IJCARS manuscript No.
(will be inserted by the editor)

ARENA: Inter-modality affine registration using evolutionary strategy

Nima Masoumi · Yiming Xiao · Hassan Rivaz

Received: date / Accepted: date

Abstract
Purpose: Image fusion of different imaging modalities renders valuable information to clinicians. In this paper, we proposed an automatic multimodal registration method to register intra-operative ultrasound images (US) to pre-operative magnetic resonance images (MRI) in the context of image-guided neurosurgery (IGNS).
Methods: We employed refined correlation ratio (CR) as a similarity metric for our intensity based image registration method. We deem MRI as the fix image (I_f) and US as the moving image (I_m) and then transform I_m to align with I_f. We utilized the covariance matrix adaptation evolutionary strategy (CMA-ES) to find the optimal affine transformation in registration of I_m to I_f.
Results: We applied our method on the publicly available retrospective evaluation of cerebral tumors (RESECT) database and Montreal Neurological Institute's brain images of tumors for evaluation (BITE) database. We validated the results qualitatively and quantitatively. Qualitative validation is conducted (by the three authors) through overlaying pre- and post-registration US and MRI to allow visual assessment of the alignment. Quantitative validation is performed by utilizing the corresponding landmarks in the databases for the pre-operative MRI and the intra-operative US. Average mean target registration error (mTRE) has been reduced from 5.40±4.27 to 2.77±1.13 in 22 patients in the RESECT database and from 4.12±2.03 to 2.82±0.72 in the BITE database. A nonparametric statistical analysis performed using the Wilcoxon rank sum test shows that there is a significant difference between pre- and post-registration mTREs with a p-value of 0.0058 ($p < 0.05$) for the RESECT database and 0.0483 ($p < 0.05$) for the BITE database.
Conclusions: The proposed fully automatic registration method significantly improved the alignment of MRI and US images and can therefore be used to reduce the misalignment of US and MRI caused by brain shift, calibration errors, and patient to MRI transformation matrix.

Keywords Image registration · Correlation Ratio · Affine Transformation · CMA-ES · RESECT Database · mTRE

1 Introduction

In medical imaging, we often have chronological images of tissues (which are usually collected with different imaging modalities) that need to be aligned [1,2]. Fusion of the information of those corresponding images is proven to provide useful information to clinicians [3–5]. Even though registration based on

N. Masoumi
PERFORM Centre, Concordia University, Montreal, Canada
Department of Electrical and Computer Engineering, Concordia University, Montreal, Canada
E-mail: n_masoum@encs.concordia.ca

Y. Xiao
Robarts Research Institute, Western University, London, Canada
E-mail: yxiao286@uwo.ca

H. Rivaz
PERFORM Centre, Concordia University, Montreal, Canada
Department of Electrical and Computer Engineering, Concordia University, Montreal, Canada
E-mail: hrivaz@ece.concordia.ca

manually selected homologous landmarks can be performed on images, the corresponding images are often misaligned due to reasons such as tissue deformation and errors in landmark selection. For example, in image-guided surgery, deformation of the organs, such as the brain can invalidate surgical plannings [6–9].

Image registration is the method, which aligns corresponding misaligned images acquired in different times and/or with different sensors [10]. One can categorize image registration methods in various classes such as automatic or interactive [11]. Automatic image registration is generally faster and avoids erroneous actions of the user [12]. Another classification can be made based on the method used: intensity-based or feature-based. Intensity-based image registration methods generally work better for smaller deformations, whereas feature-based methods generally work better if the initial misalignment is large [13, 14].

An automatic intensity-based image registration method can consist of different components. One image would be chosen as the template or fixed image (I_f). The other image is called the moving image (I_m). During the registration process, I_m should move to be registered to I_f. The movement of I_m can be restricted and modeled by a spatial transformation. A transformation type is selected based upon the application [15–17]. When there is no deformation of the object scene, we can simply use a rigid transformation, which only has six degrees of freedom [18–20]. When one image has deformation with respect to the other one, we can use transformations with more parameters for instance, affine or free-form B-spline transformations [21–23]. The image registration method should have a similarity metric to evaluate the similarity of two images after the transformation. On one end of the spectrum, the similarity metric can assume a restrictive equality relationship between image intensities and easily subtract two images as in sum of square differences (SSD). On the other end of the spectrum, it can assume a general information-based similarity between images as in mutual information (MI) [24]. Correlation ratio (CR) assumes a functional relationship between intensities of the two images and provides a compromise between these two extremes. The original CR proposed by Roche et al [25, 26], was calculated globally for the entire volumes, whereas CR used in [27] is calculated in small local patches. More importantly, unlike the original CR, we perform a binning of intensities of the reference image and calculate histograms using Parzen windows. This allows us to reliably calculate CR from small patches. The third component of registration methods, maximizes the similarity of the images by varying the parameters of the chosen transformation [25–27].

We proposed an automatic intensity-based image registration method using the refined version of CR. The proposed method is an extended version of the method (MARCEL) proposed in [28] which was itself based on RaPTOR (Robust PaTch-based cOrrelation Ratio) [29]. Our similarity metric measures similarity of the images based on corresponding patches locally. We modeled movement of I_m with affine transformation and utilized the covariance matrix adaptation evolutionary strategy (CMA-ES) [30] as the optimization approach. We applied our method on RESECT (REtroSpective Evaluation of Cerebral Tumors) database [31] to validate the results. Recent work has successfully performed US-US registration of the RESECT database [32]. In order to show the performance of our method on different MRI and ultrasound data, we also applied our method to the BITE database [33], which was collected using different ultrasound and MRI machines.

The contributions of this paper are threefold. First, ARENA uses CMA-ES, which does not need gradient of the cost function, which becomes noisy when the patch size (i.e. the number of samples) is small. Therefore, the optimization step in ARENA is less susceptible to patch-size and noise. Second, we show for the first time that US and MRI images of the RESECT database can be automatically registered. And third, although CMA-ES has been successfully used in registration of other imaging modalities [34–37], it is used for registration of MRI and US for the first time in this paper, where we show that it works even for patients where a very large initial misalignment exists between US and MRI.

The main difference between MARCEL [28] and ARENA is that MARCEL uses gradient descent optimization, while ARENA uses CMA-ES. Gradient descent optimization needs the derivative of the cost function and also requires tuning the step-size, whereas CMA-ES does not need the gradient and further does not need tuning of the optimization parameters. In addition, MARCEL was only tested on 5 subjects in the RESECT database, whereas ARENA is tested on all 22 subjects. This further justifies the use of CMA-ES and its ability to converge to the correct solution in all tested cases. Consequently, ARENA has the following improvements comparing to MARCEL. First, ARENA is less computationally expensive and is also more straightforward to implement due to the simplicity of the CMA-ES optimization method. Second, ARENA is less sensitive to parameter tuning compared to MARCEL and RaPTOR.

