Automatic Deformable MR-Ultrasound Registration for Image-Guided Neurosurgery

Hassan Rivaz, Sean Jy-Shyang Chen and D. Louis Collins

Abstract—Fusion of tracked ultrasound (US) with MR has many applications in diagnostics and interventions. Unfortunately, the fundamentally different natures of US and MR imaging modalities renders their automatic registration challenging. In this work, we present a novel algorithm for registration of 3D volumetric US and MR using Robust PaTch-based cOrrelation Ratio (RaPTOR). RaPTOR computes local correlation ratio (CR) values on small patches and adds the CR values to form a global cost function. It is therefore invariant to large amounts of spatial intensity inhomogeneity. We also propose a novel outlier suppression technique based on the orientations of the RaPTOR gradients. Our deformation is modeled with free-form cubic B-splines. We analytically derive the derivatives of RaPTOR with respect to the transformation, i.e. the displacement of the B-spline nodes, and optimize RaPTOR using a stochastic gradient descent approach.

RaPTOR is validated on MR and tracked US images of neurosurgery. Deformable registration of the US and MR images acquired respectively pre-operation and post-resection is of significant clinical significance, but challenging due to, among others, the large amount of missing correspondences between the two images. This work is also novel in that it performs automatic registration of this challenging dataset. To validate the results, we manually locate corresponding anatomical landmarks in the US and MR images of tumor resection in brain surgery. Compared to rigid registration based on the tracking system alone, RaPTOR reduces the mean initial mTRE over 13 patients from 5.9 mm to 2.9 mm, and the maximum initial TRE from 17.0 mm to 5.9 mm. Each volumetric registration using RaPTOR takes about 30 sec on a single CPU core.

An important challenge in the field of medical image analysis is the shortage of publicly available dataset, which can both facilitate the advancement of new algorithms to clinical settings and provide a benchmark for comparison. To address this problem, we will make our manually located landmarks available online.

Index Terms—Non-rigid registration, Online database, Validation database, Intra-operative ultrasound, Brain surgery, Image guided neuro-surgery, IGNS

I. INTRODUCTION

The success of the surgical resection of brain tumors depends to a large extent on the complete removal of the tumor. The proximity of many tumors to critical brain structures coupled with a poor visibility of brain tumors in the operating room renders complete removal of the tumor challenging. As a result, intra-operative tracked ultrasound (US) has gained significant momentum in neurosurgery [1], [2], [3], [4]. In image-guided neurosurgery, the preoperative MR images can be first converted to the patient space by selecting a few corresponding landmarks on the MR images and patient's head. The US probe is tracked with a position sensor, which provides the transformation of the US images to the patient space. The US and MR images, both in the patient space, should ideally be in correct alignment. However, the selection of corresponding landmarks in the MR image and patient's head is subject to errors, and hence image-to-patient registration is inaccurate. In addition, the brain tissue can deform as much as 38 mm during neurosurgery [5], and therefore the registration of pre- and intra-operative images is subject to large levels of error. Since brain tumors usually have a high contrast in the pre-operative MR images, deformable registration of these images to post-resection US can significantly reduce the extent and likelihood of residual tumor.

A critical part of an image-based registration technique is the similarity metric, which should be maximized to align the images. Popular similarity metrics include mutual information (MI), correlation ratio (CR), correlation coefficient (CC) and sum of squared differences (SSD). An elegant study of these similarity metrics is provided in [6]. Among them, MI is the most general, which assumes a statistical relationship between image intensities and was proposed for registration of the multi-modal images [7], [8], [9]. Assuming a more restrictive functional relationship $f(i)$ that maps the intensity i in one image to the other, Roche *et al.* [10] proposed the CR for multi-modal registration. If a linear relationship $f(i) = ai + b$ can be assumed, CR simplifies to CC. And finally, if $a = 1$ and $b = 0$, CC reduces to SSD. Since intensity relations in multi-modal images are not linear, CC and SSD are generally not suitable for registration of multi-modal images.

A disadvantage of MI is that it samples the entire image to establish the statistical relationship. This significantly limits the application of MI when this relationship changes spatially, e.g. in the presence of spatial intensity inhomogeneity. Local MI (LMI) [11] addresses this problem by computing multiple MI estimates at different spatial locations and averaging the results to get LMI. The spatial neighborhood needs to be large enough to contain enough samples for reliable MI estimations, which limits the locality of LMI estimates. Furthermore, the computation time of LMI linearly increases with the number of local MI estimates, making it significantly slower than MI. The spatial intensity inhomogeneity in US images is very large due to wave attenuation, and shadowing and enhancement artifacts. Furthermore, the US time-gain compensation (TGC) settings

H. Rivaz is with the Department of Electrical and Computer Engineering and the PERFORM Centre, Concordia University, Montreal, QC, H3G 1M8, Canada

S. J. Chen and D. L. Collins are with the McConnell Brain Imaging Centre, Montreal Neurological Institute, McGill University, 3801 University Street, Montreal, QC, H3A 2B4, Canada.
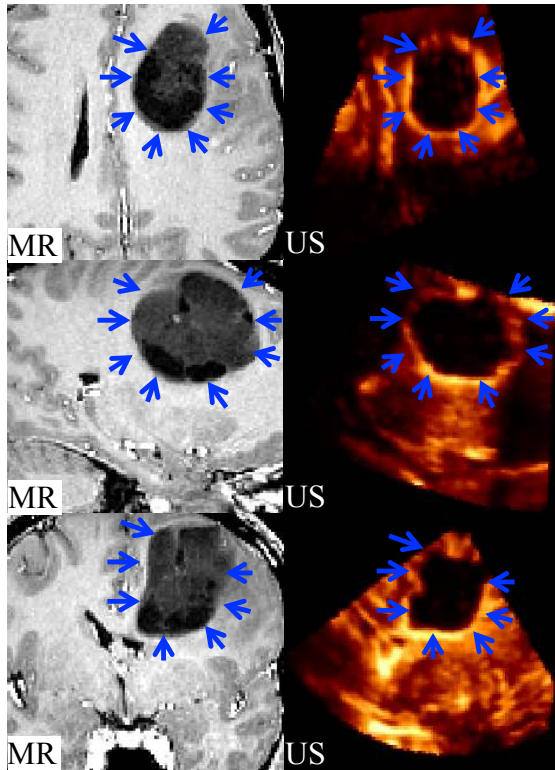
Email: hrivaz@ece.concordia.ca

Fig. 1. The tumor in the pre-operative MR and the cavity in the post-resection US images, pointed to by arrows. First to third row respectively show axial, sagittal and coronal slices.

allow the user to adjust image brightness with depth, which means that the intensity of the US image can be changed arbitrarily. This very large spatial intensity inhomogeneity makes application of LMI challenging due to its requirement for relatively large spatial neighborhoods.

CR assumes the more restrictive functional relationship compared to MI, while being less sensitive to the sample size and less computationally expensive [12], [13]. For regions with equal intensities (iso-sets) in the first image, CR essentially measures the variance in the corresponding intensity values of the second image. Therefore, it measures the dispersion of the intensity relationships from the function $f(x)$, and maximization of CR minimizes this dispersion. We will provide a more detailed overview of CR in Section II-A.

The very high level of spatial bias in US requires highly local estimation of the similarity metric. We therefore propose the RaPTOR algorithm (Robust PaTch-based cOrrelation Ratio), which computes the CR from small patches by binning of the intensities in the first image. A difference between RaPTOR and the traditional CR is this binning process: since the number of bins can be significantly less than the number of iso-sets in the image, binning allows reliable CR computation from small patches. Also, we compute the histogram using a differentiable kernel, which makes the analytic derivative of RaPTOR available for efficient optimization using the stochastic gradient descent approach [14]. An important difference of RaPTOR with LMI emanates from the difference between CR and MI: CR is an order of magnitude (about 16 to 64 times,

depending on the number of bins of the joint histogram) faster than MI [12]. This allows us to compute RaPTOR over *many* small patches and sum the results in a short time.

We automatically register pre-operative MR, acquired a few days before the surgery, to post-resection US, which is the first such study to the best of our knowledge. Here, the tumor in the MR image is replaced with the tumor cavity in the US image (see Figure 1). Image artifacts in the post-resection US image [15] also add to the difficulty of automatic registration. We therefore propose the approach based on the orientation of the gradients of RaPTOR to suppress outliers and such missing correspondences.

Our group has recently published the BITE database [16], which is available online at http://www.bic.mni.mcgill.ca/ BITE and contains three groups of US and MR images. As part of this paper, we will add a fourth group[1] that contains pre-operative MR and post-resection US images along with the corresponding anatomical landmarks. Obtaining real clinical data that enables the medical imaging field validate its algorithms is usually a big challenge, given that many image processing laboratories are not located in hospitals. Providing these data online has two important applications. First, it provides the researchers with an easily accessible database for testing and perfecting their algorithm. Second, it provides a common benchmark for comparing different algorithms.

The contributions of this paper are:

1) Proposing a CR-based measure that *locally* estimates the similarity metric from *small* patches.
2) Deriving the analytic derivative of the similarity metric and performing efficient optimization.
3) Introducing a novel metric for outlier rejection.
4) Registering pre-operative MR to post-resection US for the first time.
5) Locating corresponding anatomical landmarks between pre-operative MR and post-resection US for validation of the registration results, and providing these landmarks online. This is the first such database to the best of our knowledge.

This paper is organized as follows. We first provide an overview of related work, followed by a more in-depth description of CR. We then provide RaPTOR details and derive its gradient to perform efficient optimization. The outlier suppression based on RaPTOR gradient orientations is provided next, followed by the description of the patient data experiments and validation results.

A. Related Previous Work

Previous work has tackled registration of US and other imaging modalities. Roche *et al.* [17] performed a parametric polynomial fit of US intensity as a function of MR and its gradient to construct the functional CR similarity metric. We describe this approach in Section II-A. Arbel *et al.* [18] and Mercier *et al.* [3] segmented the MR image and assigned different intensity transformations to different MR regions to

---

[1] after acceptance of this paper

generate pseudo-US images. They then registered the pseudo-US to the real US using the ANIMAL registration technique [19]. Using a similar approach, Kuklisova-Murgasova *et al.* [20] recently proposed to register MR to a segmented atlas, generate a US-like volume from the MR volume, and finally register it with the US volume using a uni-modal block-matching technique. In contrast, RaPTOR does not require registration and segmentation of the MRI and the simulation of the pseudo-US image, and directly registers MR to US. Wein *et al.* [21] simulated pseudo-US images from CT by taking into account the physics of these modalities, and registered the pseudo-US using a CR based similarity metric. Mellor and Brady [22] and Zhang et al. [23] used MI between phase information, which is local and invariant to global intensity inhomogeneities. Ji *et al.* [24] used normalized MI (NMI) of US and MR by considering voxels throughout the volumes and hence computing a global estimate of NMI. Since US images suffer from spatial intensity inhomogeneity, estimation of global similarity measures can be unreliable. De Nigris *et al.* [4] calculated the orientation of image gradients, and used the level of their alignment as a similarity metric [25], [26]. Heinrich *et al.* [27] used self-similarity context [28], which compares neighboring patches using SSD to estimate a local description of the image features. These approaches are similar to RaPTOR in that they compute similarity metrics *locally*, and are therefore robust to spatial intensity inhomogeneities. Wein *et al.* [29] and Fuerst *et al.* [30] used correlation of US intensities and the intensity and gradient of MR similar to [17]. Ferrante and Paragios [31] registered 2D US images to 3D MR volume using discrete optimization within a Markov Random Field framework. Sun et al. [32] proposes to optimize local self-similarity features of [28] using efficient duality-based convex optimization. Feature-based registration methods are an alternative to the aforementioned intensity-based methods, and provide another avenue for estimating *local* similarity measures. Penney *et al.* [33] and Reinertsen *et al.* [34], [35] used blood vessel features to register US and MR, while Coupé et al. [36] used the hyperechogenic cerebral falx and the sulci structures in the brain US and MR images. In [37], we proposed to weight a graph-based α-MI measure based on self-similarity in the MR image, and recently, we proposed to condition MI on self-similarity [38]. We will discuss these methods and compare them to RaPTOR in terms of computational cost and resistance to outliers.

