

ORIGINAL ARTICLES

3D-SCoBeP: 3D medical image registration using sparse coding and belief propagation

Aminmohammad Roozgard, Nafise Barzigar, Pramode Verma, Samuel Cheng

University of Oklahoma, Tulsa, USA.

Correspondence: Aminmohammad Roozgard. Address: University of Oklahoma, Tulsa, USA.
Email: roozgard@gmail.com

Received: May 30, 2014

Accepted: November 16, 2014

Online Published: December 30, 2014

DOI: 10.5430/ijdi.v2n1p54

URL: <http://dx.doi.org/10.5430/ijdi.v2n1p54>

Abstract

There are various medical imaging methods which have been used broadly in clinical and medical research. Consequently, the interests in registering and finding similarities of different images for diagnosis, treatment, and the sake of basic science are increasing. As images are typically captured at different times, angles, and often by different modalities, registering (or aligning) one image with another is challenging. In general, the accuracy of registration techniques will affect the performance and robustness of all subsequent analysis. We propose an efficient 3D medical image registration method based on sparse coding and belief propagation for Computed Tomography (CT) and Magnetic Resonance (MR) imaging. We used 3D image blocks as the input features and then we employed sparse coding with a dictionary of the features to find a set of the candidate voxels. To select optimum matches, belief propagation was subsequently applied on a factor graph of voxels generated by these candidate voxels. The outcome of belief propagation was interpreted as a probabilistic map of aligning the candidate voxels to the source voxels. We compared our proposed method (3D-SCoBeP) with the state-of-the-art medical image registration, MIRT and GP-Registration algorithm. Our objective results based on Root Mean Square Error (RMSE) are smaller than those from MIRT and GP-Registration. Our results prove the effectiveness of our algorithm in registering the reference image to the source image.

Keywords

Image registration, Sparse coding, Belief propagation, Graph-cuts

1 Introduction

Image registration refers to the process of aligning two or more images obtained from different capturing modules and/or angles, and/or at different times into the same coordination system ^[1]. Registration is essential in many clinical applications including diagnosis ^[2], simulating and surgical planning ^[3]. For example, registration techniques have been used to align a Magnetic Resonance (MR) image to a Computer Tomography (CT) image ^[4,5]. In surgery, radiotherapy, or radiological intervention, preoperative medical data are used to diagnose, plan, simulate, guide, or otherwise assist a surgeon, or possibly a robot ^[6]. While the surgical procedure is performed in the coordinate system relative to the patient, the surgical plan is constructed in the coordinate system relative to the preoperative data. The spatial transformation between the plan and the preoperative data is formed by registration. Registration as a central step of processing images in the treatments, allows any voxel defined in the preoperative image to be precisely located in the patient coordinate system.

This can aid the surgeon by delineating the position of the surgical instruments relative to the ultimate target. Images of similar or different modalities need to be aligned for navigation, detection, data-fusion and visualization in medical applications [7]. Medical image registration still presents many challenges. For example, finding a one-to-one correspondence between several scans of the patient is difficult, because the body of the patient can be subject to sudden changes or the modality of the scans can be different. The first one makes the transformations between scans highly non-rigid and the last one creates significantly different images in overall appearance and resolution [8]. Many medical image registration methods have been developed in the last two decades [1-16]. These can be divided into two major categories, namely, direct and feature based matching. Direct methods use all available image data, and they result in very accurate registration if initialization points are close to target points at the beginning of the registration procedure [11]. For instance, a general purpose registration algorithm for the medical images has been developed which incorporates both geometric and intensity transformation [10]. The authors modeled the transformation with a local affine model and a global smoothness constraint. Intensity variations are also modeled with local changes in brightness, contrast and a global smoothness constraint. Moreover, Myronenko and Song used the definition of the similarity measure to propose a registration method [9]. 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 around Harris corners) to ensure reliable matching. As a result, feature-based methods are independent from an initialization point [13]. Also, feature-based registration methods obtain the transformation parameters from the set of extracted features. For example, Glocker et al. used different levels of smoothness in modeling medical images and then used Markov Random Fields (MRFs) to formulate image deformations [12]. Liu et al. [17, 18] and Elbakary et al. [19] registered multi-modal medical images using banks of local Gabor and Gaussian filters to evaluate the frequencies. The number and characteristics of filters in those works were selected empirically. Staring et al. [20] incorporated multiple image features including the intensity gradients and Hessians. They combined parametric cubic B-splines, and an iterative stochastic gradient ascent optimization [21, 22] to solve the registration problem.

Image-registration techniques based on type of deformation have been divided into two categories: “rigid” and “non-rigid”. In the rigid techniques, like [14], images are assumed to have rotation and translation only but in the non-rigid techniques, like [23], images can have restricted localized stretching. For example, in brain image registration with different modalities, a rigid body approximation is sufficient due to relatively little changes in brain shape over a short period between scans. Roche, G, et al. formulated the rigid registration problem based on general image acquisition model and cast the problem of finding a similarity measure into their maximum likelihood problem [14]. Then, they derived similarity measures for different modeling assumptions. Their experimental results concentrated on the multi-modal images of the brain. Sabuncu et al. introduced an entropy-based algorithm for registering rigid multi-modal images that incorporates spatial information [24]. Spatial feature vectors obtained from the images and a minimum spanning-tree approach were used to estimate the conditional higher-dimensional entropy. They minimized the Jensen-Renyi divergence between the learned and new joint intensity distributions with a gradient descent method.

As an example for non-rigid registration techniques, Likar et al. [23] proposed a hierarchical image subdivision strategy to perform a non-rigid registration method based on mutual information. The non-rigid matching problem was decomposed into a Thin-Plate-Spline-based (TPS) elastic interpolation of multiple local rigid registrations of sub-images. One of the sub-categories of the non-rigid image registration is topology preserving registration. In these kinds of methods, the existing structures are kept, no new structures are allowed to be added, and neighborhood relationships between the structures are preserved. For example, Musse et al. [15] proposed a parametric topology-preserving deformable image registration using the Gauss-Seidel optimization method. The Jacobian of the mapping was controlled over the domain of the transformation to ensure topology preservation. The authors derived the necessary and sufficient conditions for the

determinant of the Jacobian of such transformations to be continuously positive everywhere and applied their method to the 2D images.

Based on the dimensions of input data, the registration method can be categorized as 3D or 2D registration techniques. The 3D registration method normally applies to the registration of two tomographic datasets but the 2D registration may apply to the separate slices from tomographic data. Also, a 3D to 2D registration may help to transfer the acquired 3D data to the 2D data, to facilitate treatment planning. In [25-27], the authors developed automated intensity-based algorithms for updating a 3D position of an interventional instrument using a single plane angiogram registered to a 3D volume. Penney et al. aligned preoperative CT and intraoperative fluoroscopy images where the surface-target registration errors were of the order of 12 mm [25]. Hipwell et al. expanded the former method to registering 3D cerebral Magnetic Resonance Angiography (MRA) with 2D X-Ray angiograms [26] where their RMSE were 1.5 ± 0.9 mm for 85% of the clinical images [26]. Byrne et al. [27], extended Penney et al. work and registered 3D X-ray Digital Subtraction Angiography (3DDSA) images. Their registration method accuracy was 1.3 ± 0.6 mm in the clinical study of the two images with the same modality.

