Determining histology-MRI slice correspondences for defining MRI-based disease signatures of prostate cancer

Gaoyu Xiao a,∗, B. Nicolas Bloch b, Jonathan Chappelow a, Elizabeth M. Genega c, Neil M. Rofsky c, Robert E. Lenkinski d, John Tomaszewski f, Michael D. Feldman i, Mark Rosen i, Anant Madabhushi a

a Department of Biomedical Engineering, Rutgers, The State University of New Jersey, Piscataway, NJ 08854, USA
b Boston Medical Center, MA, USA
c South Western Medical Center, Dallas, TX, USA
d Department of Radiology, Beth Israel Deaconess Medical Center, MA, USA
∗Department of Surgical Pathology, Beth Israel Deaconess Medical Center, MA, USA
i University of Pennsylvania, 3400 Spruce Street, Philadelphia, PA, USA

ABSTRACT

Mapping the spatial disease extent in a certain anatomical organ/tissue from histology images to radiological images is important in defining the disease signature in the radiological images. One such scenario is in the context of men with prostate cancer who have had pre-operative magnetic resonance imaging (MRI) before radical prostatectomy. For these cases, the prostate cancer extent from ex vivo whole-mount histology is to be mapped to in vivo MRI. The need for determining radiology-image-based disease signatures is important for (a) training radiologist residents and (b) for constructing an MRI-based computer-aided diagnosis (CAD) system for disease detection in vivo. However, a prerequisite for this data mapping is the determination of slice correspondences (i.e. indices of each pair of corresponding image slices) between histological and magnetic resonance images. The explicit determination of such slice correspondences is especially indispensable when an accurate 3D reconstruction of the histological volume cannot be achieved because of (a) the limited tissue slices with unknown inter-slice spacing, and (b) obvious histological image artifacts (tissue loss or distortion). In the clinic practice, the histology-MRI slice correspondences are often determined visually by experienced radiologists and pathologists working in unison, but this procedure is laborious and time-consuming. We present an iterative method to automatically determine slice correspondence between images from histology and MRI via a group-wise comparison scheme, followed by 2D and 3D registration. The image slice correspondences obtained using our method were compared with the ground truth correspondences determined via consensus of multiple experts over a total of 23 patient studies. In most instances, the results of our method were very close to the results obtained via visual inspection by these experts.

© 2010 Elsevier Ltd. All rights reserved.

1. Introduction

As medical images from different modalities convey different types of information about the same anatomical organ/tissue, it is very important to correlate these images so that complementary information can be combined to better aid in disease detection and diagnosis [1,2]. For example, in order to define the radiological-imaging-based disease signature pertaining to a specific anatomical organ/tissue, it is necessary to map the ground truth of the spatial disease extent from the corresponding histology onto the in vivo radiological images. Two such applications where this ability to precisely map disease extent will be very useful are in the context of (a) training radiologist residents to identify disease signatures on in vivo imaging and (b) constructing a radiology-image-based computer-aided diagnosis (CAD) system for detecting and diagnosing disease. A specific problem considered in this paper is the training of a magnetic resonance imaging (MRI)-based CAD system for prostate cancer detection [3–7], where the ground truth for cancer extent on the ex vivo whole-mount histological images needs to be mapped to the corresponding pre-operative in vivo magnetic resonance (MR) images.

Prostate cancer (CaP) is the most common malignancy among men and the second leading cancer related cause of death after lung cancer [8,9]. Researchers have looked at non-invasive medical imaging modalities such as MRI to identify CaP patterns [10]. Some groups, including ours, have also developed MRI-based CAD systems to detect CaP in vivo [3,11]. To distinguish cancerous from non-cancerous regions in vivo via prostate MRI (T2-w or T1-w),
CaP appearance needs to be quantitatively modeled using a set of training images on which disease extent has been delineated. However, since it is generally difficult to annotate the cancerous regions directly on the pre-operative in vivo MR images, delineation of disease extent on histopathology through microscopic analysis still remains the “gold standard” [12]. For men with prostate cancer and undergoing radical prostatectomy, pre-operative MR images of the prostate gland could be registered with the ex vivo post-operative histological images. This allows for mapping the ground truth for spatial extent of CaP from histology to MRI [13,14].