Registration of images with missing correspondences is another area of related work. Chitphakdithai *et al.* [39], [40], [41] proposed a joint registration and segmentation approach for detection of the missing data, which is optimized in an expectation maximization (EM) framework. Risholm *et al.* [42] also proposed a method that alternates between registration and segmentation using level set in the region with high residuals. Daga *et al.* [43] utilized the least trimmed squares approach along with a block-matching strategy to obtain robust deformation field from interventional MR images. Finally, Kamen *et al.* [44] proposed to optimize a functional with SSD as the similarity metric for corresponding regions. For non-corresponding regions, the similarity metric of their functional will try to match the intensity distribution function of *only one*

of the images to a priori intensity distribution, and therefore no correspondence is enforced.

II. METHODS

Let the fixed and moving images be respectively $I_f(\boldsymbol{x})$, $I_m(\boldsymbol{x})\colon \boldsymbol{\Omega} \subset \mathbb{R}^d \to \mathbb{R}$, where d is the dimension of the images: $d = 3$ for 3D volumetric data, and let $\mathbf{T}$ denote the unknown transformation that aligns I_m to I_f. The registration problem can be formulated as minimizing a cost function C:

$$C = D(I_f(\boldsymbol{x}), I_m(\mathbf{T}(\phi, \boldsymbol{x})) + \frac{\omega_R}{2} trace(\nabla \mathbf{T}^\top \nabla \mathbf{T}) \quad (1)$$

where D is a dissimilarity metric, ω_R is a regularization penalty weight (divided by 2 to make the gradient of C more compact), ∇ is the gradient operator, and $\mathbf{T}(\phi, \boldsymbol{x})$ is the transformation modeled by ϕ. We choose a free-form transformation parameterized by the location of cubic B-spline nodes. Assume $d = 3$, $\phi = [\varphi_x, \varphi_y, \varphi_z]$ be the location of all the nodes, and n_x, n_y, n_z be the node spacing. The transformation is [45]:

$$\mathbf{T}(\phi, \boldsymbol{x}) = \sum_{a=0}^{3} \sum_{b=0}^{3} \sum_{c=0}^{3} B_a(\alpha) B_b(\beta) B_c(\gamma) \phi_{i+a, j+b, k+c} \quad (2)$$

where $i = \lfloor x/n_x \rfloor - 1$, $j = \lfloor y/n_y \rfloor - 1$, $k = \lfloor z/n_z \rfloor - 1$, and $\alpha = x/n_x - i - 1$, $\beta = y/n_y - j - 1$, $\gamma = z/n_z - k - 1$, and B represents B-spline basis functions (see [45] for more details). The dissimilarity metric D is commonly based on MI or CR in multi-modal registration; we use CR in this work.

A. Correlation Ratio (CR)

We start by a brief overview of the well-known work of Roche *et al.* [10], which has inspired this work. In the next section, we base the derivation of RaPTOR on this overview. Let $\mathrm{E}\langle Y|X\rangle$ be a mapping that estimates the intensity of either I_f or I_m (denoted by Y) as a function of the other image (denoted by X). Note that here, we use X and Y because either I_f or I_m can play the role of X.

$$\mathrm{Var}[Y - \mathrm{E}\langle Y|X\rangle] = \mathrm{Var}[Y] - \mathrm{Var}[\mathrm{E}\langle Y|X\rangle] \quad (3)$$

where Var and E respectively denote variance and expected value. $\mathrm{Var}[\mathrm{E}\langle Y|X\rangle]$ is the part of $\mathrm{Var}[Y]$ that is predicted by X, and $\mathrm{Var}[Y - \mathrm{E}\langle Y|X\rangle]$ is the part of Y that is functionally independent of X. A problem with minimizing $\mathrm{Var}[Y - \mathrm{E}\langle Y|X\rangle]$ is that it can try to minimize $\mathrm{Var}[Y]$ and hence the overlap between the two images. Therefore, [10] defined CR as:

$$\eta(Y \mid X) = \frac{\mathrm{Var}[\mathrm{E}\langle Y|X\rangle]}{\mathrm{Var}[Y]} = 1 - \frac{\mathrm{Var}[Y - \mathrm{E}\langle Y|X\rangle]}{\mathrm{Var}[Y]}. \quad (4)$$

Here Y is predicted from X; note that $\eta(Y \mid X) \neq \eta(X \mid Y)$, i.e. CR is asymmetric. Choosing I_f or I_m to serve as the model image X is important and depends on the imaging modalities [10]. η varies between 0 and 1: values close to 0 mean no functional relationship, while values close to 1 translate to perfect functional relationship.

Previous work has used two approaches for finding $\mathrm{E}\langle Y|X\rangle$: **(1)** A non-parametric mapping is obtained in [10] as
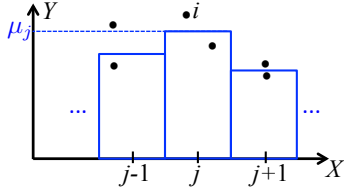
Fig. 2. CR computation by binning the X values. Each sample i contributes to two bins according to its distance from the bin centers. Maximizing CR is equivalent to minimizing the variance of samples within each bin.

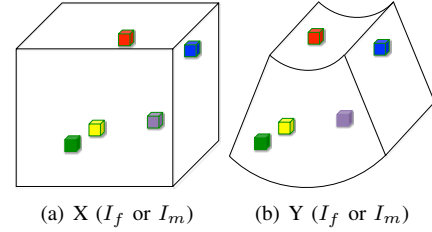(a) X (I_f or I_m) (b) Y (I_f or I_m)

Fig. 3. RaPTOR is averaged over N_p patches. At each iteration, these patches are selected randomly throughout the volumes. We use $N_p = 1000$ in our experiments; in this figure $N_p = 5$.

following. Assuming X is discrete sampled, the iso-sets of X are $\mathbf{\Omega}_i = \{\boldsymbol{x} \in \mathbf{\Omega},\ X(\boldsymbol{x}) = i\}$. The right hand side of Eq. 4 is used to calculate $1 - \eta$ by computing the variance of Y over these iso-sets. [10] then use Powell's optimization method, a robust technique which does not require computation of derivatives, to estimate the rigid transformation that aligns the two volumes, and **(2)** In [17], a polynomial parametric fitting of the form

$$\mathrm{E}\langle Y|X\rangle = \sum_{p+q \leq d} c_{pq} X^p \|\nabla X\|^q = c_{00} + c_{01}\|\nabla X\| + \cdots + c_{d0}X^d$$

is proposed, which assumes Y can be written as a polynomial function of X and the magnitude of its gradient $\|\nabla X\|$. Here, d is the degree of the polynomial and c_{pq} are the unknowns. d is set to 3, and the function is efficiently optimized for $(d+1)(d+2)/2 = 10$ unknowns.

B. RaPTOR

Our approach is similar to that of [10] in that we also perform nonparametric estimation, with the difference that we compute CR locally to achieve resistance to the large spatial intensity inhomogeneity in the US images. To this end, we perform binning of the X values instead of computing $\mathrm{E}\langle Y|X\rangle$ for iso-sets of X (see Figure 2). This gives us two advantages: First, we can use few bins which allows CR computation from few samples in small patches. Second, it allows histogram estimation using Parzen windowing, which is robust to quantization error and is smooth and differentiable. We use this property in Section II-C by deriving the derivative of the cost function analytically and performing efficient gradient-based optimization. In our histogram, each sample i contributes to two closest bins $j-1$ and j linearly, according to its distance from the bin centers. We found this linear kernel to provide a good compromise between the running time and accuracy and robustness. After some manipulation of Eq. 4, we have (see also Figure 2)

$$1 - \eta(Y \mid X) = \frac{1}{N\sigma^2}\left(\sum_{i=1}^{N} y_i^2 - \sum_{j=1}^{N_b} N_j \mu_j^2\right) \quad (5)$$

$$\mu_j = \frac{\sum_{i=1}^{N} \lambda_{i,j} y_{i,j}}{N_j}, \quad N_j = \sum_i \lambda_{i,j} \quad (6)$$

where $\lambda_{i,j}$ is the contribution of sample i to bin j, $N_j = \sum_i \lambda_{i,j}$ is the total weighted numbers of samples in bin j, N is the total number of samples, N_b is the number of bins and

$\sigma^2 = \mathrm{Var}[Y]$. The first term in the bracket, $\sum y_i^2$, is related to the variance of the Y, and the second term, $\sum N_j \mu_j^2$, is related to the variance of the bin mean values. We now average this parameter over N_p patches $\mathbf{\Omega}_i$ to estimate RaPTOR (see also Figure 3):

$$\mathrm{RaPTOR}(X, Y) = D(Y, X) = \frac{1}{N_p}\sum_{i=1}^{N_p}\left(1 - \eta(Y \mid X; \mathbf{\Omega}_i)\right) \quad (7)$$

The N_p patches can be chosen on a lattice to cover the entire volume or even have some overlap; we elaborate the patch selection process in Section II-C. RaPTOR is a *dissimilarity* metric between Y and X that varies between 0 and 1: for aligned images with functional intensity relationships, its value is close to 0.

We make two important notes here: (1) Larger $\mathbf{\Omega}_i$ patches allow more accurate estimates of η, but they also make the estimate less local. We will show in Section III-A1 that CR allows reliable computation of the similarity metric using small patches $\mathbf{\Omega}_i$, especially compared to MI. Computation of the CR using small patches make the similarity metric robust to spatial intensity inhomogeneities. (2) Larger N_p values make the RaPTOR estimate more accurate and robust at the expense of higher computational complexity. Since individual η computations are very fast, we can set N_p to the very large numbers (in the order of 1000).

We now analytically derive the derivative of RaPTOR to allow fast gradient-based optimization of the cost function.

C. Optimization

Analytically estimating the derivative of RaPTOR allows utilization of fast gradient descent based optimization techniques. To optimize Eq. 1, the derivative of the cost function with respect to ϕ can be computed using the chain rule:

$$\nabla_\phi D = \frac{\partial D}{\partial \phi} = \frac{\partial \mathbf{T}}{\partial \phi} \cdot \frac{\partial I_m}{\partial \mathbf{T}} \cdot \frac{\partial D}{\partial I_m} \quad (8)$$

where $\frac{\partial \mathbf{T}}{\partial \phi}$ is the transformation Jacobian, $\frac{\partial I_m}{\partial \mathbf{T}}$ is the image gradient, and $\frac{\partial D}{\partial I_m}$ is the derivative with respect to the intensity of all pixels. It is instructive to note the sizes of these matrices. Let d be the dimension of images (e.g. $d = 3$ for volumetric images), p the number of pixels in the image (e.g. $p = 64$ for an image of size 8×8), and n the number of nodes in the B-spline transformation model. $\frac{\partial \mathbf{T}}{\partial \phi}$ is of size $d \cdot n \times d \cdot p$, $\frac{\partial I_m}{\partial \mathbf{T}}$

is of size $d \cdot p \times p$, and $\frac{\partial D}{\partial I_m}$ is of size $p \times 1$. Therefore, $\nabla_\phi D$ is of size $d \cdot p \times 1$, i.e. the derivative of D with respect to the position of each node in each dimension. To compute $\frac{\partial D}{\partial I_m}$, we have to differentiate Eq. 7 with respect to the intensity of the pixels of the moving image. Since either I_f or I_m can be the model X as we discussed, we compute the derivative of Eq. 7 with respect to both X and Y. Starting from Eq. 5, we have

$$\frac{\partial (1-\eta)}{\partial y_i} = \frac{-\partial \sigma^2/\partial y_i}{N\sigma^4} \left(\sum_{i=1}^{N} y_i^2 - \sum_{j=1}^{N_b} N_j \mu_j^2 \right) + \frac{1}{N\sigma^2} \partial \left(\sum_{i=1}^{N} y_i^2 - \sum_{j=1}^{N_b} N_j \mu_j^2 \right) /\partial y_i$$

(9)

After performing the derivatives and some manipulations, this equation leads to:

$$\frac{\partial (1-\eta)}{\partial y_i} = \frac{2}{N\sigma^2} (y_i - \lambda_{i,j-1}\mu_{j-1} - \lambda_{i,j}\mu_j - \frac{1}{(N-1)\sigma^2}(y_i - \mu) \left(\sum_{i=1}^{N} y_i^2 - \sum_{j=1}^{N_b} N_j \mu_j^2 \right))$$

(10)

where μ is the average value of Y, and μ_j is the average value of Y over bin j (see Figure 2 and Eq. 5). We use Eqs. 7, 8 and 10 to analytically calculate the derivative of RaPTOR with respect to Y.

