Multimodal image registration of *ex vivo* 4 Tesla MRI with whole mount histology for prostate cancer detection

Jonathan Chappelow[a], Anant Madabhushi[a], Mark Rosen[b], John Tomaszeweski[c], Michael Feldman[c]

[a]Rutgers University Department of Biomedical Engineering, 599 Taylor Road, Piscataway, NJ, USA 08854

[b]Department of Radiology, [c]Department of Pathology, University of Pennsylvania, 3400 Spruce Street, Philadelphia, PA, USA 19104

ABSTRACT

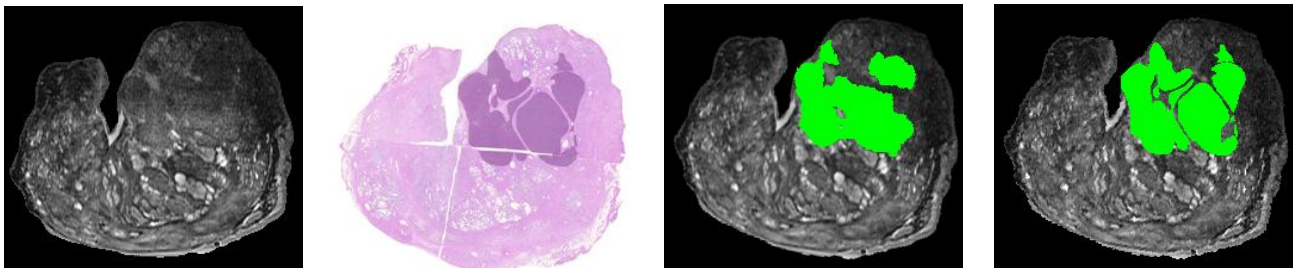
In this paper we present novel methods for registration and subsequent evaluation of whole mount prostate histological sections to corresponding 4 Tesla *ex vivo* magnetic resonance imaging (MRI) slices to complement our existing computer-aided diagnosis (CAD) system for detection of prostatic adenocarcinoma from high resolution MRI. The CAD system is trained using voxels labeled as cancer on MRI by experts who visually aligned histology with MRI. To address voxel labeling errors on account of manual alignment and delineation, we have developed a registration method called combined feature ensemble mutual information (COFEMI) to automatically map spatial extent of prostate cancer from histology onto corresponding MRI for prostatectomy specimens. Our method improves over intensity-based similarity metrics (mutual information) by incorporating unique information from feature spaces that are relatively robust to intensity artifacts and which accentuate the structural details in the target and template images to be registered. Our registration algorithm accounts for linear gland deformations in the histological sections resulting from gland fixing and serial sectioning. Following automatic registration of MRI and histology, cancer extent from histological sections are mapped to the corresponding registered MRI slices. The manually delineated cancer areas on MRI obtained via manual alignment of histological sections and MRI are compared with corresponding cancer extent obtained via COFEMI by a novel registration evaluation technique based on use of non-linear dimensionality reduction (locally linear embedding (LLE)). The cancer map on MRI determined by COFEMI was found to be significantly more accurate compared to the manually determined cancer mask. The performance of COFEMI was also found to be superior compared to image intensity-based mutual information registration.

Keywords: registration, prostate cancer, CAD, dimensionality reduction, mutual information, COFEMI, histology, MRI, multimodality, evaluation, LLE

1. INTRODUCTION

Prostatic adenocarcinoma (CAP) is the most common malignancy among men with 234,460 new cases and 27,350 deaths estimated to have occurred in 2006 (*American Cancer Society*). We have recently developed[1] computer-aided diagnosis (CAD) algorithms for early detection of CAP on high resolution magnetic resonance imaging (MRI). The core of the CAD system is a supervised classifier that is trained to identify cancer voxels on MRI via labeled instances. The supervised classifier was trained using voxels manually labeled as cancer by experts on 4 Tesla *ex vivo* MRI of prostatectomy specimens for which corresponding whole mount histology was available.[1] The experts manually delineated spatial extent of cancer on individual MRI sections (Figure 1(c)) by visually registering the corresponding MRI (Figure 1(a)) and histology sections (Figure 1(b)). This procedure was however shown to be prone to errors and subject to observer bias.[2] We have previously demonstrated that expert labeling errors can severely compromise the classification accuracy of CAD models as well as the subsequent CAD evaluation.[2] To address the issue of expert labeling errors due to manual registration, we propose a novel multimodality image registration routine called combined feature ensemble mutual information (COFEMI). We

Medical Imaging 2007: Image Processing, edited by Josien P. W. Pluim, Joseph M. Reinhardt,
Proc. of SPIE Vol. 6512, 65121S, (2007) · 1605-7422/07/$18 · doi: 10.1117/12.710558



(a) (b) (c) (d)

Figure 1. (a) *Ex vivo* prostate MRI slice, (b) corresponding histology slice with cancer ground truth (dark stain) determined by microscopic analysis of haematoxylin and eosin stained tissue, (c) determined extent on MRI by visually aligning (a) and (b), and (d) automatically mapped cancer extent from (b) onto (a) via the COFEMI method. Note dissimilarities between the manually (c) and automatically determined (d) cancer masks.

demonstrate the utility of this technique in mapping the cancer extent from whole slide histological sections onto corresponding ex vivo MRI images (Figure 1(d)).

Registration of images from diverse modalities such as histology and radiology is complicated on account of the uniquely different image characteristics of the individual modalities. Hence, the appearance of tissue and anatomical structures on MRI and histology are significantly different. Tissue morphology of the whole mount histology is also significantly altered due to uneven tissue fixation, gland slicing and sectioning resulting in duct dilation, gland deformation, and tissue loss (note the white spaces in the gland in Figure 1(b)). Some researchers have investigated the use of mutual information (MI) for intramodality registration of in vivo MR[3] and CT images[4] of the prostate. The task of automatically registering histological sections of the prostate with MRI has been addressed for small rat brain slices,[5] however the histology is not cut into quarters on account of small specimen size. The large human prostate specimens in our data sets undergo significantly more tissue loss and deformation as a result of their size and the extra quartering step.

Mutual information (MI) is widely used[6] as a similarity measure for multimodal image registration, due to the robustness of the MI measure to differences in intensity patterns between image modalities and protocols. Most MI-based registration techniques are based on the assumption that a consistent statistical relationship exists between the intensities of the two images being registered. However in cases where considerable image differences between modalities (e.g. MRI and histology) exist, and when imaging artifacts (e.g. bias field inhomogeneity in MRI)[7] are present, intensity based MI registration schemes are inadequate. The limitations of image intensity to drive the MI similarity metric has resulted in efforts to complement intensity information with additional features derived by transformation of image intensities. Several MI-based techniques utilizing image gradients,[8] cooccurrence information,[9] color[10] and image segmentations[11] have been proposed in conjunction with various formulations of the similarity metric. Several of the similarity metric formulations proposed have been *ad hoc*[10] while several do not provide sufficient justification on the choice of image features to complement image intensity.