There are two general approaches to map ex vivo the histological CaP extent to pre-operative MR images. The first, and perhaps the more intuitive approach, is to reconstruct the 3D histology volume first, then register the 3D histology volume with the 3D MR volume [16,34]. The second approach is to register each 2D histology slice to its corresponding 2D MRI slice separately [3,14,15]. In the first approach, one critical prerequisite is the accurate reconstruction of the histological volume; while in the second approach, the prerequisite is to determine histology-MRI slice correspondences. In some cases, the former prerequisite may not be achievable (as will be discussed in details below), hence the only solution is to take the second approach.

1.1. Difficulties in the reconstruction of 3D histology volume

In order to accurately reconstruct the 3D histology volume, the following two pre-requisites need to be satisfied:

1. A large number of histological slices with known inter-slice spacing are available: In [22], 89 and 28 slices of histological sections with inter-spacing varying between 140 μm and 260 μm were obtained from two data sets. In [23], the 1.5 cm thick human basal ganglia were cut into 70 μm thick sections before one out of every 10 sections were obtained. In [24], the rat brains were cut into 20–26 sections with an inter-section spacing of 0.4 mm. Similarly, in [25] whole rat brains were sectioned with a uniform slice spacing of 1 mm. Schmitt et al. [26] first cut a 170 mm long human brain into 6214 sections before choosing 450 sections for the brain reconstruction, while in [33], 121 histological sections with an inter-slice spacing of 0.72 mm were used in the 3D reconstruction of baboon brains.

2. The artifacts in the histological slices can be corrected: The artifacts are mostly in the form of tissue loss or distortion caused by prostate sectioning or Haematoxylin and Eosin (H & E) staining. In order to correct these artifacts, researchers have previously relied on (a) the use of blockface images (digital photos of the tissue slice taken right before the histological sectioning) and/or (b) extra fiducial markers to aid in the 3D histological reconstruction of the prostate [16,33–35]. For both approaches, pathologists need to carefully examine all the histological slices so as to manually place control points or fiducial markers at appropriate locations on the selected slices.

However, these two prerequisites may not be easily satisfied. This is especially true during the preparation of the prostate histological data for the purpose of training a CAD system. Ideally, to train a CAD system, it is desirable to collect a large amount of training data from as many patients as possible. However, when large amounts of histological data are required, it becomes unrealistic for the pathologists to either scan a large number of histological slices for each patient, or laboriously go through the time-consuming procedures of preparing blockface images and manually placing fiducial markers on these slices. Therefore, in actual clinical practice, the histological image preparation procedures need to be performed in a reasonable time frame. Hence in many medical centers, the prevailing protocol of preparing prostate histological data comprises of the following steps: (1) The excised prostate gland is sectioned into a number of blocks each of which is a few millimeters thick. (2) Each of the blocks is further sectioned into much thinner slices each of which is approximately 5 μm thick. (3) All these thin slices are examined under a microscope and for efficiency, typically only one representative thin slice with relatively good imaging quality is selected for each block and scanned, while the rest are discarded. Hence, the histological slices are coarsely and unevenly spaced, with indeterminate inter-slice spacing possibly ranging from 5 μm to several millimeters, making it difficult to accurately reconstruct a 3D histological volume. In this case, the only solution for mapping ex vivo the histological CaP extent to pre-operative MR images is to register each 2D histology slice to its corresponding 2D MRI slice separately. Apparently, this requires the 2D histology-MRI slice correspondences be determined at first.

1.2. Determination of the 2D slice correspondences between prostate histology and MRI

The existence of 2D histology and MRI slice correspondences has been reported in the context of human prostate data [3,27,28], as well as small animal data [15,25]. Nevertheless, as Park et al. [16] recently pointed out, because of the possible out-of-plane alignment problem (i.e. the orientation of the 2D histology slices and the axial MRI slices not being identical), there is no guarantee of the existence of a particular in vivo axial MRI slice corresponding to a particular histology slice. This in turn suggests that the exact corresponding MRI slice can only be found through a slice-to-volume registration between each 2D histology slice and the 3D MRI volume.

