Generalized Correlation Ratio for Rigid Registration of 3D Ultrasound with MR Images

Alexis Roche, Xavier Pennec, Grégoire Malandain, Nicholas Ayache, Sébastien Ourselin

INRIA

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Generalized Correlation Ratio for Rigid Registration of 3D Ultrasound with MR Images

Alexis Roche — Xavier Pennec — Grégoire Malandain —

Nicholas Ayache — Sébastien Ourselin

N° 3980

Juillet 2000

THÈME 3

ISSN 0249-6399 ISRN INRIA/RR--3980--FR+ENG

INRIA
SOPHIA ANTIPOLIS

Generalized Correlation Ratio for Rigid Registration of 3D Ultrasound with MR Images

Alexis Roche , Xavier Pennec , Grégoire Malandain ,
Nicholas Ayache , Sébastien Ourselin

Thème 3 — Interaction homme-machine,
images, données, connaissances
Projet Epidaure

Rapport de recherche n° 3980 — Juillet 2000 — 24 pages

Abstract: Automatic processing of 3D ultrasound (US) is of great interest for the development of innovative and low-cost computer-assisted surgery tools. In this paper, we present a new image-based technique to rigidly register intra-operative 3D US with pre-operative Magnetic Resonance (MR) data. Automatic registration is achieved by maximization of a similarity measure that generalizes the correlation ratio (CR). This novel similarity measure has been designed to better take into account the nature of US images. A preliminary cross-validation study has been carried out using a number of phantom and clinical data. This indicates that the worst registration errors are of the order of the MR image resolution.

Key-words: registration, ultrasound images, magnetic resonance images, correlation ratio

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles, B.P. 93, 06902 Sophia Antipolis Cedex (France)
Téléphone : 04 92 38 77 77 - International : +33 4 92 38 77 77 —Fax : 04 92 38 77 65 - International : +33 4 92 38 77 65

Rapport de corrélation généralisé pour le recalage d'images ultrasonores 3D avec des images par résonance magnétique

Résumé : Le traitement des images ultrasonores 3D est d'un intérêt considérable pour le développement d'outils innovants et bon marché en chirurgie assistée par ordinateur. Nous présentons ici une technique utilisant le contenu des images pour recaler rigidement des images ultrasonores per-opératoires avec des images pré-opératoires acquises par résonance magnétique (IRM). Le recalage est réalisé de façon automatique par maximisation d'une mesure de similarité qui généralise le rapport de corrélation. Cette nouvelle mesure a été conçue dans l'optique de mieux tenir compte de la nature des images ultrasonores. Nous proposons une étude préliminaire de validation croisée utilisant un certain nombre de données fantômes et cliniques. Les résultats indiquent que les pires erreurs de recalage sont de l'ordre de la résolution de l'IRM.

Mots-clés : recalage, images ultrasonores, images par résonance magnétique, rapport de corrélation

Contents

1 Introduction

Over the last years, the development of real-time 3D ultrasound (US) imaging has revealed a number of potential applications in image-guided surgery. 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.

The registration of US with pre-operative Magnetic Resonance (MR) images will allow the surgeon to accurately localize the course of instruments in the operative field, resulting in minimally invasive procedures. At present, few papers have been published on this particular registration problem [6]. Most of the approaches that have been proposed are based on stereotactic systems. For instance, in [10] registration is achieved by tracking the US probe with a DC magnetic position sensor. Existing image-based methods match homologous features extracted from both the US and MR data. Features are user-identified in [2], while semi-automatically extracted in [3]. More recently, Ionescu *et al* [4] registered US with Computed Tomography (CT) data after automatically extracting contours from the US using watershed segmentation.

The present registration technique expands on the correlation ratio (CR) method [15]. It is an intensity-based approach as it does not rely on explicit feature extraction. In a previous work [14], we reported preliminary results of US/MR registration by maximization of CR and mutual information (MI). While results obtained using CR were more satisfactory than when using MI, the method was still lacking precision and robustness with respect to the initialization of the transformation parameters.

In this paper, we have improved the CR method following three distinct axes: (1) using the gradient information from the MR image, (2) reducing the number of intensity parameters to be estimated, and (3) using a robust intensity distance. These extensions are presented 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 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 $f(J \circ T)$ be as similar as possible to I. In a first approach, this could be achieved by minimizing the following cost function:

$$\min_{T,f} \quad \sum_k \left[I(x_k) - f(J(T(x_k))) \right]^2 , \tag{1}$$

which integrates over the voxel positions x_k in image I. In the following, we will use the simplified notations $i_k \equiv I(x_k)$, and $j_k^{\downarrow} \equiv J(T(x_k))$, where the arrow expresses the dependence in T. 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 registration, we always choose the MR as the template.

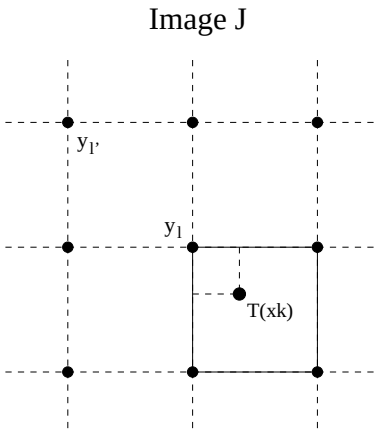
In practice, the criterion defined in eq (1) cannot be computed exactly due to the finite nature of the template image. One obvious problem is that the transformed position of a voxel will generally not match a grid point of J, such that the corresponding intensity $j_k^{\downarrow}$ is unknown. A classical approach is then to linearly interpolate $j_k^{\downarrow}$ using the eight neighbours of $T(x_k)$ in the grid of J. However, instead of interpolating the image intensity, we may directly interpolate the incremental contribution of x_k, i.e., $[i_k - f(j_k^{\downarrow})]^2$. The difference between these two approaches is illustrated in figure 1. In fact, the last method is equivalent to the so-called partial volume (PV) interpolation, originally proposed by Maes *et al* [5] in the context of joint histogram computation. We have found PV to generally outperform classical linear interpolation in terms of smoothness of the resulting registration criterion.

