



EVolution: an Edge-based Variational method for non-rigid multi-modal image registration

Baudouin Denis de Senneville, Cornel Zachiu, Mario Ries, Chrit Moonen

► To cite this version:

Baudouin Denis de Senneville, Cornel Zachiu, Mario Ries, Chrit Moonen. EVolution: an Edge-based Variational method for non-rigid multi-modal image registration. *Physics in Medicine and Biology*, 2016, 61 (20), pp.7377-7396. 10.1088/0031-9155/61/20/7377 . hal-01365989

HAL Id: hal-01365989

<https://hal.science/hal-01365989>

Submitted on 21 Dec 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

EVolution: an edge-based variational method for non-rigid multi-modal image registration

This content has been downloaded from IOPscience. Please scroll down to see the full text.

2016 Phys. Med. Biol. 61 7377

(<http://iopscience.iop.org/0031-9155/61/20/7377>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 87.91.200.174

This content was downloaded on 05/01/2017 at 13:57

Please note that [terms and conditions apply](#).

You may also be interested in:

[An improved optical flow tracking technique for real-time MR-guided beam therapies in moving organs](#)

C Zachiu, N Papadakis, M Ries et al.

[3D nonrigid medical image registration using a new information theoretic measure](#)

Bicao Li, Guanyu Yang, Jean Louis Coatrieux et al.

[Non-rigid imageregistration](#)

Yiu H Lau, Michael Braun and Brian F Hutton

[Medical image registration](#)

Derek L G Hill, Philipp G Batchelor, Mark Holden et al.

[The adaptive FEM elastic model for medical image registration](#)

Jingya Zhang, Jiajun Wang, Xiuying Wang et al.

[Performance evaluation of MIND demons deformable registration of MR and CT images in spinal interventions](#)

S Reaungamornrat, T De Silva, A Uneri et al.

[Non-rigid registration of medical images based on estimation of deformation states](#)

Bahram Marami, Shahin Sirouspour and David W Capson



EVolution: an edge-based variational method for non-rigid multi-modal image registration

B Denis de Senneville^{1,2,3}, C Zachiu¹, M Ries¹ and C Moonen¹

¹ Imaging Division, UMC Utrecht, Heidelberglaan 100, 3584 CX, Utrecht, Netherlands

² ‘Institut de Mathématiques de Bordeaux’, Université Bordeaux/CNRS UMR 5251, 351 Cours de la Libération, 33405 Talence Cedex, France

E-mail: b.desenneville@umcutrecht.nl, C.Zachiu@umcutrecht.nl, M.Ries@umcutrecht.nl and C.Moonen@umcutrecht.nl

Received 1 June 2016, revised 21 July 2016

Accepted for publication 30 August 2016

Published 3 October 2016



Abstract

Image registration is part of a large variety of medical applications including diagnosis, monitoring disease progression and/or treatment effectiveness and, more recently, therapy guidance. Such applications usually involve several imaging modalities such as ultrasound, computed tomography, positron emission tomography, x-ray or magnetic resonance imaging, either separately or combined.

In the current work, we propose a non-rigid multi-modal registration method (namely EVolution: an edge-based variational method for non-rigid multi-modal image registration) that aims at maximizing edge alignment between the images being registered. The proposed algorithm requires only contrasts between physiological tissues, preferably present in both image modalities, and assumes deformable/elastic tissues. Given both is shown to be well suitable for non-rigid co-registration across different image types/contrasts (T1/T2) as well as different modalities (CT/MRI). This is achieved using a variational scheme that provides a fast algorithm with a low number of control parameters. Results obtained on an annotated CT data set were comparable to the ones provided by state-of-the-art multi-modal image registration algorithms, for all tested experimental conditions (image pre-filtering, image intensity variation, noise perturbation). Moreover, we

³ Author to whom any correspondence should be addressed.



Original content from this work may be used under the terms of the [Creative Commons Attribution 3.0 licence](https://creativecommons.org/licenses/by/3.0/). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

demonstrate that, compared to existing approaches, our method possesses increased robustness to transient structures (i.e. that are only present in some of the images).

Keywords: non-rigid registration, multi-modal registration, variational method

(Some figures may appear in colour only in the online journal)

1. Introduction

Image registration is the process of aligning two or more images of the same scene acquired at different time instants, using different sensors and/or from a different point-of-view. Although initially proposed in the field of digital photography and video sequence processing, as medical imaging technologies progressed, it has become an increasingly important pre-processing step for medical image analysis. Currently, image registration is part of a large variety of medical applications including diagnosis, monitoring disease progression and/or treatment effectiveness and, more recently, therapy guidance (Mani and Arivazhagan 2013). Such applications usually involve several imaging modalities such as ultrasound (US), computed tomography (CT), positron emission tomography (PET), x-ray or magnetic resonance imaging (MRI), either separately or combined. Whether one or more imaging modalities are employed, image registration algorithms can be divided in two categories: mono-modal and respectively multi-modal. A variety of each has already been proposed in the medical image processing literature (Hill *et al* 2001). Mono-modal registration algorithms typically rely on the assumption that an anatomical structure is present in all images included in the registration process and also that the structures preserve, to a certain extent, their gray-level intensities. Such algorithms are well suited, for example, for tumor growth monitoring and intervention verification (Maintz and Viergever 1998). Multi-modal algorithms relax the gray-level intensity conservation constraint. Moreover, a certain variety do not even require a structure to be present in all images. Cross-modality registration algorithms are more suitable for diagnostic purposes, since different imaging modalities usually provide complementary information. For example, the excellent anatomical information provided by an MRI image could be complemented by the physiological information present in a PET scan, each modality having, in general, more vast capabilities than the other in terms of the specified provided information (Mani and Arivazhagan 2013). Compared to mono-modal algorithms, multi-modal methods are usually more complex and require additional computational resources.

A particular variety of registration algorithms are the so-called variational methods which aim at minimizing a cost function in order to find the transformation between the registered images (Weickert *et al* 2003). Since rigid/affine transformations have a limited (although important) purpose in medical image registration, the focus in this paper will be on deformable transformations. While different in nature, all variational methods usually employ cost functions that include two components: a data fidelity term and respectively a regularization term. The data fidelity term measures the similarity between the images undergoing the registration process, while the regularization term imposes constraints on the estimated deformation, since the minimization of the data fidelity term alone is usually an ill-posed problem. Variational methods are a particularly attractive solution for medical image registration since they require a low number of input/control parameters and typically involve fast numerical schemes (Ries *et al* 2010, Denis de Senneville *et al* 2015). Whether they are able to cope with images acquired through different modalities depends mostly on the similarity measure used

by the data fidelity term. While for mono-modal registration algorithms the sum of squared differences (SSD) applied directly on the images might suffice (Horn and Schunck 1981, Denis de Senneville *et al* 2011), such a measure is highly unsuitable for registering across modalities. Modality independent similarity measures are thus necessary.

Previous studies have addressed the cross-modality registration problem by pre-processing the input images into modality-independent scalar representations of themselves. Once pre-processed, the SSD was used as a similarity metric between the images under the new representation. Examples of modality independent scalar representations include the local phase, local entropy or gradient orientation (Mellor and Brady 2005, Haber and Modersitzki 2006, Wachinger and Navab 2012). While indeed showing a good performance in both registration quality and computational demands, these methods are hampered by the fact that for more challenging multi-modal registration tasks, the features present in the images under the new representation might not be discriminative enough.

Another broad category of multi-modal registration algorithms makes use of the mutual information (MI) (directly or variations) as an image similarity measure. Initially proposed independently by Viola and Wells (Viola and Wells 1997) and respectively by Maes *et al* (1997), mutual information-based multi-modal registration algorithms were met with great interest by the medical image registration community, triggering a long series of studies showcasing the great potential of such approaches (Pluim *et al* 2003). However, MI is inherently a global measure, making its local estimation a rather challenging task. For this reason, while having excellent performance for multi-modal rigid/affine realignment tasks, it becomes more difficult to apply when estimating local deformations. In addition, optimizing the mutual information between two or more images is computationally more complex and slow compared to an SSD-based similarity measure.