Many researchers incorporate smoothness (or spatial coherence) conditions by reformulating matching into an optimization problem [28,29]. For example, Tang and Chung [29] assigned a vector displacement label indicating the position in the test image to each pixel in the reference image. They used a smoothness constraint based on the first derivative to penalize sharp changes in displacement labels across pixels. Then they employed a graph-cuts method to solve that labeling problem. Moreover, Liu et al. [28] used belief propagation 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 two CT or MR images using sparse coding and belief propagation. First, we build an over complete dictionary out of all 3D features of a reference image [30]. Note that since the dictionary is constructed by padding the features directly, we only need to normalize each column. We then find a set of the 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 [31]. The best match will be selected as the candidate with the highest score. For those voxels with belief less than the threshold θ , we use graph-cuts algorithm [32] to find the proper matches. In comparing to the state-of-the-art belief propagation based registration methods, the key innovation of the proposed approach (3D-SCoBeP) is the inclusion of a preprocessing step to preselect good candidate registration points for each voxel. Belief propagation is very powerful optimization technique, but if the size of the problem increases, it is more difficult to obtain a good local optimum. This restricts the size of the search range for each voxel. In prior approaches such as SIFT-flow [28], the search range is simply chosen as a patch containing neighboring voxels around each target voxel. In contrast, a preprocessing step is used to carefully preselect candidate registration points for each voxel in the 3D-SCoBeP. Since these candidate points are selected from any voxel in the image, the search range of the 3D-SCoBeP is much larger than prior approaches and essentially covers the entire image. This is a main reason for the improvement of the 3D-SCoBeP over the prior works.

A naïve approach computes the Mean Square Error (MSE) of the input patch with each possible patch of the reference image and selects patches that have the smallest MSEs. These kinds of approaches have poor diversity which means the candidate patches are concentrated in a small region. Furthermore, the naïve approach may fail to find the true corresponding match points. It is possible that the naïve approach returns a set of candidates where all of them concentrate around a wrong point. Instead, we propose to find candidate match points using sparse coding. The intuition is that if these candidate patches are similar enough to the source patch, we should be able to construct a source patch out of good candidate patches (so they correspond to a sparse coding solution). With our technique, the sparse coding outputs the patches that can reconstruct the original patch through a linear combination. The resulting patches of sparse coding are likely to be complementary to each other and so provide a better diversity than the naïve solution.

The proposed method described here is inspired by our recent works, SCoBeP^[33] to answer the current challenges in the medical data alignment which are: 1) the patient body movements while capturing the data, 2) the patient body tissue changes due to progress of disease or treatment and 3) different sampling rates of the data because of different sampling rates along the directions. A preliminary version of this work has been reported in^[34]. Since then, much research has been done and the novel components in addition to the aforementioned works are summarized as follows: (1) Using 3D feature of input data which makes the proposed method more accurate than the other state-of-the-art methods; (2) Identifying the best similar matches based on the local and the geometric characteristics using a 3D factor graph, which efficiently trades off both characteristics optimally; (3) Employing a graph-cuts method as a post-processing method which further refines the matches obtained from the belief propagation step.

The rest of the paper is organized as follows. In the next section, we will review some basic concepts of sparse coding and belief propagation. Then in Section 3 we will introduce the concept of our 3D-SCoBeP and the inference algorithm. Finally, in Section 4, we will show our simulation results, followed by a brief conclusion in Section 5.

2 Background

In this section, we will briefly review two core techniques need in our method: sparse coding and belief propagation.

2.1 Sparse coding

Consider a signal $y \in \mathbb{R}^M$ and a fat matrix $D \in \mathbb{R}^{M \times N}$, where we say the matrix is “fat” since $M \ll N$. We are interested in representing y with the column space of $D \in \mathbb{R}^{M \times N}$, i.e., finding $\alpha \in \mathbb{R}^N$ such that $y = D\alpha$. Since D is fat, α is not unique. However, if we also restrict α to be the sparsest vector to satisfy $y = D\alpha$ (i.e., α that has fewest number of non-zero elements), then in theory there is a unique solution. Sparse coding precisely considers the aforementioned problem of finding a sparse α such that $y = D\alpha$ is satisfied. Mathematically, we can write the problem as

$$\hat{\alpha} = \arg \min \|\alpha\|_0 \text{ subject to } y = D\alpha. \quad (1)$$

However, this l_0 optimization problem is NP-complete^[35] and thus several alternative methods have been proposed to solve it^[36]. For example, when a sufficiently sparse solution actually exists, substituting the l_1 norm for the l_0 pseudo-norm in (1) as below

$$\hat{\alpha} = \arg \min \|\alpha\|_1 \text{ subject to } y = D\alpha \quad (2)$$

will still result in the same solution^[35]. Moreover, solving this modified problem is much easier since it can be readily transformed into a linear programming problem. Besides linear programming, many other suboptimal techniques have been proposed to solve (2), including orthogonal matching pursuit^[37], gradient projection^[38] and subspace pursuit^[39].

2.2 Subspace pursuit^[39]

Subspace pursuit (SSP) is a l_1 -Norm minimization method which has a reconstruction capability compared to the Linear Programming (LP) methods, and has very low reconstruction complexity of matching pursuit techniques for very sparse signals. For any sampling matrix A satisfying the restricted isometry property (RIP)^[40] with a constant parameter independent of K , the Subspace pursuit algorithm can recover arbitrary K -sparse signals exactly from its noiseless measurements.

When the measurements are inaccurate and/or the signal is not exactly sparse, the reconstruction distortion is of order a constant multiple of the measurement and/or signal perturbation energy. More precisely, for very sparse signals with $K = O(\sqrt{N})$ where N the number of columns of A , which, for instance, the computational complexity of the Subspace

pursuit algorithm is upper bounded by $O(mNK)$, but can be further reduced to $O(mN \log K)$ when the nonzero entries of the sparse signal decay slowly (For more information about the details of the Subspace pursuit, please read Ref.39).

2.3 Belief Propagation and message passing

Belief Propagation (BP) ^[31] is an approximate inference method used on graphical models such as factor graphs. It is performed by passing messages through the factor graph of a given problem. Define $N(i)$ and $N(a)$ as two sets of neighbors of a variable node “ i ” and a factor node “ a ”, respectively, and denote $m_{i \rightarrow a}$ and $m_{a \rightarrow i}$ as the forward and backward messages from node “ i ” to node “ a ”. The message itself is a vector containing current beliefs of a node mapping to all possible values. For example, $m_{a \rightarrow i}(x_i)$ can be interpreted as the belief of node “ a ” that how probable is the pixel of node i to a map to location x_i in the factor graph. Message updates for $m_{i \rightarrow a}$ and $m_{a \rightarrow i}$ will be based on the messages received by the incoming messages towards nodes i and a , respectively. More precisely, they are given by ^[31]

$$m_{i \rightarrow a}(x_i) = \prod_{b \in N(i) \setminus a} m_{b \rightarrow i}(x_i) \quad (3)$$

$$m_{a \rightarrow i}(x_i) = \sum_{x_a \setminus x_i} f(x_a) \prod_{j \in N(a) \setminus i} m_{j \rightarrow a}(x_j) \quad (4)$$

where we use $N(a) \setminus i$ to denote the neighbor of node a excluding node i .