This paper is organized as follows. In Section 2, we elaborate our method and derive the equations. In Section 3, qualitative and quantitative validations of the method are presented. In Section 4, we discuss

the advantages and disadvantages of our method and avenues for the future. And finally, we provide a brief conclusion in Section 5.

2 Methods

Let I_m and I_f be respectively the moving and fixed images. In our registration problem we fix I_f and move I_m so that it matches I_f. We transform I_m with $\mathbf{T}$. The optimal $\mathbf{T}$, when applied to I_m, for each point like $\mathbf{x}$ in the space of images, gives us the best alignment of I_f and I_m. Alignment of I_f and I_m is measured by a dissimilarity metric D. The best alignment of I_f and I_m with T corresponds to minimum achievable D. In other words, our goal is to minimize the following cost function:

$$C = D(I_f(\mathbf{x}), I_m(\mathbf{T}(\mathbf{x}))) + R(\mathbf{T}) \tag{1}$$

where $R(\mathbf{T})$ is a regularization term to enforce a smooth transformation and C is the cost function. Minimizing C by varying $\mathbf{T}$ provides the transformation that aligns the fixed and moving images.

2.1 Dissimilarity Metric

As explained in Eq. 1, D measures the alignment of input images i.e. the fixed and moving images. Since CR is an asymmetric similarity metric, the order of computing CR is important. To allow either I_m or I_f to be the first or second image in CR, we label our input images as X and Y. D in Eq. 1 and in Eq. 2 is the amended version of RaPTOR [29]. D can vary from zero to one. In case that X and Y are the same, $D = 0$. When X and Y do not have any similarity, $D = 1$. Therefore D is a dissimilarity metric. In Eq. 2, η is CR, the similarity metric proposed by Roche et al [25]. The similarity metric needs to identify corresponding features of X and Y locally. So we calculate CR in N_P corresponding pacthes of X and Y.

$$D(Y, X) = \frac{1}{N_p} \sum_{i=1}^{N_p} (1 - \eta(Y|X; \boldsymbol{\Omega_i})) \tag{2}$$

where $\boldsymbol{\Omega_i}$ reperesents the patch i space. The definition of CR in Eq. 2 is as following:

$$1 - \eta(Y|X) = \frac{1}{N\sigma^2} \left(\sum_{t=1}^{N} i_t^2 - \sum_{j=1}^{N_b} N_j \mu_j^2 \right) \tag{3}$$

$$\mu_j = \frac{\sum_{t=1}^{N} \lambda_{t,j} i_t}{N_j}, N_j = \sum_t \lambda_{t,j} \tag{4}$$

where N is the total number of voxels in Y, $\sigma^2 = Var[Y]$, i_t is the intensity of voxel number t in Y, N_b is the total number of bins, and $\lambda_{t,j}$ is the contribution of sample t in bin j as proposed in [29].

Obviously in calculation of D in Eq. 2, pacthes that have approximately the same voxel intensities or equally small variances, should be discarded because they do not include any image feature. Therefore, we apply a gamma correction on patches of X and Y as the one explained in [38] after selecting patches in X and Y to increase variance of patch intensities. The gamma correction applies an experimental transformation, $i_n = exp(\gamma i_0)$, on the voxels of patches where i_n is new intensity of the voxel. γ is the correction parameter which we set it to 50 heuristically, and i_0 is old intensity of the voxel. We normalize intensities of the patches right after the gamma correction. Then every pair of patches in which $\sigma^2 < T$ are discarded. Heuristically, we found that $T = 1$ is the best value.

2.2 Transformation

We used affine transformation to model the movement of moving image. In our formulation, no regularization is needed in Eq. 1 since the affine transformation has 12 parameters to be optimizaed, and the images provide many cues for reliably optimizing for those parameters. The affine transformation matrix

is defined as:

$$\boldsymbol{T} = \begin{bmatrix} a_1 & a_2 & a_3 & a_4 \\ a_5 & a_6 & a_7 & a_8 \\ a_9 & a_{10} & a_{11} & a_{12} \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{5}$$

As one can see in Eq. 5, the affine transformation has twelve parameters which are a_i, $1 \leq i \leq 12$. In general, these twelve parameters can be any real number.

2.3 Optimization

The explanation in the Section 2 defines the registration procedure as an optimization problem. Image registration, in general, is an ill-posed problem, and consequently entails optimizing a highly non-convex objective function [39]. In order to tackle this problem, we deployed CMA-ES as our optimizer. In Eq. 1, C is the cost of the objective function D. The affine transformation parameters a_i, $1 \leq i \leq 12$ in Eq. 5 are used by the optimization algorithm to minimize C in Eq. 1.

CMA-ES is similar to natural selection of the biological creatures [40]. In each iteration (generation) λ new candidate solutions (offsprings) $x_k^{(g+1)}$, $1 \leq k \leq \lambda$ are calculated from the best μ out of λ of the last generation (parents) $x_{i:\lambda}^{(g)}$, $1 \leq i \leq \mu$.

There are $N = 12$ degrees of freedom in the optimization established by affine transformation parameters. Hence, the parameter settings for λ and μ are $\lambda = 4 + \lfloor 3\ln(N) \rfloor$ and $\mu = \lfloor \lambda/2 \rfloor$. CMA-ES update equation for the generation g to $g+1$ is presented in Eq. 6.

$$x_k^{(g+1)} = \frac{1}{\sum_{i=1}^{\mu} w_i} \sum_{i=1}^{\mu} w_i x_{i:\lambda}^{(g)} + \sigma^{(g)} \boldsymbol{B}^{(g)} \boldsymbol{D}^{(g)} z_k^{(g+1)} \tag{6}$$

where w_i, $1 \leq i \leq \mu$ are summation weights of offsprings and they are calculated as Eq. 7.

$$w_i = \ln(\frac{\lambda + 1}{2}) - \ln(i) \tag{7}$$

In Eq. 6 $\sigma^{(g)} \in \mathbb{R}^+$ is the step size at the generation g. So called covariance matrix $\boldsymbol{C^{(g)}}$ in the generation g is a symmetric positice definite $N \times N$ and its relationship with defined parameters is presented in Eq. 8:

$$\boldsymbol{B}^{(g)} \boldsymbol{D}^{(g)} z_k^{(g+1)} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{C^{(g)}}) \tag{8}$$

For detailed explanations and equations of $\sigma^{(g)}$, $\boldsymbol{B}^{(g)}$, $\boldsymbol{D}^{(g)}$, $z_k^{(g+1)}$, and $\boldsymbol{C^{(g)}}$ one can refer to [40, 41].

