Medical Image Registration Method Based on Simulated CT

Xuqing Wang[✉], Yanan Su, Ruoyu Liu, Qianhui Qu, Hao Liu, and Yi Gu

School of Artificial Intelligence and Computer Science, Jiangnan University, 1800 Lihu Avenue, Wuxi 214122, Jiangsu, People's Republic of China
1063390860@qq.com

Abstract. In recent years, artificial intelligence (AI) has been widely used in medicine, for example, tumor screening, qualitative diagnosis, radiotherapy organ delineation, curative effect evaluation and prognosis, etc. Medical image registration is a typical problem and technical difficulty in the field of image processing. Specifically, in order to achieve the goal of information fusion, a certain spatial transformation is used to map two images in a set of image data sets to each other, so that the points corresponding to the same location in the two images are matched one by one. Most of the existing algorithms perform rigid registration based on simple features of images or non-automatic registration based on artificial markers. The registration accuracy is not ideal due to the imaging mechanism of different modal images, moreover, calculate consumption is high. In order to overcome these various problems, this essay proposes a new method of image registration based on differential homeomorphic deformation field to realize the same mode image registration between simulated CT and given CT, and that of image fusion using pixel weighted average method. Good registration and fusion imaging performance can still be achieved in the abdomen, which is extremely challenging body part in terms of medical image registration. The use of intelligent medical technology can reduce the workload of doctors and improve the accuracy of registration to make the diagnosis more Scientifically and reliably.

Keywords: Intelligent medical technology · Medical image registration · Differential homeomorphism algorithm · Pixel weighted average method

1 Introduction

In recent years, with the continuous growth of digital economy, artificial intelligence (AI) has been developing rapidly and now is deeply integrated with many application scenarios, Wise Information Technology of med (WITMED) has become one of research hotspots. Image registration technology has been widely used in computer vision, medical image processing and material mechanics and many other fields. Combining various images and displaying their information on the same image to provide multi-data and multi-information images for clinical medical diagnosis has become a technology with great application value, the accurate and efficient image matching criterion is both a key and a difficult point. For different medical applications, the research and development

of special registration algorithms is an essential direction of development. At present, the two research directions of medical image registration are full automation and high precision. In order to avoid the over-reliance on imaging specialists in clinical practice and reduce their work intensity, image registration technology ranges from complex and laborious pre-imaging registration based on the positioning devices to a semi-automated approach to human-computer interaction, and then completely completed by the computer automatic registration, whose development is very rapid, the accuracy is gradually improved. The current registration accuracy has reached the "sub-pixel" level [1].

In order to improve the registration accuracy of Computed Tomography (CT) and Magnetic Resonance Imaging (MR), which are the two most common medical images, focusing on the abdomen, this essay proposes an indirect multimodal image registration method, which converts a multimodal image registration problem into a multiple single-modal image registration problem. The method introduces the great similarity and high precision of the single-modal image registration into the multimodal image registration. Analog CT is introduced as the intermediate medium. The MR image to be registered generates an analog CT image to be registered, and the reference image is still a CT image. The method is a quasi-single-mode image registration, where the analog CT image replaces the original MR image in the traditional registration mode as the image to be registered, and the CT image is taken as the reference image. At the same time, a semi-supervised classification model (SSC-EKE) with maximum knowledge utilization ability is introduced to generate Analog CT [2, 3]. The voxels generated by this method can correspond to the that in MRI, which is helpful to improve the registration accuracy.

In clinical diagnosis and treatment of diseases, image fusion can obtain an image that shows lesions more clearly and includes more medical information of the image [4]. Medical image fusion technology plays an important role in the localization of lesions, the diagnosis and analysis of disease, the formulation of treatment plan, and the recording and research of later pathology [5]. Therefore, since the image fusion was proposed, it has become one of hotspots of research and exploration in the field of medical imaging. In this study, we adopt the spatial domain fusion method, since it is intuitive and easy to implement.

The main content distribution of this essay is as follows:

The second part mainly introduces the related work, the third and fourth parts mainly introduce the research process and results, and then the fifth part is the summary of the proposed method.