Since RaPTOR is an asymmetric dissimilarity metric, i.e. $\text{RaPTOR}(X,Y) \neq \text{RaPTOR}(Y,X)$, derivatives with respect to X and Y are also different. The derivative of Eq. 5 with respect to X can be calculated similar to that of Y. Starting from Eq. 5, we have

$$\frac{\partial (1-\eta)}{\partial x_i} = \frac{1}{N\sigma^2} \cdot \partial \left(-\sum_{j=1}^{N_b} N_j \mu_j^2 \right) /\partial x_i,$$

(11)

and from Eq. 6, we have

$$\frac{\partial \mu_j}{\partial x_i} = \frac{y_i - \mu_j}{N_j} \cdot \frac{\partial \lambda_{i,j}}{\partial x_i}, \quad \frac{\partial N_j}{\partial x_i} = \frac{\partial \lambda_{i,j}}{\partial x_i}.$$

(12)

We use Eqs. 7, 8, 11 and 12 to analytically calculate the derivative of RaPTOR with respect to X.

The gradient of D in the chain rule of Eq. 8 is estimated with respect to I_m. Therefore, one should use Eq. 10 if I_m is set to Y, or Eq. 12 if I_m is set to X. Gradient estimation is very fast using the analytic functions of the derivatives: for a patch of size 7^3 and $N_b = 32$ bins, the derivative computations of Eq. 10 with respect to all 343 intensity values takes 3×10^{-5} sec in our implementation using a 3.6 GHz processor. With $N_p = 1000$, gradient of RaPTOR in Eq. 7 is estimated in only 0.03 sec. Estimating the gradient of the dissimilarity metric is one of the most computationally expensive steps of automatic image registration.

1) Stochastic Gradient Descent Optimization: RaPTOR is computed over small cubic patches, which can cover the entire image or can even have overlap. Such densely selected boxes, however, significantly increases the computation time. Instead, RaPTOR is computed over N_p randomly selected patches and a method similar to the stochastic gradient descent optimization of [14], [46] is used to minimize it. In Figure 3, we schematically illustrate our randomized patch selection.

Our iterative stochastic optimization implementation is as follows. In every iteration, we randomly select N_p patches and compute the derivative of our cost function with respect to the B-spline nodes according to Eq. 8. Note that these N_p patches are selected randomly throughout the volumes and are not necessarily centered on the B-spline nodes. The update equation is $\phi_{t+1} = \phi_t + a_t \nabla_\phi D$. The step size is a decaying function of the iteration number t:

$$a_t = a/(A+t)^\tau,$$

(13)

with $a > 0, A \geq 0$ and $0 < \tau \leq 1$ user-defined constants [14]. The recommended values for these parameters are provided in [14]: A should be around 0.1 of the maximum number of iterations or less and τ should be more than 0.6. The value of a is user-defined and is critical as it determines the step-size [14]. If a is too small more iterations are required and it is also more likely that the optimization gets trapped in a local minima. On the other hand, the registration can diverge if a is too large. Fortunately, only the order of magnitude of a needs to be tuned, and the final registration result varies negligibly with a. We set it to values between 1 and 10^4 by multiplying it by 10 each time and evaluated the deformation at each iteration. After we found its order of magnitude, we varied it by smaller steps and finally set it to 500.

D. Outlier Suppression

Common approaches to robust estimation, such as M-estimators [47], [48], [49] and least trimmed squares [50] need to first detect outliers by inspecting the value of the similarity function they produce. For RaPTOR, as an example, one can label outliers as patches with very low CR values, much smaller than the maximum value of 1. For US-MR registration, we noticed that treating patches with low CR values as outliers does not significantly improve the registration results. We have two reason to explain this: First, assuming that the US and MR images are aligned properly, the quality of the functional relationship between the images varies significantly even for inlier data depending on tissue type and image gradient and depth in the US image. Second, the value of the CR varies at different rates with the level of misalignment, i.e. the gradient of the CR with respect to displacement depends on many image properties such as the depth of the US image, distance from focal points and the acoustic property of the tissue. Therefore, even for a constant displacement for all patches, the CR value is significantly different. The fact that brain deformation varies spatially by a considerable amount adds to the CR variability. Therefore, instead of relying on a single CR value, we propose to use multiple measurements based on the orientation of the CR gradients as follows.
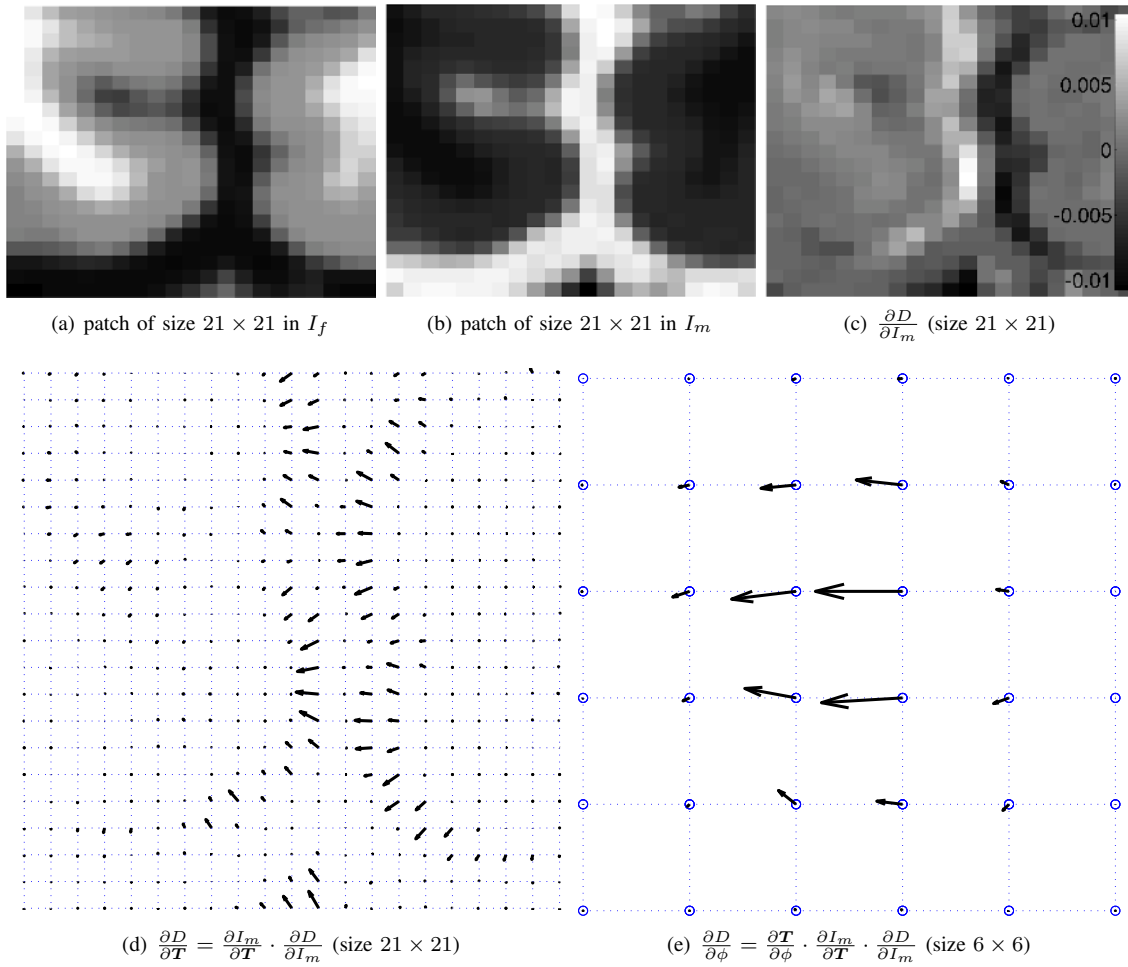
(a) patch of size 21×21 in I_f (b) patch of size 21×21 in I_m (c) $\frac{\partial D}{\partial I_m}$ (size 21×21)

(d) $\frac{\partial D}{\partial \boldsymbol{T}} = \frac{\partial I_m}{\partial \boldsymbol{T}} \cdot \frac{\partial D}{\partial I_m}$ (size 21×21)   (e) $\frac{\partial D}{\partial \phi} = \frac{\partial \boldsymbol{T}}{\partial \phi} \cdot \frac{\partial I_m}{\partial \boldsymbol{T}} \cdot \frac{\partial D}{\partial I_m}$ (size 6×6)

Fig. 4. Computing the derivatives of D (the dissimilarity metric) using the chain rule. I_f and I_m are respectively T1 and T2 images from the BrainWeb. There is a 1 pixel translation in the x direction between I_f and I_m. (c) shows the derivative of D with respect to intensities of I_m. The derivatives of D in the x and y directions at the pixel and node levels are respectively shown in (d) and (e).

Our hypothesis here is that in a small patch, the descent gradients of the CR at different pixels should point in a similar direction. We illustrate this using an example: Figure 4 shows small 21×21 pixel patches of T1 and T2 images from BrainWeb [51], which are misaligned in the x direction by 1 pixel. In the first step, we use Eq. 10 to compute the derivative of the CR with respect to the intensity of each pixel (see part (c)). If we multiply this term by image gradients in the x and y directions, we get (d), which shows the descent direction for each pixel. Note that most pixels are moving to the left, which is the ground truth transformation. Finally, multiplying the result of (d) by the transformation Jacobian $\frac{\partial \boldsymbol{T}}{\partial \Delta \phi}$, we get (e) which predicts the descent direction for each B-spline node. For this illustrative example, we set the spacing between two adjacent nodes to 7 pixels. Since the images are of size 21×21, the number of nodes for the cubic B-spline is $(21/7 + 3) \times (21/7 + 3) = 6 \times 6$ (see part (e)). Again, note that most nodes are moving to the left in this example. For small patches, the deformation of the patch is small and therefore the image derivatives inside the patch must have some agreement in terms of the direction of motion of the patch.

