Journal of Algorithms & Computational Technology

IPTA Workshop on Image Processing Techniques and Applications - Research Article

Journal of Algorithms & Computational Technology
Volume 14: 1–9

DOI: 10.1177/1748302620966691
journals.sagepub.com/home/act
SAGE

A variational joint segmentation and registration framework for multimodal images

Adela Ademaj[1], Lavdie Rada[2], Mazlinda Ibrahim[3] and Ke Chen[4]

Abstract

Image segmentation and registration are closely related image processing techniques and often required as simultaneous tasks. In this work, we introduce an optimization-based approach to a joint registration and segmentation model for multimodal images deformation. The model combines an active contour variational term with mutual information (MI) smoothing fitting term and solves in this way the difficulties of simultaneously performed segmentation and registration models for multimodal images. This combination takes into account the image structure boundaries and the movement of the objects, leading in this way to a robust dynamic scheme that links the object boundaries information that changes over time. Comparison of our model with state of art shows that our method leads to more consistent registrations and accurate results.

Keywords

Image processing, multimodal image processing, segmentation, registration, regmentation

Received 8 December 2019; revised 10 July 2020; accepted 26 September 2020

Introduction

In the last decades, the development of imaging devices significantly increased image data collections and consecutively the need for human interpretation. To reduce labour hours of specialists new image processing techniques are developed. Two main image processing techniques are image registration and segmentation. Image registration provides an understanding of the data behavior in two or more different scenarios taken at different times. Meanwhile, image segmentation is the process of partitioning objects or features considering the characteristics of each pixel. In a wide range of problems, such as comparison of images taken at different time or modalities, image registration depends on image segmentation and vice-versa. Erdt et al.[1] indicates that more than 20% of scientific research in medical imaging area requires a combined registration and segmentation scheme. The combined registration and segmentation models can be divided into two different categories: simultaneous registration and segmentation or joint segmentation and registration which lastly has been referred to as *regmentation*. In the last years, *regmentation* functional models have been introduced on monomodal imaging frameworks. These models are variational rigid registration based models,[2,3] variational nonrigid registration based models,[4,5] or atlas based models.[6] Recently, Ibrahim-Rada-Chen[7] present a linear curvature *regmentation* variational based model which gets profit on the global and local deformation and easily can cope with images which contain more than one object.

[1]Biomedical Engineering Department, Heidelberg University, Heidelberg, Germany
[2]Biomedical Engineering Department, Bahcesehir University, Istanbul, Turkey
[3]National Defence University of Malaysia, Kuala Lumpur, Malaysia
[4]Centre for Mathematical Imaging Techniques, University of Liverpool, Liverpool, UK

Corresponding author:
Lavdie Rada, Biomedical Engineering Department, Bahcesehir University, Istanbul, Turkey.
Email: lavdie.rada@eng.bau.edu.tr

In difference with monomodal images, multimodal image processing requires spatial correspondence estimation or extract certain information from images with different modalities (or protocols). Such example can be from imaging using CT and MRI protocols, which brings different image perspective into the medical field of clinical and pre-clinical diagnostics. Even though assessing image similarity becomes challenging for multimodal images, multimodal images are desired as they bring different image perspectives and properties. In difference with monomodal images registration, where simple image similarity measures between image pairs can provide a good spatial transformation, dealing with multimodalities is substantially more difficult. Registering multimodal images, which are acquired through different mechanisms and have distinct modalities, involves the alignment of both images in terms of shapes and salient components while preserving the modality of one given image. For a comprehensive overview of registration techniques in a systematic manner, the work of Sotiras et al.[8] is referred. By adding the active contour without edges model into the curvature joint registration based on MI proposed by Modersitzki,[9] a modified joint image segmentation and registration delivers results which are shown in Figure 1. The same images are processed with the new proposed joint *regmentation* model and the results shown in the last section indicate that the combination of segmentation and registration tasks into a single framework is relevant as avoids the errors produced from simultaneous tasks. The existing registration models for mono modal-images can be classified as unsupervised and supervised methods. Among different registration similarity measure, the most popular one is mutual information (MI), introduced by Viola et al.[10] Recently, learning-based approaches measuring image similarity have been proposed as well. Techniques involving KL-divergence,[11] max-margin structured learning,[12] boosting,[13] or deep learning[14] have been investigated. However, their large training sets make them limited by the lack of generalization to previously unseen object classes and the registration performance depends on numerical optimization of the optimal registration parameters. There are a few unsupervised works to solve the joint segmentation and registration problem. The work proposed by Wang and Vemuri[15] introduces a registration driven by segmentation of a reference image without varying degrees of non-rigidity. The model applies cross cumulative residual entropy as a distance measure[15] and piecewise constant Chan-Vese model for segmentation.[16] Although, the algorithm can accommodate image pairs having very distinct intensity distributions but fails in deforming large objects. Droske et. al[17] presented variational method based on a local energy density which uses a generalized motion of image morphology for multimodal image registration. A similar idea was lastly proposed by Aganj and Fischl[18] where the segmentation score was applied to register two multimodal brain images. This model is not designed for general image deformation as the model applies rigid image transformation. Recently, a joint segmentation and registration model has been proposed by Debroux and Le Guyader,[19] which is formulated based on the combination of nonlocal total variation and nonlocal shape descriptors. The fitting

