NeuroImage 202 (2019) 116094

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/neuroimage

Deformable MRI-Ultrasound registration using correlation-based attribute matching for brain shift correction: Accuracy and generality in multi-site data

Inês Machado [a,b,*], Matthew Toews [c], Elizabeth George [a], Prashin Unadkat [d], Walid Essayed [d], Jie Luo [e], Pedro Teodoro [j], Herculano Carvalho [f], Jorge Martins [b], Polina Golland [g], Steve Pieper [a,h], Sarah Frisken [a], Alexandra Golby [d], William Wells III [a,g], Yangming Ou [i,**]

[a] Department of Radiology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA
[b] Department of Mechanical Engineering, Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal
[c] Department of Systems Engineering, École de Technologie Supérieure, Montreal, Canada
[d] Department of Neurosurgery, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA
[e] Graduate School of Frontier Sciences, University of Tokyo, Tokyo, Japan
[f] Department of Neurosurgery, Hospital de Santa Maria, CHLN, Lisbon, Portugal
[g] Computer Science and Artificial Intelligence Laboratory, MIT, Cambridge, MA, USA
[h] Isomics, Inc., Cambridge, MA, USA
[i] Department of Pediatrics and Radiology, Boston Children's Hospital, Harvard Medical School, Boston, MA, USA
[j] Escola Superior Náutica Infante D. Henrique, Lisbon, Portugal

ARTICLE INFO

Keywords:
Brain shift
Intraoperative ultrasound
MR-iUS registration
Multi-site data
Brain Tumor
Surgical Guidance

ABSTRACT

Intraoperative tissue deformation, known as brain shift, decreases the benefit of using preoperative images to guide neurosurgery. Non-rigid registration of preoperative magnetic resonance (MR) to intraoperative ultrasound (iUS) has been proposed as a means to compensate for brain shift. We focus on the initial registration from MR to predurotomy iUS. We present a method that builds on previous work to address the need for accuracy and generality of MR-iUS registration algorithms in multi-site clinical data. High-dimensional texture attributes were used instead of image intensities for image registration and the standard difference-based attribute matching was replaced with correlation-based attribute matching. A strategy that deals explicitly with the large field-of-view mismatch between MR and iUS images was proposed. Key parameters were optimized across independent MR-iUS brain tumor datasets acquired at 3 institutions, with a total of 43 tumor patients and 758 reference landmarks for evaluating the accuracy of the proposed algorithm. Despite differences in imaging protocols, patient demographics and landmark distributions, the algorithm is able to reduce landmark errors prior to registration in three data sets (5.37±4.27, 4.18±1.97 and 6.18±3.38 mm, respectively) to a consistently low level (2.28±0.71, 2.08±0.37 and 2.24±0.78 mm, respectively). This algorithm was tested against 15 other algorithms and it is competitive with the state-of-the-art on multiple datasets. We show that the algorithm has one of the lowest errors in all datasets (accuracy), and this is achieved while sticking to a fixed set of parameters for multi-site data (generality). In contrast, other algorithms/tools of similar performance need per-dataset parameter tuning (high accuracy but lower generality), and those that stick to fixed parameters have larger errors or inconsistent performance (generality but not the top accuracy). Landmark errors were further characterized according to brain regions and tumor types, a topic so far missing in the literature.

1. Introduction

The proximity of many tumors to critical areas of the brain coupled with the difficulty of differentiating tumor tissue from normal brain parenchyma based on visual inspection renders complete tumor removal a very challenging task (Dimaio et al., 2006). Surgical navigation allows

* Corresponding author. 75 Francis St, Boston, MA 02115, USA.
** Corresponding author. 401 Park Drive, Boston, MA, 02115, USA.
E-mail addresses: ines.prata.machado@tecnico.ulisboa.pt (I. Machado), yangming.ou@childrens.harvard.edu (Y. Ou).

https://doi.org/10.1016/j.neuroimage.2019.116094
Received 10 May 2019; Received in revised form 18 July 2019; Accepted 9 August 2019
Available online 22 August 2019

for the use of preoperative images as a map to guide surgery (Golby, 2015). However, when the brain has shifted, preoperative MR images may not serve as an accurate guide (Gerard et al., 2017). To compensate for brain shift, non-invasive, high-resolution, but time-consuming MR image acquisitions can be registered to near-real-time intraoperative ultrasound (iUS) (Unsgaard et al., 2005; Rygh et al., 2008; Miller et al., 2011; Coupé et al., 2007).

The first focus of our work is the MR-to-iUS registration accuracy. A new correlation-based similarity measure is used on a rich set of attributes extracted at each voxel of the two images. This differs from feature-based registration algorithms, which attempt to identify and match key features in the images, such as landmarks in tumor boundaries, sulci, ventricles or blood vessels (Porter et al., 2001; Reinertsen et al., 2007; Farnia et al., 2018; Coupé et al., 2012). Since localizing key features is difficult and may be subject to errors, features existing in one image may not always appear in another image, and features are not always close to the regions of clinical interest (e.g., tumor and surroundings), our algorithm utilizes all voxels in the images (Sotiras et al., 2013). Voxel-wise approaches generally use either single scalar intensities or high-dimensional attributes extracted at each voxel to find across-image correspondences. Some intensity-based methods previously demonstrated on MR-to-iUS registration tasks use similarity measures such as Mutual Information (MI), Correlation Coefficient (CC), and Correlation Ratio (CR) (Rivaz et al., 2014; Hartov et al., 2008; Roche et al., 2000; Letteboer et al., 2003; Ji et al., 2008; Rivaz et al., 2015; Rivaz and Collins, 2015; Myronenko and Song, 2010; Fuerst et al., 2014). However, relative image intensities between MR and iUS do not always follow a linear relationship. For example, in Fig. 1, where we label corresponding structures in MR and iUS images and number them from 1 to 9. Bright regions in MR correspond to bright intensities in the iUS image, e.g., regions 2 and 5; similarly, dark regions corresponding to dark intensities, e.g., regions 7 and 8. However, bright regions, e.g. regions 1, 4, 6 and 9, or gray regions, e.g. region 3, can correspond to dark intensities. The inconsistent intensity relationship and measurement of voxel-wise

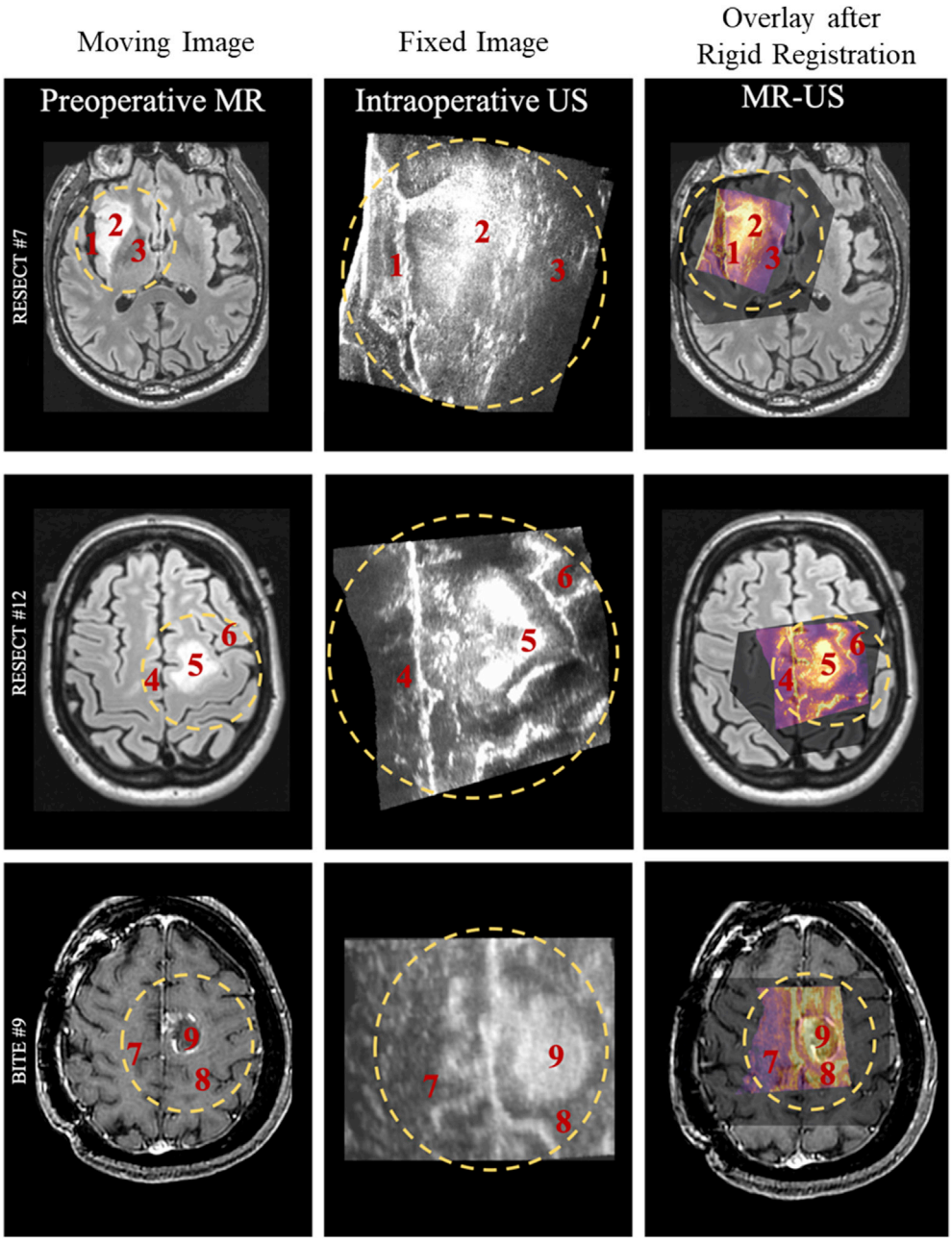


Fig. 1. MR-to-iUS registration challenges. Each row corresponds to a different clinical case. The left, middle, and right columns correspond to the preoperative MR image, iUS image, and the superposition of the two images, after rigid registration, respectively. Dashed yellow circles show the different fields-of-view of the two modalities. Corresponding structures between the two imaging modalities are numbered from 1 to 9. One can see that bright regions in the preoperative MR can correspond to bright intensities in the iUS, e.g., regions 2 and 5; similarly, dark regions in the preoperative MR can correspond to dark intensities in the iUS, e.g., regions 7 and 8. However, bright regions in the MR, e.g. regions 1, 4, 6 and 9, or gray regions, e.g. region 3, can correspond to dark intensities in the iUS image.

similarity on high-dimensional attributes in patches centered at those voxels have been studied, e.g. (Shen and Davatzikos, 2002; Ou et al., 2014; Toews and Wells, 2013; Wachinger and Navab, 2013; Wu et al., 2010; Toews and Arbel, 2003). Self-Similarity Correlation (SSC) (Heinrich et al., 2013) and Linear Correlation of Linear Combination (LC2) methods (Wein et al., 2013) showed increased accuracy in MR-to-iUS registration tasks.

We recently developed an attribute-based algorithm, Deformable Registration via Attribute Matching and Mutual-Saliency Weighting (DRAMMS) (Ou et al., 2011) that has been validated in multi-site clinical images of brain MRI to MRI registrations across subjects and across time (Diez et al., 2014; Ou et al., 2014). In contrast to SSC and LC2, which extract attributes from a patch at single scale, DRAMMS uses multi-scale and multi-orientation textures as attributes, on the basis that voxels in different brain regions may be distinctive at different scales and orientations. This current study proposes using the Normalized Correlation Coefficient (aNCC) and Correlation Ratio (aCR) on attribute vectors as similarity measures. The original DRAMMS version used the Sum of Squared Differences (aSSD) of attribute vectors but aNCC is able to model the linear relationship of attribute vectors and aCR can handle non-linear correlation of attribute vectors. Both are correlation-based similarity measures previously used on image intensities (Roche et al., 2006; Heinrich et al., 2012), though not on attribute vectors. This modified version will be referred as correlation-DRAMMS or cDRAMMS.

In addition to using a new similarity measure, we propose a transformation strategy to explicitly handle the field-of-view (FOV) mismatch between the two images. As the dashed yellow circles in Fig. 1 show, a large proportion of the content in MR is not included in the iUS image. We mask the MR image based on the FOV of the iUS image plus a buffer zone. Recent MR-to-iUS registration studies advocate masking the MR image (Rivaz and Collins, 2015; Heinrich et al., 2013; Rivaz et al., 2014; Jiang et al., 2016), and one suggests including a 3 mm radius buffer zone (Drobny et al., 2018). Our study includes studying the effets of MR masking in greater detail, optimizing the radius of the buffer zone, and evaluating the applicability of these approaches when working with multi-site data.

The second focus of this work is to evaluate the "generality" of the proposed algorithm. Our goal is to achieve consistently high accuracy in multi-site data, without per-dataset parameter tuning. We propose three strategies to achieve this goal. First, we inherit the use of multi-scale and multi-orientation attributes in the original DRAMMS framework, as single-scale and single-orientation attributes may not always generalize well to other datasets (Wein, 2018; Heinrich, 2018). Second, we skip ad-hoc pre-processing. In many MR-to-iUS registration studies, pre-processing is needed. Some methods are quite extensive and require skull stripping (Farnia et al., 2014; Farnia et al., 2016) or tissue segmentation of MR images (Hong et al., 2018; Morin et al., 2017; Palombi et al., 2018; Reinertsen et al., 2014) or segmentation and removal of bright strips in skin surfaces (Wein, 2018; Shams et al., 2018). These pre-processing steps may be non-trivial, error-prone, and may often require case-specific human intervention in tumor-bearing MR images (Drobny et al., 2018; Farnia et al., 2014; Farnia et al., 2016). Finally, we optimize key parameters, seeking a fixed framework that consistently produces good results in multi-site data.

A third feature of our work is a detailed comparison of 16 state-of-the-art MR-to-iUS registration algorithms across multi-site data. Most existing methods have been evaluated with only single-site data (Farnia et al., 2014; Farnia et al., 2016; Lindseth et al., 2003; Farnia et al., 2015), or up to two datasets that resulted in inconsistent levels of accuracy (Hong et al., 2018; Shams et al., 2018; Zhong et al., 2018), or used dataset-specific parameters (Wein, 2018; Heinrich, 2018). In this work, we used the most comprehensive multi-site dataset available to date, which includes the Brain Images of Tumors for Evaluation (BITE) database, released in 2012 (Mercier et al., 2012), the REtroSpective Evaluation of Cerebral Tumors (RESECT) database, released in 2017 (Xiao et al., 2017), and the Multimodal Imaging of Brain Shift (MIBS), a BWH proprietary dataset (Machado et al., 2018b). These 3 datasets include a total of 43 patients and 758 reference landmarks to evaluate the accuracy of the registration. Table 1 displays the multi-site data used in this work covering a wide variety of imaging protocols (e.g., scanner vendor, field-of-strength, resolution, contrast), patient demographics (age, gender, tumor type, grade, location, extent, scope) evaluation references (varying landmark locations and distributions), and levels of brain shift (0–21 mm). The diversity in the data helps to characterize the generality and accuracy of sixteen MR-to-iUS registration algorithms – the largest number of algorithms ever evaluated for registering MR and iUS images of brain tumor patients.

In addition to evaluating our registration method based on multi-site data, we provide the first quantitative evaluation of registration error according to brain region and tumor type. A radiologist physician (E.G.), with 3.5 years of experience in iUS reviewed the 758 reference landmark pairs one by one, and assigned them to five brain regions: (i) sulci/gyri, (ii) ventricles, (iii) falx and tentorium, (iv) tumor boundary, and (v) others where landmarks appear less frequently, including vessels, mid brain, pons, bone, white matter, and interpeduncular cistern. We found that the ordering of the regions (i) through (iv) in this manner reflected an increase in registration errors that is consistent across datasets. We also found that registration errors are comparable in patients with low-grade gliomas and in patients with metastatic brain tumors, but larger in patients with high-grade gliomas. Evaluating errors by brain region and tumor type provides a quantitative reference for neurosurgeons.

This study is an extension of our previous workshop version (Machado et al., 2018a), with a more detailed algorithm description, additional algorithm optimization, comprehensive validation in multi-site data, and quantification of errors according to brain region and tumor type.

2. Multi-site data

Table 1 shows the three datasets used in this study: BITE (Mercier et al., 2012), RESECT (Xiao et al., 2017), and MIBS (Machado et al., 2018a,b; Luo et al., 2018; Tempany et al., 2015). As can be seen in Table 1, there is a wide variety of (a) imaging sites, (b) patient demographics, (c) MR and US protocols, and (d) reference landmark distributions.

3. Methods

We use ultrasound as the fixed image and MR as the moving image, since our goal is to update the MR used in surgical planning to reflect brain shift observable in intra-operative ultrasound.