Although some existing slice-to-volume registration methods have been previously proposed [29–32], the registration between prostate 2D histology and 3D MRI is not readily addressable via these methods. Because of the presence of histological artifacts [10], the registration between the prostate histology and MR images often has to rely on special landmarks. Although in [10], some internal landmarks were extracted from both high resolution quartered histological and ex vivo MRI slices, these landmarks are much more difficult to identify automatically on low resolution whole-mount histological and in vivo MRI slices. For this reason, manual selection of landmarks for registration is still the most reliable and commonly adopted approach, such as in [16].

The manual selection of landmarks in the prostate images has to be done by pathologists and radiologist reviewing both histological and MRI images, a very laborious and time consuming task. In order to ease this burden on the pathologists and radiologists, it would be desirable to limit the searching space to find the landmarks. One way to achieve this is to find the slice correspondences between the histology and MR images, so that the pathologists and radiologists only need to select the landmarks on each pair of corresponding slices (if such slice correspondences do exist, such as in [27,28,3]) or on the few image slices that are in the vicinity of the corresponding slices (if the out-of-plane problem exists, such as in [16]). In both cases, the 2D histology-MRI slice correspondences need to be determined.

At first glance, a straightforward way to determine the 2D histology-MRI slice correspondences is via a brute-force, pair-wise comparison scheme, wherein each histology slice is compared with every MRI slice according to a pre-defined image similarity measure (such as mutual information (MI) [17,18]). Thus, the MRI slice determined to have the highest similarity value with respect to a specific histology slice is identified as the corresponding MRI slice.

However, this pair-wise comparison approach may only work for certain types of data in which the multi-modal or multi-protocol image differences are not very large and where both image sets have consistent and relatively small slice spacing. The establish-
2. Overview of the determination of slice correspondences

2.1. General framework

We present a computerized framework to automatically determine slice correspondence between images from histology and MRI. Our method consists of three modules (described below) which are iteratively executed until there is no further change to the slice correspondence result:

Module 1: Obtain an initial estimate of the slice correspondences via a group-wise comparison of mutual information between the histological and MR images. Here, we consider all the histology slices as a single group and all the MRI slices as another group. The goal is to find a sub-set of MRI slices that best matches the histology group. The intuition behind this scheme is that during this group-wise comparison procedure, the order of the image slices in each group is strictly maintained, thereby limiting the extent of the mismatch.

Module 2: Based on the slices correspondences obtained in Module 1, each histology slice is registered to its corresponding MRI slice to compensate the image differences between the histology and MRI slices.

Module 3: Using the estimate of the slice correspondences obtained in Module 1 and the registered histology slices obtained in Module 2, a 3D histological pseudo-volume can be roughly reconstructed through zero-padding (i.e., inserting additional zero-value slices). The 3D MRI volume is then registered to this reconstructed pseudo-volume so as to gradually alleviate the out-of-plane problem. After the registration, all the 2D MRI slices can be updated by re-slicing the registered MRI volume along the axial direction.

These three modules will be discussed in detail in Sections 3–5, respectively. With these iterative steps, we not only make use of the existence of approximate histology-MRI slice correspondences, but also take into consideration the possible out-of-alignment problem. Moreover, our scheme allows us to rectify the histological artifacts that can affect the slice correspondence result. Fig. 2 illustrates a flowchart of our complete framework to automatically determine the histology-MRI slice correspondences.

2.2. Notation

We define an image \( C = (c, f) \), where \( C \) is a 2D grid of image pixels \( c \in C \) and the intensity value at each pixel location \( c \) is denoted by \( f(c) \). Given \( m \) slices \( C^{h,1}, \ldots, C^{h,m} \) from the first image set \( S^h \), and \( n \) slices \( C^{m,1}, \ldots, C^{m,n} \) from the second image set \( S^m \) (in our case, \( n > m \)), with images in sequential order for both image sets, we formulate the determination of slice correspondences between \( S^h \) and \( S^m \) as a group-wise comparison problem. The aim is thus to choose from all possible image sub-sets \( S^A \) \( 1 \leq A \leq \binom{n}{m} \) of \( S^m \), the most similar sub-set to \( S^h \), where \( S^A = \{ C^{m,a_1}, C^{m,a_2}, \ldots, C^{m,a_n} \} \), \( \{a_1 < a_2 < \ldots < a_n\} \) indicates the number of combinations of \( n \) taken \( m \) at a time, and \( a_i \) is the index of the image slice in \( S^m \) and \( A \) is a sequential number indicating each of all possible subsets of \( S^m \). The group-wise similarity measure between \( S^h \) and \( S^m \) is denoted by \( D(S^h, S^A) \), which will be further elaborated in Section 3.2.