Another difficulty to compute eq (1) is that some points x_k may transform outside the template domain and lack eight grid neighbours. We decide not to take into account such points in the computation of the registration criterion. Doing so without particular attention, the criterion would become zero when every point 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. Justifications of this particular normalization strategy will be found in [15], while related normalization issues are discussed in [17, 18].

These practical considerations lead us to the following modification of eq (1):

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

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



Linear interpolation weights:

$$w_{kl}^{\downarrow} = \beta(T(x_k) - y_l), \quad \text{with :}$$
$$\begin{cases} \beta(z) = (1 - |z^1|)(1 - |z^2|), & \text{if } |z^1| \leq 1,\ |z^2| \leq 1 \\ \beta(z) = 0, & \text{otherwise} \end{cases}$$

Intensity interpolation:

$$\left[i_k - f(j_k^{\downarrow})\right]^2 \approx \left[i_k - f\left(\textstyle\sum_l w_{kl}^{\downarrow} j_l\right)\right]^2$$

PV interpolation:

$$\left[i_k - f(j_k^{\downarrow})\right]^2 \approx \textstyle\sum_l w_{kl}^{\downarrow} [i_k - f(j_l)]^2$$

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

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 [15]. 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 practice, the maximization of η^2 is performed using Powell's method.

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 *difference* of acoustical impedance caused by successive tissue layers. Ideally, the US signal should be high at the interfaces, and zero within homogeneous tissues.

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 will remain uniform in the remapped MR and, thus, this procedure is not able to account for intensity variations at the interfaces. To enable a better prediction, we propose to use the modulus of the MR gradient 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 that maps double-valued intensities to single-valued intensities. The MR gradient is practically computed by convolution with a Gaussian kernel [7].

At first glance, using the modulus of the MR gradient does not appear to be fully acceptable from the physics of US imaging. In fact, the US signal which is produced at an interface also depends 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.

We have not yet studied the effect of using the projected MR gradient versus the modulus. Still, we believe that there are good reasons not to take into account information from the gradient orientation, at least as a first-order approximation. Through diffraction of the ultrasound beam on interfaces, the received echo is actually less dependent on the direction propagation than would be the case with perfectly specular reflection. This, added to log-compression, tends to equalize the response values corresponding to different tissue orientations.

2.3 Parametric intensity fit

If we put 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 [14]. 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 draw 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 y_l, respectively in the original MR, M, and in $\|\nabla M\|$. We are searching for a mapping of the form:

$$f(m_l, g_l) = \sum_{p+q \leq d} \theta_{pq}\, m_l^p g_l^q, \tag{3}$$

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 appendix A that minimizing eq (2) with respect to the polynomial coefficients brings us to a weighted least square (WLS) linear regression problem. As is standard, this is solved by the method of singular value decomposition (SVD).

This polynomial fitting procedure, however, has significant extra computational cost with respect to the unconstrained fitting. Recall that, in the basic version of the CR method, f is updated for each transformation trial. Such a strategy is no longer affordable when estimating a polynomial function. Instead, the minimization of eq (2) may be performed alternatively along T and f, resulting in the following algorithm: (1) given a current transformation estimate T, find the best polynomial f and remap J accordingly; (2) given a remapped image $f(J)$, minimize $C(T, f)$ with respect to T using Powell's method; third, return to (1) if T or f has evolved.

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 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 artefacts, we do not expect this assumption to be perfectly true. Shadowing, duplication or interference artefacts may cause large variations of the US intensity from its predicted value, and this even when the images are perfectly registered. Such bad intensity matches are false negative.

The sensitivity of the registration criterion to false negative may be reduced by replacing the expression $(1/n^{\downarrow}) \sum_{k,l} w_{kl}^{\downarrow} [i_k - f(j_l)]^2$ in eq (2) with a robust scale estimate. A similar idea was developed in [8]. We here propose to build such an estimate from a one-step S-estimator [16]:

$$\hat{S}^2(T, f) = \frac{S_0^2}{K n^{\downarrow}} \sum_{k,l} w_{kl}^{\downarrow}\, \rho\left(\frac{i_k - f(j_l)}{S_0}\right), \tag{4}$$

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 implies few modifications of 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 as shown in appendix B, generally requiring no more than 5-6 iterations.

In our implementation, we have opted for the Geman-McClure ρ-function, $\rho(x) = \frac{1}{2}x^2/(1+\frac{x^2}{c^2})$, 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 normalization constant is then $K = 0.416$.

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, $|i_k - f(j_l)|$ (see appendix B for details). Due to the initial misalignment, S_0 tends to be overestimated. Thus, it may not allow to reject efficiently bad intensity matches. For that 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 report were performed within the framework of the European project ROBOSCOPE (`http://www.ibmt.fhg.de/Roboscope/home.htm`). 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 undergo sever interpolation artifacts (blurring) in areas which are distant from the probe. In the following, we will refer to US images as cubic images, but one has to keep 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. Each acquisition consists in one 3D MR image and one 3D US image.

Both balloons were initially filled with 40ml of fluid (acquisition 0). During the first series of 5 acquisitions, the ellipsoid was filled in steps of 10ml, while the complex ellipsoid was kept constant. During the second series, the ellipsoid was deflated and the complex ellipsoid filled by steps of 10ml:

Acquisition number	0	1	2	3	4	5	6	7	8	9	10
Ellipsoid volume (ml)	40	50	60	70	80	90	80	70	50	60	40
Complex ellipsoid volume (ml)	40	40	40	40	40	40	50	60	70	80	90

All MR images do have $256 \times 256 \times 124$ voxels of size $0.90 \times 0.90 \times 1\text{mm}^3$ except images 0 and 10 where the voxel size is $0.94 \times 0.94 \times 1\,\text{mm}^3$. All (Cartesian) US images have $184 \times 184 \times 184$ isotropic voxels of size 0.41 mm.

