Registration and Modeling from Spaced and Misaligned Image Volumes

Adeline Paiement, Majid Mirmehdi, *Senior Member, IEEE,* Xianghua Xie, *Senior Member, IEEE,* and Mark C. K. Hamilton

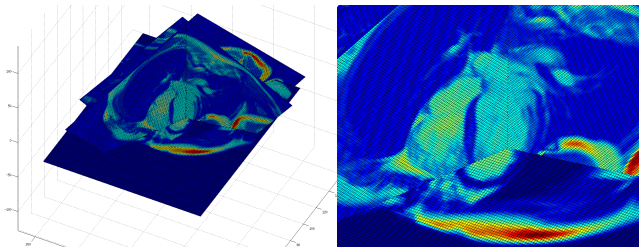
Abstract—We address the problem of object modeling from 3D and 3D+T data made up of images which contain different parts of an object of interest, are separated by large spaces, and are misaligned with respect to each other. These images have only a limited number of intersections, hence making their registration particularly challenging. Furthermore, such data may result from various medical imaging modalities and can therefore present very diverse spatial configurations. Previous methods perform registration and object modeling (segmentation and interpolation) sequentially. However, sequential registration is ill-suited for the case of images with few intersections. We propose a new methodology which, regardless of the spatial configuration of the data, performs the three stages of registration, segmentation, and shape interpolation from spaced and misaligned images simultaneously. We integrate these three processes in a level set framework, in order to benefit from their synergistic interactions. We also propose a new registration method that exploits segmentation information rather than pixel intensities, and that accounts for the global shape of the object of interest, for increased robustness and accuracy. The accuracy of registration is compared against traditional mutual information based methods, and the total modeling framework is assessed against traditional sequential processing and validated on artificial, CT, and MRI data.

Index Terms—Modeling methodologies, registration, segmentation, shape interpolation, level set methods.

1 Introduction

We present an approach for object modeling from single patient's 3D or 3D+T medical volume made up of several misaligned sequences of spaced 2D images. Sparse datasets are commonly produced to decrease the acquisition time, for example when imaging infants or requiring breath holds. Misalignments may occur between sequences produced at the same visit due to patient movements between their independent acquisitions. If, in addition, the subjects are required to repeatedly hold their breath, as is usual in cardiac MRI (a major application of our method), different amounts of air may remain in the lungs, inducing different positions of the heart in the thorax[1]. The combined effects of patient movement and inconsistent breath holds lead to rigid misalignments between the image sequences, as illustrated in Fig. 1. Object modeling from such data is typically used to e.g. establish a diagnosis or plan surgery, and it requires solving the three closely intertwined tasks of registration, segmentation, and interpolation.

We identify three cases of sequences which may be misaligned due to patient movement between successive acquisitions. Stacks of slices (where "slice" denotes the location of a 2D image in a 3D volume) may be acquired with different scanner parameters in order to highlight different properties of the object, as illustrated in Fig. 2b. They may also be acquired with different slice orientations in order to obtain satisfactory views of different aspects of the object, as in Fig. 2c. In these two cases, each of the many sequences that make up the data comprises a set of slices, acquired at the same time, and spans the whole volume. We refer to such sequences as "spatial sequences". "Temporal sequences" may also be encountered, where one sequence is defined as all the time-frames of a single slice, as in Fig. 2a. In the case of cardiac MRIs, the acquisitions of the different temporal sequences (i.e. different slice positions) are synchronized by an electrocardiogram. Therefore, the sequences can be assumed to be temporally aligned with each other, i.e. frame t in all sequences represents the same phase t in the cardiac cycle. In addition, all the images of a sequence, regardless of the type of sequence, are acquired in a single measure (and breath hold if needed), and the patient should be immobile during this short acquisition time. Thus, we can reasonably assume that the images of



(a) (b)

Fig. 1. Example of misaligned dataset – one short-axis (horizontal) and two long-axis (vertical) intersecting slices in a cardiac MRI. (a) global view of the three slices, and (b) zoom on the heart region. Only one timeframe is shown. Color denotes pixel intensity. Note the misalignment between the horizontal slice and the other two vertical slices, which is due to a shift of the heart between their respective acquisitions.

A. Paiement and M. Mirmehdi are with the Department of Computer Science, University of Bristol, Bristol, BS8 1UB, UK e-mail: csatmp@bristol.ac.uk, majid@cs.bris.ac.uk.
X. Xie is with the Department of Computer Science, University of Swansea, Swansea, SA2 8PP, UK e-mail: x.xie@swansea.ac.uk.
M.C.K. Hamilton is with the Department of Radiology, Bristol Royal Infirmary, Bristol, BS2 8HW, UK e-mail: mark.hamilton@uhbristol.nhs.uk.
Manuscript received xx, 2015; revised xx, 201x.

1. No significant deformations are induced in the case of breath holds at end-expiration, that is very common in clinical practices.

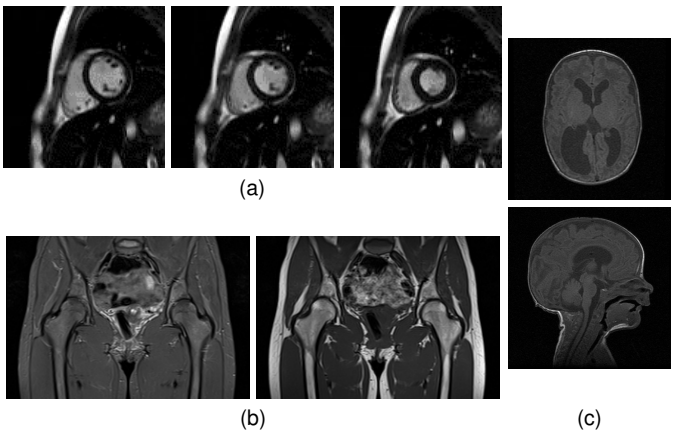


Fig. 2. Illustration of the three types of sequences that may make up a dataset. Temporal sequence: (a) three time-frames of a slice in a cine cardiac MRI. Spatial sequences: (b) central slices of two sequences acquired with different scanner parameters in an MRI of the pelvis, (c) central slices of two sequences acquired with different slice orientations in a brain MRI – top: axial, bottom: sagittal.

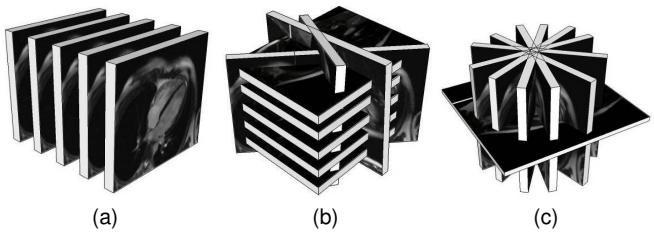


Fig. 3. Examples of sparse datasets having diverse spatial configurations – slice thickness is imposed by scanner's limitations, and spacing is adjusted as a compromise between accuracy and patient comfort. (a) simple sequence of parallel slices, (b) standard spatial configuration of a cardiac MRI: several parallel short-axis slices that span the volume plus a few long-axis slices, and (c) radial dataset. *Figure reproduced from [1].*

the same sequence are spatially aligned. However, spatial misalignments may still exist between these independently acquired sequences.

Therefore, the data needs to be registered prior to any segmentation or interpolation. This registration is challenging, because very diverse numbers, positions, and orientations of the slices may be encountered in order to account for the variety of imaged objects and of patient morphologies – see examples in Fig. 3 – and the registration method must be able to deal with arbitrary spatial configurations. In addition, the images of such datasets contain different parts of the object of interest and may have only a very limited number of intersections when their sequences have different orientations. In such cases, intensity information at slice intersections does not capture the geometry of the object, and it is therefore not reliable for driving registration, as will be demonstrated in Section 4. We argue that 3D segmentation information would provide the shape information that is needed for a robust registration. So, while 3D segmentation requires registered images, registration may also benefit from 3D segmentation.

The segmentation process itself is difficult, as it seeks to recover the 3D or 4D shape of the object from spaced 2D images. Independent segmentations of the images are not suitable since these segmentations may not be consistent with each other. Thus the 2D images are better segmented simultaneously in 3D or 4D. However, they commonly have relatively large spacings in order to shorten the acquisition and reduce patient discomfort, resulting in incomplete datasets as shown in Fig. 3. Such missing data do not offer support for 3D or 4D segmentation, and thus interpolation is necessary, either on the 2D images or on the shape of the object being segmented, after corrections of any misalignment. We showed in [1], [2] that combining 2D segmentation and shape interpolation produces more robust 3D segmentations than the sequential interpolation and segmentation approach. Clearly, the success of each of these three stages (registration, segmentation, and interpolation) depends on the robustness of the others.

We propose to model objects from spaced and misaligned volumes by combining registration, segmentation, and shape interpolation in a level set framework. The novel contributions of this work are: 1) a *new registration method* for internal alignment of volumes made up of spaced 2D slices having arbitrary relative positions and orientations and containing different parts of the object of interest. This method exploits segmentation results rather than pixel intensity, to better deal with largely spaced images that have too few intersections between them to allow computing reliable similarity measures. It also accounts for the global geometry of the object of interest, hence being more robust to large misalignments and local minima. 2) the *full integration of the three stages of registration, segmentation, and shape interpolation* into a novel and general level set framework that allows handling spaced and misaligned data with an increased robustness for all three processes. To the best of our knowledge, this is the first time that the three processes are integrated in a level set framework.

Next in Section 2, we review previous works on the analysis of spaced and misaligned image data, with particular reference to integrated methods. Then we introduce our method in Section 3. Section 4 presents validation tests and comparative results, and Section 6 concludes the paper.

2 Previous Works

Sequential approach – Traditionally, registration, segmentation, and interpolation, are performed sequentially. Usually, the different sequences of a dataset are registered first. Notably, registration based on similarity measures such as the Mutual Information (MI) and Normalized Mutual Information (NMI) [3] is very popular. Lotjonen *et al.* [4] used NMI to register by translation two orthogonal sets of largely spaced temporal sequences (i.e. 2D slices) in cardiac MRIs. This method performs in two iterative steps: first, one slice is chosen randomly, and second, it is shifted in the direction which increases its similarity measure with all of its intersecting slices. It yields satisfactory results provided that the number of intersections between the different slices is high enough to compute meaningful similarity measures. Similarly, Slomka *et al.* [5] and Elen *et al.* [6] registered, by in-plane translations and out-of-plane rotations, and by 3D translations, respectively, cardiac MRIs that comprise two orthogonal long-axis (LA) slices and a stack of short-axis (SA) slices, by optimizing the match of intensity profiles

along the intersections of the various slices. Kim *et al.* [7], [8] also optimized intensity profile matchings to rigidly align fetal brain MRI slices that form three orthogonal stacks. Woo *et al.* [9] performed MI based affine registration followed by deformation on three volumes formed by orthogonal stacks of slices. Li and Denney [10] registered intersecting and largely spaced 2D slices using the segmentation contours obtained from their preliminary independent segmentations, with the aforementioned drawback that the independent automatic 2D segmentations may not be consistent with each other.