To register the *ex vivo* MRI and histological images of the prostate, we present a new registration method called COFEMI that works by maximizing the combined mutual information (CMI) shared by an intensity image A and multiple representations $B'_1, B'_2, \ldots, B'_n$ of intensity image B in n feature spaces. COFEMI provides enhanced registration performance by complementing the intensity information in B with transformed feature images $B'_1, B'_2, \ldots, B'_n$, all derived from B, but which are not as susceptible to intensity artifacts as B and which provide additional similarity information regarding A not contained in B. Hence COFEMI is more robust to the differences between images of different modalities and less susceptible to imaging artifacts than the intensity image. The combination of information from several feature images also provides a more robust similarity metric compared to the use of a single feature image. The CMI formulation presented in this paper combines only unique information provided by each additional feature image, and has been demonstrated to be an appropriate mechanism for combining multiple semi-independent sources of information.[9] We also present a novel method for optimally selecting image features to complement image intensity for the registration scheme. Even though

Contact: Anant Madabhushi: E-mail: anantm@rci.rutgers.edu, Telephone: 1 732 445 4500 x6213

we compute nearly 300 feature images for each intensity image, our feature selection algorithm selects in an unsupervised fashion the top few informative feature representations.

Quantitative evaluation of image registration is an important component of gauging the efficacy of registration methods. Registration evaluation like with segmentation or CAD evaluation is tricky for clinical studies because a *ground truth* estimate of spatial location of the target is required. Registration accuracy is commonly gauged by overlap of accurately segmented regions in both images being registered. Hence for the datasets considered in this study, knowledge of the true spatial extent of cancer on the MRI studies would have been useful in quantitatively evaluating the performance of the COFEMI method. To our knowledge no generalizable method for evaluating registration performance in the absence of definitive ground truth of either alignment or region of interest segmentation is currently available. A technique for evaluation of segmentation accuracy in the absence of ground truth has been demonstrated which generates a surrogate of ground truth by combining multiple expert segmentations.[12] However, this approach is not practical for establishing definitive cancer extent on MRI of the prostate due to high inter- and intra-observer variability of expert cancer segmentations.[1] Thus, our second major contribution in this paper is to present a novel registration evaluation technique wherein ground truth for a target object is not required.

The main spirit behind our evaluation technique is the following. Consider registration method R_1 to align images A and B. Further consider an area a^A corresponding to a single object class on A. Let a^{B,R_1} be the corresponding registered map of a^A on B. If a^{B,R_1} has been accurately mapped onto B via method R_1 and since we know that a^A corresponds to a single object class it is reasonable that all the pixels in image B corresponding to a^{B,R_1} will cluster together in a low dimensional feature space. Objects that are far removed from the cluster represent registration errors. For the problem considered in this paper, we analyze the clustering of pixels on MRI corresponding to the cancer masks mapped from histology via the COFEMI scheme and manual mapping. A number of different feature operators are applied to the cancer mask and every pixel in the mask is characterized by a high dimensional feature vector. Locally linear embedding (LLE),[13] a non-linear dimensionality reduction method is then used to cluster the pixels within the mask on MRI in a low dimensional space. The main spirit of the non-linear dimensionality reduction method[2] is that distances between objects in the high dimensional feature space are preserved in the low dimensional embedding space. Hence those pixels on MRI corresponding to the cancer mask on histology obtained via COFEMI which are far removed from the dominant cluster are identified as registration errors, our assumption being that all cancer pixels will cluster together in the low dimensional feature space. In other words, different classes (cancer and non-cancer) will exhibit different textural characteristics and occupy separate distinct clusters. We apply this evaluation scheme to compare automatic with manual registration and determine false positive and false negative errors associated with both methods. The location and number of the outliers are used to quantify registration/segmentation accuracy and also to determine the spatial locations where the methods failed to correctly align the histological sections with the corresponding MR slices. To reiterate, the novel aspects of the work presented in this paper are,

- a novel multimodality image registration routine (COFEMI) that utilizes a generalized formulation of MI and an ensemble of feature images to overcome the inherent limitations associated with using image intensity-based similarity metrics,
- a novel method for optimally selecting only the most informative features for use in conjunction with COFEMI,
- a novel method for quantitative registration evaluation in the absence of a definitive ground truth for a target object and which a uses non-linear dimensionality reduction (LLE) method.

The rest of this paper is organized as follows. In Section 2 we describe the new COFEMI registration methodology along with the new similarity metric formulation and description of the feature extraction and selection processes necessary for COFEMI. Section 3 describes our new methodology for quantitative registration evaluation using LLE. In Section 4 we present the results of qualitative and quantitative evaluation of the COFEMI scheme. Concluding remarks and future research directions are presented in Section 5.

Table 1. List of notation and symbols used in this paper.

Symbol	Description	Symbol	Description
$\mathcal{C}^{ex}$	2D MRI scene	$\mathcal{C}^{H}$	2D Histology scene
C	2D array of pixels c	$f(c)$	Image intensity at c
$T(\mathcal{C}^{H})$	Set of cancer pixels on $\mathcal{C}^{H}$	$E(\mathcal{C}^{ex})$	Cancer pixels determined manually on $\mathcal{C}^{ex}$
Φ_{β}	Feature operator where $\beta \in \{1,\ldots,n\}$	$f^{\Phi_{\beta}}(c)$	Feature value at c for feature Φ_{β}
$\mathcal{C}^{ex,r}$	$\mathcal{C}^{ex}$ after COFEMI	$\mathcal{C}^{\Phi_{\beta}}$	Feature scene for Φ_{β}
$E(\mathcal{C}^{ex,r})$	$E(\mathcal{C}^{ex})$ on $\mathcal{C}^{ex,r}$	$T(\mathcal{C}^{ex,r})$	$T(\mathcal{C}^{ex})$ on $\mathcal{C}^{ex,r}$

2. METHODOLOGY

2.1. Notation

We define a 2D image or scene by the notation $\mathcal{C} = (C, f)$ where C is a coordinate grid, called the scene domain, and f is a function that assigns to every pixel $c \in C$ an intensity value $f(c)$. The slices of the *ex vivo* MRI are referred to as $\mathcal{C}^{ex}$, while the corresponding histology slices are denoted by $\mathcal{C}^{H}$. The list of notations and symbols used in the rest of this paper are listed in Table 1.

2.2. Data Acquisition

Prostate specimens obtained via radical prostatectomy at the Hospital at the University of Pennsylvania were embedded in 2% agar in a small Plexiglas box. The embedded gland was then imaged using a 4 Tesla (T) whole body Siemens Trio MR scanner using 2D fast spin echo, TE 126 msec, TR 3000, 15.6 khz, and 4 signal averages to obtain $\mathcal{C}^{ex}$. MRI slices and histological slices were maintained in the same plane of section to ensure slice correspondences by both leveling the embedded gland and by using a rotary knife to cut 4μm thick serial sections of the gland at 1.5 mm intervals (correlating to 2 standard MRI slices). Whole mount histological sections ($\mathcal{C}^{H}$) were stained with Hematoxylin and Eosin (H&E). Each 1.5 mm slice was cut into quarters and the histological samples were microscopically examined to determine cancer extent on $\mathcal{C}^{H}$, denoted by the set of pixels $T(\mathcal{C}^{H})$. The histological sections were scanned with a whole slide scanner and the quarters were reconstructed using Adobe Photoshop*. Histological images were converted to gray scale and resampled from their original resolution of 0.07mm/pixel (2000 x 2000 pixels) to match the resolution of the 2D MRI slice (256 x 256 pixels). Since there were roughly twice the number of MRI slices as there were histological sections for each patient study, the MRI sections corresponding to the histological sections were manually identified by an expert by visually identifying anatomical features on $\mathcal{C}^{H}$ and $\mathcal{C}^{ex}$. Two prostate studies comprising 15 and 11 histological sections were considered in this study.