2   Related Work

2.1   Laplace Support Vector Machine (LAPSVM)

The learning performance of traditional SVM depends on the quality and quantity of training samples to a great extent. When the training label data is insufficient but the quantity of non-label data is sufficient, the classification accuracy is low. Transforming the geometric structure information of edge distribution of labeled data and unlabeled data into manifold regular terms and adding them to the traditional classification supervision algorithm SVM, LAPSVM algorithm extends SVM to a semi-supervised learning algorithm.

Let be a training dataset, which contains n labeled data and n unlabeled data. The label corresponding to the tag data is denoted as y_i. The popular regularization framework can be expressed as:

$$\min_{f \in H_K} (\frac{1}{l} \sum_{i=1}^{l} V(x_i, y_i, f(x_i)) + \gamma_A \|f\|_K^2 + \gamma_I \|f\|_I^2) \tag{1}$$

where $V(\cdot)$ is the loss function, γ_A and γ_I are the coefficients of the two regularization terms.

The first term in Eq. (1) controls the empirical risk expressed as a hinge loss function, and the second item avoids over-fitting by applying a smoothing condition to the possible solutions in the regenerated kernel Hilbert space R KHS. Fitting problem, the third term is based on manifold learning to exploit the inherent geometric distribution of all data strengths and bases on manifold learning. Use the adjacency data graph to $G = (w, f)$ characterize the intrinsic manifold of the data distribution:

$$\|f\|_T^2 = \frac{1}{(u+l)^2} \sum_{i,j=1}^{l+u} (f(x_i) - f(x_j))^2 W_{ij} = \frac{1}{(u+l)^2} f^T L f \tag{2}$$

At this point $f = [f(x_1), \ldots, f(x_{l+u})]^T$, $W_{ij} \in W$, $i, j = 1, \ldots, u+1$, it represents the edge weights in the data adjacency graph. $L = S - W$ Refers to the Laplacian plot, where D represents the degree matrix, D the mid-diagonal elements $D_{ii} = \sum_{j=1}^{l+u} W_{ij}$, and the rest are 0 in matrix D.

From formula (1), the framework of Laplacian support vector machine (LapSVM) can be obtained:

$$\min_{f \in H_K} (\frac{1}{l} \sum_{i=1}^{l} (1 - y_i f(x_i)) + \gamma_A \|f\|_K^2 + \frac{\gamma_I}{(u+l)^2} f^T L f) \tag{3}$$

The solution of the Eq. (3) is $f^*(x) = \sum_{i=1}^{l+u} \alpha_i K(x, x_i)$, introducing slack variables ξ_i, $i = 1, \ldots, l$, Eq. (3) can be rewritten as:

$$\min_{\alpha \in R^{l+u}, \xi_i \in R} (\frac{1}{l} \sum_{i=1}^{l} \xi_i + \gamma_A \alpha^T K \alpha + \frac{\gamma_I}{(u+l)^2} \alpha^T K L K \alpha),$$

$$s.t. \ y_i(\sum_{j=1}^{l+u} \alpha_j K(x_i, x_j) + b) \geq 1 - \xi_i, i = 1, \ldots, l, \xi_i \geq 0, i = 1, \ldots, l. \tag{4}$$

Based on the Karush-Kuhn-Tucker (KKT) condition, the dual condition of Eq. (4) is as follow:

$$\max_{\beta \in R^l}(\frac{1}{l} \sum_{i=1}^{l} \beta_i - \frac{1}{2} \beta^T Q \beta), s.t. \sum_{i=1}^{l} \beta_i y_i = 0, 0 \leq \beta_i \leq \frac{1}{l}, i = 1, \ldots, l \tag{5}$$

where $Q = YJK(2\gamma_A I + (2\gamma_I/(u+l)^2 LK)^{-1} J^T Y$, $J = [I \ 0]$ is a $l \times (l+u)$ matrix of size $l \times (l+u)$, I is the identity matrix, $Y = diag(y_1, y_2, \ldots, y_l)$ and K is $(l+u) \times (l+u)$ the kernel matrix with the size of $(l+u) \times (l+u)$.

According to formula (5), it can be obtained that the solution of formula (4) is as follow:

$$\alpha^* = \left(2\gamma_A I + (2\gamma_I/(u+l)^2)LK\right)^{-1} J^T Y \beta^*.$$

2.2 Semi-supervised Classification Maximum Knowledge Utilization (SSC-EKE)

Semi-Supervised Classification Algorithm with Maximum Knowledge Utilization (SSC-EKE) is a semi-supervised learning algorithm [3], which effectively combines the manifold regularization mechanism of Laplacian support vector and fully mining useful knowledge in samples and a large number of unlabeled data samples to improve the classification performance, and achieved remarkable results.