Following registration, modeling an object from sets of spaced images requires interpolating the gaps in the data. Li *et al.* [11] and Pihuit *et al.* [12], amongst others, proposed to analyze sets of parallel slices by first segmenting the slices independently, and then interpolating a surface from the 2D contours. More commonly, the image slices are first interpolated in order to reconstruct a 3D volume, e.g. as in [7], [8], [9], [13], [14]. Kim *et al.* [7] addressed the problem of super-resolution from anisotropic voxel data by fusing slice information using a point spread function weighted averaging method that selects information based on estimated local blurring. The problem of the interpolation between images having different orientations is ill-posed and, most other image interpolation works apply only to one or two stacks of parallel slices. In addition, while methods such as [7] work under the assumption that image slices have identical contrasts, this assumption does not hold in general cases [2], [8], [9]. It has been demonstrated in [2] and [9] that the interpolation of intensities may be very sensitive to differences of gain and contrast in the images, thus producing interpolation artefacts which may bias the subsequent segmentation. Thus, [8], [9] introduced an extra step of intensity normalization before the volume interpolation.

This sequential approach lacks robustness, since a failure at an earlier stage is cascaded to all subsequent stages. Thus, a few attempts have been made to combine two or three of these stages in a single framework.

Integrated registration and segmentation – Prior-knowledge based segmentation usually requires the registration of a shape model to the image being segmented. Guyader and Vese [15], and Rousson, Paragios, and coworkers [16], [17] achieved this in a level set framework.

Registration and co-segmentation of overlapping images have been combined, mostly using statistical frameworks, to update the two stages simultaneously, e.g. [18] and [19] amongst others, or iteratively in turn, e.g. [20]. Yezzi *et al.* [21] presented a level set based method to register and co-segment two overlapping images or dense volumes simultaneously by minimizing an energy containing a region-based segmentation term and an alignment term. This method does not apply to our case of spaced images which contain different parts of an object of interest that may not have the same 2D geometry and topology. However, its integration strategy inspired our proposed integrated registration method for spaced data, as explained in Section 3.1. Prevost *et al.* [22] developed a similar method for co-segmenting the kidney from contrast-enhanced and 3D ultrasounds using additional prior-knowledge constraints on the recovered shape. Unal *et al.* [23] proposed a variation of the work in [21] for co-segmentation and regularized non-rigid registration of two volumes.

Integrated segmentation and shape interpolation – Methods which register and deform a model rely on the continuity of the model in the gaps between images. This can be viewed as simultaneous segmentation and shape interpolation. Amongst others, [24] used Active Shape Models (ASMs), [25] combined 2D Active Appearance Models (AAMs) and a 3D shape model, and [26] used a triangulated surface model. These methods require a training phase and lack flexibility.

In our previous works in [2] and [1], we proposed to integrate segmentation and shape interpolation in a level set framework in order to benefit from the accuracy and flexibility of level set methods. The level set segmenting surface was interpolated in the gaps in the data using a Mean Curvature Flow (MCF) in [2] and Radial Basis Functions (RBFs) in [1]. The shape interpolation strategy of [1] will be exploited in our proposed, fully integrated, framework.

When dealing with misaligned data, these integrated segmentation and shape interpolation methods require an independent registration to be performed beforehand. Such semi-sequential approaches with a separate registration stage will be shown to lack robustness, and a full integration of all three stages would be desirable. To the best of our knowledge, this has only been attempted by one other work to date which we describe next.

Integrated registration, segmentation, and shape interpolation – Zambal *et al.* [27] registered, segmented, and interpolated 3D cardiac MRI data made up of spaced parallel SA slices using a two component deformable model, relying on the continuity of this model to interpolate in the gaps. The model components were updated iteratively: 2D AAMs segmented the individual slices, and a global shape model linked the 2D AAMs together. Registration was performed at every iteration by aligning the AAMs on the global shape model. We also propose to perform registration, segmentation, and shape interpolation simultaneously, but unlike Zambal *et al.* [27] who used parametric models and prior knowledge, we combine the three stages in a level set framework, offering greater flexibility and accuracy, while requiring no prior knowledge.

3 Proposed Method

The registration method of [21] presented in Section 2 applies to the co-segmentation of dense and overlapping data only. However, we will show in Section 3.1 that this approach has physical grounds which we will use first in Section 3.2 to propose a variation that is more robust to local minima, and then in Section 3.3 to derive a new and more general registration method suited for the registration of slices with arbitrary spatial configurations and containing different parts of an object. This new registration method will be integrated in the segmentation and shape interpolation framework of [1] in Section 3.4.

3.1 Justification for the Registration Approach

As explained in Section 1, this work addresses misalignments that may be approximated by rigid motions. Let us

examine in this case the action of the registration of [21]:

$$\begin{cases} \frac{\partial C}{\partial t} = & f\mathbf{N} + |g'| \left(\hat{f} \circ g\right) \mathbf{N} - \kappa \mathbf{N} \quad (1a) \\ \frac{\partial g_i}{\partial t} = & \oint_C \left(\hat{f} \circ g\right) \left\langle \frac{\partial g}{\partial g_i}, \left((g')^{-1} |g'|\right)^T \mathbf{N} \right\rangle \quad (1b) \end{cases}$$

where C is the level set contour, $\mathbf{N}$ its normal, κ its curvature, and f and $\hat{f}$ may be seen as its speed in the target and registered image spaces respectively. g is the mapping from the registered image to the target image, with g_i its i^{th} parameter, and $|g'|$ is the determinant of g' the Jacobian matrix of g with respect to the spatial position.

Equation (1a) makes the contour evolve under the sum of the contributions of the two images. When aligning by translation $\mathbf{T}$ only, (1b) becomes

$$\frac{\partial T_i}{\partial t} = \oint_C \left\langle \mathbf{u}_i, \left(\hat{f} \circ g\right) \mathbf{N} \right\rangle, \quad (2)$$

where T_i is the i^{th} component of the alignment vector $\mathbf{T}$, and $\mathbf{u}_i$ is the unit vector oriented in the i^{th} direction. $\left(\hat{f} \circ g\right)(\mathbf{x})\,\mathbf{N}(\mathbf{x})$ is the displacement imposed on the contour at point $\mathbf{x}$ by the segmentation of the registered image. So (2) makes $\mathbf{T}$ evolve by the total segmentation displacement integrated on the contour.

When aligning by rotation around one axis only, for example the z-axis, with the rotation matrix being

$$R_\theta = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (3)$$

then, using trigonometric identities, (1b) becomes

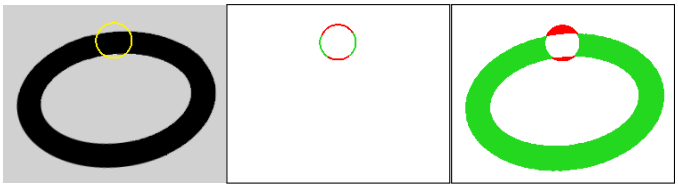
$$\begin{aligned} \frac{\partial \theta}{\partial t} &= \oint_C \left\langle R'_\theta \mathbf{x}, \left(\hat{f} \circ g\right) R_\theta \mathbf{N} \right\rangle = \oint_C \left\langle R_{\frac{\pi}{2}+\theta} \mathbf{x}, \left(\hat{f} \circ g\right) R_\theta \mathbf{N} \right\rangle \\ &= \oint_C \left\langle \mathbf{x}, \left(\hat{f} \circ g\right) R_{-\frac{\pi}{2}} \mathbf{N} \right\rangle. \end{aligned} \quad (4)$$

$\left\langle \mathbf{x}, \left(\hat{f} \circ g\right)(\mathbf{x})\, R_{-\frac{\pi}{2}} \mathbf{N}(\mathbf{x}) \right\rangle$ may be seen as the torque of a force applied to the contour at point $\mathbf{x}$, directed tangentially, and with an intensity of $\left(\hat{f} \circ g\right)(\mathbf{x})$. The sum of these torques along the contour C makes it rotate towards a better alignment with the registered image. Thus, the action of these two types of alignment (translation and rotation) can be understood as having physical grounds and it would make sense to apply them to more general cases than [21], such as non-overlapping images in a spaced volume.

3.2 Domain Extension of Image Force

In the original registration method of [21] expressed in (1b), only the points located on the level set contour are used to update the image registration by summing the pulls and torques of their "segmentation force" identified in Section 3.1. Since registration relies exclusively on these segmentation information, this implies that, before the contour converges on the edges of the object, the global geometry of the object is unknown[2], as illustrated in Fig. 4b. This makes the registration more likely to get trapped in local minima, especially in cases of high initial misalignments.

2. After convergence, the contour has the shape of the object, thus the segmentation information does contain information on the global geometry of the object.



(a) (b) (c)

Fig. 4. Speeds used for registration in the local and global variants of the registration method. (a) image and segmenting contour in yellow. (b) and (c) segmentation speeds $\hat{f} \circ g$, computed using the piecewise constant model from [28], on the contour C in (b) (local variant) and on $C \cup \Omega_{\text{diff}}$ in (c) (global variant) – red: negative; green: positive. Note that the global variant (c) exploits the extra information about the geometry of the object provided by $\hat{f} \circ g$ in Ω_{diff}.

We propose to address this issue by exploiting the "segmentation force" for registration in areas that provide a complementary description of the geometry of the object. Thus, we use the speed $\hat{f} \circ g$ both on the contour and in the domain Ω_{diff}

$$\Omega_{\text{diff}} = \left\{ \mathbf{x} \in I | \left(\hat{f} \circ g\right)(\mathbf{x}) \cdot \phi(\mathbf{x}) < 0 \right\}. \quad (5)$$

Ω_{diff} contains the relevant speeds $\hat{f} \circ g$ that describe the global geometry of the object in the registered image as well as the current geometry of C, as seen in Fig. 4c. Equation (1b) then becomes

$$\frac{\partial g_i}{\partial t} = \int_{C \cup \Omega_{\text{diff}}} \left(\hat{f} \circ g\right) \left\langle \frac{\partial g}{\partial g_i}, \left((g')^{-1} |g'|\right)^T \mathbf{N} \right\rangle. \quad (6)$$

As will be shown in Section 4, this use of additional segmentation information in Ω_{diff}, that describe the geometry of the object of interest even before it is matched by C, reduces the number of local minima which may bias the registration.

We refer to this variant of the registration method as the "global variant", as opposed to the original "local variant" which minimizes [21]'s co-segmentation energy and uses local information of the contour only. Note that "global" here refers to the domain rather than to a "global solution" of the optimization problem in [21]. Instead, this global variant may be seen as the minimization of an energy that penalizes a poor match between the shapes of the contour C and of the object in the image, similar to the energy used in [16], [17] for rigidly aligning shapes represented by two distance functions:

$$E(g) = \int_{\Omega_{\text{reg}}} \left(\phi(\mathbf{x}) - \phi_I(g(\mathbf{x}))\right)^2, \quad (7)$$

where ϕ and ϕ_I would be in our case the distance functions of contour C and of the shape in the image, respectively, and $\Omega_{\text{reg}} = \Omega_{\text{diff}}$. Indeed, by performing a similar examination of (7) as in Section 3.1, we find that

$$\frac{\partial T_i}{\partial t} = \int_{\Omega_{\text{reg}}} D(\mathbf{x}) \left\langle \mathbf{u}_i, \nabla \phi_I(g(\mathbf{x})) \right\rangle \approx \int_{\Omega_{\text{reg}}} D(\mathbf{x}) \left\langle \mathbf{u}_i, -\mathbf{N}_I \right\rangle \quad (8)$$

in the case of alignment by translation, and

$$\begin{aligned}\frac{\partial\theta}{\partial t} &= \int_{\Omega_{\text{reg}}} D(\mathbf{x}) \left\langle \mathbf{x}, R_{-\frac{\pi}{2}} R_{-\theta} \nabla\phi_I \left(g(\mathbf{x})\right)\right\rangle \\ &\approx \int_{\Omega_{\text{reg}}} D(\mathbf{x}) \left\langle \mathbf{x}, -R_{-\frac{\pi}{2}} \mathbf{N}_I \right\rangle\end{aligned} \tag{9}$$

for aligning by rotation, with $D(\mathbf{x}) = 2\left(\phi(\mathbf{x}) - \phi_I\left(g(\mathbf{x})\right)\right)$ and $\mathbf{N}_I$ the normal to the object contour in the reference space. Since the shape of the object of interest is unknown before segmentation is complete, its normal $\mathbf{N}_I$ cannot be computed in our case. In our formulation, this is not necessary and registration relies on the normal to the segmenting contour $\mathbf{N}$ instead. Nevertheless, the link of our global variant of registration with registration-based energy minimization demonstrates its theoretical foundation.