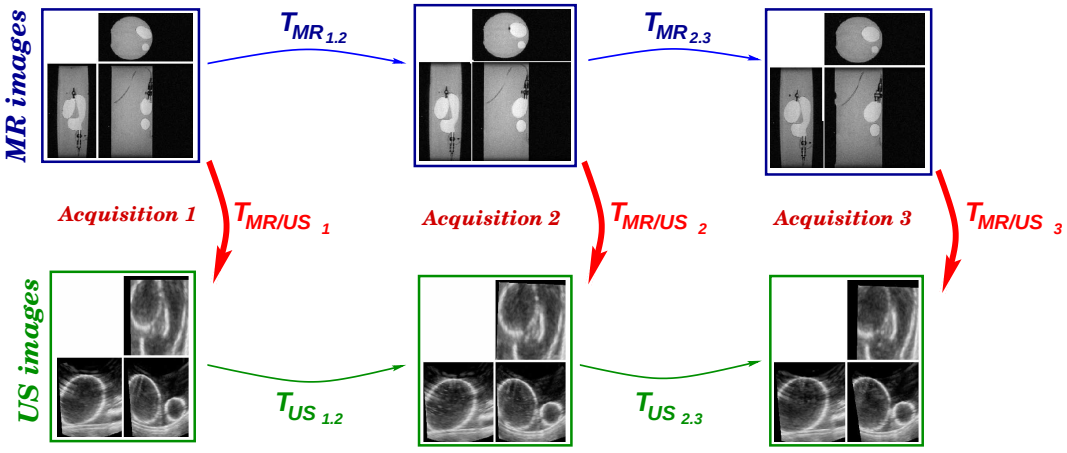


Figure 2: *Original MR and US images (before registration) of the Phantom and the rigid transformations that relate them.*

This dataset would be perfect for the validation of MR and US registration if all MR (resp. all US) images were in exactly 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 if the rigid registration of MR images seems to be quite accurate, the registration of US images is specially difficult and inaccurate (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 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%, 80%, 90%, 100%).

The MR image has $256 \times 256 \times 124$ voxels of size $0.89 \times 0.89 \times 0.9\,\text{mm}^3$. The Cartesian US images have $184 \times 184 \times 184$ isotropic voxels of size 0.29 mm.

3.3 Patient dataset

This dataset is a real surgical case: two sagittal MR T1 images with and without a contrast agent were acquired before surgery. After craniotomy (dura mater still closed), three 3D US images were acquired to precisely locate the tumor to resect. The MR and US image sizes are summarised below:

	Image dimension			voxel size		
MR1	256	256	124	0.90	0.90	1.10
MR2	256	256	124	0.94	0.94	1.10
US1	236	136	164	0.61	0.61	0.61
US2	236	136	164	0.63	0.63	0.63
US3	190	136	194	0.95	0.95	0.95
US4	180	136	188	0.95	0.95	0.95

4 Accuracy study

We computed all the MR/US registrations using the previously described algorithm. 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 transformation. Here, 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. However this user interaction could be alleviated using a calibration system such as the one described in [10]. After initialization, we observed that the algorithm found residual displacements in the range of 10 mm and 10 degrees. In all the experiments, the gradient norm of the MR image was computed by linear filtering using a Gaussian kernel with $\sigma = 1$ voxel.

4.1 Principle of the accuracy evaluation

To estimate the accuracy of the algorithm, one should ideally compare the result of a registration with a gold-standard. Up to our knowledge, there is no such gold-standard for MR/US registration. 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 on test points. What we call a registration loop is a succession of composition of transformations leading to a global transformation that is theoretically the identity. 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 below.

If we were given perfectly registered images within each modality, this loop would only be disturbed from the identity by errors on the two MR/US registrations. Since variances are additive, the variance of the observed error should roughly be: $\sigma^2_{loop} = 2\sigma^2_{MR/US}$. Unfortunately, we are not provided with a ground truth registration within each modality: we need to estimate it. This time, we are combining one MR/MR, one US/US and two MR/US registrations. Thus, the variance of the loop error will be roughly: $\sigma^2_{loop} \simeq 2\sigma^2_{MR/US} + \sigma^2_{MR/MR} + \sigma^2_{US/US}$. The *expected* MR/US accuracy is then:

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

However, what we really measure is σ_{loop}, from which we may infer the expected MR/US accuracy provided that $\sigma^2_{MR/MR}$ and $\sigma^2_{US/US}$ are faithfully estimated (see section 4.2). In any event, the worst-case or *conservative* MR/US accuracy is given by $\sigma_{MR/US} \simeq \sigma_{loop}/\sqrt{2}$. In order to minimize the influence of intra-modality

registrations errors in this figure, we need to provide very accurate MR/MR and US/US registrations. For that purpose, we designed the following algorithm.

4.2 Multiple intra-modality registration

To relate n images together, we need to estimate $n-1$ rigid registrations $\bar{T}_{i,i+1}$. To obtain a very good accuracy, we chose to register all image pairs, thus obtaining $n(n-1)$ transformations $\hat{T}_{i,j}$, and estimate the transformations $\bar{T}_{i,i+1}$ that best explain our measurements in the least-square sense, i.e. that minimizes the following criterion:

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

We used a robust variant of the left invariant distance on rigid transformations introduced in [13]: let σ_r and σ_t be typical scales on the rotation angle and on the translation magnitude and χ^2 a threshold. If (r,t) are the rotation vector and the translation of transformation T, the distance between two transformations is

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

The standard deviations σ_r and σ_t are manually adjusted to correspond roughly to the residual rotation and translation error after convergence. To obtain the minimum, we used a Newton gradient descent similar to the one described in [13], but on all transformations $\bar{T}_{i,i+1}$ together (see details in appendix C). Moreover, we end-up with an estimation of the variance of the estimated transformations, which allows us to obtain good estimations of the previously introduced $\sigma^2_{MR/MR}$ and $\sigma^2_{US/US}$.