Section 3.1 revisits the original DRAMMS framework. Sections 3.2, 3.3 and 3.4 present the three key components of our algorithm: Section 3.2, the new similarity measure; Section 3.3, the new transformation strategy; and Section 3.4, the parameter optimization in multi-site data.

3.1. Original DRAMMS framework revisited

The original DRAMMS pipeline is shown in Fig. 2. Given two images $I_1 : \Omega_1 \rightarrow \mathbb{R}$ and $I_2 : \Omega_2 \rightarrow \mathbb{R}$ in the 3D image domains $\Omega_i (i = 1, 2) \subset \mathbb{R}^3$, DRAMMS seeks a transformation T that maps every voxel $\mathbf{u} \in \Omega_1$ to its corresponding point $T(\mathbf{u}) \in \Omega_2$, by minimizing a cost function $E(T)$,

$$\min_T E(T) = \int_{\mathbf{u}\in\Omega_1} \underbrace{ms(\mathbf{u}, T(\mathbf{u}))}_{Mutual-Saliency} \cdot \underbrace{sim\left(A_1^*(\mathbf{u}),\ A_2^*(T(\mathbf{u}))\right)}_{Attribute\ Matching} d\mathbf{u} + \lambda R(T), \quad (1)$$

where $A_i^*(\mathbf{u})$ $(i = 1, 2)$ is an attribute vector that reflects the geometric texture of voxel $\mathbf{u}$ at multiple scales and orientations (four scales: at 3, 5, 9, 17 mm neighborhoods and six orientations, equally sampling 0 to π in three orthogonal planes at a voxel) (Ou et al., 2011); sim $(\cdot, \cdot)$ measures the similarity of two voxels based on their attribute vectors. The default method is to use the sum of the squared differences of attributes (aSSD).

The term $ms(\mathbf{u}, T(\mathbf{u}))$ is a continuously-valued mutual-saliency

Table 1
Detailed information for the three datasets used in this study. They cover a wide variety of patient demographics, MR and US protocols, and reference landmark distributions.

			RESECT	BITE	MIBS
Imaging Site (a)			St. Olavs University Hospital, Trondheim, Norway	Montreal Neurological Institute, Montreal, Canada	Brigham and Women's Hospital, Boston, USA
Patient and Tumor Demographics (b)	**No. of Patients (No. Male)**		22[a]	13 (9 males)	8 (5 males)
	Ages (years): Mean±SD (range)		>18[a]	52.77 ± 18.16 (23–76)	46.37 ± 15.08 (27–69)
	Tumor volume[b] (cm[c]): Mean±SD (range)		40.0 ± 44.09 (1.4–165.9) Munkvold et al. (2017)	34.65 ± 23.62 (0.2–79.2)	21.93 ± 17.90 (3.1–57)
	Tumor Diagnosis: Astrocytoma (32.6%) Oligodendroglioma (34.9%) Oligoastrocytoma (4.7%) Glioblastoma (25.6%) Metastatic Brain Tumor (2.3%)		Astrocytoma (10) Oligodendroglioma (10) Oligoastrocytoma (2)	Astrocytoma (3) Oligodendroglioma (3) Glioblastoma (7)	Astrocytoma (1) Metastasis (1) Oligodendroglioma (2) Glioblastoma (4)
	Tumor Type[c] LGG (67.4%) HGG (30.2%) MET (2.3%)		LGG (22)	LGG (4), HGG (9)	LGG (3), HGG (4), MET (1)
	Tumor Location: Left Hemisphere (51.2%) Right Hemisphere (48.8%) Frontal lobe (51.2%) Temporal lobe (20.9%) Insular lobe (9.3%) Parietal lobe (7%) Parietoccipital lobe (4.7%) Frontoparietal lobe (4.7%) Temporoparietal lobe (2.3%)		Right brain: - Frontal (5), - Temporal (3), - Insular (2), - Frontoparietal (1) Left brain: - Temporal (3) - Frontoparietal (1) - Insular (2) - Frontal (5)	Right brain: - Frontal (2), - Temporal (1), - Parietal (2) Left brain: - Parietoccipital (1) - Frontal (6), - Temporal (1)	Right brain: - Frontal (4), - Temporal (1) Left brain: - Parietal (1), - Parietoccipital (1), - Temporoparietal (1)
MR Protocol (c)	**MRI Machine**		1.5T Siemens Magnetom Avanto	1.5T GE Signa	3T Siemens Magnetom Verio
	MRI Sequences		T2-FLAIR	T2 and T1w gadolinium enhanced MR	T1 and T2-FLAIR
	Date of Acquisition		1 day before the surgery	avg. 17 days pre-surgery (range = 1–72 days)	avg. 18 days pre-surgery (range = 1–77 days)
	Time of Echo and Repetition		388 ms and 5000 ms	8 ms and 23 ms	232 ms and 2000 ms
	Flip Angle		120°	20°	120°
	Voxel Size		1.0 × 1.0 × 1.0 mm[c]	1.0 × 1.0 × 1.0 mm[c]	1.0 × 1.0 × 1.0 mm[c]
US Protocol (c)	**US Frequency Range**		5–15 MHz	4–7 MHz	3.8–10 MHz
	No. of Frames		200 to 400	200 to 600	100 to 300
	Resolution of US Reconstruction		0.14 × 0.14 × 0.14 mm[c] to 0.24 × 0.24 × 0.24 mm[c]	0.3 × 0.3 × 0.3 mm[c] Mercier et al. (2011)	0.5 × 0.5 × 0.5 mm[c]
Landmark Distribution (d)	**Landmark pairs/patient**		15–16	19–40	5–10
	No. Landmark pairs		338	355	65
	Dispersion of landmarks[d]		336.25 ± 155.38	828.28 ± 237.56	529.43 ± 194.96
	Landmark Locations	Sulci/Gyri	206	198	25
		Tumor Boundary	123	35	26
		Falx and Tentorium	8	47	14
		Ventricles	1	62	–
		Others[e]	–	13	–

[a] The RESECT database (Xiao et al., 2017) does not provide patients age and gender.
[b] Tumor volume is quantified based on manually drawn boundaries in the preoperative MR images.
[c] Tumor type is categorized in low-grade glioma (LGG), high-grade glioma (HGG) or metastasis (MET).
[d] The spatial dispersion of the 3D point cloud is measured by the trace of the covariance matrix of the coordinates of points where higher traces correspond to higher degrees of dispersion (Hladuvka, 2003; Rizzini and Caselli, 2007).
[e] "Others" where landmarks appear less frequently, including vessels, mid brain, pons, bone, white matter, and interpeduncular cistern.

weight quantifying the confidence of two voxels $\mathbf{u} \in \Omega_1$ and $T(\mathbf{u}) \in \Omega_2$ being a correspondence. Voxels with ambiguous or nonexistent correspondence (e.g., outlier regions) are assigned with a smaller mutual-saliency weight and thus down-weighted in the registration optimization, thereby reducing the impact of missing correspondences. The term $R(T)$ is a smoothness/regularization term using the Laplacian operator (square of Laplacian), also known as the bending energy, of the deformation field T (Bookstein, 1989). The weight, λ (default at 0.2 (Ou et al., 2011)), is greater if a smoother deformation is needed. We evaluate the sensitivity of this parameter (λ) in Section 5.2.3.

The original DRAMMS framework uses the diffeomorphic Free Form Deformation (FFD) (Rueckert, 1999; Heckemann et al., 2006) as its

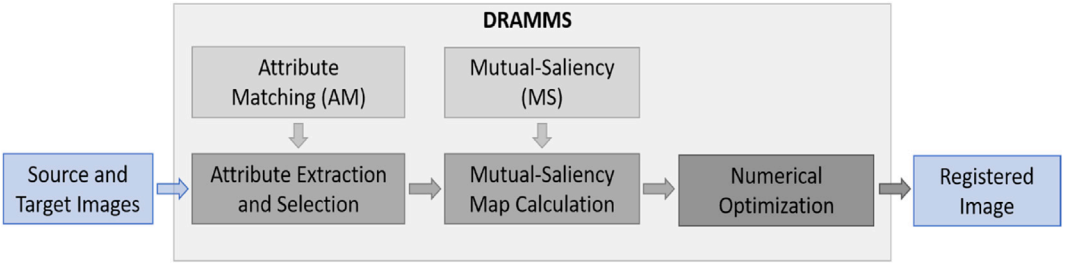


Fig. 2. The original DRAMMS framework (sketch of the framework in (Ou et al., 2011)).

transformation model and the discrete optimization (Komodakis et al., 2008; Glocker et al., 2008) strategy to optimize the deformation in this model.

We follow the original DRAMMS framework (Ou et al., 2011) for attribute extraction, normalization, and selection. Accordingly, attributes are extracted as described in Equation (1) and then normalized into [0, 256] by linearly mapping the attribute values to the minimum and maximum Gabor responses, throughout the image, and across scales and orientations. Given a similarity measure (regardless of the original or the proposed similarities), attributes are automatically selected by maximizing the similarity and reliability of matching (mutual-saliency) using the forward-inclusion-and-backward-elimination (FIBE) feature selection algorithm (Fan et al., 2007; Ou et al., 2009). Overall, attribute extraction, normalization and selection are outside the scope of this paper, but we refer the readers to (Ou et al., 2011) for more details.

3.2. Similarity measure: correlation-based attribute matching

Fig. 3 shows an example of multi-scale and multi-orientation Gabor attributes extracted from the (a) preoperative MR and (b) iUS images. High frequency attributes capture localized fine-scale edges, whereas low frequency attributes capture more regional and coarse-scale edges. Attributes from an orientation of $\pi/2$ highlight vertical edges whereas attributes oriented at π highlight horizontal edges. Regions 1, 2, and 3, in Fig. 3, show that attributes follow a more consistent relationship than image intensities. Overall, attributes at different scales (frequencies) and different orientations allows the extraction of a rich set of texture information at each voxel.

The original DRAMMS uses the Sum of Squared Differences in attribute vectors (aSSD) as the similarity measure,

$$\mathrm{sim}_{\mathrm{aSSD}}\left(A_1^{\star}(u), A_2^{\star}(T(u))\right) = \\ = \frac{1}{1+\frac{1}{D}\left\|A_1^{\star}(u) - A_2^{\star}(T(u))\right\|^2} \in [0,1], \tag{2}$$

where D is the dimension of the attribute vector.

This measure prefers identical attributes that may not always be

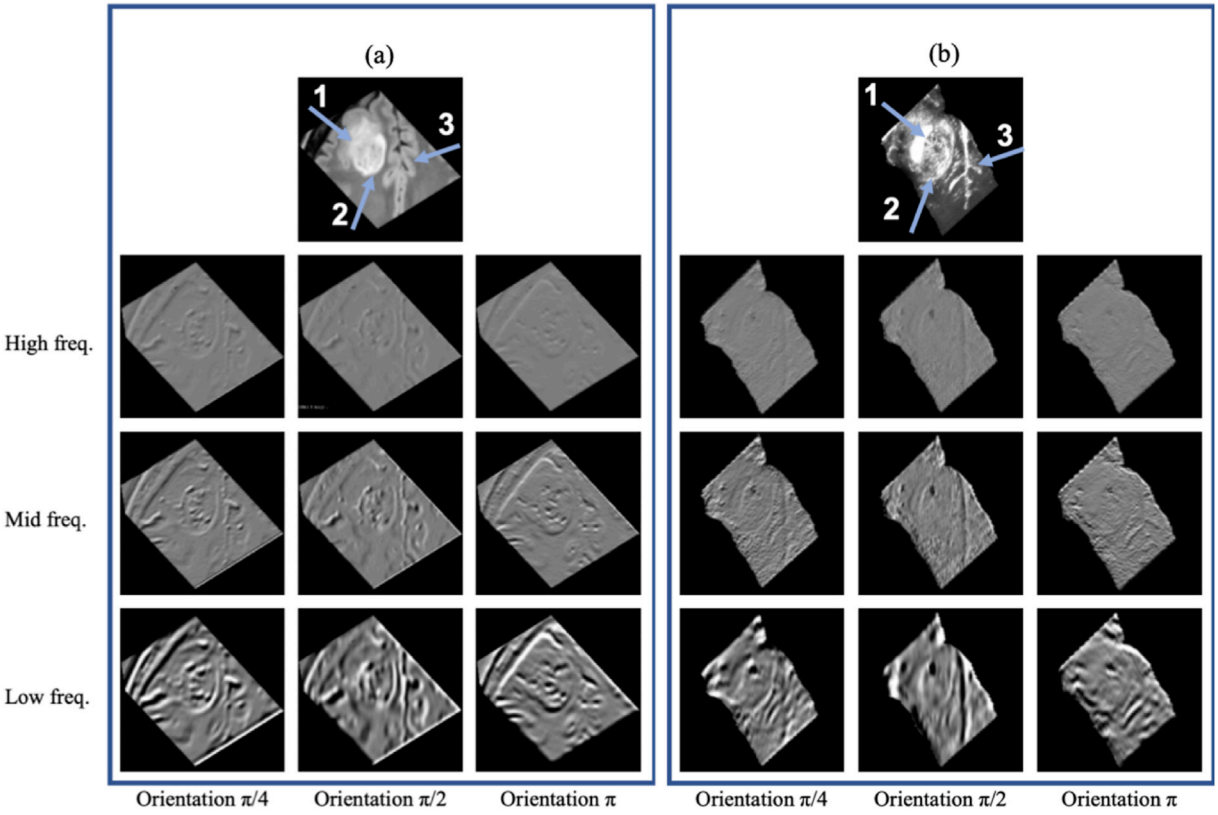


Fig. 3. Multi-scale and multi-orientation Gabor attributes in a patient's MR image (left panel) and US image (right panel) visualizing the benefit of using attributes over intensities for image matching. Regions 1–3 show the inconsistent relationship of MR and US image intensities disqualifying the use of NCC or CR directly on image intensities. Multi-scale and multi-orientation Gabor attributes are extracted from (a) preoperative MR and (b) intraoperative US to better handle the registration challenges.

available in multi-modal image data. Instead, we use the Normalized Correlation Coefficient (aNCC) and the Correlation Ratio (aCR) in attribute vectors. These two measures have been used for image intensity matching but, to the best of our knowledge, it is the first time being used for matching attribute vectors.

$$\mathrm{sim}_{\mathrm{aNCC}}\left(A_1^\star(\mathbf{u}), A_2^\star(T(\mathbf{u}))\right) = \frac{d\sum_{i=1}^{d}\left[A_{1i}^\star(\mathbf{u})A_{2i}^\star(T(\mathbf{u}))\right] - \left(\sum_{i=1}^{d}A_{1i}^\star(\mathbf{u})\right)\left(\sum_{i=1}^{d}A_{2i}^\star(T(\mathbf{u}))\right)}{\sqrt{d\sum_{i=1}^{d}\left[A_{1i}^\star(\mathbf{u})\right]^2 - \left[\sum_{i=1}^{d}A_{1i}^\star(\mathbf{u})\right]^2}\sqrt{d\sum_{i=1}^{d}\left[A_{2i}^\star(T(\mathbf{u}))\right]^2 - \left[\sum_{i=1}^{d}A_{2i}^\star(T(\mathbf{u}))\right]^2}} \tag{3}$$

where $A_{1i}^\star(\mathbf{u})$ and $A_{2i}^\star(T(\mathbf{u}))$ are d-dimensional attributes selected during the optimization process (within the end-to-end DRAMMS framework), and the index i denotes the i-th dimension element in the attribute vector.

In a similar notation, aCR is defined as

$$\mathrm{sim}_{\mathrm{aCR}}\left(A_1^\star(\mathbf{u}), A_2^\star(T(\mathbf{u}))\right) = = 1 - \frac{1}{d\sigma^2}\sum_{k=1}^{K} d_k\sigma_k^2 \tag{4}$$

where

$$\sigma^2 = \frac{1}{d}\left(\sum_{i=1}^{d}\left[A_{2i}^\star(T(\mathbf{u}))\right]^2\right) - m^2, \quad m = \frac{1}{d}\sum_{i=1}^{d}A_{2i}^\star(T(\mathbf{u})) \tag{5}$$

$$\sigma_k^2 = \frac{1}{d_k}\left(\sum_{i\in S_k}\left[A_{2i}^\star(T(\mathbf{u}))\right]^2\right) - m_k^2, \quad m_k = \frac{1}{d_k}\sum_{i\in S_k}A_{2i}^\star(T(\mathbf{u})) \tag{6}$$

$$S_k = \left\{ i \in \{1,\ldots d\}, A_{1i}^\star(\mathbf{u}) \in \max\left(A_1^\star(\mathbf{u})\cdot\left[\frac{k-1}{K}, \frac{k}{K}\right]\right)\right\} \quad \text{and } d_k = |S_k| \tag{7}$$

The attribute elements of a voxel $\mathbf{u}$ in image I_1 are divided into totally K disjoint bins $\{S_k\}$ for $k = \{1,\ldots,K\}$, with cardinality d_k, and the aCR measures the functional dependence between the attribute vectors of two voxels. We keep K equal to 32 according to the original DRAMMS version.