On the basis of LAPSVM, SSC-EKE uses MS and CS to represent two collections----must-link and cannot-link, and then defines a matrix that is as follow:

$$Q_{ij} = Q_{ji} = \begin{cases} 1/|MS|, & \forall \langle i,j \rangle \in MS \text{ or } \langle j,i \rangle \in MS; \\ -1/|CS|, & \forall \langle i,j \rangle \in CS \text{ or } \langle j,i \rangle \in CS; \\ 0 & \text{default,} \end{cases}$$

After combining this matrix with the LAPSVM part, the definition of the joint regularization formula for the manifold and pairwise constraints can be obtained:

$$\min_{f}(\Phi_{MPCJFR}(f) = (1-\tau)f^T L' f + \tau f^T Z' f = f^T ((1-\tau)L' + \tau Z')f) \qquad (6)$$

where $Z = H - Q$ and $H = diag(Q \cdot 1_{(l+u)\times 1})$, similar to the definition of L in LapSVM, so in this formula, the range of values for $f^T L' f$ and $f^T Z' f$ are same, $\tau \in 0,1)$ is the dimension trade-off coefficient, which can control their individual significance in any data scenario.

Following the SVM minimum risk structure, the expression definition of SSC-EKE can be obtained from formula (6):

$$\min_{f \in H_k} \left(\frac{1}{l} \sum_{i=1}^{1} (1 - y_i f(x_i))_+ + \gamma_A f_k^2 + \gamma_J f^T ((1-\tau)L' + \tau Z')f \right) \qquad (7)$$

On the basis of Eq. (7), a bias term b is Introduced, so that Eq. (4) can be rewritten as:

$$\min_{\alpha \in R^{l+u}, \xi_i \in R} \left(\frac{1}{l} \sum_{i=1}^{l} \xi_i + \gamma_A \alpha^T K \alpha + \gamma_J \alpha^T K((1-\tau)L' + \tau Z')K \partial \right),$$

$$s.t. \ y_i \left(\sum_{j=1}^{l+u} \alpha_j K(x_i, x_j) + b \right) \geq 1 - \xi_i, i = 1, \ldots, l, \qquad (8)$$

Lagrange multiplier is expressed by, $\beta = (\beta_1, \ldots, \beta_l)$, and then according to the KKT condition, the dual form of formula (8) is as follow:

$$\min_{\beta \in R^l} \left(\sum_{i=1}^{l} \beta_i - \frac{1}{4}\beta^T S \beta \right),$$

$$s.t. \ \sum_{i=1}^{l} \beta_i y_i = 0, 0 \leq \beta_i \leq \frac{1}{l}, \quad i = 1, \ldots, l \qquad (9)$$

According to the solution of Eq. (9), the initial solution of Eq. (8) can be expressed as:

$$\alpha^* = \frac{1}{2}\left(\gamma_A K + \gamma_J K\left((1-\tau)L' + \tau Z'\right)K\right)^{-1} P\beta^*, \tag{10}$$

$$b^* = \frac{1}{l}\sum_{i=1}^{l}\left(y_i - \sum_{j=1}^{l+u}\alpha_j^* K(x_i, x_j)\right), \tag{11}$$

Ultimately, the final form of the classification decision function of the S SC-EKE algorithm is as follow:

$$f^*(x) = \sum_{i=1}^{l+u}\alpha_i^* K(x_i, x_j) + b^*$$

2.3 Medical Image Registration Based on Phase Difference

In order to achieve the effect of registration without changing the overall topological information of the image, we use the phase difference between multiple CT image data to minimize the objective function in the differential homeomorphic space, so as to improve its accuracy and improve the effect of image registration for large deformation. The phase difference is the wavefront aberration extracted from several different images by using the wavefront detection technology.

This algorithm is mainly divided into three steps:

Deformation field calculation: transform the phase difference between the floating image and the reference image into deformation field by using orthogonal filter.