3. Group-wise comparison to estimate the slice correspondences

3.1. Overview of group-wise comparison

The group-wise comparison module corresponds to Module 1 in Fig. 2. Given 2 image sets comprising individual 2D slices \( S^h = \{C^{h,1}, \ldots, C^{h,m}\} \) and \( S^m = \{C^{m,1}, \ldots, C^{m,n}\} \), respectively, in order to determine the slice correspondences between images in sets \( S^h \) and \( S^m \), all sub-sets of \( S^m \) containing \( m \) 2D slices are compared with the 2D slices in \( S^h \). Then all these sub-sets \( S^A \), \( 1 \leq A \leq \binom{n}{m} \), are ranked in descending order based on the computed MI measure. Since the top-ranked subset may not always comprise the optimal correspondences, the slice correspondences are estimated via averaging the top \( K \) most highly ranked subsets. Our group-wise comparison scheme comprises the following steps:
Step 1: **Image sub-set generation**: From the n slices $S^n = \{C^{M,1}, \ldots, C^{M,n}\}$, generate a comprehensive list of all sub-sets $S^A$ \((1 \leq A \leq \binom{n}{m})\), each of which contains m images in sequential order.

Step 2: **Group-wise comparison between $S^n$ and $S^M$**: For each sub-set $S^A \subset S^M$, compute the group-wise image similarity, $D(S^n, S^A)$.

Step 3: **Refine the ranked list of slice correspondences**: Compute the average of the slice indices of the top ranked sub-sets in the ranked list of $S^A$. Use this result as the final slice correspondences.

These steps will be discussed in detail in the following subsections. A flowchart illustrating all these steps is shown in Fig. 3.

### 3.2. Group-wise comparison to estimate the slice correspondences

In order to determine the slice correspondences between images in $S^n$ and $S^M$, all sub-sets of $S^M$ containing m images are compared with $S^n$ using CMI. Then all these sub-sets $S^A$, \((1 \leq A \leq \binom{n}{m})\), are ranked in descending order according to the computed CMI measure. Since the top-ranked subset may not always comprise the optimal correspondences, the final slice correspondences are determined via averaging the top K most highly ranked subsets.

#### 3.2.1. Ranking image subsets based on CMI

We define the group-wise similarity measure $D$ between the image set $S^n = \{C^{H,1}, \ldots, C^{H,m}\}$ and an image sub-set $S^A = \{C^{M,a_1}, C^{M,a_2}, \ldots, C^{M,a_m}\}$ as

$$D(S^n, S^A) = \sum_{i=1}^{m} MI(C^{H,i}, C^{M,a_i}),$$

where $MI$ is mutual information between two images, $\alpha_i \in \{1, \ldots, n\}$, and $i \in \{1, \ldots, m\}$. A comprehensive list of all possible sub-sets $S^A \subset S^M$ of $S^M$ are generated, and each sub-set $S^A$ is compared with $S^n$ using Eq. (1). From the total \((\binom{n}{m})\) image sub-sets $S^A \subset S^M$, a ranked list of these image sub-sets in descending order $L = \{S^a\}$ is obtained, where $1 \leq u \leq \binom{n}{m}$, $D(S^n, S^{a_1}) > D(S^n, S^{a_2})$ when $1 \leq B_1 < B_2 \leq \binom{n}{m}$.

#### 3.2.2. Averaging top ranked subsets

**3.2.2.1. Measurement of the slice correspondence error.** One would expect that the top ranked image sub-set $S^1$ to be the most similar to the image set $S^n$. However, for highly dissimilar images such as the prostate histology and MRI with different slice spacing, the top ranked slice correspondences in $L$ may not always correspond to the correct match. For instance, assume that for $S^n = \{C^{M,a_1}, \ldots, C^{M,a_m}\}$, the ground truth for slice correspondences in $S^n$ is given by $S^* = \{C^{M,a_1}, \ldots, C^{M,a_n}\}$, where $1 \leq \epsilon_i$, $\alpha_i \leq n$. Let $V^A = [\epsilon_1 \epsilon_2 \ldots \epsilon_m]$ and $V^A = [\alpha_1 \alpha_2 \ldots \alpha_m]$ be the vectors of slice indices in $S^n$ and $S^A$, respectively. The slice correspondence error between ground truth $S^*$ and $S^n$ may be quantified via the L1-norm of the difference between the two vectors $V^A$ and $V^A$ as:

$$\psi(S^n, S^A) = \|V^A - V^A\|_1 = \frac{1}{m} \sum_{i=1}^{m} |\epsilon_i - \alpha_i|,$$

where $A \in \{1, \ldots, \binom{n}{m}\}$.

