Estimating nonrigid motion from inconsistent intensity with robust shape features

Wenyang Liu

Department of Bioengineering, University of California, Los Angeles, California 90095

Dan Ruan^{a)}

Department of Bioengineering, University of California, Los Angeles, California 90095; Department of Radiation Oncology, University of California, Los Angeles, California 90095; and Department of Biomedical Physics, University of California, Los Angeles, California 90095

(Received 24 July 2013; revised 18 October 2013; accepted for publication 24 October 2013; published 22 November 2013)

Purpose: To develop a nonrigid motion estimation method that is robust to heterogeneous intensity inconsistencies amongst the image pairs or image sequence.

Methods: Intensity and contrast variations, as in dynamic contrast enhanced magnetic resonance imaging, present a considerable challenge to registration methods based on general discrepancy metrics. In this study, the authors propose and validate a novel method that is robust to such variations by utilizing shape features. The geometry of interest (GOI) is represented with a flexible zero level set, segmented via well-behaved regularized optimization. The optimization energy drives the zero level set to high image gradient regions, and regularizes it with area and curvature priors. The resulting shape exhibits high consistency even in the presence of intensity or contrast variations. Subsequently, a multiscale nonrigid registration is performed to seek a regular deformation field that minimizes shape discrepancy in the vicinity of GOIs.

Results: To establish the working principle, realistic 2D and 3D images were subject to simulated nonrigid motion and synthetic intensity variations, so as to enable quantitative evaluation of registration performance. The proposed method was benchmarked against three alternative registration approaches, specifically, optical flow, B-spline based mutual information, and multimodality demons. When intensity consistency was satisfied, all methods had comparable registration accuracy for the GOIs. When intensities among registration pairs were inconsistent, however, the proposed method yielded pronounced improvement in registration accuracy, with an approximate fivefold reduction in mean absolute error (MAE = 2.25 mm, SD = 0.98 mm), compared to optical flow (MAE = 9.23 mm, SD = 5.36 mm), B-spline based mutual information (MAE = 9.57 mm, SD = 8.74 mm) and multimodality demons (MAE = 10.07 mm, SD = 4.03 mm). Applying the proposed method on a real MR image sequence also provided qualitatively appealing results, demonstrating good feasibility and applicability of the proposed method.

Conclusions: The authors have developed a novel method to estimate the nonrigid motion of GOIs in the presence of spatial intensity and contrast variations, taking advantage of robust shape features. Quantitative analysis and qualitative evaluation demonstrated good promise of the proposed method. Further clinical assessment and validation is being performed. © 2013 American Association of Physicists in Medicine. [<http://dx.doi.org/10.1118/1.4829507>]

Key words: motion estimation, DCE-MRI, nonrigid registration

1. INTRODUCTION

Image-based motion estimation plays an important role in various biomedical context, such as dynamic imaging and MR based radiation therapy. The problem of image registration, which estimates the deformation based on image information, either between pairs of frames or a sequence, has been intensively studied.^{1–5} To register images, a certain type of consistency condition needs to be satisfied, typically imposed on the intensity values themselves or the intensity transformations among images, leading to the development of a large body of works on minimal squared error and/or maximum mutual information.^{2,6–9} These approaches, unfortunately, are not directly applicable when the

heterogeneous contrast variations are present between image pairs or among temporal sequences, as in the imaging processes involving contrast enhancement. In this work, we focus on dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI), where a series of MR images are acquired before and after an intravenous injection of gadolinium.¹⁰ To assess renal function using DCE-MRI, pharmacokinetic models are fitted to the voxelwise enhancement curves to map the renal blood flow and obtain glomerular filtration rate.¹¹ Kidney motion caused by breathing¹² induces artifacts and inaccuracy in the voxel-based analysis^{13–16} and need to be corrected via image registration. On the other hand, the kinetics of the contrast agent leads to temporally inconsistent intensity, which makes estimating and

subsequently correcting for the physiological and physical motion challenging.

A relatively straightforward way to perform DCE-MR registration is to consider the contrast-variant images as *effectively* different modalities and apply mutual information as the matching metric.¹⁷ However, the spatially variant intensity changes in DCE-MR does not guarantee such mapping consistency, as maximized in mutual information based methods, and may lead to misalignment.^{7,18} More recently, efforts have been made to modify generic registration algorithms to address the difficulties caused by intensity and contrast inconsistencies. Merrem *et al.*¹⁹ modifies a nonrigid registration method that uses the squared cross-correlation as matching metric, by introducing an isotropic Gaussian kernel to smooth the estimated deformation field at each step, to reduce noise sensitivity and alleviate the effect of spatially variant enhancement pattern. This approach is sensitive to the kernel size, and risks misregistration if not selected appropriately. A modified demons algorithm has also been proposed for DCE-MR registration,²⁰ yet it only stabilizes the original demon for low gradient regions by adding a small modifier to the denominator but does not improve sensitivity to contrast change. Haber and Modersitzki²¹ have proposed a normalized gradient method by using the direction of the gradient as registration features. While the gradient direction is robust to contrast changes, it is sensitive to spatially varying noise which is prevalent in practice. It also faces stability issue when certain edges appear in only one of the images under registration. In addition, the efficacy of the above methods were largely assessed visually with only qualitative conclusions.

In this study, we propose and evaluate a novel method to estimate the motion of roughly persistent GOIs in the presence of intensity variations. We propose to use full shape contour as registration features and benefit from their robustness to noise and intensity changes. To quantitatively assess the registration performance, 2D and 3D nonrigid motion were simulated and used as ground-truth for evaluation purposes. We also assessed the performance of the proposed method on real MR series, yielding promising results.