Similar to aSSD ($\mathrm{sim}_{\mathrm{aSSD}}$), aNCC ($\mathrm{sim}_{\mathrm{aNCC}}$) is symmetric with respect to the two images and provides values from zero, indicating no linear dependence between the attribute vectors of two voxels and one, signifying a linearly dependence. Although aCR ($\mathrm{sim}_{\mathrm{aCR}}$) is not symmetric, it uses a value of 0 to indicate that there is no functional dependence of the target voxel's attribute vector given the source voxel's attribute vector and a value of 1 to indicate a purely deterministic dependence. The following simulated experiment is included as a demonstration.

Understanding aNCC and aCR in a Simulated Experiment. We created three data sets to simulate different relationships between the variables. These included a linear relationship ($\mathbf{y}_0 = \mathbf{x}$), a non-linear but monotonic relationship ($\mathbf{y}_1 = \mathbf{x}^2$), and a non-linear but non-monotonic relationship ($\mathbf{y}_2 = \sin(2\boldsymbol{\pi}\mathbf{x})$). They are graphed in Fig. 4 as blue, black, and gray curves, respectively. Panels *a-d* simulate the addition of 0%, 5%, 10%, and 20% noise to the y-axis, respectively. The table at the bottom of the figure displays the four similarity values calculated for each relationship according to aSSD, aNCC and aCR. This experiment demonstrates the following:

i. All aCR values are close to 1 which indicates that aCR can best capture all three types of relationship. In contrast, aNCC can capture both linear relationships and non-linear but monotonic relationships (values close to 1), but it does not capture well the non-linear and non-monotonic relationships (values only at 0.75–0.76). When two variables follow an identical or linear relationship, aSSD is suitable, but it almost completely misses non-linear relationships.
ii. aNCC is not as sensitive to non-linear non-monotonic relationships as aCR. However, aNCC is more stable ($0.77 \rightarrow 0.76$) than aCR, which quickly drops from 1.0 to 0.94 in the presence of noise.

Both aNCC and aCR can likely improve over aSSD for multi-modal MR-to-iUS registration. aCR is more sensitive and aNCC is more stable. Comparison of the experiments based on real MR and iUS data is shown in the Results section.

3.3. Transformation strategy to explicitly handle MR-iUS FOV mismatch

We used an automated transformation strategy to deal with the FOV mismatch between MR and iUS images. Without explicit skull stripping or other frequently-used semi-automated initialization, we begin from rigid registration of raw MR and iUS images, with the original FOVs. The robust rigid module in DRAMMS is used (please see Appendix for details, the "-a 4" argument in the software command). It has improved robustness in the presence of large FOV mismatch by measuring similarity in the overlap instead of the union of two images' FOVs (Ou et al., 2018).

Following the rigid transformation of MRI into iUS space, the second step is to dilate the foreground boundary of the iUS image by a radius of r, and it is used to separate out regions in the rigidly-transformed MR image, as shown in Fig. 5. Dilation ensures that following masking the rigidly-transformed MR image has FOV similar to that of the iUS image. The remaining transformation is left for the third step, the deformable process. We call this three-step strategy the "rigid + masking + deformable" strategy and compare it to "rigid only" and "rigid-+deformable" which does not involve explicit masking.

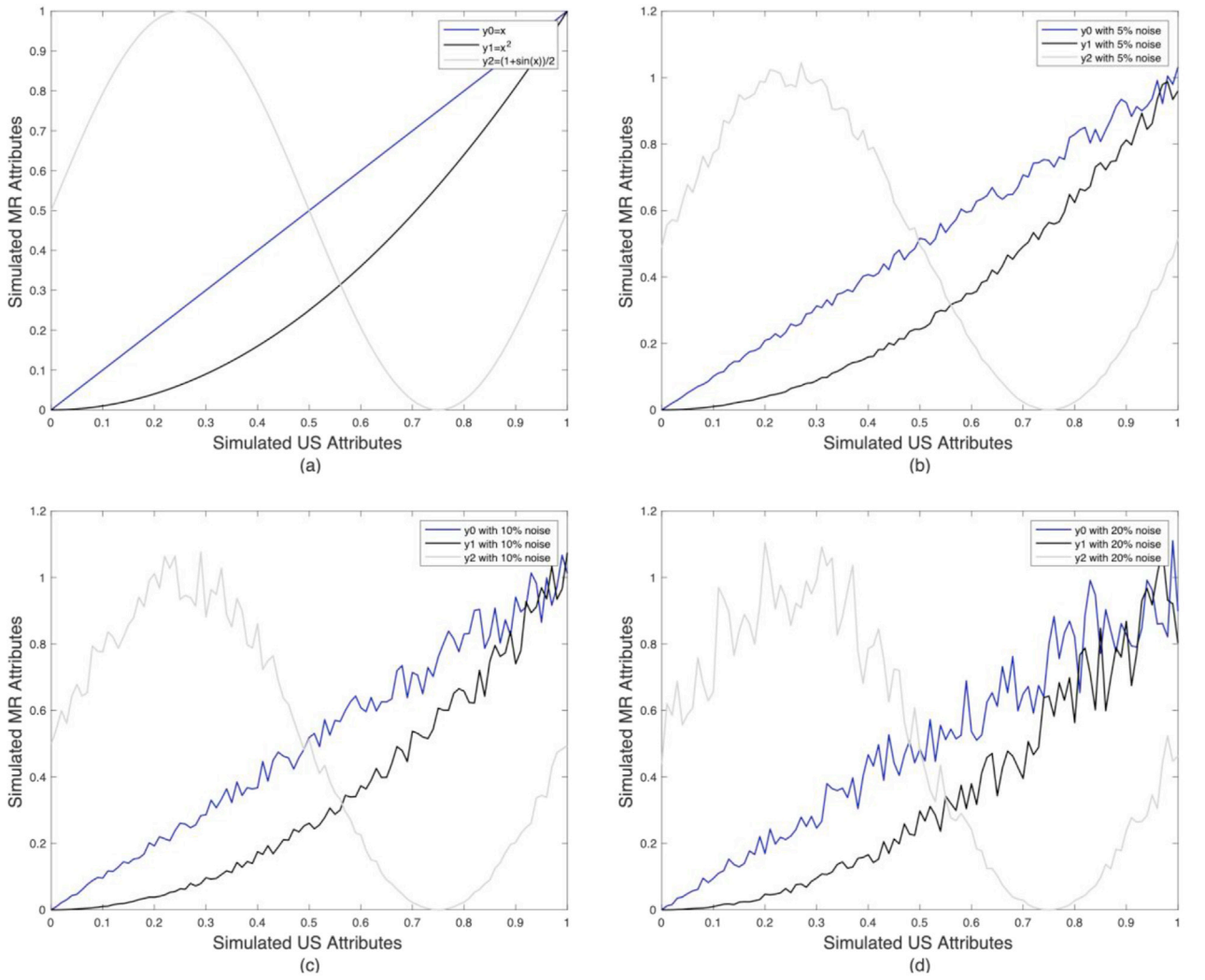
For the RESECT dataset, similar strategies were used (Drobny et al., 2018; Hong et al., 2018) in which the dilation radius r was set to 0 and 15 voxels (equivalent to 3 mm). The parameter r was set to 0 for the BITE dataset in other studies (Rivaz and Collins, 2015; Heinrich et al., 2013; Rivaz et al., 2014a,b; Jiang et al., 2016).

The following sub-section describes experiments used to find the optimal r that could be applied generally in multi-site data.

3.4. Parameter optimization for generality and consistent accuracy in multi-site data

Our goal is to find a fixed set of parameters that are consistent and highly accurate across multi-site MR-iUS clinical data. The key parameters in our framework are the following:

i. Similarity measure: we compare aSSD, which is the original DRAMMS similarity measure, to the proposed similarity measures aNCC and aCR;



Similarity Measure	*Linear relationship* ($x \leftrightarrow y_0$)	*Non-linear monotonic relationship* ($x \leftrightarrow y_1$)	*Non-linear non-monotonic relationship* ($x \leftrightarrow y_2$)
aSSD (Eqn. (2))	1.0 → 0.98 → 0.90 → 0.69	0.23 → 0.23 → 0.22 → 0.20	0.026 → 0.026 → 0.027 → 0.026
aNCC (Eqn. (3))	1.0 → 1.0 → 0.99 → 0.97	0.97 → 0.97 → 0.96 → 0.96	0.77 → 0.77 → 0.76 → 0.76
aCR (Eqn. (4))	1.0 → 1.0 → 1.0 → 1.0	1.0 → 1.0 → 1.0 → 1.0	1.0 → 1.0 → 0.98 → 0.94

Fig. 4. Comparison of aSSD, aNCC, and aCR in simulated data. Blue curves represent a linear relationship ($y_0 = x$), black curves represent a non-linear but monotonic relationship ($y_1 = x^2$), and gray curves represent a non-linear non-monotonic relationship ($y_2 = \sin(2\pi x)$). The four panels represent the addition of a) 0% b) 5% c) 10% and d) 20% noise on the y-axis, respectively. Each cell of the table shows the similarity values calculated for each level of noise in decreasing order according to the method indicated at the top of the column.

ii. Transformation strategy: we compare "rigid only", "rigid + deformable", and "rigid + masking + deformable".
iii. Dilation margin r in the proposed "rigid + masking + deformable" transformation strategy. We vary r from 1 mm to 6 mm to compensate for any registration errors that may arise from rigid registration, which, as experiments show, are typically less than 6 mm (Zimmer et al., 2019).
iv. Smoothness weight λ in Equation (1). We vary λ from 0.6 to 2.5. Smaller values of λ, typically 0.1–0.8, lead to a more aggressive inter-subject deformation (Ou et al., 2014), whereas larger values of λ, 1 to 2.5, correspond to a smoother intra-subject registration.

The similarity measure and transformation strategy are the most important parameters of the list and they are coupled. Therefore, we explore an exhaustive combination of these two parameters, test a range of r and λ values, and plot the accuracy as a function of those parameters in all three datasets.

4. Comprehensive evaluation

4.1. Accuracy metric

All three datasets, BITE, RESECT and MIBS, include expert-annotated corresponding landmarks in the MR and iUS images (see Table 1). Similar to other studies that evaluated MR-to-iUS registration accuracy (Heinrich et al., 2013; Wein et al., 2013; Masoumi, Xiao and Rivaz, 2018a), we use the mean target registration error (mTRE), which is the average distance between corresponding landmarks in each pair of MR and iUS images. Let x_i and x_i' represent expert-annotated corresponding landmark locations in the MR and iUS images, respectively, indexed by i = 1,2, …,N, and T

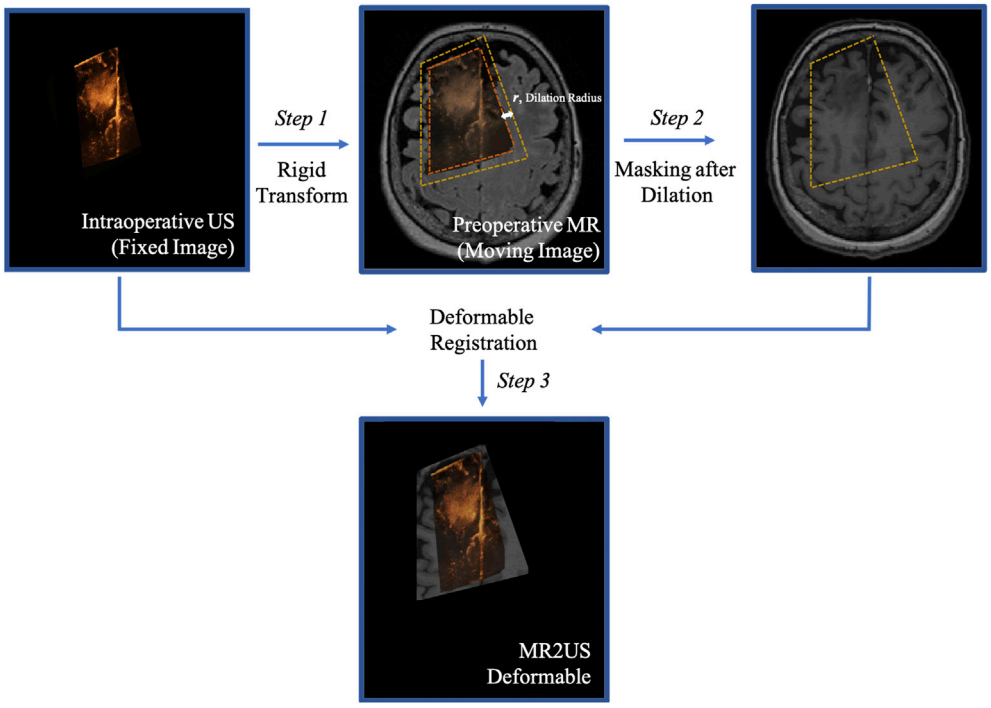


Fig. 5. Sketch of the "rigid + masking + deformable" transformation strategy. Throughout the registration process, iUS corresponds to the fixed image and the preoperative MR corresponds to the moving image. The MR image is rigidly aligned to the iUS (Step 1), masked by the foreground binary mask of the US image with a dilation margin r (Step 2), and the masked rigidly-transformed MR image patch is then deformed onto the iUS image (Step 3).

the computed MR-to-iUS transform given by a registration algorithm. The mTRE of a registration solution is:

$$\text{mTRE} = \frac{1}{N}\sum_{i}^{N}\left\|T(x_i) - x_i^{'}\right\|. \quad (8)$$

4.2. Comparison with fifteen other MR-to-iUS registration algorithms

Table 2 lists fifteen algorithms that have been previously demonstrated in MR-to-iUS registration for the correction of brain shift. These algorithms use a wide variety of image similarity measures, transformation models, and optimization strategies, some of which are also based on machine learning approaches (Zhong et al., 2018; Sun and Zhang, 2018). The algorithms in Table 2 are noted by an asterisk (*) if they have been validated with self-reported accuracies, or a hashtag (#) if they have participated in the MICCAI Challenge 2018 for Correction of Brain shift with Intra-Operative UltraSound (CuRIOUS2018), the first public platform to benchmark MRI-to-iUS registration algorithms for third-party independently evaluated accuracies (please see sub-section 4.3).

Two MR-to-iUS registration tools, SSC/Deeds (Heinrich et al., 2013) and NiftyReg (Drobny et al., 2018), are publicly and freely available at http://www.mpheinrich.de/software.html and https://cmiclab.cs.ucl.ac.uk/mmodat/niftyreg/tree/master, respectively. The SSC/Deeds registration algorithm uses different sets of parameters to register MR and iUS images from different databases. To demonstrate their algorithm in the RESECT database, the authors establish a set of parameters #1 (Heinrich, 2018). The authors establish a different set of parameters #2 (Heinrich et al., 2013) to demonstrate their algorithm in the BITE database. As SSC/Deeds is an open source software, we tested each set of parameters (#1 and #2) on all three datasets. NiftyReg has been previously tested in the RESECT database (Drobny et al., 2018). As NiftyReg is open source, we ran the software, for the first time, in BITE and MIBS databases.

We could not see the LC2 performance in the MIBS database (our proprietary database) as the software is not freely available. LC2 has been previously tested in the BITE database (Wein et al., 2013) and the RESECT database (Wein, 2018) reporting one of the best accuracies in the state-of-the-art MR-to-iUS registration algorithms (please refer to Table 4 for further details). However, LC2 uses different sets of parameters for the different MR-iUS databases, as summarized in Table 3.

4.3. Participation in open challenges held by third-parties

The Correction of Brain shift with Intra-Operative Ultrasound Challenge (http://curious2018.grand-challenge.org) was held in conjunction with the Medical Image Computing and Computer Assisted Intervention (MICCAI) conference in 2018 (Xiao et al., 2019). This Challenge included two phases. In the training phase, the participating teams had access to both MR and iUS images and to the reference landmarks for 22 brain tumor patients. Participating teams trained their algorithms and reported their own mTREs for each clinical case. In the testing phase, participating teams had access to the testing dataset which consisted of 10 new MR-iUS clinical cases but only the landmark locations in the preoperative MR images. Participating teams computed the deformed MR landmark locations in the iUS space, and submitted the locations to the organizers for independent evaluation. The goal of the challenge was to test the accuracy and generality of the algorithm for a cohort of new MR-iUS RESECT clinical cases. cDRAMMS was one of the proposed algorithms.