2.4 Patient Data

We applied the proposed image registration method on the RESECT database [31] and the BITE database [33]. The RESECT database is an open source clinical database that contains 23 surgical cases of low-grade gliomas resection operated at St. Olavs University Hospital. With the primary goal to help develop image processing techniques for brain shift correction, for each patient, the dataset provides pre-operative T1w and T2-FLAIR MRI scans, intra-operative 3D ultrasound volumes obtained before, during, and after tumor resection, and corresponding anatomical landmarks between MRI-US pairs and US-US pairs. To demonstrate our proposed algorithm, we used the pre-operative T2-FLAIR MRI and US volume before tumor resection since often this stage sets the tone for the total brain shift after craniotomy. More specifically, 22 patients from the RESECT dataset were used, where 15-16 pairs of MRI-US homologous landmarks were manually tagged. The BITE database consists of 14 patients with pre-operative T1w MRI and pre-resection US. As one of the patients' landmarks were outside the image. we excluded that patient from our experiment (as is also done in other publications [29, 42, 43]).

2.5 Registration Procedure

For each patient, we first up-sampled the MRI image ($resolution = 1 \times 1 \times 1mm^3$) to the resolution of corresponding US image because of the US images considerable higher resolution ($resolution = 0.24 \times$

$0.24 \times 0.24mm^3$). We set the US as I_m and MRI as I_f and then we implemented the image registration algorithm on each patient. For better performance of our method, we used up to four levels of Gaussian pyramids to tackle the large misalignment present in some of the cases. We use $[q/p]$ patches with the size of $7 \times 7 \times 7$ in this work to calculate CR where q is the number of nonzero voxels in the US image, p is the number of pixels in each US slice, and [.] denotes rounding a number to the next smaller integer operator.

3 Results

3.1 Qualitative Validation

By comparing the images before and after the registration, with visual inspection, we evaluated the quality of the registration. We compared the alignment of corresponding brain anatomical features for instance sulci and tumor boundaries in the MRI and US images before and after registration. Each patient data includes the brain tumor in MRI and US images. We checked whether alignment of the boundary of the tumor has improved as well.

Figure 1 shows overlaid US and MRI slices of sagittal view for Patient 5, 12, 19, and 21 in the RESECT database [31]. Columns show before and after the registration respectively. Each row corresponds to an individual patient. The arrows guide the reader to locate the improvement after the registration. The first and second rows show significant improvements in tumor and sulci region. The second and the third rows show improvements around the tumor region.

3.2 Quantitative Validation

Corresponding homologous landmarks are selected manually in the US and MRI in the RESECT database by two experts. Consider N as the total number of corresponding landmarks in US and MRI volumetric images. We used the provided landmarks to calculate mean target registration error (mTRE) [44]. mTRE for each patient is defined as Eq. 9:

$$mTRE = \frac{1}{N}\sum_{i=1}^{N} \|\boldsymbol{T}(\boldsymbol{x_i}) - \boldsymbol{x_i'}\| \tag{9}$$

where $\boldsymbol{T}$ is the optimal affine transformation derived after implementing the image registration algorithm.

Initial mTRE of each patient before registration and the number of landmarks for each patient is reported in Table 1. Each patient has N landmarks and affine transformation has twelve parameters. In this table, minimum achievable mTRE is the minimum mTRE we can achieve using an affine transformation for the registration. We made system of linear equations to find the optimal achievable affine transformation. In this system, the provided landmarks are knowns and the optimal affine transformation parameters are unknowns. Therefore, the number of knowns is more than the number of unknowns $N > 12$. We solved this overdetermined problem with least squares (LS). We reported LS solution for each patient in Table 1. It is worth mentioning that while the minimum achievable mTREs are calculated the similar way as the fiducial registration error (FRE) [45], they are not equal to FRE. FRE is the root mean square error (RMSE) and we calculate mean root square error (MRSE) so that it can be compared to the initial and final mTRE values calculated before and after registration respectively.

Recently, the Correction of Brainshift with Intra-Operative Ultrasound (CuRIOUS) 2018 Challenge (curious2018.grand-challenge.org) was held in conjunction with Medical Image Computing and Computer Assisted Intervention (MICCAI) 2018 Conference and addressed the same problem of registration of pre-resection US to MRI in the RESECT database. These papers used methods such as Multilayer Perceptron (MLP) [46], Demons [47], Linear Correlation of Linear Combination (LC^2) [48], Spatial Transformer Network (STN) using 3D Convolutional Neural Network (CNN) [49], Self-Similarity Context descriptors (SSC) [50], Scale-Invariant Feature Transform (SIFT) [51], symmetric block matching using Normalized Cross-Correlation (NCC) [52], and LC^2 using Bound Optimization BY Quadratic Approximation (BOBYQA) [53]. ARENA is compared to the first [48] and second [50] place participatns in this challenge.

ARENA improved alignments for each patient. In Table 1, initial mTRE shows rather high value of standard deviation. As in Table 1, our method had a significant improvement for standard deviation.

