

Automatic ultrasound–MRI registration for neurosurgery using the 2D and 3D LC^2 Metric



Bernhard Fuerst^{a,b,*}, Wolfgang Wein^c, Markus Müller^{a,c}, Nassir Navab^{a,b}

^a Computer Aided Medical Procedures (CAMP), Technische Universität München, Boltzmannstraße 3, 85748 Garching b. München, Germany

^b Computer Aided Medical Procedures (CAMP), Johns Hopkins University, 3400 North Charles Street, Baltimore, Maryland 21218, USA

^c ImFusion GmbH, Agnes-Pockels-Bogen 1, 80992 München, Germany

ARTICLE INFO

Article history:

Received 12 January 2014

Received in revised form 17 March 2014

Accepted 10 April 2014

Available online 2 May 2014

Keywords:

Image registration

Similarity measure

Linear correlation of linear combination

ABSTRACT

To enable image guided neurosurgery, the alignment of pre-interventional magnetic resonance imaging (MRI) and intra-operative ultrasound (US) is commonly required. We present two automatic image registration algorithms using the similarity measure Linear Correlation of Linear Combination (LC^2) to align either freehand US slices or US volumes with MRI images. Both approaches allow an automatic and robust registration, while the three dimensional method yields a significantly improved percentage of optimally aligned registrations for randomly chosen clinically relevant initializations. This study presents a detailed description of the methodology and an extensive evaluation showing an accuracy of 2.51 mm, precision of 0.85 mm and capture range of 15 mm (>95% convergence) using 14 clinical neurosurgical cases.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

Medical image registration is the process of spatially aligning images in a common coordinate space and aligning related features which exist in all images. It has been a widely investigated area in the past few decades, however remains challenging in particular for multi-modal registration. Often, different modalities complement each other well, which is relevant to a vast range of clinical applications for improving diagnosis, treatment planning, interventions, procedure follow-up, and screening. In a neurosurgical scenario which mainly motivates this work, MRI provides a good visualization of the anatomy and tumors, while US is inexpensive and allows for intra-operative use to detect and correct for brain shift after opening the skull. However, registering US and MRI images is a complex and difficult process, largely because represented information originates from very different physical properties. MRI intensities correlate with the relaxation times of the 1H nuclei, while the US intensity values represent the changes in acoustic impedance, overlaid by a significant speckle noise and various direction-dependent artifacts.

In this article, we present a new powerful set of methods based on the previously proposed LC^2 similarity measure (Wein et al., 2008), which allows for globally convergent, automatic registration of MRI and US data with clinically acceptable computation times.

2. Related work

During MRI and US registration a transformation is searched for which the alignment of the images is optimal. This requires a measure to evaluate the current alignment of the images, which is referred to as *cost functions* or *similarity measure*. Ideally, this function exhibits one distinctive extremum when the images are aligned optimally, and a nearly monotonous shape to provide optimal support in finding this extremum. In this section we will discuss several similarity measures which have been utilized for MR and US registration.

The MRI and US registration approaches using similarity measures based on sum of squared distances, Normalized Cross-Correlation (NCC), Mutual Information (MI) and normalized Mutual Information (nMI) tend to fail (Huang et al., 2005). This is caused by the very different nature of the intensity values and by structures that are not visible in one or the other imaging modality. For instance, details in MRI may lay in US shadow regions or certain materials can not be visualized by MRI (e.g. calcifications, air). Therefore, we focus on similarity measures specific to the application during MRI and US registration, which are not organ-specific and do not introduce a significant effort due to pre-processing, such as for liver vasculature presented by Penney et al. (2004).

Higher-dimensional Mutual Information (α -MI) is theoretically suited to assess US–MRI alignment based on both intensity and gradient information (in fact, an arbitrary number of features may be used). However, current approaches are neither practical

* Corresponding author. Tel.: +49(89)289 17058.

E-mail addresses: be.fuerst@tum.de (B. Fuerst), wein@imfusion.de (W. Wein), mueller@imfusion.de (M. Müller), nassir.navab@tum.de (N. Navab).

in terms of implementation effort nor computation time (Rivaz and Collins, 2012; Heinrich et al., 2013).

De Nigris et al. (2012) presented an interesting approach utilizing the alignment of high confidence gradient orientations. Anatomical boundaries characterized by the gradient orientations from the MRI and US images are used, while small regions with a high confidence for identifying anatomical boundaries were only selected from one image. However, the lack of the use of intensity values suggests that this method requires either nearly optimal data or a close initialization. Also the appearance of dominant gradients in one but not the other image, such as the skull in MRI but not in US, may lead to a poor alignment.

A different powerful method is the modality independent neighborhood descriptor (MIND) (Heinrich et al., 2012) and its extension self-similarity context (SSC) (Heinrich et al., 2013; Cifor et al., 2013), which utilize the differences of pre-defined neighborhood descriptors. They are based on a self similarity measure initially presented by Buades et al. (2005), and do not rely on the assumption of a global intensity relation. However, modality specific artifacts can not be considered and the computational effort for pre-processing is high due to the generation of voxel-wise neighborhood descriptors. Also, such self-similarity approaches tend to strongly abstract the image data, which might impact its accuracy as opposed to methods using the original image information.

