

# Medical Image registration using sparse coding and belief propagation

Aminmohammad Roozgard, Nafise Barzigar, Samuel Cheng and Pramode Verma

**Abstract**—Recently, various medical imaging such as CT and MRI imaging has been used more and more widely in clinical and medical research. As a result, there is an increasing interest in accurately relating information in different images for diagnosis, treatment, and the sake of basic science. As images are typically acquired at different times and often by different modalities, registering (or aligning) one image with another is not a simple task in general and its success will affect the effectiveness and accuracy of all subsequent analysis. We propose an efficient medical image registration method based on sparse coding and belief propagation for medical CT imaging. We used 3-D image blocks as features, and then we employed sparse coding to find a set of candidate voxels. To select optimum matches, belief propagation was subsequently applied on these candidate voxels. The outcome of belief propagation was interpreted as probabilistic map between candidate voxels and source voxel. We compared with the state-of-the-art of medical image registration, MIRT [1] and GP-Registration algorithm [2]. Our objective results based on RMSE (Root Mean Square Error) are smaller than those from MIRT and GP-Registration. Our results also proved the effectiveness of our algorithm in registering reference image to source image.

## I. INTRODUCTION

Image registration is the process of overlaying two or more images using different capturing modules into the same coordination system [3]. Registration is required in many clinical applications including diagnosis and surgical planning. For example, registration techniques have been used to align an MRI to a CT image [4]. Images of similar or differing modalities need to be aligned as a pre-processing step for many planning, navigation, detection, data-fusion and visualization tasks in medical applications [5], [6]. Medical image registration still presents many challenges. Several notable difficulties are the nonlinear transformation between images and different appearance or resolution of images.

Many medical image registration methods have been developed in the last decade [1]–[8], and they are divided into two major categories such as direct and feature-based matching. Direct methods use all available image data, and they result in very accurate registration if initialization points that are close to target points at the start of the registration procedure are available [7]. For example, in [2], a general-purpose registration algorithm for the medical images has been developed, which incorporates both geometric and intensity transformation. The geometric model assumes a locally affine and globally smooth transformation. The

intensity model accounts for local differences in contrast and brightness while imposing a global smoothness on the overall intensity differences. Also, Myronenko and Song used the definition of the similarity measure to propose a registration method [1]. They derived the similarity measure by analytically solving for the intensity correction field and its adaptive regularization. The final measure was interpreted as one that favors a registration with minimum compression complexity of the residual image between the two registered images. Feature-based registration methods, utilize invariant features (especially those around Harris corners) to ensure reliable matching. As a result, feature-based methods do not depend on initialization point [9]. In [8], Glocker *et al.* used different levels of smoothness in modeling medical images. The authors then used Markov Random Fields (MRFs) to formulate image deformations.

Many researchers incorporate smoothness (or spatial coherence) conditions by reformulating matching into an optimization problem [10]. For example, in [11], each pixel in the reference image is assigned a vector displacement label indicating the position in the test image to which it spatially corresponds. To penalize sharp changes in displacement labels across pixels, a smoothness constraint based on the first derivative was used. Graph cuts method was then used to solve that labeling problem. Also in [10], belief propagation was used to optimize cost function incorporated with smoothness constraints which encourage similar displacements of near-by pixels.

In this paper, we propose a dense, registration technique by aligning local 3-D features of two CT images using sparse coding and belief propagation. First, we build an overcomplete dictionary out of all 3-D features of a reference CT image [12]. Note that since the dictionary is constructed by padding the features directly, no extra time is spent on training. We then find a set of candidate voxels for each voxel of the source image using sparse coding out of the constructed dictionary. The match score of each candidate voxel will be evaluated taking both local and neighboring information into account using belief propagation [13]. The best match will be selected as the candidate with highest score.

The rest of the paper is organized as follows. In the next section, we will introduce the concept of our 3D-SCoBeP and the inference algorithm. In Section III, we will show our simulation results, followed by a brief conclusion in Section IV.

\*This work was supported in part by Oklahoma Tobacco Research Center under a seed grant.