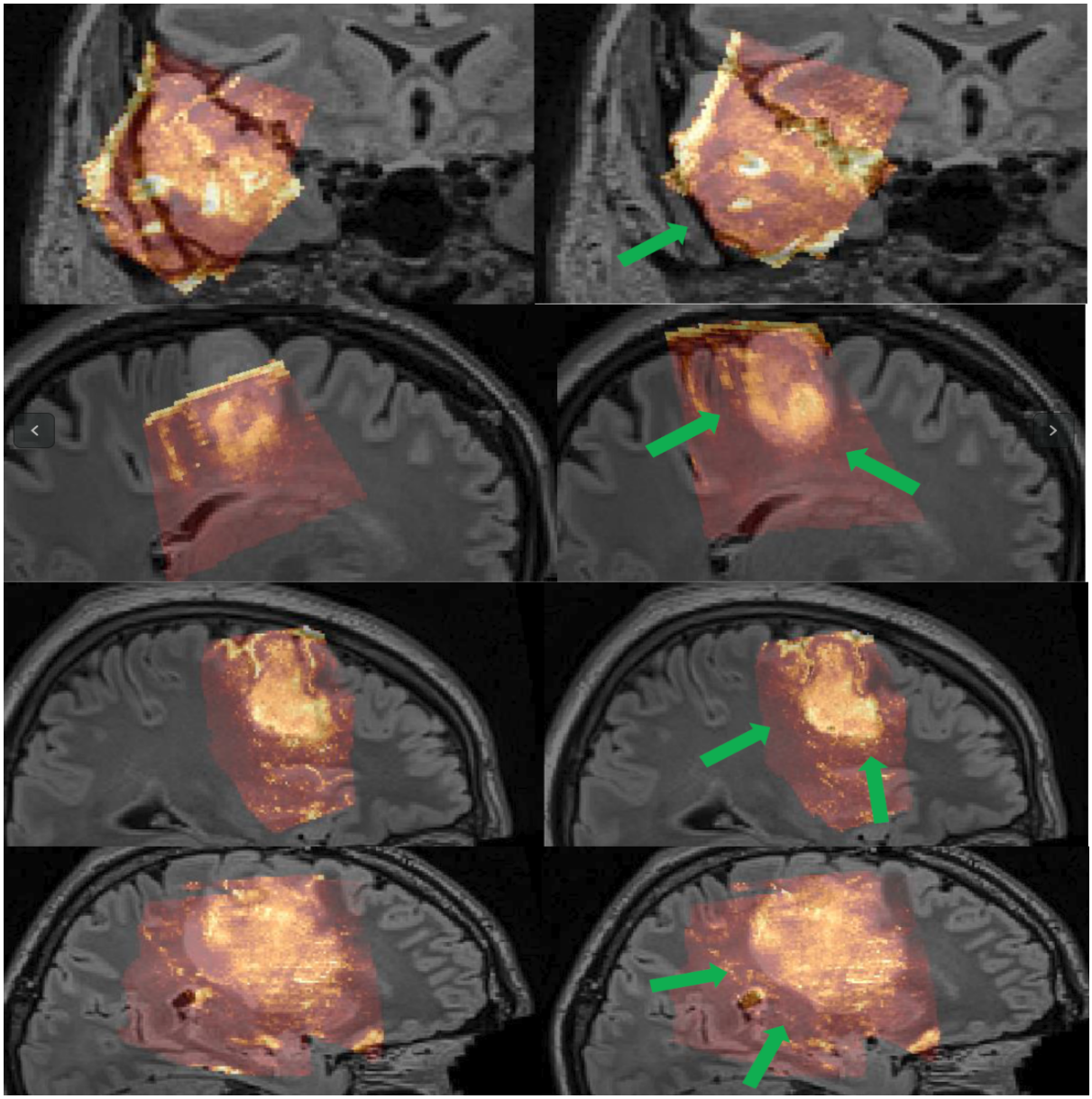

Fig. 1 From the top row, sagittal view of Patient 5, 12, 19, and 21 respectively. Columns show before and after the registration respectively. The arrows show where the images had improvements

One can interpret it as ability of the method to improve a wide range of misaligned images with high mTRE values. Figure 2 shows the data in Table 1 graphically.

Furthermore, the proposed method was applied on the BITE database as well (Table 2). The method is compared with SSC [43] and LC^2 [42]. Figure 3 shows the results of Table 2 graphically. Note that the SSC method utilized nonlinear deformable transformation with 10^7 DOFs, which allows more complex deformation than an affine transformation with 12 DOFs.

The SSc method applied on the BITE database [43] has different transformation and similarity metric than the one applied on the RESECT database [50]. In [43] authors utilized a complex transformation with 10^7 Degree of Freedoms (DOFs) whereas in [50], they applied a linear method and nonlinear methods to correct the brain shift. On the other hand the rigid registration performed with LC^2 in [42] and [48] are different from each other. In [42], the method registered 2D slices of US images to 3D MRI images using rigid registration as the initialization and the cubic spline as the deformable registration. Whereas in [48], the authors performed a 3D registration by initializing the registration with a translation, then they performed a rigid registration, and finally they applied the method with an affine transformation. In Table 1 and Table 2 we reported results of the rigid initialization before the principal registration in [42] and [48].

Table 1 Pre- and post-registration mTRE values corresponding to the RESECT database. SSC was used with linear deformation. LC^2 and ARENA were used with rigid and affine transformation respectively. The minimum achievable mTRE using an affine transformation is shown for comparison.

Patients ID No.	Initial mTRE (mm)	SSC (mm)	LC^2 (mm)	ARENA (mm)	Minimum Achievable mTRE (mm)	No. of Land-marks
1	1.82 (0.5-3.8)	1.88	1.72	1.49 (0.3-2.5)	1.10	15
2	5.68 (3.4-8.9)	2.38	2.53	3.23 (1.6-7.9)	1.11	15
3	9.58 (8.5-10.3)	1.29	1.33	3.65 (2.6-4.9)	0.80	15
4	2.99 (1.6-4.5)	1.31	1.65	2.04 (0.5-3.5)	0.95	15
5	12.02 (10.0-14.1)	1.87	1.50	4.08 (1.3-5.5)	0.93	15
6	3.27 (2.2-4.2)	1.86	1.67	1.55 (0.4-2.8)	0.75	15
7	1.82 (0.2-3.6)	1.58	1.57	1.70 (1.0-2.7)	1.22	15
8	2.63 (1.0-4.1)	2.66	1.94	1.76 (0.4-3.0)	1.08	15
12	19.68 (18.5-21.3)	1.43	1.06	5.52 (3.0-8.1)	0.91	16
13	4.57 (2.7-7.5)	3.47	3.74	2.55 (1.3-3.9)	0.90	15
14	3.03 (1.9-4.4)	1.33	1.20	2.61 (1.1-4.2)	0.89	15
15	3.21 (1.1-5.9)	2.32	1.91	2.46 (1.3-4.9)	1.33	15
16	3.39 (1.6-4.4)	1.41	1.24	1.63 (0.8-2.6)	0.91	15
17	6.39 (4.4-7.8)	1.78	1.71	3.69 (2.1-5.3)	1.03	16
18	3.56 (1.4-5.4)	1.23	1.24	1.96 (0.9-3.2)	0.78	16
19	3.28 (1.3-5.4)	2.12	2.12	2.66 (1.0-4.0)	0.86	16
21	4.55 (3.4-6.1)	1.90	1.87	3.08 (1.8-4.5)	0.68	16
23	7.01 (5.2-8.2)	1.59	1.89	3.15 (1.9-4.8)	0.69	15
24	1.10 (0.4-2.0)	1.57	1.12	1.04 (0.5-1.8)	0.70	16
25	10.06 (7.1-15.1)	3.21	2.78	4.56 (3.1-6.5)	0.87	15
26	2.83 (1.6-4.4)	1.60	1.36	2.50 (1.1-4.0)	0.98	16
27	5.76 (4.8-7.1)	1.58	1.44	3.99 (2.4-5.5)	1.03	16
Mean(μ)	5.37	1.88	1.75	2.77	0.93	-
Std(σ)	4.27	0.53	0.62	1.13	0.17	-

Table 2 Pre- and post-registration mTRE values corresponding to the BITE database. SSC was used with nonlinear deformation wheras LC^2 and ARENA were used with rigid and affine transformation respectively. The minimum achievable mTRE using an affine transformation is shown for comparison.

