Lung CT Image Registration Using Robust Point Matching

Jianbing Yi[1,2], Xuan Yang[2], Guoliang Chen[2], and Yan-Ran Li[2]

[1] College of Information Engineering, Shenzhen University, Shenzhen, Guangdong 518000, China
[2] National High Performance Computing Center at Shenzhen, College of Computer Science and Software Engineering, Shenzhen University, Shenzhen, Guangdong 518000, China

1 Introduction

Respiratory motion causes significant uncertainties in image-guided radiotherapy procedures. Image registration techniques can be used to describe the temporal changes in the position and shape of tumors by aligning the source image to the target image. So the image registration of lung CT has recently attracted considerable interest from the clinical application community. Many registration algorithms of lung CT have been proposed, in order to compare the performance of the algorithms, the EMPIRE10 team organised a challenge. The authors registered to the challenge under the team name "HPC". A robust dynamic point matching algorithm based on intensity is proposed by authors to take part in the challenge, and the method is fully automatic. Our method contains a global affine transformation and a local elastic transformation. While the size of the moving image is much different from the fixed image, zoom in or zoom out is adopted to the moving image before the registration method is performed.

2 Methods

The correspondence of landmark points is determined from the distance between the points and the similarity between local images that are centered at points at the same time. To ensure that the points in the source image correspond to the points in the target image, the virtual target points are first created and shifted based on the similarity between the local image centered at the source point and the local image centered at the virtual target point. Second, the target points are shifted by the constrained inverse function mapping the target points to the virtual target points. The source point set and shifted target point set are used to estimate the transformation function between the source image and target image. Image registration is performed in three stages:

(1) Set the source point set X_0 extracted from the lung parenchyma I_{sp}, the target point set Y_0 extracted from the lung parenchyma I_{tp}, and the transformation function f_0 between X_0 and Y_0 using robust point matching with shifting the target points such that the mapped point set $f_0(X_0)$ can be obtained.

(2) Set the source point set X_1 extracted from the boundaries of the source lung mask I_{sm}, the target point set Y_1 extracted from the boundaries of the target lung mask I_{tm}, and the transformation function f_1 between X_1 and Y_1 using robust point matching with no shifting the target points such that the mapped point set $f_1(X_1)$ can be obtained.

(3) The transformation function f between $X = \{X_0, X_1\}$ and $Y = \{f_0(X_0), f_1(X_1)\}$ can be obtained using a thin-plate spline [1]. And in order to save computing time, while the intensity difference between the moving image and the fixed image is small, the second and third stages of our method can be ignored.

Image preprocessing: To remove the influence of the rib cage, the lung parenchymas are extracted from the lung CT image using lung masks, which are created by the EMPIRE10 organisers. To improve registration performance, zoom in or zoom out is adopted to the moving image while the size of the moving image is much different from the fixed image. And to get more useful landmark points, the blood vessels of the lung are enhanced using the vascular filtering algorithm proposed by Frangi [2], and then the vascular diffusion method proposed by Perona [3] is used to improve the legibility of the vascular structures and to reduce the background.

Cost function: In our method, I_s is selected as the source image, and I_t is selected as the target image. The point sets extracted from source image I_s and target image I_t are denoted by source point set X and target point set Y, respectively. Given the source point set $X = \{x_i : x_i \in R^3, i = 1, 2, ..., K\}$ and the target point set $Y = \{y_j : y_j \in R^3, j = 1, 2, ..., N\}$. Our goal is to estimate the correspondence between X and Y, and then to compute the transformation function f. The robust point matching model minimizes the following energy function:

$$\min_{M,f} \left\{ \sum_{j=1}^{N} \sum_{i=1}^{K} m_{ij} ||y_j - f(x_i)||_2^2 + \lambda ||Lf||_2^2 + T \sum_{j=1}^{N} \sum_{i=1}^{K} m_{ij} \log m_{ij} - \zeta \sum_{j=1}^{N} \sum_{i=1}^{K} m_{ij} \right\}, \tag{1}$$

where L is the Laplacian operator that denotes the smoothness of the spatial transformation, λ and ζ are the weighting parameters to balance these terms, T is the annealing temperature during the robust point matching.