The authors are with the Department of Electrical and Computer Engineering, Oklahoma University, Tulsa, OK, 74135.  
{roozgard, barzigar, samuel.cheng, pverma}@ou.edu

## II. PROPOSED METHOD

For each voxel in the source CT image, we select from the reference CT image a set of  $n$  candidate voxels which are likely to be most similar to the target voxel.

The proposed method described here is inspired by our recent work, SCoBeP [14]. First, we extract the features from the source CT image and the reference CT image. In this paper, we focus ourselves on only using 3-D block features even though the proposed approach can generally be applied to other features (such as SIFT-features). Thus, each feature considered here is essentially a vectorized 3-D block centered around a voxel in an CT image. Second, to match the extracted features of the reference CT image to the corresponding extracted features of the source CT image, we create a dictionary which contains all feature vectors of the reference CT image and apply sparse coding to each extracted features of the source CT image. Sparse coding will reconstruct a 3-D source patch at voxel  $[i, j, k]$  as a linear combination of reference 3-D patches. Note that the obtained sparse coefficient vector should be sparse, i.e., it should be 0 for most coefficients. To select the  $n$  candidate voxels, we simply pick those corresponding to  $n$  largest coefficients in the sparse coefficient vector. We denote a set as an  $n \times 2$  matrix storing the locations of these candidate voxels and a probability vector as the length- $n$  vector storing the corresponding values of the sparse coefficient vector. Each coefficient in the probability vector serves as a prior probability of matching the 3-D source patch at  $[i, j, k]$  to a 3-D reference patch of the reference CT image taking only local characteristics into accounts but ignoring geometric characteristics of the matches. Finally, to incorporate geometric characteristics, we model the problem by a factor graph and apply belief propagation to update probabilities (for more details, see [14]).

### A. Implementation

As mentioned in Section I, in some applications especially medical imaging we need dense registration so that for each voxel of the first image (source CT image) a corresponding match voxel will be found on the second image (reference CT image). This section describes the implementation's details of our proposed registration method. The main procedure for our proposed denoising method is summarized in Algorithm 1.

#### Implementation Details:

- $\mathbf{Y} = PatchExtractor(\{\mathcal{Y}_s\}_{s=1}^{k_y})$  presents a patch extractor algorithm using  $\{\mathcal{Y}_s\}_{s=1}^{k_y}$  as a source CT image, where the results is a 4-D matrix containing the vectorized 3-D blocks. To achieve this purpose, we consider a 3-D block of size  $(2a+1) \times (2b+1) \times (2c+1)$  containing neighboring voxels around each voxel on 3-D CT image, where  $a$ ,  $b$  and  $c$  are positive integers. For each voxel  $p_{i,j,k}$  in the source CT image  $\{\mathcal{Y}_s\}_{s=1}^{k_y}$ , we vectorized the patch of  $p_{i,j,k}$  to a feature vector  $y_{i,j,k} \in \mathbb{R}^{S \times 1}$ , where  $S = (2a+1) \times (2b+1) \times (2c+1)$ .

---

**Algorithm 1** 3D-SCoBeP for medical image registration-estimate version of registered CT image  $\mathcal{Z}$

---

**Inputs :** reference CT image  $\{\mathcal{X}_s\}_{s=1}^{k_x} \in \mathbb{R}^{M \times N \times K}$ , source CT image  $\{\mathcal{Y}_s\}_{s=1}^{k_y} \in \mathbb{R}^{M \times N \times K}$ , reference CT image slices number  $k_x$ , source CT image slices number  $k_y$ , candidate voxels number  $n$

**Extract 3-D dense feature and construct dictionary :**

- $\mathbf{Y} = PatchExtractor(\{\mathcal{Y}_s\}_{s=1}^{k_y})$
- $\mathbf{X} = PatchExtractor(\{\mathcal{X}_s\}_{s=1}^{k_x})$
- $\mathcal{D} = MakeDic(\mathbf{X})$

**Find the initial estimate of candidate voxels:** For each vector  $y_{i,j,k} \in \mathbf{Y}$  do:

- $\check{\alpha}_{i,j,k} = FindSCV(\mathcal{D}, y_{i,j,k})$
- $\mathcal{L}_{i,j,k} = TopSCV(\check{\alpha}_{i,j,k}, n)$
- $\rho_{i,j,k} = Prob(\mathcal{L}_{i,j,k})$

