To be published as:

A. Roche, X. Pennec, G. Malandain, and N. Ayache.
Rigid Registration of 3D Ultrasound with MR Images: a New Approach Combining Intensity and Gradient.
IEEE Transactions on Medical Imaging,
20(10):1038–1049, October 2001.

Rigid Registration of 3D Ultrasound with MR Images: a New Approach Combining Intensity and Gradient Information

Alexis Roche Xavier Pennec Grégoire Malandain
Nicholas Ayache

INRIA - Projet Epidaure
2004 Route des Lucioles BP 93
06902 Sophia Antipolis, France

Abstract

We present a new image-based technique to rigidly register intra-operative 3D ultrasound (US) with pre-operative Magnetic Resonance (MR) images. Automatic registration is achieved by maximization of a similarity measure which generalizes the correlation ratio (CR), and whose novelty is to incorporate multivariate information from the MR data (intensity and gradient). In addition, the similarity measure is built upon a robust intensity-based distance measure, which makes it possible to handle a variety of US artifacts. A cross-validation study has been carried out using a number of phantom and clinical data. This indicates that the method is quite robust and that the worst registration errors are of the order of the MR image resolution.

Keywords: image registration, ultrasound, magnetic resonance, correlation ratio, robust estimation.

1 Introduction

Over the past few years, the development of real-time 3D ultrasound (US) imaging has revealed a number of potential applications in image-guided neurosurgery 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 developed to the same extent as other medical imaging modalities, probably due to the low signal-to-noise ratio of US images.

The registration of intra-operative US with pre-operative Magnetic Resonance (MR) images has the potential to enable the surgeon to accurately localize the trajectories of instruments in the operative field, resulting in minimally invasive procedures. To date, few papers have been published on this particular registration problem [10]. Most of the existing approaches are stereotactic-based

as they make use of a tracking device to assess the position and orientation of the US probe in real-time [24, 13, 2]. Such systems are calibrated before the intervention in order to relate the US and MR coordinates. Pure stereotactic-based approaches are intrinsically limited to rigid registration since they exploit only the probe position and orientation.

Other approaches are image-based in the sense that registration is performed using the image data itself. Such techniques are potentially able to compensate for a non-rigid transformation which may result from brain deformations or geometric distorsions in the US. They can also be used in combination with stereotactic-based registration [1, 5]. Existing image-based techniques generally consist of matching homologous features extracted from both the US and MR data. Features are user-identified in [4, 1, 5], while they are semi-automatically extracted in [6]. Fully automated feature extraction is reported in [7] for prostate images, and [17] for liver and forearm images (using color Doppler US in the last reference).

Slightlty different is the approach of King *et al* [8] in which no explicit feature detection is performed on the US. Instead, a Bayesian estimation technique is used to deform a surface extracted from the MR image according to the US intensity and gradient information. In the paper we refer to, however, experiments with phantom data only are presented.

The present registration technique expands on the correlation ratio (CR) method [20]. It is a pure intensity-based approach as it does not rely on *any* feature extraction. In a previous work [19], we reported preliminary results of US/MR registration by maximizing two different similarity measures, namely CR and mutual information (MI) [25, 9]. While results obtained using CR were more appealing than when using MI, the method still lacked precision and robustness with respect to the initialization of the transformation parameters.

In this paper, we generalize the CR method following two distinct themes. First, we observe that ultrasound imagery is essentially concerned with the interfaces between anatomical structures. Our idea is then to correlate the US intensity with both the MR intensity and the MR gradient magnitude, which leads to a bivariate extension of the CR. Secondly, we incorporate a robust intensity-based distance measure in order to prevent the bivariate CR from being biased by various ultrasound artifacts.

The bivariate CR method is described in detail in the following section. Section 3 describes the phantom and clinical data that were used in our experiments, while sections 4 and 5 propose original evaluations of the method's accuracy and robustness, respectively.

2 Method

2.1 Correlation ratio

Given two images I and J, the basic principle of the CR method is to search for a spatial transformation T and an intensity mapping f such that, by displacing J and remapping its intensities, the resulting image $f(J \circ T)$ is as similar as possible to I. In a first approach, this could be achieved

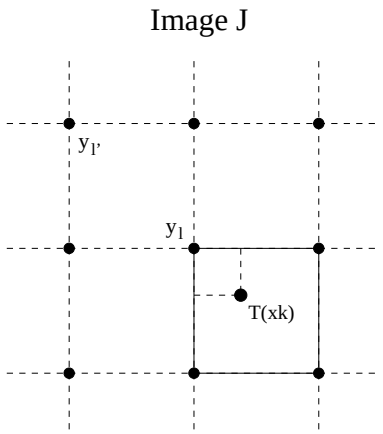
by minimizing the following cost function:

$$\min_{T,f} \quad \sum_k \Big[I(\mathbf{x}_k) - f(J(T(\mathbf{x}_k))) \Big]^2, \tag{1}$$

which integrates over the voxel positions $\mathbf{x}_k$ in image I. This formulation is asymmetric in the sense that the cost function changes when the roles of I and J are interchanged. Since the positions and intensities of J actually serve to predict those of I, we call J the "template image". In the context of US/MR registration, we always choose the MR volume as the template as it yields higher quality information with better signal-to-noise ratio. In the following, we will use the simplified notations $i_k \equiv I(\mathbf{x}_k)$, and $j_k^{\downarrow} \equiv J(T(\mathbf{x}_k))$, where the arrow expresses the dependence on T.

In practice, the criterion defined in Eq. (1) cannot be computed exactly due to the sampled nature of the template image. One obvious problem is that the transformed position of a voxel will generally not coincide with a grid point of J, such that the corresponding intensity $j_k^{\downarrow}$ is unknown. A classical approach is to linearly interpolate $j_k^{\downarrow}$ using the eight neighbours of $T(\mathbf{x}_k)$ in the grid of J. However, instead of interpolating the image intensity, we may directly interpolate the incremental contribution of $\mathbf{x}_k$, i.e., $[i_k - f(j_k^{\downarrow})]^2$. The difference between these two approaches is illustrated in Fig. 1. The last method turns out to be equivalent to the so-called partial volume (PV) interpolation, originally proposed by Maes *et al* [9] in the context of joint histogram computation. We have found PV interpolation generally outperforms intensity interpolation in terms of smoothness of the resulting registration criterion.

Another difficulty in computing Eq. (1) is that some points $\mathbf{x}_k$ may transform outside the template domain and lack eight grid neighbours. It is common *not* to take into account these points in the computation of the registration criterion. However, doing so without careful attention, the criterion would become zero when every point $\mathbf{x}_k$ transforms outside J. Hence, in order to avoid an absolute minimum when the image overlap is small, we impose the additional constraint that the variance of I be large in the overlapping region. Justification for this particular normalization



Linear interpolation weights:

$$w_{k\ell}^{\downarrow} = \beta(T(\mathbf{x}_k) - \mathbf{y}_\ell), \quad \text{with :}$$
$$\begin{cases} \beta(\mathbf{z}) = (1-|z^1|)(1-|z^2|), & \text{if } |z^1| \le 1,\ |z^2| \le 1 \\ \beta(\mathbf{z}) = 0, & \text{otherwise} \end{cases}$$

Intensity interpolation:

$$\Big[i_k - f(j_k^{\downarrow}) \Big]^2 \approx \Big[i_k - f\left(\textstyle\sum_\ell w_{k\ell}^{\downarrow} j_\ell \right) \Big]^2$$

PV interpolation:

$$\Big[i_k - f(j_k^{\downarrow}) \Big]^2 \approx \textstyle\sum_\ell w_{k\ell}^{\downarrow} \, [i_k - f(j_\ell)]^2$$

Figure 1: *Illustration of linear interpolation in the 2D case, and two related strategies of interpolating the registration criterion.*

strategy may be found in [20], while related normalization issues are discussed in [23, 25].