The contour normal $\mathbf{N}$ may be computed following the usual formula for level sets[3]: $\mathbf{N} = -\nabla\phi$. However, as $\mathbf{N}$ is now used potentially in the whole volume rather than only on contour C, we must make sure that $\nabla\phi$ is directed pointing inwards of the contour everywhere. We enforce this by adding the classic level set normalization term $sign\left(\phi\right)\left(1 - \|\nabla\phi\|\right)$ to the level set evolution function.

When registering cluttered images using the global variant, one must ensure that any background objects similar to the object of interest do not attract the contour and bias the registration. We enforce this by simply setting to zero any areas of positive speeds $\hat{f} \circ g$ which are not connected to the interior of the level set contour, so that only the object being segmented can drive registration. The local variant does not require this as it ignores the erroneous segmentation of spurious false object pixels in the background. This makes it more accurate, as was shown by our experiments.

Therefore, we recommend to use a combination of the two variants of registration in order to benefit from their two complementary advantages. The global variant should be used first to avoid local minima when there is significant misalignments. Then, when the level set contour matches the shape of the object sufficiently well, thus indicating a low risk of local minimum for registration, the local variant should be used to achieve best local fitting. We use the following shape matching condition between the contour C and the object of interest:

$$\left|\Omega_{\left(\hat{f}\circ g\right)>0} \cap \Omega_{\phi>0}\right| > \left|\Omega_{\text{diff}}\right| \ , \tag{10}$$

where $|.|$ denotes cardinality, i.e. the intersection between the shape of the contour C and the shape of the object is greater than the union of their relative complements.

3.3 Proposed Registration Method

Based on the analysis of the registration method in [21] that we performed in Section 3.1, we propose to align internally, by translation and rotation, a 3D or 4D volume made up of any number of spaced 2D slices that have arbitrary relative positions and orientations. At the same time, we segment an object in the volume and interpolate its shape where no image information is available. For each sequence n of M_n 2D images, we define a mapping function $g^n\left(\mathbf{x}\right) = R^n\mathbf{x} + \mathbf{T}^n$ (where $R = R_\chi R_\xi R_\zeta$ is the 3D rotation matrix) from the common 3D space to the local 3D space associated to the sequence. The update of g^n moves all M_n images of sequence n in unison since, as explained in the introduction, all the images of the same sequence are considered as internally registered. The evolution of the level set function ϕ and of the mapping functions g^n is given by our generalization of Yezzi *et al.* [21]'s evolution equations for the case of registration by translation and rotation, i.e.

$$f = \sum_n f^n \circ g^n \ , \tag{11}$$

$$\left\{\begin{array}{rcl} \frac{\partial T_i^n}{\partial t} &=& \frac{1}{\left|\Omega_{\text{reg}}^n\right|}\int_{\Omega_{\text{reg}}^n}\left(f^n\circ g^n\right)\left\langle \mathbf{u}_i, R^n\mathbf{N}\right\rangle \\ \frac{\partial\zeta^n}{\partial t} &=& \frac{1}{\left|\Omega_{\text{reg}}^n\right|}\int_{\Omega_{\text{reg}}^n}\left(f^n\circ g^n\right) \\ & & \left\langle R_{\chi^n}R_{\xi^n}R_{\frac{\pi}{2}+\zeta^n}\mathbf{x} + \mathbf{T}^n, R^n\mathbf{N}\right\rangle \\ \frac{\partial\xi^n}{\partial t} &=& \frac{1}{\left|\Omega_{\text{reg}}^n\right|}\int_{\Omega_{\text{reg}}^n}\left(f^n\circ g^n\right) \\ & & \left\langle R_{\chi^n}R_{\frac{\pi}{2}+\xi^n}R_{\zeta^n}\mathbf{x} + \mathbf{T}^n, R^n\mathbf{N}\right\rangle \\ \frac{\partial\chi^n}{\partial t} &=& \frac{1}{\left|\Omega_{\text{reg}}^n\right|}\int_{\Omega_{\text{reg}}^n}\left(f^n\circ g^n\right) \\ & & \left\langle R_{\frac{\pi}{2}+\chi^n}R_{\xi^n}R_{\zeta^n}\mathbf{x} + \mathbf{T}^n, R^n\mathbf{N}\right\rangle \end{array}\right. \tag{12}$$

with $\Omega_{\text{reg}}^n = \Omega_n \cap C$ for the local variant of the registration method, and $\Omega_{\text{reg}}^n = \left(\Omega_n \cap C\right) \cup \Omega_{\text{diff}}^n$ for its global variant. Ω_n is the domain of the images of sequence n within the common 3D space Ω. Adapting the definition of Ω_{diff} in (5), Ω_{diff}^n is defined as

$$\Omega_{\text{diff}}^n = \left\{\mathbf{x} \in \Omega_n \middle| \left(f^n \circ g^n\right)\left(\mathbf{x}\right)\cdot\phi\left(\mathbf{x}\right) < 0\right\} \ . \tag{13}$$

Note that Yezzi *et al.* [21] with rigid registration is a special case of (12), with the local variant of Ω_{reg}^n, two images or volumes n_1 and n_2, and $\Omega_{n_1} = \Omega_{n_2}$.

In [21], the amount of displacement applied to the registered image is proportional to its remaining misalignment with the target image. Indeed, in (2) and (4) the amount of misalignment between the registered image and the contour C is an implicit approximation of the misalignment between the two images in [21]'s case of co-segmentation through registration based regularization. Although this proportionality of displacement with misalignment may be seen as a way to improve convergence speed, it takes a much more fundamental role in the registration process in our case of aligning multiple (more than two) non-overlapping images. It inactivates or activates registration for images that are well-aligned or misaligned respectively. We noticed in our experiments that this avoids the former to be sent away, jeopardizing the whole registration and segmentation process. It also avoids the latter to stay immobile, giving the segmenting contour C the opportunity to reach its segmentation optimum in most of the image although its position is not consistent with the rest of the dataset. This would block the image to its misaligned position and also endanger the whole process. Thus, moving the sequences by amounts proportional to their remaining misalignment is essential to ensure registration behaves as expected.

However, in our case, the implicit approximation of misalignment of [21] does not hold because our 2D images contain different parts of an object, and because very few points of the segmenting contour in the images of sequence n intersect other sequences. Thus, in our first experiments

3. We use here the convention $\phi(\mathbf{x}) > 0$ inside the contour.

which relied on the implicit approximation of misalignment, registration suffered from the issues described previously and could not be performed. Therefore, we need to compute an explicit estimate of the misalignment λ_n of sequence n. We propose to do so by exploiting the contour's speeds f^n and f^m ($m \neq n$) that contain information on the shape of the object of interest in both sequences. We examine their dissimilarity at the intersections between Ω^n_{reg} and Ω^m_{reg}. When using the local variant of the registration method, $\lambda_n = \lambda_n^{\text{local}}$ is estimated as

$$\lambda_n^{\text{local}} = \frac{\left|\Omega_{(f^n \circ g^n)\cdot(f^m \circ g^m)<0} \cap C\right|}{\left|\Omega_n \cap \Omega_m \cap C\right|} , \tag{14}$$

and for the global variant, the estimation of $\lambda_n = \lambda_n^{\text{global}}$ is

$$\lambda_n^{\text{global}} = \frac{\left|\Omega_{(f^n \circ g^n)\cdot(f^m \circ g^m)<0} \cap \left(C \cup \Omega^n_{\text{diff}} \cup \Omega^m_{\text{diff}}\right)\right|}{\left|\Omega_n \cap \Omega_m \cap \left(C \cup \Omega^n_{\text{diff}} \cup \Omega^m_{\text{diff}}\right)\right|} . \tag{15}$$

These explicit estimations of misalignment λ_n are then used to modulate the amount of registration, represented as the amounts of translation (in pixels) and rotation (in degrees). This modulation is achieved through the update of the norms of $\frac{\partial \mathbf{T}^n}{\partial t}$ and $\frac{\partial \boldsymbol{\theta}^n}{\partial t}$ (with $\boldsymbol{\theta}^n = \left[\ \zeta^n,\ \ \xi^n,\ \ \chi^n\ \right]^T$ being the vector of rotation angles of R^n). Although $\left\|\frac{\partial \boldsymbol{\theta}^n}{\partial t}\right\|$ only approximates the amount of rotation of R^n, its computation is simpler than finding the rotation axis and angle of rotation of R^n, and the scheme we present next in (17) and (19) was found to satisfactorily match the amount of rotation of a sequence n to its degree of misalignment.

We use piecewise linear schemes for the computation of the amounts of shift and rotation in order to respect the principle of proportionality of [21]. We empirically established the following threshold values that provide a satisfactory behavior of our registration method for all our applications:

$$\textit{shift}^n = \begin{cases} 1 & \lambda_n \geq 0.6 \\ \text{linear } 0.1 \rightarrow 1 & 0.3 \leq \lambda_n < 0.6 \\ \min\left(\text{linear } 0 \rightarrow 0.1, \left\|\frac{\partial \mathbf{T}^n}{\partial t}\right\|\right) & \lambda_n < 0.3 \end{cases} , \tag{16}$$

$$\textit{rotation}^n = \begin{cases} r_1 & \lambda_n \geq 0.6 \\ \text{linear } r_2 \rightarrow r_1 & 0.2 \leq \lambda_n < 0.6 \\ \min\left(\text{linear } 0 \rightarrow r_2, \left\|\frac{\partial \boldsymbol{\theta}^n}{\partial t}\right\|\right) & \lambda_n < 0.2 \end{cases} , \tag{17}$$

with r_1 and r_2 representing rotations of 1° and 0.1° respectively around each axis: $r_1 = \sqrt{d \cdot 1^{\circ 2}}$ and $r_2 = \sqrt{d \cdot 0.1^{\circ 2}}$ (d being the number of degrees of freedom for rotation, i.e. the number of dimensions of the data). The modified shift and rotation vectors are then

$$\widetilde{\frac{\partial \mathbf{T}^n}{\partial t}} = \frac{\frac{\partial \mathbf{T}^n}{\partial t}}{\left\|\frac{\partial \mathbf{T}^n}{\partial t}\right\|} \cdot \textit{shift}^n , \tag{18}$$

$$\widetilde{\frac{\partial \boldsymbol{\theta}^n}{\partial t}} = \frac{\frac{\partial \boldsymbol{\theta}^n}{\partial t}}{\left\|\frac{\partial \boldsymbol{\theta}^n}{\partial t}\right\|} \cdot \textit{rotation}^n . \tag{19}$$

It is desirable to limit the amount of deformation that the level set surface must undergo at each iteration in order to adapt to the displacements of the images. We achieve this by restricting translations and rotations to their minimal useful parts, by subtracting an offset to the update of each component g_i^n of g^n. This offset is the minimum displacement $\widetilde{\frac{\partial g_i}{\partial t}}_{\min}$, in absolute values, amongst all sequences, i.e.

$$\textit{offset}_i = \widetilde{\frac{\partial g_i}{\partial t}}_{\min} , \tag{20}$$

with $\widetilde{\frac{\partial g_i}{\partial t}}_{\min}$ being such that

$$\left\|\widetilde{\frac{\partial g_i}{\partial t}}_{\min}\right\| = \min_n \left\|\widetilde{\frac{\partial g_i^n}{\partial t}}\right\| . \tag{21}$$

This offset subtraction has the additional advantage of avoiding the modeled object to drift out of the boundaries of the volume. Thus, it is not necessary to arbitrarily choose a sequence to remain fixed in the volume.