Instead of comparing images from different modalities, pseudo-US images may be generated using segmented structures from MRI (Comeau et al., 2000; Coupé et al., 2012; King et al., 2010; Kuklisova-Murgasova et al., 2012). In light of the modality-specific considerations, the most promising general strategy for robust US–MRI registration, without relying on application-specific pre-processing or segmentation, is to compare US to both the MRI intensity and its gradient, as pioneered by Roche et al. (2001), where a global polynomial intensity relationship is fitted during registration. The alternating optimization of the rigid pose and the polynomial coefficients, as well as the fact that it is a global mapping, limit the convergence range though (the requirement for a local intensity mapping is explained in detail in Wein et al. (2008)). Powerful tools for image registration are similarity measures which are invariant to local changes, such as local normalized cross-correlation (invariant wrt. local brightness and contrast). Wein et al. (2008) introduced the similarity measure Linear Correlation of Linear Combination (LC^2), which exhibits local invariance to how much two channels of information contribute to an ultrasound image. The entire method has been specially designed for US–CT registration, where a strong correlation between X-ray attenuation coefficients and acoustic impedance is known, which allows a simulation of ultrasound effects from CT. These incorporate estimates of the acoustic attenuation, multiple reflections, and shadowing, which can not directly be estimated from MRI.

In this work, we adapt the LC^2 formulation to the case of MRI–US registration in neurosurgery, and evaluate it on a publicly available database of 14 patients. We had presented initial results of this method in Wein et al. (2013). Here, we provide a more thorough description and evaluation, and also add two novel alternative implementations of the LC^2 similarity, namely a 2D GPU version and a novel natively three-dimensional approach.

3. Method

3.1. Similarity measure

The similarity measure Linear Correlation of Linear Combination (LC^2) is used to search for a transformation T which aligns two images I and J . Due to the different nature of the images, a

relationship function f is required to allow a mapping of the intensity values. As a first step towards LC^2 a general cost function is defined, which applies the relationship function f to one of the images:

$$\min_T \sum_{\vec{x} \in \Phi} (I(\vec{x}) - f(J(T(\vec{x}))))^2, \quad (1)$$

where \vec{x} is a pixel or voxel position in the image domain Φ . If f is the identity function it can be seen that (1) represents sum of squared differences (SSD). In case of LC^2 , the relationship function f is defined to be a linear combination such as $f(J(\vec{y})) = \sum_{i=1 \dots n} c_i g_i(J(\vec{y}))$, where c_i are coefficients, and \vec{y} an arbitrary position in image J . The functions g_i describe an arbitrary pre-processing of the transformed image J , for instance the computation of gradients or simulation of pseudo-ultrasound images. The choice of relationship function and pre-processing strongly depends on the nature of the underlying images. For mono-modal image registration, one could assume that a simple linear function, such as $f(J(\vec{y})) = \alpha J(\vec{y}) + \beta$, would be sufficient, which basically represents a windowing function.

In case of a multi-modal image registration the relationship is more complex. As shown by Wein et al. (2008), the relationship between X-ray computed tomography (CT) and ultrasound images can be motivated from a physics standpoint, which results in f being a linear combination of ultrasound reflection and ultrasound echogeneity simulated from CT. However, in case of MRI and US registration such a simulation can not be obtained as the physical properties that are responsible for the intensity values have little in common. Therefore, the relationship function f is based on plausible observations. We assume that the US intensity value u_i for pixel/voxel i is either correlated with the MRI intensity value $p_i = J(T(\vec{x}_i))$ or with the MRI image gradient magnitude $g_i = |\nabla p_i|$. The resulting relationship function is therefore $f(J(\vec{x})) = \alpha p_i + \beta g_i + \gamma$. This caters to the fact that ultrasound intensity values on one hand may depict different soft tissue properties (due to the varying tissue inhomogeneities and echogeneity), and on the other hand represent tissue interfaces or gradients, as illustrated in Fig. 1. Of course LC^2 is not limited to linear combinations of only two components, but for MRI/US registration we currently do not see the need for additional components.

The coefficients $c = \{\alpha, \beta, \gamma\}$ of the linear combination in the relationship function are computed during every similarity measure evaluation, as shown by Wein et al. (2007, 2008). Under the assumption that the relationship function $f(J(T\vec{x}))$ shall be an optimal fit to the image $I(\vec{x})$, its coefficients can be implicitly estimated based on an ordinary least squares formulation:

$$\hat{c} = (M^T M)^{-1} M^T U, \quad (2)$$

$$\text{where } M = \begin{pmatrix} p_1 & g_1 & 1 \\ \vdots & \vdots & \vdots \\ p_m & g_m & 1 \end{pmatrix}, \quad U = \begin{pmatrix} u_1 \\ \vdots \\ u_m \end{pmatrix},$$

where m is the number of pixels/voxels in the domain, for instance $m = |\Phi|$.

It is now possible to compute a residual between the US and MRI images, but this proves to be problematic in regions in which the images do not overlap or in which the ultrasound image does not contain any structural information (e.g. shadow regions). Therefore, the cost function (1) is modified to penalize such regions by introducing the variance of the ultrasound image, as done by Roche et al. (1998, 2000, 2001), and is formulated as a similarity measure:

$$LC^2(I, J, T) = 1 - \frac{\sum_{\vec{x} \in \Phi} (I(\vec{x}) - f(J(T(\vec{x}))))^2}{|\Phi| \text{Var}(I)}. \quad (3)$$