**Refine the candidate voxels:** For each vector  $y_{i,j,k} \in \mathbf{Y}$  do:

- $\check{\rho} = BP(\mathcal{L}, \rho)$

**Output :** a probabilistic map between the reference voxels and the source voxels

---

A 4-D source feature image  $\mathbf{Y} \in \mathbb{R}^{M \times N \times K \times S}$  is then constructed from  $y_{i,j,k}$  as follows

$$\mathbf{Y} = \{y_{i,j,k} \mid 1 \leq i \leq M, 1 \leq j \leq N, 1 \leq k \leq K\}. \quad (\text{II.1})$$

Note that,  $\mathbf{X}$  is created in the same manner as  $\mathbf{Y}$  but from reference CT image  $\{\mathcal{X}_s\}_{s=1}^{k_x}$  instead.

- $\mathcal{D} = MakeDic(\mathbf{X})$  creates a dictionary using the vectors in  $\mathbf{X}$ . To match the extracted features of the source CT image to corresponding extracted features of the reference CT image, a dictionary which contains feature vectors of  $\mathbf{X}$  is constructed. Thus, we can write  $\mathcal{D}$  as

$$\mathcal{D} = [x_{1,1,1} \dots x_{1,1,K} \ x_{1,2,K} \dots x_{1,N,K} \dots x_{M,N,K}], \quad (\text{II.2})$$

where  $x_{i,j,k}$  is a feature vector in  $\mathbf{X}$ . Note that, we normalize dictionary  $\mathcal{D}$  to guarantee the norm of each feature vector to be 1.

- $\check{\alpha}_{i,j,k} = FindSCV(\mathcal{D}, y_{i,j,k})$  finds candidate match voxels using sparse coding algorithm, where the result is the sparse coefficient vector  $\check{\alpha}_{i,j,k}$ . Mathematically, we try to solve the following sparse coding problem of finding the most sparse coefficient vectors  $\check{\alpha}_{i,j,k}$  such that

$$y_{i,j,k} = \mathcal{D}\check{\alpha}_{i,j,k}. \quad (\text{II.3})$$

Although there are several methods to solve (II.3) [15]–[17], in our work, we employ Subspace Pursuit (SP) [16] because of its computational efficiency.

- $\mathcal{L}_{i,j,k} = TopSCV(\check{\alpha}_{i,j,k}, n)$  simply picks up the  $n$  largest coefficients of  $\check{\alpha}_{i,j,k}$  as  $n$  candidates.  $\mathcal{L}_{i,j,k}$  as an  $n \times 2$  matrix stores the locations of these

candidate voxels and  $\rho_{i,j,k}$  (in  $\rho_{i,j,k} = Prob(\mathcal{L}_{i,j,k})$ ) as the length- $n$  vector stores the corresponding values of  $\mathcal{L}_{i,j,k}$ . Each coefficient in  $\rho_{i,j,k}$  serves as a prior probability of matching the source patch at  $[i, j, k]$  to a patch of  $x_{i,j,k}$  taking only local characteristics into accounts but ignoring geometric characteristics of the matches.

- $\tilde{\rho} = BP(\mathcal{L}, \rho)$  models the problem by a factor graph and apply belief propagation [13] to update probabilities  $\rho$  (for more details, see [14]). The updated probabilities  $\tilde{\rho}$  can be used for the registration of the source CT image. In our case, we assign a variable node for each voxel on the source CT image and connect each pair of neighboring voxels with a factor node. Also, we introduce one extra factor node to take care of the prior knowledge obtained in the sparse coding step for each voxel of the source CT image.

### III. EXPERIMENTAL RESULTS

In this section, we present various experiments to evaluate 3D-SCoBeP. We considered the problem of registering two slices of two CT images of one person from two different times. To evaluate the performance of our approach, we conducted tests on the data sets LIDC-IDRI [18] where the size of each slice of the CT images are  $512 \times 512$  pixels. Through out the experiments, the following parameters were used: the number of candidate voxels  $n$  is set to be 4,  $a = b = 3$  and  $c = 2$ . To synthesize the source image, we replaced each voxel of the source CT image with the selected candidate voxel from the reference CT image. In other words, we map the reference CT image onto the source image using the updated probabilities and the candidate voxels location. In our work, we select the most probable voxel after the BP step as the best match voxel. We assume that our registration method successfully finds a match for an input voxel if the most probable candidate has belief larger than a threshold  $\theta = 0.25$ . Otherwise, we assume no best match is found.