Finally, $\mathbf{T}^n$ and R^n are updated as:

$$\begin{cases} \mathbf{T}^n_{t+1} &= \mathbf{T}^n_t + dt\left(\widetilde{\frac{\partial \mathbf{T}^n}{\partial t}} - \textit{offset}_{\mathbf{T}}\right) \\ \boldsymbol{\theta}^n_{t+1} &= \boldsymbol{\theta}^n_t + dt\left(\widetilde{\frac{\partial \boldsymbol{\theta}^n}{\partial t}} - \textit{offset}_{\boldsymbol{\theta}}\right), \end{cases} \tag{22}$$

with $\textit{offset}_{\mathbf{T}} = \left[\ \textit{offset}_{T_x},\ \ \textit{offset}_{T_y},\ \ \textit{offset}_{T_z}\ \right]^T$ and $\textit{offset}_{\boldsymbol{\theta}} = \left[\ \textit{offset}_{\zeta},\ \ \textit{offset}_{\xi},\ \ \textit{offset}_{\chi}\ \right]^T$.

3.4 Proposed Integrated Registration, Segmentation, and Interpolation

The integration of this registration method with the segmentation and shape interpolation framework of [1] is performed by providing the speed f of the contour C in (11) to the level set evolution equations (23) and (24) of [1], in order to compute the evolution of the level set function ϕ both on the domains of the images and in the gaps between them:

$$\begin{aligned} \frac{d\alpha_i}{dt} &= \oint_C f(\mathbf{x})\,\psi_i(\mathbf{x}) \approx \int_\Omega \delta_\epsilon(\phi(\mathbf{x}))\, f(\mathbf{x})\,\psi(\mathbf{x}_i - \mathbf{x}) \\ &\approx \left((\delta_\epsilon(\phi)\cdot f) \star \psi\right)(\mathbf{x}_i) , \end{aligned} \tag{23}$$

$$\frac{\partial \phi}{\partial t}(\mathbf{x}) = \sum_{i=1}^{N} \frac{d\alpha_i}{dt}\psi(\mathbf{x} - \mathbf{x}_i) = \left(\frac{d\alpha}{dt} \star \psi\right)(\mathbf{x}). \tag{24}$$

δ_ε is an approximation of the Dirac function δ with width ϵ, and $\star$ denotes convolution. ψ_i, $\mathbf{x}_i$, and α_i are the N RBFs, their centers, and their coefficients, introduced in [1] to interpolate ϕ:

$$\phi(\mathbf{x}) = \sum_{i=1}^{N} \alpha_i \psi_i(\mathbf{x}) = \sum_{i=1}^{N} \alpha_i \psi(\mathbf{x} - \mathbf{x}_i) . \tag{25}$$

As in [1] we use inverse multiquadric RBFs of the form $\psi(\mathbf{x}) = \left(\|\mathbf{x}\|^2 + \gamma^2\right)^{-\frac{\beta}{2}}$, with β being the number of dimensions of the data, to ensure the decrease of the RBF is not too steep, and γ being a variable which must be set by the user according to the data in order to fine tune the flatness of the RBF.

As seen in (11), f combines all the speeds f^n that are computed for the independent sequences n. They may be produced by any segmentation method of choice. The term $\kappa\mathbf{N}$ in (1a) was dropped in (11) because the level set

framework of [1] does not require contour smoothing due to the intrinsic smoothing of the convolutions by the RBF.

As stated in Section 3.2, when using the global variant of the registration method, the classic level set normalization term needs to be used in order to ensure that $\mathbf{N}$ can be computed easily everywhere. This term does not depend on the availability of data and thus is defined everywhere and does not require interpolation. In addition, its smoothing by convolutions with the RBF is not desirable and may degrade its performance. So we add it to (24) to update ϕ directly rather than incorporating it into f in (23):

$$\frac{\partial \phi}{\partial t}(\mathbf{x}) = \left(\frac{d\alpha}{dt} \star \psi \right)(\mathbf{x}) + sign\left(\phi(\mathbf{x})\right)\left(1 - \|\nabla \phi(\mathbf{x})\|\right). \quad (26)$$

We use the same timestep dt to update C and all g^n. As in [1], it is set so that C moves by 1 pixel per iteration at the beginning of the process, and it is reduced progressively when oscillations of C are detected in order to allow a finer convergence of the registration. Convergence of the algorithm is reached at the same time as convergence of the level set segmentation, since any additional motion of an image sequence would force the segmentation to adapt to the new configuration of the data. The automatic detection of the convergence of a level set contour is not in the scope of this work, and we use the simplest scheme with fixed number of iterations in our implementation.

The overall algorithm for the proposed framework is outlined in Algorithm 1 and our software implementation is available on our project web page[4].

Algorithm 1 – Integrated Registration, Segmentation, and Shape Interpolation for Spaced and Misaligned Data

```
Initialize a contour C
Initialize all g^n to the identity transformation
Initialize a boolean global_variant to true if local minima are expected, false otherwise
repeat
    Compute dα/dt using (11) and (23)
    if global_variant = true then
        Update C using (26)
    else
        Update C using (24)
    end if
    for all image sequence n do
        Compute ∂T^n/∂t and ∂θ^n/∂t using (12)
        if global_variant = true then
            Estimate misalignment λ_n using (15)
            if condition (10) = true then
                Set global_variant to false
            end if
        else
            Estimate misalignment λ_n using (14)
        end if
        Compute shift^n and rotation^n using (16) and (17)
        Compute ∂T̃^n/∂t and ∂θ̃^n/∂t using (18) and (19)
    end for
    for all degree of freedom i do
        Compute offset_i using (20) and (21)
    end for
    for all image sequence n do
        Update T^n and R^n using (22)
    end for
until convergence
```

4 Experimental Results and Analysis

To assess and evaluate the proposed methodology, we first focus on the accuracy and robustness of the proposed *registration* approach, and then we evaluate the *3D modeling of the complete framework*.

Registration methods that are designed to register two target and reference images containing the same object e.g. [15], [16], [17], [18], [19], [20], [21], [22], [23], are not suitable for comparison as they cannot align 2D images that contain different parts of a 3D volume and of an object of interest. The methods presented in [5], [6] are limited to temporal sequences forming two orthogonal stacks of slices. The methods in [7], [8], [9] only apply to three orthogonal stacks of slices. We compare against the algorithm of [4] that was described in Section 2 and that is popular for aligning cardiac MRIs. Although this method was originally designed for two orthogonal stacks of temporal sequences and was applied to translation only, the same algorithm can be applied to more general cases. It is therefore suitable for comparison, and it is the only method we know of that can align slices that contain different parts of a 3D volume and object, and that have arbitrary spatial configurations. We refer to our extended use of [4] as "SR [4]" ("Sequential Registration [4]"). The method is based on similarity measures and in [4] it is implemented with NMI that is commonly used to register multi-modal images. We retain this similarity measure in our experiments, and, in Experiment 1 with uni-modal data, we also compare against [4] implemented with two other similarity measures that only apply to uni-modal data, namely Sum of Squared intensity Differences (SSD) and Normalized Cross-Correlation (NCC). These added tests aim at demonstrating that the lack of robustness that we find in [4] is not due to the choice of similarity measure, but rather to the method relying on a local computation of a similarity measure.

We do not evaluate the quality of the *segmentation*, since this quality is highly dependent on the segmentation algorithm, which must be chosen by the user depending on the nature of the data. The choice of an optimum segmentation algorithm is not within the scope of this work.

The *shape interpolation* contribution of our framework was already evaluated in [1]. Indeed, the combined segmentation and interpolation of [1] is now used in our fully integrated framework. In [1], we compared this integrated method against a state-of-the-art sequential image interpolation [14] followed by 3D level set segmentation. This revealed that the combined approach, by taking into account the global geometry of the objects, produces more robust and accurate results than the sequential one.

We do not repeat this demonstration in this work, and instead we focus on the synergistic benefit of integrating registration into the framework of [1]. Therefore, when assessing the performance of the *complete framework*, inte-

4. See www.cs.bris.ac.uk/~paiement/IReSISD.html.

TABLE 1
Datasets used in the experiments

Dataset	Orientation	Nb. slices	Pixel size (mm)	Slice spacing	Time-frames	Nb. datasets	ISI scheme
Artificial cardiac MRI	SA LA	4 to 19 1 to 12	1.77 to 2.08	5 mm to 20 mm –	1	14	2 steps: $\gamma = 0.7$ then 3
Artificial brain MRI	Axial (2 sets) Sagittal (1 or 2 sets)	11 to 18 13 to 20	0.71 to 0.94 0.37 to 0.90	3.3 mm to 6.6 mm 3.3 mm to 4.8 mm	1	9	Simple: $\gamma = 2$
Kidney CT	Axial Sagittal	29, 15, 10 and 8 103, 52, 35 and 26	0.78	5, 10, 15 and 20 pixels	1	4	Simple: $\gamma = 3$, 4, 5 and 6
Hip bone CT	Axial Sagittal	27, 14, 9 and 7 103, 52, 35 and 26	0.49	5, 10, 15 and 20 pixels	1	4	Simple: $\gamma = 3$, 5, 7 and 9
SA+LA cardiac MRI	SA LA	4 to 19 1 to 12	1.77 to 2.08	5 mm to 20 mm –	25	14	2 steps: $\gamma = 1$ then 4
Radial cardiac MRI	LA SA	12 1	1.77	~ 15 degrees	25	1	2 steps: $\gamma = 1$ then 4
Brain MRI	Axial (2 sets) Sagittal (1 or 2 sets)	11 to 18 13 to 20	0.71 to 0.94 0.37 to 0.90	3.3 mm to 6.6 mm 3.3 mm to 4.8 mm	1	9	Simple: $\gamma = 1.5$
Hip MRI	Axial Coronal (2 sets) Diagonal (2 sets)	19 19 15	0.74 1.64 0.86	6 mm 4.8 mm 3.3 mm	1	1	Simple: $\gamma = 1.5$

grating all stages of registration, segmentation, and interpolation, we compare against the semi-sequential approach which successively applies the registration SR [4] based on a similarity measure, and the integrated segmentation and interpolation method of [1][5]. We refer to this framework as "SR+ISI", where ISI is for "Integrated Segmentation and Interpolation". Based on the comparison results in [1], we can infer that the fully sequential method SR [4] followed successively by image interpolation and 3D segmentation would produce less robust and accurate results than SR+ISI. Thus, the comparison against SR+ISI is fairer for the sequential approach.