Ideally, one would expect that $\psi(S^n, S^A) = 0$. However, for the prostate data considered in this work, it may be that $\psi(S^n, S^A) = 0$, with $u = 1$. Fig. 4(a) shows the plot of $\psi(S^n, S^A)$, where $n = 18$ and $m = 5$. There are noticeable fluctuations along this curve, and interestingly $\psi(S^n, S^A) = 0$. Clearly, the assumption that the first
element $S^1$ from the ranked list $L$ is in correspondence with $S^H$ is incorrect. However, it is reasonable to assume that the majority of the slice correspondences for the top ranked subsets in $L$ are close to the best match. Thus we exploit the fact that a consensus in slice correspondences among the top ranked subsets in $L$ should yield the final correct set of slice correspondences. Another way to view this is by observing that the curve in Fig. 4(a) appears to have been obtained by adding noise to a smoother underlying curve. Intuitively, this prompts us to somehow smooth the “noisy” curve (Fig. 4(a)) to make it resemble a smoother monotonically increasing curve (Fig. 4(b)). At the end of this smoothing step, the newly refined subset $S^1$ ranked highest in the list $L$ should correspond to the final slice correspondences.

3.2.2.2. Averaging the slice correspondences. The algorithm for establishing a consensus among the slice correspondences in the top ranked subsets in $L$ proceeds as follows. The top $K$ most highly ranked subsets $S^u$, $u \in \{1, \ldots, K\}$ in $L$ based on the value of $D(S^A, S^H)$, $A \in \left\{ \begin{array}{c} 1, \ldots, \binom{n}{m} \end{array} \right\}$, are identified and the vectors comprising the indices of slice locations in $S^u$, $V^u = [v^u_1, v^u_2, \ldots, v^u_m]$, are determined. For each of $V^u$, $u \in \{1, \ldots, K\}$, $j \in \{1, \ldots, m\}$, we compute the average index $\hat{v}_j$ as:

$$\hat{v}_j = \left\lfloor \frac{1}{R} \sum_{u=1}^{K} v^u_j \right\rfloor,$$

where $\lfloor \cdot \rfloor$ denotes the rounding operation towards minus infinity. In this way, the final refined set of index locations for slice correspondences is obtained as the vector $V = [\hat{v}_1, \hat{v}_2, \ldots, \hat{v}_m]$. Hence the newly refined slice correspondences are obtained as $S = [c_{M,\hat{v}_1}^M, c_{M,\hat{v}_2}^M, \ldots, c_{M,\hat{v}_m}^M]$. 

---

**Fig. 2.** The complete framework to automatically determine the histology-MRI slice correspondences. The three modules are executed iteratively until there is no further change to the slice correspondence result. Module 1 will be expanded in detail in Fig. 3.
In the next stage, these estimated slice correspondences will be further improved through additional 2D and 3D registration.

4. 2D registration between the estimated corresponding slices

While the group-wise comparison approach allows for an initial estimate of histology-MRI slice correspondences, the image differences caused by the histological tissue distortion and the organ deformation during MRI scanning have not yet been compensated. Our aim is to reduce these artifacts, thereby further improving the slice correspondence result.