Since the space of rigid transformations is non-Euclidean, this minimization only gives us a local minimum around the initial values of the transformations $\bar{T}_{i,i+1}$. To improve the robustness and find a more global minimum, we compute for each index i all the possible transformations from i to $i+1$ based on two measured transformations (with the convention $T_{i,i} = \mathrm{Id}$):

$$T^k_{i,i+1} = \hat{T}_{k,i+1}\circ\hat{T}_{i,k} \qquad \text{and} \qquad (T^k_{i+1,i})^{(-1)} = \left(\hat{T}_{k,i}\circ\hat{T}_{i+1,k}\right)^{(-1)}$$

and chose among them the one that minimizes the robust distance with all the others to initialize $\bar{T}_{i,i+1}$. With this method, we observed a convergence of the multiple registration algorithm to the right transformations even with more than 50% outliers.

4.3 Phantom dataset

In this experiment, we used 8 acquisitions with different balloons 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 for the US images.

The main problem for the multiple intra-modality registration of the phantom images is that the acquisitions are intrinsically non-rigid except the outer part of the container and in a certain extent one of the two valves.

Thus, intra-modal registrations of MR images were carried out using a feature-based registration algorithm known to handle a large amount of outliers [12]. With the US images, the high amount of non-rigidity is a much more acute problem because only a small part of the container is visible (see Fig. 2). As it is very difficult to extract meaningful features on these images, we used the robust block matching technique proposed in [9].

As we are testing the *rigid* registration, we cannot register MR and US images across acquisitions. Thus, the simplest loops we can use for accuracy estimations are the $n(n-1)$ following loops: $US_i \to MR_i \to MR_j \to US_j \to US_i$. 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 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.

As we have no or very few deformations within the images, we can rigidly register all the US images onto our single MR and test the 30 following loops $US_i \to MR \to US_j \to US_i$ (only 5 of them being independent). For that, we still need to register the US images together. Since there are no deformations between the acquisitions (we only have a motion of the probe and a degradation of the image quality) the algorithm is much more efficient and accurate than for the Phantom.

4.5 Patient dataset

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. In this experiment, we use the three US images that are large enough to contain the ventricles.

The two MR images were registered using the feature based method with a very high accuracy (probably overestimated as we only have two images) and we tested the loops $US_i \to MR_0 \to MR_1 \to US_i$. As the acquisition cone of the US probe is completely within the Cartesian image (see Fig. 5), the region of interest is much smaller than the images size: we took our typical points at the corners of a $80 \times 80 \times 80$ mm^3 cube centered in the image.

4.6 Results and discussion

We put in table (1) the standard deviations of the residual rotation, of the residual translation and of the displacement of the test points for the different registration involved. Since we took the origin of the images at the center, the σ_{trans} value corresponds to the mean error at the center of the image while σ_{test} corresponds to the maximum registration error within the US image (except for the patient experiment, the test points are taken at the corners of the image).

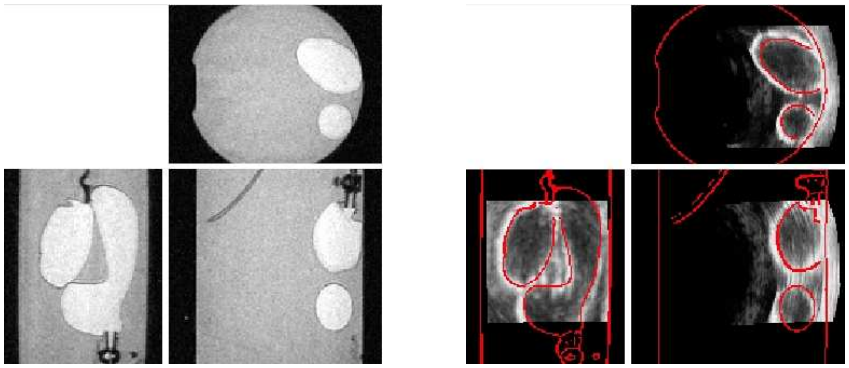


Figure 3: Example registration of the MR and US images of the phantom. Left: original MR image. Right: registered US image with the MR contours superimposed.

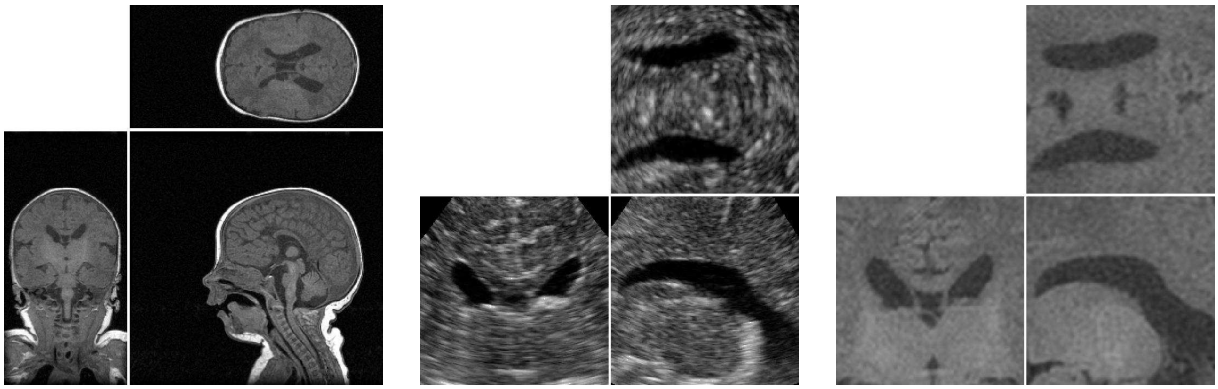


Figure 4: *Example registration of the MR and US images of the baby. From left to right: original MR T1, original US and registered MR.*