4.1 Data and Evaluation Protocol

In our experiments, we distinguish between two cases:

- Spatial sequences, as in Figs. 2b and 2c, that are made up of a stack of slices that move together during registration. In the rest of this article, we will denote such registration as "stack-wise" registration.
- Temporal sequences made up of a number of time-frames of a *single* slice position and orientation, as in Fig. 2a. In this case, the different slices which span the volume are registered with each other. We denote such registration as "slice-wise" registration.

The registration stage alone and the complete integrated framework are both evaluated quantitatively on artificial data in Section 4.2 and on artificial data made up of real images in Section 4.3. Section 4.4 presents qualitative results on real datasets. The datasets used in these experiments are presented in Table 1.

The artificial data used in Section 4.2 was created to allow the assessment of our registration method and of our integrated framework independently of the quality of segmentation. Thus, it is used to establish the *limits of performance* of our method in the case of a perfect segmentation. Indeed, as long as the segmentation is correct, the registration quality depends only on the geometry of the data. As explained in Section 4, the quality and robustness of segmentation depends strongly on the chosen segmentation algorithm and this choice is not within the scope of this work. We avoid the problem of finding a perfect segmentation algorithm by producing artificial data that are very easily segmented, i.e. images consisting of piecewise constant regions. We generate such images from a model of a left ventricle (LV) cavity of a heart and from a model of both ventricles of a neonatal brain that were created by segmenting, using the method of [1], two real MRI datasets which were visually inspected for no misalignments and requiring only a minimal amount of interpolation. Our artificial datasets were extracted from these two volumes using the positions and orientations of slices of 14 heart MRIs and 9 brain MRIs respectively. Therefore, although our artificial images are made up of piecewise constant regions, they contain the shape of real organs and have realistic positions and orientations. Since cardiac and brain MRIs usually comprise temporal and spatial sequences respectively, the artificial cardiac and brain datasets are used to evaluate the slice-wise and stack-wise registration respectively. Only one time-frame was generated for the artificial cardiac data, which is the worst case for slice-wise registration as it restricts the amount of available information.

Our quantitative evaluation of Section 4.3 is on artificial data made up of real images, in order to provide quantitative evaluation on more plausible images and segmentation quality while still having registration and interpolation groundtruths. We create datasets with artificial gaps and misalignments by extracting spaced axial and sagittal slices from full-volume real CT-scans of hip and kidney which have no misalignments. Thus, the resulting spaced datasets are made up of real images but have artificial gaps.

In Sections 4.2 and 4.3, artificial initial misalignments are induced by initializing the registration vectors $\mathbf{T}^n$ and θ^n of the individual sequences to random values. A perfect registration should move all the sequences of a dataset to the same position, which we call "reference sequence position" (RSP). Since we do not use a target sequence, the RSP may

5. Note, we did not compare with the fully integrated framework presented in [27] since the authors informed us that their code and prior (AAM and 3D shape) models were no longer available.

TABLE 2
Stack-wise registration accuracy on artificial data: mean and std of registration error (in pixels and degrees) averaged over all 9 datasets

Initial misalignment	SR [4] with SSD	SR [4] with NCC	SR [4] with NMI	Proposed
$T_{i\ \text{ini}}^{n} = \text{rand}\,[-10, 10]$	(0.50, 1.07, 0.36)	(0.06, 0.13, 0.13)	(0.02, 0.05, 0.08)	(0.05, 0.08, 0.09)
$\theta_{i\ \text{ini}}^{n} = \text{rand}\,[-10°, 10°]$	±(1.8e-3, 1.8e-2,6.7e-2)	±(4.1e-3, 3.3e-2, 7.9e-3)	±(1.2e-4, 4.3e-3, 2.3e-3)	±(3e-3, 6.7e-3, 2.5e-3)
	(0.21°, 0.23°, 1.69°)	(0.40°, 0.24°, 0.10°)	(0.05°, 0.08°, 0.05°)	(0.27°, 0.35°, 0.22°)
	±(6.7e-4, 6.0e-4, 2.5e-4)	±(4.5e-3, 7.2e-4, 1.5e-4)	±(2e-5, 3.8e-5, 2.6e-5)	±(2e-3, 2.8e-3, 7e-4)
$T_{i\ \text{ini}}^{n} = \text{rand}\,[-20, 20]$	(8.09, 8.68, 4.37)	(0.63, 2.26, 1.2)	(0.74, 2.61, 1.99)	(0.05, 0.08, 0.07)
$\theta_{i\ \text{ini}}^{n} = \text{rand}\,[-20°, 20°]$	±(132, 57.5, 15)	±(0.3, 5.39, 1.04)	±(0.33, 6.81, 2.11)	±(2e-3, 4.6e-3, 3.4e-3)
	(7.76°, 4.76°, 5.78°)	(3.34°, 2.27°, 1.14°)	(2.15°, 4.70°, 2.14°)	(0.34°, 0.40°, 0.22°)
	±(0.88, 0.22, 0.54)	±(0.13, 0.08, 0.02)	±(0.05, 0.23, 19.47)	±(2.7e-3, 3.9e-3, 6.9e-4)

take any value in the 6D space of translations and rotations. In case of misalignments, we define the RSP as the mean position of the sequences. We then compute the registration error as the average and standard deviation of the distances of the sequences to the RSP. The complete framework is evaluated by computing the Jaccard coefficients of the shapes reconstructed from the initial misalignments. To do this, the groundtruth shape is first placed at the RSP.

In Section 4.4, we use real datasets for which no registration and interpolation groundtruth is available, thus we assess the performance of the compared methods qualitatively. The datasets were obtained from hospital patients in real and routine clinical examinations. They include cardiac MRIs that comprise temporal sequences, and neonatal brain and hip MRIs that consist of spatial sequences.

We use the same segmentation algorithm for the proposed and sequential methods. As in [1], we use the piecewise constant model from [28] for our artificial images, and a method based on a piecewise model and a Parzen window estimator of pixel intensity for our real images. The level set surface is initialized as a small sphere, and its initial position initializes the segmentation models. For artificial data, since the parameters of the segmentation model are known constant intensities, the initial sphere is arbitrarily placed at the center of the volume. For real data, it is placed by the user by clicking at the center of the object on a single central image. Note that the initial sphere does not need to be positioned inside the object for all slices (and it is usually not, especially in case of large misalignments) as long as it allows a suitable initialization of the segmentation parameters. The heart datasets are processed with the scheme in two steps that was introduced in [1]. Other datasets are processed with the simple scheme. We refer the reader to [1] for details on both schemes. The schemes and values of RBF parameter γ are reported in Table 1.

4.2 Quantitative Evaluation on Artificial Data

4.2.1 Registration accuracy

Stack-wise registration – We perform two tests of stack-wise registration on the 9 artificial brain datasets, with two different ranges of initial misalignments. In the first test, the sequences are initially randomly shifted by up to 10 pixels in all directions ($T_{i\ \text{ini}}^{n} = \text{rand}\,[-10, 10]$) and randomly rotated by up to 10° along all axes ($\theta_{i\ \text{ini}}^{n} = \text{rand}\,[-10°, 10°]$, with θ_i^n the i^{th} component of θ^n). In the second test, we have $T_{i\ \text{ini}}^{n} = \text{rand}\,[-20, 20]$ and $\theta_{i\ \text{ini}}^{n} = \text{rand}\,[-20°, 20°]$. Table 2 shows registration errors for all degrees of freedom in pixels and degrees, averaged on all 9 datasets.

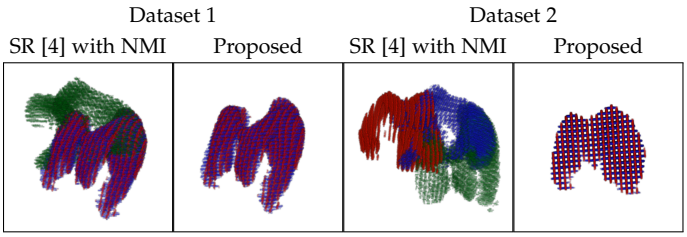


Fig. 5. Registration of artificial brain datasets that have local minima. SR [4] is implemented with NMI, and implementations with SSD and NCC suffered from the same local minima issues. Colors correspond to different sequences, and may not be all visible in case of near perfect alignment. For visualization purposes, only pixels containing the object are displayed. Due to space constraints, the groundtruth registration is not shown as it has no visible differences with the results of the proposed method.

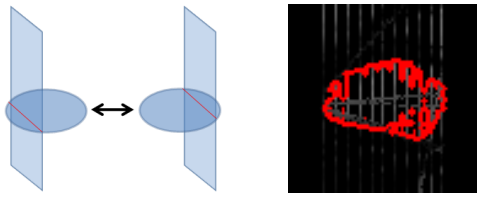
In the first test, the SR [4] method with NMI is slightly more accurate than the proposed registration method, but both methods achieve sub-pixel and sub-degree accuracy. We believe this small difference may be explained by the fact that in this "not so sparse" scenario (see Fig. 5) the many pixel intensity information that SR [4] uses both inside and outside of the object may constrain it better than the 3D shape information used by our method. In the second test with more considerable initial misalignments, the overall performance of the SR [4] method with all tested similarity measures is lower than the proposed registration method. This is due to the SR [4] method (regardless of the similarity measure used) being sometimes trapped in local minima, even in the absence of any background object. This is illustrated in Fig. 5 where two different ventricles of the brain are confused as one single object and superimposed. Conversely, the proposed registration method exploits information on the global shape of the objects of interest, and therefore is more robust to such local minima and maintains a good accuracy.

Slice-wise registration – We test slice-wise registration on the artificial cardiac datasets with alignment by translation in the SA plane first, and then by 3D translations and rotations, in both cases with a small then large range of initial misalignments. Table 3 presents the registration errors of these 4 tests, averaged on all 14 datasets.

The first two tests of SA plane translations are motivated by the fact that real cardiac MRIs acquired during a breath hold at end-expiration tend to suffer mainly from shifts in the SA plane, and have negligible shifts along the LA direction and rotations. For the first test with a small range of initial misalignments ($T_{xy_{\text{ini}}}^{n} = \text{rand}\,[-5, 5]$),

TABLE 3
Slice-wise registration accuracy on artificial data: mean and std of registration error (in pixels and degrees) averaged over all 14 datasets

Initial misalignment	SR [4] with SSD	SR [4] with NCC	SR [4] with NMI	Proposed
$T^n_{xy_{\text{ini}}} = \text{rand}\,[-5, 5]$	(0.72, 0.98) ±(0.71, 2.59)	(0.62, 0.64) ±(0.65, 0.72)	(0.64, 0.77) ±(0.65, 0.87)	(0.36, 0.39) ±(0.42, 0.2)
$T^n_{xy_{\text{ini}}} = \text{rand}\,[-10, 10]$	(1.92, 2.29) ±(8.67, 7.75)	(1.27, 1.71) ±(4.73, 7.79)	(1.87, 1.99) ±(3.72, 6.11)	(0.36, 0.36) ±(0.44, 0.18)
$T^n_{xy_{\text{ini}}} = \text{rand}\,[-5, 5]$ $T^n_{z\ \text{ini}} = \text{rand}\,[-3, 3]$ $\theta^n_{i\ \text{ini}} = \text{rand}\,[-5°, 5°]$	(1.68, 1.72, 1.72) ±(3.8, 4.12, 1.71) (3.48°, 3.33°, 3.71°) ±(0.11, 0.1, 0.13)	(1.4, 1.33, 2.09) ±(1.37, 1.82, 3.12) (3.31°, 3.44°, 4.05°) ±(0.11, 0.11, 0.16)	(1.29, 1.51, 2.02) ±(1.36, 2.16, 2.54) (3.23°, 2.96°, 3.2°) ±(0.09, 0.1, 0.09)	(1.05, 0.78, 0.88) ±(1.12, 0.68, 0.56) (3.1°, 2.97°, 2.65°) ±(0.09, 0.09, 0.07)
$T^n_{xy_{\text{ini}}} = \text{rand}\,[-10, 10]$ $T^n_{z\ \text{ini}} = \text{rand}\,[-4, 4]$ $\theta^n_{i\ \text{ini}} = \text{rand}\,[-5°, 5°]$	(3.19, 2.91, 2.82) ±(11.9, 11, 7.12) (3.87°, 3.57°, 4.11°) ±(0.14, 0.11, 0.22)	(2.56, 1.89, 2.56) ±(5.05, 3.48, 3.88) (3.95°, 3.97°, 4.44°) ±(0.17, 0.13, 0.17)	(2.91, 2.72, 2.08) ±(4.86, 4.45, 1.95) (3.04°, 2.95°, 2.95°) ±(0.08, 0.06, 0.07)	(1.2, 1.15, 1.13) ±(3.65, 1.86, 0.78) (3.12°, 2.52°, 2.79°) ±(0.09, 0.09, 0.07)



(a) (b)

Fig. 6. Two possible difficulties in slice-wise registration: (a) ambiguity in the position of a slice due to insufficient constraints, and (b) contour folding around individual slices due to a too sharp RBF.