2. METHOD

Traditional motion estimation/registration methods are often based on the assumption that voxel intensity is temporally invariant or statistically dependent (as in mutual information based method). Those assumptions generally do not hold in the presence of local contrast and intensity changes as in DCE-MRI. Therefore, features that are robust with respect to intensity and contrast variations are preferable. In this study, we propose to utilize full geometric contours as registration features. Depending on gradient in a more regularized fashion, the level set representation of the geometric contour is more robust to local perturbations caused by noise and spatially varying contrast changes.

The proposed method consists of two main steps: (1) extracting GOIs with a level set based segmentation, (2) estimating nonrigid deformation on the full domain by minimizing discrepancy of the corresponding level set functions

(LSFs) in the neighborhood of the GOIs subject to regularity conditions.

2.A. Level set based segmentation

The level set method Osher and Sethian *et al.*^{22,26} is widely used for representing evolving curves and surfaces, especially for segmentation tasks. It has the advantage of being flexible in shape representation and numerically efficient with fixed-grid calculation. In this representation, a curve C is represented implicitly as the zero level set of its corresponding LSF ϕ in a space that is one dimension higher. The LSF is often enforced to have unity gradient magnitude ($|\nabla\phi| = 1$) to represent distance to the curve of interest for numerical stability. We use the level-set based method to segment the GOIs and extract shape contours represented by the LSFs.²³ At each image frame, we perform segmentation by minimizing a variational energy that encodes (i) alignment to high image gradient regions, (ii) curve smoothness, (iii) area/volume preservation between adjacent frames, and (iv) signed distance property of the LSF.

Specifically, for a LSF $\phi : \Omega \subset \mathbb{R}^d \rightarrow \mathbb{R}$ ($d = 2$ or 3), the energy functional is defined as

$$\epsilon(\phi) = \lambda \mathcal{L}_g(\phi) + \frac{\beta}{2} \mathcal{A}(\phi) + \mu \mathcal{R}_p(\phi), \quad (1)$$

where the weights λ , β , and μ balance the various contributing terms.

Conformation of the boundary to high image gradient area and curve smoothness is encouraged by $\mathcal{L}_g(\phi)$,²³ defined as

$$\mathcal{L}_g(\phi) = \int_{\Omega} g \delta(\phi) |\nabla\phi| d\mathbf{x}. \quad (2)$$

Here δ is the Dirac Delta function, and g is an edge indicator function that takes small values at locations with large gradient:

$$g = \frac{1}{1 + |\nabla G_{\sigma} * I|^2}, \quad (3)$$

where I is the input image and G_{σ} is a Gaussian kernel with standard deviation σ . The kernel parameter σ balances robustness to noise and preservation of image gradient information.

Unlike conventional level set based segmentation methods that often push the contour to either shrink or expand, we design an area regularization component $\mathcal{A}(\phi)$ to enforce consistency of area/volume across frames

$$\mathcal{A}(\phi) = \left(\left(\int_{\Omega} H(-\phi) d\mathbf{x} \right) - \left(\int_{\Omega} H(-\phi_0) d\mathbf{x} \right) \right)^2, \quad (4)$$

where H is the Heavyside function and ϕ_0 is the LSF of the previous image frame. This is motivated by the observation that areas and volumes are expected to vary little between adjacent frames.

Distance regularization is achieved with $\mathcal{R}_p(\phi)$,²⁴ which enforces the signed distance property of the LSF:

$$\mathcal{R}_p(\phi) = \int_{\Omega} p(|\nabla\phi|) d\mathbf{x}, \quad (5)$$

where p is a double-well function defined as

$$p(s) = \begin{cases} \frac{1}{(2\pi)^2}(1 - \cos(2\pi s)), & \text{if } s \leq 1 \\ \frac{1}{2}(s - 1)^2, & \text{if } s \geq 1 \end{cases}. \quad (6)$$

The double-well function has two minimums at 0 and 1, respectively, which preserves the signed distance property ($|\nabla\phi| = 1$) near the zero level set and smoothes ($|\nabla\phi| = 0$) the locations that far away from the zero level set. It has been shown that Eq. (5) preserves the signed distance property more efficiently and continuously than the traditional reinitialization methods.^{22,25,26}

Numerically, smoothed approximations of $H(\phi)$ and $\delta(\phi)$ are used²²

$$H_\epsilon(\phi) = \begin{cases} 0, & \phi < -\epsilon \\ \frac{1}{2} + \frac{\phi}{2\epsilon} + \frac{1}{2\pi} \sin\left(\frac{\pi\phi}{\epsilon}\right), & -\epsilon \leq \phi \leq \epsilon \\ 1, & \phi > \epsilon \end{cases}, \quad (7)$$

$$\delta_\epsilon(\phi) = \begin{cases} 0, & \phi < -\epsilon \\ \frac{1}{2\epsilon} + \frac{1}{2\epsilon} \cos\left(\frac{\pi\phi}{\epsilon}\right), & -\epsilon \leq \phi \leq \epsilon \\ 1, & \phi > \epsilon \end{cases}. \quad (8)$$

By taking Gâteaux derivative, energy functional [Eq. (1)] can be minimized by the gradient flow in Eq. (9).

$$\begin{aligned} \phi_t = \lambda \delta_\epsilon(\phi) \nabla \cdot \left(g \frac{\nabla\phi}{|\nabla\phi|} \right) + \beta \delta_\epsilon(\phi) \left(\int_{\Omega} H_\epsilon(-\phi) d\mathbf{x} \right. \\ \left. - \int_{\Omega} H_\epsilon(-\phi_0) d\mathbf{x} \right) + \mu \nabla \cdot (d_p(|\nabla\phi|) \nabla\phi). \end{aligned} \quad (9)$$