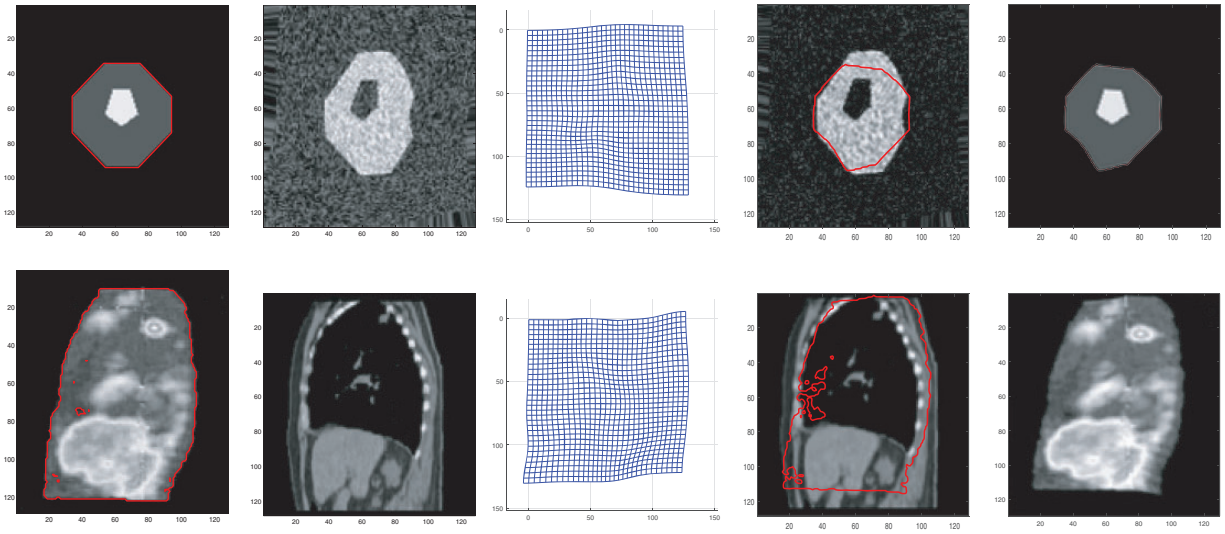

Figure 1. Simultaneous segmentation and linear curvature registration based on MI model noisy synthetic image and CT image. First column segmented template image, T & $\phi_0(\mathbf{x})$. Second column reference image R. Third column the deformation field. Forth column the transformed template image R & $\phi_0(\mathbf{x}+u(\mathbf{x}))$. Last column the template image after transformation, $T(\mathbf{x}+u(\mathbf{u}))$.

term of this model uses a weighted sum of squared difference distance measure, similar to[7] and the model depends on the many parameters such as window size, patch size, number of neighbors pixels, etc. To overcome the difficulties mentioned above we propose a new variational model which combines linear curvature, known for its ability to generate smooth transformations,[9,20] and MI as a distance measure[10,21] for the spatial transform.

The outline of this paper is as follows. In the next section, we review the existing models for monomodal image *regmentation* and multimodal image registration. Then, we introduce our proposed new model for multimodal image *regmentation*. In the last section, we show some numerical tests including comparison.

Related work

Deformable multimodal registration attempts to find spatial correspondence between image pairs that are acquired with distinct scanning protocols. In the registration process, these correspondences are estimated by determining the spatial transformation which is applied to a template image to make similar to the reference image. Generally, image registration is driven by the chosen similarity measure and the chosen deformation model. For images taken from the same modality, the image similarity usually is considered to be high, thus simple and direct image similarity measures are used, for instance, the sum of squared intensity differences (SSD). We start this section with an overview of the work introduced by Ibrahim-Rada-Chen[7] for monomodal *regmentation* of two given images. The model jointly merges the segmentation and registration into a variational function represented in a level set. This method improves the Guyader-Vese model[22] by replacing the nonlinear elastic term with linear curvature model and adding a weighted Heaviside sum of SSD term.

Let R be a reference image and T template image in a given domain $\Omega \subset R^d$, where d is the spatial dimensional of the given data. The registration problem is formulated as a transformation vector $\phi(\boldsymbol{x})$ such that $T(\phi(\boldsymbol{x})) = R(\boldsymbol{x})$. In the variational approach, image registration, the transformation is $\phi(\boldsymbol{x}) = \boldsymbol{x} + \boldsymbol{u}(\boldsymbol{x})$, with $\boldsymbol{u}(\boldsymbol{x})$ the displacement vector field $\boldsymbol{u}(\boldsymbol{x}) = (u_1(\boldsymbol{x}), u_2(\boldsymbol{x}))$. The aim of the Ibrahim-Rada-Chen[7] model is to match the contour of the template image and at the same time segment the reference image to show the deformation of displacement field lead by the segmentation process. The segmentation of the template image T is represented by the zero level line $\phi_x : \Omega \to \mathrm{R}$ which is the template contour Γ, denoted as in.[7] The joint functional proposed by Ibrahim-Rada-Chen[7] monomodal image *regmentation* is given as follows:

$$\begin{aligned}\min_{c_1,c_2,\boldsymbol{u}(\boldsymbol{x})} \mathcal{J} &= \mu \mathcal{D}^{\mathrm{SSDH}}(T, R, \phi_0(\boldsymbol{x}), \boldsymbol{u}(\boldsymbol{x})) + \lambda_1 \int_\Omega |R(\boldsymbol{x}) \\ &- c_1|^2 H_\varepsilon(\phi_0(\boldsymbol{x} + \boldsymbol{u}(\boldsymbol{x}))) \mathrm{d}\boldsymbol{x} + \lambda_2 \int_\Omega |R(\boldsymbol{x}) \\ &- c_2|^2 (1 - H_\varepsilon(\phi_0(\boldsymbol{x} + \boldsymbol{u}(\boldsymbol{x})))) \mathrm{d}\boldsymbol{x} + \alpha \mathcal{S}^{\mathrm{LC}}(\boldsymbol{u})\end{aligned} \tag{1}$$

where λ_1, λ_2, μ, α are numerical constants. $\boldsymbol{u}(\boldsymbol{x}) = (u_1(\boldsymbol{x}), u_2(\boldsymbol{x}))$ is the displacement vector field, ϕ_0 is a zero-level-set, $\mathcal{D}^{\mathrm{SSDH}}$ is the sum of squared difference distance measure

$$\mathcal{D}^{\mathrm{SSDH}} = \int_\Omega (T(\boldsymbol{x} + \boldsymbol{u}(\boldsymbol{x})) - R(\boldsymbol{x}))^2 H_\epsilon(\phi_0(\boldsymbol{x} + \boldsymbol{u}(\boldsymbol{x}))) \, \mathrm{d}\boldsymbol{x} \tag{2}$$

which is weighted by the regularised Heaviside function

$$H_\epsilon(\phi_0(\boldsymbol{x} + \boldsymbol{u}(\boldsymbol{x}))) = \frac{1}{2}\left(1 + \frac{2}{\pi}\arctan\frac{\phi_0(\boldsymbol{x} + \boldsymbol{u}(\boldsymbol{x}))}{\epsilon}\right). \tag{3}$$