Recently, a novel cross-modality registration algorithm has been proposed by Heinrich *et al* based on so-called modality independent neighborhood descriptors (MIND) (Heinrich *et al* 2012). The method transforms the input images into a vectorial representation such that a descriptor is associated to each voxel, reflecting its relationship to the surrounding voxels. The descriptors are constructed based on the notion of self-similarity as initially proposed by Buades *et al* in the design of their non-local filtering method (Buades *et al* 2005). This results in a registration algorithm that is sensitive to structural information rather than intensity. For the test cases considered in the original paper, the MIND algorithm has shown increased performance over several other popular multi-modal registration methods. Moreover, since the similarity criterion uses the SSD to measure the distance between the descriptor images, it allows for fast optimization numerical schemes to be employed. However, the method is prone to the same limitations as the non-local means filter (Buades *et al* 2005): it is sensitive to several parameters such as the local estimate of the noise variance, the patch size around the central voxel used to compute each element of the descriptors, the number of voxels included in the computation of the descriptors, and it is affected by low contrasts. In addition, if the structural information is different from one image to the other, for example, if an area that is smooth in one image and textured in the other, then the descriptors cannot be compared reliably. More precisely, the method hypothesizes that each anatomical structure in one image has its counterpart in the other. This may not be the case in several situations, including transient structures occurring in the field-of-view as a result of peristaltic activity (in case the analysis is conducted in the abdomen), the appearance/disappearance of pathologies within the analyzed organ and/or parts of the region-of-interest or its surroundings entering/exiting the field-of-view.

A different approach to cross-contrast registration are algorithms that maximize the edge alignment between the registered images. Various metrics that quantify the extent to which

the edges are aligned have already been proposed in the literature. The methods were dedicated to registering either video sequences acquired with different sensors or medical images acquired through different modalities (Pluim *et al* 2000, Sun *et al* 2004, Sutour *et al* 2015). In this approach, the underlying assumption that needs to be fulfilled is that the boundaries and details of physiologic structures show image contrast to the surrounding tissue with both image modalities. However, the studies have only gone so far as to estimate rigid/affine or coarsely deformable transformations (using splines). In the current paper, by using the principle behind the edge alignment methods, we propose EVolution: an edge-based Variational algorithm for multi-modal image registration. Our contribution is four-fold:

- By construction, the algorithm is designed to increase the robustness of the registration process against structural information variations from one image to the other.
- A patch-based approach is designed to leverage limitations arising from the above mentioned scalar representation, especially for highly challenging multi-modal scans.
- Since a variational approach is employed, the method requires a reduced number of input parameters that need to be calibrated. Moreover, the cost function we propose also renders itself compatible with fast numerical schemes, while providing a dense voxel-by-voxel deformation field.
- The benefit of using multi-CPU and GPU (graphics processing unit) architectures is evaluated.

2. Method description

2.1. Proposed EVolution method

The equations provided in the current paper refer to the 3D implementation of the algorithm. An image J is registered to the reference position given by I using a variational image registration method as follows.

2.1.1. Proposed data fidelity term. Let $\vec{\nabla}_I$ and $\vec{\nabla}_J$ be the gradient of the reference image I and the image to register J , respectively. We defined the following patch-based criterion (a patch consists in a cubic subset of the image domain, denoted by Γ , centered on one single voxel):

$$C(T) = \frac{\int_{\Gamma} |\vec{\nabla}_I(T(\vec{r})) \cdot \vec{\nabla}_J(\vec{r})| d\vec{r}}{\int_{\Gamma} \|\vec{\nabla}_I(T(\vec{r}))\|_2 \|\vec{\nabla}_J(\vec{r})\|_2 d\vec{r}} \quad (1)$$

with $T = (u, v, w)$ the spatial transformation from I to J , u , v and w the displacement vector components, and \vec{r} the spatial location.

The expression $C(T)$ can also be rewritten under the following form:

$$C(T) = \frac{\int_{\Gamma} w_T(\vec{r}) |\cos(\Delta\theta_T(\vec{r}))| d\vec{r}}{\int_{\Gamma} w_T(\vec{r}) d\vec{r}} \quad (2)$$

where $w_T(\vec{r})$ and $\Delta\theta_T(\vec{r})$ are calculated from the magnitude M and the orientation θ of the image gradient (computed using a Sobel filter) at location \vec{r} as follows:

$$\begin{aligned} w_T(\vec{r}) &= M_I(T(\vec{r})) M_J(\vec{r}) \\ \Delta\theta_T(\vec{r}) &= \theta_I(T(\vec{r})) - \theta_J(\vec{r}) \end{aligned} \quad (3)$$

Intuitively, the term $|\cos(\Delta\theta_T(\vec{r}))|$ in equation (2) favors the transformations that align the edges, regardless any possible contrast reversals: due to the absolute value of the cosine, both parallel and anti-parallel edges are considered to coincide irrespective of the gradient direction. In addition, when dealing with multi-modal images, some discontinuities may only appear in one of the two modalities, so the weight $w_T(\vec{r})$ favors strong edges that occur in both modalities. The denominator of equation (2) performs a weighted average of the score obtained for each edge.

Since, for image registration using variational methods, a minimization of the functional is mandatory, we defined the following patch-based similarity criterion $D(T)$:

$$D(T) = e^{-C(T)} \quad (4)$$

$D(T)$, which is computed individually at each spatial location \vec{r} , is a strictly positive number, which decreases as long as the alignment of I and J is improved within the local neighborhood Γ . In this manner, $D(T)$ can be employed as a data fidelity term for the proposed variational registration method.

2.1.2. Optimized variational functional. We propose minimizing the energy E given by:

$$E(T) = \int_{\Omega} D(T) + \frac{\alpha}{2} (\|\vec{\nabla} u\|_2^2 + \|\vec{\nabla} v\|_2^2 + \|\vec{\nabla} w\|_2^2) d\vec{r} \quad (5)$$

where Ω is the image coordinates domain, α a weighting factors designed to link both the data fidelity term $D(T)$ and the motion field regularity (right part of equation (5)).

At this point, it is important to underline that two user-defined parameters may impact the performance of the registration process:

- The parameter α which infers the regularity of the estimated motion field.
- The patch size p (noted p) which may infer the robustness against different structural information from the image to register to the reference (for example, on an area that is smooth on one image and textured on the other). Simultaneously, p also infers the regularity of the estimated motion field.

Throughout the rest of the manuscript, a special attention will be paid to the impact of these two parameters on the overall registration results.

2.1.3. Implemented optimization scheme. By applying the Euler–Lagrange equations on a voxel-by-voxel basis, one can derive the following system of equations for each $\vec{r} \in \Omega$:

$$\begin{cases} \frac{\partial D}{\partial u}(T) - \alpha \Delta u = 0 \\ \frac{\partial D}{\partial v}(T) - \alpha \Delta v = 0 \\ \frac{\partial D}{\partial w}(T) - \alpha \Delta w = 0 \end{cases} \quad (6)$$

where Δ denotes the Laplacian operator.

From here, we have a set of $3 \times |\Omega|$ non-linear equations with common unknowns u , v and w . The latter can be found iteratively through the following explicit fixed-point scheme:

$$\begin{cases} u_{k+1} = u_k + \Delta u_k - \alpha^{-1} D(T_k) \frac{\partial D}{\partial u}(T_k) \\ v_{k+1} = v_k + \Delta v_k - \alpha^{-1} D(T_k) \frac{\partial D}{\partial v}(T_k) \\ w_{k+1} = w_k + \Delta w_k - \alpha^{-1} D(T_k) \frac{\partial D}{\partial w}(T_k) \end{cases} \quad (7)$$

where $k + 1$ denotes the new iteration and $T_k = (u_k, v_k, w_k)$. Each operator $\Delta(\cdot)$ was numerically computed using a $3 \times 3 \times 3$ Laplacian kernel (Gonzalez and Woods 2006). For each voxel, the partial derivatives of the data fidelity term along each displacement component were computed using the following finite difference schemes:

$$\begin{cases} \frac{\partial D}{\partial u}(T_k) = \frac{D(T_k \circ S_x^+) - D(T_k \circ S_x^-)}{2} \\ \frac{\partial D}{\partial v}(T_k) = \frac{D(T_k \circ S_y^+) - D(T_k \circ S_y^-)}{2} \\ \frac{\partial D}{\partial w}(T_k) = \frac{D(T_k \circ S_z^+) - D(T_k \circ S_z^-)}{2} \end{cases} \quad (8)$$

where S_x^+ , S_y^+ and S_z^+ denote the forward shifted image transformations by one voxel along x -, y - and z - directions, respectively. S_x^- , S_y^- and S_z^- being the corresponding backward shifted image transformations. For example, $T_k \circ S_x^+$ reads as: ‘a translation of $(x, y, z) \rightarrow (x + 1, y, z)$ followed by the spatial transformation T_k ’.