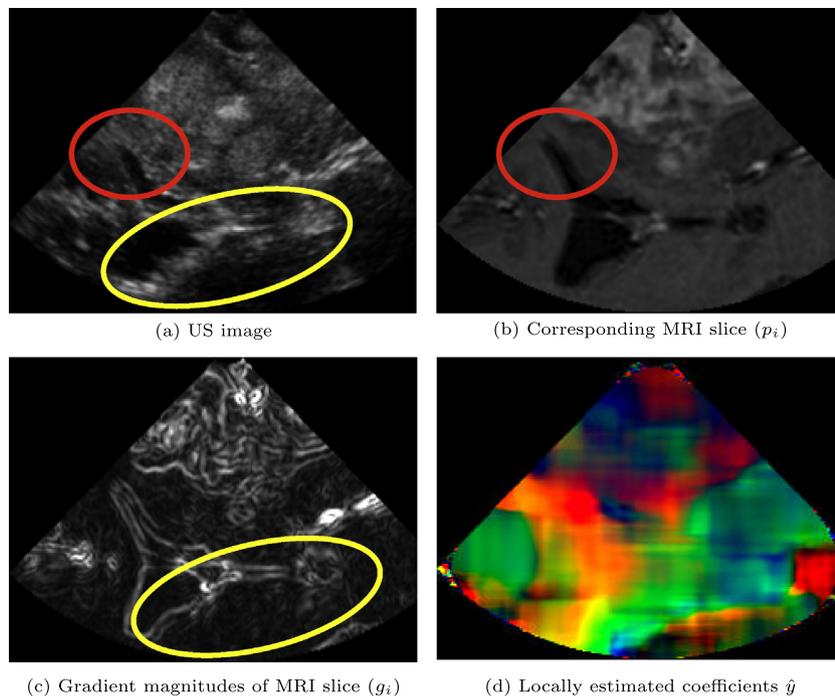


Fig. 1. The intensity values of the US image need to be expressed by a locally varying relationship function. This is indicated by the red circle, which depicts a high correspondence between US intensity values as MRI intensity values (a and b), and the yellow ellipse, which marks regions of high correspondence between the US image and the MRI gradient magnitudes (a and c). The locally estimated coefficients (for neighborhood $\Omega(\vec{x}, s)$) of the linear combination in the relationship function are visualized in (d). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

When plugging simple relationship functions in (3), it can be shown that other cost functions, such as correlation ratio (η) or Normalized Cross-Correlation (NCC), can be derived.

3.1.1. Locally normalized LC^2

The similarity measure (3) implies an image-wide constant relationship between US and MRI intensity values and MRI image gradient magnitudes. This, however, is not true in most cases, as visualized in Fig. 1. Therefore, we compute (3) for every pixel position $\vec{x} \in \Phi$ using a neighborhood $\Omega(\vec{x}, s) \subset \Phi$ of size s , and consequently replace the domain on which the coefficients \hat{y} are estimated (Eq. 2) with $\Omega(\vec{x}, s)$. The size s of the neighborhood around an image position \vec{x} defines a patch of $(2s + 1)^2$ pixels or a volume $(2s + 1)^3$ voxels in 2D or 3D respectively. This results in local similarity measures

$$LC^2(I, J, T, \vec{x}, s) = 1 - \frac{\sum_{\vec{y} \in \Omega(\vec{x}, s)} (I(\vec{y}) - f(J(T(\vec{y}))))^2}{|\Omega(\vec{x}, s)| \text{Var}(I(\vec{y}) | \vec{y} \in \Omega(\vec{x}, s))}, \quad (4)$$

where the weighted average is computed in order to obtain the image similarity measure. The weights are the local standard deviation of the image I in the neighborhood ($\sigma = \sqrt{\text{Var}(I(y) | y \in \Omega(\vec{x}, s))}$).

3.2. Similarity measure computation

3.2.1. Two-dimensional LC^2

This approach is computed on the original tracked freehand ultrasound slices, for which a 3D transformation is given. The extraction of the corresponding MRI intensity values and MRI gradient magnitudes from the given 3D volume is efficiently performed on the GPU using its hardware tri-linear interpolation capability. The accumulation of all the neighborhood intensities required for Eq. 4 is performed using a sliding-window approach

on the CPU, parallelized over rows and columns of the images on all threads of the multi-core processor. The computation time is therefore independent of the chosen LC^2 neighborhood size. While this method has been initially used in Wein et al. (2013), we have investigated an alternative full GPU implementation, where each shader accumulates the neighborhood information independently. The latter approach results in a computation dependency with respect to the neighborhood size of $O(s^2)$, and therefore offers superior performance for small neighborhood sizes.

3.2.2. Three dimensional LC^2

In order to investigate advantages of slice versus volume-based LC^2 computation, we have also implemented the similarity measure with three-dimensional neighborhood blocks on the GPU. Here, the intensity accumulation is implemented in a separable fashion for every dimension, such that the computation time scales linearly $O(s)$ with the neighborhood size s . Before the registration, the 3D freehand ultrasound data is reconstructed into a volume grid in a similar fashion as in Karamalis et al. (2009), using a quadrilateral interpolation for a good trade-off of performance and image quality.

3.3. Optimization of rigid transformation

An analytic derivation of LC^2 is difficult due to the least-squares fitting in (2) which is computed for every position in the US image. Therefore we use Bound Optimization BY Quadratic Approximation (BOBYQA) (Powell, 2009), which internally creates own derivative approximations. This results in fewer evaluations than most other search methods, and is therefore used throughout this paper. However, clinical requirements on capture range may ask for other techniques. In particular, global optimization techniques may be useful to perform a more excessive search within the specified bounds.

3.3.1. Deformable registration

After rigid registration, a free-from deformation (FFD) model using cubic splines can be fitted, where the deformation is applied on the MRI data J within the same GPU kernel which extracts MRI intensity and gradient magnitude. For that purpose, we place a configuration of $2 \times 2 \times 4$ control points within the bounding box of the registered ultrasound sweep. Then the same BOBYQA algorithm is used to optimize the displacement vectors for all control points.

4. Experiments

4.1. Clinical data and experimental setup

To evaluate our method and compare the results to other publications, we used a publicly available database containing Brain Images with Tumors for Evaluation from Montreal Neurological Institute (Mercier et al., 2012), with pre-operative T1-weighted MRI and pre-resection 3D freehand US from 14 patients. The pre-resection ultrasound has been acquired before opening the dura, and therefore only little deformation has occurred. Initial transformations and corresponding landmarks for each US–MRI pair are included (Table 1, lines 1 and 2). Therefore, we can provide ground truth evaluations, and denote the average Euclidean distance of the landmarks as Fiducial Registration Error (FRE).