2.B. Motion estimation

After segmenting the GOIs from each frame, one has access to ϕ_i —the LSF for frame i , which we use as shape features. To estimate the underlying motion field, we aim for similarity upon motion compensation on the LSFs between adjacent frames. We proceed to perform nonrigid matching to

minimize the following functional:

$$E(\omega) = \int_{\Omega_{nb}} (\phi_1(\mathbf{x}) - \phi_2(\mathbf{x} + \omega))^2 d\mathbf{x} + \alpha \int_{\Omega} \|\nabla\omega\|_F^2 d\mathbf{x}, \quad (10)$$

where ω is the motion field, ϕ_1 and ϕ_2 are the LSFs of the temporally adjacent frames, and α is the regularization weight. The data fidelity term is defined over a narrow band $\Omega_{nb} = \{\mathbf{x} : |\phi_1(\mathbf{x})| < \eta\} \cup \{\mathbf{x} : |\phi_2(\mathbf{x})| < \eta\}$, where 2η is the width of the detected organ boundary centered at the zero level set. The smoothness regularization effectively extends the boundary's motion to the rest of the domain. Minimization of the above functional yields the following Euler-Lagrange equations under Neumann boundary conditions:

$$\begin{aligned} (\phi_1(\mathbf{x}) - \phi_2(\mathbf{x} + \omega)) \nabla_i \phi_2(\mathbf{x} + \omega) + \alpha \nabla \cdot (\nabla \omega_i) = 0, \\ \text{for } i = 1, 2, \dots, d. \end{aligned} \quad (11)$$

The notation $\nabla_i \phi_2$ represents the partial derivatives of ϕ_2 along the i th direction, and ω_i is the i th vector component of ω . To solve this nonlinear system of equations, we use first-order Taylor expansion to approximate $\phi_2(\mathbf{x} + \omega)$ at each iteration:

$$\begin{aligned} \phi_2(\mathbf{x} + \omega^{n+1}) \approx \phi_2(\mathbf{x} + \omega^n) + \nabla \phi_2(\mathbf{x} + \omega^n) \\ \cdot (\omega^{n+1} - \omega^n) \end{aligned} \quad (12)$$

with $\omega^0 = 0$ as the initialization. At the $(n+1)$ th iteration, plugging Eq. (12) into Eq. (11) reduces it to the following linear system of equations for ω^{n+1} .

$$\begin{aligned} (\phi_1 - \phi_2^n - \nabla \phi_2^n \cdot (\omega^{n+1} - \omega^n)) \nabla_i \phi_2^n + \alpha \nabla \cdot (\nabla \omega_i^{n+1}) = 0, \\ \text{for } i = 1, 2, \dots, d, \end{aligned} \quad (13)$$

where $\phi_2^n = \phi_2(\mathbf{x} + \omega^n)$, $\nabla_i \phi_2^n = \nabla_i \phi_2(\mathbf{x} + \omega^n)$. This linear system of equations can be solved efficiently with various numerical methods, such as Gauss-Seidel or successive over-relaxation (SOR)²⁷ iterations.

Inspired by Horn and Shunck²⁸ we introduce an inner iteration scheme to solve Eq. (13). At the $(m+1)$ th inner iteration, the divergence $\nabla \cdot (\nabla \omega_i^{n,m+1})$ is approximated with $(\omega_i^{n,m+1} - \omega_i^{n,m})$, where the $\omega_i^{n,m+1}$ is the local weighted average of $(k^d - 1)$ neighbors for a support of k in d dimensions. This approximation leads to the fixed-point iteration scheme in Eq. (14), where n indexes the outer linear approximation iteration, and m indexes the inner fixed point iteration.

$$\omega_i^{n+1,m+1} = \frac{\nabla_i \phi_2^n [\phi_1^n - \phi_2^n + \nabla_i \phi_2^n \omega_i^n - \sum_{j=1, j \neq i}^d \phi_{2,j} (\omega_j^{n,m} - \omega_j^n)] + \alpha \overline{\omega_i^{n,m}}}{(\nabla_i \phi_2^n)^2 + \alpha}, \quad \text{for } i = 1, 2, \dots, d \quad (14)$$

We use multiscale scheme to further improve efficacy and accuracy of the proposed segmentation and motion estimation algorithm. Specifically, we created a coarse to fine image structure. By using this pyramidal structure, we are

able to perform the segmentation and motion estimation at coarser image levels and to refine the results at finer levels. Algorithm 1 provides an overview of the proposed algorithm with the multiscale implementation.

ALGORITHM I. Multiscale scheme for proposed algorithm

```

1: Input:  $I_1(\mathbf{x})$ ,  $I_2(\mathbf{x})$ ,  $\alpha$ ,  $N_{\text{scales}}$ 
2: for  $s \leftarrow N_{\text{scales}} - 1$  to 0 do
3:   Downscale images  $I_1(\mathbf{x})$ ,  $I_2(\mathbf{x})$  by  $\eta = 2^s$ 
4:   Segmentation based on Eq. (1), obtain  $\phi_1^s(\mathbf{x})$ ,  $\phi_2^s(\mathbf{x})$ 
5:   Motion estimation based on Eq. (10), obtain  $\omega^s(\mathbf{x})$ 
6:   if  $s > 0$  then
7:      $\omega^{s-1}(\mathbf{x}') = \sum \eta \omega^s(\mathbf{x}) B(\mathbf{x}'/\eta - \mathbf{x})$ 
8:      $\phi_i^{s-1}(\mathbf{x}') = \sum \eta \phi_i^s(\mathbf{x}) B(\mathbf{x}'/\eta - \mathbf{x})$ ,
9:     where  $B(\mathbf{t})$  is the bicubic interpolation kernel.
10:  end if
11: end for

```

3. EXPERIMENTS AND RESULTS

To quantitatively assess and validate the performance of the proposed method, we generated artificial but physically reasonable nonrigid motion and applied such to a reference image to simulate the physiological deformation. The baseline performance is first evaluated when intensity is consistent. Inconsistency is then imposed to test the efficacy and robustness of the proposed method. The proposed method was benchmarked against three alternative image registration methods: optical flow,²⁸ B-spline based mutual information^{2,30} D. Rueckert *et al.*,³¹ and multimodality demons methods.²⁹ After that, qualitative test was performed using real MR image sequence.