3 Proposed method

As mentioned in Section 1, in the medical image applications we need dense registration so that for each point of the *source data* a corresponding match point will be found in the *reference data*. This section describes the implementation details of our proposed registration method for the 3D medical data which is based on sparse coding and belief propagation.

First, we extract the features from the 3D reference data $\{X_s\}_{s=1}^{k_x} \in \mathbb{R}^{M \times N \times K}$ and the 3D source data $\{y_s\}_{s=1}^{k_y} \in \mathbb{R}^{M \times N \times K}$ where k_x and k_y are the numbers of the reference and the source image slices, respectively. In this paper, we focus on only using 3D block features even though the proposed approach can generally be applied to other features (such as SIFT-features ^[41] or Gabor-features ^[42]). The 3D feature block is a rectangular cube neighbor around each voxel of the 3D data which we reorder as a 1D vector. Thus, each feature considered here is essentially a vectorized 3D block centered around a voxel in a 3D data.

Second, we create a dictionary \mathcal{D} which contains all extracted feature vectors of the reference data to match to the corresponding extracted features of the source data. The dictionary includes all vectorized 1D features as its columns where all of them have been normalized. We then apply sparse coding to each extracted feature of the source data. Sparse coding will reconstruct a 3D source patch at voxel $[i, j, k]$ as a linear combination of the reference 3D patches. Note that the representation coefficients α_{ijk} 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 a $n \times 3$ matrix storing the locations of these candidate voxels and a probability vector ρ as a length- n vector storing the corresponding probabilities of the sparse coefficient vector. Each coefficient in the probability vector ρ_{ijk} serves as a prior probability of matching the 3D source patch centered around voxel $[i, j, k]$ to a 3D patch of the reference data. This probability vector is taking only local characteristics into account but ignoring geometric characteristics of the matches.

Finally, to incorporate these geometric characteristics, we model the problem by a factor graph and apply BP to identify the best matches similar to ^[43]. We consider a 3D lattice factor graph as follows: For each voxel in the source 3D data, one variable node was assigned and then we connect each variable node to its six neighbors by a factor node (see Figure 1). Also, we consider one extra factor node for each variable node to take care of prior probabilities of the candidate points. In

our model, the factor function $f(x_i, x_j)$ which can be interpreted as the local belief mapping nodes i and j to x_i and x_j can be used to impose the geometric constraint described earlier. Intuitively, since x_i and x_j are the corresponding mapped match points in the reference image of two neighboring voxels in the source image, we expect the probability of getting x_i and x_j to decrease as their distance apart increases. Therefore, in this paper, we model the function of the factor node between two particular variable nodes x_i and x_j with a Gaussian kernel as ^[31]

$$f(x_i, x_j) = e^{-\frac{\|x_i - x_j\|_2}{\sigma^2}} \quad (5)$$

where σ^2 is a parameter to control the relative strength of the geometric constraint imposed by a neighboring node. If we increase the value of σ^2 , the belief of each variable node will have less effect on its neighbors.

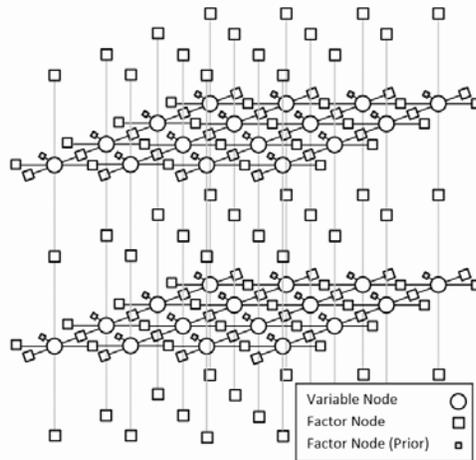


Figure 1. Three dimensional factor graph of medical data used in Belief Propagation: for each voxel in the source 3D data, one variable node was assigned to incorporate these geometric characteristics. We connect each variable node to its six neighbors by a factor node and incorporate one extra factor node to store initial probabilities. A part of two slices of medical data corresponding factor graph is shown. This factor graph can be extended on X-axis, Y-axis and Z-axis.

To synthesize the source image, we replaced each voxel of the source image with the selected candidate voxel from the reference image where its probability is more than the threshold θ . If the final maximum belief of the selected point was less than the threshold θ , we employ graph-cuts algorithm ^[32, 44] to find the correspond voxel. First, for voxels with maximum belief more than θ , we calculate the movement of each voxel of reference data in comparison to the source data and create the displacement matrix β . Then, we feed the displacement matrix β to graph-cuts algorithm to estimate the disparity of voxels with unsatisfactory beliefs. In the graph-cuts algorithm, we initialize the label of each voxel by its displacement value if its input belief is small then threshold θ and by average displacement value of all voxels otherwise. The data term is defined by a quadratic function of the distance between the current label and the desired label, and the smoothness term is defined by a linear function of the distance between the current label and its neighborhood label. The neighborhoods are the same as which we used in the BP step and the swap algorithm of graph-cuts is applied to label voxel with beliefs less than the threshold θ .

Implementation

This section describes the implementation's details of our proposed registration method. The main procedure for our proposed registration method is summarized in Algorithm 1.

Implementation Details:

$Y = \text{ExtractDenseFeature}(\{y_s\}_{s=1}^{k_y})$ presents a 3D block extractor algorithm using $\{y_s\}_{s=1}^{k_y}$ as a source data. More precisely, we consider a 3D block of size $S = (2a + 1) \times (2b + 1) \times (2c + 1)$ containing neighboring voxels around each voxel on a 3D data, where a , b and c are positive integers. For each voxel $p_{i,j,k}$ in the source data $\{y_s\}_{s=1}^{k_y}$, we vectorized the 3D block centered around the voxel $p_{i,j,k}$ to a feature vector $y_{i,j,k} \in \mathbb{R}^{S \times 1}$. A source feature $Y \in \mathbb{R}^{M \times N \times K \times S}$ is then constructed from $y_{i,j,k,t}$ as follows:

Algorithm 1 3D-SCoBeP for the medical image registration estimate version of the registered image Z

Inputs: a reference data $\{X_s\}_{s=1}^{k_x} \in \mathbb{R}^{M \times N \times K}$, a source data $\{y_s\}_{s=1}^{k_y} \in \mathbb{R}^{M \times N \times K}$, a threshold θ , the number of the candidate points n

Extract 3D dense feature and construct dictionary:

- $Y = \text{ExtractDenseFeature}(\{y_s\}_{s=1}^{k_y})$
- $X = \text{ExtractDenseFeature}(\{X_s\}_{s=1}^{k_x})$
- $\mathcal{D} = \text{MakeDic}(X)$

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

- $\hat{\alpha}_{i,j,k} = \text{FindSCV}(\mathcal{D}, y_{i,j,k})$
- $[\mathcal{L}_{i,j,k}, \rho_{i,j,k}] = \text{FindTopSCV}(n, \hat{\alpha}_{i,j,k})$

Refine the candidate voxels:

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

Find the correspond voxels:

- $[Z, \beta] = \text{Warp}(X, \hat{\rho}, \mathcal{L}, \theta)$
- if there is a voxel with probability less than θ then
 - $Z = \text{Graphcuts}(X, \beta, \theta)$

Output: the estimated version of the registered image Z

$$Y = \{y_{i,j,k,t} \mid 1 \leq i \leq M, 1 \leq j \leq N, 1 \leq k \leq K, t \in S\}. \quad (6)$$

Note that X is created in the same manner as Y but instead from the reference data $\{X_s\}_{s=1}^{k_x}$.