It was considered that the numerical scheme in equation (7) converged when the average variation of the motion magnitude from one iteration to the next was smaller than 10^{-3} voxels.

2.1.4. Coarse-to-fine scheme. A coarse-to-fine strategy was carried out, which iterated the registration algorithm from a downsampled image (4-, 8- and 16-fold downsampled images were employed for nominal original image dimensions of 128, 256 and 512 voxels, respectively), step by step to the original image resolution. To achieve a global motion regularisation, the estimate arising from lowest resolution levels (noted $T_g = (u_k^g, v_k^g, w_k^g)$) was added in the regularisation term (i.e. in the right part of equation (5)). The system of equation (7) was thus simply rewritten as follows:

$$\begin{cases} u_{k+1} = u_k + \Delta u_k + \Delta u_k^g - \alpha^{-1} D(T_k) \frac{\partial D}{\partial u}(T_k) \\ v_{k+1} = v_k + \Delta v_k + \Delta v_k^g - \alpha^{-1} D(T_k) \frac{\partial D}{\partial v}(T_k) \\ w_{k+1} = w_k + \Delta w_k + \Delta w_k^g - \alpha^{-1} D(T_k) \frac{\partial D}{\partial w}(T_k) \end{cases} \quad (9)$$

2.2. Experimental setup

The proposed EVolution method was evaluated on three data sets:

- (i) one set of thorax inhale and exhale CT-scans, for which gold-standard landmark displacements are available,

- (ii) one set of MR-scans capturing the expansion of the bladder in a healthy volunteer, for which a silver-standard motion is constructed,
- (iii) one set of paired head CT/MR-scans, for which gold-standard spatial transformations are available.

2.2.1. Thorax inhale and exhale CT-scans. We used a set of ten CT-scan pairs, freely provided by the DIR-Lab⁴ at the University of Texas (Castillo *et al* 2009, Castillo *et al* 2010a, 2010b). Each scan pair was acquired on the thorax and upper abdomen of patients treated for esophageal cancer, between inhale and exhale phase of the breathing cycle. For each CT-scan, 300 anatomical landmarks have been annotated by experts (inter-observer errors below 1 mm).

Four reliability tests were conducted in order to evaluate the registration performance under various experimental conditions:

Reliability test #1. Registration was directly applied between each pair of the original CT-scans. Major challenges arise from possible contrast variations between tissue and air induced by lung compression, motion discontinuities at the lung/rib cage interface, as well as large deformations of small features such as lung vessels, airways (Castillo *et al* 2010a).

Reliability test #2. Registration was applied on the original data set, each image being previously filtered using Gaussian filter (kernel $3 \times 3 \times 3$, $\sigma = 0.5$). In this manner, the impact of filtered anatomical structures on the overall registration accuracy is assessed.

Reliability test #3. A prior intensity perturbation was applied between I and J in order to challenge the registration against intensity contrast variation. For this purpose, the image to register J was replaced by its negative contrast in the registration process, the reference image remaining identical to the original one.

Reliability test #4. The robustness against noise was investigated by adding a Gaussian white noise ($\sigma = 10\%$ of the maximum image intensity) on both I and J , prior to the registration process.

2.2.2. Bladder filling MR-scans. A series of 3D MR-images were acquired on a healthy volunteer, capturing the expansion of the urinary bladder over time. In order to sample the shape and size of the bladder from void to partially full in a relatively short amount of time, a volunteer preparation protocol was devised. The volunteer was instructed not to consume any liquids 6h prior to the MR-acquisition session such that, at the time of the study, they would be mildly dehydrated. Before the start of the experiment, the volunteer was allowed to drink a self-chosen quantity of water (~ 350 ml). It was speculated that the mild dehydration would stimulate renal activity, leading to a rapid expansion of the bladder. Approximately every 7 min, a pair of 3D T1-T2-weighted images was acquired on the lower abdomen of the volunteer, over a total duration of ~ 40 min. The MR-acquisition sequences were optimized to mitigate the fact that each anatomical structure in the T1-weighted images has its counterpart in the T2-weighted ones. The T1-weighted sequence employed the following parameters: TE = 2 ms, TR = 4.3 ms, image matrix $192 \times 192 \times 100$, 10° flip angle, with an isotropic voxel size of $2 \times 2 \times 2$ mm³. For the T2-weighted acquisition, the employed parameters were as follows: TE = 130 ms, TR = 1000 ms, reconstructed image size $512 \times 512 \times 133$, 90° flip angle, with a voxel size of $0.5 \times 0.5 \times 1.5$ mm³. Out of the entire data set, two pairs of

⁴ www.dir-lab.com

T1-T2-weighted images (~ 30 min apart) were selected for the role of reference and respectively image to register. A silver-standard motion field was estimated using a common contrast for both the reference and the image to register (the T1-weighted images were employed) using the following mono-modal optical flow (OF) functional:

$$E_{\text{OF}}(T) = \int_{\Omega} |I_x u + I_y v + I_z w + I_t| + \alpha (\|\vec{\nabla} u\|_2^2 + \|\vec{\nabla} v\|_2^2 + \|\vec{\nabla} w\|_2^2) d\vec{r} \quad (10)$$

where $I_{x,y,z,t}$ are the spatio-temporal partial derivatives of the image voxel intensity. A complete description of the associated non-rigid registration framework can be found in Zachiu *et al* (2015). The obtained OF-motion field was taken as a reference for the evaluation of the proposed EVolution method applied on cross-contrast images (the T1- and respectively the T2-weighted images were employed for the reference and the image to register).

2.2.3. Paired head CT/MR-scans. We used a set of two CT/MR-scan pairs, freely provided by the retrospective image registration evaluation project (RIRE)⁵ at the National Institutes of Health, Vanderbilt University (West *et al* 1996, 1997). For this purpose, three 3D images of the head of a common volunteer were employed: a CT-scan (voxel size of $0.65 \times 0.65 \times 4 \text{ mm}^3$, image matrix $512 \times 512 \times 29$) and a pair of T1-T2-weighted MR images (voxel size of $1.25 \times 1.25 \times 4 \text{ mm}^3$, image matrix $256 \times 256 \times 26$). The performance of the proposed EVolution method for CT/T1-MRI and CT/T2-MRI registrations was assessed using goldstandard spatial transforms defined using a prospective, marker-based technique, as described in Fitzpatrick *et al* (1998).

2.2.4. Implementation of the MIND non-rigid registration algorithm. The identical experimental setup was also applied on the MIND non-rigid registration method (Heinrich *et al* 2012). We recall that the latter is a modality invariant local image descriptor based on the notion of self-similarity derived from the NL-means algorithm. A descriptor, based on the local similarities, is associated to each voxel \vec{r} of the image I :

$$\text{MIND}(I, \vec{r}, \gamma) = \frac{1}{n} \exp\left(-\frac{D_p(I, \vec{r}, \vec{r} + \gamma)}{V(I, \vec{r})}\right) \quad (11)$$

where n is a normalization constant, $\gamma \in \Gamma$ defines a search region in which the patches are compared, D_p is the voxel-wise square distance between patches of size p , and $V(I, \vec{r})$ is an estimation of the local variance in order to account for noise perturbations. The MIND descriptor associates a vector of size $|\Gamma|$ to each pixel of the image I . The descriptors of each modality can subsequently be compared in order to provide the following voxel-wise data fidelity metric:

$$S(T) = \frac{1}{|\Gamma|} \sum_{\gamma \in \Gamma} |\text{MIND}(I, \vec{r}, \gamma) - \text{MIND}(J, \vec{r}, \gamma)| \quad (12)$$

This data-fidelity term is then combined to a L^2 regularization term, similarly to equation (5). The implemented non-rigid registration framework was identical to the one described in Heinrich *et al* (2012). However, in order to make a proper comparison with the estimates provided by our proposed EVolution algorithm, we employed the same coarse-to-fine scheme and convergence criterion as described in section 2.1.4. Note that, for the purpose of this study, $\Gamma = [-1; 1]^3 \cap \mathbb{Z}^3$ and $p = 3 \times 3 \times 3$ were found to be a good compromise between algorithm computational performance and quality of the motion estimates (the interested reader is referred to Heinrich *et al* (2012) for a complete analysis of the two parameters Γ and p).