2.3. Manual Mapping of Cancer from Histology to MRI

Expert segmentations of cancer on *ex vivo* MRI, denoted by the set of pixels $E(\mathcal{C}^{ex})$, are obtained by manually mapping $T(\mathcal{C}^{H})$ onto $\mathcal{C}^{ex}$. Note that the manual delineation of cancer $E(\mathcal{C}^{ex})$ on $\mathcal{C}^{ex}$ was guided by the spatial extent of cancer $T(\mathcal{C}^{H})$ on corresponding $\mathcal{C}^{H}$ along with visual cues from the anatomical features of the prostate section.

2.4. COFEMI

Shannon entropy, $S(\mathcal{C}^{A})$, is a measure of information content of an image $\mathcal{C}^{A}$, whereas the joint entropy $S(\mathcal{C}^{A}\mathcal{C}^{B})$ of two images $\mathcal{C}^{A}$ and $\mathcal{C}^{B}$ describes the information gained by combined simultaneous knowledge of both images.[6] The MI of a pair of images or random variables is commonly defined in terms of marginal and joint entropies by $I_2(\mathcal{C}^{A},\mathcal{C}^{B}) = S(\mathcal{C}^{A}) + S(\mathcal{C}^{B}) - S(\mathcal{C}^{A}\mathcal{C}^{B})$, wherein $I_2(\mathcal{C}^{A},\mathcal{C}^{B})$ is maximized by minimizing joint entropy $S(\mathcal{C}^{A}\mathcal{C}^{B})$ and maintaining the marginal entropies $S(\mathcal{C}^{A})$ and $S(\mathcal{C}^{B})$. Hence $I_2(\mathcal{C}^{A},\mathcal{C}^{B})$ is used to describe the interdependence of multiple variables, or graylevels of a set of images,[6] and it is assumed that the global MI maximum will occur at the point of precise image registration, when uncertainty about $\mathcal{C}^{A}$ is maximally explained

*http://www.adobe.com/products/photoshop/

by $\mathcal{C}^B$. However, when $\mathcal{C}^A$ and $\mathcal{C}^B$ are very different, uncertainty about one image given the other increases and more information is often required to determine the optimal alignment of $\mathcal{C}^A$ and $\mathcal{C}^B$. The presence of intensity artifacts (e.g. MRI inhomogeneity and posterior-acoustic shadowing in ultrasound), and if $\mathcal{C}^A$ and $\mathcal{C}^B$ correspond to different image modalities, intensity-based registration methods to align $\mathcal{C}^A$ and $\mathcal{C}^B$ prove inadequate. COEFEMI operates by maximization of combined mutual information (CMI) of the image $\mathcal{C}^A$ with a feature ensemble associated with $\mathcal{C}^B$. These feature images represent higher-order statistical and texture representations of the MR intensity image that (1) contain additional MI with $\mathcal{C}^A$ and (2) provide additional information regarding $\mathcal{C}^B$. The CMI of image $\mathcal{C}^A$ with two images $\mathcal{C}^B$ and $\mathcal{C}^{B'}$ has been previously defined[14] as $I_2(\mathcal{C}^A, \mathcal{C}^B\mathcal{C}^{B'}) = S(\mathcal{C}^A) + S(\mathcal{C}^B\mathcal{C}^{B'}) - S(\mathcal{C}^A\mathcal{C}^B\mathcal{C}^{B'})$, and represents the reduction in uncertainty about $\mathcal{C}^A$ given simultaneous knowledge of both $\mathcal{C}^B$ and $\mathcal{C}^{B'}$. This can be extended to obtain a generalized formulation of CMI for $\mathcal{C}^A$ given an image ensemble $\mathcal{C}^B\mathcal{C}^{B'_1}\mathcal{C}^{B'_2}\cdots\mathcal{C}^{B'_n}$ as,

$$I_2(\mathcal{C}^A, \mathcal{C}^B\mathcal{C}^{B'_1}\mathcal{C}^{B'_2}\cdots\mathcal{C}^{B'_n}) = S(\mathcal{C}^A) + S(\mathcal{C}^B\mathcal{C}^{B'_1}\mathcal{C}^{B'_2}\cdots\mathcal{C}^{B'_n}) - S(\mathcal{C}^A\mathcal{C}^B\mathcal{C}^{B'_1}\mathcal{C}^{B'_2}\cdots\mathcal{C}^{B'_n}) \quad (1)$$

where $\mathcal{C}^{B'_1}\mathcal{C}^{B'_2}\cdots\mathcal{C}^{B'_n}$ are feature images directly derived by transformation of image intensities in $\mathcal{C}^B$. This formulation of CMI incorporates only the unique information of images in an ensemble, thus enhancing but not overweighting the similarity metric with redundant information. The boundedness of CMI is illustrated in Figure 2 for ensembles of feature images $\mathcal{C}^{B'_1}\mathcal{C}^{B'_2}\cdots\mathcal{C}^{B'_n}$ of increasing dimensionality. CMI approaches an asymptote (Figure 2) equal to $S(\mathcal{C}^A)$, the total information content of $\mathcal{C}^A$. Although joint entropy estimates become increasingly overestimated as the ensemble size increases,[6] Figure 2 illustrates that the formulation in Equation 1 behaves in a monotonically-increasing but asymptotic manner with inclusion of additional feature representations into the ensemble. Examples of these feature images for a MRI slice $\mathcal{C}^{ex}$ are shown in Figure 3.

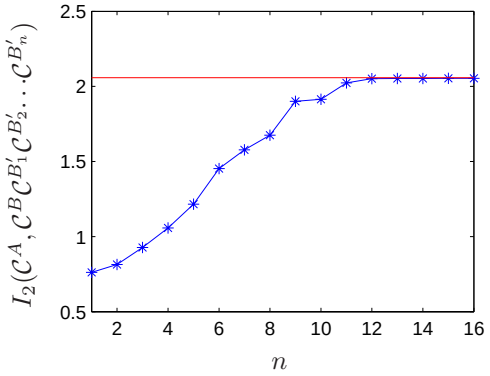


Figure 2. Bounded behavior of CMI for feature ensembles of increasing dimensionality. Note that $I_2(\mathcal{C}^A, \mathcal{C}^B\mathcal{C}^{B'_1}\mathcal{C}^{B'_2}\cdots\mathcal{C}^{B'_n})$ increases monotonically and reaches an asymptote at the value of $S(\mathcal{C}^A)$. Hence by incorporating additional feature scenes more uncertainty regarding $\mathcal{C}^A$ is explained.