3.A. Brief overview of benchmark methods

3.A.1. Optical flow

The optical flow method, first introduced by Horn and Shunck,²⁸ aims to estimate a smooth motion field ω that matches intensity values between reference I_1 and target images I_2 , by minimizing the energy $\int (I_1(\mathbf{x}) - I_2(\mathbf{x} + \omega))^2 + \alpha_{OF} \|\nabla \omega\|^2 d\mathbf{x}$, where α_{OF} is the smoothness coefficient. In this study, we implemented the optical flow method in MATLAB and set $\alpha_{OF} = 0.5$.

3.A.2. B-spline based mutual information

The mutual information based method^{2,30} maximizes the statistical dependency between two images. It is a widely used approach for multimodality registrations. Our benchmark implementation was based on a multiscale B-spline parameterized formulation for nonrigid motion estimation³¹ from Medical Image Registration Toolbox (MIRT). We used 64 histogram bins and 10×10 initial control point meshes, which is subsequently refined. In minimizing $-\text{MI}(I_1, I_2 \circ T) + \alpha_{\text{MI}} \frac{1}{V} \int \|\nabla T\|^2 d\mathbf{x}$, the regularization weight α_{MI} was set to 1.

3.A.3. Multimodality demons

The multimodality demons²⁹ is a combination of demons and mutual information methods. It transforms one image representation to another by maximizing mutual information, and

then perform demonslike registration while regularizing the smoothness of the motion field. We used the multimodality implementation of Ref. 29 with the default parameters.

For all the implementations, three hierarchical image levels were used. We used a brute force parameter searching scheme to find the optimal parameters near the suggested values.^{28,29,31}

3.B. Parameter setup and computation complexity

In the segmentation step, one needs to set three parameters λ , β , μ to balance the contribution from edge driving force, area and distance property conservation, respectively, and the time step Δt for curve evolution. Our experiments indicate that the proposed method is insensitive to the choices of λ in the range of (3,30) and μ when $\mu \Delta t < 0.25$ is satisfied, and we used fixed values in all our validation tests. The area weight β , on the other hand, depends on motion amplitude: when motion is small between adjacent frames, β should be relaxed to be zero; for large temporal motions, larger value needs to be used. For all experiments in this study, we fixed the parameters to be $\lambda = 10$, $\beta = 0.01$, $\mu = 0.2$, $\Delta t = 1$.

In the motion estimation step, one needs to set the parameter α to weigh the Tikhonov regularization term, and the width η for the narrowband region. Based on our tests, the model is insensitive to the choice of α ranging from 2 to 5 and η ranging from 11 to 21 voxels. In our implementation, we fixed α to be 2.5 and the width of narrow band η to be 15 voxels for all experiments.

The average running time for segmenting 2D image with size 256×256 was about 5 s, and 60 s for segmenting 3D image volume with size $320 \times 320 \times 32$ when the previous frame/volume segmentation was used as initialization. The computation time for estimating motion between 2D images of size 256×256 was about 30 s, and 120 s for comparing 3D volumes with size $320 \times 320 \times 32$. The algorithm was implemented with MATLAB 2012b on Mac OS 10.84, running on a macbook pro laptop with quad-core i7 2.3 GHz, 8 GB RAM. No specific code optimization schemes were used.

3.C. Experiments on simulated nonrigid 2D motion

We start with a 2D EPI MR image of the kidney with $\text{TR} = 137$ ms, $\text{TE} = 44$ ms, $\text{FOV} = 390 \times 390$ mm² and voxel size of 1.52×1.52 mm². The kidneys under quiet respiration have predominant motion in the superior-inferior (SI) direction with a mean magnitude of 19 mm.³² To mimic this behavior, we simulated a motion field parameterized by B-splines,³³ with an average motion magnitude of 15 mm in SI direction and 8 mm in left-right (LR) direction around the kidneys in the abdominal MR images, as depicted in Fig. 1. Figure 2 shows the segmentation results for reference and target images, and the corresponding LSF representations. Figure 3 illustrates the registration performance. Figures 3(a) and 3(b) compares the preregistration and postregistration agreement of the reference and target images, using a checkerboard presentation which alternately displays the intensity squares from either image. Figures 3(c) and 3(d) report the absolute

FIG. 1. Simulated nonrigid motion field with B-spline parameterization.

intensity difference map to quantify the registration accuracy. The checkerboard in Fig. 3(b) shows significantly improved continuity on kidney boundaries compared to Fig. 3(a), while

Fig. 3(d) demonstrates a large reduction of intensity discrepancies compared to Fig. 3(c), demonstrating the efficacy of the proposed registration method. Table I reports the error statistics in terms of mean absolute error ($MAE = \frac{1}{|\Omega|} \sum_i |e_i|$), root means squared error ($RMSE = \sqrt{\frac{1}{|\Omega|} \sum_i e_i^2}$), standard deviation (SD), mean, median, and max error in the neighborhood of kidneys, showing voxel-size comparable registration accuracy. The directional error is evaluated by taking the difference between the simulated ground-truth and the estimated deformation field $e_i(\mathbf{x}) = \omega_i(\mathbf{x}) - \hat{\omega}_i(\mathbf{x})$ for voxels $\mathbf{x}$ in the proper neighborhood of the GOI. The aggregated $nD(n = 2$ or 3) error is measured by $e(\mathbf{x}) = \sqrt{\sum_{i=1}^n (\omega_i(\mathbf{x}) - \hat{\omega}_i(\mathbf{x}))^2}$.