We now proceed to compare 3D-SCoBeP with other approaches; Figs. III.1 and III.2 show the output of our proposed method compared to two the state-of-the-art methods; the MIRT [1] and GP-Registration [2]. Figs. III.1(a) and III.2(a) corresponds the reference CT image and Figs. III.1(b) and III.2(b) to the source CT image. Figs. III.1(c)-(e) and III.2(c)-(e) show results using MIRT [1], GP-Registration [2] and our proposed method. The warped images using MIRT and GP-Registration with highlighted artifacts are shown in (h) and (i), respectively. The estimated images generated from 3D-SCoBeP with highlighted areas are shown in (j).

To quantify our registration performance, we used the root mean square error (RMSE) measure between the true and estimated transformations:

$$\varepsilon_{RMSE} = \sqrt{(1/N) \sum \|\tau - \hat{\tau}\|^2}, \quad (III.1)$$

where  $N$  is the number of voxels in the reference and  $\tau$  and  $\hat{\tau}$  are the source image and the estimated transformation respectively. However, the root mean square error (RMSE)

can not qualify the accuracy of the registration methods perfectly. It can only give a rough estimation of similarity between estimated image and source image. In the term of RMSE, we compare the source image with the output of 3D-SCoBeP, MIRT and GP-Registration and the results is shown under (c), (d) and (e) of Figs. III.1 and III.2.

### IV. CONCLUSIONS

In conclusion, we have proposed in this paper an efficient registration method based on a sparse coding and belief propagation. Our technique performs registration by first running sparse coding over an overcomplete dictionary constructed from the reference image to gather possible match candidates. Belief propagation is then applied to eliminate bad candidates and to select optimum matches. The experimental result illustrates that our proposed algorithm compares favorably with the high accuracy MIRT method by Myronenko and Song [1] and the state-of-the-art GP-Registration by Periaswamy and Farid [2].

### REFERENCES

- [1] A. Myronenko and X. Song, "Intensity-based image registration by minimizing residual complexity," *Medical Imaging, IEEE Transactions on*, vol. 29, no. 11, pp. 1882–1891, 2010.
- [2] S. Periaswamy and H. Farid, "Elastic registration in the presence of intensity variations," *Medical Imaging, IEEE Transactions on*, vol. 22, no. 7, pp. 865–874, 2003.
- [3] B. Zitova and J. Flusser, "Image registration methods: a survey," *Image and vision computing*, vol. 21, no. 11, pp. 977–1000, 2003.
- [4] J. Schnabel, D. Rueckert, *et al.*, "A generic framework for non-rigid registration based on non-uniform multi-level free-form deformations," in *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2001*. Springer, 2010, pp. 573–581.
- [5] J. Maintz and M. Viergever, "A survey of medical image registration," *Medical image analysis*, vol. 2, no. 1, pp. 1–36, 1998.
- [6] A. Roozgard, S. Cheng, and H. Liu, "Malignant nodule detection on lung ct scan images with kernel rx-algorithm," in *IEEE-EMBS International Conference on Biomedical and Health Informatics*, Hong Kong, Shenzhen, 2012.
- [7] H. Shum and R. Szeliski, "Systems and experiment paper: Construction of panoramic image mosaics with global and local alignment," *International Journal of Computer Vision*, vol. 36, no. 2, pp. 101–130, 2000.
- [8] B. Glocker, N. Komodakis, G. Tziritas, N. Navab, and N. Paragios, "Dense image registration through mrf's and efficient linear programming," *Medical Image Analysis*, vol. 12, no. 6, pp. 731–741, 2008.
- [9] M. Brown and D. Lowe, "Automatic panoramic image stitching using invariant features," *International Journal of Computer Vision*, vol. 74, no. 1, pp. 59–73, 2007.
- [10] C. Liu, J. Yuen, and A. Torralba, "Sift flow: Dense correspondence across scenes and its applications," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 33, no. 5, pp. 978–994, 2011.
- [11] T. Tang and A. Chung, "Non-rigid image registration using graph-cuts," in *Proceedings of the 10th international conference on Medical image computing and computer-assisted intervention—Volume Part I*. Springer-Verlag, 2007, pp. 916–924.
- [12] J. Mairal, F. Bach, J. Ponce, and G. Sapiro, "Online learning for matrix factorization and sparse coding," *The Journal of Machine Learning Research*, vol. 11, pp. 19–60, 2010.
- [13] F. Kschischang, B. Frey, and H. Loeliger, "Factor graphs and the sum-product algorithm," *Information Theory, IEEE Transactions on*, vol. 47, no. 2, pp. 498–519, 2001.
- [14] A. Roozgard, N. Barzigar, S. Cheng, and P. Verma, "Dense image registration using sparse coding and belief propagation," in *Signal Processing and Communication Systems (ICSPCS), 2011 5th International Conference on*. IEEE, 2011, pp. 1–5.