We achieve this by registering each histology slice to its estimated corresponding MRI slice. During the registration, the transformation is done only to the histology slices and not to the MRI slices. This is because there is no tissue loss or distortion in the MRI slices. In addition, we only perform a 2D affine registration instead of an elastic registration between each pair of corresponding histology and MRI slices. This is because within an iterative...
framework (such as the one we employ), an elastic registration at this stage will likely overfit each histology slice to a specific MRI slice, making further refinement of the slice correspondences unlikely. We choose the MI measure between the two images as the target function, and the simplex method is used in the optimization. The 2D registration procedure can then be formulated as

$$T_{2D}^i = \arg \max_T [MI(C^{M,i}, \Phi(C^{H,i}, T))],$$

where \(i \in \{1, \ldots, m\}\), \(T\) stands for a coordinate transformation, and \(\Phi\) represents a generic image transformation. Applying \(T_{2D}^i\) to \(C^{H,i}\) gives the registered image \(C^{H,i,2D}\):

$$C^{H,i,2D} = \Phi(C^{H,i}, T_{2D}^i).$$

Fig. 5 illustrates an example of the 2D registration between a histology slice and its estimated (via group-wise comparison) corresponding MRI slice. It can be seen that using the MRI slice as the reference image, the differences between histology and the MRI slices have been reduced, while at the same time, with the affine registration, the histology slice is not overly deformed to make it fit the MRI slice perfectly, so that the histology slice may end up being matched to a different MRI slice in subsequent iterations. This 2D registration module corresponds to Module 2 in Fig. 2.

5. 3D registration between the histology and MRI data

The slice correspondences obtained so far using the group-wise comparison and the 2D registration are based on the assumption that the orientation of the 2D histology slices and MRI slices are identical. This assumption is that there is no out-of-plane alignment [16], an issue that can only be resolved in the 3D space. We previously discussed that the two main obstacles in the 3D reconstruction of the prostate histological volume are (a) the limited number of histological slices with indeterminate inter-slice spacing, and (b) histological image artifacts in the form of tissue loss or distortion.

Group-wise comparison and 2D registration steps (in Modules 1 and 2) enable a coarse estimate of the histological inter-slice spacing (because the MRI inter-slice spacing is known) as well as to rectify the histological image artifacts to some extent. Therefore even though an accurate 3D reconstruction of the histological volume is still not attainable, we can still reconstruct a 3D histological pseudo-volume, where the limited histology slices are interlaced with additional zero-value slices (a process called zero-padding). With this histological pseudo-volume, 3D registration between the histology and MRI data can be performed, thereby allowing the histology slices to be matched to some non-axial MRI slices.

We obtain the the MRI volume \(Z^M\) by stacking all the 2D MRI slices \(C^{M,1}, \ldots, C^{M,n}\), since there is no inter-slice deformation in the MRI data. The histological pseudo-volume \(Z^H\) is obtained through zero-padding. As in the previous module, we still perform a 3D affine registration instead of an elastic registration in order to avoid the problem of overfitting. However, this time we choose to impose the 3D transformation to the MRI volume since there is no zero-value slice in the MRI volume. During the registration, the MI between the two volumes is once again chosen as the target function, and the simplex method is adopted as the optimization method. This 3D registration procedure is formulated as:

$$T_{3D} = \arg \max_T [MI(Z^H, \Phi(Z^M, T))],$$

and by applying \(T_{3D}\) to \(Z^M\) we get the registered MRI volume \(Z^M_{\text{registered}}\), where

$$Z^M_{\text{registered}} = \Phi(Z^M, T_{3D}).$$

Fig. 6 illustrates an example of this 3D registration procedure. For visualization purpose, surface rendering is used to represent the 3D volume. This 3D registration module corresponds to Module 3 in Fig. 2. After the registration, the registered MRI volume is re-sliced along the axial direction so that all the axial MRI slices are updated. These updated MRI slices together with the registered histology slices obtained in Module 2 are then used as input for the
group-wise comparison module (Module 1), so that the slice correspondences can be further refined. These 3 modules described in Sections 3–5 are iteratively executed until there is no further change to the slice correspondence result.

6. Experimental results

Our method was quantitatively evaluated in terms of its ability to determine the slice correspondences between $S^H$ and $S^M$ for 23 patient studies; all patient studies having been previously anonymized and de-identified. Table 1 gives a summary of the data used in these experiments. 3 Tesla endorectal T2-w MR images of the prostate were obtained in vivo with an image resolution 0.27 mm/pixel and slice spacing 2.2 mm. The prostate region of interest was manually extracted in these images by an expert radiologist. Following radical prostatectomy, the prostate histology slices were prepared using the protocol described in Section 1.1. All the 2D histological images were then resized by the same scale so that the histological images were of the same size as the MR images. In the experiments, the RGB color histological images were also converted to grayscale before the image features were computed.