Correspondence matrix: We define a fuzzy correspondence matrix $M \in R^{(K+1)\times(N+1)}$ between X and Y as

$$m_{ij} = \frac{1}{T} \exp[-\frac{\alpha ||y_j - f(x_i)||_2^2 + (1-\alpha)(1 - Corr(W(I_s, x_i), W(I_t, y_j)))}{2T}], \tag{2}$$

where T is the annealing temperature, α is the balance coefficient of two terms, f is the transformation function between two point sets, $||y_j - f(x_i)||_2^2$ is the distance between the target point y_j and the mapped source point $f(x_i)$, $W(I_s, x_i)$ and $W(I_t, y_j)$ are the local images centered at x_i in image I_s and at y_j in image I_t, respectively, and $Corr$ is the correlation function between the local images $W(I_s, x_i)$ and $W(I_t, y_j)$. When the local images $W(I_s, x_i)$ and $W(I_t, y_j)$ are similar, the term $Corr(W(I_s, x_i), W(I_t, y_j))$ is close to 1; when the the term $Corr(W(I_s, x_i), W(I_t, y_j))$ is small, the image intensity around x_i in the source image is clearly different from the image intensity around y_j in the target image.

Similarity measures: The similarity measures contain the point set shapes and local images that are centered at the points. The similarity measure of the point set shapes can be defined as

$$C = \sum_{j=1}^{N} \sum_{i=1}^{K} m_{ij} ||y_j - f(x_i)||_2^2, \tag{3}$$

where f is the transformation function between two point sets, $||y_j - f(x_i)||_2^2$ is the distance between the target point y_j and the mapped source point $f(x_i)$.

The similarity measure of local images that are centered at the points can be defined as

$$C'_{ij} = \frac{1}{T} \exp[-\frac{(1 - Corr(W(I_s, x_i), W(I_t, y_j)))}{2T}], \tag{4}$$

where $W(I_s, x_i)$ and $W(I_t, y_j)$ are the local images centered at x_i in image I_s and at y_j in image I_t, respectively, and $Corr$ is the correlation function between the local images $W(I_s, x_i)$ and $W(I_t, y_j)$. The similarity measures of local images that are centered at points directly influence the estimation of the fuzzy correspondence matrix M and indirectly influence minimizing the energy function.

Optimization Strategy: A general strategy for solving the minimization problem model (1) is to use **EM** method with two steps: 1) In **E** step, the correspondence matrix M is computed by the proposed equation (2) to determine the possibility of correspondence between the points. 2) In **M** step, By dropping the terms that are independent of f, and simplifying the distance $\sum_{j=1}^{N} m_{ij} ||y_j - f(x_i)||_2^2 \approx \| \sum_{j=1}^{N} m_{ij} y_j - f(x_i)\|_2^2$, equation (1) can be rewritten by

$$\min_f \left\{ \sum_{i=1}^{K} ||z_i - f(x_i)||_2^2 + \lambda ||Lf||_2^2 \right\}, \tag{5}$$

where $z_i = \sum_{j=1}^{N} m_{ij} y_j$ is the weighted target point, which can be treated as the virtual target point. The virtual target point set is defined as $Z = [z_1, \ldots, z_K]^\top$

and it satisfies with $Z = MY$. The spatial transformation f is obtained between the source points X and the virtual target points Z [4].

Shifting the virtual target points: The virtual target point is shifted according to the similarity between the local image that is centered at the virtual target point and the local image that is centered at the source point, which will minimize the deviation of the corresponding point position and improve the registration accuracy of the images. We denote the shifted given virtual target point z_i by $z_i' = z_i + \Delta z_i$, where Δz_i is the displacement of z_i. z_i' is selected as the point that minimizes the intensity difference of two local images $W(I_t, z_i')$ centered at z_i' in image I_t and $W(I_s, x_i)$ centered at x_i in image I_s. The search area of z_i' is limited to a local region centered at z_i with $\|\Delta z_i\|_\infty \leq \delta$ at each iteration. δ is the shifting range, it must be small and is set to be 1 in our experiment.