2.5. Feature Extraction

We compute a total of $n = 284$ unique feature images from each $\mathcal{C}^{ex}$. These feature images comprise (i) gradient, (ii) first order statistical and (iii) second order statistical texture features. These feature representations were chosen since they had been demonstrated to be able to discriminate between the cancer and non-cancer classes.[1] In addition several of these were found to be feature representations less susceptible to image artifacts such as inhomogeneity. We calculated the feature scenes $\mathcal{C}^{\Phi_\beta} = (C, f^{\Phi_\beta})$ for each $\mathcal{C}^{ex}$ by applying the feature operators Φ_β, $\beta \in \{1, \ldots, n\}$ within a local neighborhood associated with every pixel c in $\mathcal{C}^{ex}$. Hence $f^{\Phi_\beta}(c)$ is the feature value associated with feature operator Φ_β at pixel c. Figure 3(c)-(h) show first and second order statistical feature scenes $\mathcal{C}^{\Phi_\beta}$ calculated from a slice $\mathcal{C}^{ex}$ (Figure. 3(a)) for several Φ_β. These feature images are of very different appearance as a result of the different texture operators and window sizes used in their calculation.

A. Gradient Features

Gradient features are calculated using steerable and non-steerable linear gradient operators. Eleven non-steerable

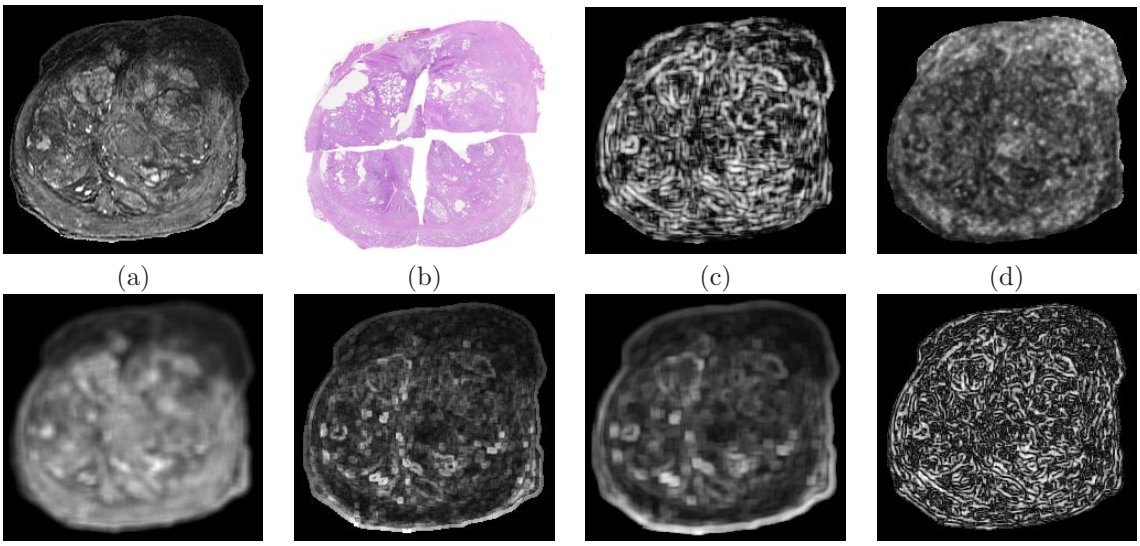


Figure 3. (a) $\mathcal{C}^{ex}$ and feature scenes $\mathcal{C}^{\Phi_\beta}$ corresponding to (c) Haralick correlation (8×8), (d) Haralick inv. diff. moment (5×5), (e) mean intensity (5×5), (f) range (5×5), (g) range (8×8), (h) and Haralick correlation (5×5), where 5×5 and 8×8 refer to the window sizes within which the local textural operations were applied. Figure 3(b) shows the histological section $\mathcal{C}^{H}$ corresponding to $\mathcal{C}^{ex}$ in (a).

gradient features were obtained using Sobel, Kirsch and standard derivative operations. Gabor gradient operators[15] comprising the steerable class of gradient calculations were defined for every pixel $c = (x, y)$,

$$f^{Gab}(c) = e^{\frac{(x'+y')}{\sigma}} \cos(2\pi\lambda x') \tag{2}$$

where λ is the spatial frequency (scale) of the sinusoid and σ is the variance of the Gaussian envelope, where $f^{Gab}(c)$ represents the magnitude of the Gabor filter response at pixel c. The orientation of the filter is affected by φ through the coordinate transformations: $x' = x\cos\varphi + y\sin\varphi$ and $y' = x\cos\varphi - y\sin\varphi$. Gabor gradient features were calculated on 6 scales, 8 orientations and 3 window sizes.

B. First Order Statistical Features

Four first order statistical features for window sizes of 5x5 and 8x8 pixels, including mean, median, standard deviation, and range, and were calculated as described in Ref. 1.

C. Second Order Statistical Features

To calculate the second order statistical feature scenes, we define a $\kappa \times \kappa$ square neighborhood $\mathcal{N}_{c,\kappa}$ associated with every pixel $c \in C$ and centered on c. Hence any pixel $d \in \mathcal{N}_{c,\kappa}$ is a κ neighbor of c. We now compute a $M \times M$ cooccurrence matrix $P_{d,c,\kappa}$ associated with $\mathcal{N}_{c,\kappa}$, where M is the maximum gray scale intensity in $\mathcal{C}$. The value at any location $[i, j]$, where $i, j \in \{1, \ldots, M\}$, in the matrix $P_{d,c,\kappa}[i, j]$ represents the frequency with which two distinct pixels $u, v \in \mathcal{N}_{c,\kappa}$ with associated image intensities $f(u) = i$, $f(v) = j$ are separated by distance d. We further define a normalized $M \times M$ cooccurrence matrix $\hat{P}_{c,\kappa}$ where the value of the matrix at any location $[i, j]$ is given as,

$$\hat{P}_{c,\kappa}[i, j] = \frac{1}{|d|} \sum_{d} P_{d,c,\kappa}[i, j] \tag{3}$$

The criteria for cooccurrence was defined by $d \in \{1, 2\}$ in order to provide an enriched description of the spatial graylevel dependence within $\mathcal{C}^B$. A total of 13 Haralick features, energy, entropy, inertia, correlation, inverse difference moment, information correlation measures 1 and 2, sum average, sum variance, sum entropy, different average, difference variance and difference entropy were extracted at every pixel $c \in C$, from $\hat{P}_{c,\kappa}$, for $\kappa \in \{5, 8\}$, and $M \in \{64, 128, 256\}$.

2.6. Feature Selection

While for each scene, $\mathcal{C}^B$, nearly 300 feature images $\mathcal{C}^{B'_1}\mathcal{C}^{B'_2}\cdots\mathcal{C}^{B'_n}$ are computed, COFEMI requires only a subset of the most informative features to achieve accurate registration. We represent an ensemble of k images taken from a total of n feature images as $\pi_k = \mathcal{C}^{B'_{\alpha_1}}\mathcal{C}^{B'_{\alpha_2}}\cdots\mathcal{C}^{B'_{\alpha_n}}$ where $\{\alpha_1, \alpha_2, \ldots, \alpha_k\} \in \{1, \ldots, n\}$. Since it is not possible to estimate the joint histogram for the maximally informative ensemble of n semi-independent descriptive features, an ensemble of $k << n$ features must be generated. However, a brute force approach to determining an optimal π_k for even $k = 5$ from $n = 284$ features involves computing $\binom{284}{5}$ ensembles which is well over 10^{10}. Consequently we propose a novel feature selection algorithm to select the optimal feature ensemble. In practice, ensembles of more than 5 features cause the gray level histogram to become too dispersed to provide a meaningful joint entropy estimate, hence for this problem we limit $k = 5$ features. Although a joint entropy estimate becomes overestimated even for a single additional image (e.g. $S(\mathcal{C}^A\mathcal{C}^B\mathcal{C}^{B'_1})$), joint entropy for increasing numbers of images are still *meaningful* and functional in a similarity metric.

