# 1 Relevance and Impact of Proposed Work

Ultrasonography is a relatively low cost diagnostic imaging modality that is widely available in many hospitals and community clinics. It would thus be practical and cost effective to perform successive ultrasound exams on patients to assess effectiveness of treatment and recovery from soft tissue lesions and inflammations. However, to make reliable comparisons between two time-seperated ultrasound scans accurate image registration is required to quantify subtle but important changes in the lesion. Investigation of ultrasound image registration is relatively new and not widespread. Only recently, with single voxel mutual information techniques, has such registration appeared to be very promising. The significantly improved registration methods outlined in this proposal may be a critical key to making routine serial ultrasound studies of soft tissues practical and informative.

## 1.1 Potential Impact on Ultrasound Imaging

New cases of breast cancer are diagnosed in over 190,000 americans per year. It is the second leading cause of cancer death in women and one out of nine women will develop breast cancer during their lifetime [1]. Survival rates can be greatly enhanced by early detection of asymptomatic breast lesions. However in mammographic screening of women with dense breasts, up to 45% of cancers are currently missed by community practitioners [33]. Our research to improve registration accuracy could aid diagnosis of benign and malignant masses by facilitating visual or automatic comparison with the structures in the same region at the time of a previous examination. Image volume registration may aid tracking of response to cancer therapy as well. Success in these goals should lead to earlier detection and diagnosis of many breast cancers, as well as reduced pain and uncertainty for those with suspicious lesions and more appropriate dosages and selections of therapies.

Sports and recreational injuries to muscles and other soft tissues, such as tendon tears, ruptures, and dislocations, affect many thousands of americans every year. The incidence of debilitating trauma to muscolo-skeletal tissues is expected to increase as our nation enters a new era of uncertainty possibly involving terrorist acts and other acts of violence. Recovery from such trauma is a slow process during which periodic musculoskeletal ultrasonography can provide crucial information affecting judgements on treatment planning, secondary diagnosis, and the need for surgical intervention [2]. Accurate registration of soft tissue ultrasound image volumes will prove essential for efficient detection and quantification of changes that are indicative of processes such as: decreased bursal inflammation, decreased fluid collection around joints, or reductions of cyst or lesion sizes.

## 1.2 Broader Impact of Work

The registration methods that we propose to investigate are widely applicable to a spectrum of areas extending well beyond our focus on 3D ultrasound imaging. Other medical imaging modalities such as magnetic resonance imaging (MRI), positron emission tomography (PET), single photon emission tomography (SPECT), X-ray computed tomography (CT), rely on accurate image registration for both clinical applications, such as CT-guided radiation therapy and treatment planning, and research applications, such as PET brain activiation studies. Image

registration is also crucial in the rapidly progressing area of biological imaging including registering within and between modalities such as computational sectioning microscopy (COSM), confocal microscopy, and atomic force microscopy (AFM). Our research is also directly applicable to geo-registration applications involving fusion of information from electro-optical (EO), acoustic, and radar sensors for remote detection of agricultural yields, minefields, and underground facilities. Finally, by adding a fourth dimension (time) to our image source model our methods can be implemented in video registration applications for homeland security such as: face recognition, gait recognition, and camera surveillance.

## 2 Problem Statement and Background

New techniques developed by our ultrasound (UL) group in Basic Radiological Sciences (see http://www.ultrasound.med.umich.edu/) are ready for evaluation and refinement in clinical studies and include: registration of serial studies in single and multiple scan modes, creation of expanded 3D fields of view, a moderately fast, manual method of segmenting ellipsoidal volumes, quantification of vascularity in regions of interest, and improved UL imaging by compounding (combining multiple views for improved image contrast to noise ratio) using image volume based registration (IVBaR)[7-10]. Work is also progressing on computer aided diagnosis (CAD) of UL masses [34], with potential links between UL and x-ray CAD. What is needed is a much more accurate and robust registration method to provide the needed confidence in the registration for studies with a wider range of image quality. Segmentation of artifacts, such as acoustic shadows and reverberations, is particularly important, for the artifacts can be the dominant "information" in some cases, and yet are affected strongly by subtle changes in the UL view. Figure 1 shows examples of three typical but very different breast scans which illustrate the challenge of image registration for this application.