- $\mathcal{D} = \text{MakeDic}(X)$ creates a dictionary \mathcal{D} using the vectors of X . Later, the dictionary \mathcal{D} is used to match the extracted features of the source data to the corresponding extracted features of the reference data. We can write \mathcal{D} as

$$\mathcal{D} = [x_{1,1,1} \dots x_{1,1,K} \dots x_{1,N,K} \dots x_{M,N,K}], \quad (7)$$

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

- $\hat{\alpha}_{i,j,k} = \text{FindSCV}(\mathcal{D}, y_{i,j,k})$ finds the candidate match voxels using the sparse coding algorithm, where $\hat{\alpha}_{i,j,k}$ is a sparse vector. Mathematically, we try to solve the following sparse coding problem to find the most sparse coefficient vector $\hat{\alpha}_{i,j,k}$ (see Figure 2) such that

$$y_{i,j,k} = \mathcal{D} \hat{\alpha}_{i,j,k}. \quad (8)$$

Although there are several methods to solve (8) ^[37-39], in our work, we employ Subspace Pursuit (SP) ^[39] because of its computational efficiency.

- $[\mathcal{L}_{i,j,k}, \rho_{i,j,k}] = \text{FindTopSCV}(n, \hat{\alpha}_{i,j,k})$ picks up the n largest coefficients of $\hat{\alpha}_{i,j,k}$ as n candidates. $\mathcal{L}_{i,j,k}$ as an $n \times 3$ matrix stores the locations of these candidate voxels and $\rho_{i,j,k}$ as a 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 centered around the voxel $x_{i,j,k}$. After finding the candidate locations $\mathcal{L}_{i,j,k}$ and their initial probabilities $\rho_{i,j,k}$ for each voxel, we concatenate the results and construct following matrices:

$$\mathcal{L} = [\mathcal{L}_{1,1,1} \dots \mathcal{L}_{1,1,K} \dots \mathcal{L}_{1,N,K} \dots \mathcal{L}_{M,N,K}] \tag{9}$$

$$\rho = [\rho_{1,1,1} \dots \rho_{1,1,K} \dots \rho_{1,N,K} \dots \rho_{M,N,K}] \tag{10}$$

which we will use to apply belief propagation at the next step.

- $\hat{\rho} = BP(\mathcal{L}, \rho)$ models the problem by a factor graph and applies belief propagation^[31] to update probability ρ . The updated probability $\hat{\rho}$ can be used to register the reference data onto the source data. In our case, we assign a variable node for each voxel in the source data 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 data (for more details, see^[33]).
- $[\mathbf{Z}, \beta] = Warp(\mathbf{X}, \hat{\rho}, \mathcal{L}, \theta)$ returns the registered image \mathbf{Z} and a displacement matrix β which contains the movement of each voxel of reference data. This matrix can be used to refine the result of the voxels with a probability less than θ .
- $\mathbf{Z} = Graphcuts(\mathbf{X}, \beta, \theta)$ applies the graph-cuts method to find the displacement of the voxels with a probability less than θ and returns the registered image \mathbf{Z} . The displacement matrix β keeps the movement of the voxels and marks the area with a probability less than θ . We feed the matrix β to graph-cuts algorithm and initialize the label of each voxel by the average displacement value of all voxels if it is marked in the matrix β and by its displacement value otherwise. We use a quadratic function of the distance between the current label and the desired label as the data term and a linear function of the distance between the current label and the neighboring label as the smoothness term. The graph-cuts algorithm updates the matrix β and we use the new displacement matrix to find the corresponding voxels.

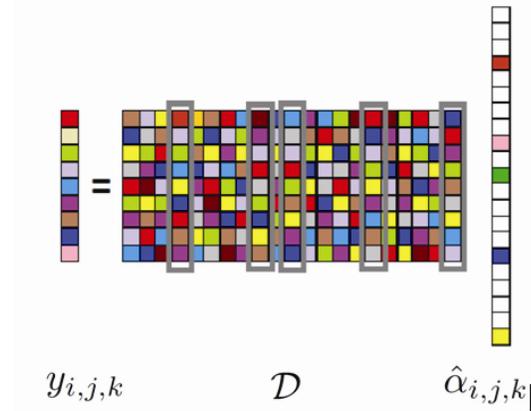


Figure 2. Sparse representation of a feature vector $y_{i,j,k}$ with a dictionary: $\hat{\alpha}_{i,j,k}$ as a sparse vector constructs the feature vector $y_{i,j,k}$ using a few columns (highlighted in gray) of dictionary \mathcal{D} .

4 Experimental results

The utility and novelty of our medical image registration algorithm lies in the fact that it can handle images captured not just from a single plane but also from different planes. Hence, in this section, we studied the performance of our method for both cases and compare it with the different registration methods.

In a brief statement, we presented two experiments in this section: *the 3D CT image registration taken along a same direction* in Section IV-A where we consider the problem of registering two lung CT images of one person from two different times, and *3D MR image registration taken along different directions* in Section IV-B where two brain MR

images were captured along the X-Z (sagittal) and the X-Y (transverse) planes. We implemented the 3D-SCoBeP algorithm in Matlab and tested it on a Pentium 3 GHz (11-GB RAM) machine.

4.1 3D CT image registration taken in same directions

To evaluate the performance of our approach, we conducted tests on the data sets LIDC-IDRI^[45] where the size of each slice of the CT images is 512×512 voxels. Throughout the experiments, the following parameters were used: the number of the 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 $\theta = 0.3$. Otherwise, we assume no “best match” is found. The threshold θ can be chosen empirically which was the way that we chose in Figure 4 and 5 to express the results.