Algorithm $CMIfeatures$
Input: $\mathcal{C}^B$, k, $\mathcal{C}^{B'_1}\mathcal{C}^{B'_2}\cdots\mathcal{C}^{B'_n}$.
Output: π_k.
begin
0. Initialize π_k, Ω as empty ques;
1. *for* $u = 1$ to k
2. *for* $v = 1$ to n
3. *If* $\mathcal{C}^{B'_v}$ is present *then*
3. Insert $\mathcal{C}^{B'_v}$ into π_k; $\Omega[v] = S(\mathcal{C}^B\pi_k)$;
4. Remove $\mathcal{C}^{B'_v}$ from π_k;
5. *endif*;
5. *endfor*;
6. $m = \text{argmax}_v\, \Omega[v]$; Insert $\mathcal{C}^{B'_m}$ into π_k;
7. *endfor*;
end

The main spirit behind the algorithm is to compute the Shannon entropy between $\mathcal{C}^B$ and π_k when a single feature image $\mathcal{C}^{B'_v}$ is inserted into π_k in a pairwise fashion over all $v \in \{1, \ldots, n\}$. At the end of each set of n iterations, the feature image $\mathcal{C}^{B'_v}$ determined to be most informative is inserted into the queue π_k and removed from further consideration. The process is repeated until $k-1$ additional unique feature images are inserted into the ensemble.

2.7. COFEMI Registration

Scenes $\mathcal{C}^A$ and $\mathcal{C}^B$ are aligned using an affine transformation with a total of five parameters for rotation (θ_z), translation (dx, dy), and scaling (ς_x, ς_y). The affine transformation matrix ($\mathbf{M}$) is constructed by combination of three coordinate transformation matrices corresponding to rotation in the plane of the image ($\mathbf{R_z}$), translation in the x and y directions ($\mathbf{T_{xz}}$), and horizontal and vertical scaling ($\mathbf{S_{xy}}$). The individual affine transformation matrices are composed in terms of the five parameters ($\theta_z, dx, dy, \varsigma_x, \varsigma_y$) by common geometric coordinate transformations equations. Homogeneous transformation matrices and coordinate representations were utilized to apply translations in the same matrix operation as scaling and rotation. Thus for a homogeneous coordinate $c_h = (x, y, z, 1)^T$, the transformed coordinate is determined by $c'_h = \mathbf{M} * c_h$, where $*$ indicates matrix multiplication. Deformations are applied in the order of rotation, scaling, and then translation by the product $\mathbf{M} = \mathbf{Txy} * \mathbf{Sxy} * \mathbf{R_z}$. Both the COFEMI and image intensity-based MI registration techniques utilize NN interpolation. Intensity based MI registration is achieved by optimizing (via the Nedler Mead simplex algorithm) the affine transformation, $\mathbf{M}^*$, by maximization of MI of $\mathcal{C}^A$ with $\mathcal{C}^B$. The COFEMI procedure operated by maximization of CMI of $\mathcal{C}^A$ and ensemble $\mathcal{C}^B\pi_k$. The transformation $\mathbf{M}^*$ for which $I_2(\mathcal{C}^A, \mathcal{C}^B_m\pi_{k,m})$ is maximized is determined as the optimal affine transformation between $\mathcal{C}^A$ and $\mathcal{C}^B$. $\mathcal{C}^B_m$ and $\pi_{k,m}$ represent the different affine transformations M applied to $\mathcal{C}^B$ and corresponding ensemble π_k. The algorithm is presented below.

Algorithm $COFEMIreg$
Input: $\mathcal{C}^A$, $\mathcal{C}^B$.
Output: $\mathcal{C}^{B,r}$, $\mathbf{M}^*$.
begin

0. Obtain set of feature images $\mathcal{C}^{B'_1}\mathcal{C}^{B'_2}\cdots\mathcal{C}^{B'_n}$ for $\mathcal{C}^B$;
1. Obtain π_k, the optimal set of features via $CMIfeatures(\mathcal{C}^B)$;
2. Obtain $\mathbf{M}^* = \text{argmax}_m[I_2(\mathcal{C}^A, \mathcal{C}^B_m \pi_{k,m})]$;
3. Apply $\mathbf{M}^*$ to $\mathcal{C}^B$ to obtain $\mathcal{C}^{B,r}$;

end

Application of algorithm $COFEMIreg$ to $\mathcal{C}^H$ and $\mathcal{C}^{ex}$ results in $\mathcal{C}^{ex,r}$. Subsequently, known cancer extent on $\mathcal{C}^H$ is directly transferred to the newly generated $\mathcal{C}^{ex,r}$ to obtain the new cancer mask $T(\mathcal{C}^{ex,r})$ for each $\mathcal{C}^{ex,r}$. In this manner, cancer extent on MRI is determined for each $\mathcal{C}^{ex}$, $\mathcal{C}^H$ pair. Calculating CMI of k features results in a k-dimensional joint histogram quickly necessitates the use of a sparse matrix representation even with 16 graylevel bins. Hence, we utilize a sparse matrix addressing scheme to store the high dimensional joint histograms required to estimate CMI.

3. REGISTRATION EVALUATION

Given two scenes $\mathcal{C}^A$ and $\mathcal{C}^B$ to be registered where the correct transformation is not known but a segmented region $T(\mathcal{C}^A)$ is available on $\mathcal{C}^A$, registration accuracy is gauged by evaluating pixel similarity within $T^r(\mathcal{C}^B)$, the corresponding map of $T(\mathcal{C}^A)$ on $\mathcal{C}^B$ as determined by registration. Our method is based on the assumption that pixels from different regions will exhibit different textural properties, the essence of these differences being captured by a high dimensional set of texture features of each pixel. Identifying errors in the mapped area begins with calculating a high dimensional feature space representation $(f^{B'_1}(c), f^{B'_2}(c), \ldots, f^{B'_n}(c))$ of the pixels of $T^r(\mathcal{C}^B)$, from which a low-dimensional clustering of pixels in $T^r(\mathcal{C}^B)$ is obtained by non-linear dimensionality reduction using the LLE algorithm.[13] LLE operates by assuming that points in a local neighborhood of feature space are linearly organized in order to produce weighting vectors that describe each pixel in terms of it's neighbors. Using the local linear weighting vectors, LLE produces a neighborhood preserving mapping to create a low dimensional representation of the global structure of the entire feature space while preserving local structure. The result is a new orthogonal set of embedding coordinates, $\psi(c)$ for each $c \in T^r(\mathcal{C}^B)$, in which discriminative analysis can be performed more easily. Thus, application of LLE to the feature space of $T^r(\mathcal{C}^B)$ produces a low dimensional representation in which the relationships between pixels in the original feature space are retained by the patterns of clustering in the embedding space. Adjacent pixels in low-dimensional feature spaces form clusters representing objects of the same class. Hence, mislabeled pixels in $T^r(\mathcal{C}^B)$ occur far from the dominant cluster in the embedding space produced by LLE.

