CONSTRUCTION OF A LINEAR UNBIASED DIFFEOMORPHIC PROBABILISTIC LIVER ATLAS FROM CT IMAGES

Wei Xiong1*, S. H. Ong2, Qi Tian1, Guozhen Xu2†, Jiayin Zhou3, Jiang Liu1, S. K. Venkatash3

1Institute for Infocomm Research, A-STAR, Singapore 138632
2Department of ECE, National University of Singapore, Singapore 117576
3Department of Diagnostic Radiology, National University of Singapore, Singapore 119074

ABSTRACT

The construction of probabilistic liver atlases has received little attention in the past. Existing methods are based on landmarks and are sensitive to their choices and placements. We propose an iterative landmark-free method based on dense volumes to construct linear unbiased diffeomorphic probabilistic atlases from liver CT images. The linear averaging of the transformed images is set as the common target space followed by pairwise diffeomorphic registrations to warp all images to the target using a recent-proposed efficient deformation approach during each iteration cycle. Iterative pairwise registrations are directly used to handle possible large deformations without the need for an extra step to remove global deformations such as the use of affine transformations in traditional methods. Compared with those approaches estimating the unbiased atlas and the transformations groupwise simultaneously, the current method is more efficient. The efficiency and the convergence of our method have been demonstrated experimentally by validation using 25 CT liver sets.

Index Terms— Probabilistic atlas, liver, linear unbiased, diffeomorphic, CT

1. INTRODUCTION

Computational anatomy as an emerging discipline has gained considerable attention in the past ten years. One of the core tasks is to build a probabilistic atlas of an anatomical structure encoding probabilities of anatomic variability while retaining both spatial and densitometric variance. With more and more datasets used to construct the atlas, it becomes even more representative than conventional atlases without variance information. Probabilistic atlases provide valuable and unique information for medical image interpretation, segmentation and registration as well as group variation studies. Motivated by liver radiotherapy treatment and transplantation planning as well as group comparisons, here we improve our earlier study on the construction of a probabilistic atlas of the liver [1] using CT images. Although there are many efforts on the construction of probabilistic atlases of brain [2, 3], and most of the methods are extendable to liver atlas construction, there have been very few actual demonstrations of liver atlases [1, 4, 5, 6, 7, 8].

Existing liver probabilistic atlases are built based on landmarks. Park et al. [6, 7] use 36 expert-labeled landmark points (17 on the liver) followed by affine transformations using thin-plate splines (TPS) for interpolation. Ref. [8] extracts the bone and diaphragm first and then normalizes the liver position, size, and shape by adjusting the cross-sectional area of bone frame to a standard rectangular position and warping the diaphragm to a plane based on a TPS method. Then it creates a normalized likelihood image of liver by projecting the pre-segmented liver regions into a 3-D space. In our earlier work [1] we have followed [6] and built an abdominal probabilistic atlas using landmarks yet labelled by non-experts. We have observed that the registration and hence the atlas is very sensitive to the choice of the target, the accurate localization and the placement of the landmarks. Also, the accuracy and reliability of the points of the atlas tend to decrease for points further away from the landmarks [2]. The liver shape is highly variable and the texture is almost homogeneous in CT images and therefore automatic localization of the landmarks inside the liver is difficult. If landmarks are chosen outside the liver, then their influences are inevitable on the representativeness and the reliability of the atlas of the liver alone. In addition, the landmark-based atlases are biased to the chosen landmarks and the choice of suitable landmarks requires further study. In the current work we have decided to build the liver atlas based on the entire volume.

Typically when building an atlas using multiple subjects, a common target (or reference, or fixed) image space has to be chosen and all other training (or moving) images are mapped onto it. This introduces bias towards to the chosen image as different choices of target subjects would lead to a different atlas which makes the atlas irreproducible if another image is chosen as the target [3, 9, 10, 11, 12]. Minimizing or even removing such bias is thus highly desirable. Various choices of the target subject have been proposed [3, 9, 10, 11, 12]. Basically unbiased atlases can be based on linear [3, 9, 10] or nonlinear averaging of training images [11, 12]. Meanwhile, the deformation can be pairwise and separated from the averaging [11] or simultaneous groupwise estimation of the unbiased atlas and the transformations [12]. Generally, traditional linear unbiased atlas using pairwise deformations is computationally more convenient than those which either pose complicated registration problems for constraints of smooth deformation fields [3] or the need to solve large optimizations involving all training subjects with slow convergence, say, 200 iterations [12].