TABLE 4
Influence of γ on slice-wise registration accuracy

RBF flatness	Average mean and std of registration error (in pixels)
$\gamma = 0.5$	(0.569, 0.466)±(0.448, 0.237)
$\gamma = 1$	(0.353, 0.331)±(0.245, 0.202)

while all methods achieve sub-pixel accuracy, the proposed registration method obtains results that are almost twice as precise as the best results of the SR [4] method. The second test (second row of Table 3) is a quite extreme case, with shifts of up to 10 pixels, equivalent to around 20 mm given the image resolution. However, the performance of the proposed registration method remains as good as in the first test, while the accuracy of the SR [4] method is strongly degraded for all 3 similarity measures.

A similar behavior is observed for the tests with full 3D translations and rotations, although the overall accuracy of all methods is reduced in this more difficult registration case. For a small range of initial misalignments (third row of Table 3), the two methods achieve a similar accuracy, with the proposed registration method being slightly more accurate than the best results of the SR [4] method. In the fourth row of Table 3 with a larger range of misalignments, the proposed registration method maintains a good accuracy while the average errors of the SR [4] method are higher.

Note that the use of three different similarity measures (NMI, SSD, and NCC) did not significantly change the results of SR [4] in all 4 tests. We believe that the differences in the results between SR [4] and the proposed method on the artificial cardiac datasets is due to the fact that SR [4] uses the intensity at the intersections of images to compute a similarity measure, and therefore requires a minimum number of intersection points in order to obtain meaningful similarity measures. Indeed, in [4] the datasets were made up of two stacks of 6 and 8 parallel SA and LA slices respectively, both spanning the whole LV of the heart, thus the numbers of intersections in our datasets are significantly lower. On the contrary, the proposed registration method uses *global* 3D segmentation results rather than *local* pixel intensity, and therefore it can handle datasets where the LA slices are less numerous (e.g. 2 or 3) and cross the imaged object at various angles and locations.

We note in the last two rows of Table 3 and on Fig. 8 that the rotation accuracy of our method seems, visually and subjectively, less satisfactory than for its translation – although always within a 3° error. This may be explained by considering the nearly cylindrical shape of the modeled object which leaves more degrees of freedom to the rotation of the slices for recovering an approximated cylinder shape.

A difficulty which sometimes arises when an image intersects very few other images, is that a position ambiguity can happen, as illustrated in Fig. 6a, where two different positions can be viewed as equally correct. Such under-constrained situations happened several times during our slice-wise registration tests, and a few datasets had one or two of their LA slices caught in a local minimum at an ambiguous position, especially in the third and fourth tests. This resulted in a wider distribution of the slices around their mean position in these datasets, which explains the large average standard deviations presented in the last two rows of Table 3. Most of the time, this did not prevent the contour to recover the proper shape of the object since both ambiguous positions were acceptable in a topological point of view, as will be illustrated in Section 4.2.2.

The flatness of the RBF has been found to have an influence on the registration results. When the RBF is too sharp, the contour can sometimes fold around a slice, as in Fig. 6b, trapping the slice in its current misalignment. This problem can be solved by using a flatter RBF, as illustrated by Table 4 which presents average registration errors when aligning the artificial heart datasets in the SA plane using two different RBF flatness parameters (note that contrary to the previous tests, only one value for γ is used per processing). We see that the flat RBF (with $\gamma = 1$) avoids contour folding traps, and thus yields better results than the sharp RBF (with $\gamma = 0.5$). In the 4 previous tests, the segmentation and interpolation stages of the integrated framework required the use of a sharp and a flat RBF successively, therefore the second processing step with a flat RBF corrected any possible folding of the contour.

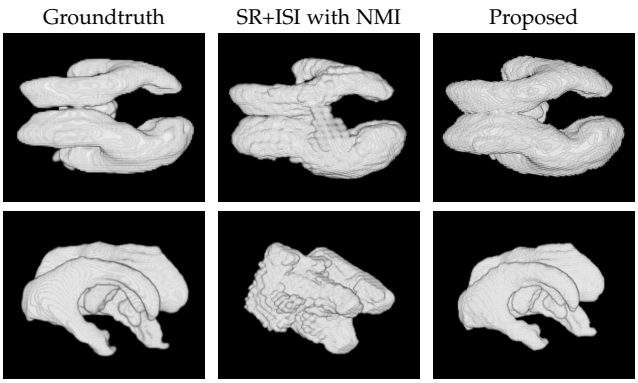


Fig. 7. Shape reconstructed from the two datasets of Fig. 5. Top: dataset 1, view from the top. Bottom: dataset 2, view from a front-right position.

4.2.2 Complete framework accuracy

Table 6 shows the Jaccard coefficients averaged on all datasets for each test and both methods.

Stack-wise registration – In the first test with small initial misalignments, both the SR+ISI method, with NCC and NMI, and the proposed method recover the shape of the brain model satisfactorily, and yield similar mean Jaccard coefficients (first row of Table 6). SR+ISI with SSD performs less well due to its poorer registration. For larger initial misalignments (second row of Table 6), the SR [4] method sometimes falls into local minima, regardless of the similarity measure, due to its use of local shape information only, as was illustrated in Fig. 5 and discussed in Section 4.2.1. In such cases, the resulting reconstructed shapes, shown in the second column of Fig. 7, suffer from the contradictory information provided by the misaligned sequences, thus significantly decreasing the mean Jaccard coefficients. On the contrary, the use of global shape information by the proposed registration method makes it more robust to such local minima and allows the framework to recover satisfactory shapes (third column of Fig. 7) and to maintain the same accuracy.

Slice-wise registration – The proposed framework recovers better shapes than SR+ISI in all four tests, due to its more accurate registration (4 last rows of Table 6). In particular, as seen in the last 4 rows of Table 6, it is not impaired by larger initial misalignments in tests 2 and 4 and yields mean Jaccard coefficients similar to those of tests 1 and 3 with smaller misalignments. In addition, in tests 3 and 4, the more significant number of degrees of freedom (3 translations and 3 rotations) makes the registration problem less constrained than in tests 1 and 2, and causes some slices to have several acceptable positions, as was discussed in Section 4.2.1 and illustrated in Fig. 6a. However, the proposed framework recovers quite accurate shapes even in such cases, as shown on the right of Fig. 8. Indeed, it manages to place the slices in positions which make them contribute to the overall shape of the object while limiting any conflicts with other slices. The SR+ISI method however, attempts to align the slices without taking into account the global shape of the object, and therefore obtains poor registration and shape recovery in such cases, as illustrated in the middle columns of Fig. 8.

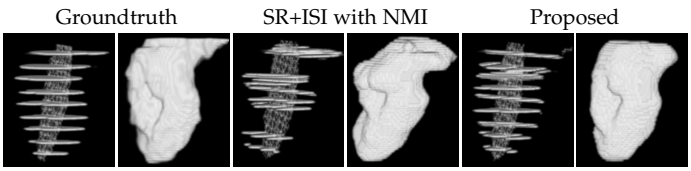


Fig. 8. Shape reconstruction from a dataset containing slices with ambiguous poses. Left: slices after registration. Note that only one LA (vertical) slice is available, hence the numerous ambiguities in the pose of the SA slices. Right: reconstructed 3D shape.

4.3 Quantitative Evaluation on Artificial Datasets Made up of Real Images

In this experiment we show that our proposed framework is not limited to images consisting of piecewise constant regions, but can also deal with real images from a variety of modalities that may exhibit large variations in appearance and various levels of noise.

4.3.1 Registration accuracy

Stack-wise registration – We perform two tests, with a small and larger range of initial misalignments respectively. In the first test, only a small ROI around the object of interest is used by both methods in order to avoid possible local minima caused by background objects and thus test the accuracy of both methods. In the second test, the entire images are used in order to evaluate the robustness to local minima due to cluttered backgrounds. Table 5 presents the average registration errors sorted by slice spacings.

The results show the same trend as the corresponding experiment on artificial data. For small initial misalignments and slice spacings (first 5 rows of Table 5), both methods generally achieve sub-pixel and sub-degree accuracy. However, for larger spacings or misalignments (last 5 rows of Table 5), the SR [4] method falls into local minima, especially when the image contains several objects of similar intensities, as illustrated in Fig. 9a. Therefore, its average accuracy in the second test is significantly reduced. The proposed registration method avoids such issues due to its global approach and yields satisfactory registration for all our datasets except one, which contains an object with a rather symmetrical shape (Fig. 9b). This symmetry causes an ambiguity on the pose of the sequences, similar to that discussed in Section 4.2 and illustrated in Fig. 6a. Thus, the method results in a registration error for this dataset of (1.76, 11.75, 2.16) pixels and (9.25°, 0.23°, 16.23°), but the recovered shape is acceptable, as shown in Fig. 9e. This case increases the overall error of the proposed registration method in test 2 to (0.32, 1.67, 0.58) pixels and (1.3°, 0.2°, 2.2°). If we discount it, the mean error becomes (0.08, 0.34, 0.49) pixels and (0.41°, 0.23°, 0.35°), thus the method can yield a similar accuracy than in the first test with smaller initial misalignments if no local minimum occurs due to shape symmetry.