Patients ID No.	Initial mTRE (mm)	SSC (mm)	LC^2 (mm)	ARENA (mm)	Minimum Achievable mTRE (mm)	No. of Land-marks
2	6.50 (1.9-9.1)	1.81 (0.8-2.7)	1.73	3.65 (1.2-6.4)	1.39	35
3	9.38 (6.3-14.6)	2.60 (1.2-3.9)	2.76	2.95 (2.1-5.9)	2.17	40
4	3.93 (1.0-6.1)	1.78 (0.8-2.7)	1.96	2.61 (1.0-4.40)	1.55	32
5	2.62 (0.5-6.9)	2.30 (0.5-4.0)	2.14	2.40 (0.3-6.2)	1.82	31
6	2.30 (0.2-4.4)	1.83 (0.8-2.7)	1.94	1.81 (0.3-3.6)	1.45	37
7	3.04 (0.3-6.3)	2.41 (1.3-3.4)	2.33	2.74 (1.4-4.4)	1.98	19
8	3.65 (0.0-8.4)	2.40 (1.1-3.6)	2.87	2.95 (1.2-7.5)	2.20	23
9	5.09 (2.5-7.6)	2.41 (0.8-3.9)	2.81	3.63 (0.9-6.2)	2.23	21
10	2.99 (0.8-5.3)	1.68 (0.8-2.5)	2.06	2.25 (0.4-4.3)	1.40	25
11	1.52 (0.6-3.5)	2.17 (0.8-3.4)	2.18	1.46 (0.3-3.2)	1.27	25
12	3.70 (0.9-7.0)	3.31 (1.3-5.3)	2.67	3.05 (1.3-6.2	2.04	21
13	5.15 (1.5-10.4)	3.31 (1.7-4.8)	3.58	4.00 (0.8-8.3)	2.66	23
14	3.77 (1.2-5.7)	2.41 (0.9-3.8)	2.48	3.17 (0.6-6.7)	2.24	23
Mean(μ)	4.12	2.12	2.52	2.82	1.88	-
Std(σ)	2.03	1.29	0.87	0.72	0.43	-

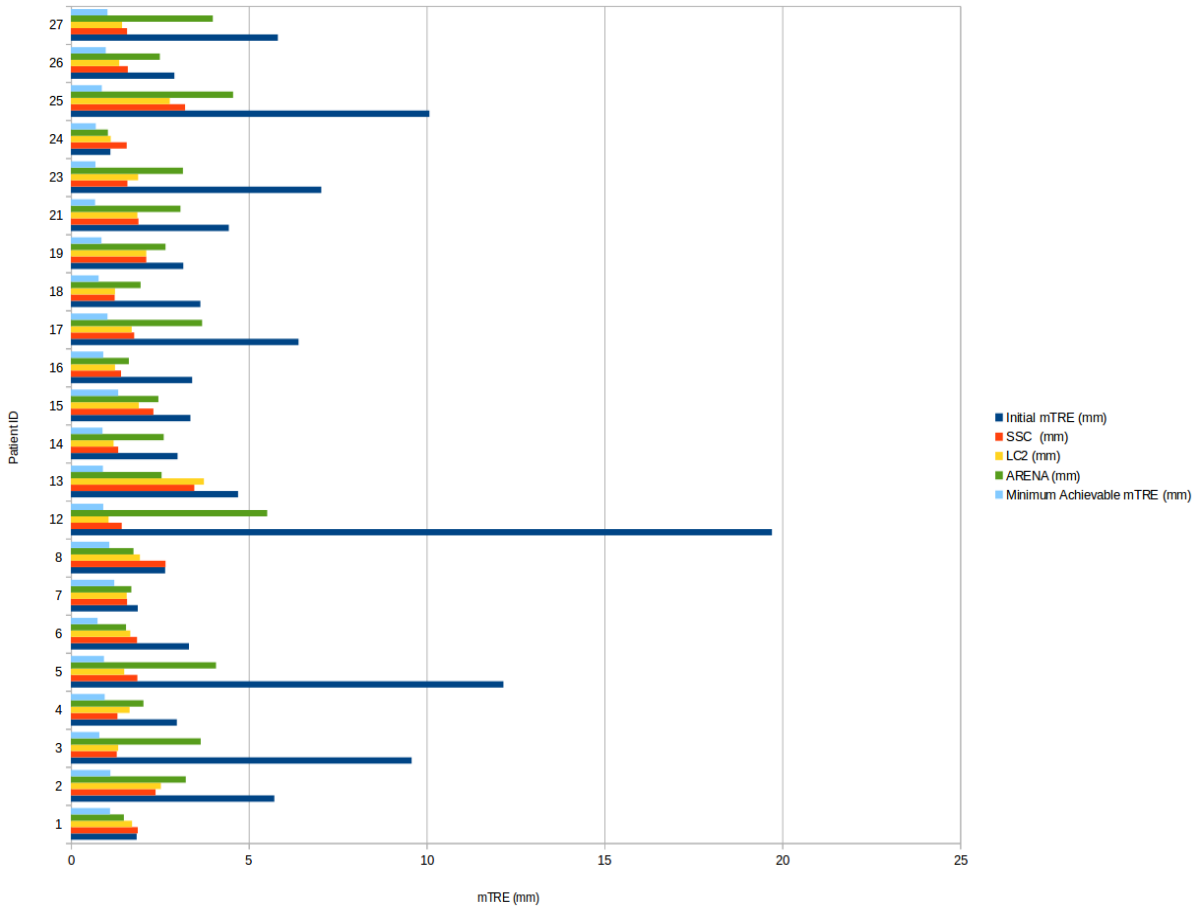

Fig. 2 Comparison of mTREs before and after the registration of patients in the RESECT database for SSC using linear transformation, LC^2 using rigid transformation, and ARENA. The minimum achievable mTRE by affine transformation is included as well.

In addition to the validation method, we did a statistical analysis of our results. We used the Wilcoxon rank sum test which is a nonparametric statistical analysis method [54]. In this test, the null hypothesis H_0 is: the method did not have improvement in mTRE. Using the data in Table 1, the null hypothesis is $\mu = 5.40$. The alternative hypothesis H_1 would be $\mu \geqslant 5.40$. Using the initial mTRE before the image registration and the results of ARENA in Table 1, we achieved the p-value of 0.0058 by applying Wilcoxon rank sum test. Considering the conventional significance level of $\alpha = 0.05$, $p = 0.0058$ shows that not only we reject H_0 and H_1, but also with %99.42 confidence we improved the result. With the same setting using the data in Table 2 and the mTREs of pre- and post-registration for ARENA, the p-value of 0.0483 ($p < 0.05$) has been achieved.

4 Discussion and Future Work

We showed the minimum achievable mTRE values with an affine transformation to provide a lower bound for mTRE values. We have not used these values to optimize and improve ARENA. We achieved mTRE values that are very close to this minimum value in some patients (e.g., in Patient 24). However, the average minimum achievable mTRE is 0.93mm, which is in the order of inter-observer variability (below $0.5mm$ [31]) in landmark selection. Therefore, it is expected that our final mTRE values be larger than the minimum achievable error.