4.4. Qualitative analysis of registration errors

A radiologist (E.G.) visually checked the locations of 758 landmark pairs one by one, and assigned them into five regions: (i) sulci/gyri, (ii) ventricles, (iii) falx and tentorium, (iv) tumor boundary and (v) "others" where landmarks appear less frequently, including vessels, mid brain, pons, bone, white matter ,and interpeduncular cistern. The information allows us to quantify registration errors by tumor type and by brain region, a topic so far missing in the literature.

Table 2
Major components of the sixteen MR-to-iUS registration algorithms that have been previously tested in iUS-based brain shift correction. An asterisk (*) in some cells of the second column denotes algorithms with self-reported accuracies and a hashtag (#) means that the authors have participated in the CuRIOUS 2018 challenge for the third-party independently reported accuracies. The symbol (–) under the "Publicly Released" column means that the algorithm has not yet been publicly released. Abbreviations: CMA-ES – Covariance matrix adaptation evolutionary strategy; FFD – Free Form Deformation based on cubic B-splines; SGD – Stochastic Gradient Descent; DO – Discrete Optimization; BOBYQA – Bound Optimization by Quadratic Approximation; CCOD – Closest commuting operator distance.

Author, Publication Yr.	Algorithm	Similarity Measure	Deformation Model	Optimization Strategy	Publicly Released
Nigris et al. (2013)	Gradient orientation (*)	Local gradient orientation	Rigid	CMA-ES[a]	–
Rivaz et al. (2014)	SeSaMI (*)	α-MI	FFD	SGD	–
Rivaz et al. (2014)	CoCoMi (*)	Contextual Conditioned Mutual Information	FFD	SGD[a]	–
Rivaz et al. (2015)	RaPTOR (*)	CR	FFD	SGD	–
Jiang et al. (2016)	miLBP (*)	miLBP	FFD	DO	–
Masoumi, Xiao, & Rivaz (2018b)	MARCEL (*)	CR	Affine	SGD	–
Hong et al. (2018)	SS + Demons (#)	Demons (Vercauteren et al., 2007)			–
Drobny et al. (2018)	Symmetric Block-Matching Based Approach (NiftyReg) (#)				Yes
Heinrich (2018)	SSC (#)	Self-similarity context metric	FFD	DO	Yes
Wein (2018)	LC2 (#)	LC2	Demons	BOBYQA	Partially[b]
Shams et al. (2018)	LC2+P (#)	LC2+P[c]	FFD	BOBYQA	–
Masoumi et al. (2018a)	ARENA (*)	CR	Affine	CMA-ES[a]	–
Zimmer et al. (2019)	Laplacian Commutators (*)	CCOD	FFD	SGD	–
Zhong et al. (2018)	MLP (#)	Learning-based Approaches			–
Sun & Zhang (2018)	CNN + STN (#)				–
cDRAMMS (#)		aNCC and aCR	Rigid + Masking + FFD	DO	Yes

[a] The authors used a similar version of stochastic gradient descent optimization (for further details please refer to (Rivaz et al., 2014a,b)).
[b] The similarity calculator is freely released, whereas the whole registration is implemented in the ImFusion SDK. This platform is not open source.
[c] The registration includes a LC2-based rigid registration and a non-rigid step with "LC2+P″ where P is a pixel weighting term (for further details please refer to (Shams et al., 2018)).

5. Results

Section 5.1 reports the computational time of the proposed MR-to-iUS registration algorithm. Section 5.2 presents the optimization of the key parameters. Section 5.3 compares cDRAMMS to other sixteen MR-to-iUS registration algorithms. Section 5.4 discusses the mTREs in anatomical-annotated-landmarks for various brain regions and tumor types.

5.1. Computation time

cDRAMMS runs in a single thread on a single CPU. It requires 5–15 min to register a typical pair of MR and iUS images. We ran cDRAMMS hundreds of times in parallel using the Partners Enterprise Research Infrastructure & Services (ERIS) Unix cluster to optimize the parameter settings across multiple image pairs and multiple patient data sets. It has more than 380 compute nodes, 7000 CPU cores, and a total of 56 TB RAM memory. Jobs submitted to the ERIS cluster were handled by the Load Sharing Facility (LSF) job scheduler for high-performance computing in the Unix environment.

5.2. Optimization of key parameters

5.2.1. Similarity measures and transformation strategies

Fig. 6 shows the mTREs for the various combinations of similarity measures and transformation strategies, explained in Section 3.2 and 3.3. Each subplot refers to a specific dataset: BITE (top left), RESECT (top right), MIBS (bottom left) and all three datasets together (bottom right). The number of clinical cases (samples) in each dataset is shown in bold. For each dataset, each subplot refers to a specific transformation strategy: "rigid", "rigid + deformable", and "rigid + masking + deformable", from left to right, respectively. For each transformation strategy, each color, red, green and blue, corresponds to a different similarity measure, aSSD, aNCC and aCR, respectively. The y-axis shows a box plot with the mTRE for each combination (maximum, minimum, interquartile range and median). The pair-wise T-test is used to determine statistically significant difference. The observations below are generally applicable to all three

Table 3
Parameter settings used by publicly-available tools for different MR-iUS databases. SSC/Deeds and NiftyReg are publicly-available, whereas LC2 only has the similarity metric tool freely-available.

Registration Tool (Reference)	Variable	for BITE Dataset	for RESECT Dataset
SSC (used for BITE in (Heinrich et al., 2013) and for RESECT in (Heinrich, 2018))	Number of coarse-to-fine pyramid levels	3	5
	Control point spacing	{6, 5, 4}mm in 3 levels	{4, 3.5, 3, 2.5, 2}mm in 5 levels
	Search range	{12, 5, 2}mm in 3 levels	{6.4, 6.125, 6, 4.125, 8}mm in 5 levels
	Discretization of search space	[6, 5, 4] intervals in 3 levels	[5, 4, 3, 2, 1] intervals in 5 levels
	Patch size	$3 \times 3 \times 3$ voxels $= 1.5 \times 1.5 \times 1.5$ mm	$3 \times 3 \times 3$ voxels $= 1.5 \times 1.5 \times 1.5$ mm
	Weight for regularizer	0.5	1.6
NiftyReg (used for RESECT in (Drobny et al., 2018))	Coarse-to-fine pyramid levels	–	2
	Iterations per level	–	10 (coarse level), 5 (fine level)
	Block size	–	$4 \times 4 \times 4$ voxels
	% of blocks	–	25% (blocks with highest intensity variations)
LC2 (used for BITE in (Wein et al., 2013); for RESECT in (Wein, 2018))	Patch size	$7 \times 7 \times 7$ voxels	$7 \times 7 \times 7$ voxels
	Transformation	Free-form Cubic B-Spline	Rigid

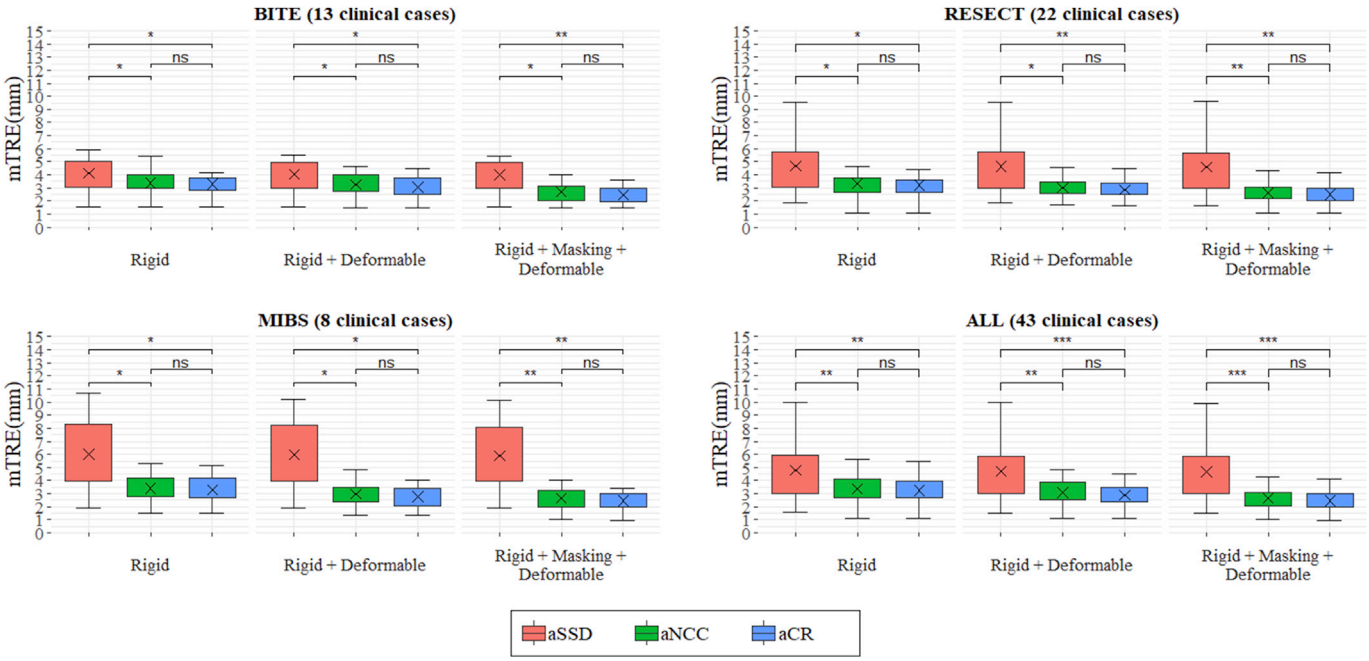


Fig. 6. Distributions of mTRE for different combinations of similarity measures and transformation strategies. Pair-wise T-test is used to determine statistically significant difference: $*$: $p < 0.05$; $**$: $p < 0.01$; $***$: $p < 0.001$, $****$: $p < 0.0001$ and ns : $p > 0.05$ (not significant).

datasets, suggesting consistent performance of cDRAMMS across the multi-site data.

Comparing aSSD, aNCC and aCR, we found:

i. Regardless of transformation strategies, correlation-based attribute matching aNCC (in green) and aCR (in blue) have significantly smaller mTREs than aSSD (in red).
ii. Comparing the two different correlation-based attribute matching: aCR (in blue) leads to significantly lower errors than aNCC (in green) in the BITE (top left) and RESECT (top right) datasets, and lower, but statistically-equivalent errors, in the MIBS dataset (bottom left).
iii. The observations above echo results in the simulated data (Fig. 4). MR and iUS image characteristics are better model if we relax the requirement of equality relationship (aSSD) of attribute vectors to a non-linear relationship (aNCC and aCR) of attribute vectors. aNCC can model consistent non-linear relationships of attributes. When, further relaxted to aCR, it can model inconsistent non-linear relationships among attributes, leading to a further decrease in registration errors.
iv. Comparing transformation strategies, we found that landmark errors decrease when moving from "rigid" to "rigid + deformable" and then to "rigid + masking + deformable".

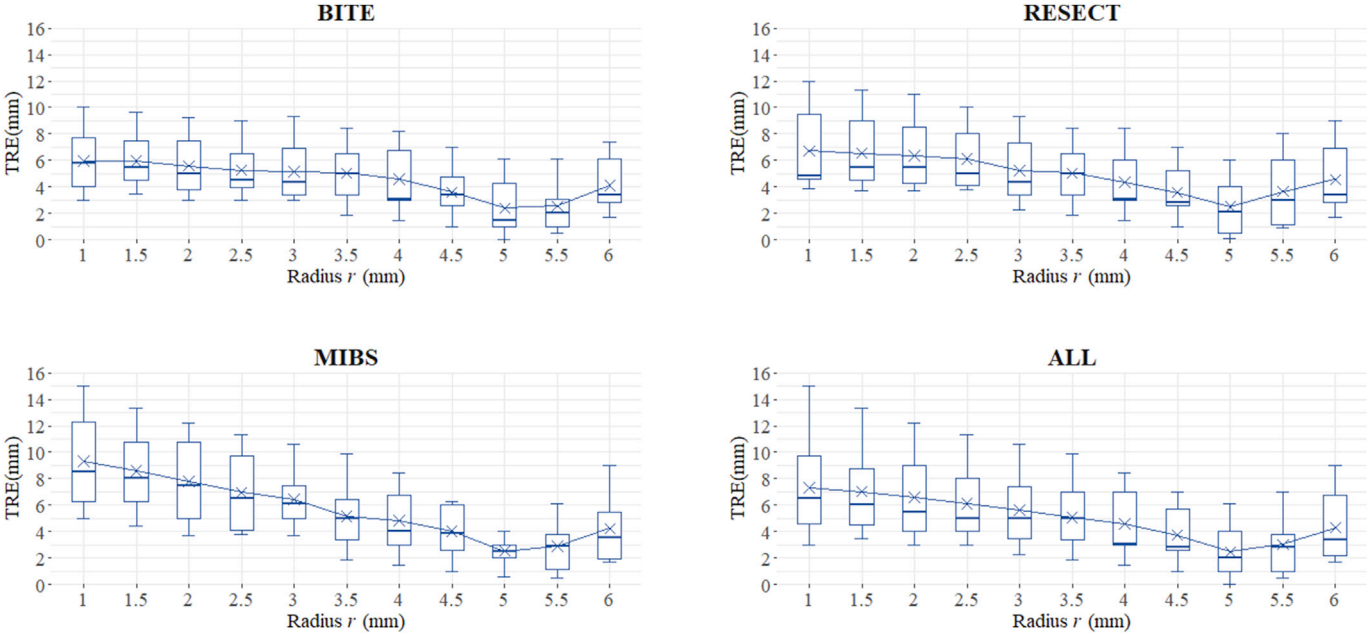


Fig. 7. Registration error (mTRE) as a function of dilation radius r. There are four panels: one for each of three datasets and one for all the data combined. We used aCR as the similarity measure and "rigid + masking + deformable" as the transformation strategy. Pair-wise T-test is used to determine statistically significant difference:$*$: $p < 0.05$; $**$: $p < 0.01$; $***$: $p < 0.001$; $****$: $p < 0.0001$ and ns : $p > 0.05$ (not significant).

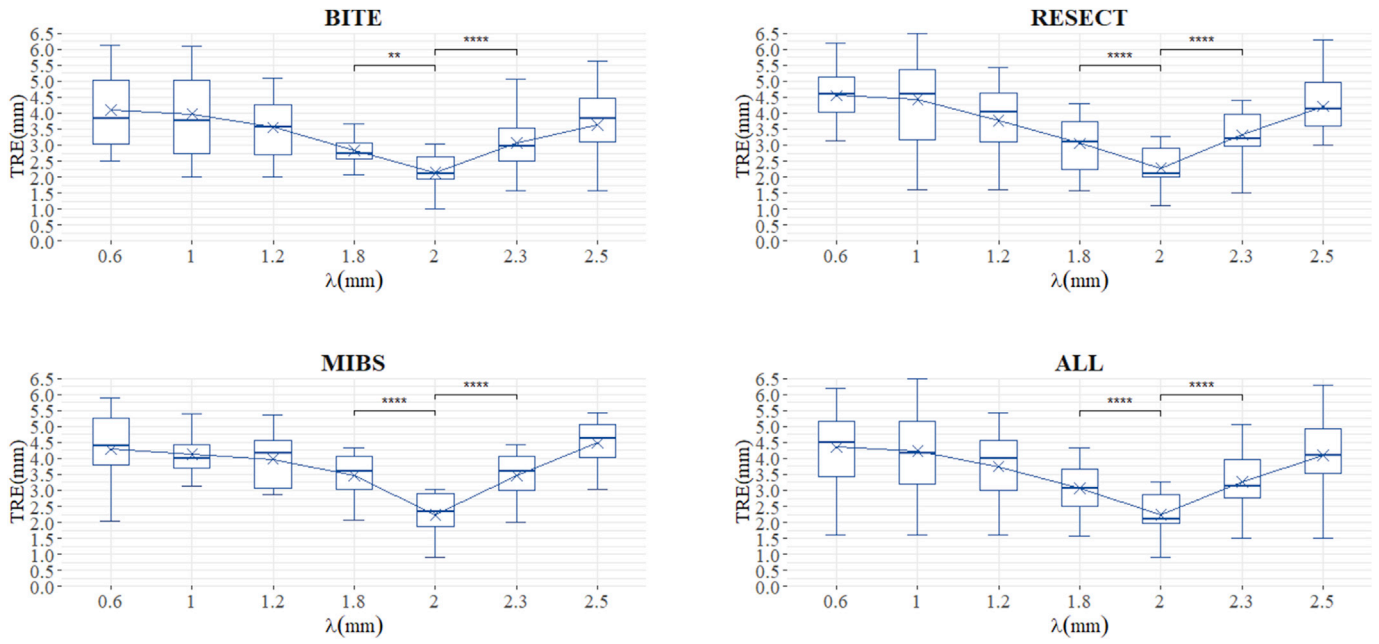


Fig. 8. TRE at various deformation smoothness levels. There are four panels: one for each of three datasets and one for all the data combined. We used aCR as the similarity measure and "rigid + masking + deformable" as the transformation strategy. Pair-wise T-test is used to determine statistically significant difference: $*: p < 0.05$; $**: p < 0.01$; $***: p < 0.001$; $****: p < 0.0001$ and ns: $p > 0.05$ (not significant).