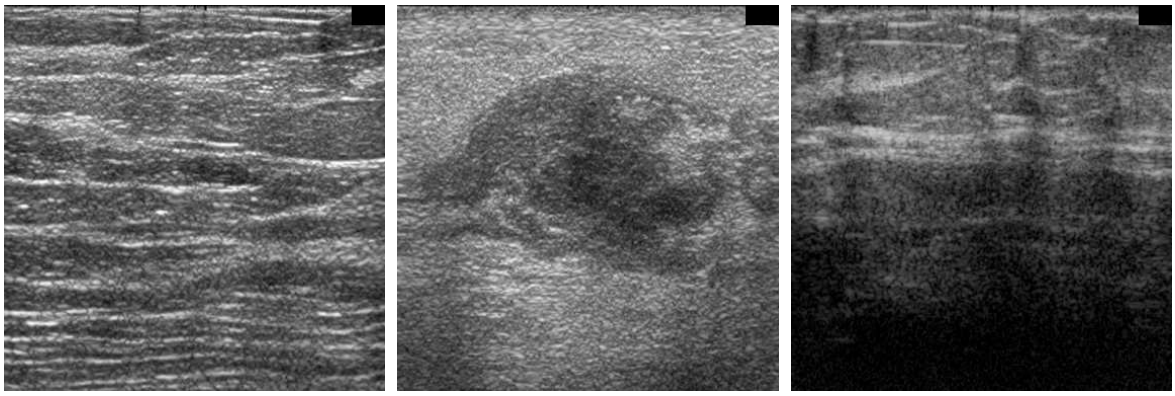

Figure 1: *2D slices from three 3D ultrasound breast scans. From top to bottom are: case 151, case 142 and case 162. The image slice chosen from Case 151 (image slice 40) had significant connective tissue structure. Case 142 (slice 35) had a distinct malignant tumor, while Case 162 (slice 60) was degraded due shadowing.*

## 2.1 Past Approaches to Image Registration

Registration of an image to a database of rotated and translated exemplars constitutes an important image retrieval and indexing application which arises in biomedical imaging, digital libraries, georegistration, and other areas. Let $X_0$ be a the reference image and consider a secondary image $X_1$ and sequence of deformed secondaries $X_i = T_i(X_1)$, $i = 1, 2, \ldots$. Let $Z_0$, $Z_i$ be feature vectors extracted from $X_0$, $X_i$ and define the joint histogram $f_{0,i}(z_0, z_i)$ and the marginal histograms $f_0(z_0)$, $f_i(z_i)$. The similarity between features $Z_0$ and $Z_i$ can be gauged by the difference between $f(z_0, z_i)$ and the product $f(z_0)f(z_i)$ which measures statistical dependence. Alternative measures of dependency are cross correlation, higher order cross-moments like joint kurtosis, and mutual-information (MI) methods based on the entire feature histogram.

The simple single-pixel mutual-information method introduced by Viola and Wells [36] is an example of the general MI approach. Their technique has been refined and adapted by several medical image registration groups including a group at The University of Michigan (UM) Dept. of Radiology [27, 28, 23]. In all of these groups the MI-based technique have supplanted previous correlation methods due to its superior performance. The UM method, called MIAMI Fuse© functions as follows. A randomly or intelligently chosen set of voxels, called control points, are selected in the secondary image and associated with knot positions on a thin-plate spline. The bending resistance of the thin plate spline causes a smooth deformation of the entire voxel grid on the secondary image as the control points are moved around. A sequence of deformations is defined by successive movements of the control point positions giving rise to the sequence of deformed secondary images $X_i = T_i(X_1)$, $i = 1, 2, \ldots$ obtained by interpolation of $X_1$ onto the deformed grid. Let $X_0$ and $X_i$ be $N \times N$ images ($N^2$ pixels). In single-pixel registration techniques such as MIAMI-fuse©, for each iteration $i$ a set of feature pairs $\{(Z_0(k), Z_i(k))\}_{k=1}^{N^2}$ are defined as the successive pairs of gray levels (typically ranging from 0-255) of the lexicographically indexed primary and deformed secondary images, respectively. A two dimensional gray level empirical histogram $\hat{f}_{0,i}(z_0, z_i)$, $z_0, z_i \in \{0, \ldots, 255\}$, is then generated from these feature pairs $\{(Z_0(k), Z_i(k))\}_{k=1}^{N^2}$ yielding an extended point cloud in the plane if the images are not registered and a straight line if they are perfectly registered (see Figure 2). A good feature matching criterion would be a measure of the degree of independence between $Z_0$ and $Z_i$, or equivalently the degree to which the $\hat{f}_{0,i}(z_0, z_i)$ can be factored into the product $\hat{f}_0(z_0)\hat{f}_i(z_i)$, where $\hat{f}_0(z_0)$ and $\hat{f}_i(z_i)$ are the marginals. The Shannon MI is one such feature matching criterion

$$\mathrm{MI}(\hat{f}_{0,i}) = \sum_{z_0} \sum_{z_i} \hat{f}_{0,i}(z_0, z_i) \ln \hat{f}_{0,i}(z_0, z_i)/(\hat{f}_0(z_0)\hat{f}_i(z_i)). \tag{1}$$

Registration occurs when this registration function is maximized, and the search over deformations $T_i$ for this maximum is performed by an iterative optimization algorithm.