In this work, we proposed to use a simple affine transformation to correct for brain shift. Nevertheless, non-linear transformations offer more flexibility and allow us to recover the deformation more accurately. Before employing affine transformation we used simple translation, rigid transformation, and rigid transformation with scaling parameters. We notice that none of them is able to improve mTRE for all patients. Affine transformation was the least general transformation model that could give us signifi-

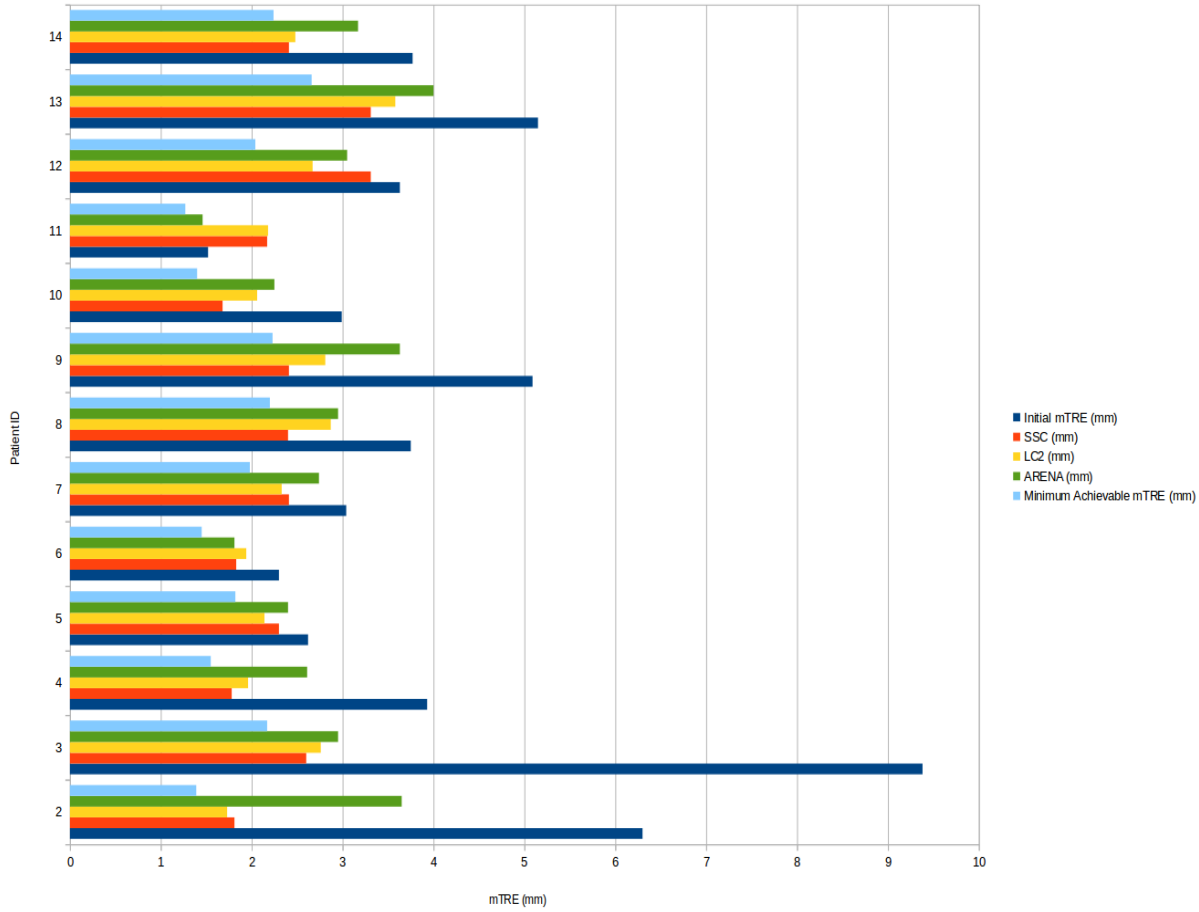

Fig. 3 Comparison of mTREs before and after the registration of patients in the BITE database for SSC using nonlinear transformation, LC^2 using rigid transformation, and ARENA. The minimum achievable mTRE by affine transformation is included as well.

cant improvement in mTRE. Affine transformation is simpler and faster than non-linear transformations, and practical in a wide range of applications.

CR and its derivatives RaPTOR and ARENA are asymmetric similarity metrics, meaning that reversing the order of images changes the similarity value and likely the results. We set the US and MRI as moving and fixed images respectively since this provided better results for ARENA. Since ARENA uses affine transformation, it can be simply inverted if clinicians prefer to deform the MRI to align with US.

Symmetric image registration methods are generally considered superior to asymmetric techniques [55, 56]. However, we found that asymmetric method used in this paper is superior to a symmetric version of ARENA. In addition, changing the role of the fixed and moving image substantially degrades the results. One reason is that CR is an asymmetric similarity measure and order of the images substantially changes its value. Because affine transformation is an invertible transformation, we can move the MRI image by inverting the transformation.

Image registration with affine transformation has a good performance for structural images. But for functional data, such as tractography, nonlinear deformation is necessary to preserve the continuity of the tracts [57]. Except for the patients with large initial mTREs in the RESECT and BITE database, ARENA has a close performance to the SSC and LC^2. ARENA was utilized with exactly the same settings for both databases, which was not the case for either SSC or LC^2 [42, 43, 48, 50]. In addition, LC^2 method is computationally expensive and was implemented on GPU. The population size λ in CMA-ES optimization is an important factor, and a larger λ leads to better performance by costing more computation time.

CMA-ES implementation in MATLAB is not optimized and it is relatively slow with conventional CPUs. More specifically, for each hierarchical level the optimization takes $2-5$ minutes. Nevertheless, it is fast enough in IGNS settings where neurosurgeons generally spend about 10-20 min between collection

of US images and resection of the tumor. For the next step, we plan to implement our method with GPU in order to further accelerate the registration process. Finally, we aim to further test our method on more datasets in different applications.

5 Conclusion

Herein, we presented ARENA, an affine registration method to align US and MRI volumetric images. We applied our method on the RESECT database and the BITE database, validated our method qualitatively and quantitatively, and compared to two recently published registration methods. The qualitative results show that the registered images have improvements in alignment of salient image features. ARENA has consistently improved the mTRE in all patients in both databases, and is therefore a potentially promising registration method for use during IGNS.

Acknowledgements This work is funded by Natural Science Engineering Council of Canada (NSERC) grant RGPIN-2015-04136.

6 Compliance with Ethical Standards

Conflict of interest The authors declare that there is no conflict of interest.
Ethical standard All procedures perfomed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Declaration of Helsinki and its later amendments or comparable ethical standards.
Informed consent Informed consent was obtained from all participants included in the study.