4.1.1. Two dimensional LC^2

The MRI volumes were used as provided, while the higher-resolution US images were down-sampled such that their pixel sizes is smaller than twice the size of an MRI voxel. This guarantees that information provided by MRI voxels is never discarded when the tri-linear interpolation is used. Furthermore, US slices were skipped to avoid overlapping planes, resulting in an average distance between the slices of <1.5 mm or less due to slower scanning in the areas of interest.

4.1.2. Three dimensional LC^2

For the 3D experiments the freehand ultrasound data was first reconstructed into a cartesian volume grid with an isometric resolution of 0.3 mm, and afterwards further re-sampled as needed. We found that down-sampling the US volumes by the factor of three provides a good trade-off between fast computation times, a total

load of <2 GB memory on GPU, a high capture range and good accuracy. In addition, speckle noise is removed due to the smoothing effect. The used US volumes yield an isometric resolution of 0.9 mm. The MRI volumes were used as provided.

4.1.3. System specifications

All registration attempts were performed utilizing the parallel processing capabilities of the Graphics Processing Unit (GPU) of a workstation with an Intel i7-3770 CPU with 8 threads and a Nvidia GeForce GTX Titan GPU with 2688 cores and 6 GB memory.

4.2. Numerical analysis of LC^2 configuration

All convergence analyses were performed by carrying out 100 randomly initialized transformations (± 10 mm/ $^\circ$ in all 6 parameters) for each choice of a parameter.

4.2.1. Convergence analysis in terms of US slice spacing

Consistently good results are obtained when performing a registration using the 2D approach with an *slice spacing* of <5 mm, where *slice spacing* refers to the average euclidean distance between the centers of tracked ultrasound slices. For deformable registration, we chose <1.5 mm to make sure we are not missing even smallest structures.

4.2.2. Convergence analysis in terms of neighborhood size

The sensitivity of the 2D method has been investigated using both CPU and GPU implementation of the similarity measure. For neighborhood sizes 2–24 the accuracy is similar, while the percentage of successful registration attempts peak around 8 and 9, as depicted in Fig. 2(a) and (b). Furthermore, it can be seen that the two implementations yield nearly the same results. Therefore, we used neighborhood $s = 9$ (hence $m = (2 \cdot 9 + 1)^2 = 361$ pixels) as a fair trade-off between convergence and accuracy for all further experiments. Reported results are computed using the CPU implementation.

The 3D method requires an independent convergence analysis, as the neighborhood size describes a volume rather than a 2D patch. The accuracy is nearly constant for neighborhood sizes 2–7, which also exhibits the highest percentage of successful registrations (Fig. 2c). To balance computation time, which scales linearly with the neighborhood size, and the performance, a neighborhood size $s = 3$ (hence $m = (2 \cdot 3 + 1)^3 = 343$ voxels) is

Table 1

Overview of clinical data (Mercier et al., 2012), previous published results (De Nigris et al., 2012; Rivaz and Collins, 2012), and results using our method for rigid and deformable registration including computation times. The presented registration results using 2D and 3D LC^2 are based on 100 randomly initialized registration attempts for each patient and method.

Patient	1	2	3	4	5	6	7	8	9	10	11	12	13	14	Mean
<i>Dataset overview and related methods</i>															
1 # of Tags	37	35	40	32	31	37	19	23	21	25	25	21	23	23	–
2 Initial FRE (mm)	4.93	6.30	9.38	3.93	2.62	2.30	3.04	3.75	5.09	2.99	1.52	3.70	5.15	3.77	4.18 \pm 5.20
3 US spacing (mm)	0.24	0.42	0.23	0.20	0.25	0.17	0.24	0.18	0.18	0.22	0.16	0.18	0.21	0.19	0.22 \pm 0.20
4 FRE in (De Nigris et al., 2012) (mm)	4.89	1.79	2.73	1.68	2.12	1.81	2.51	2.63	2.7	1.95	1.56	2.64	3.47	2.94	2.53 \pm 0.87
5 FRE in (Rivaz and Collins, 2012) (mm)	–	2.05	2.76	1.92	2.71	1.89	2.05	2.89	2.93	2.75	1.28	2.67	2.82	2.34	2.57 \pm 0.82
<i>Registration results using 2D LC^2</i>															
6 FRE (mm)	4.82	1.73	2.76	1.96	2.14	1.94	2.33	2.87	2.81	2.06	2.18	2.67	3.58	2.48	2.52 \pm 0.87
7 SD (mm)	0.01	0.01	0.01	0.01	0.02	0.01	0.05	0.30	0.02	0.00	0.03	0.15	0.05	0.04	0.05 \pm 0.08
8 Duration (sec)	5.9	8.3	11.1	5.7	7.1	8.2	18.2	8.6	6.0	23.4	17.3	25.8	8.1	7.0	11.5 \pm 6.8
9 Convergence (%)	54	66	44	60	75	38	51	19	54	76	82	94	53	44	57.9 \pm 19.6
10 FRE def. (mm)	4.95	1.64	2.43	1.91	2.26	2.2	2.52	3.64	2.65	2.09	1.76	2.45	3.71	2.76	2.64 \pm 0.9
11 Time def. (sec)	158	141	279	92	133	166	563	312	76	675	597	93	106	282	262 \pm 204
<i>Registration Results using 3D LC^2</i>															
12 FRE (mm)	4.86	1.70	2.55	1.73	1.96	1.83	2.32	2.68	2.74	2.09	1.81	2.71	3.44	2.45	2.49 \pm 0.84
13 SD (mm)	0.01	0.00	0.00	0.01	0.01	0.01	0.02	0.03	0.01	0.01	0.02	0.03	0.02	0.01	0.01 \pm 0.01
14 Duration (sec)	2.61	2.91	2.89	1.30	2.15	2.08	2.87	2.68	1.49	2.58	2.23	2.64	1.24	2.79	2.32 \pm 0.59
15 Convergence (%)	96	85	88	89	90	80	81	76	80	91	100	94	86	92	87.7 \pm 6.8