Figure 3 shows the result of the 3D-SCoBeP and MIRT^[9] with a 3D perspective. In this figure, we decided to show only a part of the CT images because the inside details of the lung are more important than the tissue around it. Figure 3(a) and 3(b) are the reference and the source CT image, respectively. Figure 3(c) shows the result of 3D-SCoBeP where we used the voxel of the reference data to synthesize the source data. We created one RGB image where its first channel was assigned to the intensity of source CT image and its third channel was equal to 255. We assigned the reference CT image intensity, the MIRT result and the 3D-SCoBeP result to the second channel, respectively. Therefore, Figure 3(d) corresponds to the initial state and Figure 3(e) and 3(f) are final state of the MIRT and the 3D-SCoBeP. Note that in Figure 3(c), we display a pure result of the proposed method which only the voxels with a probability more than the threshold θ was shown, therefore there are some dark voxels in this figure. For those voxels with probability less than θ , the analysis of motion fields of neighborhood voxels could be used to estimate their motions. It means, one can extract the motions of voxels which have probability more than θ in each direction and apply the graph cuts method^[32], median filter, or moving average to estimate the motion of voxels with probability less than threshold θ .

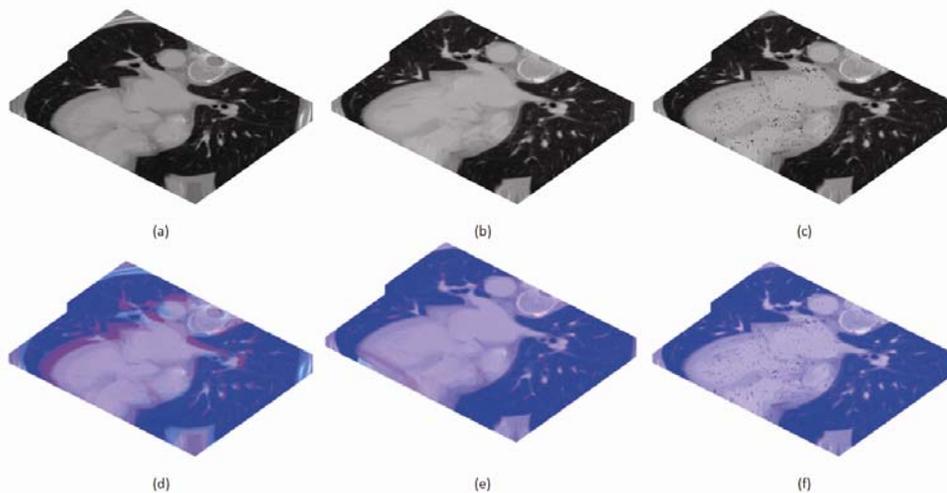


Figure 3. Result of 3D-SCoBeP on Lung CT images. (a) The reference CT image; (b) The source CT image; (c) The 3D-SCoBeP result; (d) The comparison between correspondent voxel between the source and the reference; (e) The comparison between correspondent voxel between the source and the MIRT result; (f) The comparison between correspondent voxel between the source and the 3D-SCoBeP result; In (d)-(e) we used a RGB image where the first channel of the image was assigned to the source image intensity and the second channel to the reference, MIRT, 3D-SCoBeP results, respectively.

We now proceed to compare the 3D-SCoBeP with other approaches; Figure 4 and 5 show the output of our proposed method compared to two of the state-of-the-art methods: the MIRT^[9] and GP-Registration^[10]. In these figures, we select only one slice of CT image to show the weaknesses and strengths of each technique. Figure 4(a) and 5(a) correspond to the reference CT image and Figure 4(b) and 5(b) correspond to the source CT image. Figure 4(c)-(e) and 5(c)-(e) show results using MIRT^[9], GP-Registration^[10] and 3D-SCoBeP. The warped images using MIRT with highlighted artifacts are shown in Figure 4(h) and 5(h) and the warped images using GP-Registration with highlighted artifacts are shown in Figure 4(i) and 5(i). The estimated images generated from the 3DSCoBeP with highlighted areas are shown in 4(j) and 5(j).

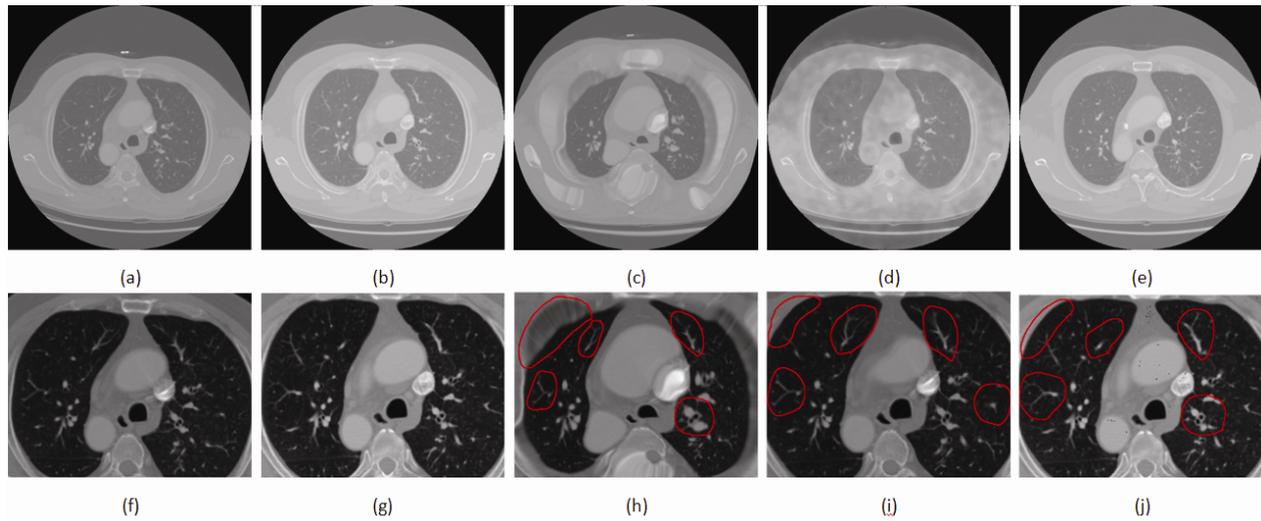


Figure 4. Registration result of the lung CT images that were captured with six months gap. (a) Source image; (b) Reference image; (c) MIRT^[9] [RMSE: 24.31]; (d) GP-Registration^[10] [RMSE: 28.78]; (e) 3D-SCoBeP [RMSE: 21.73]; (f) Source image (zoom in); (g) Reference image (zoom in); (h) MIRT^[9] (zoom in); (i) GP-Registration^[10] (zoom in); (j) 3D-SCoBeP