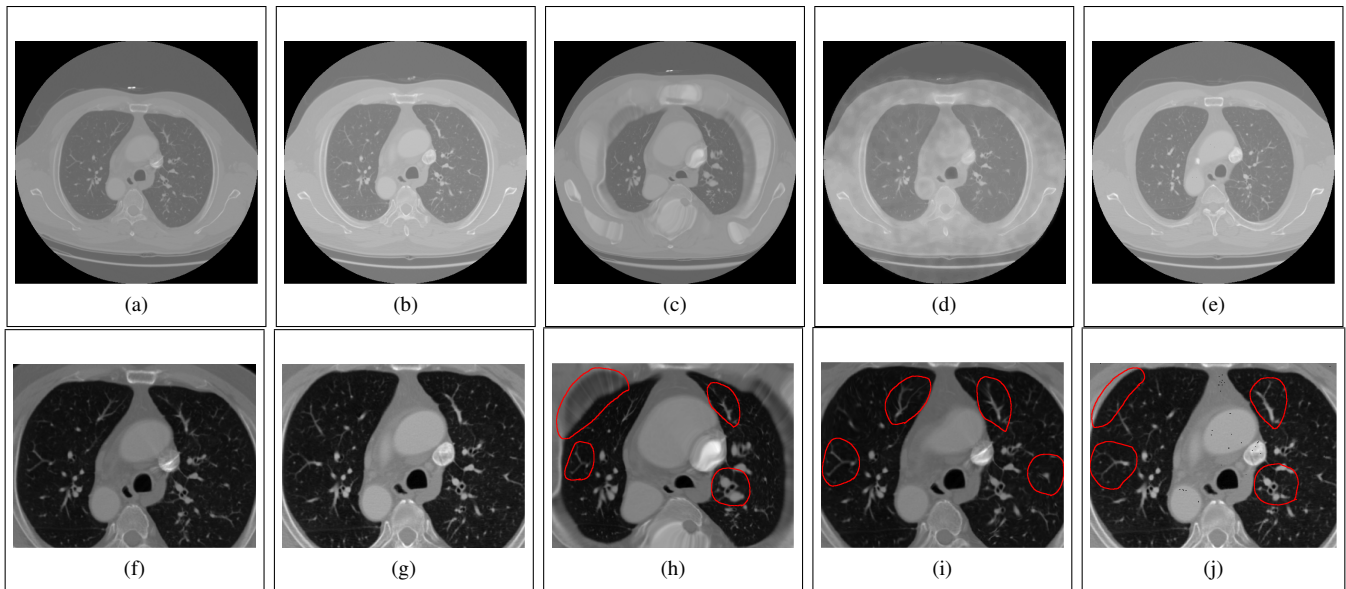


Fig. III.1: Synthetic experiment 1. We register the reference image (a) onto the source image (b) (RMSE in brackets). (a) Source image; (b) Reference image; (c) MIRT [1] [RMSE: 24.3060]; (d) GP-Registration [2] [RMSE: 28.7820]; (e) 3D-SCoBeP [RMSE: 21.9342]; (f) Source image (Zoom in); (g) Reference image (Zoom in); (h) MIRT [1] (zoom in); (i) GP-Registration [2] (zoom in); (j) 3D-SCoBeP (zoom in).