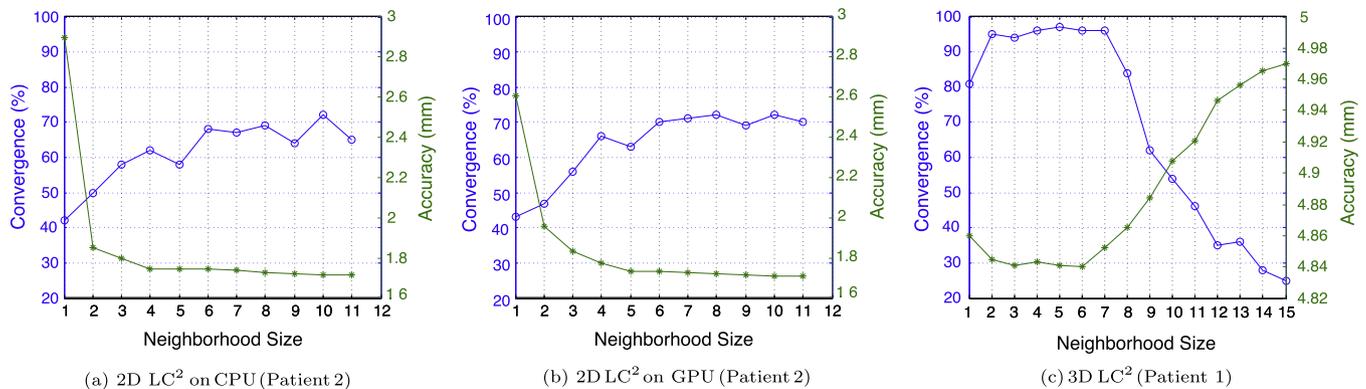


Fig. 2. Relationship between convergence (blue), accuracy (green) and the neighborhood size s for the three different LC^2 methods. Differences between the two 2D implementations (a and b) are caused by different numerical behavior in the implementation (e.g. on the CPU everything is done with double precision, on the GPU only the final matrix inversion), not by differences in the methodology. The convergence analysis indicates that a neighborhood size of $s = 9$ ($m = 361$ pixels) yields a fair trade-off between convergence, accuracy and computation time. The typical behavior of the 3D approach is shown in (c), indicating $s = 3$ ($m = 343$ voxels) being a good choice of the neighborhood size. Note that patient 2 and 1 show the typical behavior of the 2D and 3D approaches respectively and are therefore plotted here. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

used for all further experiments. Overly large patches result in a global mapping of MRI intensity and gradient, removing the main advantage of LC^2 over other methods (robustness wrt. local changes of intensity-gradient relationship).

4.2.3. Gradient magnitudes vs. directed gradients

Finally, we have investigated the effect of using the dot product of the MRI gradient g with the US beam direction, instead of g directly. This reduces the influence of vertical gradients, similar to the US simulation presented by Wein et al. (2008). Interestingly, this results in 10–25% more outliers (the cost function becomes more non-linear due to the added directional dependence).

4.3. Registration results

The resulting errors for all 14 patient data sets are nearly the same for registrations using gradient orientation alignment (De Nigris et al., 2012), 2D LC^2 , or 3D LC^2 , as depicted in Table 1, lines 4, 6, 12 respectively. Rivaz and Collins (2012) report slightly higher errors when applying a costly deformable registration which requires several hours (line 5). The increased FRE coincides with our findings when using the faster 2D LC^2 during deformable registrations (line 10). This indicates that performing deformable registration does not provide any significant benefit when applied to mostly rigid data sets. However, we are convinced that the change of landmark errors induced by deformable registration or the difference between the 2D and 3D approach of LC^2 lay within the range of the fiducial localization error (FLE) of the data. Examples of initially aligned and registered images are shown in Fig. 3.

4.4. Accuracy, precision and capture range

Some initial alignments yield significant errors (e.g. patients 2, 3, 9, 13), which are reduced by all algorithms listed in Table 1. Therefore, an analysis of the capability of our algorithm to reach the optimum under all conditions is necessary.

Trials with each 100 randomly initialized transformations (± 10 mm/ $^\circ$ in all 6 parameters) were performed for all 14 patients using the 2D and 3D LC^2 approach, resulting in a total of 2800 registration attempts. Comparing the LC^2 similarity measure with the final FRE shows in all cases that the best transformation corresponds to the highest similarity and that the misalignments are clearly separated yielding a significantly lower similarity. This demonstrates that both LC^2 algorithms allow for global registration in a realistic clinical setup. Fig. 4 shows the results, including the

percentages of the converged optimizations. The average errors (*accuracy*) and standard deviation (SD; *precision*) are listed in Table 1.

The *capture range* describes the range of initial FRE values for which >95% of the registration approaches are successful. When using the 2D approach, it can be observed that 95% of the experiments converge within an initial FRE of 9 mm. This capture range is significantly increased to 15 mm when using the 3D approach. Both sets of experiments are based on the aforementioned randomly initialized studies using BOBYQA. Fig. 5(a) and (b) depict the capture range for selected patients, while (c) shows the total percentage of outliers vs. initial FRE for both algorithms.

Since the gradient orientation alignment (GOA) method (De Nigris et al., 2012) yields similar FRE values, we implemented it to the best of our knowledge and re-ran the aforementioned randomized trials with it. We obtain >90% outliers and further investigation into the cost function properties revealed that only a minor local optimum is present. A possible explanation is, that without further heuristics the GOA method would line up strong gradients from e.g. dura mater or skull; besides, using only gradients larger than a threshold limits the image content considered. While we believe these to be general issues, it has to be acknowledged that better results would probably be obtained by the original authors, e.g. by changing implementation details such as resolution, smoothing and interpolation. Precision and capture range were not reported in their work though, as unfortunately too often the case.