⁵ www.insight-journal.org/rire/

2.2.5. Assessment of the motion estimation process. Concerning the data set of thorax CT-scans, the performance were assessed using the target registration error (TRE) of anatomical landmarks as follows (Maurer *et al* 1997):

$$\text{TRE} = \sqrt{(x_I + u(\bar{r}_I) - x_J)^2 + (y_I + v(\bar{r}_I) - y_J)^2 + (z_I + w(\bar{r}_I) - z_J)^2} \quad (13)$$

where $\bar{r}_I = (x_I, y_I, z_I)$ and $\bar{r}_J = (x_J, y_J, z_J)$ denote the anatomical landmark coordinates in I and J , respectively.

Concerning the data set of bladder MR-scans and the paired head CT/MR-scans, the quality of the estimated motion was assessed using the voxel-wise error in flow endpoint (FEP) computed as follows (Baker *et al* 2011):

$$\text{FEP} = \sqrt{(u - u_{\text{ref}})^2 + (v - v_{\text{ref}})^2 + (w - w_{\text{ref}})^2} \quad (14)$$

($u_{\text{ref}}, v_{\text{ref}}, w_{\text{ref}}$) being the reference OF-motion estimate.

2.3. Hardware and implementation

Our test platform was an Intel 2.5 GHz i7 workstation (8 cores) with 32 GB of RAM. The GPU was a Quadro K2100M card with 2 GB of dynamic random-access memory (NVIDIA, Santa Clara, CA, USA). Both CPU and GPU implementations of the proposed EVolution algorithm were realized and tested. The CPU implementation was performed in C++ and parallelized through multi-threading. The GPU implementation was realized using the compute unified device architecture (CUDA) (NVIDIA 2008).

3. Results

3.1. Thorax inhale and exhale CT-scans

A visualization of a registration result can be seen in figure 1 (empirically established values of $\alpha = 0.5$ and $p = 11$ were employed). In this visualization, the source image is shown in magenta while the reference image is shown in green. Where the images align a gray scale image emerges. In the unregistered case on the top, magenta and green areas can clearly be seen indicating that the morphology is not aligned. In the registered case to the bottom, these colored areas have almost disappeared indicating that the images have been successfully registered.

Figure 2 shows cumulative distributions of the target registration error (TRE) over 3000 landmarks (300 landmarks for 10 CT-scans) obtained before and after registration, during the ‘reliability test #2’ (common default values of $\alpha = 0.5$ and $p = 11$ were again employed for the EVolution method, and $\alpha = 0.1$ was used for the MIND algorithm, as advised by Heinrich *et al* (2012)). Staircasing effects are observable on the curve obtained before registration, due to the voxel precision of the expert landmark positioning. However this was not the case for the curves associated to the MIND and the proposed EVolution methods, due to the real number precision of the motion estimates. Only a marginal difference can be observed between the MIND and the proposed EVolution methods for all tested experimental conditions. This tendency is confirmed for all tested experimental conditions, as shown in table 1. For both algorithms, no negative impact was observed when the image to register was replaced by its corresponding negative contrast image (‘reliability test #3’). It is also noticeable that both algorithms performed better on pre-filtered images (‘reliability test #4’). For comparison purposes, table 2 provides several findings reported in Heinrich *et al* (2012) using various

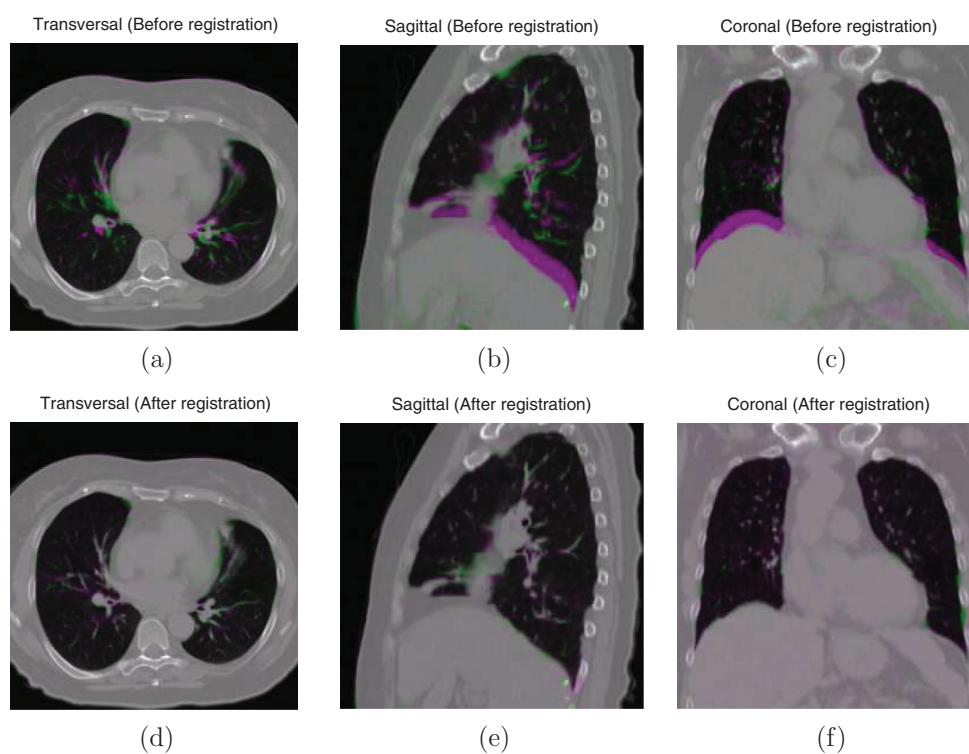


Figure 1. Example of the registration of a thorax CT-scan (CT-scan #5) using the proposed EVOlution method. Top: before registration and bottom: after registration using the proposed EVOlution method. The target image is displayed in green and the source image in magenta (complementary colour).

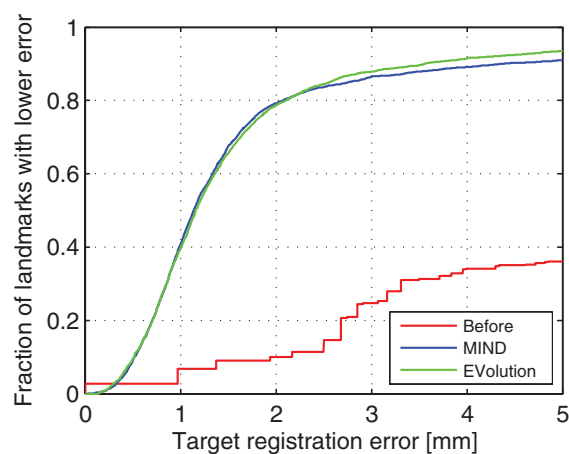


Figure 2. Registration of 10 cases of CT-scans. The plot shows the cumulative distribution of target registration error.

Table 1. Target registration error (in millimeters) obtained over the 10 cases of thorax CT-scans for all tested experimental conditions (reliability tests #1–4).

| Reliability test | Target registration error (mm) | |
|---|---------------------------------------|---------------------------------------|
| | MIND | EVolution |
| Original data (Reliability test #1) | 2.33 ± 3.54 (0.81, 1.18, 1.91) | 2.07 ± 3.23 (0.78, 1.17, 1.85) |
| Filtered data (Reliability test #2) | 2.18 ± 3.45 (0.78, 1.14, 1.77) | 1.96 ± 2.99 (0.78, 1.16, 1.80) |
| Cross-contrast (Reliability test #3) | 2.33 ± 3.55 (0.81, 1.18, 1.91) | 2.08 ± 3.24 (0.78, 1.18, 1.85) |
| Noisy data (Reliability test #4) | 2.33 ± 3.55 (0.81, 1.18, 1.91) | 2.01 ± 3.14 (0.77, 1.16, 1.82) |

Note: The first line of each cell reports the mean \pm standard deviation of the TRE, and the second line provides the first, second and third quantiles (0.25, 0.5, 0.75).