Formally, let the vector $\frac{\partial D}{\partial \boldsymbol{T}}$ be $\frac{\partial I_m}{\partial \boldsymbol{T}} \cdot \frac{\partial D}{\partial I_m}$, i.e. the descent direction for each pixel in the patch (see also Figure 4 (d)). To keep or discard a patch, we compute the unitless metric r:

$$r = \min \left\{ \frac{\mathrm{Var}(\frac{\partial D}{\partial T_x})}{\left\langle \frac{\partial D}{\partial T_x} \right\rangle^2}, \frac{\mathrm{Var}(\frac{\partial D}{\partial T_y})}{\left\langle \frac{\partial D}{\partial T_y} \right\rangle^2}, \frac{\mathrm{Var}(\frac{\partial D}{\partial T_z})}{\left\langle \frac{\partial D}{\partial T_z} \right\rangle^2} \right\} \quad (14)$$

where T_x, T_y and T_z are x, y, z components of the $\boldsymbol{T}$, $\langle \cdot \rangle$ denotes mean and Var is the variance. The numerator shows the variance in the orientation of the gradients of different voxels. The denominators are small at uniform regions with low gradients, and are relatively large at textured regions with high gradients. The denominators are very large at textured regions where the descent directions of individual pixels agree, resulting in large average values. A small ratio for each of the three values shows an agreement in terms of the orientation of the translation of the patch in the three x y z directions. Picking the minimum value means the patch voxels agree on the translation orientation in at least one of the three directions, and therefore solves most issues associated with the so called aperture problem. We illustrate in Figure 5 two patches that do not match, i.e. an example for outlier data. There is no
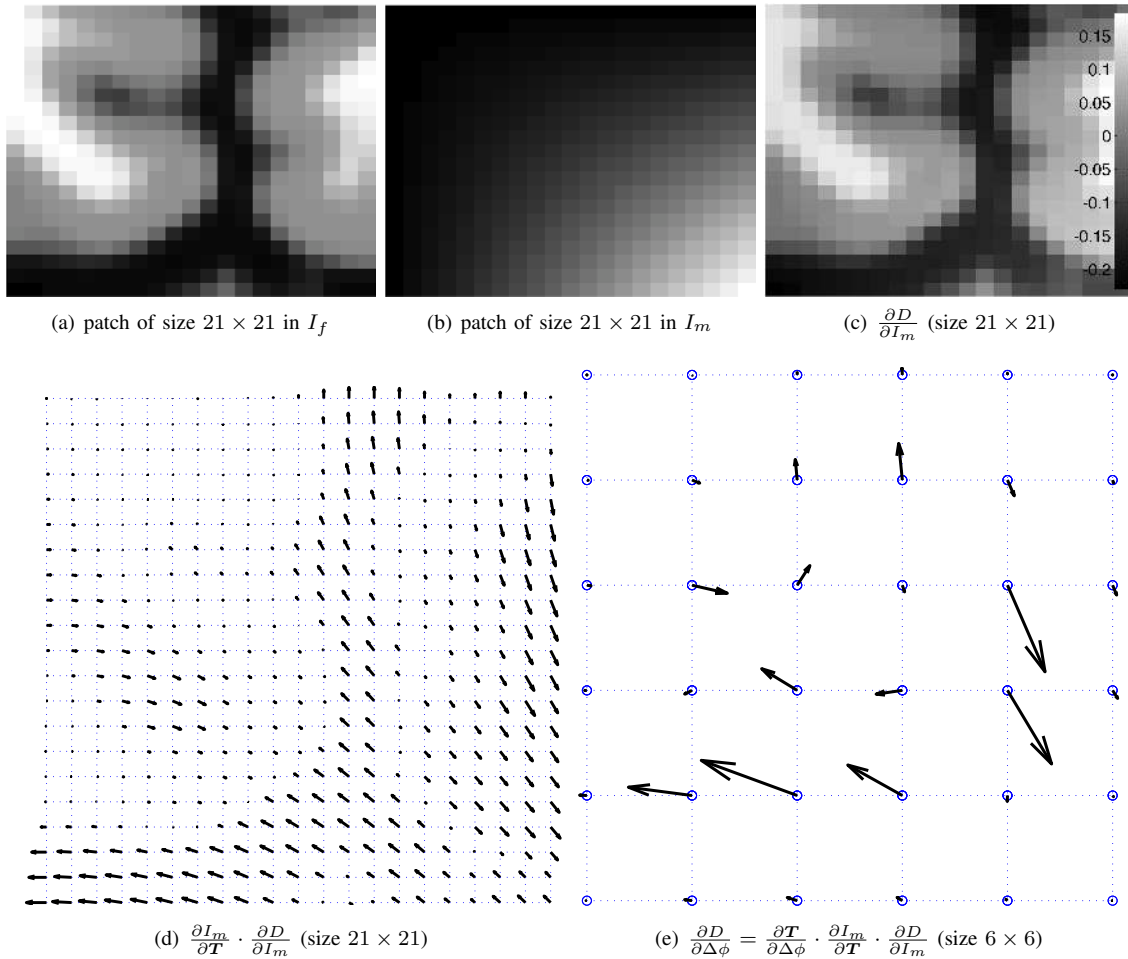
(a) patch of size 21×21 in I_f

(b) patch of size 21×21 in I_m

(c) $\frac{\partial D}{\partial I_m}$ (size 21×21)

(d) $\frac{\partial I_m}{\partial \boldsymbol{T}} \cdot \frac{\partial D}{\partial I_m}$ (size 21×21)

(e) $\frac{\partial D}{\partial \Delta \phi} = \frac{\partial \boldsymbol{T}}{\partial \Delta \phi} \cdot \frac{\partial I_m}{\partial \boldsymbol{T}} \cdot \frac{\partial D}{\partial I_m}$ (size 6×6)

Fig. 5. For two patches that do not correspond, the derivative orientations are in different orientations. The image patch in (b) has no correspondence with (a). Compare the results with Figure 4. Please see the text for more details.

agreement among the patch pixels in the descent direction of the translation; in part (d) the arrows for different pixels span almost $360°$. We discard patches that produce an r value of more than a threshold T. We set the T value in Section III-B2.

III. RESULTS

RaPTOR estimates the similarity measure locally and is robust against outlier data. Therefore, it is suitable for registration of post-resection US and pre-operative MR. In this section, we present results of deformable registration of simulated images with very high spatial bias and outliers. We compare the results to LMI, where entropies are estimated using Parzen window histogram estimation. We finally present the patient data, our landmark selection procedure and the results of non-rigid registration.

A. Simulation Results

1) Number of Samples: In our first experiment, we show the effect of number of samples on CR and MI computation. Roche *et al.* [12] showed, using MR, CT (computed tomography) and PET (positron emission tomography) images of the brain, that CR gives significantly smaller TRE values

compared to MI with few samples. In an experiment similar to [52], we show that CR, in fact, needs fewer samples to converge. We generate two signals I_f and I_m which contain *independent* white noise with equal variances as shown in Figure 6(a). We generate different signals with lengths between $7^3 = 343$ to 10^5, with 100 different instances at each length, and compute the average MI and CR values as shown in (b). This plot shows that with 32 and 64 bins, CR converges faster to the 0 ground truth value. To see the reason, note that [7], [8]:

$$\mathrm{MI}(I_f, I_m) = \mathrm{H}(I_f) + \mathrm{H}(I_m) - \mathrm{H}(I_f, I_m) \qquad (15)$$

We plot the two terms $\mathrm{H}(I_f) + \mathrm{H}(I_m)$ and $\mathrm{H}(I_f)$ and $\mathrm{H}(I_m)$ in Figure 6(c). We see that the joint entropy $\mathrm{H}(I_f, I_m)$, which is performed at higher dimension compared to the marginal entropies $\mathrm{H}(I_f)$ and $\mathrm{H}(I_m)$, needs many more samples for convergence. This is in fact related to the so called "curse of dimensionality". From (b) and (c), we also see that the convergence is much slower with 64 bins compared to 32 bins. These plots show the disadvantage of MI compared to CR in requiring more samples. The advantage of MI is that is does not assume a restrictive functional relationship, which is not guaranteed to exist, and therefore in general

(a) I_f and I_m random signals   (b) MI and CR, normalized to the same maximum value   (c) Individual and joint entropies
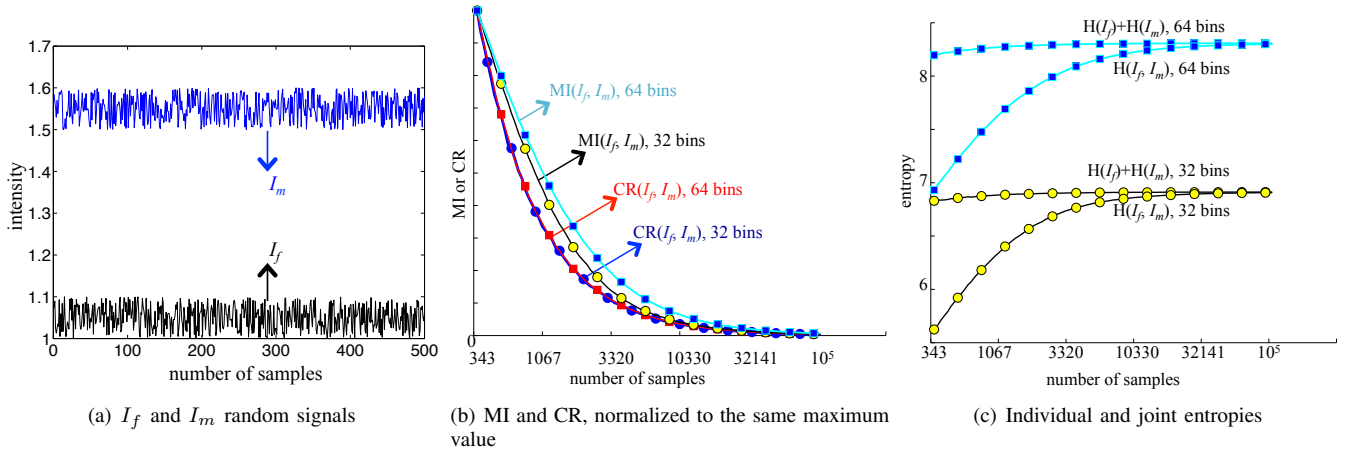
Fig. 6. The CR, MI and entropy values for randomly generated signals with independent noise. The MI and CR between the signals in (a) must be zero since intensities are independent. Two different number of bins of 32 and 64 are used, marked respectively by circles and asterisks. 343 samples corresponds to the number of voxels in a patch of size 7^3. In (b) the CR values for 32 and 64 bins are relatively close. The entropy plot in (c) shows that the joint entropy $H(I_f, I_m)$ needs more samples, compared to the marginal entropies $H(I_f)$ and $H(I_m)$, to converge to its final value. It also shows that by increasing the number of bins more samples are required for $H(I_f, I_m)$ to converge.

works better than CR. For the case of US-MR registration, we found that the disadvantage of the restrictive functional relationship assumed in CR is outweighed by its locality. Finally, note that although these plots show slower convergence of MI for higher bin number, increasing the number of bins reduces the entropy estimation bias. Assuming $h(x)$ denotes the continuous differential entropy and $H(x^\Delta)$ denotes the entropy measured with histogramming of bin-size Δ, one can show [53]: $\lim_{\Delta \to 0} \left(H(x^\Delta) + \log(\Delta) \right) = h(x)$. This equation shows that the bias increases with the increasing the bin size, a disadvantage with having few bins.

2) Outlier rejection: In the final simulation experiment, we test the outlier rejection technique on simulated images. Two simple 2D images I_f and I_m are shown in Figure 7 (a) and (b). Except for the parts inside the red (for I_f) and cyan (for I_m) rectangles, the images consist of uncorrelated multiplicative strong noise. The parts inside the rectangles are identical, but are displaced in the vertical direction by one pixel. The goal is to find outliers. We simulate 500 different I_f and I_m images similar to the ones shown, with different instances of noise. We then compute the RaPTOR metric for all patches of size 5^2 inside the images with 16 bins. Part (c) shows the average CR values over the entire 500 pairs of images. Although the pixels inside the rectangle are not outliers, they have a low average CR. Also, the outlier region outside the box has generally large CR values, which should be the case for inlier data. Therefore, an outlier rejection technique based on the CR values will reject more of the inlier region and accept the outlier region. Part (d) shows that the r value (calculated according to Eq. 14), however, has significantly lower values inside the rectangle and therefore can be used for outlier suppression.

B. Patient data

The clinical data from 13 patients were acquired at the Montreal Neurological Institute, as part of a study approved by the Montreal Neurological Institute and Hospital Review Ethics Board. The pre-operative MR images are gadolinium-enhanced T1 weighted and were acquired approximately 2 weeks before the surgery. In the operating room, a standard patient registration procedure using 9 facial landmarks was performed to allow navigation based on the preoperative MR images. The post-resection US images are obtained using an HDI 5000 (Philips, Bothell, WA) with a P7-4 MHz phased array transducer.