The term $\mathcal{S}^{\mathrm{LC}}(\boldsymbol{u})$ in (1) is the curvature regulariser term

$$\mathcal{S}^{\mathrm{LC}}(\boldsymbol{u}) = \int_\Omega (\Delta u_1)^2 + (\Delta u_2)^2 \, \mathrm{d}\boldsymbol{x}. \tag{4}$$

The values of c_1 and c_2 in equation (1) present the average intensity values inside and outside the boundary $\phi_0(\boldsymbol{x})$ in the reference image. By minimizing the functional (1) containing the SSD term and linear curvature term, we get:

$$\begin{aligned} c_1 &= \frac{\int_\Omega R(\boldsymbol{x}) H_\varepsilon(\phi_0(\boldsymbol{x} + \boldsymbol{u})) \, \mathrm{d}\boldsymbol{x}}{\int_\Omega H_\varepsilon(\phi_0(\boldsymbol{x} + \boldsymbol{u})) \, \mathrm{d}\boldsymbol{x}}, \\ c_2 &= \frac{\int_\Omega R(\boldsymbol{x})(1 - H_\varepsilon(\phi_0(\boldsymbol{x} + \boldsymbol{u}))) \, \mathrm{d}\boldsymbol{x}}{\int_\Omega 1 - H_\varepsilon(\phi_0(\boldsymbol{x} + \boldsymbol{u})) \, \mathrm{d}\boldsymbol{x}}. \end{aligned} \tag{5}$$

In multilevel representation, the registration problem is solved using quasi-Newton method by updating the deformation field $\boldsymbol{u}(\boldsymbol{x})$ on each level. Even though the model shows good performance, it cannot cope with multimodal images.

Referring to registration for multimodality, different techniques have been introduced.[23–25] A significant work, proposed by Modersitzki,[9] based on mutual information and curvature regularizer, competes the other registration models. The joint functional for multimodal registration is given as follows:

$$J(\boldsymbol{x}+\boldsymbol{u}(\boldsymbol{x})) = \mathcal{D}[T,R] + \alpha\mathcal{S}^{\mathrm{LC}} \tag{6}$$

where $\mathcal{S}^{\mathrm{LC}}$ is the curvature regularization term, α is a numerical constant and $\mathcal{D}[T,R]$ presents the distance measures considered as functionals in multimodal image pair T and R. Modersitzki[9] suggested to use mutual information (MI) (shortly reviewed below as CASE 1) or Normalized Gradient Field (NGF) (shortly reviewed below as CASE 2) to measure the similarity distance between two multimodal images.

CASE 1: The joint functional based on MI, $\mathcal{D}[T,R] = \mathcal{D}^{\mathrm{MI}}$

$$J(\boldsymbol{x}+\boldsymbol{u}(\boldsymbol{x})) = \mathcal{D}^{MI} + \alpha\mathcal{S}^{\mathrm{LC}} \tag{7}$$

where $\mathcal{D}^{MI}$ presents mutual information as a distance measure between the given images T and R and it is given with the formula:

$$\begin{aligned}\mathcal{D}^{\mathrm{MI}} = &\int_{\Omega}\rho_{[T]}\log\rho_{[T]}dt + \int_{\Omega}\rho_{[R]}\log\rho_{[R]}dr \\ &- \int_{\Omega^2}\rho_{[T,R]}\log\rho_{[T,R]}d(t,r)\end{aligned} \tag{8}$$

where $\rho_{[T]}(t)$ and $\rho_{[R]}(r)$ are marginal densities which are expressed as:

$$\rho_{[T]}(t) = \int_{\Omega}\rho_{[T,R]}(t,r)dr,\ \ \rho_{[R]}(r) = \int_{\Omega}\rho_{[T,R]}(t,r)dt \tag{9}$$

and $\rho_{[T,R]}(t,r)$ presents joint gray value distributions. Its discretized form is expressed as:

$$\rho_{[T,R]}(t,r) = \frac{1}{m}\sum_{j=1}^{m}K(t-T(\boldsymbol{x}+\boldsymbol{u}(\boldsymbol{x})),\sigma)K(r-R(\boldsymbol{x}),\sigma) \tag{10}$$

with σ that stands for the width of Parzen density estimator and K for kernel and m presents the size of the sample image. The kernel positions in template and reference image are represented with t and r respectively in the gray value space for $j=1$.

CASE 2: The joint functional based on NGF, $\mathcal{D}[T,R] = \mathcal{D}^{\mathrm{NGF}}$

$$J(\boldsymbol{x}+\boldsymbol{u}(\boldsymbol{x})) = \mathcal{D}^{NGF} + \alpha\mathcal{S}^{\mathrm{LC}}, \tag{11}$$

$\mathcal{D}^{NGF}$ presents Normalized Gradient Force distance measure for two given multimodal images, T and R which is defined as follows

$$\mathcal{D}^{\mathrm{NGF}} = \int_{\Omega}1 - (n[T(\boldsymbol{x}+\boldsymbol{u}(\boldsymbol{x}))]^{T}n[R(\boldsymbol{x})])^2 d\boldsymbol{x} \tag{12}$$

where

$$\begin{aligned}&n[T(\boldsymbol{x}+\boldsymbol{u}(\boldsymbol{x}))]^{\mathrm{T}}n[R(\boldsymbol{x})] \\ &\approx\left(\frac{\nabla T_i}{\sqrt{\|\nabla T_i\|^2+\eta^2}}\right)^{\mathrm{T}}\left(\frac{\nabla R_i}{\sqrt{\|\nabla R_i\|^2+\eta^2}}\right)\end{aligned} \tag{13}$$

where η is an important constant parameter because it determines the edge and what has to be within the specified noise level. Moreover, the terms ∇T_i and ∇R_i present gradient intensity of template image and reference image, respectively which are computed as

$$\nabla T_i = [(\partial_1^h T(\boldsymbol{x}+\boldsymbol{u}(\boldsymbol{x})))_i, (\partial_2^h T(\boldsymbol{x}+\boldsymbol{u}(\boldsymbol{x})))_i]^T, \tag{14}$$

$$\nabla R_i = [(\partial_1^h R(\boldsymbol{x}))_i, (\partial_2^h R(\boldsymbol{x}))_i] \tag{15}$$