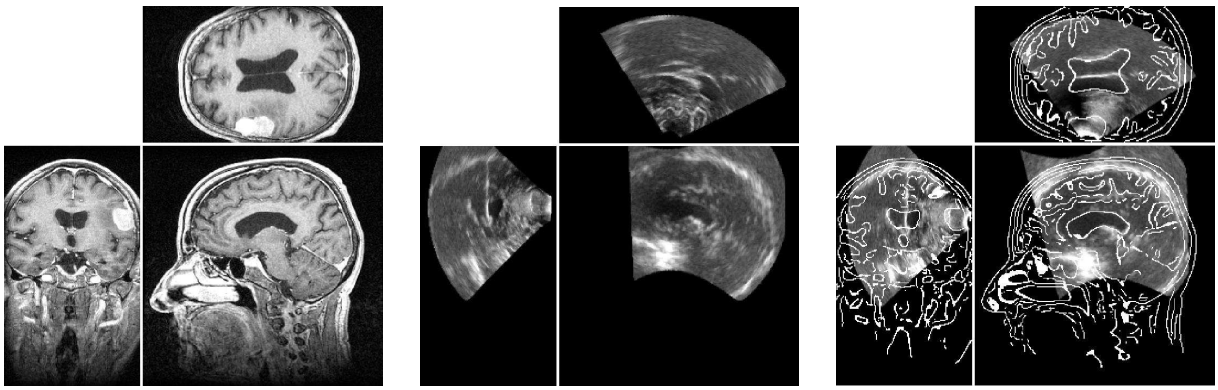


Figure 5: *Example registration of the MR and US images of the patient. From left to right: MR T1 image with a contrast agent, manual initialization of the US image registration, and result of the automatic registration with the MR contours superimposed.*

		σ_{rot} (deg)	σ_{trans} (mm)	σ_{test} (mm)
Phantom dataset MR: $0.9 \times 0.9 \times 1$ mm^3 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 MR: 0.9^3 mm^3 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 MR: $0.9 \times 0.9 \times 1.1$ mm^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: *Accuracy results. These figures have to be compared with the voxel sizes of the MR images.*

The results on the Phantom show that the MR/US registration accuracy is of the order of the MR resolution. One could probably expect a better 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 under 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 corner error of 2.61 mm ($\sigma_{\text{trans}} = 0.84$ mm) while the loops involving the two larger US images (real size $170 \times 130 \times 180$, voxels 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. Another explanation could be a misestimation of the sound speed for this small US acquisition leading to a false voxel size and once again the violation of the rigidity assumption.

5 Robustness study

5.1 Principle

The goal of this section is to study the consistency of our algorithm with respect to varying the initial transformation parameters. For that purpose, we chose one representative US and MR images in each dataset and registered them manually using an interactive matching tool. This gives us a rough "ground truth" transformation which is independent of any automated registration algorithm. Notice, 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 Δt with random direction and constant magnitude $\|\Delta t\| = 20$ mm. These

values hopefully correspond to a large 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:

- the standard correlation ratio (CR) [15],
- the correlation ratio using the MR gradient norm image as a template,
- mutual information (MI) as implemented in [5],
- mutual information using the MR gradient norm image as a template,
- the generalized correlation ratio (GCR) as described in section 2, but without the use of a robust intensity distance (which corresponds to the choice $\rho(x) = x^2$),
- the generalized correlation ratio 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).

Doing this, we try to answer the question: "if we had used a different initialization, would the algorithm have found the same result?" 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 decided. This has not to be confused with the *accuracy* evaluation described in section 4, where we characterized the variability of the registration result with respect to the variability of the data. Here, we aim at characterizing the variability of the registration result with respect to the variability of the initial transformation parameters, no matter this result corresponds to a good registration or not.

5.2 Computing the robustness and precision

In order to quantify both robustness and precision, we compute for each set of registrations a mean transformation. 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 [11] : 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 [11]. However, this algorithm only gives us a local minimum around the starting point. Thus, to obtain the global minimum, we repeated this minimization for every transformation T_i as the 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 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 should note that computing the Fréchet mean transformation maximizes the success rate since a success always account for less than a failure in the Fréchet minimization.

5.3 Discussion of results

The results of the robustness analysis are given in table 2. Experiments were reproduced with different χ^2 values (from 10 to 40) without significative differences.

Firstly, we should notice that the magnitude of the initial transformation perturbation corresponds to an upper bound of the acceptable initialization. This explains why no algorithm converges in 100 % of the cases.

The second 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 results only for the

	Criterion	Success rate	Precision for successes		Visual inspection
			rot (deg)	trs (mm)	
Phantom dataset	MI (intensity)	39 %	0.40	0.27	poor
	MI (grad norm)	80 %	0.32	0.17	fair
	CR (intensity)	52 %	0.43	0.25	poor
	CR (grad norm)	71 %	0.25	0.15	fair
	GCR / square distance	76 %	0.08	0.04	ok
	GCR / Geman-McClure	76 %	0.14	0.09	ok
Baby dataset	MI (intensity)	94 %	0.03	0.02	ok
	MI (grad norm)	85 %	0.05	0.02	ok
	CR (intensity)	14 %	0.32	0.10	poor
	CR (grad norm)	79 %	0.02	0.01	ok
	GCR / square distance	68 %	0.12	0.01	ok
	GCR / Geman-McClure	71 %	0.02	0.01	ok
Patient dataset	MI (intensity)	29 %	0.53	0.25	ok
	MI (grad norm)	0 %	-	-	-
	CR (intensity)	90 %	0.45	0.17	ok
	CR (grad norm)	0 %	-	-	-
	GCR / square distance	85 %	0.39	0.11	ok
	GCR / Geman-McClure	55 %	0.43	0.08	ok

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

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 GCR measure (using either the square intensity distance or the Geman-McClure distance) performs at the best level with respect to precision, and yields acceptable success rates. It seems that the former is a bit more robust than the latter: the 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). In general, using the robust version of GCR tends to narrow the attraction basin of the algorithm, thus making it less robust to initialization. However, we observed from practical experience that registration results using the robust GCR seem to be generally more accurate than when using GCR with the square distance. Although further comparative evaluation is needed, we believe at present that it is safe to use the robust GCR provided that 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 set of experiments, 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 seems to be a correlation between the accuracy and robustness performances of the different registration criteria, which is not obvious *a priori*. Here again, the two GCR measures are the only ones to provide satisfying results in all experiments.

6 Conclusion

We have presented 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 nature of US images. Incidentally, we believe that the generalized CR could be considered in other registration problems where conventional similarity measures fail.