Recent developments in synergizing between the work on constructing statistical atlases of shape variation and that on non-rigid registration focus on building models of diffeomorphic deformation fields which maps one image into another with a dense correspondence, called diffeomorphisms, which is smooth and invertible so that every point in one image has a corresponding point in the other [13]. It has been found that the smoother interpolation schemes such
as TPSs are prone to tearing space [13]. There is a variety of ways of constructing diffeomorphic maps between pairs of images. Recently, Vercauteren et al. [14] propose an efficient non-parametric image registration method, called diffeomorphic demons (DD), which works on the entire space of displacement fields. The solution has been proved to be smooth and efficient. However, whether DD can be used to handle large deformations was not demonstrated.

As a first attempt in this preliminary work, we adopt the classical pairwise deformation approach to construct a linear unbiased diffeomorphic probabilistic atlas of liver. Unlike traditional average models studying the pure morphological differences up to an affine transformation [3, 11], we will not perform such an extra transformation. Instead, we directly utilize DD [14] to handle large deformations iteratively. We demonstrate the efficiency of our method using 25 liver CT datasets.

2. METHODOLOGY

The construction of a probabilistic atlas comprises two phases: find the optimal representative of the group of images and statistically summarize the spatial and intensity variations. We first introduce the basic flow of unbiased multiple-subject registration and then describe the diffeomorphic deformation approach for pairwise registrations, followed by variation probability analysis.

2.1. Basic flow

The first phase is to estimate a template image that is the best representative for the population on the infinite dimensional space of diffeomorphisms. Generally this can be considered as a registration optimization problem involving multiple images. Formally, given N image intensity images \( \{x_i\}_{i=1}^N \) defined in a very large Euclidian vector space \( \mathbb{R}^q \) (with \( q \) being the maximum number of voxels), we want to find \( \hat{\mu} \) such that

\[
\hat{\mu} = \arg \min_{\mu \in \mathbb{R}^q} \sum_{i=1}^N d(x, x_i)^2
\]  

(1)

where \( d \) is a distance metric. If we choose the simple Euclidian distance, then the unbiased estimation is the algebraic mean of the population, i.e.,

\[
\mu = \frac{1}{N} \sum_{i=1}^N x_i.
\]  

(2)

Now considering the possible deformation \( h_i \) for each image \( x_i \) with respect to the target, we have, at each iteration \( j \),

\[
\mu^{(j)} = \frac{1}{N} \sum_{i=1}^N I_i(h^{(j)}_{i}(x_i))
\]  

(3)

Here for the original training images, \( j = 0 \). Further, we use

\[
\omega(h^{(j)}_{i}(x), \mu^{(j)}) \rightarrow h^{(j+1)}_{i}(x)
\]  

(4)

to denote the \( i \)th image \( x \) deformed by \( h_i \) for each iteration \( j \) to the space of \( \mu^{(j)} \) and result in a new warped image \( h^{(j+1)}_{i}(x) \), derived from \( x \). The algorithm flow is as follows:

1. Step 1: Let \( j = 0 \).
2. Step 2: Compute \( \mu^{(j)} \) using the Procrustes method, i.e., superposition of all sets of points. If \( j = 0 \), use Eq. (2), else use Eq. (3).
3. Step 3: For each image \( I^{(j)}_{i} = I_{i}(h^{(j)}_{i}(x_i)) \), find the optimal diffeomorphic deformation \( \hat{h}^{(j)}_{i} \) to warp \( I^{(j)}_{i} \) to \( \mu^{(j)} \), i.e.,

\[
\omega(\hat{h}^{(j)}_{i}(x), \mu^{(j)}) \rightarrow \hat{h}^{(j+1)}_{i}(x)
\]

4. Step 4: Let \( j = j + 1 \); go to Step 2 until convergence or the maximum number of iterations is achieved.

The details of Step 3 are described below. Note that this step is just a registration problem of two images. Hence we will use a series of independent pairwise registrations for large deformations in an iterative fashion for atlas construction from multiple images. The final optimal group representative image is the output of the last iteration.