The NGF can be considered as the L2 norm of R, the residual of the alignment of the normalised gradients of two images at a pixel position $\boldsymbol{x}$,

$$r_{\boldsymbol{x}}^h = 1 - (n[T(\boldsymbol{x}+\boldsymbol{u}(\boldsymbol{x}))]^{T}n[R(\boldsymbol{x})])^2 \tag{16}$$

for discrete images T and R of size $N\times N$ using finite difference method. The images are discretised on a uniform mesh using vertex centred discretisation where $\boldsymbol{x}_{i,j}$ denotes the pixel position or on a non-uniform mesh with finite difference method. The gradient is calculated using

$$\begin{aligned}\partial_{x_1}T^h(\boldsymbol{x}_{i,j}) &= \frac{T^h(\boldsymbol{x}_{i+1,j}) - T^h(\boldsymbol{x}_{i-1,j})}{2h},\ \ \partial_{x_2}T^h(\boldsymbol{x}_{i,j}) \\ &= \frac{T^h(\boldsymbol{x}_{i,j+1}) - T^h(\boldsymbol{x}_{i,j-1})}{2h},\end{aligned}$$

where the first order central finite difference scheme is used to approximate the first order derivatives. The discretized form of NGF distance measure:

$$\mathcal{D}^{\mathrm{NGF}} = h^2\sum_{i=1}^{N^2}r_i \tag{17}$$

with the supposition that the spacial discretization is $h_1 = h_2 = h$. In short words, NGF is based on the alignment of the edges in the reference and template images. The gradient is normalised with its magnitude.

The proposed new joint *regmentation* model for multimodality

Our joint functional model for multimodal image *regmentation* combines an active contour without edges, a curvature regulariser, and a mutual information distance measure. Our proposed *regmentation* functional is given with the formula:

$$\begin{aligned}&\min_{c_1,c_2,\boldsymbol{u}(\boldsymbol{x})} \mathcal{J} = \lambda_1 \int_\Omega |R(\boldsymbol{x}) - c_1|^2 H_\epsilon(\phi_0(\boldsymbol{x}+\boldsymbol{u}(\boldsymbol{x}))) \, \mathrm{d}\boldsymbol{x} \\ &+\lambda_2 \int_\Omega |R(\boldsymbol{x}) - c_2|^2 (1 - H_\epsilon(\phi_0(\boldsymbol{x}+\boldsymbol{u}(\boldsymbol{x})))) \, \mathrm{d}\boldsymbol{x} + \gamma \mathcal{S}^{\mathrm{LC}}(\boldsymbol{u}) \\ &+\mu \mathcal{D}^{\mathrm{MIH}}(T, R, \phi_0(\boldsymbol{x}), \boldsymbol{u}(\boldsymbol{x}))\end{aligned} \tag{18}$$

where $\mathcal{S}^{\mathrm{LC}}$ is the linear curvature, $\mathcal{D}^{\mathrm{MIH}}$ represents the mutual information distance term which is weighted by the parameter μ. The $\mathcal{D}^{\mathrm{MIH}}$ term is evaluated as the mutual information between foreground of the template and the reference, TH_ϵ and RH_ϵ, respectively:

$$\begin{aligned}\mathcal{D}^{\mathrm{MIH}} = &\int_\Omega \rho_{[TH_\epsilon]} \log \rho_{[TH_\epsilon]} dt + \int_\Omega \rho_{[RH_\epsilon]} \log \rho_{[RH_\epsilon]} dr \\ &- \int_{\Omega^2} \rho_{[TH_\epsilon, RH_\epsilon]} \log \rho_{[TH_\epsilon, RH_\epsilon]} d(t,r)\end{aligned}$$

The joint functional is solved with respect to the displacement field using discretise then optimise approach based on the quasi-Newton method in multilevel framework for faster implementation. The grid points are located at the center of the cell $\Omega^h = \{\boldsymbol{x}_{i,j} = (x_{1,i}, x_{2,j}) = ((i-0.5)h, (j-0.5)h)\}$, $1 \le i,j \le N$. We re-define the solution vector $\boldsymbol{U} = \begin{bmatrix}\boldsymbol{u}_1 \\ \boldsymbol{u}_2\end{bmatrix}_{2N^2 \times 1}$, $\boldsymbol{x} = \begin{bmatrix}\boldsymbol{x}_1 \\ \boldsymbol{x}_2\end{bmatrix}_{2N^2 \times 1}$, where $\boldsymbol{u}_1$, $\boldsymbol{u}_2$ present displacement vectors and $\boldsymbol{x}_1$, $\boldsymbol{x}_2$ position vectors. Thus, the discretised form of the joint functional in (18), by a finite difference method is:

$$\begin{aligned}\min_{c_1,c_2,\boldsymbol{U}} \mathcal{J} = &\lambda_1 \sum_{i,j=1}^{N} |R(\boldsymbol{x}_{i,j}) - c_1|^2 H_\varepsilon(\phi_0(\boldsymbol{x}_{i,j} + \boldsymbol{u}(\boldsymbol{x}_{i,j}))) \\ &+ \lambda_2 \sum_{i,j=1}^{N} |R(\boldsymbol{x}_{i,j}) - c_2|^2 (1 - H_\varepsilon(\phi_0(\boldsymbol{x}_{i,j} + \boldsymbol{u}(\boldsymbol{x}_{i,j})))) \\ &+ \mu \eta_t \eta_r \sum_{i=1}^{n_t} \sum_{j=1}^{n_r} \rho_{i,j} \log(\rho_{i,j} + \varepsilon) \\ &+ \gamma \sum_{l=1}^{2} \sum_{i,j=1}^{N} (-4u_l(\boldsymbol{x}_{i,j}) + u_l(\boldsymbol{x}_{i+1,j}) + u_l(\boldsymbol{x}_{i-1,j}) \\ &+ u_l(\boldsymbol{x}_{i,j+1}) + u_l(\boldsymbol{x}_{i,j-1}))^2\end{aligned} \tag{19}$$

where $\eta_t = (t_n - t_0)/n_t$ and $\eta_r = (r_n - r_0)/n_r$, n_t and n_r present the number of grid. The ϵ is just a small numerical constant to prevent extra considerations such as $''0log0''$. Furthermore, we are using homogeneous Neumann boundary conditions approximated by one side differences $u_l(\boldsymbol{x}_{i,1}) = u_l(\boldsymbol{x}_{i,2}), u_l(\boldsymbol{x}_{1,j}) = u_l(\boldsymbol{x}_{2,j})$, $u_l(\boldsymbol{x}_{i,N-1}) = u_l(\boldsymbol{x}_{i,N}), u_l(\boldsymbol{x}_{N-1,j}) = u_l(\boldsymbol{x}_{N,j}), l = 1, 2$.