The 2D US images are acquired at a depth of 80 mm and are of size 390×520 pixels. Each pixel is 0.2061 mm in each direction. Reflective spheres rigidly fixed to a TA003 tracker (Traxtal Technologies Inc., Toronto, Ontario, Canada) and a Polaris infrared optical system (Northern Digital, Waterloo, Ontario, Canada) were used to provide the three locations and angles (6 DOF) of each image, as shown in Figure 8. Image acquisition, tracking, and synchronization of the tracking and imaging data are performed using the Intra-operative Brain Imaging System (IBIS) program, developed in our group by Simon Drouin and Anka Kochanowska [54]. US images are also captured using IBIS through a Pinnacle (Pinnacle Systems, Mountain View, CA) PCTVTM frame-grabbing card. Figure 9 (a) shows the 2D ultrasound images acquired after the resection, with each image positioned according to its location from the tracking system. More than 200 slices were acquired from each patient while the ultrasound probe was slowly moved in the out-of-plane direction of the US images.

For each patient, we reconstruct two US volumes using the pixel based method of [55] with two isotropic voxel sizes of 0.5 mm and 1 mm. The slow sweep guarantees that there are no holes in the 3D reconstructed data. Figure 9 (b) shows the reconstruction with the 1 mm voxel size. We use the volume with the small 0.5 mm voxel size for accurate landmark selection, which is elaborated in the next section. The volume with the larger voxel size of 1 mm is utilized for automatic registration. This large voxel size minimizes the effect of speckles, because, for every voxel, multiple measurements
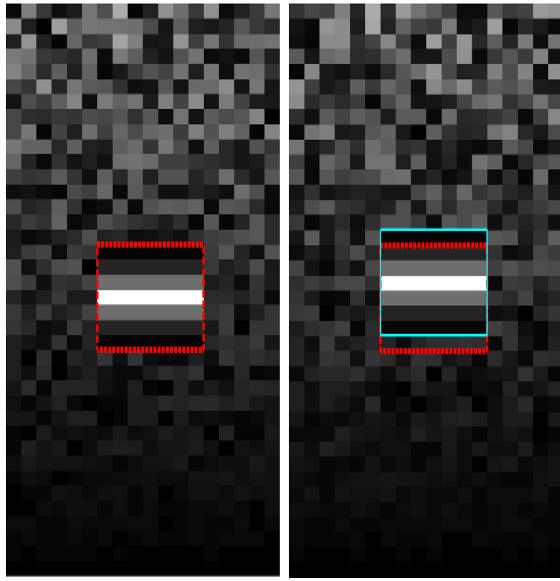
(a) I_f (b) I_m
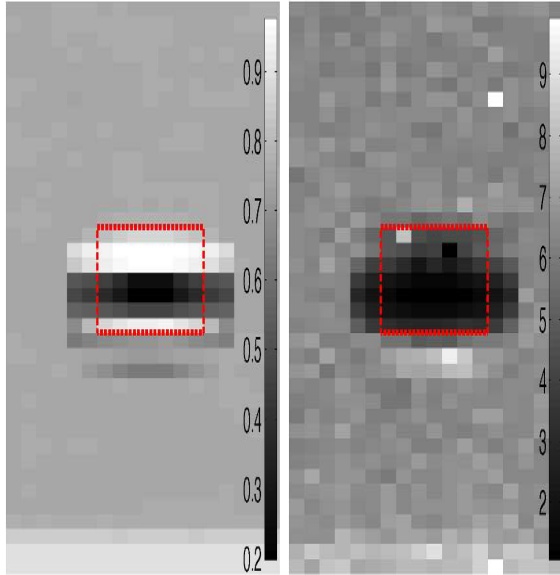
(c) η (d) std of gradient orientations

Fig. 7. Average CR values versus the orientation of gradient derivatives in simulated images. In (a) and (b), the regions respectively inside the red hashed and solid cyan are inliers, while regions outside the boxes are randomly generated. The inlier box has moved by 1 pixel vertically as can be seen in (b). The average CR values over 500 different instances of the I_f and I_m are shown in (c). For most parts of the region in the inlier region, the CR is very low (in dark color); much lower than the outlier region. In (d) the std of the gradient orientation is very low inside the inlier region and very high outside of it. We therefore use this value to find outliers.

from different images are available [56]. We therefore do not perform any post-processing, such as median filtering or Gaussian blurring of the US volumes.

1) Landmark Selection: To validate the results, an expert (the second author) has selected corresponding anatomical landmarks in each of the 13 patients. The landmarks are anatomical structure visible in ultrasound such as sulci bifurcations, vessels, choroid plexus and septa. These landmarks

Fig. 8. Acquisition of tracked 2D US slices after opening the dura. The tracker is rigidly attached to the US probe and is tracked with a Polaris camera in 3D while images are acquired.
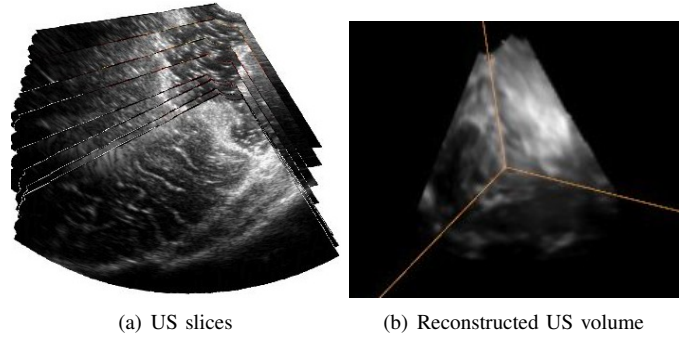
(a) US slices   (b) Reconstructed US volume

Fig. 9. Tracked US slices are acquired during the operation, which we reconstruct to 3D. More than 200 slices are acquired from each patient; we are only showing few of them here. In (b), three orthogonal cross sections of the US volume are shown.

are used to calculate mean target registration errors (mTRE) to assess the accuracy of the registration [57]. The mTRE of n corresponding marks at locations $\boldsymbol{x}$ and $\boldsymbol{x}'$ in the two images is calculated according to

$$\text{mTRE} = \frac{1}{n} \sum_{i}^{n} \|\boldsymbol{T}(\boldsymbol{x}_i) - \boldsymbol{x}_i'\| \qquad (16)$$

where $\boldsymbol{T}$ is the transformation and $\|\cdot\|$ is the norm 2 operator. The mean initial mTRE over 13 patients is 5.9 mm (see Table II for detailed TRE values for each patient).

To measure the accuracy of the manual landmark selection, the landmarks on the MR images were selected a second time in six patients. Table I shows the mTRE values of the first and second selection process. The mean distance between the two sets of landmarks in the *pre-operative* MR images is 1.5 mm with a std of 0.2 mm. Using the two sets of landmarks in the MR images, two measurements for mTRE were calculated between the *pre-operative MR* and *post-resection US* images, giving a mTRE mean absolute difference of 0.3 mm as shown in Table I. Comparing these numbers to the mean initial mTRE of 8.2 mm, the landmarks are accurate for validation purposes.

2) Parameter Selection: There is a tradeoff between the locality of the CR measurement and its accuracy when setting the patch size. Using one of our patient datasets, we found that a patch size of 7^3 voxels (voxels are of size 1 mm in all three dimensions) gives good results. This patch size depends

Patient	Landmarks	Mean dist.	mTRE1	mTRE2
P1	13	1.4	7.2	6.8
P2	11	1.8	8.9	8.5
P3	19	1.5	8.7	8.4
P4	20	1.2	4.6	4.8
P5	21	1.2	3.6	3.8
P6	16	1.6	3.6	3.3
mean	17	1.4	6.1	5.9

on the resolution and voxel size of the US and MR images. For example, if we oversample the images, the physical size of the patch in mm should remain the same, resulting in more voxels in the patch. The size of the patch can be increased if the similarity metric does not need to be very local, for example when registering T1 and T2 images that suffer from small intensity inhomogeneity.

The improvement with using more number of patches N_p diminishes after a certain point, while the computation time increases linearly with N_p. We found that $N_p = 1000$ provides a good compromise between the running time and accurate results.

The number of hierarchical levels and the spacing between the B-spline nodes are two important parameters that should be set. The number of levels depends on frequency content of the images (i.e. smooth or textured) and on the amount of deformation. More levels should be used for smooth images and large deformations to avoid getting trapped in local minima and to speed the optimization. We found that two levels are enough to register the MR and US images of the BITE database. The spacing between B-spline nodes depends on the complexity of the underlying deformation field. Too many nodes complicate the optimization problem and can cause overfitting. We found that 20 mm spacing is good for modeling the brain shift in the BITE database. Since the number of nodes (and hence the number of deformation unknowns) in the coarse level is $1/8^{th}$ of the fine level, we respectively use $N_p = 125$ and $N_p = 1000$ in the two levels. This multi-level optimization increases the capture range, reduces the computation time and reduces the number of local minima.

In stochastic gradient descent optimization, the gradient estimate at the optimal solution is not necessarily zero [14]. Therefore, convergence is ensured by decreasing the step-size with iteration (Eq. 13). We set the number of iterations to 50 in each of the two levels. Given the total runtime of 30 sec, each iteration takes 0.3 sec in average. A very small percentage of this time is spent on estimating the similarity metric and its derivative: for $N_p = 1000$ patches, $7^3 = 343$ voxels in each patch, and 32 bins, calculating RaPTOR and its derivative with respect to the intensity of all voxels takes only 0.03 sec. This is mostly because our registration implementation (i.e.

interpolation, warping, etc.) is not optimized for speed, and therefore substantial speedups can be gained.

We have derived the derivative of RaPTOR with respect to both Y (Eq. 10) and X (Eqs. 11 and 12). Therefore, I_m can be set to either X or Y. In addition, I_m can be set to either MR or US, which provides 4 total combinations for I_m (or equivalently 4 combinations for I_f). Using one of our patient datasets, we found that setting the US image as I_m and Y gives the best results. Therefore, the MR image is set to I_f and X.

The tumor region in the MR has generally high intensity gradient, while the cavity region in US is generally smooth (see Figure 10 as an example). To utilize this to help find outliers, we calculate the sum of the derivatives of I_{US} and I_{MR} inside all patches of the same size as CR computations. Such summation can be performed efficiently using convolution:

$$G_{US} = \|\nabla I_{US}\| * B, \quad G_{MR} = \|\nabla I_{MR}\| * B \qquad (17)$$

where $\|\cdot\|$ represent the magnitude, $*$ the convolution, B a box kernel which is 1 inside the patch and 0 elsewhere, and G_{US} and G_{MR} the summation of the gradient magnitudes of US and MR images over patches. We then compute the ratio

$$r_g = \frac{G_{MR}}{G_{US}}. \qquad (18)$$

Patches with large r_g represent regions with low gradient in the US (such as the resection cavity) and high gradient in the MR (such as the tumor). We use this equation, along with Eq. 14 to find outliers as patches with $r \cdot r_g > T$. Using one of our patient datasets, we found that $T = 1$ gives good results; we used the same value for all other cases. This value can change for different US transducers, and therefore should be adjusted using a sample dataset.

3) Registration Results: We first illustrate the gradient descent direction of individual pixels for inlier and outlier patches in MR-US registration (Figure 11). To allow easy visualization, we use 2D MR and US images, 2D patches of size 21×21 pixels, and 2D cubic B-spline nodes in this example. In (c) and (d), the blue patches correspond and therefore individual pixels have some agreement in the descent direction, while in (e) and (f) pixels point to random directions since the red patches do not correspond. Note that B-spline transformation has an inherent smoothing property caused by averaging the vectors from multiple pixels, and therefore (d) and (f) remove some of the spurious displacements in (c) and (e). Nevertheless, some erroneous displacements survive the averaging step and show in the B-spline nodes in (f).