5.2.2. *Dilation radius* **r**

Other studies set r = 0 mm (Rivaz and Collins, 2015; Heinrich et al., 2013; Rivaz et al., 2014a,b; Jiang et al., 2016) or r = 3 mm (Drobny et al., 2018). We fixed aCR as the similarity measure and "rigid + masking + deformable" as the transformation strategy and varied the radius r from 1 mm to 6 mm. Each subplot in Figs. 7 and 8 represents a different dataset: BITE (top left), RESECT (top right), MIBS (bottom left) and all data combined (top right). Each box plot shows the minimum, maximum, interquartile range and median of the TRE of all landmark pairs of all clinical cases of that specific dataset.

i. Landmark errors are sensitive to the dilation radius r. Therefore, minor FOV differences between MR and iUS images will affect the registration accuracy.
ii. The optimal r value appears at 5 mm. The dilation of the US mask allows to compensate for errors in rigid MR-to-iUS registration that are, on average, approximately 3 mm but can be as large as 4.5–5.5 mm (Nigris et al., 2013).

The parameter r reaches the same optimal value in all datasets, suggesting generality of our overall algorithm across multi-site data.

5.2.3. *Smoothness weight* **λ**

The default λ is 0.2 and has been shown to work well for aggressive across-patient registration (Ou et al., 2014). We expect λ = 1.5–2.5 for smoother deformations as MR and iUS images are from the same patient. Fig. 8 shows that λ = 2 is observed as the optimum weight. The optimum is the same for all datasets, suggesting consistency of our algorithm across multi-site data.

5.3. *Comparison with fifteen additional MR-to-iUS registration algorithms*

Table 4 summarizes the mTREs of pre- and post-registration for sixteen MR-to-iUS registration algorithms. In Table 4, cells are empty if the mTREs are not reported in the literature and the software is not publicly available. We used the mTREs reported by authors themselves in 26 out of 32 cells, as noted by citations next to the mTREs in Table 4. The remaining 6 cells in Table 4 correspond to mTREs obtained by running the publicly-available algorithms (NiftyReg and SSC/Deeds) and using the parameters disclosed by the authors (listed in Table 3). For each database in Table 4, the top three ranked algorithms in each subset are noted in bold, underlined, and in italic text, respectively.

Accuracy. SSC/Deeds, LC2 and cDRAMMS present the lowest mTREs among known results in the three datasets. Although the initial mTREs ranges from 0 to 21 mm (an average of 4.18–6.41 mm for the three datasets), the top ranking methods can reduce the errors to approximately 2 mm, and in some datasets, to as low as 1–5 mm.

LC2 decreases the initial mTREs to 1.75 mm and 1.57 mm for the RESECT training (Wein, 2018) and testing (Xiao et al., 2019) datasets, respectively. It reduces the initial mTRE in the BITE database to 2.52 mm (Wein et al., 2013). LC2 uses different sets of input parameters for each dataset. Please refer to Table 3 to see the different sets of parameters. We were not able to report the accuracy of LC2 in the MIBS dataset. LC2 shows the best performance in the RESECT testing dataset when compared to the other six algorithms that participated in the CuRIOUS Challenge 2018. SSC/Deeds reduces the initial mTREs to 1.67 mm and 1.87 mm for the RESECT training (Heinrich, 2018) and testing (Xiao et al., 2019) datasets, respectively, using the set of parameters #1. It reduces the initial mTRE in the BITE database to 2.34 mm (Heinrich et al., 2013), using a different set of parameters, set of parameters #2. Please refer to Table 3 to see the differences in each set. As SSC/Deeds is publicly-available, we ran the software with both sets of parameters (set of parameters #1 and #2) in the MIBS database and obtained a post-registration mTRE equal to 3.61 mm (set of parameters #1) and 2.91 mm (set of parameters #2). cDRAMMS lowers the initial mTREs for all three datasets to 2.08–2.28 mm, using a fixed set of parameters.

Generality. cDRAMMS ranks first in the BITE and MIBS datasets, and third in the RESECT dataset. This is based on using the same set of input parameters for all three datasets. Using flexible parameters for different databases, SSC/Deeds ranks second in both RESECT and MIBS and third in BITE. LC2 ranks second in RESECT's training subset, first in RESECT's testing subset, and fourth in the BITE dataset. MLP, a deep learning algorithm, ranks first in RESECT's training subset, and fifth among those results submitted for the unseen testing subset of the RESECT dataset.

Table 4
Pre- and post-registration mTREs for sixteen algorithms in three datasets. The top three ranked algorithms in each subset are noted in bold, underlined, and in italic text, respectively. The results are mostly obtained from manuscripts published by the authors of the algorithms (please see the citation right after each result), except for SSC/Deeds (set of parameters #1) in the columns of BITE and MIBS; SSC/Deeds (set of parameters #2) in the columns of RESECT and MIBS and NiftyReg in the columns of BITE and MIBS. These results were obtained by running their publicly-available software.

	BITE	RESECT		MIBS
		Training	Testing	
Initial mTRE (mm)	4.18 ± 1.97 Mercier et al., 2012	5.37 ± 4.27 Xiao et al. (2017)	6.41 ± 4.46 Xiao et al. (2019)	6.18 ± 3.38
ARENA	2.82 ± 0.71 Masoumi et al., 2018a	2.77 ± 1.13 Masoumi et al. (2018a)	–	–
CoCoMi	3.22 ± 1.70 Rivaz et al. (2014)	–	–	–
CNN + STN	–	3.91 ± 0.53 Sun & Zhang (2018)	–	–
Gradient Orientation	2.57 ± 0.74 Nigris et al. (2013)	–	–	–
Laplacian Commutators[a]	3.01 ± 1.22 Zimmer et al. (2019)	–	–	–
LC2 (Parameter #1)	–	*1.75 ± 0.62* Wein (2018)	**1.57 ± 0.55** Xiao et al. (2019)	–
LC2 (Parameter #2)	2.52 ± 0.87 Wein et al. (2013)	–	–	–
LC2+P	–	4.6 ± 3.4 Shams et al. (2018)	6.62 ± 2.80 Xiao et al. (2019)	–
MARCEL	–	2.32±0.68[b] Masoumi et al. (2018b)	–	–
RaPTOR	2.65±0.98[c] Rivaz et al. (2015)	–	–	–
miLBP	2.15 ± 1.10 Jiang et al. (2016)	–	–	–
MLP	–	**1.21 ± 0.55** Zhong et al. (2018)	5.72 ± 2.82 Xiao et al. (2019)	–
SeSaMI	2.44 ± 0.40 Rivaz et al. (2014)	–	–	–
SSC (Parameter #1)	*2.18 ± 0.37*	1.67 ± 0.54 Heinrich (2018)	1.87 ± 0.51 Xiao et al. (2019)	3.61 ± 0.82
SSC (Parameter #2)	2.34 ± 0.52 Heinrich et al. (2013)	2.52 ± 0.60	–	2.91 ± 0.88
SS + Demons	–	5.60 ± 3.94 Hong et al. (2018)	6.55 ± 4.55 Xiao et al. (2019)	–
NiftyReg	2.67 ± 0.58	2.90[d] Drobny et al. (2018)	3.24 ± 3.63 Xiao et al. (2019)	*3.18 ± 0.62*
Ours (cDRAMMS)	**2.08 ± 0.37**	2.28 ± 0.71	*2.19 ± 0.87* Xiao et al. (2019)	**2.24 ± 0.78**

[a] This group used 9 (numbers 2, 3, 5, 6, 10–14) out of 13 clinical cases.
[b] MARCEL was validated in 5 out of 22 RESECT cases. They reported an initial mTRE equal to 5.13 ± 2.78 mm (N = 5).
[c] RaPTOR was validated in a previous version of the BITE dataset in which the mTRE pre-registration was 5.92 ± 3.15 mm and the number of landmarks per case ranged from 9 to 21.
[d] This group reported only the mean TRE for all patients in the RESECT training subset.

MLP shows a far lower mTRE (1.21 ± 0.55 mm) in the RESECT training subset but the errors in the RESECT testing subset are 5.72 ± 2.82 mm. The lowest mTRE averages among the three datasets are 1.57 mm for LC2, 1.67 mm for SSC, and 2.08 mm for cDRAMMS. The highest mTRE averages among the three datasets are 2.52 mm for LC2, 3.61 mm for SSC, and 2.28 mm for cDRAMMS.

5.4. Additional analysis of results

5.4.1. Qualitative assessment by physicians

Fig. 9 shows the MR-to-iUS registration results for three different clinical cases: RESECT (Case #23), BITE (Case #6) and MIBS (Case #1). First, second, third, and fourth rows shows the preoperative MR, the intraoperative US, the superposition of both images before deformable registration and the superposition of both images after deformable registration, respectively. Blue and white arrows point out the tumor boundaries, yellow arrows the falx, and red arrows the sulci. One can see that these regions are better aligned following cDRAMMS registration (bottom row) compared to those before registration (third row).

A neurosurgeon in training (P.U) and an experienced neurosurgeon (W.E.) with two and nine years of clinical practice, respectively, and each with four years of intraoperative ultrasound experience, visually assessed the results of the registration in different brain regions such as the tumor boundary, sulci, vessels, choroid plexus, falx, and ventricles. Table 5 shows their analysis that no registrations were "bad" (grossly visible misregistration), 44% or 49%, respectively, of registrations were "good" (minor visible misalignments), and 51% or 56%, respectively, of the registrations were "great" (negligible or nearly undetectable misregistration), respectively for each surgeon.

5.4.2. Corrected brain shift

Fig. 10 compares the initial and final mTRE for all 43 patients (13 brain tumor patients from the BITE database, in red, 22 brain tumor patients from the RESECT database, in blue, and 8 brain tumor patients from the MIBS database, in green). Regardless of clinical case and database, and regardless of the scale of the initial shift (x-axis, ranging from 0 to 21 mm), the results show the consistent low mTRE (y-axis) after registration ranging from 0 to 4 mm and, for most patients, around 2 mm.

5.4.3. Distribution of registration errors according to brain regions

Fig. 11 presents color coded registration errors in various brain regions of nine brain tumor patients, three from each dataset. Color

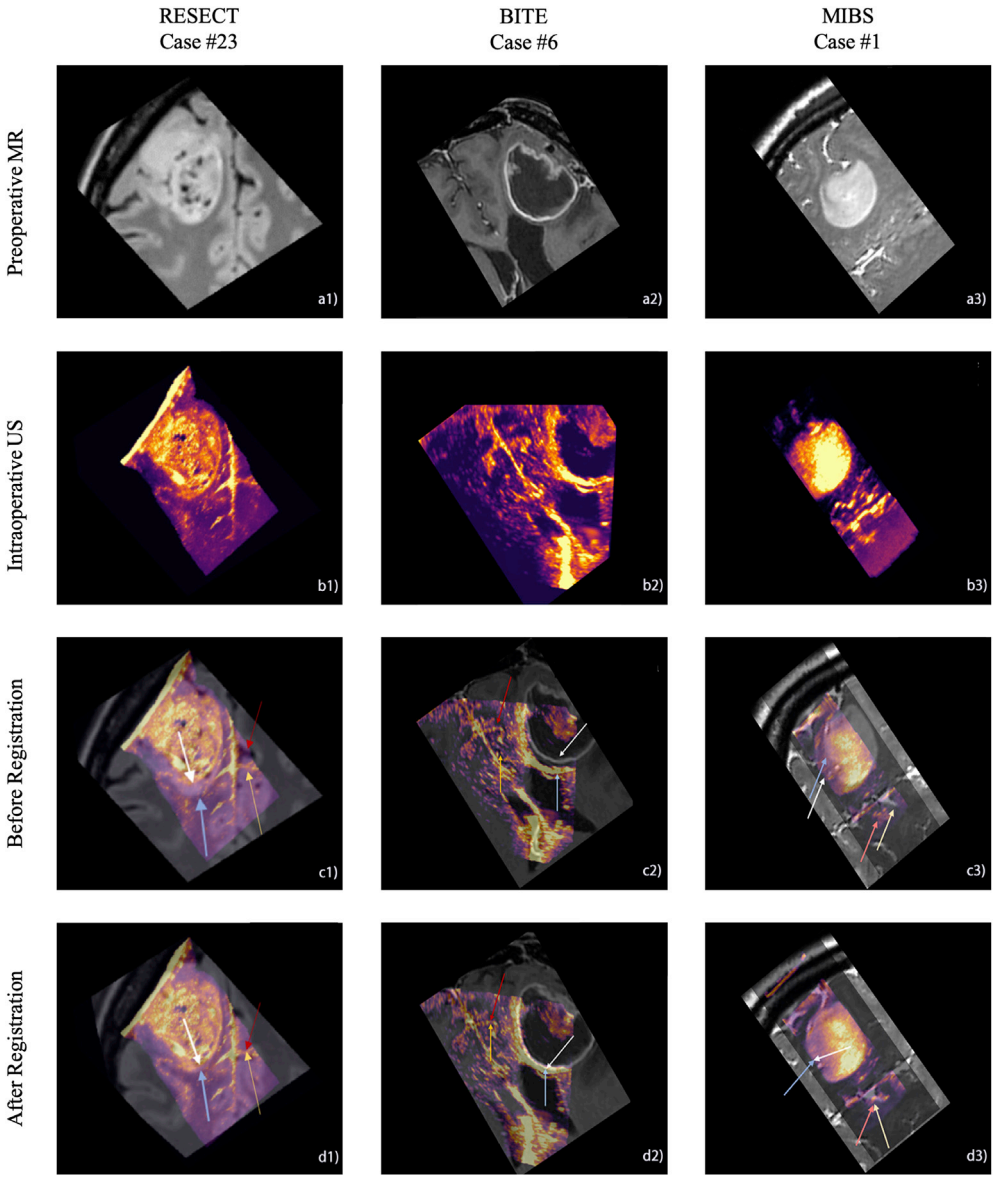


Fig. 9. Preoperative MR to iUS registration in three brain tumor cases: RESECT Case #23, BITE Case #6 and MIBS Case #1. Row designations: a) preoperative T2-FLAIR MR; b) iUS; c) MR superimposed on iUS before registration; d) MR superimposed on iUS after registration. Arrows indicate the alignment of the tumor boundary (blue and white) and the sulci/falx (red and yellow) between the two modalities, before and after registration.

Table 5
Visual grading of the alignment by two neurosurgery fellows. A total of 43 clinical cases from three different MR-iUS datasets were assessed.

Grader	Grading		
Neurosurgeon	Bad	Good	Great
P.U.	0	21 (49%)	22 (51%)
W.E.	0	19 (44%)	24 (56%)

encodes the magnitude of errors according to the color bar on the right: blue for small (0–1 mm) and red for large errors (3 mm or above). In the selected patients, the errors are smaller in sulci and gyri (0.5–1.5 mm) than in tumor boundaries (2–3.5 mm).

This is echoed in Fig. 12, which shows the statistics for pre- and post-registration landmark errors in all datasets according to brain region. Each subplot refers to a specific dataset: BITE (top left), RESECT (top right), MIBS (bottom left) and all three datasets together (bottom right). Different colors represent different brain regions (sulci/gyri in green, tumor boundary in blue, falx and tentorium in red and ventricles in yellow) with dark and light for pre- and post-registration errors, respectively. In the tumor boundary, the post-registration average landmark error is approximately 2.7 mm; in ventricles, approximately 2.7 mm; in falx and tentorium, approximately 2 mm; and in sulci and gyri, approximately 1.4 mm.

5.4.4. *Distribution of errors according to tumor grade*

Fig. 13 shows the statistics for pre- and post-registration landmark errors in all datasets according to tumor type. Each subplot refers to a specific dataset: BITE (top left), RESECT (top right), MIBS (bottom left) and all three datasets together (bottom right). Different colors represent different grades of brain tumor (green for LGG, blue for HGG and yellow for metastasis). High-grade gliomas have higher landmark errors after registration, averaging 3.1–3.3 mm, than low-grade gliomas, which average approximately 1.25 mm. The registration error for metastatic brain tumor is approximately 2.1 mm.