Minimizing equation (19) brings a system of nonlinear equation with unknown $\mathbf{U}$:

$$\triangle \mathcal{J} = 0 \tag{20}$$

where

$$\begin{aligned}\Delta\mathcal{J} = &\lambda_1 \sum_{i,j=1}^{N} |R(\boldsymbol{x}_{i,j}) - c_1|^2 H_\epsilon(\phi_0(\boldsymbol{x}_{i,j} + \boldsymbol{u}(\boldsymbol{x}_{i,j})) \\ &+ \lambda_2 \sum_{i,j=1}^{N} |R(\boldsymbol{x}_{i,j}) - c_2|^2 (1 - H_\epsilon(\phi_0(\boldsymbol{x}_{i,j} + \boldsymbol{u}(\boldsymbol{x}_{i,j}))) \\ &+ \mu \mathcal{D}^{\mathrm{MIH}} + \gamma \mathbf{u}^T \mathbf{B} \mathbf{u}\end{aligned} \tag{21}$$

The gradient of the regularization term is computed as the multiplication of $\mathbf{B}$ matrix (constant matrix) of size $2N^2 \times 2N^2$ that contains the coefficients of $\mathbf{U}$. The matrix $\mathbf{B}$ is written as

$$\mathbf{B} = \begin{bmatrix} L^T L & 0 \\ 0 & L^T L \end{bmatrix}, \tag{22}$$

where L is a block tridiagonal matrix from the regularization term. For finding the solution of our minimization problem, we start with zero initial guess, $\boldsymbol{U} = 0$, solving

$$\boldsymbol{H} \delta \boldsymbol{U} = -\boldsymbol{G} \tag{23}$$

where

$$\boldsymbol{G} = \nabla_u \mathcal{J} \tag{24}$$

for $\delta \boldsymbol{U}$ and update $\boldsymbol{U} \leftarrow \boldsymbol{U} + \tau \delta \boldsymbol{U}$ with τ as the Armijo line search parameter.[26] $\boldsymbol{H}$ and $\boldsymbol{G}$ are the Hessian and gradient matrix for the functional $\mathcal{J}$ in equation (19) with respect to the displacement vector $\boldsymbol{U}$. The algorithm for the proposed *regmentation* model based on MI is given in Algorithm 1 where the multilevel approach is adapted for fast computation of the model.

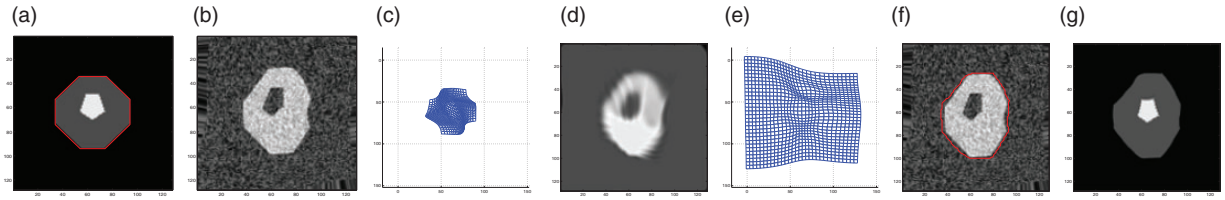

Figure 2. Comparison between monomodal image *regmentation*[7] and the proposed model. The monomodal joint *regmentation* fails to register with parameters $\lambda_1 = \lambda_2 = \mu = 1$ and $\alpha = 5$ (in (c) and (d)). The proposed model successfully deforms the given template image with $\lambda_1 = \lambda_2 = \gamma = 0.0005$ and $\mu = 2.0e-05$ (in (e), (f) and (g)). The MI distance difference of our proposed model before is $D^{MIH} = -0.22806$ which successfully minimizes the distance after to $D^{MIH} = -0.53424$. (a) T and $\phi_0(\mathbf{x})$, (b) R, (c) SSD $\mathcal{F} = -0.0994$, (d) SSD $T(\mathbf{x}+u(\mathbf{x}))$, (e) MIH $\mathcal{F} = 0.27148$, (f) MIH R and $\phi_{(\mathbf{x}+u(\mathbf{x}))}$, (g) MIH $T(\mathbf{x}+u(\mathbf{x}))$.

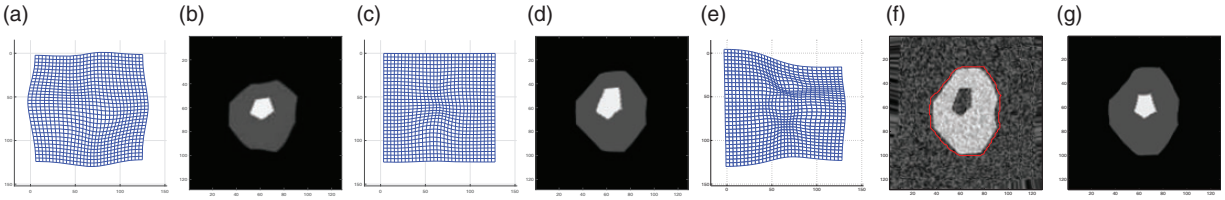

Figure 3. Comparison between the linear curvature models for synthetic images using MI and NGF[28] (refer to (a),(b) and (c),(d) for MI and NGF results of derformation field and registered template image, respectively). The proposed model is compared to both of these models and shows that successfully deformes the given template image. In contrast, curvature model using MI gets stuck in local minima which is demonstrated by the distance measurements before registration $D^{MI} = -1.0303$ and after the registration $D^{MI} = -1.0426$. Additionally, the curvature model using NGF does not show large deformation with regularizer parameter $\alpha = 2$. MI $\mathcal{F} = -0.5835$ (a) MI $T(\mathbf{x}+u(\mathbf{x}))$ (b) NGF $\mathcal{F} = -0.7465$ (c) NGF $T(\mathbf{x}+u(\mathbf{x}))$ (d) MIH $\mathcal{F} = 0.27148$ (e) MIH R and $\phi_{(\mathbf{x}+u(\mathbf{x}))}$ (f) MIH $T(\mathbf{x}+u(\mathbf{x}))$ (g)