These practical considerations lead us to the following modification to Eq. (1):

$$C(T, f) = \frac{\sum_{k,\ell} w_{k\ell}^{\downarrow} [i_k - f(j_\ell)]^2}{n^{\downarrow} \mathrm{Var}(I^{\downarrow})}, \tag{2}$$

where the terms depending on T are marked with an arrow: $w_{k\ell}^{\downarrow}$ are the linear interpolation weights, $n^{\downarrow}$ is the number of points $\mathbf{x}_k$ such that $T(\mathbf{x}_k)$ has eight neighbours in the template grid, and $\mathrm{Var}(I^{\downarrow})$ is the intensity variance computed over these points.

If no constraint is imposed on the intensity mapping f, an important result is that the optimal f at fixed T has an explicit form that is very fast to compute [20]. The minimization of Eq. (2) may then be performed by travelling 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 our implementation, the maximization of η^2 is performed using Powell's method [18].

2.2 Bivariate correlation ratio

Ultrasound images are commonly said to be "gradient images" since they enhance the interfaces between anatomical structures. The physical reason is that the amplitudes of the US echos are proportional to the *difference* between acoustical impedances caused by successive tissue layers. Ideally, the US signal should be high at the interfaces, and zero within homogeneous tissues. In reality, US reflections also occur within tissues due to "small" inhomogeneities which are almost invisible in MR. At this scale, there are significant interferences between the ingoing and outgoing pulsed waves, resulting in speckle patterns. As a consequence, homogenous tissue regions generally appear in the US with non-zero mean intensity and strong texture.

As stated above, the CR method tries to predict the intensities of the US by remapping those of the MR. Hence, uniform regions of the original MR remain uniform in the remapped MR and, thus, this procedure is not able to account for intensity variations at tissue boundaries. To enable a better prediction, we propose to use the MR gradient magnitude as an additional explanatory variable. In other terms, our template image J is now a vectorial image, $J = (M, \|\nabla M\|)$, M standing for the MR image, and we search for a function $f : \mathbb{R}^2 \to \mathbb{R}$ that maps double values to single values.

Such a mapping does not appear to be entirely adequate from a physical standpoint. On the one hand, it is clearly not able to account for texture in the US, for which speckle is responsible. This is a problem we will return to in section 2.4. In addition, using the magnitude of the MR gradient may not be ideal. In fact, the US signal which is produced at an interface depends also on the tissue orientation with respect to the scan line. Thus, perhaps a more appropriate choice than $\|\nabla M\|$ would be the dot product, $(\nabla M).\mathbf{u}$, where $\mathbf{u}$ is the scan direction. The main difficulty in using this last expression is that $\mathbf{u}$ is unknown before registration since it depends on the position

of the US probe in the MR coordinate system. A possible solution to overcome this problem is to iteratively estimate $\mathbf{u}$ using the current estimated transformation. However, in a straightforward implementation, this implies recomputing the "$(\nabla M).\mathbf{u}$" image at each transformation trial, entailing a massive increase of the computation time. An efficient implementation of this idea is an issue we have not yet tackled.

Nevertheless, we believe that ignoring the gradient orientation is acceptable, at least as a first-order approximation. As a result of diffraction of the ultrasound beam on interfaces, the received echo is actually less anisotropic than would be the case with perfectly specular reflection. Moreover, even if impedance boundaries cannot be detected isotropically, in practice all we require is that there are sufficiently many edge points whose gradient directions are approximately orthogonal to the US beam direction.

An illustration of our bivariate model is presented in Fig. 2. We rigidly registered one US/MR brain volume pair using the present method. Fig. 2 (c) displays the predicted US image corresponding to the final stage of registration. The intensity fit was computed using a polynomial function (see section 2.3 below). For comparison, we represent in Fig. 2 (d) the prediction obtained when the second explanatory variable is chosen as the MR gradient projected onto the US scan line (the US scan line was computed from the registration result).

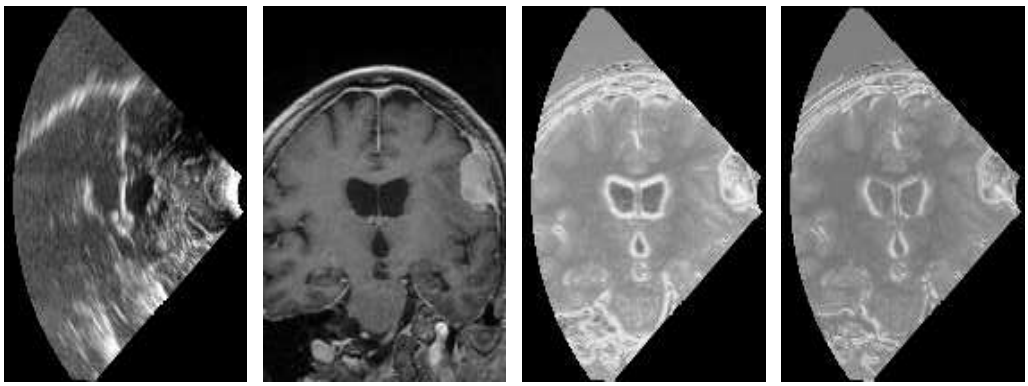
Although the images in Fig. 2 (c) and Fig. 2 (d) are reasonably similar, the latter is seen to better account for the drop in US signal at boundaries which are not normal to the beam direction (this is especially clear around the ventricles). Needless to say, however, both of these predicted images are only *roughly* comparable to the US image. In particular, they do not simulate the textural appearance of speckle. Also, since our functional model does not take into account US attenuation, impossible predicted values are found in some regions, especially outside the skull. These limitations suggest introducing a robust generalization of the bivariate CR, as will be done in section 2.4.

2.3 Parametric intensity fit

If we impose no special constraint on the mapping f to be estimated, then f is described by as many parameters as there are distinct intensity values in the template image [19]. That approach makes sense as long as the number of intensity classes in J is small with respect to the number of voxels used to make an estimate. In our case, J is a double-valued image (with, in general, floating precision encoding of the MR gradient component), and the number of parameters to be estimated becomes virtually infinite.

We will therefore restrict our search to a polynomial function f. Let m_ℓ and g_ℓ denote the intensity of the voxel with coordinates $\mathbf{y}_\ell$, respectively in the MR image, M, and in the gradient norm image, $\|\nabla M\|$. We search for a mapping of the form:

$$f(m_\ell, g_\ell) = \sum_{p+q \leq d} c_{pq}\, m_\ell^p g_\ell^q, \tag{3}$$



(a) (b) (c) (d)

Figure 2: *Visual assessment of the bivariate functional model. (a) US plan. (b) Corresponding MR slice after 3D registration. (c) Predicted US in terms of the MR intensity and MR gradient magnitude. (d) Predicted US in terms of the MR intensity and MR gradient projected onto the US scan line.*

where d is the specified polynomial degree. The number of parameters describing f then reduces to $(d+1)(d+2)/2$. In all the experiments presented below, the degree was set to $d = 3$, implying that 10 coefficients were estimated. It is shown in [21] that minimizing Eq. (2) with respect to the polynomial coefficients yields a weighted least square (WLS) linear regression problem. As is standard, this is solved by the singular value decomposition (SVD) method.