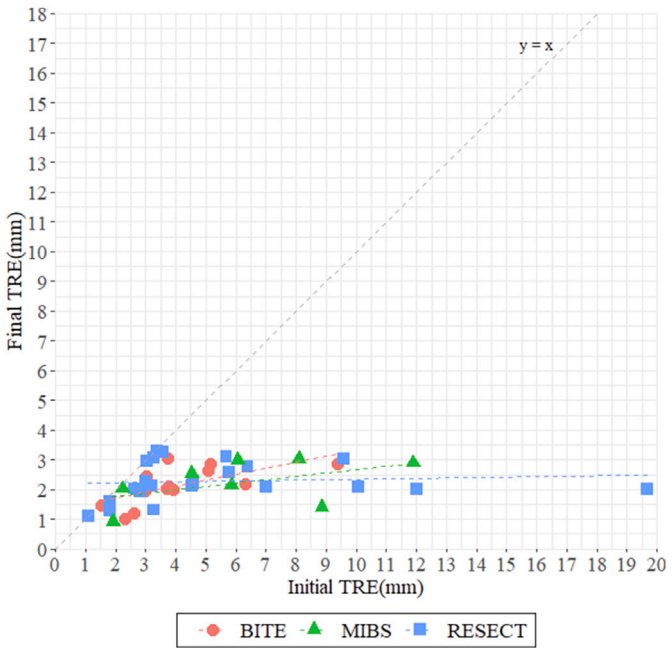


Fig. 10. Brain shift represented by the pre- and post-registration mTREs. Red circles, green triangles, and blue squares represent brain tumor cases from the BITE, MIBS, and RESECT datasets, respectively.

6. Discussion

Accuracy and generality are two major issues in registering preoperative MR to intraoperative US images for neurosurgical guidance. This study demonstrates that using correlation-based attribute matching and explicit normalization of the FOVs of preoperative MR and iUS improves the registration accuracy. These new features are now publicly available in the DRAMMS software (https://www.nitrc.org/projects/dramms). The recent availability of multi-site data allows us to optimize key parameters and test the generality of the proposed algorithm. We show that a fixed set of optimized parameters can have consistent performance in multi-site data. We found that landmark errors are larger in the tumor boundary and tentorium than in sulci and gyri and larger in HGG brain tumor patients than in LGG or metastatic brain tumor patients.

Both simulated and real clinical data show that the correlation ratio of attributes (aCR) can match brain regions with inconsistent attribute values. Normalized correlation coefficient of attributes (aNCC) can match attributes of both linear and non-linear relationships so long as the relationship is consistent or monotonic.

According to the post-registration landmark errors, attribute-matching based algorithms, including LC2, SSC/Deeds, NiftyReg and cDRAMMS, are among the top methods in accuracy and generality. This highlights the merit of using high-dimensional texture attributes, beyond one-dimensional intensity of a voxel. Texture attributes render each voxel more distinctive than the intensity information. Our current framework is open to other attributes. Other example attributes include but are not limited to moments (Shen and Davatzikos, 2002), local histogram analysis (Shen, 2007), local entropy and Laplacian attributes (Wachinger and Navab, 2012), rotationally invariant difference-of-Gaussian filter responses (Toews and Arbel, 2009), uniform spherical region descriptor (Liao and Chung, 2012) and rotation-invariant attributes based on alpha stable filter banks (Liao and Chung, 2010).

In order to normalize the MR image's FOV to that of the iUS image, we dilated the iUS binary mask to compensate for errors in the initial rigid alignment. Errors in rigid MR-to-iUS alignment are approximately 3 mm, but can be as much as 4.5–5.5 mm or even higher (Nigris et al., 2013). We believe this explains our observation that the optimal dilation of the FOV of the iUS image should be approximately 5 mm in the datasets studied here.

There is ongoing debate as to whether MR to pre-resection US registration should be rigid (Coupé et al., 2012; Letteboer et al., 2003; Nigris et al., 2013), affine (Masoumi et al., 2018a; Masoumi et al., 2018b), or deformable (Rivaz et al., 2015; Wein, 2018; Heinrich, 2018). Our results show that predurotomy brain shift is a largely rigid and highly regularized (smooth) non-rigid process.

Generality remains a major issue for multi-site data. As Table 4 shows, LC2 is highly accurate for the RESECT dataset – the average mTREs across RESECT training and testing patients are 1.57 ± 0.62 and 1.75 ± 0.55 mm, respectively. It uses a second set of parameters to maintain a high level of accuracy in the BITE dataset, 2.52 ± 0.87 mm. SSC/Deeds achieves an mTRE of 1.67 ± 0.54 mm and 1.87 ± 0.51 mm (the second-lowest error among algorithms) in the RESECT training and testing datasets, respectively and 2.18 ± 0.37 mm (the third-lowest error) in the BITE dataset. However, the same parameters lead to a larger mTRE, at 3.61 ± 0.82 mm for the MIBS dataset. Changing to another set of parameters allows SSC/Deeds to achieve a lower mTRE (2.91 ± 0.88 mm), which ranks the second lowest errors in the MIBS dataset. In contrast, cDRAMMS scores relatively consistently in all three datasets while using a fixed set of parameters. The average mTRE equals 2.08 mm in the BITE dataset (lowest error among algorithms), 2.19–2.28 mm in RESECT (training and testing, respectively, third lowest errors among methods), and 2.24 mm in MIBS (lowest error among methods). The standard deviation is 0.37 mm, 0.71–0.87 mm, and 0.78 mm, respectively in those datasets, which shows the stability across all patients. It was further demonstrated that cDRAMMS registration is stable within a narrow range (1–3.5 mm) across patients and datasets, irrespective of landmark distribution, tumor grade, and the initial mTRE (sometimes as great as approximately 20 mm). Prior to the dura opening, patient positioning, anesthesiologic drugs (which affect cerebrovascular volume), and physical brain retraction (e.g., with spatulas) cause the brain to deform, which invalidates the estimated rigid transformation by neuronavigation systems. Even for predurotomy US images, the initial maximum target registration error is 9.38 mm (Case #3), 19.68 mm (Case #12) and 11.89 mm (Case #2), respectively for BITE, RESECT and MIBS. cDRAMMS was able to reduce the maximum TREs to 2.9 mm, 3.6 mm and 3 mm, respectively for each dataset.

The deep learning-based MLP algorithm shows great promise in the RESECT training dataset as it was able to reduce landmark errors to an average of 1.21 mm (Zhong et al., 2018). However, landmark errors increased to 5.72 mm in the testing dataset. It leaves the question open as to whether refined network designs and greater availability of training data will improve accuracy of deep learning-based algorithms as has been observed in other across subject healthy brain image registration tasks (Yang et al., 2017; Wang et al., 2017; de Vos et al., 2017). This future direction requires not only algorithm development, but also multi-site collaboration for releasing more MR-iUS data. However, when large-scale multi-site data is available, it becomes difficult for experts to annotate thousands of landmark pairs. The proposed cDRAMMS algorithm can offer voxel-wise correspondences at a consistent level of accuracy for hundreds of thousands of voxels. Therefore, future deep learning algorithms can first learn the large number of voxel-wise correspondences found by an highly accurate algorithm, and then fine tune the deep learning neural network on expert-annotated landmark pairs.

In addition to our efforts to achieve accuracy and generality, our evaluation of sixteen algorithms across three datasets for iUS-based brain shift correction is the most comprehensive completed so far. The CuRIOUS 2018 Challenge is the first independent evaluation of MR-to-iUS registration algorithms in a well-organized open platform that consists of training and testing data (Xiao et al., 2019) and our work references results from that Challenge (Table 4, results for the RESECT dataset). We added to it by: (a) using two additional datasets to evaluate not only the accuracy but also the generality of the sixteen algorithms and (b) further assigning brain regions to manually-annotated landmark

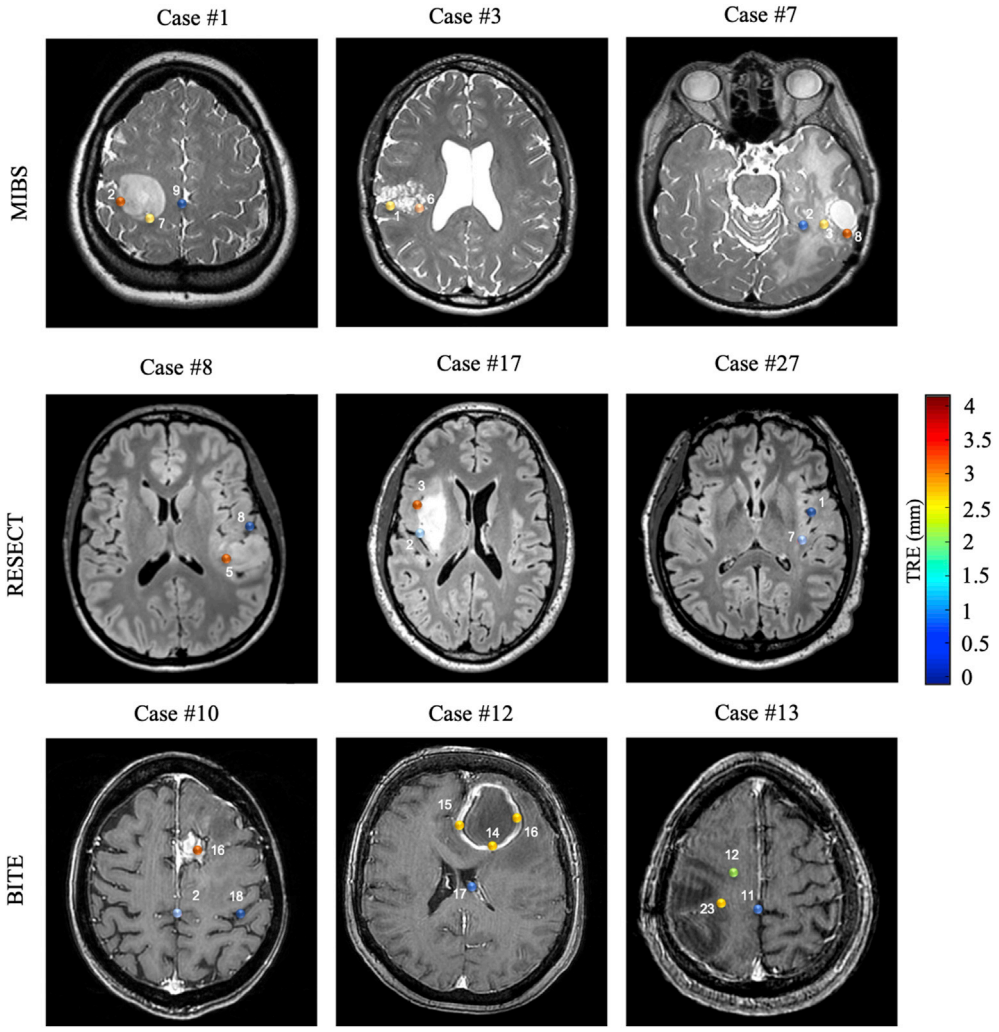


Fig. 11. Landmark errors in representative axial MR slices for nine brain tumor patients from three datasets. Color encodes the magnitude of errors: blue for small (0–1 mm) and red for large errors (3 mm or above).

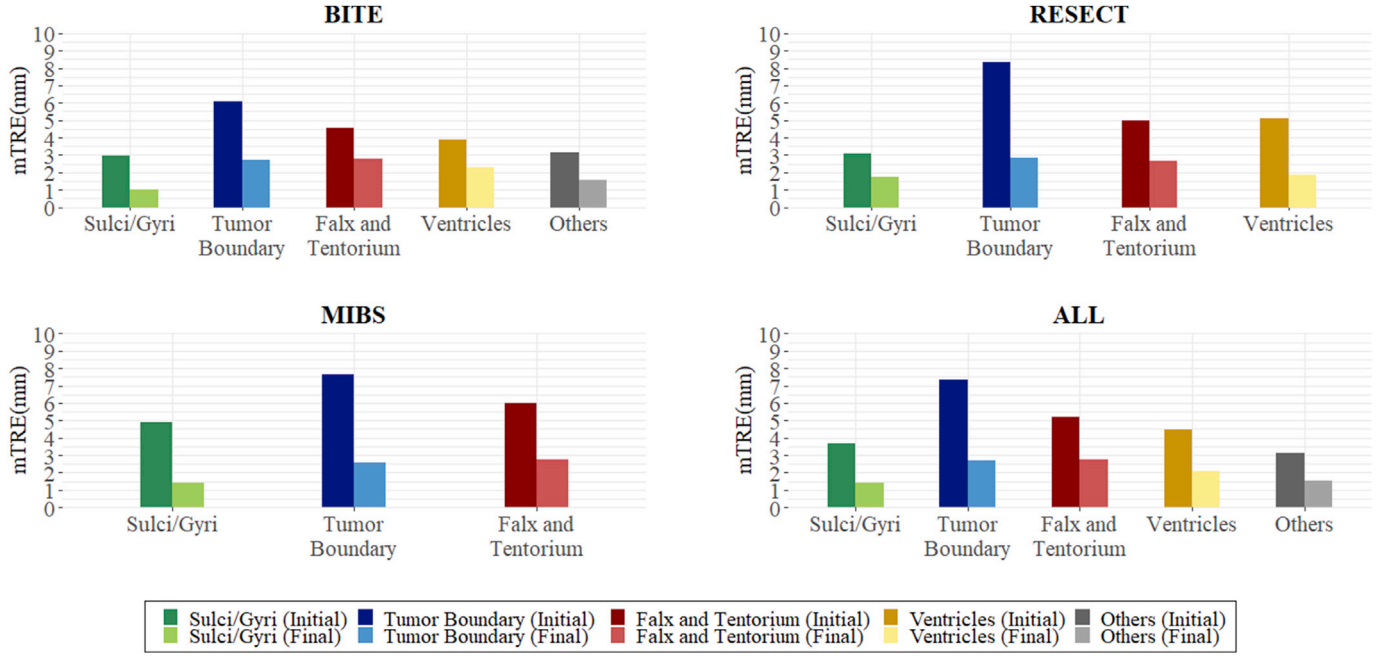


Fig. 12. Landmark errors in various brain regions. Different colors indicate different brain regions, with dark and light representing pre- and post-registration errors, respectively.

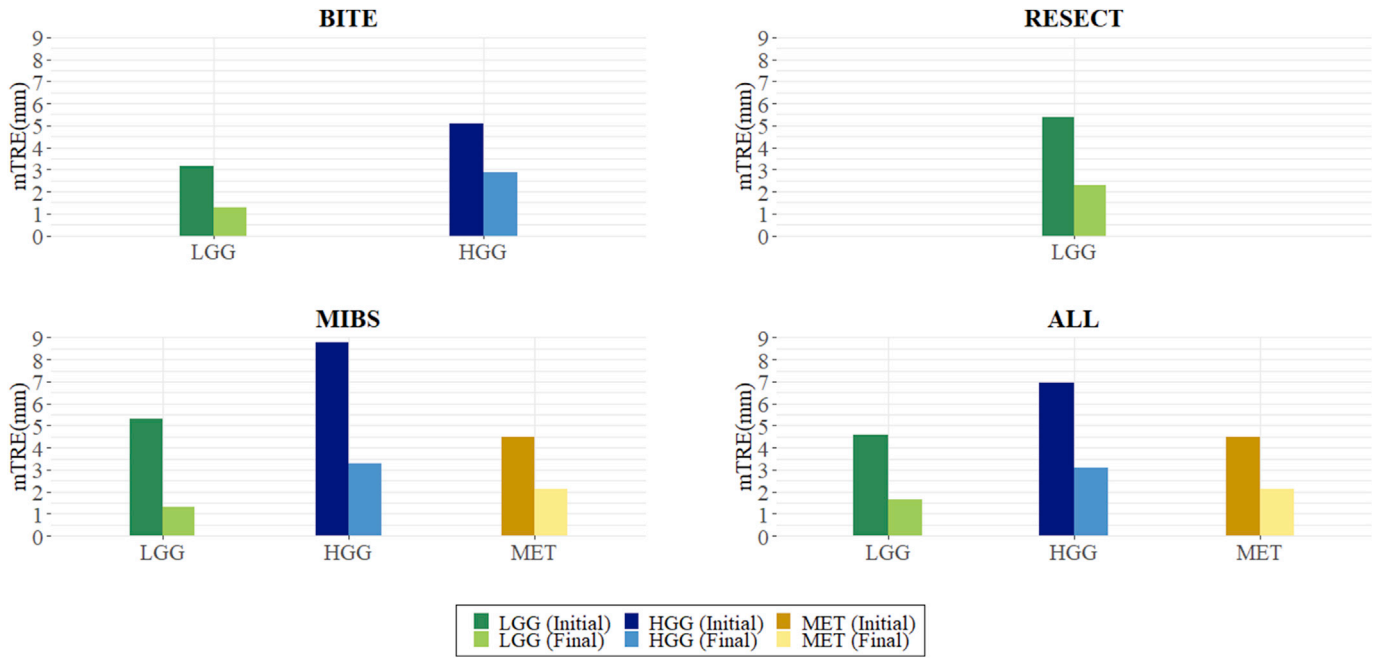


Fig. 13. Pre- and post-registration landmark errors for different grades of brain tumors.

pairs. This led to the first report of registration accuracy according to the extent of brain shift (Fig. 10), brain regions (Figs. 11 and 12), and tumor grade (Fig. 13).