The randomized experiments with the 2D and 3D LC^2 versions have been compared using the Mann–Whitney U -test. Generally, a p -value of ≤ 0.01 is considered *very significant* (Nuzzo, 2014). In fact, the 3D approach shows such a *very significant* increased convergence in all but one patients (patient 12 unchanged at 94%). With the exception for patient 12, all computed p -values are below 6×10^{-3} , indicating that the chance of the improvement being coincidentally observed is less than 0.6%.

4.5. Computation time

To compare the performance of the 2D CPU, 2D GPU and 3D GPU approaches, we have measured the average computation times. All randomly initialized registrations have been run on the same hardware. The reported times do not include the file loading, ultrasound volume reconstruction and down-sampling.

The first, original 2D implementation uses the GPU merely to perform the MRI slice extraction, while the similarity is computed

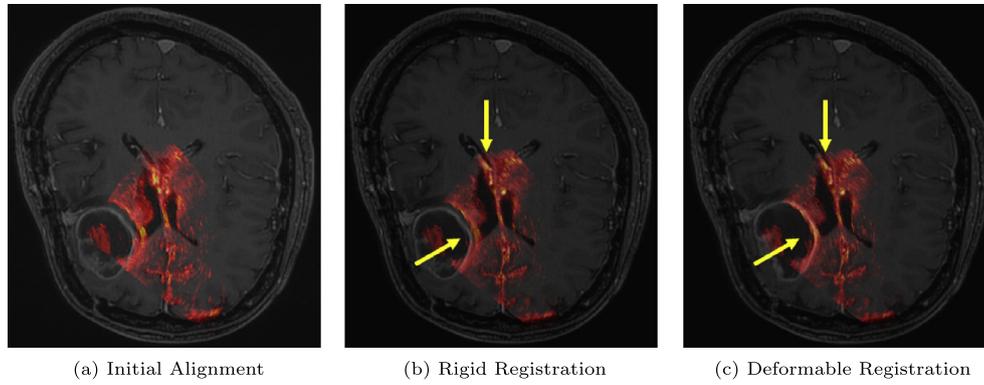


Fig. 3. Superimposed US on axial MRI slice to show the registration result of patient 6. The FRE improves by 0.11 mm from rigid to deformable registration.

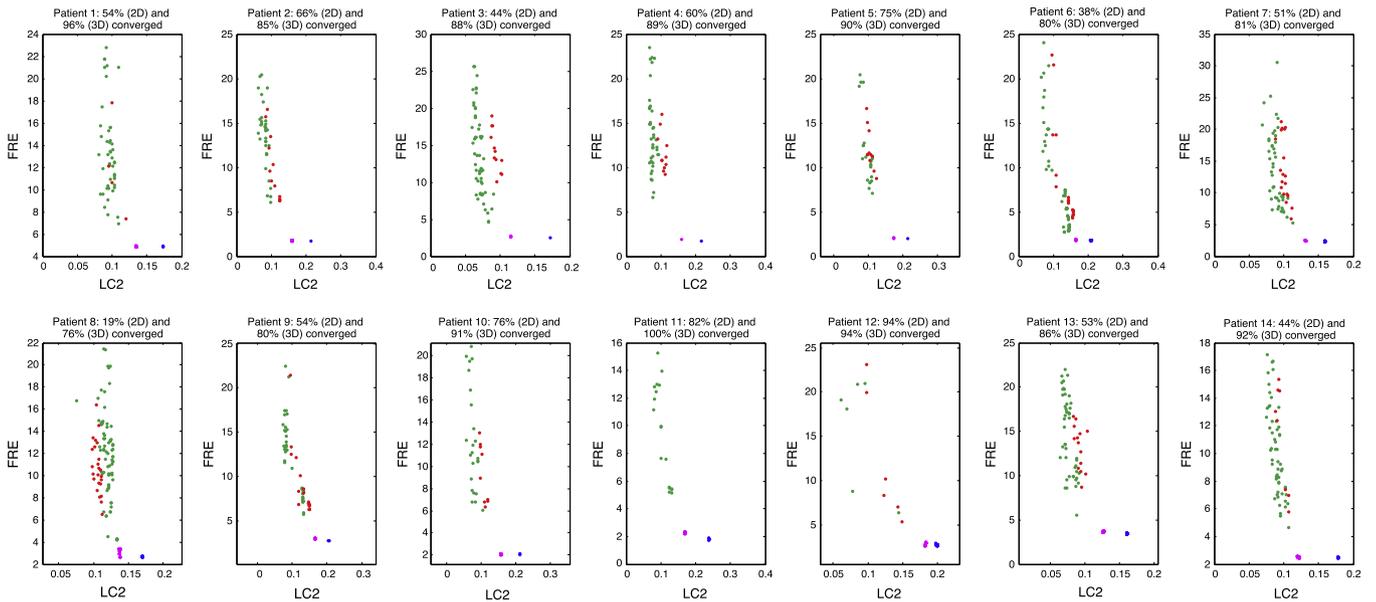


Fig. 4. Registration experiments with 100 randomly initialized transformations for all patient data sets. The converged results (2D: magenta, 3D: blue) are clearly separated from the failed registration attempts (outliers: 2D: green, 3D: red). Convergence is defined based on typical patient-specific FRE values as reported in Table 1. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

