workstations. In Sofie Norager, editor, *Proc. of HealthGrid'03*, Lyon, January 2003. European Commission, DG Information Society.

[52] Radu Stefanescu, Xavier Pennec, and Nicholas Ayache. Grid enabled non-rigid registration with a dense transformation and a priori information. In Randy E. Ellis and Terry M. Peters, editors, *Proc. of MICCAI'03, Part II*, volume 2879 of *LNCS*, pages 804–811, Montreal, November 2003. Springer Verlag.