Our evaluation includes sixteen registration algorithms/tools that have been previously demonstrated in iUS-based brain shift correction. Many general-purpose algorithms such as ANTs (SyN) (Avants et al., 2008), IRTK (Rueckert, 1999), Diffeomorphic Demons (Vercauteren et al., 2009), FNIRT (Andersson et al., 2010), ART (Ardekani et al., 2005) or others evaluated in (Ou et al., 2014; Klein et al., 2009) were not considered in this work. MR-to-iUS registration is a challenging task due to the vast differences in physics of iUS and MR imaging modalities and requires modifications of general-purpose algorithms. For example, NiftyReg, DRAMMS and SSC/Deeds are all general-purpose algorithms, but they have been adapted to the MR-to-iUS registration task by adding a FOV normalization step, followed by dilating the MR by some defined margin (Drobny et al., 2018; Heinrich et al., 2013). Specific adaptions

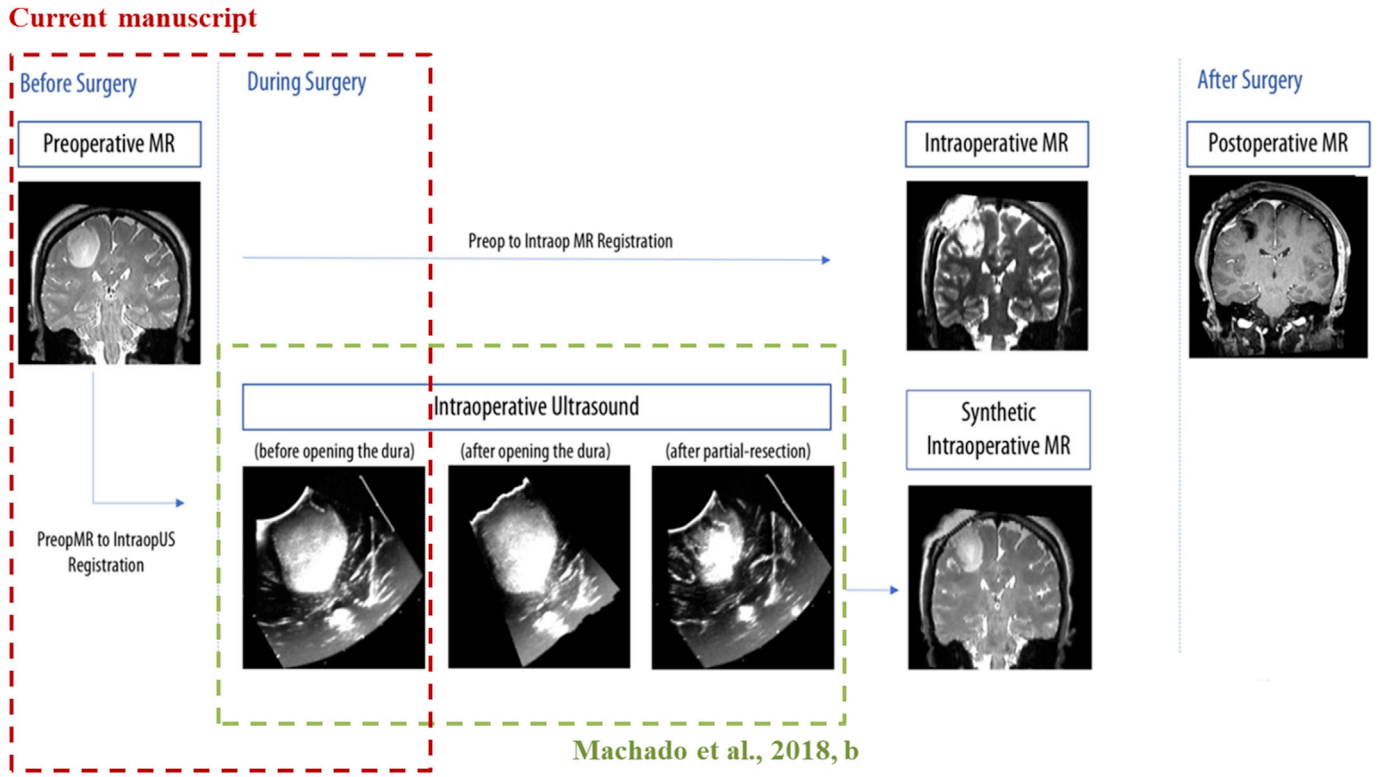


Fig. 14. Sketch of our ultimate and future goal. This system is composed by: i) preoperative MR to pre-resection US registration; ii) pre-resection to post-resection US registration (Machado et al., 2018a,b) and iii) preoperative to intraoperative MR registration and comparison to a "synthetic" intraoperative MR obtained with iUS.

(parameters, preprocessing steps) vary across these general-purpose algorithms and across datasets (Drobny et al., 2018; Heinrich et al., 2013). Therefore, in this paper, we included general-purpose algorithms only if they had already been adapted and optimized specifically for MR-to-iUS registration – this means, NiftyReg (adapted from (Modat et al., 2014) to (Drobny et al., 2018)), SSC (adapted from (Heinrich et al., 2013) to (Heinrich, 2018)), and DRAMMS (adapted from (Ou et al., 2011) to this manuscript). The remaining 15 algorithms that were evaluated in this paper are specifically designed for MR-to-iUS registration tasks. The inclusion criterion is similar to those in other evaluation studies for MR-iUS registration (Drobny et al., 2018) (Heinrich, 2018) (Wein, 2018).

Speed is an additional factor to be considered in order to assess clinical utility of an algorithm. cDRAMMS takes tens of minutes (single-thread single-CPU desktop) whereas several other algorithms may require only two to 5 min (Masoumi et al., 2018a), one to 2 min (Shams et al., 2018), or even tens of seconds (Heinrich, 2018; Zhong et al., 2018). cDRAMMS uses cubic B-spline-based free form deformation and discrete optimization and recent studies have shown that the same components can be optimized so that the overall computation time is reduced to only a few seconds (Jiang et al., 2016; Heinrich, 2018). This is an additional area for future research.

A further limitation and subject for future work is how to best evaluate registration accuracy. This study, like most brain tumor MR-iUS registration studies, quantifies distances between manually annotated landmark correspondences. However, expert landmarks are subject to human error, are non-uniformly distributed in the image, and may, in extreme cases, not be representative of errors in a clinical target such as the tumor edges (Wein, 2018; Fitzpatrick, 2009). A more rigorous validation would be the exploration of the Dice overlap on the actual segmentation of the tumor and surrounding key regions (Wein, 2018), and the properties of the deformation field (Fitzpatrick, 2009; Rohlfing, 2012). Evaluating registration accuracy in terms of the overlaps of key corresponding regions/structures is the norm in brain MRI-to-MRI registration (Ou et al., 2014; Klein et al., 2009). However, segmenting iUS images is a challenging problem that is still unsolved and thus not available to us for this study. Automatic segmentation of US is an important area of future work.

Recent progress has been made toward improving the clinical utility of MR-to-iUS registration algorithms. A research direction is filling the gap between the laboratory-based algorithm development and actual clinical bedside navigation system. CustusX (Askeland et al., 2016), published in 2016, represents a step in this direction. In a 2018 study (Iversen et al., 2018), the authors of CustusX teamed up with authors of the LC2 algorithm, and tested the performance of a commercial MR-to-iUS registration software, ImFusion GmbH (Munich, Germany), on the CustusX platform. They reported a mean landmark error mTRE of 4.47 mm (median at 2.72 mm) (Iversen et al., 2018). Future work includes disseminating laboratory-developed software into clinical platforms for further evaluation and improvement in clinical settings.

The proposed registration is part of a larger system for image-guided neurosurgery, sketched in Fig. 14, where the ultimate goal is to compensate for brain shift using iUS and iMRI modalities. The system contains at least three key registration components: (a) pre-operative MRI to intra-operative, pre-resection US (Fig. 14, red box), the focus of this manuscript as it typically sets the basis of brain shift for the rest of the surgery; (b) intra-operative pre-resection US and intra-operative post-resection US (Fig. 14, green box), the focus of our recent publication (Machado et al., 2018b); and finally (c) preoperative MR and intra-operative MR followed by comparison with the synthetic intraoperative MR (obtained from intraoperative US images), which will be our future direction.

7. Conclusions

We present a correlation-based attribute matching algorithm for MR-to-iUS registration in image-guided neurosurgery and show its generality and accuracy in multi-site clinical data. Our work shows promise for an algorithm applicable to multi-site data, that achieves consistent accuracy, using a fixed set of parameters. The generality of this algorithm is based on (a) extracting multi-scale and multi-orientation attribute vectors, (b) using robust correlation-based similarity measures on those attribute vectors, (c) explicitly handling field-of-view differences, and (d) thoroughly optimizing key parameters in multi-site data. This brings us a step closer to the clinical use of registration for neurosurgery assistance. It motivates the continuing development of GPU-powered, faster, more sophisticated learning-based algorithms for larger-scale multi-site data.

Acknowledgements

The authors would like to acknowledge the financial support to IM from the Portuguese Foundation for Science and Technology under the references PD/BD/105869/2014 and IDMEC/LAETA UID/EMS/50022/2013. The authors also acknowledge the financial support to YO from the Faculty Development Award of Harvard Medical School and Boston Children's Hospital, and the Career Development Award from Thrasher Research Fund (THF13411). This study was also supported by the National Institutes of Health Grants P41-EB015898, P41-EB015902 and R01-NS049251, and a Natural Sciences and Engineering Research Council of Canada Discovery grant. We thank the organizers of the CuRIOUS 2018 Challenge, Drs. Ingerid Reinertsen, Hassan Rivaz, Yiming Xiao and Matthieu Chabanas, for organizing the first open challenge for MR-US registration, and for their insightful discussions. We thank Enterprise Research Infrastructure & Services (ERIS) at Partners Healthcare for their in-depth support and for the provision of the ERIS-ONE high performance cluster computing environment.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.neuroimage.2019.116094.

Conflicts of interest

The authors declare that there are no conflicts of interest.

Compliance with ethical standards

All procedures performed in these studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committees and with the 1964 Declaration of Helsinki and its amendments or comparable ethical standards.

Informed consent

Informed consent was obtained from all participants included in the study.

References

Andersson, J.L.R., Jenkinson, M., Smith, S., 2010. Non-linear Registration, Aka Spatial Normalisation. FMRIB Technical Report.

Ardekani, B.A., Guckemus, S., Bachman, A., Hoptman, M.J., Wojtaszek, M., Nierenberg, J., 2005. Quantitative comparison of algorithms for inter-subject registration of 3D volumetric brain MRI scans. J. Neurosci. Methods. https://doi.org/10.1016/j.jneumeth.2004.07.014.

Askeland, C., Solberg, O.V., Bakeng, J.B.L., Reinertsen, I., Tangen, G.A., Hofstad, E.F., et al., 2016. CustusX: an open-source research platform for image-guided therapy. Int. J. Comput. Assist. Radiol. Surg. https://doi.org/10.1007/s11548-015-1292-0.

Avants, B.B., Epstein, C.L., Grossman, M., Gee, J.C., 2008. Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. Med. Image Anal. https://doi.org/10.1016/j.media.2007.06.004.

Bookstein, F.L., 1989. Principal warps: thin-plate splines and the decomposition of deformations. In: IEEE Transactions on Pattern Analysis and Machine Intelligence. https://doi.org/10.1109/34.24792.

Coupé, P., Hellier, P., Morandi, X., Barillot, C., 2007. A probabilistic objective function for 3D rigid registration of intraoperative us and preoperative MR brain images. In: 2007 4th IEEE International Symposium on Biomedical Imaging: from Nano to Macro - Proceedings. https://doi.org/10.1109/ISBI.2007.357103.

Coupé, P., Hellier, P., Morandi, X., Barillot, C., 2012. 3D rigid registration of intraoperative ultrasound and preoperative MR brain images based on hyperechogenic structures. Int. J. Biomed. Imaging. https://doi.org/10.1155/2012/531319.

de Vos, B.D., Berendsen, F.F., Viergever, M.A., Staring, M., Išgum, I., 2017. End-to-end unsupervised deformable image registration with a convolutional neural network. In: Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). https://doi.org/10.1007/978-3-319-67558-9_24.

Diez, Y., Oliver, A., Cabezas, M., Valverde, S., Martí, R., Vilanova, J.C., et al., 2014. Intensity based methods for brain MRI longitudinal registration. A study on multiple sclerosis patients. Neuroinformatics. https://doi.org/10.1007/s12021-013-9216-z.

Dimaio, S.P., Archip, N., Hata, N., Talos, I., Warfield, S.K., Majumdar, A., et al., 2006. Image-guided neurosurgery at Brigham and Women's hospital. In: IEEE Engineering in Medicine and Biology Magazine. https://doi.org/10.1109/MEMB.2006.1705749.

Drobny, D., Vercauteren, T., Ourselin, S., Modat, M., 2018. Registration of MRI and iUS data to compensate brain shift using a symmetric block-matching based approach. In: Simulation, Image Processing, and Ultrasound Systems for Assisted Diagnosis and Navigation, pp. 172–178. https://doi.org/10.1007/978-3-030-01045-4_21.

Fan, Y., Shen, D., Gur, R.C., Gur, R.E., Davatzikos, C., 2007. COMPARE: classification of morphological patterns using adaptive regional elements. IEEE Trans. Med. Imaging. https://doi.org/10.1109/TMI.2006.886812.

Farnia, P., Ahmadian, A., Shabanian, T., Serej, N.D., Alirezaie, J., 2014. A hybrid method for non-rigid registration of intra-operative ultrasound images with pre-operative MR images. Eng. Med. Biol. Soc. (EMBC) 5562–5565. https://doi.org/10.1109/EMBC.2014.6944887.

Farnia, P., Ahmadian, A., Shabanian, T., Serej, N.D., Alirezaie, J., 2015. Brain-shift compensation by non-rigid registration of intra-operative ultrasound images with preoperative MR images based on residual complexity. Int. J. Comput. Assist. Radiol. Surg. 10 (5). https://doi.org/10.1007/s11548-014-1098-5.

Farnia, P., Makkiabadi, B., Ahmadian, A., Alirezaie, J., 2016. Curvelet based residual complexity objective function for non-rigid registration of pre-operative MRI with intraoperative ultrasound images. In: Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS), 2016-Octob, pp. 1167–1170. https://doi.org/10.1109/EMBC.2016.7590912.

Farnia, P., Najafzadeh, E., Ahmadian, A., Makkiabadi, B., Alimohamadi, M., Alirezaie, J., 2018. Co-sparse analysis model based image registration to compensate brain shift by using intra-operative ultrasound imaging. In: 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 1–4.

Fitzpatrick, J.M., 2009. Fiducial registration error and target registration error are uncorrelated. In: Medical Imaging 2009: Visualization, Image-Guided Procedures, and Modeling. https://doi.org/10.1117/12.813601.

Fuerst, B., Wein, W., Müller, M., Navab, N., 2014. Automatic ultrasound-MRI registration for neurosurgery using the 2D and 3D LC2 Metric. Med. Image Anal. https://doi.org/10.1016/j.media.2014.04.008.

Gerard, I.J., Kersten-Oertel, M., Petrecca, K., Sirhan, D., Hall, J.A., Collins, D.L., 2017. Brain shift in neuronavigation of brain tumors: a review. Med. Image Anal. https://doi.org/10.1016/j.media.2016.08.007.

Glocker, B., Komodakis, N., Tziritas, G., Navab, N., Paragios, N., 2008. Dense image registration through MRFs and efficient linear programming. Med. Image Anal. https://doi.org/10.1016/j.media.2008.03.006.

Golby, A., 2015. Image-guided Neurosurgery. Academic Press.

Hartov, A., Roberts, D.W., Paulsen, K.D., 2008. A comparative analysis of coregistered ultrasound and magnetic resonance imaging in neurosurgery. Neurosurgery 62 (3 Suppl. 1). https://doi.org/10.1227/01.neu.0000317377.15196.45.

Heckemann, R.A., Rueckert, D., Aljabar, P., Hammers, A., Hajnal, J.V., 2006. Diffeomorphic Registration Using B-Splines. https://doi.org/10.1007/11866763_86.

Heinrich, M.P., 2018. Intra-operative ultrasound to MRI fusion with a public multimodal discrete registration tool. In: Simulation, Image Processing, and Ultrasound Systems for Assisted Diagnosis and Navigation, pp. 159–164.

Heinrich, Mattias P., Jenkinson, M., Bhushan, M., Matin, T., Gleeson, F.V., Brady, S.M., Schnabel, J.A., 2012. MIND: modality independent neighbourhood descriptor for multi-modal deformable registration. Med. Image Anal. https://doi.org/10.1016/j.media.2012.05.008.