Testings were performed on several phantom and clinical data, and accuracy was evaluated using an original method that does not require the knowledge of a ground truth. We estimated the worst registration errors (errors at the US corners) to be of the order of 1 millimeter.

In our experiments, we used US images stored in Cartesian format. This obviously does not help registration owing to the fact that: 1) intensities of voxels distant from the probe are unreliable, 2) resampling on a cubic lattice artificially increases the concentration of such voxels, consequently increasing their influence on the

registration criterion. We believe that using polar US images as inputs to our registration algorithm could significantly improve both its accuracy and robustness.

Further developments also include non-rigid registration, in order to correct for false distance artefacts in US, as well as for tissue deformations due to brain shift and operative manipulations.

Acknowledgements

The datasets used in this report were provided by ISM, Salzburg, Austria, for the US datasets, and the Max Planck Institute for Psychiatry, AG-NMR, Munich, Germany, for the MR datasets. 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).

References

[1] D.R.K. Brownrigg. The weighted median filter. *Communications of the ACM*, 27(8):807–818, August 1984.

[2] 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.

[3] 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.

[4] 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*, volume 1679 of *LNCS*, pages 768–777, Cambridge, UK, September 1999. Springer Verlag.

[5] 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.

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

[7] O. Monga, R. Deriche, G. Malandain, and J.-P. Cocquerez. Recursive filtering and edge tracking: two primary tools for 3-D edge detection. *Image and Vision Computing*, 9(4):203–214, August 1991.

[8] 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.

[9] S. Ourselin, A. Roche, S. Prima, and N. Ayache. Block Matching: a General Framework to Improve Robustness of Rigid Registration of Medical Images. In *MICCAI'2000*, Pittsburgh, USA, October 2000.

[10] 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.

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

[12] X. Pennec, N. Ayache, and J.P. Thirion. Chap. 31: Landmark-based registration using features identified through differential geometry. In *Handbook of Medical Imaging*. Academic Press, 2000. in press.

[13] 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)*, number 1496 in LNCS, pages 1107–1114, Cambridge, USA, October 1998. Springer Verlag.

[14] 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.

[15] A. Roche, G. Malandain, X. Pennec, and N. Ayache. The Correlation Ratio as a New Similarity Measure for Multimodal Image Registration. In *Proc. MICCAI'98*, volume 1496 of *LNCS*, pages 1115–1124, Cambridge, MA, USA, October 1998. Springer Verlag.

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

[17] 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.

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

Appendices

A Least squares parametric fitting

At a fixed transformation T, consider the problem defined in section 2.3, that is minimizing the criterion,

$$C(T,f) = \frac{\sum_{k,l} w_{kl}^{\downarrow} \left[i_k - f(j_l)\right]^2}{n^{\downarrow} \mathrm{Var}(I^{\downarrow})}, \tag{6}$$

among the set of functions f that can be expressed as linear combinations of specified basis functions ϕ_p. The unknown f is thus parameterized by a finite set of coefficients θ_p:

$$f(j) = \sum_p \theta_p \phi_p(j).$$

A polynomial fitting corresponds to the special case where the basis functions are chosen as monoms, i.e., $\phi_p(j) = j^p$ if J is a real valued image and, more generally, functions of the form $\phi_p(j) = \left(j^1\right)^{k_1}\left(j^2\right)^{k_2}\ldots\left(j^d\right)^{k_d}$ if J is a d-dimensional vectorial image with components $(j^1, j^2, \ldots, j^d)$ at a given voxel position. Also, since the denominator in eq (6) is clearly independent of f, our problem amounts to solving for the coefficients vector $\Theta = (\theta_1, \theta_2, \ldots)^t$ that minimizes the numerator:

$$\hat{\Theta} = \arg\min_{\Theta} \sum_{k,l} w_{kl}^{\downarrow} \left[i_k - \sum_p \theta_p \phi_p(j_l)\right]^2,$$

which is nothing but a weighted least squares linear regression problem. Since the criterion is quadratic in Θ, a necessary and sufficient condition for Θ to be a solution is $\partial C/\partial\Theta = 0$, yielding:

$$\forall q, \quad \sum_{k,l} w_{kl}^{\downarrow} \phi_q(j_l) \left[\sum_p \theta_p \phi_p(j_l) - i_k\right] = 0, \tag{7}$$

or equivalently, by rewriting this expression in algebraic form,

$$A\Theta = B, \tag{8}$$

where A is the square matrix with coefficients $A_{pq} = \sum_{k,l} w_{kl}^{\downarrow} \phi_p(j_l)\phi_q(j_l)$, and B is the column vector with components $B_p = \sum_{k,l} w_{kl}^{\downarrow} i_k \phi_p(j_l)$. Provided that A is not singular, the solution is unique and it is given by $\hat{\Theta} = A^{-1}B$. Owing to the fact that A is symmetric, an efficient method to compute A^{-1} which has proven to be numerically stable is the singular value decomposition (SVD). In case A is singular (unlikely to occur in practice), there is no longer a unique solution but the SVD will still provide one since it actually computes the pseudo-inverse, $A^{\dagger}$, so that the equality $AA^{\dagger} = I$ holds in all circumstances.

B Robust parametric fitting

As an extension of the previous least squares regression, consider now the more general regression problem defined by the minimization of the cost function introduced in section 2.4,

$$C(T,f) = \frac{S_0^2}{Kn^{\downarrow}\mathrm{Var}(I^{\downarrow})} \sum_{k,l} w_{kl}^{\downarrow}\, \rho\left(\frac{i_k - f(j_l)}{S_0}\right), \tag{9}$$

where ρ is the objective function associated with a given M-estimator, and S_0, K are known constants. Without loss of generality, we may assume that $S_0 = 1$, which amounts to the variable change $\rho(x) \leftarrow \rho(x/S_0)$. Here again, the denominator is seen to play no role in the minimization with respect to f (at fixed T), so that our problem boils down to finding:

$$\hat{\Theta} = \arg\min_{\Theta} \sum_{k,l} w_{kl}^{\downarrow} \rho\left(i_k - \sum_p \theta_p \phi_p(j_l)\right),$$