Table 2. Summary of several findings reported in the literature using the employed set of 10 CT-scans.

| Registration method | Target registration error (mm) |
|---|--|
| Before | 8.46 ± 6.58 (3.11, 6.97, 12.55) |
| Sum of squared differences (Pluim <i>et al</i> 2003) | 2.73 ± 3.72 (0.89, 1.44, 2.85) |
| SSD of entropy images (Pluim <i>et al</i> 2003) | 2.86 ± 4.91 (0.86, 1.33, 2.33) |
| Normalised mutual information (Pluim <i>et al</i> 2003) | 2.97 ± 4.22 (0.91, 1.42, 2.67) |
| Conditional mutual information (Loeckx <i>et al</i> 2010) | 3.06 ± 4.10 (1.00, 1.59, 2.85) |
| MIND (Heinrich <i>et al</i> 2012) | 2.14 ± 3.71 (0.77, 1.16, 1.79) |

existing registration methods and the identical set of CT-scans (the interested reader is also referred to Heinrich *et al* (2012) for a complete description of the associated non-rigid registration framework employed to generate the results). It can be observed that the TRE obtained using our implementation of the MIND algorithm are on average almost identical to the ones reported in the original paper (2.18 mm for our implementation against 2.14 mm for the original). The standard deviation of the TRE was even marginally improved using our implementation (3.45 mm against 3.71 mm). We believe that this is due to the fact that our convergence criterion, which had to be similar to the one employed for the proposed EVolution method (see section 2.1.3), is more severe in our implementation.

Figure 3 analyzes the sensitivity to parameter α of the MIND algorithm and respectively the proposed EVolution method. It is important to indicate that, for $\alpha < 0.1$, high instabilities of the numerical schemes were observed for both the MIND and the EVolution approaches. It can be observed that the performance of the MIND algorithm deteriorates as α increases. $\alpha = 0.1$ was thus found the optimal calibration for the MIND algorithm, which matches the

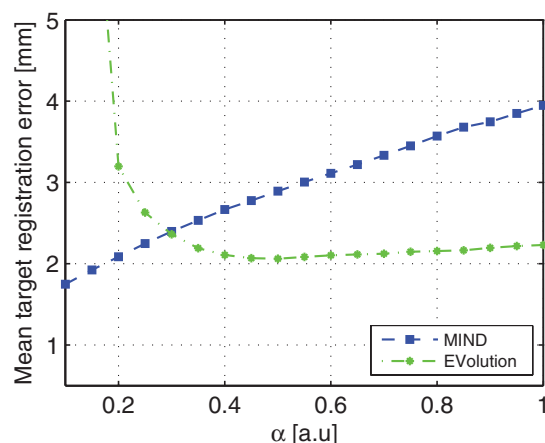


Figure 3. Example (CT-scan #5) of the sensitivity to the α parameter of the MIND and the EVolution algorithms. We recall that the size of the patch for the EVolution method was fixed to cover a dimension of $11 \times 11 \times 11$ voxels.

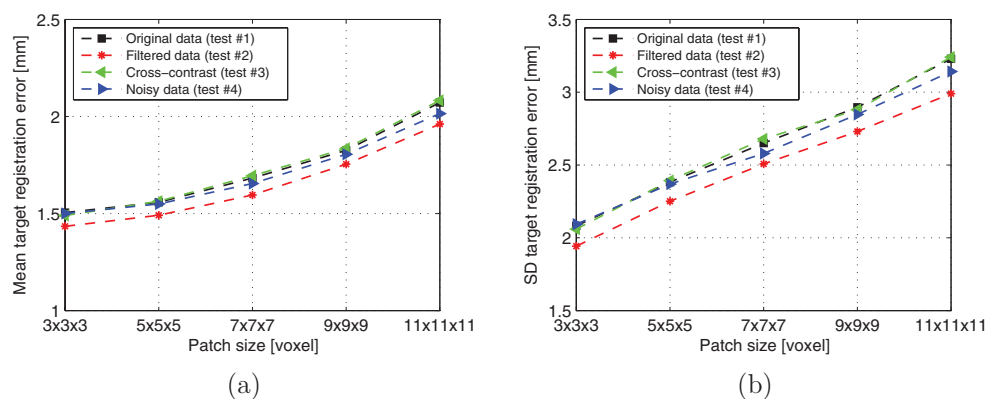


Figure 4. Analysis of the impact of the patch size on the performance of the EVolution method. We recall that the regularization parameter α was here fixed to a value of 0.5.

findings reported in Heinrich *et al* (2012). $\alpha = 0.1$ was thus employed as a default value throughout the manuscript for the MIND algorithm. Using the EVolution method, the TRE, as a function of α , exhibits a flat zone: a range of values for α allowing an accurate registration could thus be determined. This justifies our choice to set a value of 0.5 as a default parameter for α for the proposed EVolution algorithm in the scope of this study.

Figure 4 analyzes the impact of the patch size on the performance of the proposed approach. The TRE was calculated over the 10 cases of the CT-scan data set (300 expert landmarks per case). The mean and standard deviation of the TRE can thus be directly compared to the values reported in tables 1 and 2. It can be observed that the performance of the EVolution method could be further improved for all tested experiment conditions by simply reducing the patch size. It is also important to indicate that a patch restrained to the dimension of one voxel resulted in high instabilities in the employed numerical scheme.

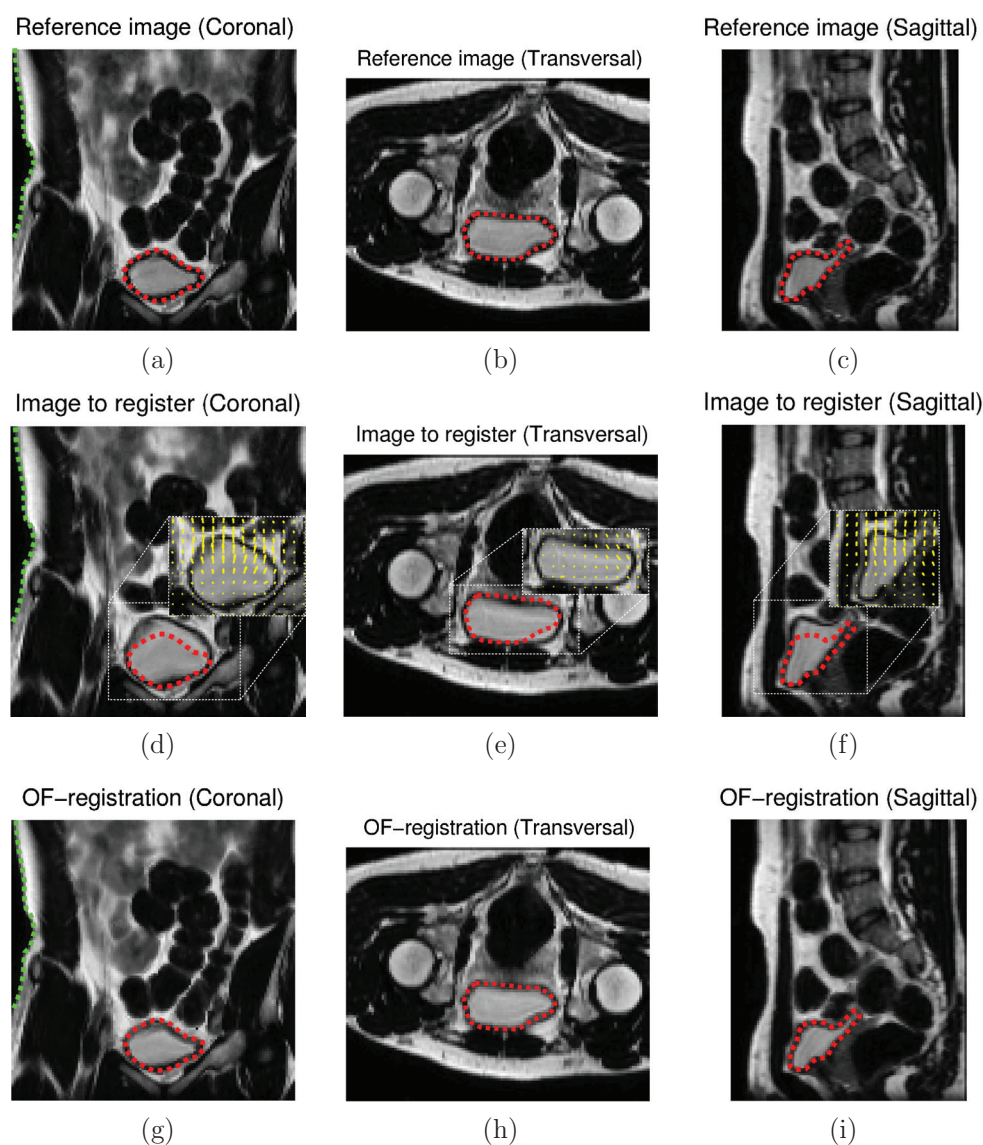


Figure 5. Construction of a sliver-standard motion for the cross-contrast MR-data set: an OF-based algorithm was applied on the T1-weighted data for both the reference and the image to register. Manually defined ROI encompassing the bladder and the thorax are illustrated using red and green dashed lines in order to improve the visual inspection of the quality of the OF registration. The inserts in (d)–(f) depict the estimated OF-motion field vectors within the bladder.