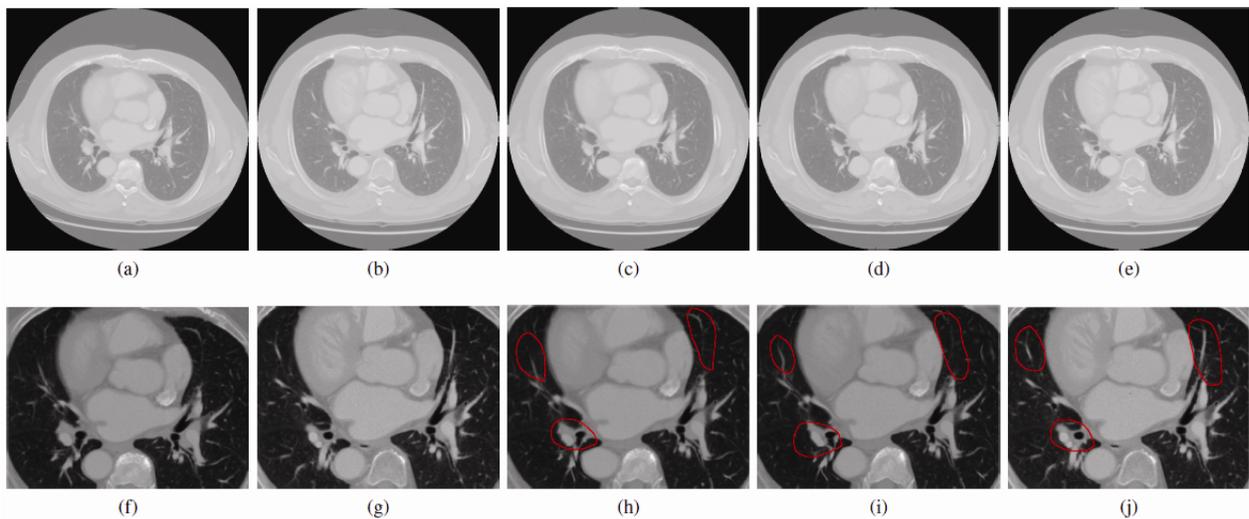


Figure 5. Registration result of the lung CT images that were captured with three months gap. (a) Source image; (b) Reference image; (c) MIRT^[9] [RMSE: 7.71]; (d) GP-Registration^[10] [RMSE: 7.38]; (e) 3D-SCoBeP [RMSE: 4.18]; (f) Source image (zoom in); (g) Reference image (zoom in); (h) MIRT^[9] (zoom in); (i) GP-Registration^[10] (zoom in); (j) 3D-SCoBeP (zoom in).

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

$$\epsilon_{RMSE} = \sqrt{(1/N) \sum \|\tau - \hat{\tau}\|^2}, \quad (11)$$

where N is the number of voxels in the reference and τ and $\hat{\tau}$ are the source image and the estimated transformation respectively. In the term of the RMSE, we compare the source image with the output of the 3D-SCoBeP, MIRT and GPRegistration and the results are shown under (c), (d) and (e) of Figure 4 and 5. However, in Figure 4, the RMSE values for the MIRT and GP-Registration are 24.31 and 28.78, respectively. Although, The MIRT generates the result with the lower RMSE value, the GP-Registration preserves the structures better. Our proposed technique shows an improved qualitative and quantitative result. The 3D-SCoBeP preserves the structure as precise as GP-Registration and also has the less RMSE value, 21.73, in comparison to two other methods. We marked the critical areas with red color to emphasize the weaknesses and strengths of each technique. However, we quantified our registration performance and the other methods using the root mean square error (RMSE) for each figure. In Figure 5, all three methods show the same structure preservation but the RMSE value of the 3D-SCoBeP is much less than the RMSE value of MIRT and GP-Registration.

4.2 3D MR image registration taken in different directions

In this section, we are trying to align the medical data shown in Figure 6 and 7 where the brain MR images captured in parallel to the X-Z (sagittal) and the X-Y (transverse) planes, respectively. The X-axis is from left to right along the column direction in Fig. 6 and is from anterior to posterior along each slice in Figure 7. The Z-axis and Y-axis are from the first slice to the last slice along the plane direction in Figure 6 and 7, respectively.

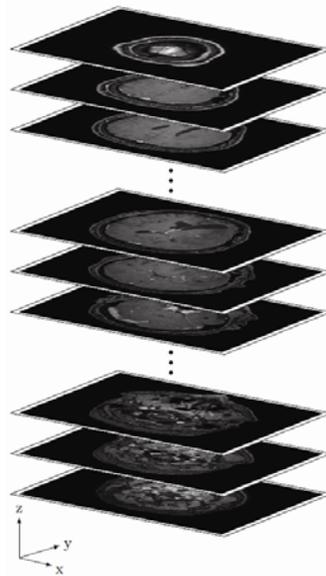


Figure 6. Brain MR images captured in parallel to X-Y (transverse) plane

Here, the captured data are slices of brain image, where each one has 320×320 voxels. In this case, we use 200 slices which were captured in parallel to the X-Y (sagittal) plane as the reference data and 180 slices in parallel to the X-Z (transverse) plane as the source data. For registering this kind of 3D data, we first map each reference and source to a template data in size of $320 \times 320 \times 320$ using B-Spline interpolation^[46].

Note that we set $a = b = c = 7$ and keep all the parameters as the same as previous section. Then, we apply the 3DSCoBeP on the interpolated data. Figure 8(a) and 8(b) show the reference and the source data, respectively. Figure 8(c) shows the output of 3D-SCoBeP. In Figure 8(d), we present the motion field in one of the slices in parallel to the X–Y (sagittal) plane which the darker points have less movement (minimum three pixels movement) and the lighter points are the area with more displacement (maximum sixteen pixels movement). In the other words, the selected area has three pixels translation in compared to the reference data. Also, the distance between the slices in Y direction is not uniform. These slices are closer to each other in the area that is darker in Figure 8(d). Note that all selected voxels' probabilities are bigger than the threshold θ .

The computational complexity of 3D-SCoBeP can be determined by considering the following three steps: 1) extracting dense features and constructing dictionary, 2) finding candidate points via sparse coding, and 3) applying BP. Assume the size of the test and reference images are the same and both have m^3 pixels. The required time of feature extraction will be $O(m^3)$. As for dictionary construction, the only time needed is for the normalization of each column, which requires $O(m^3)$ amount of time. Thus the total time complexity of the first step is $O(m^3)$. In the second step of the 3D-SCoBeP, the time complexity of SP is $O(\log(f) m^6)$ [47], where f is the number of iteration for finding the sparse vector. Since we have to repeat the process of finding candidate points for all m^3 feature vectors, the time complexity of finding candidate points by SP is $O(\log(f) m^6)$. In the third step, the time complexity of BP in our factor graph is $O(vn^3m^6)$, where v is the number iterations before converging. Consequently, if the 3D-SCoBeP uses SP, its time complexity will be $O(m^3 + \log(f) m^6 + vn^3m^6)$. The complexity associated with the second step takes 90% of the overall complexity of 3DSCoBeP.