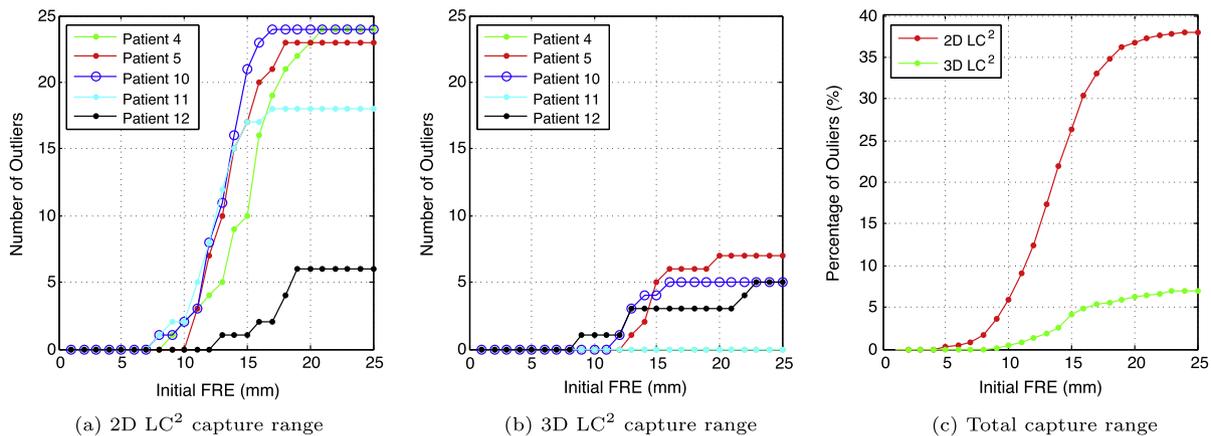


Fig. 5. The capture range describes the relation between outliers and initial FRE. For selected patients this relation is plotted in (a) and (b) for the 2D and 3D LC^2 approach respectively. The total capture range of the algorithm for all patients is plotted in (c). It can be seen that the capture range (less than 5% outliers) for 2D LC^2 is 9 mm, while it is increased to 15 mm for the 3D approach.

using a sliding-window approach on the CPU. The average run time is 11.6 sec. In the second 2D implementation we have performed the similarity evaluation using a GPU based filtering strategy,

which reduces the average run time by only 5%, with quadratic dependency on the LC^2 neighborhood size. Here, the similarity measure computation does not utilize the full capabilities of the

GPU yet, since only a parallelization within the (small) ultrasound frames is used.

Finally, we measured the average registration time for the GPU based 3D approach. Due to the implementation as separable filter, the performance scales linearly with the neighborhood size. Using a neighborhood size $s = 3$ an average run time of 2.32 ± 0.59 sec can be observed.

5. Discussion and conclusion

5.1. Choice of method

The LC^2 computation on the original freehand ultrasound 2D images yields a number of advantages. First, an offline volume reconstruction step is avoided, which might reduce the quality of the original ultrasound image information. Considering the limited voxel resolution of the MRI data this, however, does not pose a problem since we need to further down-sample the ultrasound data anyway. More importantly, a slice-based approach may immediately start looking for the correct alignment in real-time once the first frames are obtained. Besides, an optimization of the calibration transformation or compensation of tracking errors may be computed on-the-fly, which would otherwise require re-computation of the 3D ultrasound volume.

The 3D volume-based approach yields a superior capture range, which can be explained by the fact that the LC^2 neighborhood stretches into the third dimension, therefore allowing even more consistent local matching of the relationship function f . This comes with the trade-off of higher computing resource requirements. While our implementation as fully parallelized separable filters in every dimension on the GPU ensures efficient computation, it requires a significant memory footprint (<2 GB for US voxel size 0.9 mm, 2–3 GB for US voxel size 0.6 mm).

Finally, a comparison of the same 2D approach on CPU and GPU depends very much on the actual implementation. In our case, the CPU method is completely independent of the neighborhood size s , whereas the GPU version has a squared dependence. The latter has room for improvement in terms of both a separable computation and parallelization over the US frames. It is also important to note that numerically equally stable behavior on the GPU is only obtained when the actual computation of Eq. (2) is performed with double floating-point precision. This in turn mandates a certain choice (and cost) of GPU hardware.

5.2. Performance

As opposed to other proposed methods such as De Nigris et al. (2012); Heinrich et al. (2013), the LC^2 method uses the full amount of available image information from both modalities, by locally matching ultrasound intensities to both MRI intensity and gradient magnitude. A direct comparison is unfortunately not possible, since other studies lack an evaluation of precision and capture range (which is crucial for developing an automatic registration in a given clinical context). The accuracy in terms of the point-based registration error is consistent with other studies, and is always <1 mm higher than the best possible rigid fit of the point correspondences themselves. Deformable registration does neither significantly increase or decrease the errors, however the visual alignment generally improves (see Fig. 3).

The computation time of the 2D and 3D GPU implementations allow for 100–500 cost function evaluations per second. Hence global registration within a clinically realistic bounding box of the pose parameters is possible. To our knowledge, this has not been shown in related work to date.

The LC^2 similarity measure has a single parameter, namely the neighborhood size s . We have shown that both the 2D and 3D variants work well within a fairly large range of s . However, as common in image registration scenarios, further tunable parameters arise in the pre-processing (e.g. down-sampling, volume reconstruction) of the image data, as well as the choice and configuration of a non-linear optimization algorithm.

5.3. Future work

While our approach adequately solves the problem of image-based MRI–US registration for correcting brain shift, some further work is required to address resection follow-up. Missing anatomical correspondence in the area of a tumor resection site might result in unstable registration results of our straight-forward free-form deformation model, when looking at the registration of pre- and post-resection ultrasound data. In order to provide a clinically meaningful solution in this context, however, a detailed discussion with physicians will be required in the first place, to establish how such data before and after the procedure shall be transformed and compared.