One of the patients (patient P10 in Table II) has a relatively small initial mTRE value of 2.3 mm. We calculate the optimum rigid registration (by minimizing the mTRE) and transform the MR image with the optimal rigid registration. The new mTRE is 2.2 mm. We then move the MR image in x y directions by ± 12 voxels (voxel size is 1^3 mm) and calculate both LMI and RaPTOR cost functions at each displacement. The results are shown in Figure 12. The minimum cost of RaPTOR is at (0,0), while the minimum cost for LMI is around $(x,y)=(-5,5)$. The combination of very high spatial bias and outliers cause
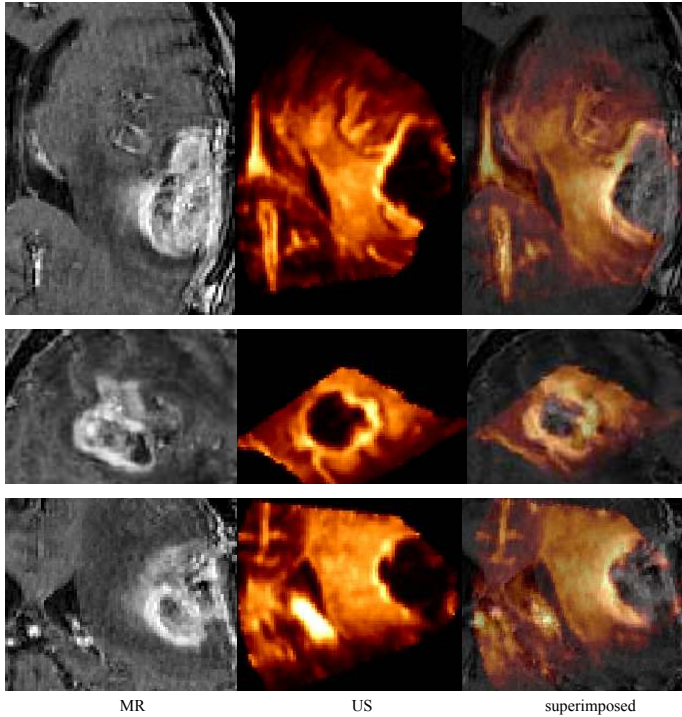
Fig. 10. The tumor region generally has texture in the MR image, while the cavity is generally uniform in the US. First, second and third rows are respectively axial, sagittal and coronal slices.

(a) MR   (b) US

(c) Gradient descent, pixels   (d) Gradient descent, nodes

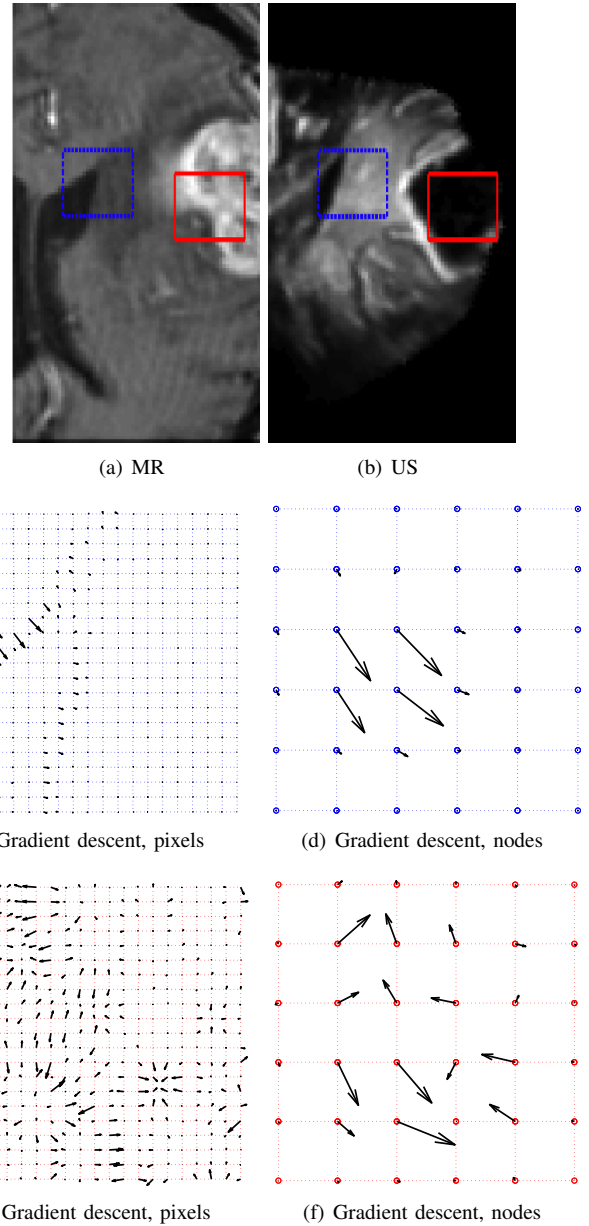(e) Gradient descent, pixels   (f) Gradient descent, nodes

Fig. 11. The gradient descent directions in MR-US registration. The blue dashed and red solid patches in (a) and (b) are respectively corresponding and non-corresponding patches. Second and third rows respectively correspond to the gradient descent directions for the blue and red patches.

the LMI to fail. Note that this set of US and MR volumes only provide a bronze standard, as there is still non-rigid deformation between the two volumes with mTRE of 2.2 mm. We believe the relatively smooth minimum at (0,0) in (b) is in fact caused by these small deformations: at small deformations around (0,0) some parts of the US and MR volumes align, causing the minimum at 0 displacement to be relatively blurry. Also, note the large ± 12 mm capture range of RaPTOR in either direction.

We show the RaPTOR displacement field, i.e. the brain shift, in Figure 13. The magnitude of the displacements is between 0.5 mm and 3.9 mm. Note that the maximum deformation is located close to the tumor/resection cavity where the ultrasound probe is positioned. This is in accordance with the intuition that the highest displacement should happen at the surface where the brain tissue is the least constrained.

Figure 14 shows the alignment of the tumor and the resection cavity before and after registration. The alignment of the longitudinal fissure, ventricles and the tumor/cavity boundaries are substantially improved after nonlinear registration. This improvement can potentially help the neurosurgeon remove residual tumor, while minimizing damage to healthy brain tissue.

The final mTRE results for all patients are shown in Table II, with the last three rows showing a summary of results over the 13 patients. RaPTOR reduces the average mTRE from 5.9 mm to 2.9 mm, and the maximum TRE from 17 mm (in patient 11) to 8.3 mm (in patient 2). The difference between the initial and RaPTOR mTRE values is statistically significant,

with a p-value of 0.01 using t-test. The results of Patient 11, with an mTRE of 14 mm and maximum TRE of 17 mm shows that RaPTOR also has a large capture range. The minimum, average and maximum TRE values are also shown in Figure 15, which highlights the significant reduction of TRE values with RaPTOR.

To test the sensitivity of the results to the patch size, we repeat the registration experiments using patches of size 5^3 and 9^3 mm. The resulting average mTRE values over all 13 patients are receptively 2.9 mm and 3.0 mm. These numbers are very close to 2.9 mm average mTRE in Table II, which is obtained by a patch size of 7^3 mm. Finally, we perform the registration experiments with $N_p = 500$ and $N_p = 2000$. The
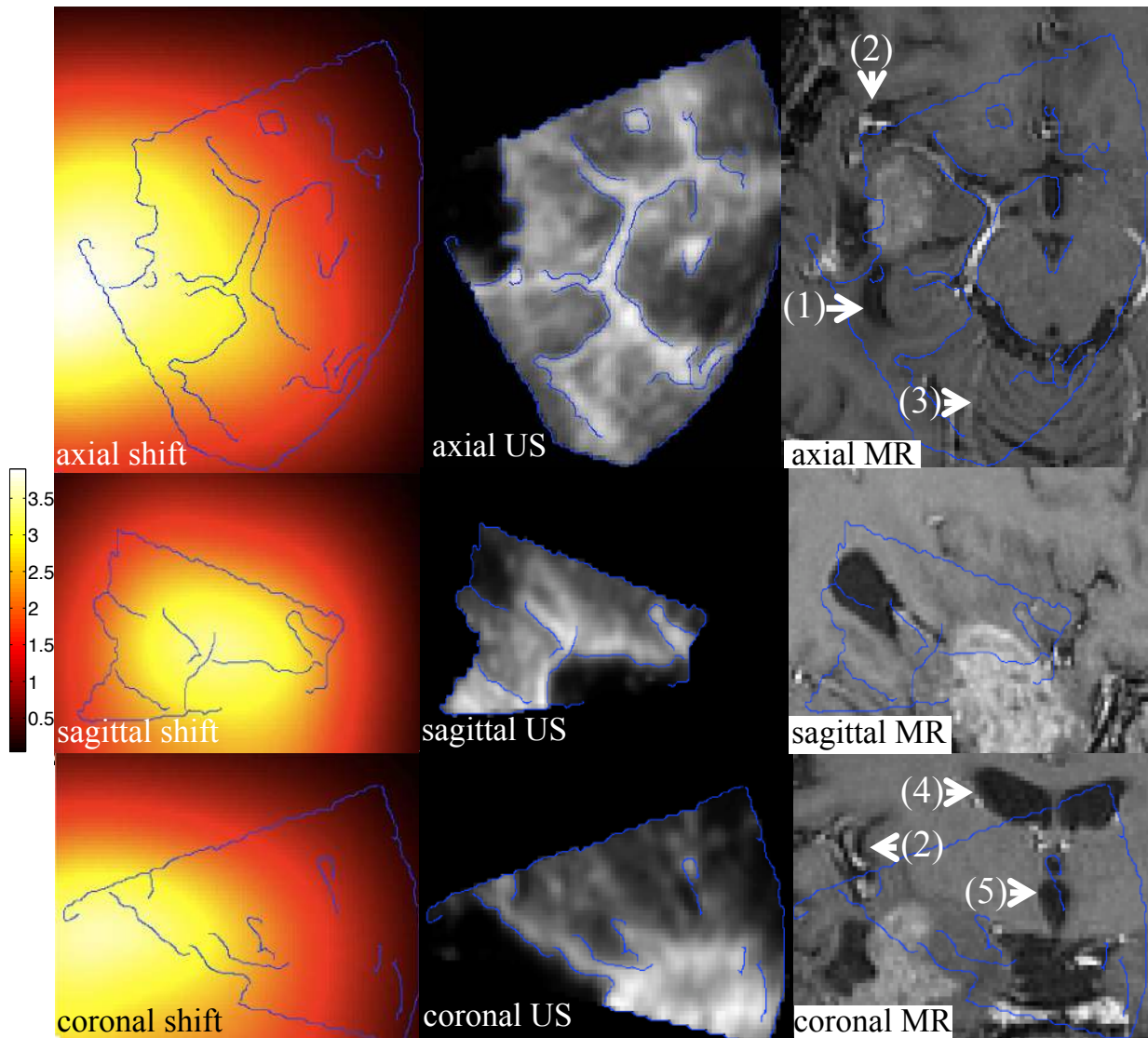
Fig. 13. The brain shift (in the first column) in mm, along with the corresponding US and MR images. The annotated structures are: (1) temporal horn of the lateral ventricles, (2) insular cortex, (3) hippocampus, (4) lateral ventricle, (5) third ventricles. The blue contours are automatically calculated from the US images using Canny edge detector and are shown to help the reader compare the brain shift, US and MR images.
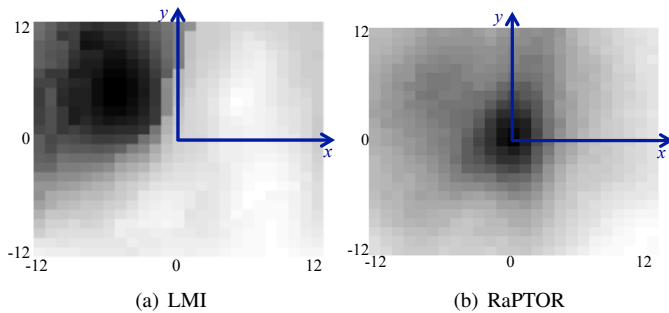

Fig. 12. Two cost functions (represented by image intensities) for a range of ± 12 voxels (each voxel is 1 mm) displacement of the MR image in the x and y directions. The minimum cost is expected to be at $(x,y)=(0,0)$, i.e. at 0 translation.

resulting average mTRE values are 2.96 mm and 2.90 mm,

again similar to 2.9 mm of Table II. These results show the size and number of the patches can be varied considerably while the final mTRE result varies negligibly.