Shifting the target points: When only the virtual target point set Z is shifted without shifting the target point set Y, the transformation function f cannot be improved further. To address the above issue, we further shift the target point set Y based on the shift of the virtual target point set Z. Due to the virtual target point set obtained by $Z = MY$, when the virtual target point set Z is shifted to Z', Y should be shifted to Y' for satisfying with $Z' = MY'$. Denoting $Z' = Z + \Delta Z$ and $Y' = Y + \Delta Y$, the problem is changed to solve $\Delta Z = M\Delta Y$. we use the generalized inverse matrix of M to solve ΔY as

$$\Delta Y = \begin{cases} M^+ \Delta Z & \|\Delta Y\|_\infty \leq \delta \\ 0 & \text{otherwise} \end{cases}, \tag{6}$$

where M^+ is the Moore-Penrose inverse, which is a special type of a generalized inverse and is referred to as the pseudoinverse of matrix M. The parameter δ must be small as 1. The shifted Y' needs to be close to Y to preserve the object shape.

Model of transformation: Under the proposed robust dynamic point matching algorithm, the spatial transformation function f between the source point set X and the target point set Y is estimated by minimizing energy function using regularized thin-plate splines [1] with two regularized parameters λ_1 and λ_2, and the source point set X and the shifted virtual target point set Z' are selected as the corresponding point pair sets.

The parameter settings: The initial anneal temperature T_0 is set to the largest square distance of all point pairs. We generally set ρ to be 0.93 (normally between [0.9,0.99]) so that the annealing process is slow enough for the algorithm to be robust but not too slow. The regularization parameters λ_1 and λ_2 both have annealing schedules ($\lambda_i = \lambda_i^{init}\rho, i = 1, 2$). To provide more freedom for the affine transformation, λ_2^{init} is set to be considerably smaller than λ_1^{init}. While this approach works in practice, it is inelegant; thus, we generally set $\lambda_1^{init} = 3$, $\lambda_2^{init} = 0.03$ in our method [4]. And the window size of the similarity measure of local images that are centered at points is set to be $23 \times 23 \times 11$.

The details of the implementation :

Algorithm 1: Lung CT Image Registration Using Robust Point Matching

---

Input: the initial anneal temperature T_0, the annealing rate ρ, the regularized parameters λ_1 and λ_2, the source image I_s, and the target image I_t.
Output: the transformation function f.
begin
Step1: Extract the source point set X from the source image I_s;
Step2: Extract the target point set Y from the target image I_t;
Step3: Calculate the corresponding matrix M using equation (2);
Step4: Calculate the virtual target point set Z and shift it as $Z^{'}$;
Step5: Solve the transformation function f based on the correspondence between the source point set X and the shifted virtual target point set Z';
Step6: Update the target point set Y;
Step7: If the stop criterion(e.g., the distance between the mapping point sets $f(X)$ and Y under the special distance) is not satisfied, go to step 3; otherwise, output registration results.
end

---

3 Runtime

Extracting landmark points and calculating the correlation coefficient between the source point set and the target point set are performed on a GPU server with 512 CUDA cores and 16 GB of memory, the other experiments are performed on a computer with a 3.2-GHz quad-core processor and 10 GB of memory. On average the computation time of our method is approximately 125 minutes, of which 13 minutes are spent to the preprocessing.

References

[1] Fred L. Bookstein: Principal warps: thin-plate splines and the decomposition of deformations. IEEE Transactions on Pattern Analysis and Machine Intelligence 11 (1989) 567-585

[2] Alejandro F. Frangi, Wiro J. Niessen, Koen L. Vincken, Max A. Viergever: Multi-scale vessel enhancement filtering. Lecture Notes in Computer Science 1496 (1998) 130-137

[3] P. Perona, J. Malik: Scale-space and edge detection using anisotropic diffusion. IEEE Transactions on Pattern Analysis and Machine Intelligence 12 (1990) 629-639

[4] Haili Chui, Anand Rangarajan: A new point matching algorithm for non-rigid registration. Computer Vision and Image Understanding 89 (2003) 114-141