For these experiments, the radiologist and the pathologist determined the ground truth slice correspondences through visual inspection of all the histology and MRI slices. Specifically, they opened two windows on the computer screen for the two types of images, respectively. By scrolling up and down the images in each window, the two experts visually compared the similarity in appearance between every pair of images based on their medical expertise before reaching an agreement on the final slice correspondences.

We also compared the performance of our method with the pair-wise comparison approach (described in Section 1.1) for determining slice correspondences. The error norm $\Psi$ described in Eq. (2) was used to compare the two approaches against the ground truth. Thus for each pair of $S^H$ and $S^M$, the errors of our method ($\Psi_{GW}(S^H, S^M)$) and the pair-wise comparison approach ($\Psi_{PW}(S^H, S^M)$) were obtained and compared. Note that since the pair-wise comparison often matches more than one histology slices to a single MRI slice, the 2D and 3D registration as implemented in our method was not done following the pair-wise comparison, because these additional registration procedures not only cannot correct this multi-to-one matching mistake, but also are likely to re-enforce this mistake.

![Fig. 6. Example of the 3D affine registration between the histology pseudo-volume and the MRI volume. Surface renderings of the (a) MRI volume, (b) histology pseudo-volume, and (c) the two volumes in a common coordinate system post 3D registration.](image)

**Fig. 7** shows the histology-MRI slice correspondences for one of the patient studies (Study #14). **Fig. 7(a)–(d)** shows all the 4 histology slices from this patient study. **Fig. 7(e)–(h)** shows the corresponding MRI slices determined by our method. The corresponding MRI slices determined by the pair-wise comparison approach are shown in **Fig. 7(i)–(l)**, while the ground truth of the corresponding MRI slices established via a consensus of experts are shown in **Fig. 7(m)–(p)**. Note that the number in the top right corner of each corresponding MR image is the index of this MRI slice. It can be seen that for this patient study, our result is at the most only one slice away from the ground truth, while multiple histology slices are matched to a common MRI slice when the pair-wise comparison approach is used (**Fig. 7(i)–(l)**).

The quantitative analysis of the experimental results on all the 23 patients is illustrated in **Fig. 8**, where a bar plot of $\Psi_{GW}$ and $\Psi_{PW}$ for all the patient studies is shown. Across 23 patient studies, the values of $\Psi_{GW}$ range between 0 and 2.7, while the values of $\Psi_{PW}$ range between 1.7 and 5.2. A paired student t-test was also performed, the null hypothesis being that no significant difference exists between these two types of errors. For a two-tailed paired t-test, a $p$-value of $8.15 \times 10^{-11}$ was obtained.

7. Discussion

Mapping the spatial disease extent in a certain anatomical organ/tissue from histology images to radiological images is important in defining the disease signature in the radiological images. We consider the specific application of an MRI-based CAD system for the detection and diagnosis of prostate cancer, where prostate cancer extent from histology needs to be mapped to MRI to train the CAD system. The disease mapping can be done through 3D registration between the histological and the MRI volumes. A prerequisite in this process is the accurate 3D reconstruction of the histological volume in advance, which typically requires a large number of finely sectioned histology slices and/or time-consuming procedures to correct the histological artifacts (as discussed in Section 1.1). However, for the specific application of training a CAD system, a large amount of training data from a large number of patients is usually needed. Consequently, the complex procedures that are necessary to reconstruct the 3D histological volume are usually not feasible within the time and labor constraints imposed by routine clinical practice. In this case, the only solution is to register each 2D histology slice to its corresponding 2D MRI slice separately, which requires the 2D histology-MRI slice correspondences be determined in advance. For this purpose, we have come up with a framework comprising iterative implementation of group-wise comparison, 2D and 3D registration. Experimental results show the slice correspondences determined by our method are close to the ground truth correspondences determined by the experts. In addition, a statistical t-test indicates that our method significantly outperforms the pair-wise comparison approach.