A comparison of two labels $T^1(\mathcal{C}^B)$ and $T^2(\mathcal{C}^B)$ on the same scene $\mathcal{C}^B$ but obtained by different means is accomplished by simultaneously calculating embedding coordinates from pixels in both labels. It should be noted that $T^1(\mathcal{C}^B)$ and $T^2(\mathcal{C}^B)$ may in fact be the results of any registration or segmentation techniques for which a comparison is desired. The following general algorithm is presented for paired comparison of any two region mappings by identifying false positive and false negative errors for each region $T^1(\mathcal{C}^B)$ and $T^2(\mathcal{C}^B)$. Since clustering is performed on pixels of the combined region mask, errors are determined without regard to the pixels' originating labels.

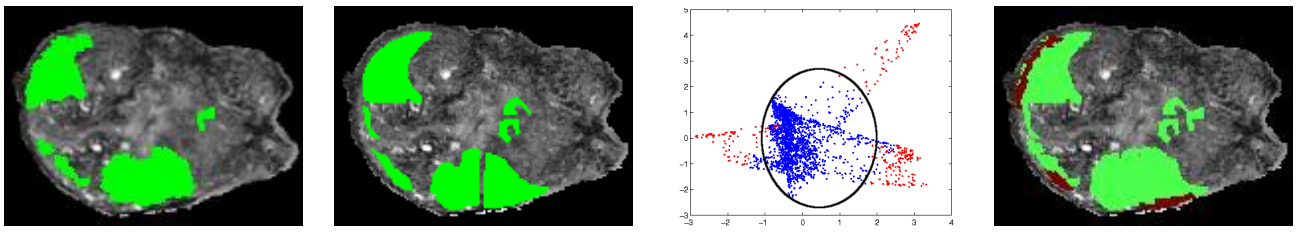
Algorithm $EvalLLE$
Input: $T^1(\mathcal{C}^B)$, $T^2(\mathcal{C}^B)$.
Output: $U^{FP}_{T_1}, U^{FP}_{T_2}, U^{FN}_{T_1}, U^{FN}_{T_2}$.
begin

0. Initialize $V, \hat{V}, U^{FP}_{T_1}, U^{FP}_{T_2}, U^{FN}_{T_1}, U^{FN}_{T_2}$ to empty ques;
1. $V = T^1(\mathcal{C}^B) \cup T^2(\mathcal{C}^B)$;
2. Obtain n feature values $f^{B'_1}(c), f^{B'_2}(c), \ldots, f^{B'_n}(c)$ for each $c \in V$;
3. Obtain embedding coordinates $\psi(c)$ via LLE for all $c \in V$;
4. Obtain cluster centroid $\hat{\psi} = \frac{1}{|V|}\sum_{c \in V} \psi(c)$;

5. *for* each $c \in V$ do
6. If $\left\| \hat{\psi} - \psi(c) \right\| > \Delta$ then
7. If $c \in T^1(\mathcal{C}^B)$ then insert c into $U_{T_1}^{FP}$;
8. If $c \in T^2(\mathcal{C}^B)$ then insert c into $U_{T_2}^{FP}$;
9. *endif*
10. *endfor*
11. $\hat{V} = V - (U_{T_1}^{FP} \cup U_{T_2}^{FP})$;
12. *for each* $c \in \hat{V}$ do
13. If $c \in T^1(\mathcal{C}^B)$ then insert c into $U_{T_1}^{FN}$;
14. If $c \in T^2(\mathcal{C}^B)$ then insert c into $U_{T_2}^{FN}$;
15. *endfor*

end

The ground truth for cancer on $\mathcal{C}^{ex,r}$ determined by COFEMI registration ($T(\mathcal{C}^{ex,r})$) is then compared with manually determined ground truth ($E(\mathcal{C}^{ex,r})$) using the *EvalLLE* algorithm to identify errors in each cancer mask. To compare the masks, the same transformation $\mathbf{M}^*$ determined for each registered slice is used to map each $E(\mathcal{C}^{ex})$ to $\mathcal{C}^{ex,r}$ to obtain $E(\mathcal{C}^{ex,r})$, the manually determined cancer mask on the same MR scene as $T(\mathcal{C}^{ex,r})$. The procedure for detection of mislabeled pixels in the cancer labels is based on the observation that clusters of closely spaced pixels in $\psi(c)$ belong to the same tissue class. Hence, outlier pixels from the dominant cluster, which represents cancer in this study, are determined to be of a dissimilar class (i.e. non-cancer) and thus mislabeled. In this manner, we use the *EvalLLE* algorithm to identify false positives and a subset of false negatives associated with $E(\mathcal{C}^{ex,r})$ or $T(\mathcal{C}^{ex,r})$ (Figure 4(a) and (b), respectively). False positives errors in each cancer label are determined in steps 5-10 via the embedding space (Figure 4(c)) of the combined cancer label. All false positive errors determined by this step are shown on the combined label in Figure 4(d). After FP errors are removed from the combined label in step 11, steps 12-15 of the algorithm determine false negative errors for the two cancer labels. The FP and FN estimates provide two measures of accuracy of $E(\mathcal{C}^{ex,r})$ and $T(\mathcal{C}^{ex,r})$, thus gauging the accuracy of the procedures that established their positions on $\mathcal{C}^{ex,r}$. A cancer label with few mislabeled pixels is indicative of accurate labeling and in the case of $T(\mathcal{C}^{ex,r})$ an accurate registration. The evaluation of COFEMI registration in this application is based on the premise that (i) the source cancer mask $T(\mathcal{C}^H)$ on $\mathcal{C}^H$ is without errors and has relatively homogeneous appearance and (ii) the true unknown mapping of $T(\mathcal{C}^H)$ on $\mathcal{C}^{ex}$ is relatively homogeneous in the high dimensional feature space and hence the dominant cluster corresponds to cancer.



(a) (b) (c) (d)

Figure 4. Identification of mislabeled pixels by cluster analysis of combined cancer mask $\mathcal{C}^{ex,r}$. Texture features from (a) $E(\mathcal{C}^{ex,r})$ and (b) $T(\mathcal{C}^{ex,r})$ are combined to generate (c) LLE of feature space of combined tumor mask with outlier pixels (shown in red) determined by distance thresholding in the first 3 embedding eigenvectors. (b) Combined mask on $\mathcal{C}^{ex,r}$ shown with outliers (false positives) in dark red and correctly labeled cancer pixels in green.

4. RESULTS

Quantitative and qualitative evaluation was performed on a total of 26 image sections from 2 separate patient studies. Intensity-based MI was performed using 128 graylevel bins. It was found that 16 graylevel bins was most appropriate for COFEMI registration using multiple features.

4.1. Qualitative Results

Figure 5 shows the boundary of the prostate on histology overlaid onto the MRI for 3 consecutive MR image slices. COFEMI registration with a feature ensemble comprising range, correlation, and standard deviation features (automatically chosen by $CMIfeatures$) performed with accuracy consistently superior to both intensity-based MI and CMI with randomly chosen features. Figure 5 illustrates this performance improvement for registration of 3 corresponding histology and MRI slices that were difficult to register using conventional MI. The feature ensemble used for COFEMI is clearly robust (Figure 5(g)-(i)) to the inhomogeneity of tissue intensities visible in the MR intensity images, which has caused the misregistration associated with conventional MI (Figure 5(d)-(f)). The green boundaries in Figure 5 represent the boundary of $\mathcal{C}^{H}$ overlaid onto $\mathcal{C}^{ex,r}$, while red accentuates the boundary of the prostate in $\mathcal{C}^{ex,r}$. These boundaries would coincide with ideal registration.