When the image volumes are identical except for rotation, translation and small deformations, the MI will exhibit a very sharp peak over $i = 1, 2, \ldots$ for that deformation $T_i$ which achieves perfect registration of the images, as in the perfectly registered example shown in the left panel of Fig 2. However the assumption that there exists a perfect registration is unrealistic in practical cases where cancer detection, diagnosis and treatment response are the goal. In the latter case, the reference and secondary images are genuinely somewhat different from each other due to biological changes and differences in the imaging, such as speckle, positioning of

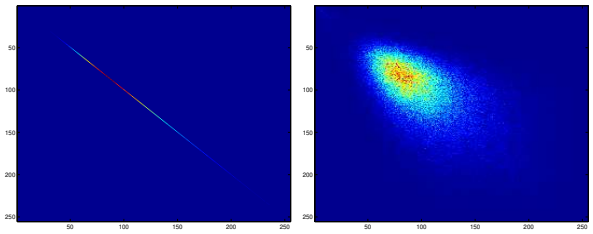

Figure 2: *Single-pixel gray-level histogram $\hat{f}_{0,i}$ for registering a single slice to rotated rotated version of itself (relative rotation angle = 5%). The slice is taken from the ultrasound breast image shown in first panel of Fig 1. Horizontal and vertical axes range over the gray levels (0-255). Observe that a straight line is obtained (panel at right) when perfect registration occurs.*

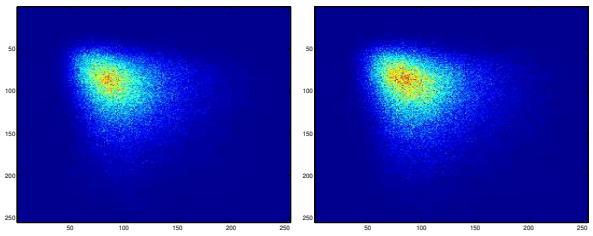

Figure 3: *Same as in Fig. 2 except that primary and secondary images are taken from adjacent slices separated by 2mm. Due to the presence of speckle we do not obtain a straight line even when images are registered (left panel).*

the tissues during compression, and angle dependence of UL scattering from tissue boundaries. In addition, the tissues are distorted out of a given image plane as well as within it leading to poor performance of slice-by-slice (2D) registration. As a result, even registered images will not be perfectly correlated (see Fig. 3) and the MI peak will not be as sharp as in the idealized situation. The combination of fully 3D feature registration techniques and more sensitive feature matching measures that we propose in this grant proposal can overcome these problems.

# 3 Proposed Methods

We take a novel approach to image registration that gets around the disadvantages of previous single pixel- or voxel-based registration techniques. The key to our approach is the inclusion of higher order image features combined with generalized information divergence criteria for feature matching. We will use inductive learning to select the most relevant and robust features (curves, edges, textures and simple spatial relations) from a large set of previously collected scans, e.g. of breast or Achilles tendon. These features are organized in an efficient hierarchical database, called a randomized feature tree. To register a secondary image to a reference image, coincidence histograms of the features captured by the feature tree for the reference and secondary images are constructed. The coincidence histograms are used to register the images by optimizing a new class of registration functions over the space of deformations of the secondary image. This class of registration functions includes the mutual $\alpha$-information and the (negative) $\alpha$-Jensen difference which will be estimated using a combination of feature density estimation and a novel minimal graph matching technique developed by Hero and co-workers [26, 15]. Our framework

specializes to the standard (MIAMI-Fuse©) registration method when the features are single-pixel gray levels and $\alpha = 1$ for which the mutual $\alpha$-information is equivalent to Shannon MI. The generality of our approach has the following advantages: 1) use of optimized feature matching measures can lead to much more stable registration function having much higher discrimination capability; and 2) use of higher order features can capture non-local spatial information which is ignored by previous algorithms.

### 3.1 Optimal Feature Matching Criteria

We adopt a systematic decision-theoretic framework for selecting the registration function. In this framework we assume that the feature point cloud, e.g. either panel of Fig. 3, is a random realization of some unknown underlying joint density $f_{0,i}$. A good registration function will satisfy the following criteria: 1) high average accuracy over an ensemble of feature realizations from joint density $f_{0,i}$, corresponding to accuracy averaged over a large representative image database; 2) high accuracy for a single realization of feature points from $f_{0,i}$, corresponding to accuracy for a single pair of images; 3) robustness to outliers, i.e. spurious features that would normally confound registration; 4) fast algorithm implementation.