whenever f is parameterized in the same fashion as in appendix A. Unless ρ is chosen as $\rho(x) = x^2$, as was previously the case, the criterion is generally not quadratic with respect to Θ. Therefore, determining an analytical solution is hopeless[1]. In practice, one seeks a solution that fulfills the necessary condition $\partial C/\partial\Theta = 0$, yielding:

$$\forall q, \quad \sum_{k,l} w_{kl}^{\downarrow} W(r_{kl}) \phi_q(j_l) \left[\sum_p \theta_p \phi_p(j_l) - i_k\right] = 0, \tag{10}$$

where r_{kl} are the residuals defined by $r_{kl} = \sum_p \theta_p \phi_p(j_l) - i_k$, and W is the weighting function associated with ρ, i.e., by definition: $W(x) = (d\rho/dx)/x$. The reason for writing this last expression using these definitions is to give a flavour of the commonly adopted resolution strategy. Ignoring the fact that the residuals, r_{kl}, depend on Θ, eq (10) becomes equivalent to eq (7), condition to replace the original weights $w_{kl}^{\downarrow}$ with new ones, $w_{kl}^{\downarrow} W(r_{kl})$, which incorporate the non-quadratic nature of the ρ-function. Then, solving eq (10) amounts to solving a linear system, which still may be achieved by means of the SVD method. However, since the residuals actually depend on Θ, a single iteration is not sufficient to provide an acceptable solution. The process has to be iterated until convergence, updating the weights at each step in terms of the residuals corresponding to the current estimate of Θ. As shown in [16], this iterative weighted least squares procedure is guaranteed to converge to a solution of eq (10). The initial estimate for Θ is practically chosen as the result of the standard least squares regression.

Estimation of S_0

To ensure that this robust regression does a good job, it is important that the constant S_0 be tuned adequately. A too large value would let the fitting function be spoiled by outliers, whereas a too small value would act as rejecting too many inliers, and again making the resulting fit unreliable. Ideally, S_0 should be a robust scale estimate, i.e., an estimate of the typical error of fit. Thus, finding an optimal value for S_0 would require the fitting function to be known; on the other hand, the regression method needs S_0 to work.

[1] Also notice that proofs of uniqueness are subject to the additional assumption that ρ is strictly convex, a requirement that is not met, e.g., by the Geman-McClure function.

The common trade-off is to use the standard least squares solution as a rough initial guess of the fitting function and then computing S_0 from the corresponding residuals using some robust scale estimator. A good choice is the median of absolute deviations (MAD),

$$\text{MAD} = 1.4826 * \text{median}\{r_1, r_2, \ldots, r_n\},$$

where the normalization factor 1.4826 ensures that the MAD be an unbiased estimate of the residuals' standard deviation when they are normally distributed. Here, our problem is in fact slightly more complex since the residuals are supplied with two indices (k, l) and weighted by $w^{\downarrow}_{kl}$ (as an effect of PV interpolation, see section 2.1). Therefore, we use a straightforward generalization of the MAD estimator based on a weighted median [1]:

$$S_0 = 1.4826 * \text{wmedian}\{r_{kl}; w^{\downarrow}_{kl}\}.$$

C Multiple registration: a Newton gradient descent algorithm

This appendix explains how to optimize the multiple registration criterion of section 4.2. Since it is heavily based on the work presented in [13] and [11], the reader is invited to consider this appendix as a continuation of these papers, especially for the notations.

C.1 Formalization of the criterion

Let $f^T = (r^T, t^T)$ be a rigid transformation represented by is rotation vector and translation (i.e. in the principal chart), and Λ a metric (typically a diagonal matrix with $1/\sigma_r^2$ for the rotation and $1/\sigma_t^2$ for the translation). We consider here the following left invariant distance:

$$\text{dist}^2(f_1, f_2) = \text{dist}^2(f_2^{(-1)} \circ f_1, \text{I}_d) = (f_1^{(-1)} \circ f_2)^T.\Lambda.(f_1^{(-1)} \circ f_2)$$

For the multiple registration, we are looking for $n-1$ rigid transformations $\bar{f}_{(1)} \dots f_{(i)} \dots \bar{f}_{(n-1)}$ from image i to image $i+1$. Let F be the vector of all these unknowns. Using the following recursive function, one can construct any transformation from image i to image j:

$$g_{(i,j)} = \begin{cases} \bar{f}_{(j-1)} \circ g_{(i,j-1)} & \text{if } j > i+1 \\ \bar{f}_{(i)} & \text{if } j = i+1 \\ g_{(j,i)}^{(-1)} & \text{if } j < i \end{cases}$$

We are looking for the set of transformations F that minimizes the sum of square distances from $g_{(i,j)}$ to the measured transformations $\hat{f}_{(i,j)}$. Introducing the error vectors $z_{(i,j)} = \hat{f}_{(i,j)}^{(-1)} \circ g_{(i,j)}$, the criterion is

$$C(F) = \sum_{i,j} \text{dist}^2 \left(g_{(i,j)}, \hat{f}_{(i,j)} \right) = \sum_{i,j} z_{(i,j)}^T.\Lambda.z_{(i,j)}.$$

C.2 The Newton gradient descent

As we are dealing with a least-square criterion, the derivative can be approximated by:

$$\Phi = \left(\frac{\partial C}{\partial F}\right)^T = \sum_{i,j} \left(\frac{\partial z_{(i,j)}}{\partial F}\right)^T.\Lambda.z_{(i,j)} \qquad \text{and} \qquad H = \left(\frac{\partial^2 C}{\partial F^2}\right) \simeq \sum_{i,j} \left(\frac{\partial z_{(i,j)}}{\partial F}\right)^T.\Lambda.\left(\frac{\partial z_{(i,j)}}{\partial F}\right)$$

The Newton gradient descent consists in minimizing at each time-step the second order Taylor expansion of the criterion, which mean that we should update the current value of F by the increment $\delta F = -H^{(-1)}.\Phi$. Since we are not in a vector space but on a Riemannian manifold, we have to correct for the metric distortions and the updating formula becomes for each rigid transformation $\bar{f}_i$:

$$\bar{f}_i \longleftarrow \bar{f}_i \circ \left(J_L(\bar{f}_i)^{(-1)}.\delta \bar{f}_i \right)$$