This polynomial fitting procedure, however, has significant additional computational cost with respect to the unconstrained fitting. Recall that, in the basic version of the CR method, f is updated at each transformation stage. Such a strategy is no longer affordable when estimating a polynomial function. Instead, the minimization of Eq. (2) may be performed alternately along T and f, resulting in the following algorithm: **1.** given a current transformation estimate T, find the best polynomial function f and remap the MR image accordingly; **2.** given the remapped MR, $f(M, \|\nabla M\|)$, minimize $C(T, f)$ with respect to T using Powell's method; **3.** return to **1** if T or f has changed.

The 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 minimum 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. As discussed in section 2.2, we do not expect this assumption

to be strictly valid. Speckle patterns, attenuation, and other US artefacts may cause large variations of the US intensity from its predicted value, and this remains so even when the images are perfectly registered. From a registration standpoint, such bad intensity matches result in false negatives.

The sensitivity of the registration criterion to false negatives may be reduced by replacing in Eq. (2) the quadratic error function, $(1/n^{\downarrow}) \sum_{k,\ell} w^{\downarrow}_{k\ell} \left[i_k - f(j_\ell) \right]^2$, with a robust scale estimate. A similar idea was developed in [11]. Here we propose to build such an estimate from a one-step S-estimator [22]:

$$\hat{S}^2(T, f) = \frac{S_0^2}{K n^{\downarrow}} \sum_{k,\ell} w^{\downarrow}_{k\ell} \, \rho \left(\frac{i_k - f(j_\ell)}{S_0} \right), \tag{4}$$

where ρ is the objective function corresponding to a given M-estimator, K is a normalization constant to ensure consistency with the normal distribution, and S_0 is some initial guess of the scale. The new registration criterion is then:

$$C(T, f) = \frac{\hat{S}^2(T, f)}{\mathrm{Var}(I^{\downarrow})}. \tag{5}$$

This criterion requires few modifications to our alternate minimization strategy. As a function of T, it may still be minimized by means of Powell's method. As a function of f, the solution is found by a simple iterative WLS procedure [21], generally requiring no more than 5-6 iterations to converge. In our implementation, we have opted for the Geman-McClure ρ-function, $\rho(x) = x^2/(1 + \frac{x^2}{c^2})$, for its computational efficiency and good robustness properties, to which we always set an empirical cut-off value of $c = 3.5$ (the normalization constant is then $K \approx 0.8$).

Initially, the intensity mapping f is estimated in a non-robust fashion. The starting value S_0 is then computed as the weighted median absolute deviation of the corresponding residuals, $|i_k - f(j_\ell)|$ (see [21] for details). Due to the initial misalignment, S_0 tends to be overestimated and, thus, some bad intensity matches may still bias the registration criterion. For this reason, we reset S_0 at each new iteration, i.e., after completing one minimization along T and one minimization along f.

3 Data

The experiments related in this article were performed within the framework of the EC-funded ROBOSCOPE project. The goal is to assist neuro-surgical operations using real-time 3D ultrasound images and a robotic manipulator arm. The operation is planned on a pre-operative MRI and 3D US images are acquired during surgery to track in real-time the deformation of anatomical structures. In this context, the rigid registration of the pre-operative MR with the first US image (dura mater still closed) is a fundamental task to relate the position of the surgical instruments with the actual anatomical structure. This task being determinant for the global accuracy of the system, different datasets were acquired to simulate the final image quality and to perform accuracy evaluations.

It should be emphasized that all the US images provided in this project were stored in Cartesian format, which means that the actual (log-compressed) ultrasound signal is resampled on a regular cubic lattice. As a consequence, the images are subject to severe interpolation artifacts, especially

in areas with low spatial resolution (far from the probe). In the following, we will refer to the US images as cubic images, but one has to bear in mind that this is somewhat artificial. Notably, the voxel size in Cartesian US images should not be confused with the real spatial resolution, which is in fact spatially dependent.

3.1 Phantom dataset

Within ROBOSCOPE, ISM developed an MR and US compatible phantom made of two balloons, one ellipsoid and one ellipsoid with a "nose" (complex ellipsoid), that can be inflated with known volumes in order to simulate deformations (see Fig. 3). Each acquisition consists of one 3D MR image and one 3D US image.

Both balloons were initially filled with 40 ml of fluid. During the first series of 5 acquisitions, the ellipsoid was filled in steps of 10 ml, while the complex ellipsoid was kept constant. During the second series, the ellipsoid was deflated and the complex ellipsoid filled in steps of 10 ml. Each MR image has $256 \times 256 \times 124$ voxels of size $0.9 \times 0.9 \times 1$ mm^3. Each (Cartesian) US image has $184 \times 184 \times 184$ voxels of size 0.41 mm^3.

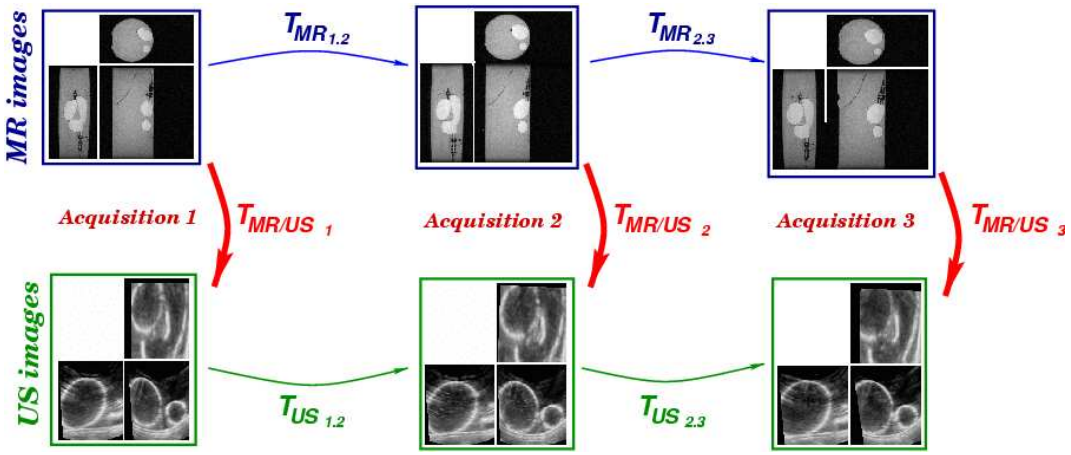


Figure 3: *Original MR and US images (before registration) of the Phantom and the rigid transformations that relate them.*

This dataset would be ideal for the validation of MR and US registration if all MR (resp. all US) images were exactly in the same coordinate system. Since the US probe cannot enter the MR machine, this is impossible: the phantom has to be moved and the US probe removed between MR acquisitions. Thus we have to register the MR (resp. the US) images together. Unfortunately, there were no rigid markers inserted in the phantom and the US/US rigid registrations are much less accurate than the MR/MR rigid registrations (see section 4).