3.2. Bladder filling MR-scans

Figure 5 illustrates the first step of the experimental protocol designed to assess the proposed EVOlution algorithm on the bladder filling MR-data set: a sliver-standard motion is constructed using an optical flow (OF) algorithm (equation (10)) applied on the T1-weighted data for both the reference and the image to register. The inserts in figures 5(d)–(f) show a

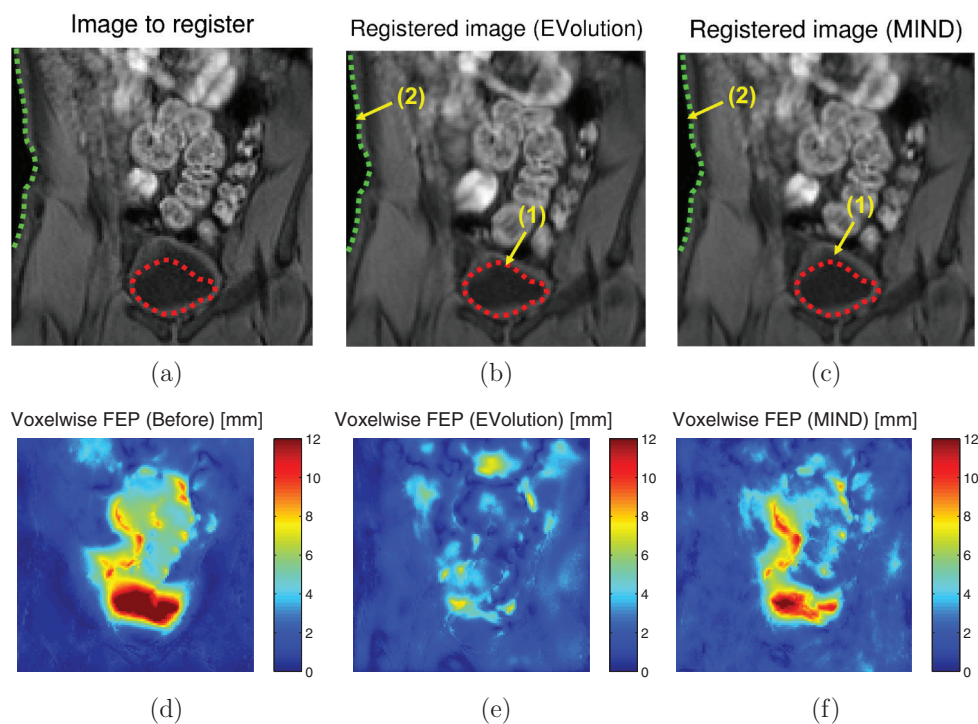


Figure 6. Cross-contrast registration of the T2-weighted bladder image (the reference image being the T1-weighted image displayed in figures 5(a)–(c)). The coronal image is displayed before registration (left column), using the EVolution method (middle column, $\alpha = 0.5$, patch size = $11 \times 11 \times 11$ voxels) and using the MIND algorithm (right column, $\alpha = 0.1$). (a)–(c) Display the anatomical images. Manually defined ROI encompassing the bladder (arrow (1)) and the thorax outline (arrow (2)) are shown using red and green dashed lines in order to improve the visual inspection of the quality of the cross-contrast registration. (d)–(f) Display the corresponding voxel-wise FEP maps.

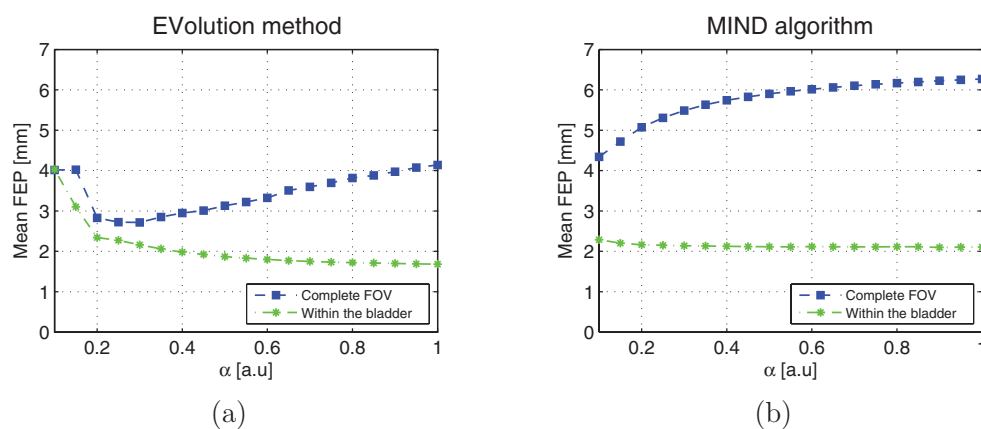


Figure 7. Impact of the α parameter on the performance of the proposed EVolution (a) and the MIND (b) algorithms for the bladder filling MR-scans.

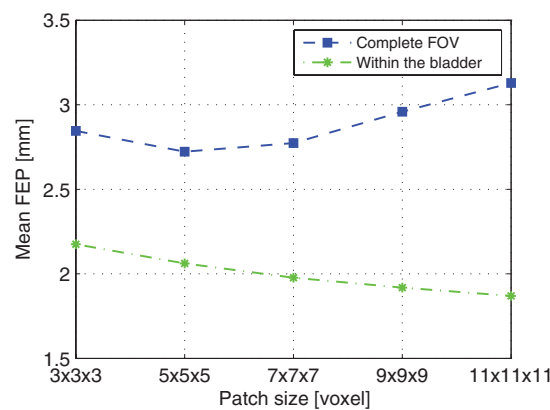


Figure 8. Impact of the patch size on the performance of the proposed EVolution method for the bladder filling MR-scans.

spatially regular motion field. Moreover, a good correspondence has been found between the bladder on the registered volumetric image of figures 5(g)–(i) and the region of interest manually defined on the reference data.

A visualization of cross-contrast registration results can be seen in figure 6. Default calibration parameters were employed for both the MIND and the proposed EVolution methods.

This visualization shows a close correspondence of the contour of the co-registered bladder with the manually defined bladder boundary when the proposed algorithm is employed (see figures 6(a)–(c)). The averaged FEP in the bladder decreased from 6.33 mm before registration to 3.12 mm and 4.34 mm using the proposed EVolution and the MIND algorithm, respectively (see figures 6(d)–(f)).

Figure 7 shows the impact of the parameter α on the registration within the bladder (blue curve) as well as the complete image field-of-view (green curve), using the EVolution method (figure 7(a)) and the MIND algorithm (figure 7(b)). It can be observed that the behavior of both algorithms in the bladder was similar than the one reported in figure 3. Using the MIND algorithm, the default value of $\alpha = 0.1$ optimized locally the results in the bladder but also provided the worst global performance within the complete image field-of-view. Moreover, while high α values improved the accuracy on the complete image field-of-view (wherein the high motion smoothness then coped nicely with the small global motion of the abdomen), this also had an opposite impact in the bladder (wherein the mean FRE necessarily converged toward the one obtained before registration). On the other hand, using the proposed EVolution method, a range of values could be found which optimized results in both the complete field-of-view as well as locally in the bladder (see arrows (1) and (2) in figure 6(b) and corresponding locations in the error map of figure 6(e)).

Figure 8 shows the impact of the patch size on the performance of the proposed EVolution method. The behavior of these two curves match closely the behavior of the green and the blue curves displayed in figure 7(a). It is important to mention that a patch restrained to the dimension of one voxel resulted in high instabilities in the employed numerical scheme.