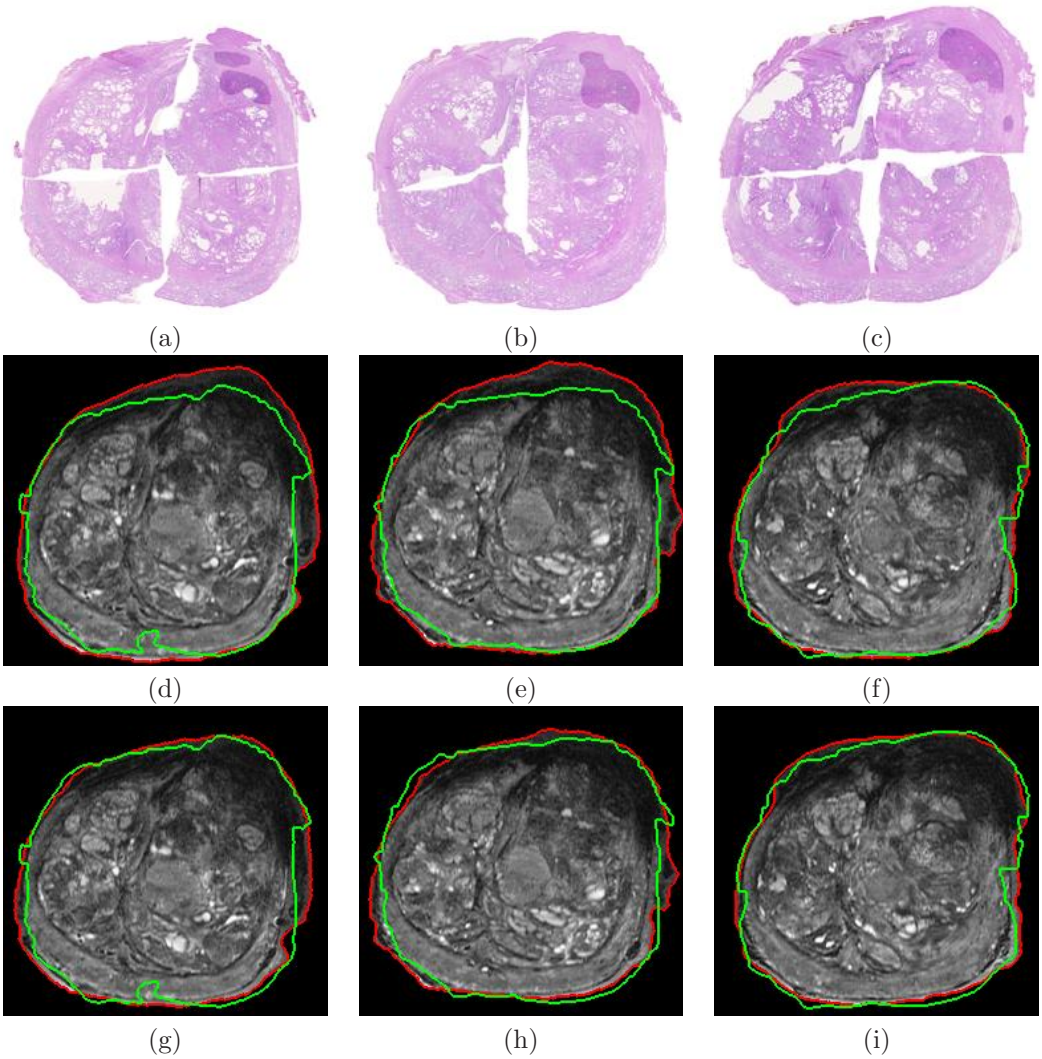


Figure 5. Performance comparison of (d)-(f) intensity-based MI and (g)-(i) feature driven COFEMI for registration of three consecutive slices of $\mathcal{C}^{ex}$ with (a)-(c) corresponding histological slices from $\mathcal{C}^{H}$ demonstrates improved accuracy with COFEMI. Green contours represent the boundary of corresponding histology section overlaid onto the MR slice. Red contours accentuate the boundary of the prostate in the MRI slice. Serious misregistration is visible on the upper gland borders of (d)-(f) conventional MI registration results due to inhomogeneity of the tissues appearance on MR intensity images that is not present on the corresponding histology. These errors are greatly diminished on (g)-(i) COFEMI registration results. Note the considerable differences in the overall appearance of prostate histology and MRI in terms of structural details and distortions to the histology from tissue loss and slicing.

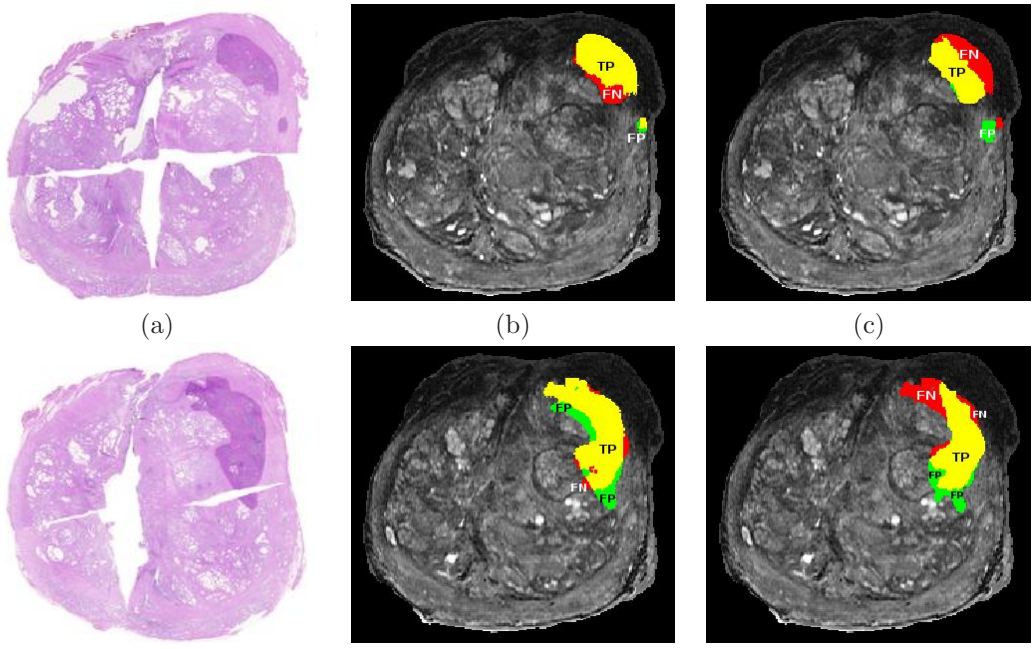


Figure 6. Identification of mislabeled pixels by outlier detection on LLE clustering using *EvalLLE*. (a),(d) cancer ground truth from histology mapped to (b),(e) $\mathcal{C}^{ex,r}$ to obtain $T(\mathcal{C}^{ex,r})$. (c),(f) $E(\mathcal{C}^{ex,r})$ shows more FN compared to $T(\mathcal{C}^{ex,r})$ as a result of errors in the visual mapping process. TP are shown in yellow, FP in green, and FN in red. Note that while registration errors exist in (e) due to severe histology deformations in (d), FN extent is still much less than in (f) and FP extent is comparable.