IV. DISCUSSIONS

The results of Figure 14 show that the quality of the MR-US alignment substantially improves with RaPTOR. This registration is of significant clinical importance for multiple reasons. First, the contrast of brain tumors is often low in the US images, and therefore fusion of the MR can improve tumor visualization. Second, Surgicel and hemorrhage can further deteriorate the quality of the post-resection US images. Finally, MR images show exquisite details of brain anatomy and most neurosurgeons are trained with this modality. Therefore, accurate registration of MR and post-resection US can potentially reduce the presence of residual tumors. We plan to integrate RaPTOR with our neuro-navigation system IBIS, and use it

(a) US   (b) initial MR   (c) registered MR

(d) US   (e) initial MR   (f) registered MR

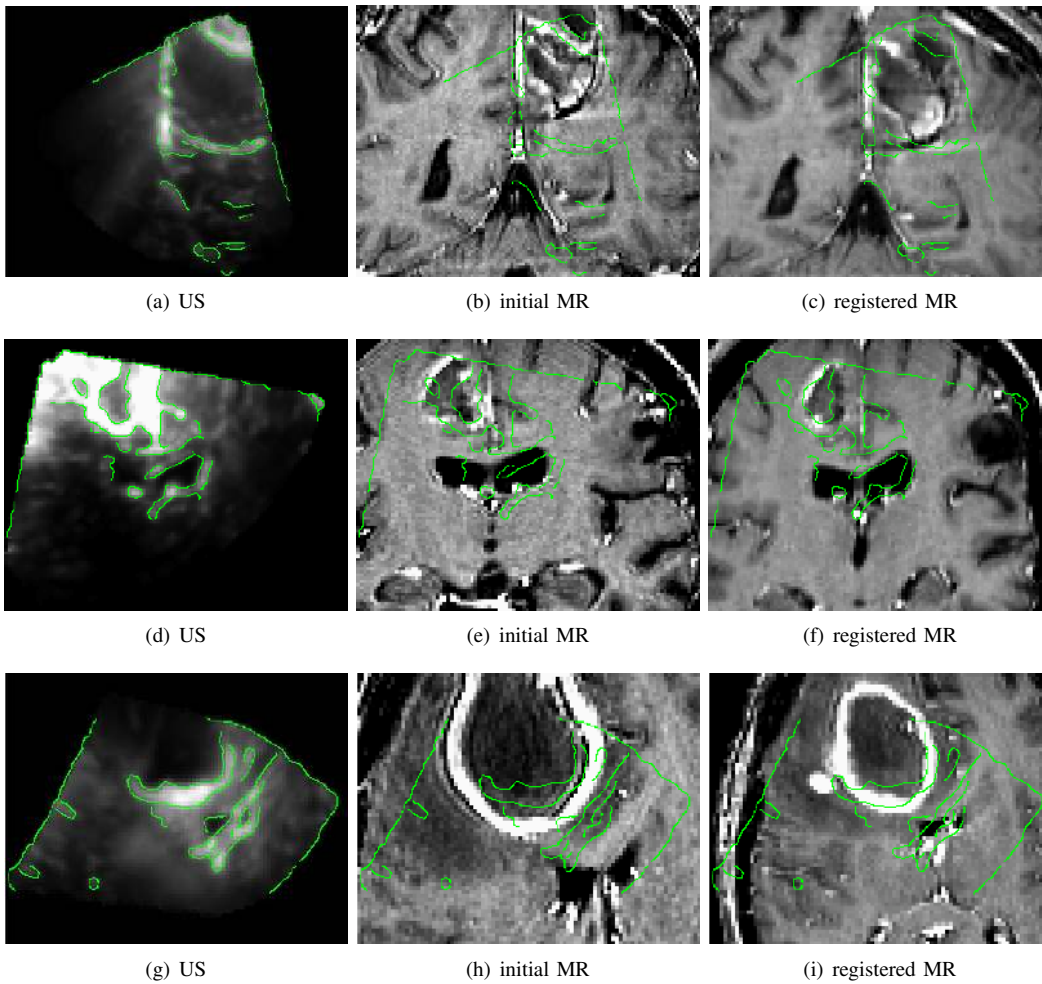(g) US   (h) initial MR   (i) registered MR

Fig. 14. US and MR images before and after 3D registration in three patients. The green contours are automatically calculated from the US images using Canny edge detector, and are overlain on MR images to allow visual assessment of the alignment. The initial registration in the second column is obtained using the tracking information. The alignment of the ventricles and tumor is substantially improved after the registration in the third column.

in the OR to assess its robustness, accuracy and potential improvement on the outcome of the surgery.

RaPTOR estimates very local similarity metrics over small patches, and averages the results over many patches. It can therefore perform accurate non-rigid registration of US and MR images, whose intensities are only related in very local scales. The combination of the efficient gradient decent optimization (instead of generally slower non-gradient based methods) and analytic derivation of the similarity metric (instead of numerical/finite difference estimation of the derivative) makes RaPTOR computationally efficient. Finally, the use of the stochastic gradient descent method means that RaPTOR is only estimated in a small subset of the image voxels, leading to its near real-time performance for non-rigid registration of volumetric data. Further speed-ups can be achieved using parallel implementation on the graphics card [58].

This work improves on our previous work [37], [38] in two fronts. First, it runs in only 30 sec (implemented in MATLAB and MEX functions), compared to 2 hr in [37] and 10 min in [38] (similar implementations in MATLAB and MEX functions). Second, it suppresses the effect of outliers

by exploiting the orientations of the descent direction for individual voxels.

Part of our criteria for finding areas of non-correspondence is based on looking for regions with low gradient in US and high gradient in MR (Eq. 18). This assumption may not improve the results for low grade glioma (LGG) tumors that have low contrast in MR. An area of future work is to perform manual segmentation of the tumor and resection cavity and learn some intensity distribution function for tumor in MR and the resection cavity in US, similar to the work of Kamen *et al.* [44]. This intensity distribution can then be used to further improve the detection of the tumor and the resection cavity.

We will provide our manual landmarks available online, and create a fourth group of images in the BITE database, which can significantly speed the translation of future algorithms from bench to bedside. This data also can be used for easy comparison of future registration techniques.

Several features of RaPTOR can lead to its widespread use in the operating room (OR). First, RaPTOR uses a local similarity metric, and therefore is robust against the large

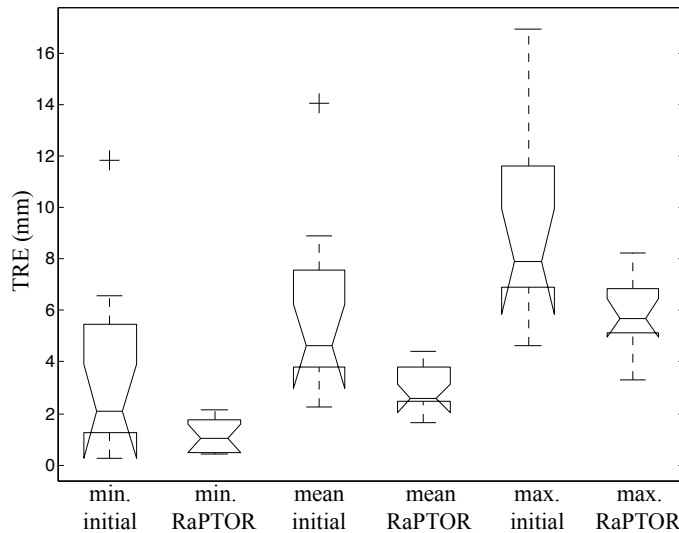| Patient	Landmarks	Initial	RaPTOR
P1	13	7.2 (5.2-10.6)	**3.8 (2.1-5.7)**
P2	11	8.9 (6.2-13.8)	**4.1 (1.7-8.3)**
P3	19	8.7 (6.6-12.7)	**2.9 (0.9-6.9)**
P4	20	4.6 (1.1-11.3)	**2.5 (1.0-6.8)**
P5	21	3.6 (**0.4**-5.6)	**2.5** (0.9-**4.1**)
P6	16	3.6 (1.4-7.9)	**2.5 (0.4-5.4)**
P7	17	6.4 (3.6-7.9)	**2.7 (1.5-5.4)**
P8	15	5.1 (1.3-8.3)	**2.5 (0.4-6.8)**
P9	12	4.2 (2.4-6.5)	**2.3 (0.5-5.2)**
P10	16	2.3 (**0.2**-4.6)	**1.6** (0.4-**3.3**)
P11	9	14.0 (11.8-17.0)	**4.4 (1.8-8.1)**
P12	13	4.6 (1.7-7.1)	**1.3 (2.1-6.1)**
P13	12	3.8 (2.1-7.0)	**1.3 (1.1-4.9)**
mean	15	5.9 (3.4-9.3)	**2.9 (1.1-5.9)**
std	3.7	3.2 (3.3-3.6)	**0.8 (0.6-1.4)**
p-value			**0.01**

Fig. 15. The boxplot of minimum, mean and maximum TRE values over 13 patients from Table II.

spatial inhomogeneity in the US images. Second, it is based on correlation ratio, a well-established similarity measure that is relatively easy to implement. Third, it is very fast, due to the use of analytic derivative of the similarity measure and efficient stochastic gradient descent optimization. Faster runtime may be achieved by optimizing the implementation and exploiting graphics card.

V. CONCLUSION

We presented RaPTOR, an algorithm for non-rigid registration of challenging clinical images of pre-operative MR and post-operative US. We analytically derived the derivatives of RaPTOR and optimized it using efficient stochastic gradient descent optimization. We also proposed a novel intuitive technique for minimizing the effect of outliers. We will provide our data available online, which we hope, speeds the translation of future registration techniques to the operating rooms.

REFERENCES

[1] M. Letteboer, P. Willems, M. Viergever, and W. Niessen, "Brain shift estimation in image-guided neurosurgery using 3-d ultrasound," *IEEE Trans Med. Imag.*, vol. 52, pp. 267–276, 2005.

[2] I. Reinertsen, A. Jakola, P. Friderichsen, F. Lindseth, O. Solheim, T. Selbekk, and G. Unsgård, "A new system for 3d ultrasound-guided placement of cerebral ventricle catheters," *International journal of computer assisted radiology and surgery*, vol. 7, no. 1, pp. 151–157, 2012.

[3] L. Mercier, V. Fonov, C. Haegelen, R. Maesstro, K. Petrecca, and D. L. Collins, "Comparing two approaches to rigid registration of three-dimensional ultrasound and magnetic resonance images for neurosurgery," *Compt Aided Surg*, vol. 7, no. 1, pp. 125–136, 2012.

[4] D. De Nigris, D. L. Collins, and T. Arbel, "Multi-modal image registration based on gradient orientations of minimal uncertainty," *IEEE Trans. Medical Imag*, vol. 31, no. 12, pp. 2343–2354, 2012.

[5] A. Nabavi, P. Black, and et al., "Serial intraoperative magnetic resonance imaging of brain shift," *Neurosurgery*, vol. 48, no. 4, pp. 787–798, 2001.

[6] A. Roche, G. Malandain, N. Ayache, and S. Prima, "Towards a better comprehension of similarity measures used in medical image registration," in *Medical Image Computing and Computer-Assisted Intervention*. Springer, 1999, pp. 555–566.

[7] W. Wells, P. Viola, H. Atsumid, S. Nakajimae, and R. Kikinis, "Multimodal volume registration maximization of mutual information," *Med Imag Anal*, vol. 1, no. 1, pp. 35–51, 1996.

[8] F. Maes, A. Collignon, D. Vandermeulen, M. G, and P. Suetens, "Multimodality image registration by maximization of mutual information," *IEEE Trans. Medical Imag*, vol. 16, no. 2, pp. 187–198, 1997.

[9] J. Pluim, J. Maintz, and M. Viergever, "Mutual-information-based registration of medical images: a survey," *IEEE Trans. Medical Imag.*, vol. 22, no. 8, pp. 986–1004, 2003.