Slice-wise registration – We initially shift all the slices of a dataset independently in the (x,y)-directions, by a random amount bounded by the slice spacing of the dataset: $T^n_{xy_{\text{ini}}} = \text{rand}\,[-SP, SP]$ with SP the slice spacing, and $T^n_{z\,\text{ini}} = 0$ and $\theta^n_{i\,\text{ini}} = 0°$. Registration results are presented in Table 7. This test is different to the slice-wise registration tests on ar-

TABLE 5
Stack-wise registration accuracy on real data: mean error (in pixels and degrees) averaged over all datasets for each test and slice spacing

Initial misalignment	Slice spacing	SR [4]	Proposed
$T_{i\ \text{ini}}^{n} = \text{rand}[-10, 10]$, $\theta_{i\ \text{ini}}^{n} = \text{rand}[-10°, 10°]$	5 pixels	(0, 0, 0.11) (0.01°, 0.001°, 0.001°)	(0.01, 0.01, 0.12) (0.02°, 0.02°, 0.02°)
	10 pixels	(0, 0, 0.11) (0.002°, 0.001°, 0.001°)	(0.01, 0.06, 0.13) (0.06°, 0.09°, 0.1°)
	15 pixels	(0.001, 0.001, 0.11) (0.01°, 0.003°, 0.001°)	(0.09, 0.05, 0.08) (0.11°, 0.16°, 0.1°)
	20 pixels	(0.11, 0.56, 0.42) (3.33°, 0.47°, 1.95°)	(0.08, 0.28, 0.25) (1.39°, 0.26°, 1.09°)
	Average	**(0.03, 0.14, 0.19) (0.84°, 0.12°, 0.49°)**	**(0.05, 0.1, 0.14) (0.4°, 0.13°, 0.33°)**
$T_{i\ \text{ini}}^{n} = \text{rand}[-30, 30]$, $\theta_{i\ \text{ini}}^{n} = \text{rand}[-20°, 20°]$	5 pixels	(0.001, 0.004, 0.12) (0.02°, 0.004°, 0.003°)	(0.03, 0.03, 0.14) (0.09°, 0.04°, 0.11°)
	10 pixels	(0.001, 0.002, 0.11) (0.01°, 0.001°, 0.003°)	(0.05, 0.07, 0.16) (0.1°, 0.08°, 0.02°)
	15 pixels	(8.66, 12.04, 5.15) (5.39°, 1.14°, 4.07°)	(0.90, 5.88, 1.17) (4.67°, 0.26°, 8.15°)
	20 pixels	(3.08, 7.54, 24.95) (8.07°, 3.02°, 8.04°)	(0.3, 0.69, 0.85) (0.32°, 0.41°, 0.49°)
	Average	**(2.94, 4.9, 7.58) (3.37°, 1.04°, 3.03°)**	**(0.32, 1.67, 0.58) (1.3°, 0.2°, 2.2°)**

TABLE 6
Jaccard coefficients obtained on artificial data

Test	SR+ISI (SSD)	SR+ISI (NCC)	SR+ISI (NMI)	Proposed
Stack-wise 1	0.85±0.25	0.96±0.02	**0.97±0.00**	0.96±0.01
Stack-wise 2	0.60±0.42	**0.83±0.29**	0.82±0.31	**0.96±0.02**
Slice-wise 1	0.87±0.08	0.89±0.06	0.89±0.05	**0.91±0.06**
Slice-wise 2	0.78±0.14	0.82±0.12	0.82±0.12	**0.91±0.05**
Slice-wise 3	0.78±0.08	0.77±0.06	0.77±0.07	**0.85±0.04**
Slice-wise 4	0.68±0.10	0.69±0.09	0.65±0.09	**0.83±0.07**

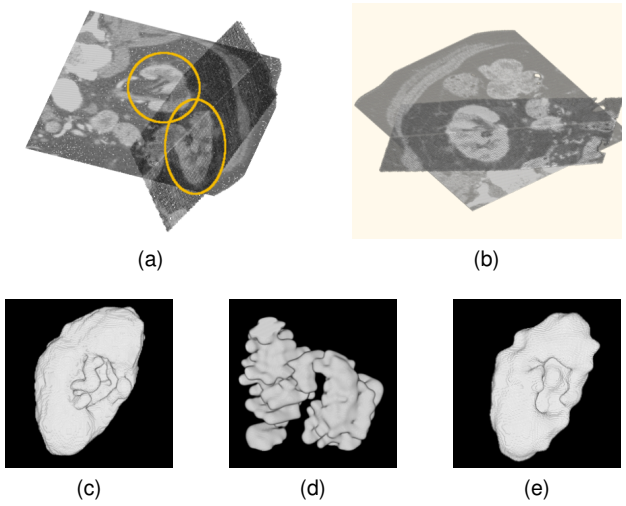


Fig. 9. Registration and shape reconstruction from a kidney CT-scan made up of two spatial sequences with large (15 pixels) gaps, severe initial misalignment, and local minima due to cluttered background. Top: registration result – for visualization purposes, only one central slice of each spatial sequence is displayed. (a) SR [4] and (b) proposed registration methods. Bottom: reconstructed shape – (c) original shape reconstructed from the volume with groundtruth alignment, (d) SR+ISI, (e) proposed framework.

tificial data, since in Section 4.2 the datasets were made up of one stack plus a varying small number of roughly orthogonal slices having arbitrary positions and orientations, whilst here the datasets are composed of two orthogonal stacks of parallel slices. This spatial configuration contains more intersections between the slices and is the configuration for which the method in [4] was originally designed. Therefore, both SR [4] and the proposed registration method perform similarly well for slice spacings and initial misalignments of up to 15 pixels, achieving sub-pixel accuracy. In the last test,

TABLE 7
Slice-wise registration accuracy on real data: mean and std of registration error (in pixels) averaged over all datasets for each test

Slice spacing & shift max	SR [4]	Proposed
5 pixels	(0.34, 0.54) ±(0.22, 0.69)	(0.12, 0.2) ±(0.01, 0.03)
10 pixels	(0.34, 0.92) ±(0.14, 3.58)	(0.17, 0.35) ±(0.02, 0.06)
15 pixels	(0.27, 0.27) ±(0.02, 0.02)	(0.26, 0.77) ±(0.13, 0.71)
20 pixels	(4.22, 3.45) ±(27.84, 27.25)	(0.48, 0.84) ±(0.19, 0.56)
Average	**(1.29, 1.29) ±(7.05, 7.89)**	**(0.26, 0.54) ±(0.09, 0.34)**

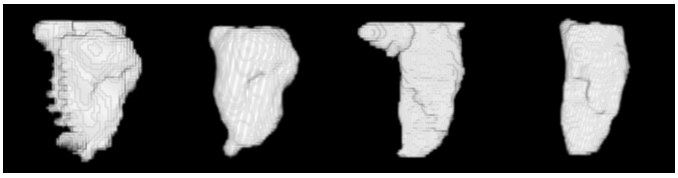
TABLE 8
Jaccard coefficients yielded on real data

Test	SR+ISI	Proposed
Stack-wise 1	0.94±0.05	0.94±0.04
Stack-wise 2	0.72±0.35	0.93±0.05
Slice-wise	0.88±0.13	0.93±0.04

with 20 pixels spacing and maximum misalignments, the SR [4] method encounters local minima problems when objects of the background are confused with the object of interest. At the same time, the number of intersection points becomes less suitable to compute similarity measures. As a result, its average registration error increases dramatically. The proposed registration method does not suffer from these problems and maintains a sub-pixel accuracy.

4.3.2 *Complete framework*

The Jaccard coefficients for both methods, averaged on all datasets for each test, are presented in Table 8. When the registration yielded by the SR+ISI method and the proposed framework are similarly accurate, as in the first stack-wise registration test, both frameworks reconstruct satisfactory shapes with similar Jaccard coefficients. However, when the SR [4] method falls into local minima because it confuses background objects with the object of interest, as in Fig. 9a, the two sequences which make the dataset provide incoherent information to the segmenting level set surface, lowering the Jaccard coefficient to 0.12 in the case of Fig. 9d. The proposed framework does not suffer from this problem due to its use of global shape information. Indeed, as explained in Section 4.3.1, the only local minimum its registration encountered was caused by the symmetry of the object. However, the registration method takes into account the global shape of the object and attempts to reduce the discrepancy in the segmentation information that the images



(a) SR+ISI (b) Proposed (c) SR+ISI (d) Proposed

Fig. 10. Registration, segmentation, and shape interpolation of the LV cavity of the heart from two real 4D cardiac MRIs. Bumps in (a) and (c) are due to misaligned slices. The bump at the bottom of shape (b) is not due to a misaligned slice but is a leak of the segmentation surface which could be avoided by choosing a more robust segmentation algorithm.

provide. Therefore, in such cases, the proposed framework still recovers a plausible shape, as illustrated in Fig. 9e where it obtained a Jaccard coefficient of 0.83. Thus, the mean Jaccard coefficient for the second stack-wise registration test remains as high as for smaller initial misalignments.

Similarly, in the slice-wise registration test, the SR [4] method encountered local minima problems for the largest initial misalignments, and the reconstructions of the full framework suffer from the resulting badly aligned sequences. On the contrary, the greater robustness of the proposed registration method to local minima offered a better support to the shape reconstruction of the full framework, and provides a better average Jaccard measure.

4.4 Qualitative Evaluation on Real Data

The brain and hip datasets have only very small initial misalignments of less than 5 pixels and a few degrees. The registration and 3D reconstruction results were inspected visually, and we subjectively assessed that both methods perform similarly well on these datasets.

The range of initial misalignments of the temporal sequences of the cardiac MRIs is more important, with shifts of 1 to 10 pixels. We found that registration by full translations and rotations produced a similarly good quality of alignment on these data than registration by translation in the SA plane only. This is supposedly due to the sequences being acquired at end-expiration, which is a position where the heart undergoes negligible or no rotation and shift along its long-axis. Therefore, we show in Fig. 10 results obtained using translation in the SA plane only.

The datasets in Fig. 10 suffered from severe initial misalignments of 5 to 10 pixels. The SR+ISI method performs less satisfactorily in such cases and yields even poorer results than with the artificial heart datasets. We explain this by the higher complexity of the data whose background contained several objects which can confuse the SR [4] registration method. On the contrary, the proposed registration method registers the datasets very well in all cases and the complete framework recovers plausible shapes of the LV cavity of the heart.

5 Discussion

We now present a cohesive review of the strengths and limitations of our registration method and its proposed implementation. The main advantage of this method against the co-segmentation via registration based regularization method of Yezzi *et al.* [21] is that it can handle sparse volumes made up of any number of images located in arbitrary planes which contain different parts of the object being segmented. The only requirement on image location is that the images have to intersect each other, although one image does not need to intersect all other images.

A challenge of this more general application field is that the implicit estimation of remaining misalignment, that holds in Yezzi *et al.*'s special case of images containing the same parts of the object, cannot be used in the general case. One can intuitively see that it makes sense to register images by amounts corresponding to their respective misalignment, so that images that are strongly misaligned get strongly moved towards a better alignment, while ones that are well aligned remain undisturbed. Our experiments showed that our general formulation can compute the direction of alignment as well as Yezzi *et al.*'s method. However, the remaining misalignment, and therefore amount of registration to be applied to the images, cannot be computed reliably in cases where the images contain different parts of the object and have very few intersection points. Therefore, we proposed a scheme in Section 3.3 to (a) estimate this remaining misalignment (Eqs. (14) and (15)) and (b) account for it in the amount of registration applied to the individual images (Eqs. (16), (17), (18) and (19)).

There are two main limitations to the scheme we propose in (16), (17), (18) and (19) for setting the amount of registration. Firstly, the amount of registration is approximated by two separate treatments of translation and rotation. In some of our tests, we even performed registration by translation only, sometimes adding rotation after a number of iterations. This approach is quite appropriate to our case of cardiac MRIs acquired during breath-holds, since the initial amounts of misalignment by translation and by rotation are quite unbalanced, with misalignment by translation being predominant. Our method might benefit from a joint treatment of translation and rotation for a simpler and more general application. Furthermore, although we did not encounter gradient descent issues in our tests, a coupled treatment of translation and rotation, such as by quasi-Newton approaches [29], has been suggested as a way to ensure that no such issue arise in other cases.