3.2 Baby dataset

This clinical dataset was acquired to simulate the degradation of the US images quality with respect to the number of array transducer elements used. Here, we have one MR T1 image of a baby's head and 5 transfontanel US images with different percentages of elements used (40%, 60%, 70%, 80%, 90%, 100%). The MR image has $256 \times 256 \times 124$ voxels of size $0.9\,\mathrm{mm}^3$. The Cartesian US images have $184 \times 184 \times 184$ voxels of size $0.29\,\mathrm{mm}^3$. Note the US images were truncated from the original acquisition, as can be observed in Fig. 5.

3.3 Patient dataset

This dataset is an actual surgical case: two sagittal 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 were acquired to precisely locate the tumor to resect. The MR images have $256 \times 256 \times 124$ voxels of size $0.9 \times 0.9 \times 1.1\,\mathrm{mm}^3$. The US images have various dimensions with voxel sizes ranging from $0.17\,\mathrm{mm}^3$ to $0.95\,\mathrm{mm}^3$.

4 Accuracy study

We computed all the MR/US registrations using the bivariate CR algorithm summarized by Eq. 5. Since the location of the US probe was linked to the pathology and its orientation was arbitrary (the rotation was occasionaly more than 90 degrees), it was necessary to provide a rough initial estimate of the transformation. Here, this was done using an interactive tool that allows to draw lines in the images and match them. This procedure was carried out only to get a rough initialization, but we always made sure that a slight misalignment was still visible.

In all the experiments, the gradient norm of the MR image was computed by convolution with a Gaussian derivarive [3] with $\sigma = 1$ voxel. The minimization of the registration criterion using Powell's method took of the order of 5-10 minutes on a standard PC (Pentium II running at 450MHz with 500MB of RAM). After manual initialization, the algorithm found residual displacements that were in the range of 10 mm and 10 degrees.

Fig. 4, 5, 6 show registration examples corresponding respectively to the phantom, the baby, and the patient datasets (see section 3). More results, including movies, are accessible on the internet from the INRIA website [1].

4.1 Principle of the accuracy evaluation

Ideally, the method's accuracy should be assessed by comparing the result of a registration with a gold standard. Such a gold standard may be obtained by relating both the MR and US coordinate systems to the same physical coordinate system. For MR, this may be done efficiently by attaching fiducial markers to the patient (or, more easily, the phantom), and then matching the positions

[1] http://www-sop.inria.fr/epidaure/personnel/pennec/Demos/Roboscope/

of the markers in the image with their physical coordinates. However, this solution is not easily applicable in the case of US as the detection of markers is prone to inaccuracy. A better solution is probably to track the US probe using an external device [24, 13, 2]. Unfortunately, no such device was available in the ROBOSCOPE consortium at the time when the datasets were acquired and, thus, no gold standard is available. Moreover, the clinical datasets (the patient's and the baby's) were not originally intended for a registration evaluation purpose.

To get around this problem, our main idea is to use several MR and/or US images to compute registration loops and assess residual errors. What we call a registration loop is a series of transformation compositions leading to a global transformation that is, ideally, the identity matrix.

A typical loop is a sequence of the form $US_i \rightarrow MR_i \rightarrow MR_j \rightarrow US_j \rightarrow US_i$ in the case of the Phantom data (see Fig. 3). If we were given perfectly registered images within each modality, this loop would only be disturbed from the identity by errors due to the two MR/US registrations. Since variances are additive, the variance of the observed error should roughly be: $\sigma^2_{loop} = 2\sigma^2_{MR/US}$.

In our case, the situation is slightly more complex as we actually need to estimate the intra-modality transformations (see section 4.2). This also implies that we must take into account errors that occur at each stage of the loop. Since we are combining one MR/MR, one US/US and two MR/US registrations, the variance of the loop error will be: $\sigma^2_{loop} \simeq 2\sigma^2_{MR/US} + \sigma^2_{MR/MR} + \sigma^2_{US/US}$. We thus define the *expected* accuracy as:

$$\sigma_{MR/US} \simeq \sqrt{\frac{1}{2}\left(\sigma^2_{loop} - \sigma^2_{MR/MR} - \sigma^2_{US/US}\right)} \tag{6}$$

What we are able to measure is σ_{loop}, from which we may infer the expected MR/US accuracy provided that $\sigma_{MR/MR}$ and $\sigma_{US/US}$ are faithfully estimated. For that purpose, we designed the multiple registration algorithm described in section 4.2. Notice that Eq. (6) furnishes at least an upper bound of the MR/US error, given by $\sigma_{MR/US} \simeq \sigma_{loop}/\sqrt{2}$, which corresponds to the case where the intra-modalityregistrations are perfect. In the following, this worst-case value will be referred to as the *conservative* MR/US error.

4.2 Multiple intra-modality registration

To relate n images together, we need to estimate $n-1$ rigid transformations $\bar{T}_{i,i+1}$. The principle of multiple registration is that an optimal accuracy will be reached when registering all image pairs, and thus obtaining $n(n-1)$ transformations $\hat{T}_{i,j}$. We may then compute the transformations $\bar{T}_{i,i+1}$ that best explain the measurements in the least-square sense, i.e., that minimize the following criterion:

$$C(\bar{T}_{1,2}, \ldots \bar{T}_{n-1,n}) = \sum_{i,j} \mathrm{dist}^2(\bar{T}_{i,j}, \hat{T}_{i,j}) \quad \text{where} \quad \bar{T}_{i,j} = \begin{cases} \bar{T}_{j-1,j} \circ \ldots \circ \bar{T}_{i,i+1} & \text{if } j > i \\ \bar{T}^{(-1)}_{j,i} & \text{if } j < i \end{cases} \tag{7}$$

In this work, the 'dist' function was chosen as a robust variant of the left invariant distance on rigid transformations [16]. Let χ^2 be a certain threshold and let σ_r and σ_t be rough initial

estimates of the rotation and translation standard deviations. If a transformation T is represented by a rotation vector r and a translation vector t, then the distance between two transformations is given by:

$$\mathrm{dist}^2(T_1, T_2) = \mathrm{dist}^2(T_2^{(-1)} \circ T_1, \mathrm{Id}) = \min\left(\|r_2^{(-1)} \circ r_1\|^2/\sigma_r^2 + \|t_1 - t_2\|^2/\sigma_t^2 \ , \ \chi^2 \right),$$

where $r_2^{(-1)}$ is the rotation vector associated with the inverse of the rotation defined by r_2.

The minimization of Eq. (7) is carried out using a Newton gradient descent whose implementation details may be found in [21]. We end up with estimates of the rigid transformations $\bar{T}_{i,i+1}$ as well as estimates of their variance, from which we can compute the intra-modality registration errors, $\sigma^2_{MR/MR}$ and $\sigma^2_{US/US}$.

4.3 Phantom dataset

In this experiment, we used 8 acquisitions with different balloon volumes, each acquisition consisting of one 3D MR and one 3D US image. However, we cannot directly compare the MR/US registrations as the phantom is moved between the acquisitions. Thus, the first step is to rigidly register all the MR images together and similarly the US images.