3.2.1. Paired head CT/MR-scans. A visualization of a CT/MRI registration result can be seen in figure 9 (default values of $\alpha = 0.5$ and $p = 11$ were used). The employed visualization is identical to that described in section 3.1. In the registered case shown on the right, it can be observed that the skull, which features a hyper-intense signal in the CT-scan (magenta), matches closely the spacial location of the scalp and the brain, the latter being visible only in

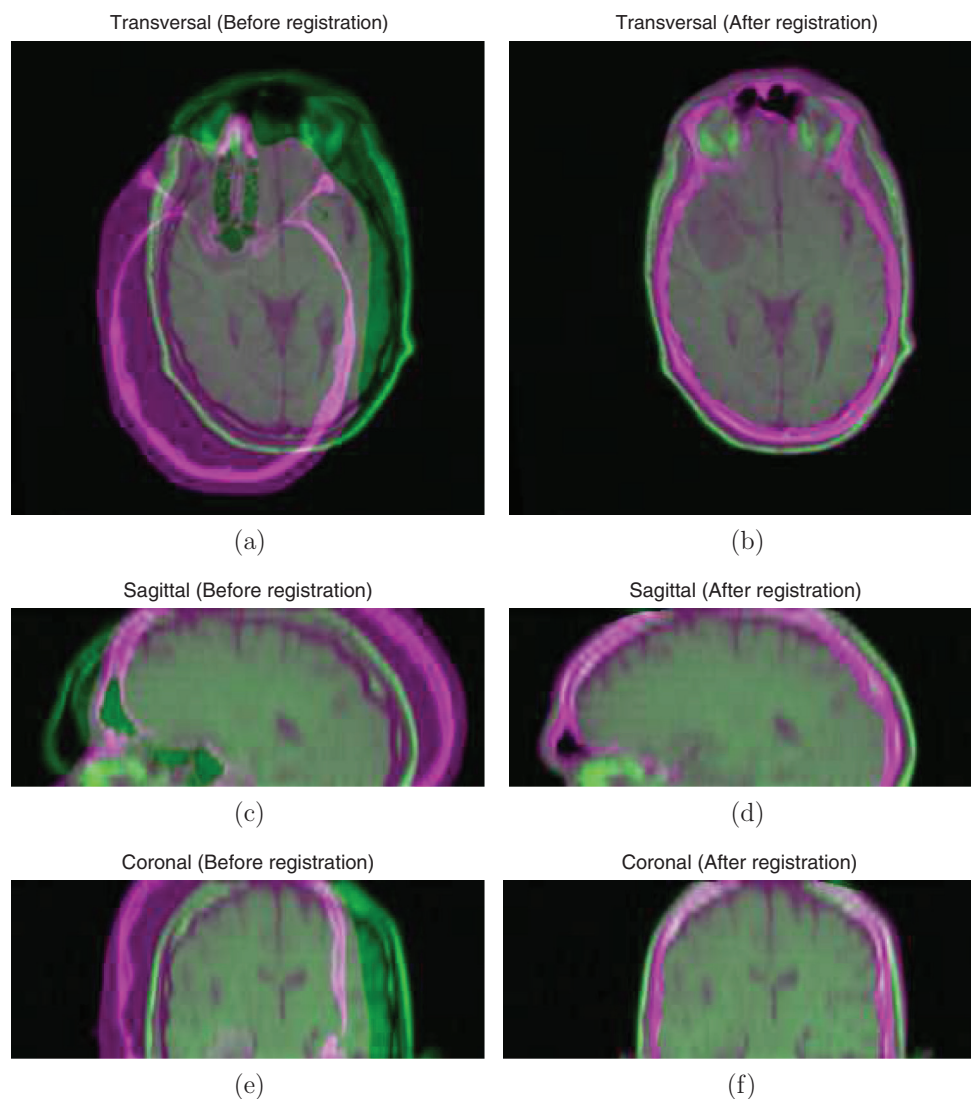


Figure 9. Example of CT/T1-MRI registration of the brain using the proposed EVOlution method. Left: before registration and right: after registration using the proposed EVOlution method. The CT-image is displayed in magenta and the T1-weighted MRI in green (complementary colour).

the T1-weighted image (green). This indicates that the images have been successfully registered. This visual inspection is confirmed by the table 3, which analyses of the voxel-wise error in the flow endpoint for the two CT/MRI paired scans. For each CT/T1 and CT/T2 scenario, an average FEP below the voxel-size was obtained, indicating a sub-voxel registration accuracy.

3.2.2. Benchmark. A benchmark of our CPU and GPU implementations of the proposed EVOlution method for different patch size is provided in table 4. As compared to the CPU implementation, a great acceleration could be obtained using the GPU, especially for high patch size. It is interesting to note that less than a minute was necessary to register two volumes of $128 \times 128 \times 128$ voxels (i.e. the Bladder MR-scan).

Table 3. Motion estimation error (in millimeters) obtained for the registration of the two paired head CT/MR-scans (i.e. CT/T1-MRI and CT/T2-MRI).

| Registered image modalities (Modality #1/Modality #2) | Motion estimation error (mm) | |
|--|---|---------------------------------------|
| | Before registration | EVolution |
| CT-scan/T1-weighted MRI | 41.24 ± 2.66 (39.27, 41.27, 43.27) | 3.30 ± 1.84 (1.93, 2.98, 4.24) |
| CT-scan/T2-weighted MRI | 49.32 ± 2.54 (47.34, 49.25, 51.24) | 3.87 ± 2.24 (2.04, 3.42, 5.36) |

Note: The first line of each cell reports the mean \pm standard deviation of the FEP, and the second line provides the first, second and third quantiles ([0.25, 0.5, 0.75]).

Table 4. Computation time in seconds obtained using our implementation of the proposed EVolution method on our test platform (multi-CPU and GPU implementations) for different patch size (see section 2.3).

| Patch size (voxels) | Thorax CT-scan #5 ($256 \times 256 \times 106$) | | Bladder MR-scan ($128 \times 128 \times 128$) | |
|--------------------------|--|-----|--|-----|
| | CPU | GPU | CPU | GPU |
| $3 \times 3 \times 3$ | 106 | 68 | 34 | 19 |
| $5 \times 5 \times 5$ | 195 | 98 | 56 | 22 |
| $7 \times 7 \times 7$ | 331 | 134 | 114 | 29 |
| $9 \times 9 \times 9$ | 693 | 219 | 179 | 41 |
| $11 \times 11 \times 11$ | 1094 | 320 | 300 | 59 |

4. Discussion

Using variational image-based registration methods, the negative impact of appearing/disappearing anatomical structures (for example induced by specific anatomical structures featured by the involved imaging modalities, peristaltic activity or any potential image artifacts) can typically be circumvented by increasing the weight of a spatial motion regularization term (see the improvement of the FEP with increasing α values on the green curves of figures 7(a) and (b)). This, however, may have a negative impact on locally moving structures (see the deterioration of the FEP with increasing α values on the blue curve of figures 7(a) and (b)). Using our method, the robustness against structure appearance/disappearance is, by construction, not only taken into account by the motion regularization term, but also in the proposed data fidelity term. The reason is that our criterion converts the image into an ‘edge representation’, and only strong edges that occur in both modalities are favored, due to the use of the weight $w_T(\vec{r})$ in equation (2).

It can be observed in figure 9 that the skull is properly registered by the EVolution algorithm, although it features hyper- and hypo-intense signals in CT- and MR- images, respectively. This arises from the fact that the absolute value of the cosine in equation (2) favors the alignment of both parallel and anti-parallel edges, irrespective of the gradient direction. As a consequence, in this particular example, the strong contrast of the cortical bone in the CT-image and the hyperintense signal of the subcutaneous fat layers in the MR-images dominate the overall convergence process of the variational algorithm.