The false positive (FP) and false negative (FN) errors for $T(\mathcal{C}^{ex,r})$ and $E(\mathcal{C}^{ex,r})$ are shown in Figure 6. We found that $T(\mathcal{C}^{ex,r})$ represented a better spatial map of cancer extent on $\mathcal{C}^{ex,r}$ compared to $E(\mathcal{C}^{ex,r})$ in terms of FP and FN. Considerably larger FN extent is visible for $E(\mathcal{C}^{ex,r})$ in both Fig. 6(c) and (f), while FP extent is comparable.

4.2. Quantitative Results

The results are compared in terms of false positive fraction (FPF), defined as FPF $= |\mathrm{FP}|/N_L$, where N_L represents the total pixels in the individual label, and false negative fraction (FNF), defined here as FNF $= |\mathrm{FN}|/N_V$, where N_V represents the number of pixels in the combined label. In Table 2, we present for each set of corresponding $\mathcal{C}^{ex}$ and $\mathcal{C}^{H}$ scenes the average FP and FN fractions determined by the non-linear cluster analysis evaluation technique. We also list the p-value of the paired t-tests for the null hypothesis that the observed measures of FPF and FNF are the same for $T(\mathcal{C}^{ex,r})$ and $E(\mathcal{C}^{ex,r})$. For $N = 26$ slices, $T(\mathcal{C}^{ex,r})$ has a lower FPF and FNF than $E(\mathcal{C}^{ex,r})$, and significantly lower ($p < 0.05$, paired) FNF. Thus, the COFEMI determined cancer mask performed as well or better than expert determined labels in terms of FPF and FNF.

Table 2. Mean false positive fraction (FPF) and false negatives fraction (FNF), as defined in text, over all slices in both prostate histology-MRI data sets. The cancer mask mapped from histology via COFEMI registration outperforms manual cancer mapping in terms of FPF and FNF. The FNF values associated with the COFEMI mask $T(\mathcal{C}^{ex,r})$ are significantly lower than for $E(\mathcal{C}^{ex,r})$.

Error type	COFEMI	Manual	$p(N = 26)$
FPF	0.1033	0.1209	0.0657
FNF	0.1967	0.2729	0.0016

5. CONCLUDING REMARKS

In this paper, we present a new registration method called COFEMI and demonstrated it's application in aligning *ex vivo* MRI and histology images of the prostate. The COFEMI technique makes use of a new formulation of MI that incorporates additional information in the form of feature transformations of the original intensity image. By making use of multiple textural feature images that are uniquely informative and less susceptible to image intensity artifacts, we achieve robust multimodal registration of images from different modalities. We also demonstrate a novel feature selection scheme as part of COFEMI for identifying an optimal ensemble of features to drive the registration. The performance improvement of COFEMI over conventional MI is demonstrated for the registration of 26 pairs of *ex vivo* MRI and histology images of the prostate in the task of establishing ground truth for cancer on MRI. In addition, we qualitatively and quantitatively demonstrate improvement in determining spatial extent of cancer on MRI from histology via COFEMI compared to manual mapping through visual alignment of the image pair. As such, we expect these direct mappings of histological ground truth obtained via COFEMI to provide a more accurate surrogate for ground truth on the MRI, and thus improve CAD performance by improving prior feature distributions used in training of the supervised classifier.

To address registration evaluation in the absence of a gold standard for registration accuracy, we develop a novel quantitative evaluation procedure that identifies registration errors as those pixels that are far separated from the dominant class cluster in a lower dimensional embedding space. The evaluation method is general and applicable to both registration and segmentation evaluation. While the registration routines used here utilize rigid rather than elastic deformations, the registration methodologies provide a robust rigid registration that is required to perform successful non-rigid registration. In future work, we intend to evaluate our methods on a larger set of studies.

REFERENCES

1. A. Madabhushi, M. D. Feldman, *et al.*, "Automated detection of prostatic adenocarcinoma from high-resolution ex vivo mri," *IEEE Trans. Med. Imag.* **24**, pp. 1611–1625, December 2005.
2. A. Madabhushi, J. Shi, *et al.*, "Graph embedding to improve supervised classification and novel class detection: application to prostate cancer," in *Medical Image Computing and Computer-Assisted Intervention 2005*, J. S. Duncan and G. Gerig, eds., pp. 729–737, 2005.
3. A. Bharatha, M. Hirose, *et al.*, "Evaluation of three-dimensional finite element-based deformable registration of pre- and intraoperative prostate imaging," *Medical Physics* **28**(12), pp. 2551–2560, 2001.
4. M. Foskey *et al.*, "Large deformation three-dimensional image registration in image-guided radiation therapy," *Phys. Med. Biol.* **50**, pp. 5869–5892, 2005.
5. M. Jacobs, J. Windham, *et al.*, "Registration and warping of magnetic resonance images to histological sections," *Med. Phys.* **26**(8), pp. 1568–1578, 1999.
6. J. Pluim, J. Maintz, *et al.*, "Mutual-information-based registration of medical images: A survey," *IEEE Trans. Med. Imag.* **22**, pp. 986–1004, August 2003.
7. A. Madabhushi and J. Udupa, "Interplay between intensity standardization and inhomogeneity correction in mr image processing," *IEEE Trans. Med. Imag.* **24**(5), pp. 561–576, 2005.
8. J. Pluim, J. Maintz, *et al.*, "Image registration by maximization of combined mutual information and gradient information," *IEEE Trans. Med. Imag.* **19**, pp. 809–814, August 2000.
9. D. Rueckert, M. Clarckson, *et al.*, "Non-rigid registration using higher-order mutual information," **3979**, pp. 438–447, SPIE Med. Imag., 2000.
10. J. Boes and C. Meyer, "Multi-variate mutual information for registration," in *Medical Imaging and Computer-Assisted Intervention*, **1679**, pp. 606–612, MICCAI, 1999.
11. C. Studholme, D. Hill, *et al.*, "Incorporating connected region labelling into automatic image registration using mutual information," in *Math. Methods in Biomed. Image Analysis*, **3979**, pp. 23–31, IEEE, 1996.
12. S. Warfield, K. Zou, *et al.*, "Simultaneous truth and performance level estimation (staple): an algorithm for the validation of image segmentation," *IEEE Trans. Med. Imag.* **23**(7), pp. 903–921, 2004.
13. S. Roweis and L. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *Science* **290**, pp. 2323–2326, 2000.
14. H. Matsuda, "Physical nature of higher-order mutual information: Intrinsic correlations and frustration," *Phys. Rev. E* **62**, pp. 3096–3102, Sep 2000.
15. S. y. Lu, J. Hernandez, *et al.*, "Texture segmentation by clustering of gabor feature vectors," in *International Joint Conference on Neural Networks, 1991*, pp. 683–688, 1991.