2.2. Finding the optimal diffeomorphic deformations

Now we describe the method to find the optimal deformation \( \hat{h}^{(j)}_{i} \) in Step 3 above. We have chosen DD for the task. To be complete, we outline the algorithm below. Given a fixed image \( F \) and a moving image \( M \), the deformation \( s(p) \) refers to the spatial mapping of points \( p \) from the moving image space to the fixed image space. The similarity of the two images is defined by

\[
\lambda(F, M \circ s) = \frac{||F - M \circ s||^2}{2 \sigma_F^2} = \frac{\sum_{p \in \Omega_p} |F(p) - M(s(p))|^2}{2|\Omega_p|},
\]  

(5)

where \( M \circ s \) indicates image \( M \) is deformed by using transformation \( s \) and \( \Omega_p \) is the region of overlap between \( F \) and \( M \circ s \). Minimizing the similarity directly over a space of dense deformation fields would be unstable. To avoid this, a regularization term \( \theta(s) \) can be introduced to form a global energy

\[
E(s) = \frac{1}{\sigma_F^2} \lambda(F, M \circ s) + \frac{1}{\sigma_T^2} \theta(s),
\]  

(6)

where \( \sigma_T \) accounts for the noise on the image intensity and \( \sigma_T \) controls the amount of regularization [14]. However, the mixing of the similarity and the regularization terms leads in general to computationally intensive optimization steps. The optimization can be efficient by introducing a hidden variable \( c \) for the correspondence between points of the two images into the global energy

\[
E(c, s) = \frac{1}{\sigma_F^2} \lambda(F, M \circ c) + \frac{1}{\sigma_T^2} ||c - s|| + \frac{1}{\sigma_T^2} \theta(s),
\]  

(7)

with \( \sigma_T \) accounting for a spatial uncertainty on the correspondences [14]. DD was designed based on this idea. In critical computation steps a few compositions replace an addition of free form deformations, making the implementation more efficient, in addition to being diffeomorphic. An implementation of the algorithm can be found in ITK [15].

2.3. Variation probability analysis

After finding the optimal representative group mean image, we check whether there is a liver point instance at each voxel of the registered image space. If there is, the number of instances is incremented by one. After scanning all the space, the higher the number of liver instances found, the higher the probability. In this way we construct the liver probabilistic atlas.
3. EXPERIMENTAL RESULTS

3.1. Experimental Setup

We use $N = 25$ CT liver datasets of varying spacings and sizes. They are first normalized to give a common size of $441 \times 441 \times 48$ voxels with voxel size $1 \times 1 \times 5$ mm$^3$. Registration performance is measured by using the mean square metric and the mutual information between the registered image and the fixed image. The mean square metric between the intensity difference of two images $A$ and $B$ is given by

$$\text{MSE}(A, B) = \frac{1}{q_1} \sum_{i=1}^{q_1} (a_i - b_i)^2,$$

where $a_i$, $b_i$ is the intensity of the $i$th pixel of $A$ and $B$, respectively, and $q_1$ is the total number of pixels considered. We denote the marginal probability density functions of the intensities of $A$ and $B$ by $p_1$ and $p_2$, respectively, and their joint probability density function by $p_{AB}$. The entropy of an image is defined by

$$H_i = -\sum_k p_i(k) \log_2 p_i(k), i = 1, 2.$$  \hspace{1cm} (9)

Similarly, we can define the joint entropy $H_{1,2}$ by using $p_{AB}$. Now the mutual information MI can be defined by

$$\text{MI} = H_1 + H_2 - H_{1,2}.$$  \hspace{1cm} (10)

3.2. Results

We have implemented our method using liver expert-labeled CT image datasets in 10 iterations. For each cycle of iterations $j$, $j = 1, \ldots, 10$, we record the performance indexes (MSE and MI) between each registered image $h^{(j)}(x)$ and the mean $\mu^{(j)}$. The respective averages of the MSE and MI measures of all datasets for this iteration are also computed as performance indexes. We have used $N = 5, 10, 15, 20, 25$ datasets to construct 5 respective probabilistic atlases. Of all the datasets involved, their average of MSEs and MI s for $N = 10$ are denoted by MSE10 and MI10, respectively. Similarly, we can define the measures for other cases, say, MSE25 and MI25 are for $N = 25$.