[10] A. Roche, G. Malandain, X. Pennec, and N. Ayache, "The correlation ratio as a new similarity measure for multimodal image registration," *Medical Image Computing Computer Assisted Intervention (MICCAI)*, pp. 1115–1124, 1998.

[11] S. Klein, U. van der Heide, I. Lips, M. van Vulpen, M. Staring, and J. Pluim, "Automatic segmentation of the prostate in 3D MR images by atlas matching using localized mutual information," *Med. Phys.*, vol. 35, no. 4, pp. 1407–1417, 2008.

[12] A. Roche, G. Malandain, N. Ayache, and X. Pennec, "Multimodal image registration by maximization of the correlation ratio," *INRIA-3378*, pp. 1–42, 1998.

[13] S. Ourselin, A. Roche, S. Prima, and N. Ayache, "Block matching: A general framework to improve robustness of rigid registration of medical images," *Medical Image Computing Computer Assisted Intervention (MICCAI)*, pp. 557–566, 2000.

[14] S. Klein, M. Staring, and J. Pluim, "Evaluation of optimization methods for nonrigid medical image registration using mutual information and b-splines," *IEEE Trans. Imag Proc*, vol. 16, no. 12, pp. 2879–2890, December 2007.

[15] T. Selbekk, A. S. Jakola, O. Solheim, T. F. Johansen, F. Lindseth, I. Reinertsen, and G. Unsgård, "Ultrasound imaging in neurosurgery: approaches to minimize surgically induced image artefacts for improved resection control," *Acta neurochirurgica*, vol. 155, no. 6, pp. 973–980, 2013.

[16] L. Mercier, R. F. Del Maestro, K. Petrecca, D. Araujo, C. Haegelen, and D. L. Collins, "Online database of clinical mr and ultrasound images of brain tumors," *Medical Physics*, vol. 39, p. 3253, 2012.

[17] A. Roche, X. Pennec, G. Malandain, and N. Ayache, "Rigid registration of 3-d ultrasound with mr images: a new approach combining intensity and gradient information," *IEEE Trans Med Imag*, vol. 20, no. 72, pp. 291–237, 2001.

[18] T. Arbel, X. Morandi, R. Comeau, and D. L. Collins, "Automatic non-linear mri-ultrasound registration for the correction of intra-operative brain deformations," *Compt Aided Surg*, vol. 9, no. 4, pp. 123–136, 2004.

[19] D. L. Collins, P. Neelin, T. M. Peters, and A. C. Evans, "Automatic 3d intersubject registration of mr volumetric data in standardized talairach space," *Journal of computer assisted tomography*, vol. 18, no. 2, pp. 192–205, 1994.

[20] M. Kuklisova-Murgasova, A. Cifor, R. Napolitano, A. Papageorghiou, G. Quaghebeur, M. A. Rutherford, J. V. Hajnal, J. Alison Noble, and J. A. Schnabel, "Registration of 3d fetal neurosonography and mri," *Medical Image Analysis*, 2013.

[21] W. Wein, S. Brunke, A. Khamene, M. Callstrom, and N. Navab, "Automatic ct-ultrasound registration for diagnostic imaging and image-guided intervention," *Medical Imag Analysis*, vol. 12, pp. 577–585, 2008.

[22] M. Mellor and M. Brady, "Phase mutual information as a similarity measure for registration," *Medical image analysis*, vol. 9, no. 4, pp. 330–343, 2005.

[23] W. Zhang, M. Brady, H. Becher, and A. Noble, "Spatio-temporal (2d+t) non-rigid registration of real-time 3d echocardiography and cardiovascular mr image sequences," *Physics Med Biol*, vol. 56, pp. 1341–1360, 2011.

[24] S. Ji, Z. Wu, A. Hartov, D. Roberts, and K. Paulsen, "Mutual-information-based image to patient re-registration using intraoperative ultrasound in image-guided neurosurgery," *Med. Phys.*, vol. 35, no. 10, pp. 4612–4624, 2008.

[25] J. Pluim, J. Maintz, and M. Viergever, "Image registration by maximization of combined mutual information and gradient information," *IEEE Trans. Medical Imag.*, vol. 19, no. 8, pp. 809–814, 2000.

[26] E. Haber and J. Modersitzki, "Intensity gradient based registration and fusion of multi-modal images," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2006*. Springer, 2006, pp. 726–733.

[27] M. P. Heinrich, M. Jenkinson, B. W. Papież, M. Brady, and J. A. Schnabel, "Towards realtime multimodal fusion for image-guided interventions using self-similarities," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2013*. Springer, 2013, pp. 187–194.

[28] M. Heinrich, M. Jenkinson, M. Bhushan, T. Matin, F. Gleeson, M. Brady, and J. Schanbel, "MIND: Modality independent neighbourhood descriptor for multi-modal deformable registration," *Medical Image Analysis*, vol. 16, no. 7, pp. 1423–1435, 2012.

[29] W. Wein, A. Ladikos, B. Fuerst, A. Shah, K. Sharma, and N. Navab, "Global registration of ultrasound to mri using the lc2 metric for enabling neurosurgical guidance," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2013*. Springer, 2013, pp. 34–41.

[30] B. Fuerst, W. Wein, M. Müller, and N. Navab, "Automatic ultrasound-MRI registration for neurosurgery using the 2D and 3D LC^2 metric," *Medical Image Analysis*, 2014.

[31] E. Ferrante and N. Paragios, "Non-rigid 2d-3d medical image registration using markov random fields," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2013*. Springer, 2013, pp. 163–170.

[32] Y. Sun, J. Yuan, M. Rajchl, W. Qiu, C. Romagnoli, and A. Fenster, "Efficient convex optimization approach to 3d non-rigid mr-trus registration," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2013*. Springer, 2013, pp. 195–202.

[33] G. Penney, J. Blackall, M. Hamady, T. Sabharwal, A. Adam, and D. Hawks, "Registration of freehand 3d ultrasound and magnetic resonance liver images," *Med Imag Anal*, vol. 8, no. 1, pp. 81–91, 2004.

[34] I. Reinertsen, M. Descoteaux, K. Siddiqi, and D. L. Collins, "Validation of vessel-based registration for correction of brain shift," *Medical image analysis*, vol. 11, no. 4, pp. 374–388, 2007.

[35] I. Reinertsen, F. Lindseth, G. Unsgaard, and D. Collins, "Clinical validation of vessel-based registration for correction of brain-shift," *Med. Imag. Anal.*, vol. 11, no. 6, pp. 673–684, December 2007.

[36] P. Coupé, P. Hellier, X. Morandi, and C. Barillot, "3D rigid registration of intraoperative ultrasound and preoperative mr brain images based on hyperechogenic structures," *Journal of Biomedical Imaging*, vol. 2012, p. 1, 2012.

[37] H. Rivaz and D. L. Collins, "Self-similarity weighted mutual information: A new nonrigid image registration metric," *Medical Image Computing Computer Assisted Intervention (MICCAI)*, pp. 91–98, 2012.

[38] H. Rivaz, Z. Karimaghaloo, V. Fonov, and D. L. Collins, "Nonrigid registration of ultrasound and mri using contextual conditioned mutual information," *IEEE Trans. Medical Imag*, vol. 33, no. 3, pp. 708–725, March 2014.

[39] N. Chitphakdithai and J. Duncan, "Non-rigid registration with missing correspondences in preoperative and post-resection brain images," *MICCAI*, pp. 367–374, 2010.

[40] N. Chitphakdithai, K. P. Vives, and J. S. Duncan, "Registration of brain resection mri with intensity and location priors," in *Biomedical Imaging: From Nano to Macro, 2011 IEEE International Symposium on*. IEEE, 2011, pp. 1520–1523.

[41] N. Chitphakdithai, V. Chiang, and J. Duncan, "Tracking metastatic brain tumors in longitudinal scans via joint image registration and labeling," *MICCAI STIA*, pp. 124–136, 2012.

[42] P. Risholm, E. Samset, I. Talos, and W. Wells, "A non-rigid registration framework that accommodates resection and retraction," *IPMI*, pp. 447–458, 2009.

[43] P. Daga, M. Modat, C. Micallef, L. Mancini, M. White, M. J. Cardoso, N. Kitchen, A. McEvoy, J. Thornton, T. Yousry, D. Hawkes, and D. Ourselin, "Near real time brain shift estimation for interventional mri suite," *High-Performance MICCAI*, 2010.

[44] A. Khamene, D. Zikic, M. Diallo, T. Boettger, and E. Rietzel, "A novel intensity similarity metric with soft spatial constraint for a deformable image registration problem in radiation therapy," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2009*. Springer, 2009, pp. 828–836.

[45] D. Rueckert, L. I. Sonoda, C. Hayes, D. L. Hill, M. O. Leach, and D. J. Hawkes, "Nonrigid registration using free-form deformations: application to breast mr images," *Medical Imaging, IEEE Transactions on*, vol. 18, no. 8, pp. 712–721, 1999.

[46] S. Klein, J. Pluim, M. Staring, and M. Viergever, "Adaptive stochastic gradient descent optimisation for image registration," *Int J Comput Vis*, vol. 81, pp. 227–239, 2009.

[47] P. Huber, *Robust statistics*. Wiley, New York, 1997.

[48] G. Hager and P. Belhumeur, "Efficient region tracking with parametric models of geometry and illumination," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 10, pp. 1025–1039, Oct. 1998.

[49] H. Rivaz, E. M. Boctor, M. A. Choti, and G. D. Hager, "Real-time regularized ultrasound elastography," *IEEE Trans Med Imaging*, vol. 30, no. 4, pp. 928–945, Apr 2011.

[50] J. V. Manjón, J. Tohka, G. García-Martí, J. Carbonell-Caballero, J. J. Lull, L. Martí-Bonmatí, and M. Robles, "Robust mri brain tissue parameter estimation by multistage outlier rejection," *Magnetic Resonance in Medicine*, vol. 59, no. 4, pp. 866–873, 2008.

[51] D. L. Collins, A. Zijdenbos, V. Kollokian, J. Sled, N. Kabani, C. Holmes, and A. Evans, "Design and construction of a realistic digital brain phantom," *IEEE Trans. Medical Imag*, vol. 17, no. 3, pp. 463–468, 1998.

[52] A. Andronache, M. V. Siebenthal, G. Szekely, and P. Cattin, "Non-rigid registration of multi-modal images using both mutual information and cross-correlation," *Medical Imag Analysis*, vol. 12, no. 1, pp. 3–15, 2008.

[53] J. Beirlant, E. D. L. Györfi, and E. van der Meulen, "Nonparametric entropy estimation: an overview," 2001.

[54] L. Mercier, R. Maestro, K. Petrecca, A. Kochanowska, S. Drouin, C. Yan, A. Janke, S. Chen, and D. L. Collins, "New prototype neuronavigation system based on preoperative imaging and intraoperative freehand ultrasound: System description and validation," *IJ CARS*, vol. 6, no. 4, pp. 507–522, 2011.

[55] O. Solberg, F. Lindseth, H. Torp, R. Blake, and N. H, "Freehand 3D ultrasound reconstruction algorithms: A review," *Neurosurgery*, vol. 33, no. 7, pp. 991–1009, 2007.

[56] H. Rivaz, E. M. Boctor, M. A. Choti, and G. D. Hager, "Ultrasound elastography using multiple images," *Medical image analysis*, vol. 18, no. 2, pp. 314–329, 2014.

[57] P. Jannin, J. M. Fitzpatrick, D. J. Hawkes, X. Pennec, R. Shahidi, and M. W. Vannier, "Validation of medical image processing in image-guided therapy," *Neurosurgery*, vol. 21, no. 2, pp. 1445–1449, 2002.

[58] P. Daga, G. Winston, M. Modat, M. White, L. Mancini, M. J. Cardoso, M. Symms, J. Stretton, A. W. McEvoy, J. Thornton, C. Micallef, T. Yousry, D. Hawkes, J. Duncan, and S. Ourselin, "Accurate localization of optic radiation during neurosurgery in an interventional mri suite," *Medical Imaging, IEEE Transactions on*, vol. 31, no. 4, pp. 882–891, 2012.