Algorithm 1 Joint Regmentation Method Algorithm for Multimodal Images: $(\mathcal{D}^{\mathrm{MIH}}, \mathcal{S}^{\mathrm{LC}}, \phi(\mathbf{x}), u(\mathbf{x})) \leftarrow JRMI(R, T, \lambda_1, \lambda_2, \mu, \boldsymbol{U}, \phi^0)$

Step 1. Initialize the level set function $\phi(\mathbf{x})$ to be a binary function as follows:

$$\phi^0 = \phi(\mathbf{x}, t=0) = \begin{cases} -\rho & \text{if } \mathbf{x} \in \Omega_0 - \partial\Omega_0 \\ 0 & \text{if } \mathbf{x} \in \partial\Omega_0 \\ \rho & \text{if } \mathbf{x} \in \Omega - \Omega_0 \end{cases}$$

where $\rho > 0$ is a constant, Ω_0 is a subset in the image domain Ω and $\partial\Omega_0$ is the boundary of Ω_0, and zero inital guess for the displacement field

$$\boldsymbol{U} = 0.$$

Step 2. Finding deformation field using multilevel strategy:
for $level = coarseLevel : finestLevel$ **do**
* Compute the deformation field on each level $\boldsymbol{U}^{(level)} \leftarrow Registration(R^{(level)}, T^{(level)}, \boldsymbol{U}^{(level)}, \phi^0)$, using Quasi-Newton method as follows:
• update c_1 and c_2 using equation (5)
• update $\boldsymbol{U}^{(level)}$ and solve equation (23)
• continue until convergence criterion is satsified
* $\boldsymbol{U}^{(level)}$ is interpolated to each next finer *level* after registration
end for

Numerical results and conclusions

In this section, we present several examples for synthetic as well as real data in comparison with the linear curvature model based on MI applied on a synthetic image and a set of real images. In addition, the results of our proposed joint *regmentation* are compared to image registration applying mutual information as a distance measure and image segmentation simultaneously. The image pairs used in all our experiments have significantly different intensity profiles, deformation and presence of noise. In each iteration we compute the Jacobian matrix of the transformation,

$$J = \begin{bmatrix} 1 + \dfrac{\partial u_1}{\partial x_1} & \dfrac{\partial u_1}{\partial x_2} \\ \dfrac{\partial u_2}{\partial x_1} & 1 + \dfrac{\partial u_2}{\partial x_2} \end{bmatrix} \tag{25}$$

and the minimal value of Jacobian matrix $\mathcal{F} = \min(\det(J))$ is also calculated to make sure there is no folding or cracking in the deformation field if its

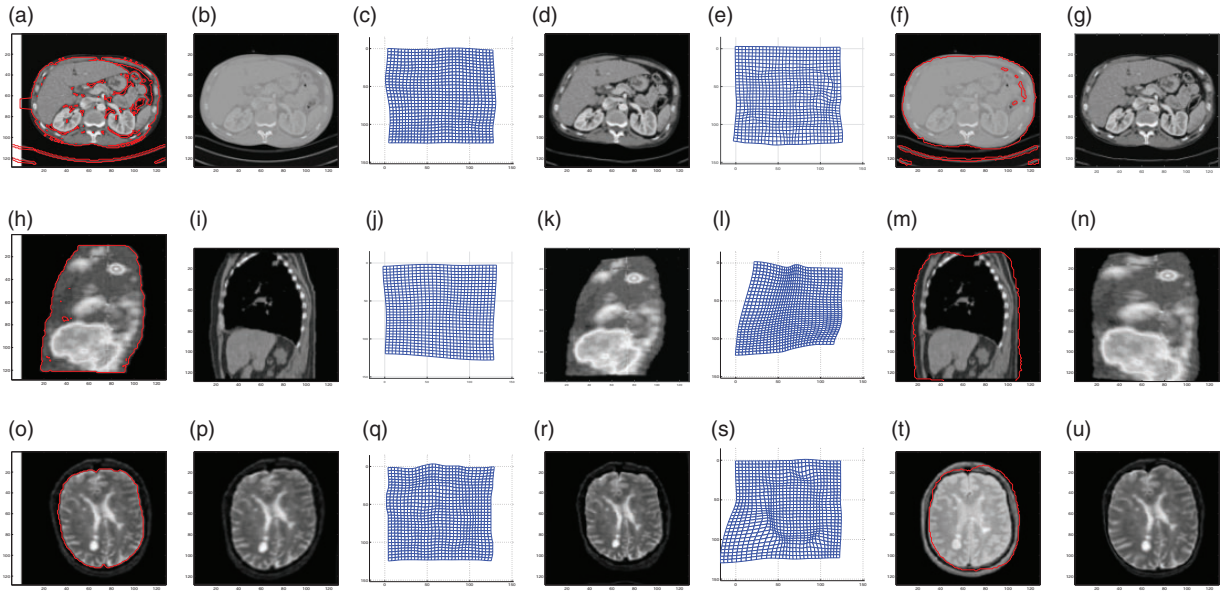

Figure 4. Comparison between linear curvature model using MI[28] and the proposed model for real images data. T & $\phi_0(\mathbf{x})$ (a) R (b) $\mathcal{F} = -0.3094$ (c) $T(\mathbf{x}+u(\mathbf{x}))$ (d) $\mathcal{F} = 0.53174$ (e) R & $\phi_0(\mathbf{x}+u(\mathbf{x}))$ (f) $T(\mathbf{x}+u(\mathbf{x}))$ (g) T & $\phi_0(\mathbf{x})$ (h) R (i) $\mathcal{F} = -0.4770$ (j) $T(\mathbf{x}+u(\mathbf{x}))$ (k) $\mathcal{F} = 0.56764$ (l) R & $\phi_0(\mathbf{x}+u(\mathbf{x}))$ (m) $T(\mathbf{x}+u(\mathbf{x}))$ (n) T & $\phi_0(\mathbf{x})$ (o) R (p) $\mathcal{F} = -0.3856$ (q) $T(\mathbf{x}+u(\mathbf{x}))$ (r) $\mathcal{F} = 0.46249$ (s) R & $\phi_0(\mathbf{x}+u(\mathbf{x}))$ (t) $T(\mathbf{x}+u(\mathbf{x}))$ (u)