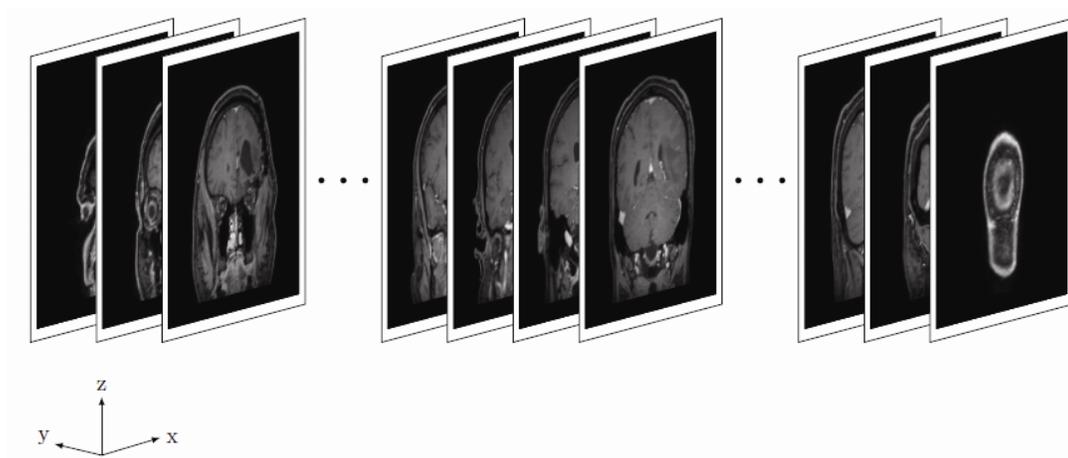


Figure 7. Brain MR images captured in parallel to X–Z (sagittal) plane

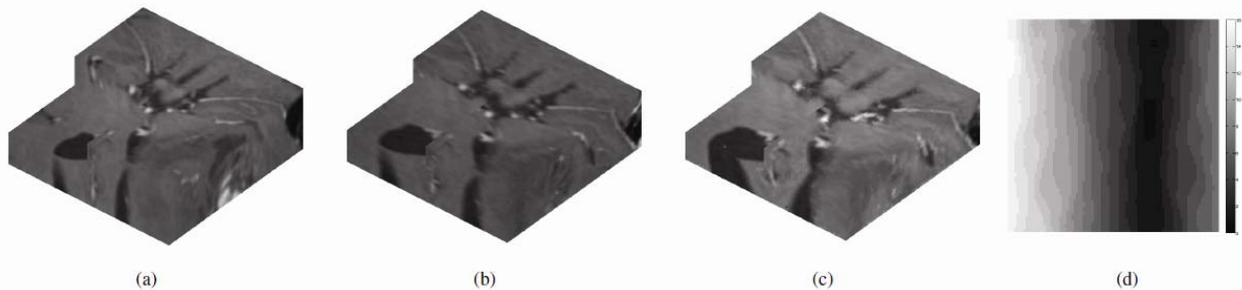


Figure 8. Result of 3D-SCoBeP on Brain CT images. (a) The reference CT image where the CT slices were taken in parallel to X–Z (sagittal) plane; (b) The source CT image where the CT slices were taken in parallel to X–Y (transverse) plane; (c) The 3D-SCoBeP result; (d) motion field in the X–Y (transverse) plane for one selected slice.

5 Discussion and conclusion

In conclusion, we have proposed an effective registration method based on a sparse coding and belief propagation. The proposed method can be used for both rigid and non-rigid registration. Our technique executes registration by first running sparse coding over an over-complete dictionary out of 3D features of 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 with the high accuracy MIRT method by Myronenko and Song^[9] and the state-of-the-art GP-Registration by Periaswamy and Farid^[10] over the CT and MR images. In addition, we used a 3D factor graph for the entire 3D data instead of using a 2D factor graph per slice. This new strategy along with using graph-cut method that refines the disparity of each voxel enables us to register 3D data captured parallel to the sagittal plane into 3D data captured parallel to the transverse plane.

Acknowledgment

The authors would like to thank Mrs. Renee Wagenblatt for editing and correcting grammar mistakes in previous drafts.

References

- [1] Zitova B, Flusser J. Image registration methods: a survey *Image and vision computing*. 2003; 21(11): 977-1000. [http://dx.doi.org/10.1016/S0262-8856\(03\)00137-9](http://dx.doi.org/10.1016/S0262-8856(03)00137-9)
- [2] Roozgard, Cheng S, Liu H. 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. PMID:23366098
- [3] Bro-Nielsen M. Medical image registration and surgery simulation. IMM-DTU PhD thesis. 1996.
- [4] Schnabel J, Rueckert D, 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*. Springer. 2001; 573-81.
- [5] Hill D, Hawkes D, Crossman J, et al. Registration of mr and ct images for skull base surgery using point-like anatomical features. *British journal of radiology*. 1991; 64(767): 1030-35. PMID:1742584 <http://dx.doi.org/10.1259/0007-1285-64-767-1030>
- [6] Lanfranco, Castellanos A, Desai J, et al. Robotic surgery: a current perspective. *Annals of Surgery*. 2004; 239(1): 14. PMID:14685095 <http://dx.doi.org/10.1097/01.sla.0000103020.19595.7d>
- [7] Maintz J, Viergever M. A survey of medical image registration. *Medical image analysis*. 1998; 2(1): 1-36. [http://dx.doi.org/10.1016/S1361-8415\(01\)80026-8](http://dx.doi.org/10.1016/S1361-8415(01)80026-8)
- [8] Periaswamy S, Farid H. Medical image registration with partial data. *Medical Image Analysis*. 2006; 10(3): 452-64. PMID:15979375 <http://dx.doi.org/10.1016/j.media.2005.03.006>
- [9] Myronenko, Song X. Intensity-based image registration by minimizing residual complexity. *Medical Imaging, IEEE Transactions on*. 2010; 29 (11): 1882-91, PMID:20562036 <http://dx.doi.org/10.1109/TMI.2010.2053043>
- [10] Periaswamy S, Farid H. Elastic registration in the presence of intensity variations. *Medical Imaging, IEEE Transactions*. 2003; 22(7): 865-74. PMID:12906240 <http://dx.doi.org/10.1109/TMI.2003.815069>
- [11] Shum H, Szeliski R. Systems and experiment paper: Construction of panoramic image mosaics with global and local alignment. *International Journal of Computer Vision*. 2000; 36(2): 101-30. <http://dx.doi.org/10.1023/A:1008195814169>
- [12] Glocker, Komodakis N, Tziritas G, et al. Dense image registration through mrfs and efficient linear programming. *Medical Image Analysis*. 2008; 12(6): 731-41. PMID:18482858 <http://dx.doi.org/10.1016/j.media.2008.03.006>
- [13] Brown M, Lowe D. Automatic panoramic image stitching using invariant features. *International Journal of Computer Vision*. 2007; 74(1): 59-73. <http://dx.doi.org/10.1007/s11263-006-0002-3>
- [14] Roche, Malandain G, Ayache N, et al. Unifying maximum likelihood approaches in medical image registration. *Inria*. 1999.
- [15] Musse A, Heitz F, Armspach J. Topology preserving deformable image matching using constrained hierarchical parametric models. *Image Processing, IEEE Transactions*. 2001; 10(7): 1081-93, PMID:18249681 <http://dx.doi.org/10.1109/83.931102>
- [16] Cain S, Hayat M, Armstrong E. Projection-based image registration in the presence of fixed-pattern noise. *Image Processing, IEEE Transactions*. 2001; 10(12): 1860-72. PMID:18255526 <http://dx.doi.org/10.1109/83.974571>
- [17] Liu J, Vemuri B, Marroquin J. Local frequency representations for robust multimodal image registration. *Medical Imaging, IEEE Transactions*. 2002; 21(5): 462-69. PMID:12071617 <http://dx.doi.org/10.1109/TMI.2002.1009382>
- [18] Collignon A, Vandermeulen D, Suetens P, et al. 3d multimodality medical image registration using feature space clustering in *Computer Vision, Virtual Reality and Robotics in Medicine*. Springer. 1995; 193-204. <http://dx.doi.org/10.1007/BFb0034948>