(1) Deformation field calculation: Transform the phase difference between the floating image and the reference image into deformation field by using orthogonal filter. During each iteration, updating the deformation field (D_u) is through the function (θ) calculated from the floating image (m) driven by the reference image (f) and the deformation field (D_a), Δa and Δu represents deformation field D_a, D_u. The calculation form is:

$$D_u \leftarrow \theta(f, m \circ \Delta a)D_a$$

(2) Deformation field superposition: The differential homeomorphism algorithm is used to map the deformation field exponentially and add the new deformation field to the original deformation field. After calculating the deformation field, increase the total displacement D_a by updating the field D_u:

$$D_a \leftarrow \Phi(D_a, D_u)$$

(3) Deformation field regularization: The main purpose is to obtain smooth transformation and reduce the influence of image noise on the registration output. Its mathematical expression is: $D_a \leftarrow \Psi(D_a)$. Operation Ψ is achieved by applying a low-pass filter to each component of the deformation field. The specific implementation is carried out by using the convolution of the deformation field through the Gaussian

kernel. At the same time, considering the deterministic mapping mentioned above, in order to ensure the high certainty of the position, the corresponding relationship between the displacement vector and its corresponding determined position is maintained, and the deterministic mapping is regularized as the deformation field mapping.

2.4 Pixel Weighted Average Image Fusion Algorithm

Multimodal medical image fusion is based on the registration of two or more different modes of medical images, which can achieve the purpose of synthesizing different modes of medical images [6].

The spatial domain fusion method is to directly process the pixel values of the image and fuse them in the gray space of the image pixels [7–9]. The spatial domain fusion methods suitable for medical image processing mainly include: the pixel gray value selection method, the pixel weighted average method [10], the pixel insertion method and the method based on Neural Network [4]. After comprehensive consideration of the effect and time of several algorithms, we choose to use the pixel weighted average method [5], which can process the fusion between different modal images in real time, and simply and intuitively express the characteristics of different modal images. The mathematical expression is as follows:

$$F(\text{i}, j) = w_1 A(\text{i}, j) + W_2(\text{i}, j)$$

$$w_1 + w_2 = 1$$

3 Experimental Details

3.1 Generate Analog CT

This article mainly discusses the accuracy of multimodal medical image registration with simulated CT as the intermediate medium, so we choose to use the current method with better effect to generate simulated CT of lower abdomen to generate simulated CT [8]. The following is a brief description of this method:

The first stage is to extract features from MR images. Because there is some noise in MR images, extracting feature data directly will have an unpredictable impact on subsequent experiments. Therefore, a filter matrix is set to smooth MR images to reduce the impact of noise. We obtained abdominal texture features from Dixon-water sequence, Dixon-fat sequence, Dixon-IP sequence and Dixon-OP sequence [11]. Considering the low bone and air signals in the mDixon sequence, the mesh generation strategy is used to obtain the location information. Finally, texture features and position information are combined to obtain seven dimensional feature vectors.

The second stage is to generate simulated CT. Read the circled data, adjust the super parameters by grid search and cross validation, and train eight SSC-EKE classifiers. After inputting the required image, about 400 points are extracted from it to obtain the average

probability. Combined with the seven dimensional characteristics of the image, all voxel tissue types are estimated by KNN nearest neighbor method, and the corresponding CT values are given. The values of bone, fat, air and soft tissue were set as 380, −98, 700 and 32 [12] respectively.

3.2 Image Registration

Based on the simulated CT image generated in Sect. 3.1, this method changes the floating image into simulated CT, and the reference image is still the real CT image. The phase difference between simulated CT images and CT images will produce corresponding deformation on MR images.

The first step is to transform the phase difference between analog CT and real CT into deformation field by using orthogonal filter.

The second step is to use the differential homeomorphism algorithm to map the deformation field exponentially and superimpose the new deformation field with the original deformation field.

The third step is to use the Gaussian kernel to regularize the deformation field by using the convolution of the deformation field, so as to obtain a smooth transformation and reduce the impact of noise.

4   Experimental Results

4.1   Environment Configuration

The experimental environment and configuration of the system are as follows: the operating system is Windows 11, the CPU model is AMD Ryzen 54600H, the CPU memory size is 16 GB, the GPU model is NVIDIA Geforce GTX 1650, and the GPU memory size is 11 GB. The project operation effect may be different under different configurations.