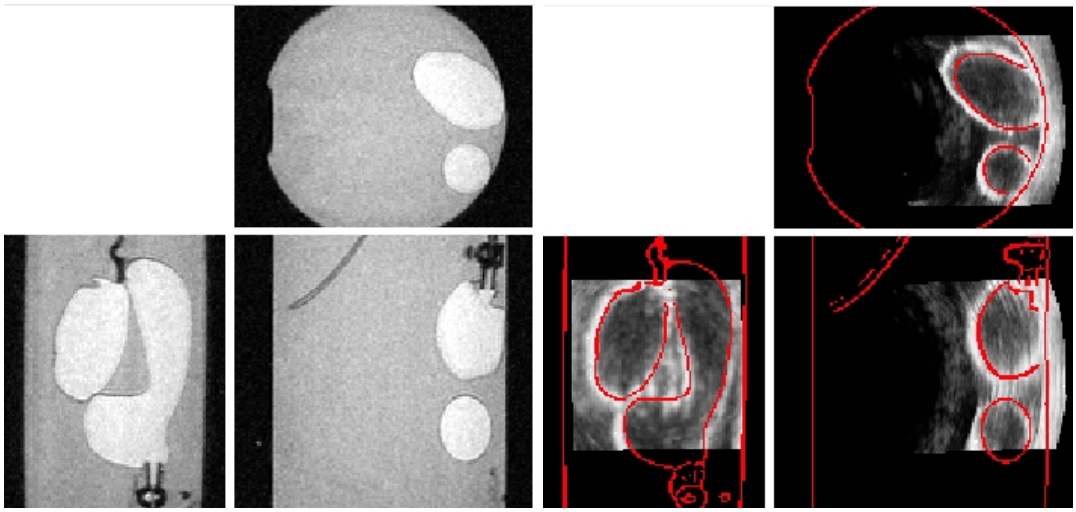


Figure 4: *Example registration of the MR and US images of the phantom. Left: original MR image. Right: registered US image with the MR contours superimposed.*

The main problem for the multiple intra-modality registration of the phantom images is that the acquisitions are intrinsically non-rigid except for the outer part of the container and to a certain extent one of the two valves. The MR/MR registrations were carried out using the "crest lines" method, a feature-based technique known to handle a large amount of outliers [15]. In the case of US images, as it is very difficult to extract meaningful features, we used a block-matching algorithm based on local maximizations of the correlation coefficient of intensities [12].

The loops we used for the accuracy estimation are the $n(n-1)$ following ones: $US_i \rightarrow MR_i \rightarrow MR_j \rightarrow US_j \rightarrow US_i$. Of course, only $n-1$ loops are independent but since the ideal value is known (the identity) there is no need to correct the estimation for the number of free parameters.

4.4 Baby dataset

This dataset consists of one MR T1 image of a baby's head and 5 transfontanel US images. In this case, we have no or very few deformations between acquisitions. Therefore, we can rigidly register all the US images onto our single MR and test the 30 following loops $US_i \rightarrow MR \rightarrow US_j \rightarrow US_i$ (only 5 of them being independent). For that, we still need to register the US images together. Our results suggest that the algorithm is much more efficient and accurate than for the phantom dataset. This is probably due to the fact that the rigidity assumption is better verified here.

4.5 Patient dataset

This dataset contains two pre-operative MR T1 images (with and without contrast agent enhancement) and a set of 3D US images acquired during surgery. However, most of these US correspond to "zooms" on the tumor area and, thus, have a very small field of view. We only use the three US images that are large enough to contain the ventricles as we could not register the others.

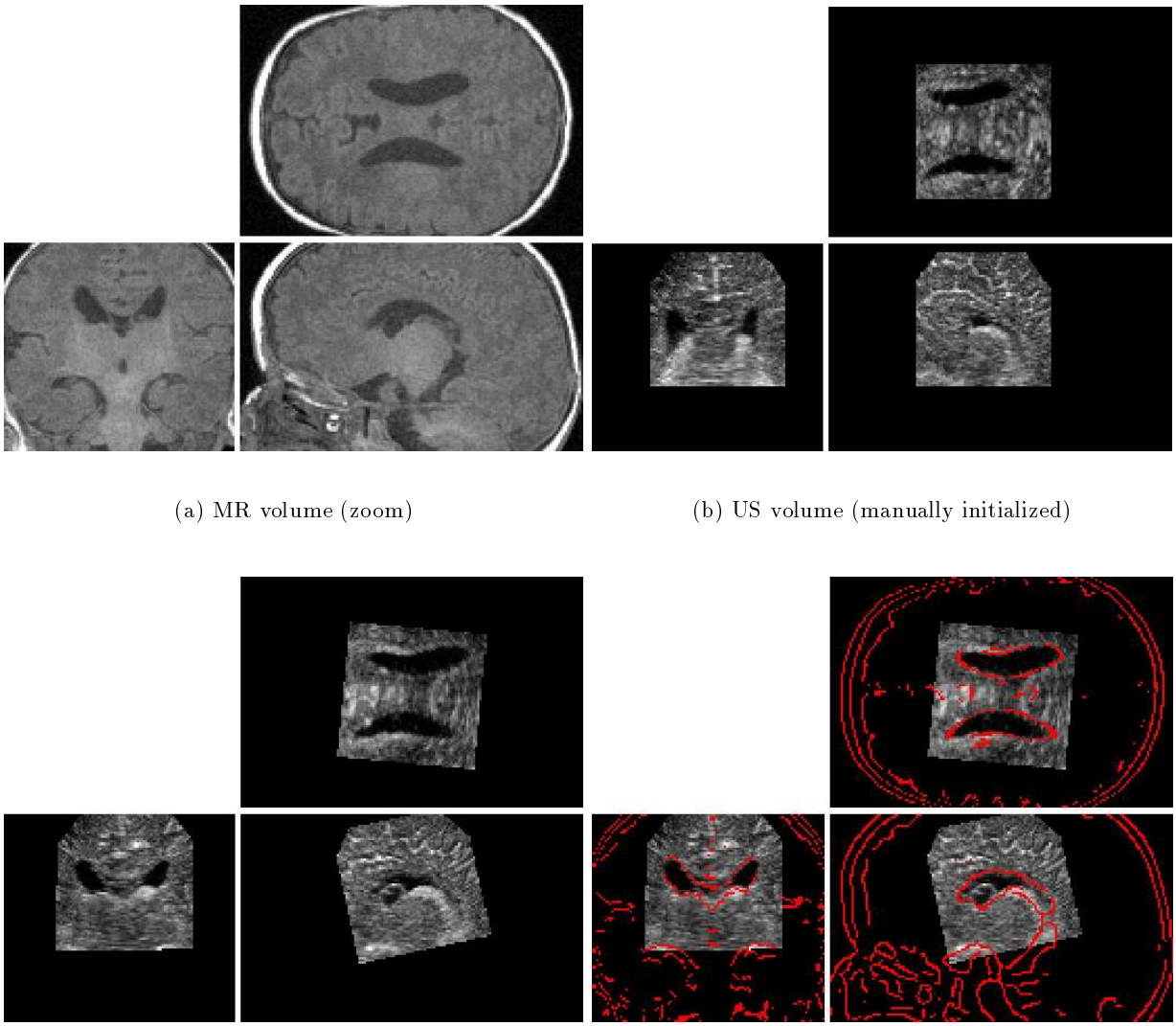
Here again, the two MR images were registered using the crest lines method [15] with very high accuracy (probably overestimated as we only have two images). We tested the loops $US_i \rightarrow MR_0 \rightarrow MR_1 \rightarrow US_i$. Since the US acquisition cone is completely within the Cartesian image (see Fig. 6), the region of interest is much smaller than the image size: we took our test points at the corners of a $80 \times 80 \times 80$ mm^3 cube centered in the Cartesian US image.