- [19] Elbakary M, Sundareshan M. Accurate representation of local frequency using a computationally efficient gabor filter fusion approach with application to image registration. *Pattern recognition letters*. 2005; 26(14): 2164-73.
<http://dx.doi.org/10.1016/j.patrec.2005.03.035>
- [20] Staring M, van der Heide U, Klein S, et al. Registration of cervical mri using multifeature mutual information. *Medical Imaging, IEEE Transactions*. 2009; 28(9): 1412-21. PMID:19278929 <http://dx.doi.org/10.1109/TMI.2009.2016560>
- [21] Klein S, Staring M, Pluim J. Evaluation of optimization methods for nonrigid medical image registration using mutual information and b-splines. *Image Processing, IEEE Transactions*. 2007; 16(12): 2879-90. <http://dx.doi.org/10.1109/TIP.2007.909412>
- [22] Klein S, Pluim J, Staring M. Adaptive stochastic gradient descent optimisation for image registration. *International journal of computer vision*. 2009; 81(3): 227-39. <http://dx.doi.org/10.1007/s11263-008-0168-y>
- [23] Likar, Pernus F. A hierarchical approach to elastic registration based on mutual information. *Image and Vision Computing*. 2001; 19(1): 33-44.
- [24] Sabuncu M, Ramadge P. Using spanning graphs for efficient image registration. *Image Processing, IEEE Transactions*. 2008; 17(5): 788-97. PMID:18390383 <http://dx.doi.org/10.1109/TIP.2008.918951>
- [25] Penney G, Batchelor P, Hill D, et al. Validation of a two-to three-dimensional registration algorithm for aligning preoperative ct images and intraoperative fluoroscopy images. *Medical physics*. 2001; 28: 1024. PMID:11439472
<http://dx.doi.org/10.1118/1.1373400>
- [26] Hipwell J, Penney G, McLaughlin R, et al. Intensity-based 2-d-3-d registration of cerebral angiograms. *Medical Imaging, IEEE Transactions*. 2003; 22(11): 1417-26. PMID:14606675 <http://dx.doi.org/10.1109/TMI.2003.819283>
- [27] Byrne J, Colominas C, Hipwell J, et al. Assessment of a technique for 2d-3d registration of cerebral intra-arterial angiography. *British journal of radiology*. 2004; 77(914): 123-28. PMID:15010384 <http://dx.doi.org/10.1259/bjr/27339681>
- [28] Liu, Yuen J, Torralba A. Sift flow: Dense correspondence across scenes and its applications. *Pattern Analysis and Machine Intelligence, IEEE Transactions*. 2011; 33(5): 978-94.
- [29] Tang T, Chung A. 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*. SpringerVerlag. 2007; 916-24.
- [30] Mairal J, Bach F, Ponce J, et al. Online learning for matrix factorization and sparse coding. *The Journal of Machine Learning Research*. 2010; 11: 19-60.
- [31] Kschischang F, Frey B, Loeliger H. Factor graphs and the sumproduct algorithm. *Information Theory, IEEE Transactions*. 2001; 47(2): 498-519. <http://dx.doi.org/10.1109/18.910572>
- [32] Boykov Y, Veksler O, Zabih R. Fast approximate energy minimization via graph cuts. *Pattern Analysis and Machine Intelligence, IEEE Transactions*. 2001; 23(11): 1222-39. <http://dx.doi.org/10.1109/34.969114>
- [33] Barzigar N, Roozgard A, Cheng S, et al. Scobep: Dense image registration using sparse coding and belief propagation. *Journal of Visual Communication and Image Representation*. 2013; 24(2): 137-147. *Sparse Representations for Image and Video Analysis*. [Internet]. Available from <http://www.sciencedirect.com/science/article/pii/S1047320312001319>
- [34] Roozgard A, Barzigar N, Cheng S, et al. Medical image registration using sparse coding and belief propagation in *Engineering in Medicine and Biology Society (EMBC). 2012 Annual International Conference of the IEEE*. IEEE, 2012; 1141-44.
- [35] Baraniuk R. Compressive sensing *IEEE Signal Processing Magazine*. 2007; 24(4): 118-120.
<http://dx.doi.org/10.1109/MSP.2007.4286571>
- [36] Pang Y, Li X, Yuan Y. Robust tensor analysis with l1-norm. *Circuits and Systems for Video Technology, IEEE Transactions*. 2010; 20(2): 172-78. <http://dx.doi.org/10.1109/TCSVT.2009.2020337>
- [37] Yang A, Sastry S, Ganesh A, et al. Fast algorithms and an application in robust face recognition: A review in *Image Processing (ICIP). 17th IEEE International Conference on*. IEEE. 2010; 1849-52.
- [38] Maleh R, Gilbert A, Strauss M. Sparse gradient image reconstruction done faster. *Image Processing. 2007. ICIP 2007. IEEE International Conference*. IEEE. 2007; 2: II-77.
- [39] Dai W, Milenkovic O. Subspace pursuit for compressive sensing signal reconstruction. *Information Theory, IEEE Transactions*. 2009; 55(5): 2230-49. <http://dx.doi.org/10.1109/TIT.2009.2016006>
- [40] Baraniuk R, Davenport M, DeVore R, et al. A simple proof of the restricted isometry property for random matrices. *Constructive Approximation*. 2008; 28(3): 253-263. <http://dx.doi.org/10.1007/s00365-007-9003-x>
- [41] Lowe D G. Object recognition from local scale-invariant features. *Computer Vision. The Proceedings of the Seventh IEEE International Conference*. IEEE. 1999; 2: 1150-57.
- [42] Feichtinger H G, *Gabor analysis and algorithms: Theory and applications*. Birkhauser, 1998.
<http://dx.doi.org/10.1007/978-1-4612-2016-9>
- [43] Cheng S, Stankovic V, Stankovic L. Improved sift-based image registration using belief propagation. *Conf. Acoustics, Speech and Signal Proc. Taiwan*. 2009; 2909-12.

- [44] Fulkerson B, Vedaldi A, Soatto S. Class segmentation and object localization with superpixel neighborhoods. Computer Vision. IEEE 12th International Conference. 2009; 670-677.
- [45] N. C. Institute. The cancer imaging archive. 2011; September. [Internet]. Available from: <https://wiki.cancerimagingarchive.net/display/Public/LIDC-IDRI>
- [46] Lehmann TM, Gonner C, Spitzer K. Addendum: B-spline interpolation in medical image processing. Medical Imaging. IEEE Transactions. 2001; 20(7): 660-65. PMID:11465471 <http://dx.doi.org/10.1109/42.932749>
- [47] Dai W, Milenkovic O. Subspace pursuit for compressive sensing: Closing the gap between performance and complexity. DTIC Document. Tech. Rep. 2008.