References

1. S. Damas, O. Cordón, and J. Santamaría, "Medical image registration using evolutionary computation: An experimental survey," *IEEE Computational Intelligence Magazine*, vol. 6, no. 4, pp. 26–42, 2011.
2. J. Ma, H. Zhou, J. Zhao, Y. Gao, J. Jiang, and J. Tian, "Robust feature matching for remote sensing image registration via locally linear transforming," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 12, pp. 6469–6481, 2015.
3. A. P. James and B. V. Dasarathy, "Medical image fusion: A survey of the state of the art," *Information Fusion*, vol. 19, pp. 4–19, 2014.
4. S. Li, X. Kang, L. Fang, J. Hu, and H. Yin, "Pixel-level image fusion: A survey of the state of the art," *Information Fusion*, vol. 33, pp. 100–112, 2017.
5. Y. Yang, Y. Que, S. Huang, and P. Lin, "Multimodal sensor medical image fusion based on type-2 fuzzy logic in nsct domain," *IEEE Sensors Journal*, vol. 16, no. 10, pp. 3735–3745, 2016.
6. A. J. Golby, *Image-guided neurosurgery*. Academic Press, 2015.
7. L. Besharati Tabrizi and M. Mahvash, "Augmented reality–guided neurosurgery: accuracy and intraoperative application of an image projection technique," *Journal of neurosurgery*, vol. 123, no. 1, pp. 206–211, 2015.
8. C. R. Maurer and J. M. Fitzpatrick, "A review of medical image registration," *Interactive image-guided neurosurgery*, vol. 1, pp. 17–44, 1993.
9. I. J. Gerard, M. Kersten-Oertel, K. Petrecca, D. Sirhan, J. A. Hall, and D. L. Collins, "Brain shift in neuronavigation of brain tumors: A review," *Medical image analysis*, vol. 35, pp. 403–420, 2017.
10. S. Nag, "Image registration techniques: A survey," *arXiv preprint arXiv:1712.07540*, 2017.
11. J. A. Maintz and M. A. Viergever, "A survey of medical image registration," *Medical image analysis*, vol. 2, no. 1, pp. 1–36, 1998.
12. M. Gong, S. Zhao, L. Jiao, D. Tian, and S. Wang, "A novel coarse-to-fine scheme for automatic image registration based on sift and mutual information," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 7, pp. 4328–4338, 2014.
13. H. J. Johnson and G. E. Christensen, "Consistent landmark and intensity-based image registration," *IEEE transactions on medical imaging*, vol. 21, no. 5, pp. 450–461, 2002.
14. B. Zitova and J. Flusser, "Image registration methods: a survey," *Image and vision computing*, vol. 21, no. 11, pp. 977–1000, 2003.
15. G. Chandrashekar and F. Sahin, "A survey on feature selection methods," *Computers & Electrical Engineering*, vol. 40, no. 1, pp. 16–28, 2014.
16. D. Rueckert and P. Aljabar, "Nonrigid registration of medical images: Theory, methods, and applications [applications corner]," *IEEE Signal Processing Magazine*, vol. 27, no. 4, pp. 113–119, 2010.
17. A. Sotiras, C. Davatzikos, and N. Paragios, "Deformable medical image registration: A survey," *IEEE transactions on medical imaging*, vol. 32, no. 7, pp. 1153–1190, 2013.
18. C. X. Yan, B. Goulet, J. Pelletier, S. J.-S. Chen, D. Tampieri, and D. L. Collins, "Towards accurate, robust and practical ultrasound-ct registration of vertebrae for image-guided spine surgery," *International journal of computer assisted radiology and surgery*, vol. 6, no. 4, pp. 523–537, 2011.