In order to mitigate the fact that the image features under the new representation might not be discriminative enough (especially in the calculation of the partial derivatives of equation (8)), the defined normalized criterion performs a weighted average of the score obtained for each edge that occur in both modalities within patches. The patch size is thus a free parameter which can be increased to improve the robustness against anatomical structure without counterpart between the image to register and the reference one. However, it must be also underlined that, similarly to the parameter α , the patch size also impacts the spatial regularity of the estimated motion field. Decreasing these two parameters generally results in a poor estimation of the displacement, due to numerical instabilities. Identically, for values increasing toward infinity, the tested reliability criteria indicated a poor estimation of the displacement, as the smoothness of motion constrains the velocity amplitude estimation. As a consequence, a good compromise of the choice of both parameters is essential for a reliable and accurate co-registration. Using the proposed default user-defined parameters (i.e. $\alpha = 0.5$ and a patch size of $11 \times 11 \times 11$ voxels), results obtained using the proposed method were comparable to those obtained using the MIND algorithm, and this for all tested experimental conditions (image pre-filtering, image intensity variation, noise perturbation). It is interesting to note in figure 4 that the proposed EVolution method could systematically be further improved, however for this specific 10 CT-scans data set, by simply decreasing the patch size. Using default calibration parameters, the proposed method outperforms the existing MIND algorithm on the bladder filling scenario, for which structures in the image to register haven't counterparts in the reference image.

It can be noticed that the addition of noise perturbation somewhat improved registration results, especially for high patch size (see the blue curve in figures 4(a) and (b)). The presence of noise indeed alleviated possible divisions by values close to 0 in equation (1) which, in turn, slightly stabilized the numerical process. A typical alternative method to get around this would be to add a very small positive constant ϵ in the denominator of equation (1) ($\epsilon = 0.1$ provided results similar to the ones reported by the blue curves of figure 4).

5. Conclusion

An image registration tool box including complementary multi-modal algorithms is a necessary prerequisite in the field of medical imaging, with respect to the ever growing need to align images of a same scene acquired at different time instants, using different sensors and/or from a different point-of-view.

In the current work we proposed a non-rigid multi-modal registration method that aligns edges in both modalities using a variational scheme that provides a fast algorithm with a low number of parameters to tune. Using the proposed default user-defined parameters (i.e. $\alpha = 0.5$ and a patch size of $11 \times 11 \times 11$ voxels), results obtained using the proposed EVolution method were comparable to those obtained using the MIND algorithm, and this for all tested experimental conditions (image pre-filtering, image intensity variation, noise perturbation). This demonstrates that both methods performed with comparable accuracy and precision for the general case. Using identical calibration parameters, it was also demonstrated that the proposed method outperforms existing approaches on a specific scenario for which structures in the image to register do not have counterparts in the reference image.

It must be underlined that, in their excursion due to respiration, the lung, liver and kidneys slide on the thoracic and abdominal walls which, by comparison, manifest a reduced amount of motion. The impact of the arising shearing effects on the motion regularization term may not be negligible, and will need to be addressed in future works. Moreover, future studies will

challenge the proposed method for various non-rigid multi-sensor registration scenarios (e.g. MRI/echography, PET/MRI, etc...).

Acknowledgment

The authors thank the ITEA (project 12026, SoRTS) and the European Research Council (project ERC-2010-AdG-20100317, Sound Pharma) for funding.

References

- Baker S, Scharstein D, Lewis J P, Roth S, Black M J and Szeliski R 2011 A database and evaluation methodology for optical flow *Int. J. Comput. Vis.* **92** 1–31
- Buades A, Coll B and Morel J M 2005 A review of image denoising algorithms, with a new one *J. Multiscale Model. Simul.* **4** 490–530
- Castillo E, Castillo R, Guerra R, Johnson V, McPhail T, Garg A and Guerrero T 2009 A framework for evaluation of deformable image registration spatial accuracy using large landmark point sets *Phys. Med. Biol.* **54** 1849–70
- Castillo E, Castillo R, Martinez J and Guerrero T 2010b Ventilation from four-dimensional computed tomography: density versus jacobian methods *Phys. Med. Biol.* **55** 4661–85
- Castillo E, Castillo R, Martinez J, Shenoy M and Guerrero T 2010a Four-dimensional deformable image registration using trajectory modeling *Phys. Med. Biol.* **55** 305–27
- Denis de Senneville B, El Hamidi A and Moonen C T W 2015 A direct PCA-based approach for real-time description of physiological organ deformations *IEEE Trans. Med. Imaging* **34** 974–82
- Denis de Senneville B, Ries M, Maclair G and Moonen C 2011 MR-guided thermotherapy of abdominal organs using a robust PCA-based motion descriptor *IEEE Trans. Med. Imaging* **30** 1987–95
- Fitzpatrick J M, West J B and Maurer C R 1998 Predicting error in rigid-body, point-based registration *IEEE Trans. Med. Imaging* **17** 694–702
- Gonzalez R and Woods R 2006 *Digital Image Processing* 3rd edn (Englewood Cliffs, NJ: Prentice-Hall)
- Haber E and Modersitzki J 2006 Intensity gradient based registration and fusion of multi-modal images *MICCAI* pp 726–33
- Heinrich M, Jenkinson M, Bhushan M, Matin T, Gleeson F, Brady S and Schnabel J 2012 MIND: modality independent neighbourhood descriptor for multi-modal deformable registration *Med. Image Anal.* **16** 1423–35
- Hill D, Batchelor P, Holden M and Hawkes D 2001 Medical image registration *Phys. Med. Biol.* **46** R1–45
- Horn B and Schunck B 1981 Determining optical flow *Artif. Intell.* **17** 185–203
- Loeckx D, Slagmolen P, Maes F, Vandermeulen D and Suetens P 2010 Nonrigid image registration using conditional mutual information *IEEE Trans. Med. Imaging* **29** 19–29
- Maes F, Collignon A, Vandermeulen D, Marchal G and Suetens P 1997 Multimodality image registration by maximization of mutual information *IEEE Trans. Med. Imaging* **16** 187–98
- Maintz J and Viergever M 1998 A survey of medical image registration *Med. Image Anal.* **2** 1–36
- Mani V and Arivazhagan S 2013 Survey of medical image registration *J. Biomed. Eng. Technol.* **1** 8–25
- Maurer C, Fitzpatrick J, Wang M, Galloway R, Maciunas R and Allen G 1997 Registration of head volume images using implantable fiducial markers *IEEE Trans. Med. Imaging* **16** 447–62
- Mellor M and Brady M 2005 Phase mutual information as a similarity measure for registration *Med. Image Anal.* **9** 330–43
- NVIDIA 2008 *NVIDIA CUDA: Compute Unified Device Architecture Programming Guide* 2.0 edn (Santa Clara: NVIDIA Corporation)
- Pluim J, Maintz J and Viergever M 2000 Image registration by maximization of combined mutual information and gradient information *IEEE Trans. Med. Imaging* **19** 809–14
- Pluim J, Maintz J and Viergever M 2003 Mutual-information-based registration of medical images: a survey *IEEE Trans. Med. Imaging* **22** 986–1004
- Ries M, Denis de Senneville B, Roujol S, Berber Y, Quesson B and Moonen C 2010 Real-time 3D target tracking in MRI guided focused ultrasound ablations in moving tissues *Magn. Reson. Med.* **64** 1704–12

- Sun Y, Jolly M and Moura J 2004 Integrated registration of dynamic renal perfusion MR images *ICIP* (IEEE) pp 1923–6
- Sutour C, Aujol J-F, Deledalle C A and Denis de Senneville B 2015 Edge-based multi-modal registration and application for night vision devices *J. Math. Imaging Vis.* **53** 131–50
- Viola P and Wells W III 1997 Alignment by maximization of mutual information *Int. J. Comput. Vis.* **24** 137–54
- Wachinger C and Navab N 2012 Entropy and laplacian images: structural representations for multi-modal registration *Med. Image Anal.* **16** 1–17
- Weickert J, Bruhn A, Papenberg N and Brox T 2003 Variational optic flow computation: from continuous models to algorithms *Int. Workshop on Computer Vision and Image Analysis (Las Palmas de Gran Canaria)* ed L Alvarez
- West J *et al* 1996 Comparison and evaluation of retrospective intermodality image registration technique *SPIE Medim'96, SPIE 2710* pp 332–47
- West J *et al* 1997 Comparison and evaluation of retrospective intermodality image brain image registration techniques *J. Comput. Assist. Tomogr.* **21** 554–66
- Zachiu C, Papadakis N, Ries M, Moonen C T W and Denis de Senneville B 2015 An improved optical flow tracking technique for real-time MR-guided beam therapies in moving organs *Phys. Med. Biol.* **60** 9003