3.D. Experiments on 3D dataset

We conducted similar experiment on 3D MR datasets. Figure 4 illustrates the sample slices of T2 weighted MR images, acquired with $TR = 700$ ms, $TE = 71$ ms, $FOV = 400 \times 400$ mm², array size of 320×320 , slice-thickness of 6 mm, and 32 slices in total in the sequence. Figure 5(a)

FIG. 2. 2D level set based segmentation results and extracted GOI features from target and reference images: (a) segmented contour with zero level set in agreement with reference kidney boundaries, (b) segmented contour with zero level set agrees with target kidney boundaries, (c) GOIs of the reference image represented by a narrowband LSF, (d) GOIs of the target image represented by a narrowband LSF. The segmented contours of target and reference images are overlaid in (a) (b).

FIG. 3. 2D motion estimation results in checkerboard presentation and absolute intensity difference maps: (a) checkerboard before registration, (b) checkerboard after registration, showing pronounced improvement of boundary continuity, (c) intensity difference map before registration, (d) intensity difference map after registration, showing significantly improved agreement between the target kidney boundary and reference kidney as well as reduced intensity difference.

presents the segmentation results of left kidney and the simulated nonrigid motion. Figures 5(b) and 5(c) compares the kidney positions before and after registration, where improved alignment from registration can be appreciated. Error statistics are summarized in Table II, showing subvoxel accuracy.

3.E. Robustness test on images with synthetic intensity and contrast variations

We tested the robustness of the proposed method with the presence of synthetic intensity and contrast changes. Figure 6 shows the simulation flow for contrast and inten-

sity variations. The same deformation field as in Sec. 3.C was used. The proposed method was applied to the reference and target images in Fig. 6 and compared against alternative registration methods using optical flow, B-spline based mutual information, and multimodality demons.^{28,31,29} The quantitative error comparisons are reported in Fig. 7. The proposed method may perform slightly inferiorly when consistency is present, possibly due to limited information captured by the contour. However, when intensity is inconsistent, it enjoyed the pronounced benefit of being robust. The checkerboard representations of difference methods are presented in Fig. 8, showing significant improvements of boundary continuities from the proposed method compared to the others.

3.F. Experiments on real MR image sequences

Finally, we tested the proposed method on real time EPI coronal MR image sequence with $TR = 137$ ms, $TE = 44$ ms, and $FOV = 390 \times 390$ mm². Figure 9(a) presents the estimated motion field using the 1st and 5th image frame in the series. Figures 9(b) and 9(c) provide checkerboard illustrations showing registration significantly improves

TABLE I. Error statistics of 2D experiments on simulated motion showing estimation errors comparable to voxel size (1.52 mm $\times$ 1.52 mm).

	MAE	RMSE	SD	Mean	Median	Max
LR	1.63	1.92	1.92	0.04	-0.08	4.03
SI	1.26	1.53	1.42	0.57	0.81	3.45
2D	2.25	2.45	0.99	2.25	2.24	5.07

FIG. 4. Sample slices of reference 3D MR data: T2 weighted MR images with $TR = 700\text{ms}$, $TE = 71\text{ ms}$, $FOV = 400 \times 400\text{ mm}^2$, array size of 320×320 and slice-thickness of 6 mm .

FIG. 5. 3D segmentation and motion estimation results: (a) left kidney segmentation result overlaid with simulated nonrigid motion field, (b) overlays for reference kidney and target kidney before registration, (c) overlays of reference kidney and target kidney after registration, showing improved alignment.

TABLE II. Error statistics of 3D experiments of simulated nonrigid motion, showing sub-voxel accuracy (voxel size = 0.8 mm × 0.8 mm × 6 mm).

	MAE	RMSE	SD	Mean	Median	Max
LR	0.38	0.50	0.38	0.33	0.25	1.08
SI	0.41	0.55	0.53	0.14	0.07	2.03
AP	0.75	0.99	0.96	−0.26	−0.17	2.80
3D	1.09	1.24	0.58	1.09	1.04	2.88

continuity of contours across the alternating squares, demonstrating the validity of proposed registration. Figures 10(a) and 10(b) presents the mean boundary motion from applying the proposed method to the first 80 frames (with a duration of approximately 10.2 s) of the MR sequence. It exhibits an approximate two-cycle behavior, which agrees with visual inspection of the image sequence and the general knowledge about typical respiratory with 4–6 s duration per cycle. Figure 10(c) presents the estimated temporal motion trajectory of the renal boundary from the first 10 frames, showing the algorithm's ability to capture local deformations.

3.G. Test of sensitivity and robustness towards optimization parameter selection noise levels

Finally, we tested the robustness of the proposed registration method under various optimization parameter settings

and noise conditions. Figure 11 reports the results of performing the registration algorithm with different optimization parameters to balance the terms in Eq. (1). Within a reasonably large range, the segmentation results vary little.

To test the sensitivity to image quality, we have simulated Gaussian noise with standard deviation one tenth of the image intensity range, reducing the power SNR to −1.63dB. A spatially heterogeneous intensity modulation pattern was further imposed to simulate the contrast variation. Figure 12 shows that segmentation results only degrade slightly with the addition of this severe noise.