4.6 Discussion

Table 1 shows the standard deviations computed from the different registration loops described above. These values correspond respectively to the residual rotation, the residual translation and the average displacement of eight test points that were taken at the corners of the Cartesian US image (except for the patient images, see section 4.5). Since we put the origin of the images at the center, the σ_{trans} value corresponds to the mean error at the center of the US image while σ_{test} corresponds roughly to the maximum registration error within the region of interest defined by the US acquisition cone.

The results of the Phantom dataset show that the MR/US registration accuracy is of the order of the MR resolution. One could probably expect more conservative accuracy by acquiring larger US images including some rigid landmarks for multiple US/US registration. One finds the same type of results for the other datasets: slightly lower than the MR voxel size for the baby dataset and a bit larger for the patient dataset.

However, when we look more carefully at the patient results, we find that the loop involving the smallest US image (real size $150 \times 85 \times 100$ mm, voxel size 0.63^3 mm^3) is responsible for a



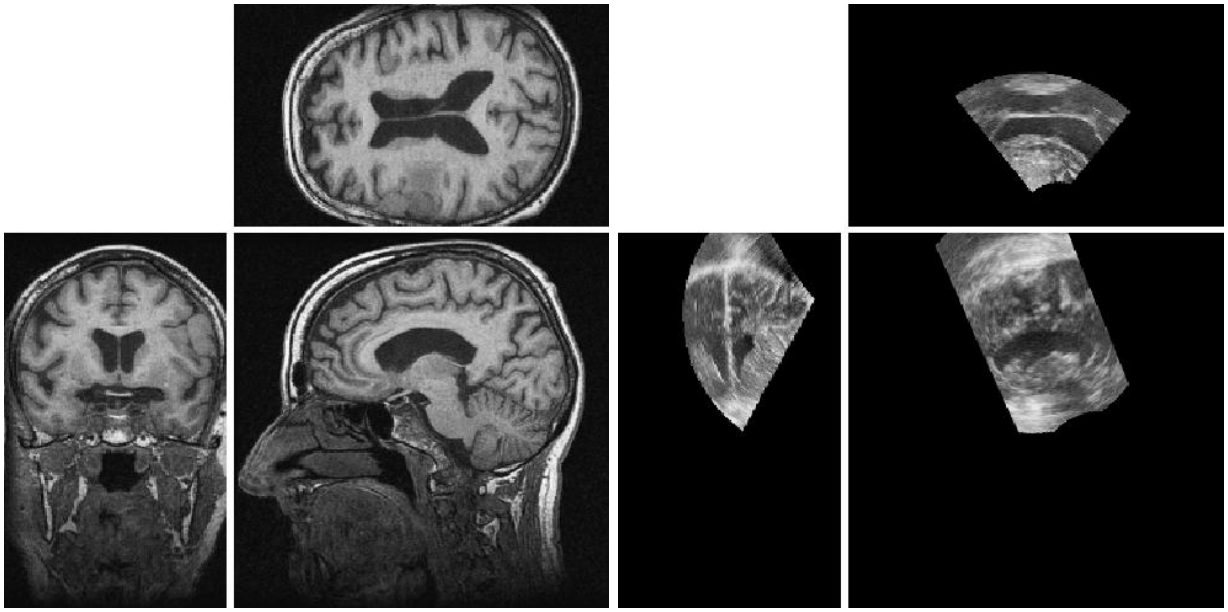
(a) MR volume (zoom)

(b) US volume (manually initialized)

(c) Registered US volume

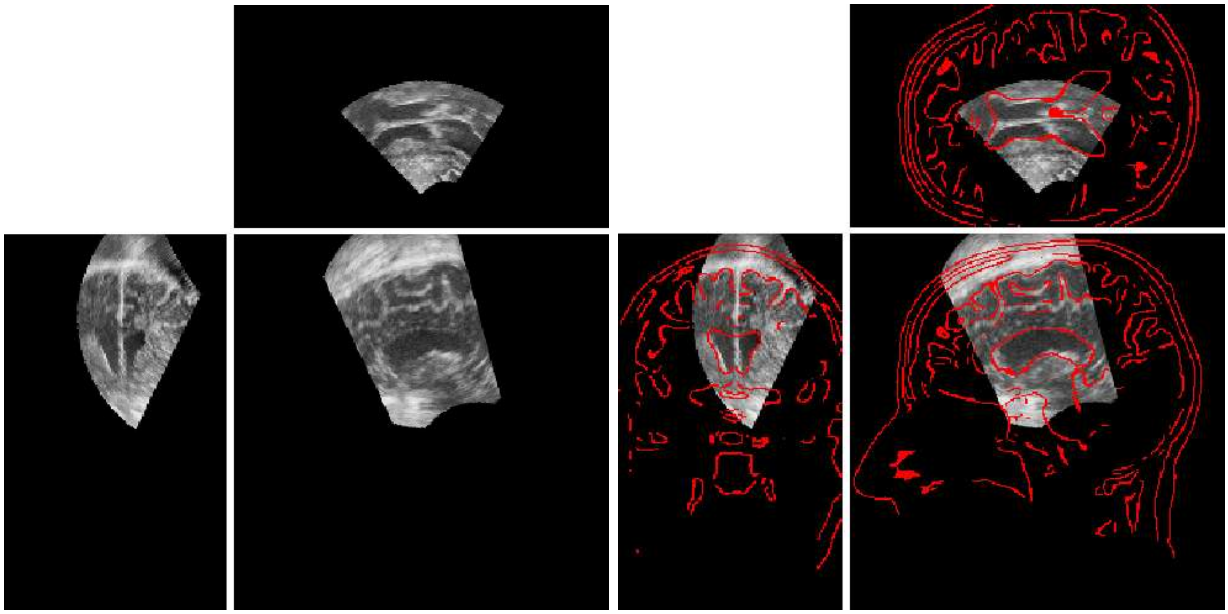
(d) Registered US volume with MR contours overlaid

Figure 5: *Example registration of the baby US/MR images. Note that the US volume is truncated with respect to the original acquisition.*



(a) Pre-operative MR

(b) Intra-operative US (manually initialized)



(c) Registered US volume

(d) Registered US volume with MR contours overlaid

Figure 6: *Example registration of the patient US/MR images. The US corresponds to the smallest acquisition cone among the dataset. This is the image we have found to be responsible for the largest registration error.*

		σ_{rot} (deg)	σ_{trans} (mm)	σ_{test} (mm)
Phantom dataset 8 MR: $0.9 \times 0.9 \times 1$ mm^3 8 US: 0.41^3 mm^3	Multiple MR/MR	0.06	0.10	0.13
	Multiple US/US	0.60	0.40	0.71
	Conservative MR/US	1.15	1.01	1.46
	Expected MR/US	1.06	0.97	1.37
Baby dataset 1 MR: 0.9^3 mm^3 5 US: 0.3^3 mm^3	Multiple US/US	0.10	0.06	0.12
	Conservative MR/US	1.21	0.36	0.90
	Expected MR/US	1.21	0.36	0.89
Patient dataset 2 MR: $0.9 \times 0.9 \times 1.1$ mm^3 3 US: 0.63^3 and 0.95^3 mm^3	MR/MR	0.02	0.02	0.03
	Conservative MR/US	1.57	0.58	1.65
	Expected MR/US	1.57	0.58	1.65

Table 1: Registration errors estimated from registration loops. See section 4.1 for the definitions of *expected error* and *conservative error*.

corner error of 2.61 mm ($\sigma_{\text{trans}} = 0.84$ mm) while the loops involving the two larger US images (real size $170 \times 130 \times 180$, voxel size 0.95^3 mm^3) do have a much smaller corner error of about 0.84 mm ($\sigma_{\text{trans}} = 0.39$ mm). We suspect that a non-rigidity in the smallest US could account for the registration difference between the two MR images. One may notice, however, that this registration error is not visually obvious (see Fig. 6).