C.3 Computing the derivatives

We are now left with the computation of the derivative of each error vector:

$$\frac{\partial z_{(i,j)}}{F} = \frac{\partial \left(\hat{f}_{(i,j)}^{(-1)} \circ g_{(i,j)} \right)}{\partial g_{(i,j)}} . \frac{\partial g_{(i,j)}}{\partial F}$$

Using once again the chain rule to derive the function $g_{i,j}$, we obtain the following recursive computation of the derivatives:

$$\frac{\partial g_{(i,j)}}{\partial F} = \begin{cases} \text{if } j > i+1: & \dfrac{\partial(\bar{f}_{(j-1)} \circ g_{(i,j-1)})}{\partial \bar{f}_{(j-1)}} . \dfrac{\partial \bar{f}_{(j-1)}}{\partial F} + \dfrac{\partial(\bar{f}_{(j-1)} \circ g_{(i,j-1)})}{\partial g_{(i,j-1)}} . \dfrac{\partial g_{(i,j-1)}}{\partial F} \\ \text{if } j = i+1: & \dfrac{\partial \bar{f}_{(i)}}{\partial F} \\ \text{if } j < i: & \dfrac{\partial g^{(-1)}_{(j,i)}}{\partial g_{(j,i)}} . \dfrac{\partial g_{(j,i)}}{\partial F} \end{cases}$$

Lastly, let $I_{(i)}$ be the $6 \times 6(n-1)$ matrix where only the the ist 6 by 6 block is non null and equal to the identity. We have:

$$\frac{\partial \bar{f}_{(i)}}{\partial F} = \left[\frac{\partial \bar{f}_{(i)}}{\partial \bar{f}_{(1)}} ; \ldots ; \frac{\partial \bar{f}_{(i)}}{\partial \bar{f}_{(i)}} ; \ldots ; \frac{\partial \bar{f}_{(i)}}{\partial \bar{f}_{(n-1)}} \right] = [0 ; \ldots ; \mathrm{Id} ; \ldots ; 0] = I_{(i)}$$

C.4 Using a robust distance

The simplest modification of a least-square problem to make it robust is to use an M-estimator, i.e. to use $\rho(\text{dist})$ instead of just the squared distance with a suitable function ρ. We use here the parabolic ρ-function up to a given threshold because of the simplicity of its derivatives:

$$\rho(x) = \begin{cases} x^2 & \text{if } x^2 < \chi^2 \\ \chi^2 & \text{otherwise} \end{cases} \qquad \frac{\partial \rho}{\partial (x^2)} = \begin{cases} 1 & \text{if } x^2 < \chi^2 \\ 0 & \text{otherwise} \end{cases} \qquad \frac{\partial^2 \rho}{\partial (x^2)^2} = 0$$

Thus, we only have to consider the error vectors $z_{(i,j)}$ that have a norm less than χ^2 to compute the derivatives.

C.5 Computing the uncertainty

Assuming that all measured transformations $\hat{f}_{(i,j)}$ do have the same type of noise: $\hat{f}_{(i,j)} = g_{(i,j)} \circ (\mathrm{Id}\,,\, \Sigma)$, one can estimate the covariance matrix Σ directly from the residuals. Correcting for the $n-1$ estimated parameters among the $N = n(n-1)$, one have: $\Sigma = \frac{1}{N-n+1} \sum_{i,j} z_{(i,j)} . z^T_{(i,j)}$. In fact, it is often sufficient to assume a diagonal covariance with a the same variance $\sigma_r^2/3$ for each rotation coordinates and the same $\sigma_t^2/3$ for each translation coordinates. Let $\theta_{(i,j)}$ be the norm of the rotation part of $e_{(i,j)}$ and $d_{(i,j)}$ the norm of its translation part:

$$\sigma_r^2 = \frac{1}{N-n+1} \sum_{i,j} \theta^2_{(i,j)} \qquad \text{and} \qquad \sigma_t^2 = \frac{1}{N-n+1} \sum_{i,j} d^2_{(i,j)}$$

Now, the uncertainty on F is given by:

$$\Sigma_{FF} = H^{(-1)} . \left(\sum_{i,j} \frac{\partial \Phi}{\partial z_{(i,j)}} . \Sigma \frac{\partial \Phi}{\partial z_{(i,j)}}^T \right) . H^{(-1)} \qquad \text{with} \qquad \frac{\partial \Phi}{\partial z_{(i,j)}} \simeq \sum_{i,j} \left(\frac{\partial z_{(i,j)}}{\partial F} \right)^T . \Lambda$$

Assuming that the metric Λ is correctly set to correspond to the measured noise ($\Lambda \simeq \Sigma^{(-1)}$), the uncertainty simplifies to $\Sigma_{FF} = H^{(-1)}$. From this covariance matrix on all measurements $\bar{f}_{(i)}$ together, one can now recompute the covariance on any transformation $g_{(i,j)}$ and thus the uncertainty on typical points on the image (the $\sigma^2_{MR/MR}$ and $\sigma^2_{US/US}$ of section 4.2).

Unité de recherche INRIA Sophia Antipolis

2004, route des Lucioles - B.P. 93 - 06902 Sophia Antipolis Cedex (France)

Unité de recherche INRIA Lorraine : Technopôle de Nancy-Brabois - Campus scientifique

615, rue du Jardin Botanique - B.P. 101 - 54602 Villers lès Nancy Cedex (France)

Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)

Unité de recherche INRIA Rhône-Alpes : 655, avenue de l'Europe - 38330 Montbonnot St Martin (France)

Unité de recherche INRIA Rocquencourt : Domaine de Voluceau - Rocquencourt - B.P. 105 - 78153 Le Chesnay Cedex (France)

Éditeur

INRIA - Domaine de Voluceau - Rocquencourt, B.P. 105 - 78153 Le Chesnay Cedex (France)

http://www.inria.fr

ISSN 0249-6399