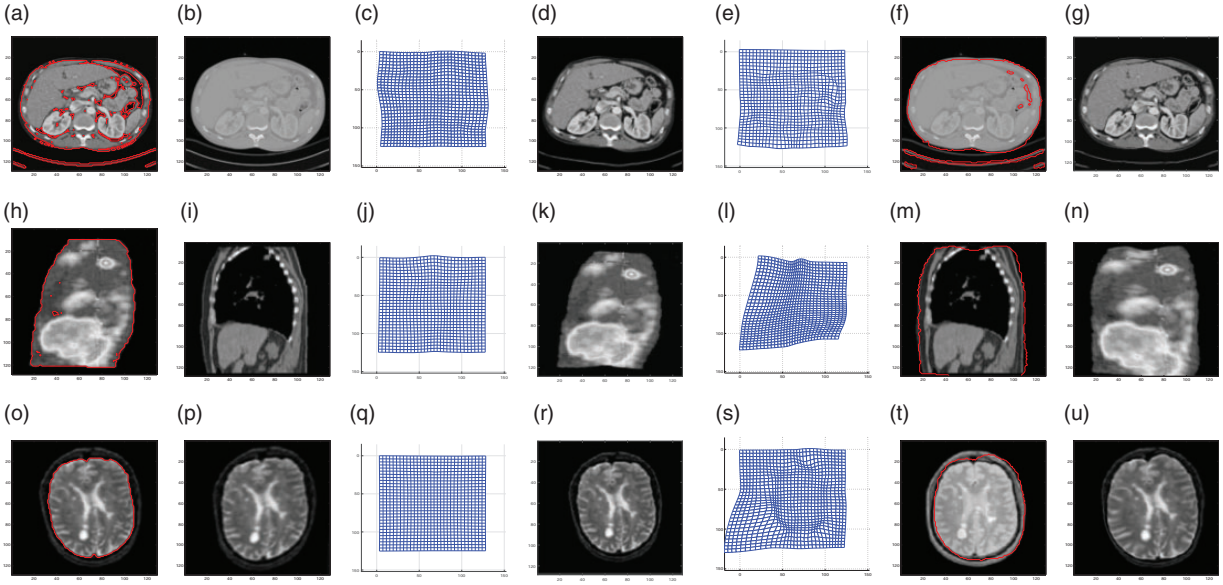

Figure 5. Comparison between linear curvature model using NGF[28] and the proposed model for real images data. T & 0 f (x) (a) R (b) F = –0.3094 (c) T(x + u(x)) (d) F = 0.56233 (e) R & 0 f (x+u(x)) (f) T(x + u(x)) (g) T & 0 f (x) (h)R (i) $\mathcal{F} = -0.3926$ (j) $T(\mathbf{x}+u(\mathbf{x}))$ (k) $\mathcal{F} = 0.56764$ (l) R & $\phi_0(\mathbf{x}+u(\mathbf{x}))$ (m) $T(\mathbf{x}+u(\mathbf{x}))$ (n) T & $\phi_0(\mathbf{x})$ (o) R (p) $\mathcal{F} = -0.5284$ (q) $T(\mathbf{x}+u(\mathbf{x}))$ (r) $\mathcal{F} = 0.46249$ (s) R & $\phi_0(\mathbf{x}+u(\mathbf{x}))$ (t) $T(\mathbf{x}+u(\mathbf{x}))$ (u)

Table 1. The first, second and third columns show the linear curvature model[17] parameters and distance measurements before and after, while the other columns are referring to our model, shown in Figure 4.

Figure 4	α	before $\mathcal{D}^{\mathrm{MI}}$	after $\mathcal{D}^{\mathrm{MI}}$	λ_1	λ_2	μ	γ	before $\mathcal{D}^{\mathrm{MIH}}$	after $\mathcal{D}^{\mathrm{MIH}}$
Chest	0.05	–1.4938	–1.3204	0.01	0.01	0.01	0.00045	–1.1677	–1.6247
Thorax	0.5	–0.2478	–0.2608	2	2	0.1	2	–0.40906	–0.88796
Brain	0.003	–1.1994	–1.2234	0.002	0.002	0.002	0.000035	–1.4477	–1.5911

value is greater than 0. Figures 2 and 3 present the results of the proposed model in comparison with Ibrahim et al.[7] *regmentation* model, and linear curvature image registration model based on MI and NGF proposed by Modersitzki et al.[28] The results show that the Ibrahim et al.[7] model, which uses the SSD similarity measure, is not capable to register and segment the given synthetic multimodal images whereas the proposed model can successfully cope with it. The fail of the method proposed by Ibrahim et. al[7] is expected as the model is not designed for multimodal images. On the other hand, the linear curvature registration model which applies MI, proposed by Modersitzki et al.,[28] provides larger deformation than the same model using NGF, referring to Figure 3. Even though, this model still has a poor registration in comparison with the proposed method, shown in Figure 2 last three columns.

Figures 4 and 5 show the results on real multimodal data images. In the first row of both figures, we show comparison results between the proposed model and linear curvature model[28] of two chest images (T1 and T2 weighted images). The second and the third row have the thorax (PET and CT images[28]) and brain images with high deformation and presence of noise of a T2 weighted CT image for the template and a MRI image for the reference. We clearly see that the proposed model delivers good results, referring to the fifth, sixth, and seventh column of Figure 4, whereas the linear curvature model[28] shown in the third and fourth column stuck in local minima due to highly non convexity of MI functional. Our multimodal joint *regmentation* provides accurate segmentation and registration in comparison also to linear curvature based on NGF model.[28] These results are shown in Figure 5. The gradient field distance measure involves second order image derivatives, as a consequence there can be a problem dealing with noisy images. Table 1 shows the parameters and distances measurements before and after the performance of linear curvature model[28] and our model. We note that the parameter α, λ_1, λ_2, μ, and γ are tuned according to the paired modalities. It can be noticed a similarity for those parameters in between the chest image case where T1 and T2 weighted CT images are involved and the image of the brain with T2 weighted CT image for the template and an MRI image for the reference, whereas those parameters change for the thorax image where PET image for the template is paired with a CT image for the reference. This drawback is expected similar to all the other inverse problems. In conclusion, the proposed model is suitable for jointly segmenting and registering images with intensity difference, severe deformation, and presence of noise. The model avoids in this way the need of pre-or post-processes such as pre-registration step which might be required for the segmentation task or vice-versa.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work is partially supported by MOHE of Malaysia under grant RACER/1/2019/ICT01/UPNM/1.