4. DISCUSSIONS

In this paper, we have proposed and validated a novel motion estimation method based on geometric contour features that are robust to intensity and contrast variations. When image intensity and contrast are consistent, the proposed method performed slightly inferior (MAE = 2.25 mm) to the benchmark methods using optical flow (MAE = 1.23 mm), maximal mutual information (MAE = 2.88 mm), and multi-modality demons (MAE = 1.13 mm), but still comparable to voxel size. This is likely caused by the fact that the proposed method only relies on limited amount of information around the extracted contours. When synthetic contrast and intensity variations were introduced, the proposed method showed

Step 1. Changing image contrast with the gamma function

Step 2. Introducing intensity inhomogeneity by multiplying with the spatially varying function

Step 3. Introducing non-rigid motion by deforming with the simulated motion field

FIG. 6. Simulation workflow for adding contrast and intensity variations and nonrigid deformation to the reference image: Step 1. Changing image contrast by applying the gamma function to the reference image: $I_1(\mathbf{x}) = I_0(\mathbf{x})^\gamma$ with $\gamma = 0.6$; Step 2. Introducing spatially varying intensity inhomogeneity to $I_1(\mathbf{x})$ by multiplying it with spatially varying function $M(\mathbf{x})$: $I_2(\mathbf{x}) = I_1(\mathbf{x}) * M(\mathbf{x})$; Step 3. Introducing nonrigid motion effect to $I_2(\mathbf{x})$ by deforming it with the simulated motion field $T(\mathbf{x})$: $I_3(\mathbf{x}) = I_2(\mathbf{x}) \circ T(\mathbf{x})$.

FIG. 7. Performance comparison with benchmark alternatives in terms of mean absolute error (MAE) and standard deviation (SD): when there is no in intensity variations, the proposed method shows slightly inferior result; in the presence of intensity changes, it shows much improved robustness.

robust estimation results (MAE = 2.25 mm) and outperformed benchmark registration methods using optical flow (MAE = 9.23 mm), maximal mutual information (MAE = 9.57 mm), and multimodality demons (MAE = 10.07 mm).

When applied to clinical image sequence where ground-truth motion is inaccessible, the proposed method generated visually appealing results that are consistent with manual inspection of the sequence. The proposed method is also robust with respect to optimization parameter selection and the specific SNR conditions.

The use of contour shape features endows the proposed method its robustness to contrast variation and noise—it alleviates the gradient magnitude's sensitivity to contrast variations and gradient direction's sensitivity to noise. The combined regularization energy controlling the curve length and enclosing area provides improved robustness for the extracted contour features, compared to isolated high gradient points. The use of this new feature also has its associated limitations. In particular, the registration relies on data fidelity defined on a narrow band around the segmented contours and the motion for voxels away from this band is effectively obtained by "propagating" via the regularization term. Therefore, the method will fail to capture any subtle local motion pattern far from the GOIs. This is not a severe drawback since the major motion estimation interest normally lies within or around the GOI. However, when globally good motion estimation is

FIG. 8. Comparison of registration performance in the presence of intensity and contrast variation between benchmark and proposed method: (a) optical flow, (b) mutual information, (c) multimodality demons, (d) proposed. Arrows in (a)–(c) indicate observable misregistration locations and arrows in (d) indicate the corresponding locations for the proposed method with improved registration performance.

FIG. 9. Motion estimation results using real MR image sequences: (a) segmented contours (reference in magenta and target in cyan) and estimated motion field, showing downward and expansive motion of the kidneys, consistent with visual examination of the sequence; (b) checkerboard visualization preregistration; (c) checkerboard postregistration, where arrows are used to indicate locations with significant improvement, compared to the same locations indicated in (b).

FIG. 10. Temporal evolution of the estimated kidney motion. (a) mean displacement in the superior-inferior direction, (b) mean displacement in the left-right direction, (c) temporal anatomical boundary trajectories reconstructed from estimated motion for the first ten images frames, with corresponding points connected and coded in the same color.

FIG. 11. Test of sensitivity to parameter value selections by varying length λ and area β weighting coefficients: the segmentation results vary little for $\lambda \in (3, 20)$ and $\beta \in (0, 0.1)$, indicating the proposed method's insensitivity of parameter selections. (a) $\lambda = 3$, $\beta = 0.01$, (b) $\lambda = 10$, $\beta = 0.01$, (c) $\lambda = 20$, $\beta = 0.01$, (d) $\lambda = 10$, $\beta = 0$, (e) $\lambda = 10$, $\beta = 0.005$, (f) $\lambda = 10$, $\beta = 0.1$.

FIG. 12. Test of robustness to intensity variations and noise: the segmented contour only degraded slightly as both intensity variations and significant noise were introduced, when the same fixed optimization parameter values were used. (a) Noise free, (b) with additive gaussian noise, (c) with intensity variations, and (d) with both gaussian noise and intensity variations.

desired, one can extend the current method to use multiple GOIs, and use multiphase level set presentation,³⁴ for example, to segment and depict a set of contours simultaneously that provide sufficient coverage for the domain of interest.

With the registration relying on distance function representation of the segmented contours, a potential concern is the dependence of segmentation accuracy. There are several aspects to be considered. First, we have utilized good initialization of the LSF with the segmentation results from temporally adjacent frames to encourage more robust and consistent results. Second, we have chosen the level set based segmentation for its flexibility and its natural compatibility into variational optimization framework. In principle, one has the option of using alternative methods to generate contours, e.g., semi-interactive contouring schemes for possibly higher segmentation accuracy, and then deriving distance functions as input to the registration module. Finally, the major goal of the proposed method is motion estimation, not segmentation *per se*. Therefore, as long as the segmented contours are close enough to the GOIs and the segmentation results are consistent among the frames, the motion estimation results are not affected. In some sense, the segmented contours should be considered as “surrogate” descriptors for the organ or structure boundaries that sharing the same motion.

Note that sometimes GOI may become partially inconspicuous in one frame out of the registration pairs, when contrast kinetic makes it indistinguishable from its surrounding at that particular frame. When such inconspicuousness is severe enough to distort the segmented contour and make it inconsistent with the other contours in the sequence, the proposed method would fail. We are in the process of incorporating prior knowledge on both shape characterization^{35–37} and deformation fields to improve stability and robustness in these challenging situations.