5 Robustness study

5.1 Principle

The goal of this section is to study the robustness of our algorithm with respect to varying the initial transformation parameters. For that purpose, we chose one representative US/MR image pair in each dataset and we registered them manually using our interactive matching tool. This gives us a rough "ground truth" transformation which is independent of any registration algorithm. Notice that, although we performed the manual registration as carefully as possible, the "ground truth" does not need to be extremely accurate.

We then performed a number of automatic registrations by initializing the algorithm with random perturbations from the "ground truth" position: a rotation vector δr with random direction and constant magnitude $\|\delta r\| = 15$ degrees, and a translation vector δt with random direction and constant magnitude $\|\delta t\| = 20$ mm. One may hope that these values correspond to the upper bound of the manual registration errors that can be made by a non-expert. For each random transformation, registrations were alternatively performed using six different registration criteria:

- CR: the standard, monovariate correlation ratio as described in [20], using the MR image as

a template,

- CR^g: the same as CR but with the MR gradient norm image as a template,
- MI: mutual information as implemented in [9], combining the US image and the MR image,
- MI^g: mutual information combining the US image and the MR gradient norm image,
- BCR_{quad}: the bivariate CR as described in section 2, but using the quadratic intensity distance ($\rho(x) = x^2$) instead of the Geman-McClure ρ-function,
- BCR : the bivariate CR as used in the accuracy study (see section 4), i.e., using the Geman-McClure ρ-function.

We performed 200 registrations for each registration criterion and for each MR/US pair (yielding a total of $200 \times 6 \times 3 = 3600$ registrations).

Here, we aim at estimating the variability of the registration result with respect to the variability of the initial transformation parameters, no matter wether or not this result corresponds to a good registration. Doing this, we try to answer the question: "if we had used a different initialization, would the algorithm have converged to the same solution?" In other words, we want to characterize: (1) the *robustness*, i.e., the ability of the algorithm to find the "same" minimum of the registration criterion from different initializations, and (2) the *precision*, i.e., the residual variation of the solution when convergence to the same minimum is declared. There is no doubt that the precision is strongly related to the interpolation method, which in this case is PV interpolation for each criterion.

5.2 Computing the robustness and precision

In order to quantify both robustness and precision, we compute a mean transformation for each set of registrations. If the distance of a transformation to this mean is less than a given threshold, the algorithm is declared to be successful and this transformation is used to compute the repeatability (i.e., the variance with respect to the mean transformation), otherwise it is considered as a failure. More precisely, we used a robust version of the Fréchet mean transformation estimation presented in [14] : using the same robust version of the left invariant distance between transformations as in section 4.2, the Fréchet mean rigid transformation is defined as:

$$\bar{T} = \arg\min_{T} \left(\sum_i \mathrm{dist}^2(T_i, T) \right)$$

This minimization is performed using a gradient descent described in [14]. However, this algorithm only gives us a local minimum around the starting point. Thus, to obtain the global minimum, we repeated the minimization by choosing any transformation T_i as a starting point, and we kept the best one.

Within this framework, the success rate is defined as the proportion of transformations that have a distance to $\bar{T}$ less than χ^2, and the precision values are computed on successful transformations using:

$$\sigma_{rot}^2 = \frac{1}{N_{sucess}} \sum_i \|\bar{r}^{(-1)} \circ r_i\|^2 \qquad \text{and} \qquad \sigma_{trans}^2 = \frac{1}{N_{sucess}} \sum_i \|\bar{t} - t_i\|^2$$

By the way, we note that computing the Fréchet mean transformation maximizes the success rate since a success always accounts for less than a failure in the Fréchet minimization.

5.3 Results and discussion

The results of the robustness analysis are reported in table 2. Experiments were reproduced with different χ^2 values (from 10 to 40) without significative differences. We note that the success rate values are probably highly dependent on the amplitude of the initial transformation perturbation, which in this case was rather large (20 mm and 15 degrees).

The main observation is that no measure based only on the MR intensity or only on the MR gradient norm is robust for every dataset. The CR measure using intensity provides consistent

	Criterion	Success rate	Precision for successes		Visual inspection
			rot (deg)	trs (mm)	
Phantom dataset	MI	39 %	0.40	0.27	poor
	MI^g	80 %	0.32	0.17	fair
	CR	52 %	0.43	0.25	poor
	CR^g	71 %	0.25	0.15	fair
	BCR_{quad}	76 %	0.08	0.04	ok
	BCR	76 %	0.14	0.09	ok
Baby dataset	MI	94 %	0.03	0.02	ok
	MI^g	85 %	0.05	0.02	ok
	CR	14 %	0.32	0.10	poor
	CR^g	79 %	0.02	0.01	ok
	BCR_{quad}	68 %	0.12	0.01	ok
	BCR	71 %	0.02	0.01	ok
Patient dataset	MI	29 %	0.53	0.25	ok
	MI^g	0 %	-	-	-
	CR	90 %	0.45	0.17	ok
	CR^g	0 %	-	-	-
	BCR_{quad}	85 %	0.39	0.11	ok
	BCR	55 %	0.43	0.08	ok

Table 2: Robustness results for $\sigma_r = 0.2$ degrees, $\sigma_t = 0.1$ mm and $\chi^2 = 18$.

results only for the patient images, while MI using intensity is unstable except for the baby images. When using the gradient norm information, these measures are useless for the patient images, although stable for the phantom and baby images. This suggests that combining both the intensity and gradient norm information is crucial in terms of robustness.

In all cases, the BCR measures (using either the square intensity distance or the Geman-McClure distance) perform best with respect to precision, and yield acceptable success rates. It seems that $\mathrm{BCR}_{\mathrm{quad}}$ is a bit more robust than BCR: their success rates are comparable for the phantom and baby datasets, but mostly in favour of the former for the patient dataset. This may sound self-contradictory since the advantage of using the Geman-McClure function is precisely to achieve some robustness properties. However, one should not confuse the robustness to initialization with the robustness involved in the use of an S-estimator, which acts as reducing the sensitivity of the registration measure to intensity artifacts (see section 2.4).

A general observation is that quadratic versions of the correlation ratio seem to yield a wider attraction basin than those based on a robust intensity distance. However, we have also observed in practice that registration results using BCR tend to be visually more accurate than when using $\mathrm{BCR}_{\mathrm{quad}}$. Although further quantitative evaluation is needed, we believe at present that using BCR is safer provided a fine initialization has been achieved.

Of course, these figures do not tell us anything about the ability of the different measures to find a *good* registration. They only provide information regarding the consistency of the results with respect to the initial transformation parameters. For each US/MR pair, we inspected the registration result corresponding to the mean transformation as computed in section 5.2. Registration was declared 'ok', 'fair', or 'poor', respectively if there was no obvious misalignment, if a slight misalignment could be seen, or if the result was clearly wrong.

From this subjective evaluation (also reported in table 2), we observe that there is a correlation between the accuracy and robustness performances of the different registration criteria, which was not obvious *a priori*. Here again, the two bivariate measures are the only ones to provide satisfactory results across experiments.