ORCID iD

Lavdie Rada https://orcid.org/0000-0002-2688-4962

References

1. Erdt M, Steger S and Sakas G. Regmentation: a new view of image segmentation and registration. *J Radiat Oncol Inform* 2012; 4: 1–23.
2. Dydenko I, Friboulet D and Magnin I. A variational framework for affine registration and segmentation with shape prior: application in echocardiographic imaging. In: VLSM *workshop-ICCV*, Los Alamitos, California, 2003, pp.201–208.
3. Yezzi A, Zollei L and Kapur T. A variational framework for joint segmentation and registration. In: *Proceedings IEEE workshop on Mathematical methods in biomedical image analysis (MMBIA)*, Los Alamitos, California, 2001, pp.44–51.
4. Alvarez L, Weickert J and Sanchez J. Reliable estimation of dense optical flow fields with large displacements. *Int J Comput Vision* 2000; 39: 41–56.
5. Unal G and Slabaugh G. Coupled PDES for non-rigid registration and segmentation. *IEEE Comput Soc Conf Compu Vision Pattern Recogn* 2002; 1: 37–42.
6. Gooya A, Pohl KM, Bilello M, et al. Joint segmentation and deformable registration of brain scans guided by a tumor growth model. In: Fichtinger G, Martel A and Peters T (eds). *International Conference on Medical image computing and computer-assisted intervention (MICCAI)*, Berlin, Heidelberg: Springer, 2011, pp.532–540.
7. Ibrahim M, Chen K and Rada L. An improved model for joint segmentation and registration based on linear curvature smoother. *J Algorithms Comput Technol* 2016; 10: 314–324.
8. Sotiras A, Christos D and Paragios N. Deformable medical image registration: a survey. *IEEE Trans Med Imaging* 2013; 32: 1153–1190.
9. Modersitzki J. *Flexible algorithms for image registration*. Philadelphia, PA: SIAM Publications, 2009.
10. Viola P and Wells III, WM. Alignment by maximization of mutual information. *International journal of computer vision (IJCV)*, 1997; 24: 137–154.

11. Guetter C, Xu FS and Hornegger J. *Learning based non-rigid multi-modal image registration using Kullback-Leibler divergence*. In: Duncan JS and Gerig G. (eds) *Medical image computing and computer-assisted intervention (MICCAI)*. Berlin Heidelberg: Springer, 2005, pp.255–262.
12. Lee D, Hofmann M, Steinke F, et al. Learning similarity measure for multi-modal 3D image registration. In: *IEEE Computer society conference on computer vision and pattern recognition (CVPR)*. Miami, FL, USA: IEEE, 2009, pp.186–193.
13. Michel F, Bronstein M, Bronstein A and Paragios N. Boosted metric learning for 3D multi-modal deformable registration. *Proceedings of the 8th International symposium on biomedical imaging (ISBI): from nano to macro*. Chicago, Illinois, USA: IEEE, 2011, pp.1209-1214.
14. Simonovsky M, Gutierrez-Becker B, Mateus D, et al. A deep metric for multimodal registration. In: Ourselin S, Joskowicz L, Sabuncu M, Unal G and Wells W. (eds) Medical image computing and computer-assisted intervention (MICCAI). Athens, Greece: Springer, 2016, pp.10–18.
15. Wang F and Vemuri BC. Simultaneous registration and segmentation of anatomical structures from brain MRI. In: Duncan JS and Gerig G. (eds) Medical image computing and computer-assisted intervention (MICCAI). Berlin Heidelberg: Springer, 2005, pp.17–25.
16. Chan T and Vese L. An active contour model without edges. In: Nielsen M, Johansen P, Olsen OF, et al. (eds) *Scale-space theories in computer vision*. Berlin, Heidelberg: Springer Berlin Heidelberg, 1999, pp.141–151.
17. Droske M and Rumpf M. Multiscale joint segmentation and registration of image morphology. *IEEE Trans Pattern Anal Mach Intell* 2007; 29: 2181–2194.
18. Aganj I and Fischl B. Multimodal image registration through simultaneous segmentation. *IEEE Signal Process Lett* 2017; 24: 1661–1665.
19. Debroux N and Guyader CL. A joint segmentation/registration model based on a nonlocal characterization of weighted total variation and nonlocal shape descriptors. *SIAM Imaging Sciences* 2018; 11: 957–990.
20. Fischer B and Modersitzki J. Curvature based image registration. *J Math Imaging Vision* 2003; 18: 81–85.
21. Maes F, Collignon A, Vandermeulen D, et al. Multimodality image registration by maximization of mutual information. *IEEE Trans Med Imaging* 1997; 16: 187–198.
22. Guyader L and Luminita A. A combined segmentation and registration framework with a nonlinear elasticity smoother. *Comput Vis Image Understand* 2011; 115: 1689–1709.
23. Jiang D, Shi Y, Chen X, et al. Fast and robust multimodal image registration using a local derivative pattern. *Med Phys* 2017; 44: 497–509.
24. Ruhaak J, Konig L, Hallmann M, et al. Intensity gradient based registration and fusion of multimodal images. In: *IEEE 10th international symposium on biomedical imaging*. Piscataway: IEEE, 2013, pp.572–575.
25. Shakir H, Ahsan ST and Faisal N. Multimodal medical image registration using discrete wavelet transform and Gaussian pyramids. In: *IEEE International conference on imaging systems and techniques*. Piscataway: IEEE, 2015, pp.1–6.
26. Nocedal J and Wright S. *Numerical optimization*. New York: Springer, 1999.
27. Modersitzki. J. *Numerical methods for image registration*. Oxford: Oxford University Press, 2004.