The proposed method may also be challenged when the underlying presumptions in designing the regularizations are violated. In particular, we used a Tikhonov type regularization for motion field smoothness. When the motion under study is nonsmooth, such as sliding, one may need to either relax such assumption by decreasing the smoothness weight α or adjust the design of the regularization to preserve discontinuity or sliding.^{38–40}

In this study, we simulated the pairwise intensity and contrast changes to verify the robustness of the proposed method to such variations. We are in the process of incorporating more precise and realistic pharmacokinetic functions⁴¹ for sequence simulation with quantitative clinical verifications.

With the advantage of robustness to contrast and intensity variations, the proposed method is applicable to other problems involving contrast change, such as multimodality registration, dynamic contrast enhanced CT, and subtraction imaging as in angiography.

5. CONCLUSIONS

In summary, we have developed a novel method to estimate nonrigid motion for geometries of interest. The proposed method is based on robust contour shape features that

preserves good consistency in the presence of intensity and contrast variations. Quantitative analyses were made on 2D and 3D datasets showing promising results. Future work will consider incorporating shape and motion priors for better segmentation and registration. Parallel computation is being implemented to speed up the execution. A clinical protocol has been established to collect free-breathing DCE-MRI image sequences of the abdominal region to further validate the proposed method, and collaboration has been established with clinicians to evaluate its performance under different image quality condition and motion patterns.

ACKNOWLEDGMENTS

This work is in part supported by AACR, TRDRP, and NIH R01 grant. The authors thank Dr. M. Gach at University of Pittsburgh for motivating this work on kidney DCE study and Dr. Peng Hu for helpful discussions. They thank Mr. Weber Shao for editorial help and acknowledge Dr. D. Low, Dr. P. Kupelian, and Dr. M. Steinberg for departmental support.

^{a)} Author to whom correspondence should be addressed. Electronic mail: druan@mednet.ucla.edu

¹ L. G. Brown, “A survey of image registration techniques,” *ACM Comput. Surv.* **24**(4), 325–376 (1992).

² F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens, “Multimodality image registration by maximization of mutual information,” *IEEE Trans. Med. Imaging* **16**(2), 187–198 (1997).

³ J. Maintz and M. A. Viergever, “A survey of medical image registration,” *Med. Image Anal.* **2**(1), 1–36 (1998).

⁴ J. P. Pluim, J. A. Maintz, and M. A. Viergever, “Mutual-information-based registration of medical images: A survey,” *IEEE Trans. Med. Imaging* **22**(8), 986–1004 (2003).

⁵ B. Zitova and J. Flusser, “Image registration methods: a survey,” *Image Vision Comput.* **21**(11), 977–1000 (2003).

⁶ X. Pennec, P. Cachier, and N. Ayache, “Understanding the ‘demon’s algorithm’: 3D non-rigid registration by gradient descent,” In Taylor, C., Colchester, A. (eds) MICCAI 1999. LNCS Vol. 1679, pp. 597–605. Springer, Heidelberg (1999).

⁷ G. P. Penney *et al.*, “A comparison of similarity measures for use in 2D-3D medical image registration,” *IEEE Trans. Med. Imaging* **17**(4), 586–595 (1998).

⁸ T. Vercauteren, X. Pennec, A. Perchant, and N. Ayache, “Diffeomorphic demons: Efficient non-parametric image registration,” *NeuroImage* **45**(1), S61–S72 (2009).

⁹ W. M. Wells, P. Viola, H. Atsumi, S. Nakajima, and R. Kikinis, “Multimodal volume registration by maximization of mutual information,” *Med. Image Anal.* **1**(1), 35–51 (1996).

¹⁰ C. Hayes, A. R. Padhani, and M. O. Leach, “Assessing changes in tumor vascular function using dynamic contrast-enhanced magnetic resonance imaging,” *NMR Biomed.* **15**(2), 154–163 (2002).

¹¹ S. P. Sourbron, H. J. Michaely, M. F. Reiser, and S. O. Schoenberg, “MRI-measurement of perfusion and glomerular filtration in the human kidney with a separable compartment model,” *Invest. Radiol.* **43**(1), 40–48 (2008).

¹² F. Lietzmann, F. G. Zöllner, U. I. Attenberger, S. Haneder, H. J. Michaely, and L. R. Schad, “DCE-MRI of the human kidney using blade: A feasibility study in healthy volunteers,” *J. Magn. Reson Imaging* **35**(4), 868–874 (2012).

¹³ U. Attenberger, S. Sourbron, H. Michaely, M. Reiser, and S. Schoenberg, “Retrospective respiratory triggering renal perfusion MRI,” *Acta Radiol.* **51**(10), 1163–1171 (2010).

¹⁴ B. D. de Senneville, I. A. Mendichovszky, S. Roujol, I. Gordon, C. Moonen, and N. Grenier, “Improvement of MRI-functional measurement with automatic movement correction in native and transplanted kidneys,” *J. Magn. Reson Imaging* **28**(4), 970–978 (2008).