Fig. 1 presents these performance indexes against the number of iterations. As we iterate more times for the optimization, the MSEs decrease monotonically while MIs increase monotonically. Hence, given enough iterations, the iterations will converge. In particular, for MSE5, MSE10 and MSE15, their curves tend to be horizontal, i.e., already converged. Furthermore, the more datasets we use, the higher their MIs are. Hence with more and more datasets for training, we can obtain more information and more representative liver atlases.

Figs. 2 and 3 illustrate, for $N = 10$ and $N = 25$, respectively, the mean images $\mu^{(j)}$, $j = 1, 5, 10$, and the corresponding probabilistic atlases in upper, middle and lower rows, respectively. In each of the figures, the left three columns are the mean images viewed from three angles and the right three columns are their respective probabilistic atlases in order. Each column is for a particular view of observations. When visualizing the atlases, we render the image with brightness proportional to the probability of each voxel at which it might be a liver point so that the higher the probability the brighter it appears.

We notice that, with more iterations, the fine parts of the liver become more apparent. Meanwhile, comparing the corresponding items in Figs. 2 and 3, we also notice that the mean images in Fig. 2 are smoother than the respective images in Fig. 3. In other word, with more datasets as training images, the mean images will contain finer details and hence more representative of actual livers and their variations. Thus given enough numbers of data and iterations, we could find the most representative image of the group of livers.

Timing wise, it takes about 12 minutes on average to complete a pairwise deformation optimization on a 2.83GHz Dual-core computer with 4GB RAM. To construct a probabilistic atlas using 25 sets in 10 iterations, we need about 50 hours.

3.3. Comparison with a biased atlas

We have also constructed a biased probabilistic atlas using the 25 datasets. In this method, we subjectively choose a dataset, say $x_1$, as
the target image and deform the other 24 datasets successively and incrementally in the space spanned by the respective output image previously registered. We have iterated the registrations five times. Fig. 4 illustrates the registered images (in the left three columns) and the respective atlases (in the right three columns) after 1, 3, and 5 iterations. Very fine details can be observed in the mean images. However, they are biased to the chosen target image.

Fig. 4. The mean images and respective probabilistic atlases of liver resulted from a biased construction method using 25 datasets for iteration 1 (upper), 3 (middle) and 5 (lower row). Three views are shown in the columns.

4. DISCUSSION AND CONCLUSION

In the framework of computational anatomy, a probabilistic atlas is a probability encoded map of anatomic variability of an anatomical structure retaining both spatial and densitometric variance. Although there has much research effort into constructing brain atlases, there has been less attention given to the construction of liver atlases especially with diffeomorphic probabilistic atlases based on dense deformations, which are more robust to noise than those based on landmarks.

In this work, we have proposed a new method to generate a linear unbiased diffeomorphic probabilistic atlas of the liver from 25 sets of CT images. Compared with existing constructed liver atlases, our registrations are based on dense deformation fields without manually-labeled and noise-sensitive landmarks. Our iterative approach has been validated by using 25 liver CT images and we have experimentally demonstrated the convergence. In our experiments, convergence can be observed within tens iterations, which is more efficient than those methods using simultaneous groupwise estimation of the unbiased atlas and the transformations.

Most recently linear averaging is found to be fine with small deformations however not for large deformations directly as under large deformation settings the space of transformations is not a vector space but rather the infinite dimensional group of diffeomorphisms [12]. Hence nonlinear unbiased atlases based on intrinsic averaging have been proposed and demonstrated for brains. This is one of the future directions we can explore. Further, it is not clear theoretically whether we can use diffeomorphic registrations directly handle large deformations without any affine transformations to find the optimal group representative. Finally, the convergence property of the current method will be further studied and tested with more datasets and iterations.

5. REFERENCES