In other clinical application areas, such as oncological diagnosis in the abdomen, it would be desirable to look into extensions of our deformation model to incorporate physical constraints of the complex deformations due to respiratory motion and patient positioning. The same holds true for potential applications in the case of prostate MRI–US fusion (Weiss et al., 2010). The incorporation of bio-mechanical deformation models might be particularly suited to address such difficult non-linear registration problems.

Acknowledgements

The authors affiliated with Technische Universität München are partially supported by the EU 7th Framework Program (FP7/2007–2013 and FP7/ICT-2009–6) under Grant Agreements Nos. 256984 (EndoTOPPET-US) and 270460 (ACTIVE).

References

- Buades, A., Coll, B., Morel, J.M., 2005. A non-local algorithm for image denoising. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE, pp. 60–65.
- Cifor, A., Risser, L., Heinrich, M.P., Chung, D., Schnabel, J.A., 2013. Rigid registration of untracked freehand 2d ultrasound sweeps to 3d ct of liver tumours. In: Abdominal Imaging. Computation and Clinical Applications. Springer, pp. 155–164.
- Comeau, R.M., Sadikot, A.F., Fenster, A., Peters, T.M., 2000. Intraoperative ultrasound for guidance and tissue shift correction in image-guided neurosurgery. *Med. Phys.* 27, 787.
- Coupé, P., Hellier, P., Morandi, X., Barillot, C., 2012. 3D rigid registration of intraoperative ultrasound and preoperative MR brain images based on hyperechogenic structures. *J. Biomed. Imag.* 2012, 1.
- De Nigris, D., Collins, D., Arbel, T., 2012. Fast and robust registration based on gradient orientations: case study matching intra-operative ultrasound to pre-operative MRI in neurosurgery. *IPCAI*, 125–134.
- Heinrich, M.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.* 16, 1423–1435.
- Heinrich, M.P., Jenkinson, M., Papież, B.W., Brady, M., Schnabel, J.A., 2013. Towards realtime multimodal fusion for image-guided interventions using self-similarities. In: *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2013*. Springer, pp. 187–194.
- Huang, X., Hill, N., Ren, J., Guiraudon, G., Boughner, D., Peters, T., 2005. Dynamic 3D ultrasound and MR image registration of the beating heart. *Med. Image Comput. Comput.-Assist. Interv.* 2005, 171–178.
- Karamalis, A., Wein, W., Kutter, O., Navab, N., 2009. Fast hybrid freehand ultrasound volume reconstruction. In: *Proc. of SPIE*, pp. 726114–1.
- King, A., Rhode, K., Ma, Y., Yao, C., Jansen, C., Razavi, R., Penney, G., 2010. Registering preprocedure volumetric images with intraprocedure 3-d ultrasound using an ultrasound imaging model. *IEEE Trans. Med. Imag.* 29, 924–937.
- Kuklisova-Murgasova, M., Cifor, A., Napolitano, R., Papageorghiou, A., Quaghebeur, G., Noble, J., Schnabel, J., 2012. Registration of 3D fetal brain US and MRI. *MICCAI*, 667–674.

- Mercier, L., Del Maestro, R., Petrecca, K., Araujo, D., Haegelen, C., Collins, D., 2012. Online database of clinical MR and ultrasound images of brain tumors. *Med. Phys.* 39, 3253.
- Nuzzo, R., 2014. Scientific method: statistical errors. *Nature* 506, 150–152.
- Penney, G., Blackall, J., Hamady, M., Sabharwal, T., Adam, A., Hawkes, D., et al., 2004. Registration of freehand 3D ultrasound and magnetic resonance liver images. *Med. Image Anal.* 8, 81–91.
- Powell, M.J., 2009. The BOBYQA Algorithm for Bound Constrained Optimization without Derivatives. Cambridge Report NA2009/06, University of Cambridge.
- Rivaz, H., Collins, D., 2012. Self-similarity weighted mutual information: a new nonrigid image registration metric. *MICCAI*, 91–98.
- Roche, A., Malandain, G., Pennec, X., Ayache, N., 1998. The correlation ratio as a new similarity measure for multimodal image registration. In: *Medical Image Computing and Computer-Assisted Intervention–MICCAI98*. Springer, pp. 1115–1124.
- Roche, A., Pennec, X., Rudolph, M., Auer, D., Malandain, G., Ourselin, S., Auer, L.M., Ayache, N., 2000. Generalized correlation ratio for rigid registration of 3d ultrasound with mr images. In: *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2000*. Springer, pp. 567–577.
- Roche, A., Pennec, X., Malandain, G., Ayache, N., 2001. Rigid registration of 3-D ultrasound with MR images: a new approach combining intensity and gradient information. *IEEE Trans. Med. Imag.* 20, 1038–1049.
- Wein, W., Khamene, A., Clevert, D., Kutter, O., Navab, N., 2007. Simulation and fully automatic multimodal registration of medical ultrasound. *Medical Image Computing and Computer-Assisted Intervention 2007*, 136–143.
- Wein, W., Brunke, S., Khamene, A., Callstrom, M., Navab, N., 2008. Automatic CT-ultrasound registration for diagnostic imaging and image-guided intervention. *Med. Image Anal.* 12, 577.
- Wein, W., Ladikos, A., Fuerst, B., Shah, A., Sharma, K., Navab, N., 2013. Global registration of ultrasound to MRI using the LC^2 metric for enabling neurosurgical guidance. In: *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2013*. Springer, pp. 34–41.
- Weiss, C., Seitz, M., Herrmann, K., Graser, A., Kiefer, B., Requardt, M., Fehre, J., Nanke, R., Diallo, M., Kamen, A., Wein, W., 2010. MR–US fusion for targeted prostate biopsy. In: *International Society for Magnetic Resonance in Medicine (ISMRM) Proceedings*.