4.2   A Subsection Sample

This chapter focuses on measuring the performance of the proposed method through performance indicators of some internationally recognized technical indicators. The specific performance indicators are shown in the following figure (Table 1):

Table 1. Registration specifications.

Technical specifications	Metric values (reference)
LP (Local-Phase)	1.9874E+06
CC (Cross-correlation)	0.2808
MI (Mutual-information)	0.7124

The technical indicators used to evaluate the quality of the images after registration have the following specific meanings:

LP index: expressed as local phase difference, this index is used to describe the registration effect between two images. The lower the value is, the better the registration effect is.

CC index: indicates the degree of correlation between the image and the reference image and the degree of matching in the relative position between the images. It is also used to describe the registration effect between the image and the CT image, the closer a value is to 0, the better the registration.

MI index: indicates the mutual information between the image to be registered and the reference image, and is also used to describe the registration effect between the MR image and the CT image. The closer the value is to 1, the better the registration effect.

4.3 Evaluation of Results

We recruited 8 subjects using a protocol approved by the Cleveland Medical Center Institutional Review Board to obtain their lower abdomen as well as CT data. Using the known method to generate the corresponding simulated CT, and then based on the simulated CT multi-modality medical image registration, thus indirectly achieve the purpose of registration with CT.

The following table shows the three index values of this method and the traditional multimodal registration method, through observation, we can find that the registration method based on simulated CT can achieve smaller phase difference and maximum mutual information than that based on local phase difference and maximum mutual information. It can be seen that the registration method based on simulated CT is an effective registration method (Fig. 1).

sub	MI			CC			Local-phase		
	sCT based	local-phase based	MI based	sCT based	local-phase based	MI based	sCT based	local-phase based	MI based
1	0.8999	0.8856	0.8970	−0.0492	−0.0973	−0.0574	1.8482E+06	2.0750E+06	1.9769E+06
2	0.9112	0.8995	0.9097	−0.0405	−0.0759	−0.0480	1.8152E+06	1.9874E+06	1.9118E+06
3	0.8950	0.8836	0.8916	−0.0438	−0.0740	−0.0523	2.1804E+06	2.3242E+06	2.2672E+06
4	0.9037	0.8986	0.9025	−0.0502	−0.0595	−0.0471	1.9619E+06	2.0465E+06	2.1489E+06
5	0.9036	0.8916	0.8999	−0.0429	−0.0696	−0.0498	2.0608E+06	2.1319E+06	2.2813E+06
6	0.8970	0.8878	0.8931	−0.0504	−0.0640	−0.0561	2.1183E+06	2.2428E+06	2.2093E+06
7	0.9069	0.8962	0.9036	−0.0362	−0.0702	−0.0471	2.0154E+06	2.2432E+06	2.1325E+06
8	0.8808	0.8775	0.8761	−0.0562	−0.0753	−0.0650	2.0942E+06	2.3002E+06	2.2815E+06

Fig. 1. Algorithm performance index value.

From Fig. 2, we can find that the multimodal registration method based on simulated CT can more accurately identify the position, shape and tissue and organs of the bones compared with the two traditional registration methods mentioned above. Competently meet the registration requirements of the lower abdomen.