Secondly, the proposed update scheme in (16) and (17) is very simple and potentially non-optimum, and its parameters were determined empirically. We simply set the amount of registration to be proportional to the estimated misalignment in the general regime, with a saturation at 1 pixel or 1 degree per iteration to avoid risks of divergence of the gradient descent scheme. We also use either a slower linear scheme or an unaltered amount of registration (whichever is smaller) close to perfect alignment in order to allow for finer registration. The rational for potentially using the unaltered amount of registration in that case is that when images are becoming well aligned, towards the end of the full registration and segmentation process, the 3D segmenting contour starts matching the global shape of the object reasonably well, making the implicit estimation of misalignment valid again. The threshold values on the estimated misalignment were determined empirically. It was found that the method is not very sensitive to these values, and they have not been optimised. This piecewise

linear scheme may be replaced by a more elaborate and/or parameter-free one with little impact on the principles of our proposed registration method.

Finally, the optimisation scheme used here is purposefully simple in order to keep the focus on our proposed registration method. We use a simple gradient descent scheme, together with the aforementioned independent processing of translations and rotations, and piece-wise linear scheme for setting the amount of registration. A more elaborate numerical optimisation method could be used to implement our registration strategy and may provide better performance, notably in term of convergence speed.

6 Conclusion and Future Work

We have presented an integrated registration, segmentation, and shape interpolation framework to model objects from 3D and 4D volumes made up of spaced and misaligned slices having arbitrary relative positions. We first proposed a new methodology for rigid registration of images that contain different parts of an object. This novel method combines a global approach for better handling of local minima, and a local approach for a better final accuracy. Second, a new level set framework was also proposed, that integrates this registration method and our previous segmentation and shape interpolation method for handling spaced data. It can exploit segmentation results for better robustness of both registration and shape interpolation to large gaps in the data and to limited number of intersections between the images. The framework was validated on artificial data and tested on real MRI and CT scans. In particular, registration is particularly well suited for handling both local minima and fewer image intersections, and it obtained more robust and accurate results than SR [4] in these cases. It yielded similar results to SR [4] in easier scenarios. The complete framework performed significantly better than the sequential approach of registration followed by segmentation and shape interpolation.

As in [21], our local variant of registration could be used for non-rigid registration if derived from the non-rigid formulation in [21]. However, as in [21] also[6], a regularization method should be chosen with care to avoid arbitrary deformations, and overall our registration method is better suited for rigid registration.

Both registration and shape interpolation rely heavily on segmentation results and this is a limitation of our method. In particular, if a background object is wrongly segmented in an image, then this image cannot be registered properly. Therefore it is important to use reliable segmentation algorithms. The proposed framework can make use of any segmentation scheme, and the choice should be made by the user based on the characteristics of the data. Future work includes investigating robust segmentation algorithms for specific modeling problems.

Acknowledgments

The CT-scans datasets used in the experiments were from the Laboratory of Human Anatomy and Embryology, University of Brussels (ULB), Belgium.

6. see Section 3.4 of [21]

References

[1] A. Paiement, M. Mirmehdi, X. Xie, and M. Hamilton, "Integrated segmentation and interpolation of sparse data," IEEE Trans. Image Processing, vol. 23, pp. 110–125, 2014.

[2] ——, "Simultaneous level set interpolation and segmentation of short and long-axis MRI," in MIUA, 2010, pp. 267–271.

[3] C. Studholme, D. Hill, and D. Hawkes, "An overlap invariant entropy measure of 3D medical image alignment," PR, vol. 32, pp. 71–86, 1999.

[4] J. Lotjonen, M. Pollari, S. Kivisto, and K. Lauerma, "Correction of movement artifacts from 4-D cardiac short- and long-axis MR data," in MICCAI, 2004, pp. 405–412.

[5] P. Slomka, D. Fieno, A. Ramesh, V. Goyal, H. Nishina, L. Thompson, R. Saouaf, D. Berman, and G. Germano, "Patient motion correction for multiplanar, multi-breath-hold cardiac cine MR imaging," Journal of Magnetic Resonance Imaging, vol. 25, no. 5, pp. 965–973, 2007.

[6] A. Elen, F. Maes, J. Bogaert, and P. Suetens, "3D breath hold related motion correction of multi-view 2D functional MRI," in MICCAI Workshop on analysis of functional medical images, 2008, pp. 49–56.

[7] K. Kim, P. Habas, F. Rousseau, O. Glenn, A. Barkovich, and C. Studholme, "Intersection based motion correction of multisclice MRI for 3-D *in* utero fetal brain image formation," IEEE Trans. Medical Imaging, vol. 29, no. 1, pp. 146–158, 2010.

[8] K. Kim, P. Habas, V. Rajagopalan, J. Scott, J. Corbett-Detig, F. Rousseau, A. Barkovich, O. Glenn, and C. Studholme, "Bias field inconsistency correction of motion-scattered multislice MRI for improved 3D image reconstruction," IEEE Trans. Medical Imaging, vol. 30, no. 9, pp. 1704–1712, 2011.

[9] J. Woo, E. Murano, M. Stone, and J. Prince, "Reconstruction of high-resolution tongue volumes from MRI," IEEE Trans. Biomedical Engineering, vol. 59, pp. 3511–3524, 2012.

[10] T. Li and T. Denney, "Registration of short and long-axis images in cine cardiac MRI," in Proc. Intl. Soc. Mag. Reson. Med., vol. 14, 2006, p. 1208.

[11] G. Li, S. Lu, Z. Min, J. Li, X. Xu, and E. Song, "3D prostate boundary reconstruction from 2D TRUS images," in ICBBE, 2007, pp. 940–943.

[12] A. Pihuit, O. Palombi, and M.-P. Cani, "Reconstruction implicite de surfaces 3D à partir de régions 2D dans des plans parallèles," in Journées de l'AFIG, 2009.

[13] X. Yuan and X. Yuan, "Fusion of multi-planar images for improved three-dimensional object reconstruction," Comp. Med. Im. Graph., vol. 35, pp. 373–382, 2011.

[14] L. Cordero-Grande, G. Vegas-Sanchez-Ferrero, P. C. de-la Higuera, and C. Alberola-Lopez, "A Markov Random Field approach for topology-preserving registration: application to object-based tomographic image interpolation," IEEE Trans. Image Processing, vol. 21, 2012.

[15] C. L. Guyader and L. Vese, "A combined segmentation and registration framework with a nonlinear elasticity smoother," CVIU, vol. 115, no. 12, pp. 1689–1709, 2011.

[16] M. Rousson and N. Paragios, "Shape priors for level set representations," in ECCV, 2002, pp. 78–92.

[17] N. Paragios, M. Rousson, and V. Ramesh, "Matching distance functions: A shape-to-area variational approach for global-to-local registration," in ECCV, 2002, pp. 775–789.

[18] C. Xiaohua, M. Brady, and D. Rueckert, "Simultaneous segmentation and registration for medical image," in MICCAI, 2004, pp. 663–670.

[19] Y. Gong and M. Brady, "Texture-based simultaneous registration and segmentation of breast DCE-MRI," in Int. Workshop Digital Mammography, 2008, pp. 174–180.

[20] Y. Sun, M.-P. Jolly, and J. Moura, "Integrated registration of dynamic renal perfusion MR images," in IEEE Symp. Image Processing, 2004, pp. 1923–1926.

[21] A. Yezzi, L. Zollei, and T. Kapur, "A variational framework for integrating segmentation and registration through active contours," Medical Image Analysis, vol. 7, pp. 171–185, 2003.

[22] R. Prevost, R. Cuingnet, B. Mory, J.-M. Correas, L. Cohen, and R. Ardon, Joint co-segmentation and registration of 3D ultrasound images, ser. LNCS, 2013, vol. 7917, pp. 268–279.

[23] G. Unal and G. Slabaugh, "Coupled PDEs for non-rigid registration and segmentation," in CVPR, vol. 1, 2005, pp. 168–175.

[24] H. van Assen, M. Danilouchkine, M. Dirksen, J. Reiber, and B. Lelieveldt, "A 3-D active shape model driven by fuzzy inference: Application to cardiac CT and MR," IEEE Trans. Information Technology in Biomedicine, vol. 12, pp. 595–605, 2008.
[25] S. Zambal, J. Hladuvka, K. Buhler, and A. Neubauer, "A fully automatic system for segmentation and analysis of the left and right ventricles of the heart using a bi-temporal two-component model," in Proc. CARS, 2007, pp. 93–94.
[26] J. Lotjonen, V. Jarvinen, B. Cheong, E. Wu, S. Kivisto, J. Koikkalainen, J. Mattila, H. Kervinen, R. Muthupillai, F. Sheehan, and K. Lauerma, "Evaluation of cardiac biventricular segmentation from multiaxis MRI data: A multicenter study," MRI, vol. 28, pp. 626–636, 2008.
[27] S. Zambal, J. Hladuvka, and K. Buhler, "Improving segmentation of the left ventricle using a two-component statistical model," in MICCAI, 2006, pp. 151–158.
[28] T. Chan and L. Vese, "Active contours without edges," IEEE Trans. Image Processing, vol. 10, pp. 266–277, 2001.
[29] J. Nocedal and S. J. Wright, Numerical Optimization, 2nd ed. Springer, 2006, ch. 6.

Adeline Paiement received the diplome d'ingénieur (M.Sc.) from Télécom Physique Strasbourg, and the M.Sc. in Physics (with commendation) from Université Louis Pasteur, Strasbourg, France, in 2008. She completed a Ph.D. degree (with commendation) in Computer Vision in 2013 at the University of Bristol, where she is now a postdoctoral researcher. Her research interests include the analysis of 3D/4D sparse data and level set methods.

Majid Mirmehdi received the BSc (Hons) and PhD degrees in Computer Science in 1985 and 1991 respectively, from the City University, London. He is currently a Professor of Computer Vision in the Department of Computer Science at the University of Bristol, and the Graduate Dean and Head of the Graduate School of Engineering. His research interests include natural scene analysis and medical imaging, and he has more than 150 refereed conference and journal publications in these and other areas. He is a Fellow of the International Association for Pattern Recognition. He is Editor-in-Chief of IET Computer Vision journal and an Associate Editor of the Pattern Analysis and Applications journal. He is a member of the IET, Senior Member of IEEE, and serves on the Executive Committee of the British Machine Vision Association.

Xianghua Xie received the M.Sc. (with commendation) and Ph.D. degrees in computer science from the University of Bristol, Bristol, U.K., in 2002 and 2006, respectively. He is currently an Associate Professor with the Department of Computer Science, Swansea University, Swansea, U.K. He was previously a Research Associate with the Department of Computer Science, University of Bristol between 2006 and 2007. His current research interests are medical imaging, video analysis, texture analysis, image segmentation, surface inspection, deformable models, and human pose estimation and tracking. He is a member of BMVA and a senior member of IEEE.

Mark Hamilton received his medical degree from the University of Dundee in 1991, Membership of the Royal College of Physicians (1995) and Fellowship of the Royal College of Radiologists (2000). He was cardiac radiology fellow in Auckland, New Zealand 2003-2005. Since then he has been consultant cardiac radiologist at United Hospitals Bristol, and clinical lead in cardiac magnetic resonance imaging. His research collaborates with the Bristol Cardiovascular NIHR Biomedical Research Unit and the Department of Computer Science, Bristol University. He has research interests in computed tomography and cardiac magnetic resonance imaging.