6 Conclusion

We have developed a novel similarity measure for 3D US/MR registration. It is a bivariate and robust generalization of the standard correlation ratio measure. The assumption underlying the bivariate CR is that the US signal may be approximated by a function of both the MR intensity and the MR gradient magnitude. This model does not account for some important aspects of the US physics (in particular, speckle and attenuation). However, since the bivariate CR may be defined in terms of a robust intensity distance, the functional assumption does not need to hold throughout the whole image overlap.

Our implementation of the bivariate CR using Powell's optimization method was successful in rigidly registering a number of US/MR volume pairs from phantom and clinical data. The

registration accuracy was estimated using an original approach that does not require the knowledge of ground truth. We found the worst registration errors (maximum errors in the region of interest defined by the US cone) to be of the order of 1 millimeter. Moreover, the bivariate CR was shown to significantly outperform the conventional CR and MI measures in terms of robustness.

We believe that the method, and especially its robustness, may still be improved in several ways. As discussed in section 2.2, the bivariate CR could be enhanced by incorporating information from the MR gradient orientation with respect to the US scan line. Another interesting problem is how to sample the US image. In our experiments, we considered US volumes as regular lattices, which was probably not optimal given the interpolation problems associated with resampling. Other sampling strategies may be more appropriate, e.g. polar sampling.

Finally, in the context of image-guided surgery, the ultimate goal of US/MR registration is to correct for tissue deformations that arise due to the brain shift and operative manipulations. Therefore, further developments should also include non-rigid registration. Due to the lack of information in US images, this probably involves strong spatial constraints. The definition of such constraints will be the key to a non-rigid implementation of the bivariate CR.

Acknowledgements

The US datasets were provided by ISM, Salzburg, Austria, and the MR datasets were provided by the Max Planck Institute for Psychiatry, AG-NMR, Munich, Germany. We are grateful to our colleagues Dorothee Auer (Max Planck Institute), Michael Rudolph (ISM), and Pr. Ludwig Auer (ISM), for their dedication in this project.

This work was partially supported by la Région PACA (France), and 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).

Last but not least, we would like to thank Djamal Boukerroui and Mike Brady for their input.

References

[1] 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 Troccaz J., Grimson E., and Mösges R., editors, *First Joint Conference on Computer Vision, Virtual Reality and Robotics in Medicine and Medical Robotics and Computer-Assisted Surgery (CVRMed-MRCAS'97)*, volume 1205 of *Lecture Notes in Computer Science*, pages 459–466, 1997.

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

[3] R. Deriche. Recursively Implementing the Gaussian and Its Derivatives. In *Proc. Second International Conference On Image Processing*, pages 263–267, Singapore, September 1992.

[4] 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. In *Applications of Computer Vision in Medical Image Processing*, volume 1124 of *Excerpta Medica – Int. Congress Series*, pages 225–230. Elsevier, 1996.

[5] 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 *Second International Conference on Medical Image Computing And Computer-Assisted Intervention (MICCAI'99)*, volume 1679 of *Lecture Notes in Computer Science*, pages 920–927, 1999.

[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 *Visualization in Biomedical Computing*, volume 2359 of *SPIE Proceedings*, pages 252–259, Rochester, MN, USA, October 1994. SPIE Press.

[7] G. Ionescu, S. Lavallée, and J. Demongeot. Automated Registration of Ultrasound with CT Images: Application to Computer Assisted Prostate Radiotherapy and Orthopedics. In *Second International Conference on Medical Image Computing And Computer-Assisted Intervention (MICCAI'99)*, volume 1679 of *Lecture Notes in Computer Science*, pages 768–777, Cambridge, UK, September 1999. Springer Verlag.

[8] A. King, J. Blackall, G. Penney, P. Edwards, D. Hill, and D. Hawkes. Bayesian Estimation of Intra-operative Deformation for Image-Guided Surgery Using 3-D Ultrasound. In S. Delp, A.M. DiGioia, and B. Jaramaz, editors, *Third International Conference on Medical Image Computing And Computer-Assisted Intervention (MICCAI'00)*, volume 1935 of *Lecture Notes in Computer Science*, pages 588–597, Pittsburgh, October 11-14 2000. Springer Verlag.

[9] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens. Multimodality Image Registration by Maximization of Mutual Information. *IEEE Transactions on Medical Imaging*, 16(2):187–198, 1997.

[10] J. B. A. Maintz and M. A. Viergever. A survey of medical image registration. *Medical Image Analysis*, 2(1):1–36, 1998.

[11] C. Nikou, F. Heitz, J.-P. Armspach, and I.-J. Namer. Single and multimodal subvoxel registration of dissimilar medical images using robust similarity measures. In *Medical Imaging 1998: Image Processing (MI'98)*, volume 3338 of *SPIE Proceedings*, pages 167–178. SPIE Press, April 1998.

[12] S. Ourselin, A. Roche, S. Prima, and N. Ayache. Block Matching: a General Framework to Improve Robustness of Rigid Registration of Medical Images. In S. Delp, A.M. DiGioia,

and B. Jaramaz, editors, *Third International Conference on Medical Robotics, Imaging And Computer Assisted Surgery (MICCAI'00)*, volume 1935 of *Lecture Notes in Computer Science*, pages 557–566, Pittsburgh, USA, October 11-14 2000. Springer Verlag.

[13] 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.

[14] X. Pennec. Computing the mean of geometric features - application to the mean rotation. Technical Report 3371, INRIA, March 1998.

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

[16] X. Pennec, C.R.G. Guttmann, and J.P. Thirion. Feature-based registration of medical images: Estimation and validation of the pose accuracy. In *First International Conference on Medical Image Computing And Computer-Assisted Intervention (MICCAI'98)*, number 1496 in Lecture Notes in Computer Science, pages 1107–1114, Cambridge, MA, USA, October 1998. Springer Verlag.

[17] B.C. Porter, D.J. Rubens, J.G. Strang, S. Totterman, and K.J. Parker. Three-Dimensional Registration and Fusion of Ultrasound and MRI Using Major Vessels as Fiducial Markers. *IEEE Transactions on Medical Imaging*, 20(4):354–359, April 2001.

[18] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling. *Numerical Recipes in C*. Cambrige University Press, 2nd edition, 1992.

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

[20] A. Roche, G. Malandain, X. Pennec, and N. Ayache. The Correlation Ratio as a New Similarity Measure for Multimodal Image Registration. In *First International Conference on Medical Robotics, Imaging And Computer Assisted Surgery (MICCAI'98)*, volume 1496 of *Lecture Notes in Computer Science*, pages 1115–1124, Cambridge, MA, USA, October 1998. Springer Verlag.

[21] A. Roche, X. Pennec, G. Malandain, N. Ayache, and S. Ourselin. Generalized Correlation Ratio for Rigid Registration of 3D Ultrasound with MR Images. Technical Report 3980, INRIA, July 2000.

[22] Peter J. Rousseeuw and Annick M. Leroy. *Robust Regression and Outlier Detection*. Wiley Series In Probability And Mathematical Statistics, first edition, 1987.

[23] C. Studholme, D. L. G. Hill, and D. J. Hawkes. An overlap invariant entropy measure of 3D medical image alignment. *Pattern Recognition*, 1(32):71–86, 1998.

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

[25] P. Viola and W. M. Wells. Alignment by Maximization of Mutual Information. *International Journal of Computer Vision*, 24(2):137–154, 1997.