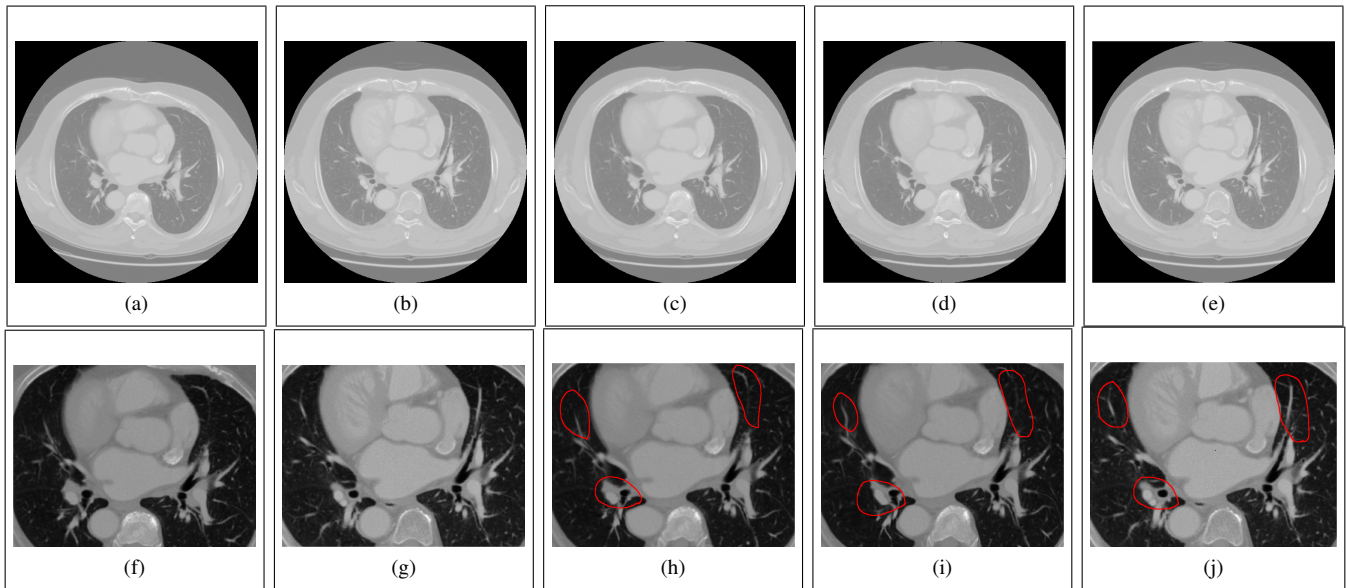


Fig. III.2: Synthetic experiment 2. We register the reference image (a) onto the source image (b) (RMSE in brackets). (a) Source image; (b) Reference image; (c) MIRT [1] [RMSE: 7.7130]; (d) GP-Registration [2] [RMSE: 7.3828]; (e) 3D-SCoBeP [RMSE: 4.2334]; (f) Source image (Zoom in); (g) Reference image (Zoom in); (h) MIRT [1] (zoom in); (i) GP-Registration [2] (zoom in); (j) 3D-SCoBeP (zoom in).

- [15] A. Yang, S. Sastry, A. Ganesh, and Y. Ma, "Fast 1-minimization algorithms and an application in robust face recognition: A review," in *Image Processing (ICIP), 2010 17th IEEE International Conference on*. IEEE, 2010, pp. 1849–1852.
- [16] W. Dai and O. Milenkovic, "Subspace pursuit for compressive sensing signal reconstruction," *Information Theory, IEEE Transactions on*, vol. 55, no. 5, pp. 2230–2249, 2009.
- [17] R. Maleh, A. Gilbert, and M. Strauss, "Sparse gradient image reconstruction done faster," in *Image Processing, 2007. ICIP 2007. IEEE International Conference on*, vol. 2. IEEE, 2007, pp. II–77.
- [18] N. C. Institute, "The cancer imaging archive,"

<https://wiki.cancerimagingarchive.net/display/Public/LIDC-IDRI>, September 2011. [Online]. Available: <https://wiki.cancerimagingarchive.net/display/Public/LIDC-IDRI>