As described in Section 5, because of the possible out-of-plane alignment problem, a 3D histological pseudo-volume of the

<table>
<thead>
<tr>
<th>Number of studies</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Modalities</strong></td>
</tr>
<tr>
<td><strong>Number of slices in each study</strong></td>
</tr>
<tr>
<td>23</td>
</tr>
<tr>
<td>$S^H$: Prostate histology (whole mount)</td>
</tr>
<tr>
<td>$S^M$: Prostate 3 T endorectal T2 in vivo MRI</td>
</tr>
</tbody>
</table>

**Table 1** Description of data sets used for the experiment.
Fig. 7. Examples of slice correspondence results obtained via different approaches. (a)–(d) The prostate histology slices, (e)–(h) corresponding MRI slices determined by our method, (i)–(l) corresponding MRI slices determined by pair-wise comparison approach, and (m)–(p) expert determined correspondences which serve as the ground truth. The number in the top right corner of (e)–(p) refers to the index of each MRI slice. For this patient study, our result is at most only 1 slice away from the ground truth, while in the result of pair-wise comparison, multiple histology slices are matched to a common MRI slice and the slice differences are larger. (The dotted curves in the histology slice are the cancer spatial distribution outlined by the pathologist.)

Fig. 8. Plots of $\Psi_{GW}$ and $\Psi_{PW}$ for the 23 patient studies used in our experiment. The number along the X-axis indicates each individual patient study. The values of $\Psi_{GW}$ range between 0 and 2.7, while the values of $\Psi_{PW}$ ranges between 1.7 and 5.2. This indicates superior performance of our method over the pair-wise comparison approach.
prostate is constructed through zero-padding. With this pseudo-volume, a 3D registration between the histology and the MRI data can be performed. However, we would like to point out that the 3D pseudo-volume reconstructed in this way contain some zero-value slices, while the 3D registration only utilizes the voxels on the non-zero slices. Therefore, improvement over the slice correspondences resulting from this 3D registration procedure is only moderate in most cases. This is also empirically illustrated in our experimental results, which suggest that in most cases, there is only a small change in the slice correspondences after the 3D registration. The group-wise comparison (Module 1) appears to play a more important role in the determination of slice correspondences.

The group-wise comparison is done on the segmented histology and MRI slices. While the segmentation of prostate on the histological images is relatively straightforward, the segmentation of prostate on the in vivo T2-w MRI images is more challenging. So far in all our experiments, the segmentation of the prostate on MR images was done by an experienced radiologist. It should, however, be pointed out that this manual segmentation may be subject to some inter-observer variability. In case this inter-observer variability is large for some special patient studies, the slice correspondences result obtained using our group-wise scheme might vary with the quality of MRI segmentation by a different radiologist. A possible solution is that in face of these extreme cases where the radiologist is very doubtful of the exact contours of the prostate in the MR images, instead of using our method, the slice correspondences will be determined by the consensus of the radiologist and the pathologist. In the next stage, we will focus on removing the dependency of our algorithm on segmented prostate images.

8. Concluding remarks

In this paper, we presented a complete framework to automatically determine slice correspondence between images from histology and MRI. Our method consists of three modules. In the first module, an initial estimate of the slice correspondences is obtained via a group-wise comparison of the image similarity. Based on these group slices correspondences, in the second module, each histological slice is registered to its corresponding MRI slice in order to compensate the image differences caused by the possible histological artifacts and prostate organ deformation. In the third module, a 3D histological pseudo-volume is reconstructed, and 3D registration between the histology pseudo-volume and the MRI volume is then performed to gradually correct for the out-of-plane alignment problem. These three modules are executed iteratively until there is no further change to the slice correspondence result. The prostate histology-MRI slice correspondences obtained using our method were compared with the ground truth slice correspondences determined via consensus of multiple experts. The accuracy of our method was in most instances close to the results given by the experts. Additionally, our method was found to be better than the pair-wise comparison approach. Future work will involve using the slice correspondences determined via our scheme to drive a multi-modal registration method for accurate mapping of disease signatures from ex vivo histopathology onto corresponding in vivo MRI in order to train and evaluate a CAD system.

Conflict of interest statement

The authors do not have any financial and personal relationships with other people or organizations that could inappropriately influence the research presented in this paper.

Acknowledgments

This work was made possible via grants from the Wallace H. Coulter Foundation, National Cancer Institute (Grant Nos. R01CA136535-01, R01CA140772-01, and R03CA143991-01), and the Cancer Institute of New Jersey. The authors would also like to acknowledge Arjun Kalyanpur and Johnson Henry from Teleradiology Solutions, Bangalore, India for their assistance in providing the expert segmentations.

References