19. S. Gill, P. Abolmaesumi, G. Fichtinger, J. Boisvert, D. Pichora, D. Borshneck, and P. Mousavi, "Biomechanically constrained groupwise ultrasound to ct registration of the lumbar spine," *Medical image analysis*, vol. 16, no. 3, pp. 662–674, 2012.
20. I. Hacihaliloglu, A. Rasoulian, R. N. Rohling, and P. Abolmaesumi, "Local phase tensor features for 3-d ultrasound to statistical shape+ pose spine model registration," *IEEE transactions on Medical Imaging*, vol. 33, no. 11, pp. 2167–2179, 2014.
21. G. Balakrishnan, A. Zhao, M. R. Sabuncu, J. Guttag, and A. V. Dalca, "An unsupervised learning model for deformable medical image registration," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 9252–9260, 2018.
22. O. Weistrand and S. Svensson, "The anaconda algorithm for deformable image registration in radiotherapy," *Medical physics*, vol. 42, no. 1, pp. 40–53, 2015.
23. B. Zhao, G. E. Christensen, J. Hyun Song, Y. Pan, S. E. Gerard, J. M. Reinhardt, K. Du, T. Patton, J. M. Bayouth, and G. D. Hugo, "Tissue-volume preserving deformable image registration for 4dct pulmonary images," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pp. 41–49, 2016.
24. F. Maes, D. Loeckx, D. Vandermeulen, and P. Suetens, "Image registration using mutual information," in *Handbook of Biomedical Imaging*, pp. 295–308, Springer, 2015.
25. A. Roche, G. Malandain, N. Ayache, and X. Pennec, *Multimodal image registration by maximization of the correlation ratio*. PhD thesis, INRIA, 1998.
26. A. Roche, X. Pennec, M. Rudolph, D. Auer, G. Malandain, S. Ourselin, L. M. Auer, and N. Ayache, "Generalized correlation ratio for rigid registration of 3d ultrasound with mr images," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 567–577, Springer, 2000.
27. H. Rivaz and D. L. Collins, "Deformable registration of preoperative mr, pre-resection ultrasound, and post-resection ultrasound images of neurosurgery," *International journal of computer assisted radiology and surgery*, vol. 10, no. 7, pp. 1017–1028, 2015.
28. N. Masoumi, Y. Xiao, and H. Rivaz, "Marcel (inter-modality affine registration with correlation ratio): An application for brain shift correction in ultrasound-guided brain tumor resection," in *International MICCAI Brainlesion Workshop*, pp. 55–63, Springer, 2017.
29. H. Rivaz, S. J.-S. Chen, and D. L. Collins, "Automatic deformable mr-ultrasound registration for image-guided neurosurgery," *IEEE transactions on medical imaging*, vol. 34, no. 2, pp. 366–380, 2015.
30. N. Hansen and A. Ostermeier, "Adapting arbitrary normal mutation distributions in evolution strategies: The covariance matrix adaptation," in *Evolutionary Computation, 1996., Proceedings of IEEE International Conference on*, pp. 312–317, IEEE, 1996.
31. Y. Xiao, M. Fortin, G. Unsgård, H. Rivaz, and I. Reinertsen, "Retrospective evaluation of cerebral tumors (resect): a clinical database of pre-operative mri and intra-operative ultrasound in low-grade glioma surgeries," *Medical physics*, 2017.
32. I. Machado, M. Toews, J. Luo, P. Unadkat, W. Essayed, E. George, P. Teodoro, H. Carvalho, J. Martins, P. Golland, and S. Pieper, "Non-rigid registration of 3d ultrasound for neurosurgery using automatic feature detection and matching," *International journal of computer assisted radiology and surgery*, pp. 1–14, 2018.
33. L. Mercier, R. F. Del Maestro, K. Petrecca, D. Araujo, C. Haegelen, and D. L. Collins, "Online database of clinical mr and ultrasound images of brain tumors," *Medical physics*, vol. 39, no. 6Part1, pp. 3253–3261, 2012.
34. S. Klein, M. Staring, and J. P. Pluim, "Evaluation of optimization methods for nonrigid medical image registration using mutual information and b-splines," *IEEE transactions on image processing*, vol. 16, no. 12, pp. 2879–2890, 2007.
35. S. Winter, B. Brendel, I. Pechlivanis, K. Schmieder, and C. Igel, "Registration of ct and intraoperative 3-d ultrasound images of the spine using evolutionary and gradient-based methods," *IEEE Transactions on Evolutionary Computation*, vol. 12, no. 3, pp. 284–296, 2008.
36. R. H. Gong and P. Abolmaesumi, "2d/3d registration with the cma-es method," in *Medical Imaging 2008: Visualization, Image-Guided Procedures, and Modeling*, vol. 6918, p. 69181M, International Society for Optics and Photonics, 2008.
37. Y. Otake, M. Armand, R. S. Armiger, M. D. Kutzer, E. Basafa, P. Kazanzides, and R. H. Taylor, "Intraoperative image-based multiview 2d/3d registration for image-guided orthopaedic surgery: incorporation of fiducial-based c-arm tracking and gpu-acceleration," *IEEE transactions on medical imaging*, vol. 31, no. 4, pp. 948–962, 2012.
38. E. Reinhard, W. Heidrich, P. Debevec, S. Pattanaik, G. Ward, and K. Myszkowski, *High dynamic range imaging: acquisition, display, and image-based lighting*. Morgan Kaufmann, 2010.
39. B. Fischer and J. Modersitzki, "Ill-posed medicineâĂŤan introduction to image registration," *Inverse Problems*, vol. 24, no. 3, p. 034008, 2008.
40. N. Hansen and A. Ostermeier, "Completely derandomized self-adaptation in evolution strategies," *Evolutionary computation*, vol. 9, no. 2, pp. 159–195, 2001.
41. "Cma-es in matlab - yarpiz."
42. W. Wein, A. Ladikos, B. Fuerst, A. Shah, K. Sharma, and N. Navab, "Global registration of ultrasound to mri using the lc 2 metric for enabling neurosurgical guidance," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 34–41, Springer, 2013.
43. M. P. Heinrich, M. Jenkinson, B. W. Papież, M. Brady, and J. A. Schnabel, "Towards realtime multimodal fusion for image-guided interventions using self-similarities," in *International conference on medical image computing and computer-assisted intervention*, pp. 187–194, Springer, 2013.
44. P. Daga, G. Winston, M. Modat, M. White, L. Mancini, M. J. Cardoso, M. Symms, J. Stretton, A. W. McEvoy, J. Thornton, and C. Micallef, "Accurate localization of optic radiation during neurosurgery in an interventional mri suite," *IEEE transactions on medical imaging*, vol. 31, no. 4, pp. 882–891, 2012.
45. J. M. Fitzpatrick, "Fiducial registration error and target registration error are uncorrelated," in *Medical Imaging 2009: Visualization, Image-Guided Procedures, and Modeling*, vol. 7261, p. 726102, International Society for Optics and Photonics, 2009.
46. X. Zhong, S. Bayer, N. Ravikumar, N. Strobel, A. Birkhold, M. Kowarschik, R. Fahrig, and A. Maier, "Resolve intraoperative brain shift as imitation game," in *Simulation, Image Processing, and Ultrasound Systems for Assisted Diagnosis and Navigation*, pp. 129–137, Springer, 2018.
47. J. Hong and H. Park, "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, Springer, 2018.

48. W. Wein, "Brain-shift correction with image-based registration and landmark accuracy evaluation," in *Simulation, Image Processing, and Ultrasound Systems for Assisted Diagnosis and Navigation*, pp. 146–151, Springer, 2018.
49. L. Sun and S. Zhang, "Deformable mri-ultrasound registration using 3d convolutional neural network," in *Simulation, Image Processing, and Ultrasound Systems for Assisted Diagnosis and Navigation*, pp. 152–158, Springer, 2018.
50. M. P. Heinrich, "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, Springer, 2018.
51. I. Machado, M. Toews, J. Luo, P. Unadkat, W. Essayed, E. George, P. Teodoro, H. Carvalho, J. Martins, and P. Golland, "Deformable mri-ultrasound registration via attribute matching and mutual-saliency weighting for image-guided neurosurgery," in *Simulation, Image Processing, and Ultrasound Systems for Assisted Diagnosis and Navigation*, pp. 165–171, Springer, 2018.
52. D. Drobny, T. Vercauteren, S. Ourselin, and M. Modat, "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, Springer, 2018.
53. R. Shams, M.-A. Boucher, and S. Kadoury, "Intra-operative brain shift correction with weighted locally linear correlations of 3dus and mri," in *Simulation, Image Processing, and Ultrasound Systems for Assisted Diagnosis and Navigation*, pp. 179–184, Springer, 2018.
54. J. D. Gibbons and S. Chakraborti, "Nonparametric statistical inference," in *International encyclopedia of statistical science*, pp. 977–979, Springer, 2011.
55. B. B. Avants, N. J. Tustison, G. Song, P. A. Cook, A. Klein, and J. C. Gee, "A reproducible evaluation of ants similarity metric performance in brain image registration," *Neuroimage*, vol. 54, no. 3, pp. 2033–2044, 2011.
56. M. Modat, M. J. Cardoso, P. Daga, D. Cash, N. C. Fox, and S. Ourselin, "Inverse-consistent symmetric free form deformation," in *International Workshop on Biomedical Image Registration*, pp. 79–88, Springer, 2012.
57. Y. Xiao, L. Eikenes, I. Reinertsen, and H. Rivaz, "Nonlinear deformation of tractography in ultrasound-guided low-grade gliomas resection," *International journal of computer assisted radiology and surgery*, vol. 13, no. 3, pp. 457–467, 2018.