Heinrich, Mattias Paul, Jenkinson, M., Papiez, B.W., Brady, S.M., Schnabel, J.A., 2013. Towards realtime multimodal fusion for image-guided interventions using self-similarities. In: Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). https://doi.org/10.1007/978-3-642-40811-3_24.

Hladuvka, J., 2003. Establishing point correspondence on training set boundaries. Tech. Rep. 40 (2003), 1–15. Retrieved from. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.6, 3459&rep=rep1&type=pdf.

Hong, J., Park, H., Hong, J., Park, H., 2018. Non-linear approach for MRI to intra-operative us registration using structural skeleton. In: Simulation, Image Processing, and Ultrasound Systems for Assisted Diagnosis and Navigation, pp. 138–145. https://doi.org/10.1007/978-3-030-01045-4_16.

Iversen, D.H., Wein, W., Lindseth, F., Unsgård, G., Reinertsen, I., 2018. Automatic intraoperative correction of brain shift for accurate neuronavigation. World Neurosurg. https://doi.org/10.1016/j.wneu.2018.09.012.

Ji, S., Wu, Z., Hartov, A., Roberts, D.W., Paulsen, K.D., 2008. Mutual-information-based image to patient re-registration using intraoperative ultrasound in image-guided neurosurgery. Med. Phys. 35 (10), 4612–4624. https://doi.org/10.1118/1.2977728.

Jiang, D., Shi, Y., Yao, D., Wang, M., Song, Z., 2016. miLBP: a robust and fast modality-independent 3D LBP for multimodal deformable registration. In: International Journal of Computer Assisted Radiology and Surgery. https://doi.org/10.1007/s11548-016-1407-2.

Klein, A., Andersson, J., Ardekani, B.A., Ashburner, J., Avants, B., Chiang, M.C., et al., 2009. Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. Neuroimage. https://doi.org/10.1016/j.neuroimage.2008.12.037.

Komodakis, N., Tziritas, G., Paragios, N., 2008. Performance vs computational efficiency for optimizing single and dynamic MRFs: setting the state of the art with primal-dual strategies. Comput. Vis. Image Understand. https://doi.org/10.1016/j.cviu.2008.06.007.

Letteboer, M.M.J., Viergever, M.A., Niessen, W.J., 2003. Rigid registration of 3D ultrasound data of brain tumours. Int. Congr. Ser. https://doi.org/10.1016/S0531-5131(03)00305-4.

Liao, S., Chung, A.C.S., 2010. Feature based nonrigid brain MR image registration with symmetric alpha stable filters. IEEE Trans. Med. Imaging. https://doi.org/10.1109/TMI.2009.2028078.

Liao, S., Chung, A.C.S., 2012. Nonrigid brain MR image registration using uniform spherical region descriptor. IEEE Trans. Image Process. https://doi.org/10.1109/TIP.2011.2159615.

Lindseth, F., Kaspersen, J.H., Ommedal, S., Langø, T., Bang, J., Hokland, J., et al., 2003. Multimodal image fusion in ultrasound-based neuronavigation: improving overview and interpretation by integrating preoperative MRI with intraoperative 3D ultrasound. Comput. Aided Surg. 8 (2), 49–69. https://doi.org/10.3109/10929080309146040.

Luo, J., Frisken, S., Machado, I., Zhang, M., Pieper, S., Golland, P., et al., 2018. Using the variogram for vector outlier screening: application to feature-based image registration. Int. J. Comput. Assist. Radiol. Surg. https://doi.org/10.1007/s11548-018-1840-5.

Machado, I., Toews, M., Luo, J., Unadkat, P., Essayed, W., George, E., et al., 2018a. Deformable MRI-ultrasound registration via attribute matching and mutual-saliency weighting for image-guided neurosurgery. In: Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). https://doi.org/10.1007/978-3-030-01045-4_20.

Machado, I., Toews, M., Luo, J., Unadkat, P., Essayed, W., George, E., et al., 2018b. Non-rigid registration of 3D ultrasound for neurosurgery using automatic feature detection and matching. Int. J. Comput. Assist. Radiol. Surg. https://doi.org/10.1007/s11548-018-1786-7.

Masoumi, N., Xiao, Y., Rivaz, H., 2018a. ARENA: inter-modality affine registration using evolutionary strategy. Int. J. Comput. Assist. Radiol. Surg. https://doi.org/10.1007/s11548-018-1897-1.

Masoumi, N., Xiao, Y., Rivaz, H., 2018b. MARCEL (Inter-Modality Affine Registration with CorrELation Ratio): an application for brain shift correction in ultrasound-guided brain tumor resection. In: Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). https://doi.org/10.1007/978-3-319-75238-9_5.

Mercier, L., Del Maestro, R.F., Petrecca, K., Araujo, D., Haegelen, C., Collins, D.L., 2012. Online database of clinical MR and ultrasound images of brain tumors. Med. Phys. https://doi.org/10.1118/1.4709600.

Mercier, L., Del Maestro, R.F., Petrecca, K., Kochanowska, A., Drouin, S., Yan, C.X.B., et al., 2011. New prototype neuronavigation system based on preoperative imaging and intraoperative freehand ultrasound: system description and validation. Int. J. Comput. Assist. Radiol. Surg. https://doi.org/10.1007/s11548-010-0535-3.

Miller, D., Benes, L., Sure, U., 2011. Stand-alone 3D-ultrasound navigation after failure of conventional image guidance for deep-seated lesions. Neurosurg. Rev. https://doi.org/10.1007/s10143-011-0314-9.

Modat, M., Cash, D.M., Daga, P., Winston, G.P., Duncan, J.S., Ourselin, S., 2014. Global image registration using a symmetric block-matching approach. J. Med. Imaging. https://doi.org/10.1117/1.jmi.1.2.024003.

Morin, F., Courtecuisse, H., Reinertsen, I., Le Lann, F., Palombi, O., Payan, Y., Chabanas, M., 2017. Brain-shift compensation using intraoperative ultrasound and constraint-based biomechanical simulation. Med. Image Anal. 40, 133–153. https://doi.org/10.1016/j.media.2017.06.003.

Munkvold, B.K.R., Bo, H.K., Jakola, A.S., Reinertsen, I., Berntsen, E.M., Unsgard, G., et al., 2017. Tumor volume Assessment in low-grade gliomas: a comparison of preoperative magnetic resonance imaging to coregistered intraoperative 3-dimensional ultrasound recordings. Neurosurgery. https://doi.org/10.1093/neuros/nyx392.

Myronenko, A., Song, X., 2010. Intensity-based image registration by minimizing residual complexity. IEEE Trans. Med. Imaging. https://doi.org/10.1109/TMI.2010.2053043.

Nigris, D. De, Collins, D.L., Arbel, T., 2013. Fast rigid registration of pre-operative magnetic resonance images to intra-operative ultrasound for neurosurgery based on high confidence gradient orientations. Int. J. Comput. Assist. Radiol. Surg. https://doi.org/10.1007/s11548-013-0826-6.

Ou, Y., Akbari, H., Bilello, M., Da, X., Davatzikos, C., 2014. Comparative evaluation of registration algorithms in different brain databases with varying difficulty: results and insights. IEEE Trans. Med. Imaging. https://doi.org/10.1109/TMI.2014.2330355.

Ou, Y., Shen, D., Zeng, J., Sun, L., Moul, J., Davatzikos, C., 2009. Sampling the spatial patterns of cancer: optimized biopsy procedures for estimating prostate cancer volume and Gleason Score. Med. Image Anal. https://doi.org/10.1016/j.media.2009.05.002.

Ou, Y., Sotiras, A., Paragios, N., Davatzikos, C., 2011. DRAMMS: deformable registration via attribute matching and mutual-saliency weighting. Med. Image Anal. https://doi.org/10.1016/j.media.2010.07.002.

Ou, Y., Zöllei, L., Da, X., Retzepi, K., Murphy, S.N., Gerstner, E.R., et al., 2018. Field of view normalization in multi-site brain MRI. Neuroinformatics. https://doi.org/10.1007/s12021-018-9359-z.

Palombi, O., Morin, F., Payan, Y., Chabanas, M., Le Lann, F., Courtecuisse, H., Reinertsen, I., 2018. Resection-induced brain-shift compensation using vessel-based methods. https://doi.org/10.1117/12.2293640.

Porter, B.C., Rubens, D.J., Strang, J.G., Smith, J., Totterman, S., Parker, K.J., 2001. Three-dimensional registration and fusion of ultrasound and MRI using major vessels as fiducial markers. IEEE Trans. Med. Imaging. https://doi.org/10.1109/42.921484.

Reinertsen, I., Lindseth, F., Unsgaard, G., Collins, D.L., 2007. Clinical validation of vessel-based registration for correction of brain-shift. Med. Image Anal. 11 (6), 673–684. https://doi.org/10.1016/j.media.2007.06.008.

Reinertsen, Ingerid, Lindseth, F., Askeland, C., Iversen, D.H., Unsgård, G., 2014. Intra-operative correction of brain-shift. Acta Neurochir. https://doi.org/10.1007/s00701-014-2052-6.

Rivaz, H., Chen, S.J.S., Collins, D.L., 2015. Automatic deformable MR-ultrasound registration for image-guided neurosurgery. IEEE Trans. Med. Imaging 34 (2). https://doi.org/10.1109/TMI.2014.2354352.

Rivaz, H., Collins, D.L., 2015. Deformable registration of preoperative MR, pre-resection ultrasound, and post-resection ultrasound images of neurosurgery. Int. J. Comput. Assist. Radiol. Surg. 10 (7), 1017–1028. https://doi.org/10.1007/s11548-014-1099-4.

Rivaz, H., Karimaghaloo, Z., Collins, D.L., 2014a. Self-similarity weighted mutual information: a new nonrigid image registration metric. Med. Image Anal. https://doi.org/10.1016/j.media.2013.12.003.

Rivaz, H., Karimaghaloo, Z., Fonov, V.S., Collins, D.L., 2014b. Nonrigid registration of ultrasound and MRI using contextual conditioned mutual information. IEEE Trans. Med. Imaging. https://doi.org/10.1109/TMI.2013.2294630.

Rizzini, D.L., Caselli, S., 2007. Improved mixture representation in real-time particle filters for robot localization. In: EMCR.

Roche, A., Pennec, X., Rudolph, M., Auer, D.P., Malandain, G., Ourselin, S., et al., 2000. Generalized correlation ratio for rigid registration of 3D ultrasound with MR images. Miccai. https://doi.org/10.1007/978-3-540-40899-4_58.

Roche, Alexis, Malandain, G., Pennec, X., Ayache, N., 2006. The correlation ratio as a new similarity measure for multimodal image registration. https://doi.org/10.1007/bfb0056301.

Rohlfing, T., 2012. Image similarity and tissue overlaps as surrogates for image registration accuracy: widely used but unreliable. IEEE Trans. Med. Imaging. https://doi.org/10.1109/TMI.2011.2163944.

Rueckert, D., 1999. Nonrigid registration using free-form deformations: application to breast mr images. IEEE Trans. Med. Imaging. https://doi.org/10.1109/42.796284.

Rygh, O.M., Selbekk, T., Torp, S.H., Lydersen, S., Hernes, T.A.N., Unsgaard, G., 2008. Comparison of navigated 3D ultrasound findings with histopathology in subsequent phases of glioblastoma resection. Acta Neurochir. https://doi.org/10.1007/s00701-008-0017-3.

Shams, R., Boucher, M.A., Kadoury, S., 2018. Intra-operative brain shift correction with weighted locally linear correlations of 3D US and MRI. In: Simulation, Image Processing, and Ultrasound Systems for Assisted Diagnosis and Navigation, pp. 179–184.

Shen, D., 2007. Image registration by local histogram matching. Pattern Recognit. https://doi.org/10.1016/j.patcog.2006.08.012.

Shen, D., Davatzikos, C., 2002. HAMMER: hierarchical attribute matching mechanism for elastic registration. IEEE Trans. Med. Imaging. https://doi.org/10.1109/TMI.2002.803111.

Sotiras, A., Davatzikos, C., Paragios, N., 2013. Deformable medical image registration: a survey. IEEE Transactions on Medical Imaging. https://doi.org/10.1109/TMI.2013.2265603.

Sun, L., Zhang, S., 2018. Deformable MRI-ultrasound registration using 3D convolutional neural network. In: Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). Springer, pp. 129–137.

Tempany, C.M.C.C., Jayender, J., Kapur, T., Bueno, R., Golby, A., Agar, N., Jolesz, F.A., 2015. Multimodal imaging for improved diagnosis and treatment of cancers. Cancer 121 (6), 817–827. https://doi.org/10.1002/cncr.29012.

Toews, Arbel, 2003. Entropy-of-likelihood feature selection for image correspondence. In: Proceedings of the 9th IEEE International Conference on Computer Vision (ICCV), pp. 1041–1047. https://doi.org/10.1109/iccv.2003.1238464.

Toews, M., Arbel, T., 2009. Detection, localization, and sex classification of faces from arbitrary viewpoints and under occlusion. IEEE Trans. Pattern Anal. Mach. Intell. https://doi.org/10.1109/TPAMI.2008.233.

Toews, M., Wells, W.M., 2013. Efficient and robust model-to-image alignment using 3D scale-invariant features. Med. Image Anal. https://doi.org/10.1016/j.media.2012.11.002.

Unsgaard, G., Selbekk, T., Brostrup Müller, T., Ommedal, S., Torp, S.H., Myhr, G., et al., 2005. Ability of navigated 3D ultrasound to delineate gliomas and metastases - comparison of image interpretations with histopathology. Acta Neurochir. https://doi.org/10.1007/s00701-005-0624-1.

Vercauteren, T., Pennec, X., Perchant, A., Ayache, N., 2007. Non-parametric diffeomorphic image registration with the demons algorithm. In: Medical Image Computing and Computer-Assisted Intervention – MICCAI 2007. https://doi.org/10.1007/978-3-540-75759-7_39.

Vercauteren, T., Pennec, X., Perchant, A., Ayache, N., 2009. Diffeomorphic demons: efficient non-parametric image registration. Neuroimage. https://doi.org/10.1016/j.neuroimage.2008.10.040.

Wachinger, C., Navab, N., 2012. Entropy and Laplacian images: structural representations for multi-modal registration. Med. Image Anal. https://doi.org/10.1016/j.media.2011.03.001.

Wachinger, C., Navab, N., 2013. Simultaneous registration of multiple images: similarity metrics and efficient optimization. IEEE Trans. Pattern Anal. Mach. Intell. https://doi.org/10.1109/TPAMI.2012.196.

Wang, S., Kim, M., Wu, G., Shen, D., 2017. Scalable high performance image registration framework by unsupervised deep feature representations learning. In: Deep Learning for Medical Image Analysis. https://doi.org/10.1016/B978-0-12-810408-8.00015-8.

Wein, W., 2018. Brain-shift correction with image-based registration and landmark accuracy evaluation. In: Simulation, Image Processing, and Ultrasound Systems for Assisted Diagnosis and Navigation. Springer, pp. 146–151.

Wein, W., Ladikos, A., Fuerst, B., Shah, A., Sharma, K., Navab, N., 2013. Global registration of ultrasound to MRI using the LC2metric for enabling neurosurgical guidance. In: Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). https://doi.org/10.1007/978-3-642-40811-3_5.

Wu, G., Yap, P.T., Kim, M., Shen, D., 2010. TPS-HAMMER: improving HAMMER registration algorithm by soft correspondence matching and thin-plate splines based deformation interpolation. Neuroimage. https://doi.org/10.1016/j.neuroimage.2009.10.065.

Xiao, Y., Fortin, M., Unsgärd, G., Rivaz, H., Reinertsen, I., 2017. REtroSpective Evaluation of Cerebral Tumors (RESECT): a clinical database of pre-operative MRI and intra-operative ultrasound in low-grade glioma surgeries: a. Med. Phys. https://doi.org/10.1002/mp.12268.

Xiao, Y., Rivaz, H., Chabanas, M., Fortin, M., Machado, I., Ou, Y., et al., 2019. Evaluation of MRI to Ultrasound Registration Methods for Brain Shift Correction: the CuRIOUS2018 Challenge. ArXiv E-Prints arXiv:1904.10535.

Yang, X., Kwitt, R., Styner, M., Niethammer, M., 2017. Quicksilver: fast predictive image registration – a deep learning approach. Neuroimage. https://doi.org/10.1016/j.neuroimage.2017.07.008.

Zhong, X., Bayer, S., Ravikumar, N., Strobel, N., Birkhold, A., Kowarschik, M., et al., 2018. Resolve intraoperative brain shift as imitation game. In: Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 11042 LNCS. Springer, pp. 129–137. https://doi.org/10.1007/978-3-030-01045-4_15.

Zimmer, V.A., González Ballester, M.Á., Piella, G., 2019. Multimodal image registration using Laplacian commutators. Inf. Fusion. https://doi.org/10.1016/j.inffus.2018.09.009.