5   Conclusion

In this paper, we introduce simulated CT as an intermediate medium for multimodal registration of MR and CT images, and select two traditional registration methods for
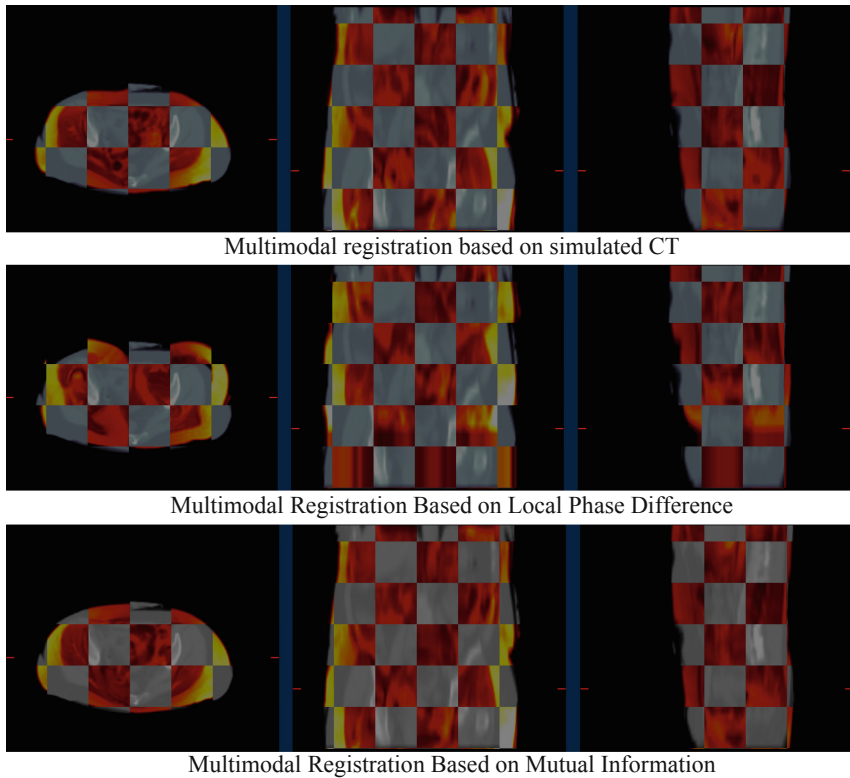
Multimodal registration based on simulated CT

Multimodal Registration Based on Local Phase Difference

Multimodal Registration Based on Mutual Information

Fig. 2. Comparison of registration results of each method.

comparison. The traditional multi-modal registration problem is transformed into single-modal registration, and the high-precision effect of single-modal registration is used to indirectly achieve high-precision registration between multi-modalities and obtain high-quality registration images. After comparing various fusion algorithms, we choose the pixel weighted average algorithm, which is simple, intuitive and easy to implement. It can realize the fusion of two images with different degrees of fusion, and the fusion effect is better.

References

1. Alafeef, M., Fraiwan, M.: On the diagnosis of idiopathic Parkinson's disease using continuous wavelet transform complex plot. J. Ambient Intell. Hum. Comput. **10**, 2805–2815 (2019). https://doi.org/10.1007/s12652-018-1014-x
2. Iizuka, S., Simo-Serra, E., Ishikawa, H.: Globally and locally consistent image completion. ACM Trans. Graph. (TOG) **36**, 1–14 (2017)
3. Guo, K., Cao, R., Kui, X., Ma, J., Kang, J., Chi, T.: LCC: towards efficient label completion and correction for supervised medical image learning in smart diagnosis. J. Netw. Comput. Appl. **133**, 51–59 (2019)

4. Raja, N.S.M., Fernandes, S.L., Dey, N., et al.: Contrast enhanced medical MRI evaluation using Tsallis entropy and region growing segmentation. J. Ambient Intell. Hum. Comput. (2018). https://doi.org/10.1007/s12652-018-0854-8
5. Alvén, J., Norlén, A., Enqvist, O., Kahl, F.: Überatlas: fast and robust registration for multi-atlas segmentation. Pattern Recogn. Lett. **80**, 249–255 (2016)
6. Fernandez-de-Manuel, L., et al.: Organ-focused mutual information for nonrigid multimodal registration of liver CT and Gd–EOB–DTPA-enhanced MRI. Med. Image Anal. **18**(1), 22–35 (2014)
7. Ahmad, S., Khan, M.F.: Multimodal non-rigid image registration based on elastodynamics. Vis. Comput. **34**(1), 21–27 (2018)
8. Qian, P., Xi, C., et al.: SSC-EKE: semi-supervised classification with extensive knowledge exploitation. Inf. Sci.: Int. J. **422**, 51–76 (2018)
9. Engelen, J.E.V., Hoos, H.H.: A survey on semi-supervised learning. Mach. Learn. **109**, 373–440 (2019)
10. Wald, L.: Some terms of reference in data fusion. IEEE Trans. Geosci. Remote Sens. **37**(3), 1190–1193 (1999)
11. Eggers, H., Brendel, B., Duijndam, A.: Dual-echo Dixon imaging with flexible choice of echo times. Magn. Reson. Med. **65**(1), 96–107 (2011)
12. Schneider, W., et al.: Correlation between CT numbers and tissue parameters needed for Monte Carlo simulations of clinical dose distributions. Phys. Med. Biol. **45**(2), 459–478 (2000). https://doi.org/10.1088/0031-9155/45/2/314