- ¹⁵S. Li, F. G. Zöllner, A. D. Merrem, Y. Peng, J. Roervik, A. Lundervold, and L. R. Schad, "Wavelet-based segmentation of renal compartments in DCE-MRI of human kidney: initial results in patients and healthy volunteers," *Comput. Med. Imaging Graph.* **36**(2), 108–118 (2012).
- ¹⁶M. Notohamiprodjo *et al.*, "Measuring perfusion and permeability in renal cell carcinoma with dynamic contrast-enhanced MRI: A pilot study," *J. Magn. Reson Imaging* **31**(2), 490–501 (2010).
- ¹⁷F. G. Zöllner, R. Sance, P. Rogelj, M. J. Ledesma-Carbayo, J. Rørvik, A. Santos, and A. Lundervold, "Assessment of 3D DCE-MRI of the kidneys using non-rigid image registration and segmentation of voxel time courses," *Comput. Med. Imaging Graph.* **33**(3), 171–181 (2009).
- ¹⁸J. Pluim, J. Maintz, and M. Viergever, "Image registration by maximization of combined mutual information and gradient information," *IEEE Trans. Med. Imaging* **19**(8), 809–814 (2000).
- ¹⁹A. D. Merrem, F. G. Zöllner, M. Reich, A. Lundervold, J. Rørvik, and L. R. Schad, "A variational approach to image registration in dynamic contrast-enhanced MRI of the human kidney," *J. Magn. Reson Imaging* **31**(5), 771–777 (2012).
- ²⁰X. Yang, P. Ghafourian, P. Sharma, K. Salman, D. Martin, and B. Fei, "Non-rigid registration and classification of the kidneys in 3D dynamic contrast enhanced DCE MR images," *SPIE Medical Imaging*, International Society for Optics and Photonics (IEEE, San Diego, CA, 2012), p. 83140B.
- ²¹E. Haber and J. Modersitzki, "Intensity gradient based registration and fusion of multi-modal images," *Methods Inf. Med.* **46**(3), 292–299 (2007).
- ²²S. Osher and R. Fedkiw, *Level Set Methods and Dynamic Implicit Surfaces* (Springer, New York, NY, 2003), Vol. 153.
- ²³S. Kichenassamy, A. Kumar, P. Olver, A. Tannenbaum, and A. Yezzi, Jr., "Conformal curvature flows: from phase transitions to active vision," *Arch. Ration. Mech. Anal.* **134**(3), 275–301 (1996).
- ²⁴C. Li, C. Xu, C. Gui, and M. D. Fox, "Distance regularized level set evolution and its application to image segmentation," *IEEE Trans. Image Process.* **19**(12), 3243–3254 (2010).
- ²⁵V. Estellers, D. Zosso, R. Lai, S. Osher, J.-P. Thiran, and X. Bresson, "An efficient algorithm for level set method preserving distance function," *IEEE Trans. Image Process.* **21**(12), 4722–4734 (2012).
- ²⁶J. A. Sethian, *Level Set Methods and Fast Marching Methods: Evolving Interfaces in Computational Geometry, Fluid Mechanics, Computer Vision, and Materials Science* (Cambridge University Press, Cambridge, UK, 1999), Vol. 3.
- ²⁷D. M. Young, *Iterative Solution of Large Linear Systems* (Dover Publications, Mineola, NY, 2003).
- ²⁸B. K. Horn and B. G. Schunck, "Determining optical flow," *Artif. Intell.* **17**(1), 185–203 (1981).
- ²⁹D.-J. Kroon and C. H. Slump, "MRI modality transformation in demon registration," in *Proceedings of IEEE International Symposium on Biomedical Imaging: From Nano to Macro, 2009. ISBI'09* (IEEE, New York, NY, 2009), pp. 963–966.
- ³⁰P. Viola and W. M. Wells III, "Alignment by maximization of mutual information," *Int. J. Comput. Vis.* **24**(2), 137–154 (1997).
- ³¹D. Rueckert, L. I. Sonoda, C. Hayes, D. L. Hill, M. O. Leach, and D. J. Hawkes, "Nonrigid registration using free-form deformations: Application to breast MR images," *IEEE Trans. Med. Imaging* **18**(8), 712–721 (1999).
- ³²S. Davies, A. Hill, R. Holmes, M. Halliwell, and P. Jackson, "Ultrasound quantitation of respiratory organ motion in the upper abdomen," *Br. J. Radiol.* **67**(803), 1096–1102 (1994).
- ³³M. Unser, A. Aldroubi, and M. Eden, "Fast B-spline transforms for continuous image representation and interpolation," *IEEE Trans. Pattern Anal. Mach. Intell.* **13**(3), 277–285 (1991).
- ³⁴L. A. Vese and T. F. Chan, "A multiphase level set framework for image segmentation using the mumford and shah model," *Int. J. Comput. Vis.* **50**, 271–293 (2002).
- ³⁵T. F. Chan and W. Zhu, "Level set based shape prior segmentation," in *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2005* (IEEE, New York, NY, 2005), Vol. 2, pp. 1164–1170.
- ³⁶D. Cremers, S. J. Osher, and S. Soatto, "Kernel density estimation and intrinsic alignment for shape priors in level set segmentation," *Int. J. Comput. Vis.* **69**(3), 335–351 (2006).
- ³⁷M. Rousson and N. Paragios, "Prior knowledge, level set representations & visual grouping," *Int. J. Comput. Vis.* **76**(3), 231–243 (2008).
- ³⁸T. Rohlfing, C. R. Maurer, Jr., D. A. Bluemke, and M. A. Jacobs, "Volume-preserving nonrigid registration of mr breast images using free-form deformation with an incompressibility constraint," *IEEE Trans. Med. Imaging* **22**(6), 730–741 (2003).
- ³⁹D. Ruan, "Directionally selective regularization for sliding preserving medical image registration," *Nuclear Science Symposium Conference Record (NSS/MIC), 2009 IEEE* (IEEE, New York, NY, 2009), pp. 2936–2939.
- ⁴⁰D. Ruan, S. Esedoglu, and J. A. Fessler, "Discriminative sliding preserving regularization in medical image registration," *Proceedings of IEEE International Symposium on Biomedical Imaging: From Nano to Macro, 2009. ISBI'09* (IEEE, New York, NY, 2009), pp. 430–433.
- ⁴¹N. Hackstein, J. Heckrodt, and W. S. Rau, "Measurement of single-kidney glomerular filtration rate using a contrast-enhanced dynamic gradient-echo sequence and the Rutland-Patlak plot technique," *J. Magn. Reson Imaging* **18**(6), 714–725 (2003).