Optimal registration functions which satisfy the above criteria will be developed and studied in a systematic three step process. First we study theoretically optimal *omniscient* registration functions which have access to the ensemble statistics of the image database, i.e., the exact feature density function $f_{0,i}$. These ensemble statistics will depend on the medical imaging application. Then, we explore methods of optimal empirical estimation of an omniscient registration function based on a single pair of images. Optimal estimation strategies will also depend on the medical imaging application. Finally we will explore fast and robust implementations of these empirical registration function estimates using graph matching techniques.

#### 3.1.1 Omniscient Registration Functions

The above considerations have led us to propose the class of $\alpha$-information criteria ($\alpha \in [0,1]$) and their close relatives, the $\alpha$-Jensen differences.

As above let $f_{0,i}(z_0, z_i)$ be the true (ensemble) joint density between feature vectors of the primary and deformed secondary images. The (ensemble) $\alpha$-information is defined as

$$D_\alpha(f_{0,i} \parallel f_0 f_i) \quad = \quad \frac{1}{\alpha - 1} \log \int f_{0,i}^{\alpha}(z_0, z_i) f_0^{1-\alpha}(z_0) f_i^{1-\alpha}(z_i) dz_i dz_0. \tag{2}$$

When specialized to various values of $\alpha$ the $\alpha$-information reduces to well known information measures such as the Battacharya-information ($\alpha = 1/2$), and Shannon MI used in MIAMI-Fuse© ($\alpha = 1$). The $\alpha$-information is a special case of "$f$-information" which has been investigated in a recent paper [32] for medical image registration and also generalizes the mutual-information method of Viola and Wells [36]. However, as shown by Hero and his co-workers [15], the $\alpha$-divergence is the most appropriate member of the $f$-information class as it is directly related to the minimum attainable decision error incurred when trying to discriminate between an independent (factorizable) and a dependent (nonfactorizable) feature density $f_{0,i}(z_0, z_i)$ based on a finite number of feature samples. In particular, large deviations theory asserts that there exists a theoretically optimal value of $\alpha$ that optimizes the discrimination performance. This

value of $\alpha$ is generally not equal to the value $\alpha = 1$ corresponding to the standard Shannon MI and depends on the class of feature densities $f_{0,i}$ under consideration.

A closely related quantity to be investigated is the (ensemble) $\alpha$-Jensen difference. The $\alpha$-Jensen difference is of interest due to the existence of simple low complexity estimators using minimal spanning graphs over the feature pairs $\{(Z_0(k), Z_i(k))\}_{k=1}^{N^2}$. The $\alpha$-Jensen difference is

$$\triangle H_\alpha(\beta, f_0, f_1) \triangleq H_\alpha(\beta f_0 + (1-\beta) f_1) - [\beta H_\alpha(f_0) + (1-\beta) H_\alpha(f_1)], \quad \alpha \in (0,1), \tag{3}$$

where

$$H_\alpha(f) = \frac{1}{1-\alpha} \int f^\alpha(z) dz$$

is the $\alpha$-entropy of the probability density $f$ and $\beta \in [0,1]$ is a primary-secondary mixing parameter. We have chosen $\beta$ propportional to the ratio of the total number of feature incidences in the primary image relative to the secondary image. The $\alpha$-Jensen difference is minimized when $f_0 = f_1$ taking on a value of zero there. Minimzing the $\alpha$-Jensen difference was originally proposed by us [26, 25] for geo-registration applications including SAR and electro-optical radar imaging A generalization was recently proposed by He, Ben-Hamza and Krim [14] for registering an arbitrary number of image modalities. The $\alpha$-Jensen difference was proposed several years before by one of the PI's long time collaborators O. Michel for classifying image differences in the time-frequency plane [29].

### 3.1.2 Empirical Registration Functions

In practice the joint feature density $f_{0,i}$ is unknown so that the above ensemble registration functions are unimplementable. The statistical problem of registration can be stated as follows: based on a single realization $\{(Z_0(k), Z_i(k))\}_{k=1}^{N^2}$ of joint features, e.g. the point cloud illustrated in either of the panels of Fig. 3 corresponding to a single image pair, empirically estimate the ensemble $\alpha$-information or the ensemble $\alpha$-Jensen difference. This estimation problem is closely related to entropy estimation which has a long history in the statistics and information theory communities [13, 22, 3, 35, 5, 7, 30].

The approach to empirical estimation that will be taken here is motivated by our recent work in entropy estimation [20, 16, 15]. We will investigate two different techniques: density plug-in techniques and minimal graph techniques. Each of these methods has its own complementary strengths and weaknesses which justify considering both of them for registration applications.

Density plug-in methods are in widespread use for estimation of Shannon mutual-information and other entropy functions [36, 28, 5]. These methods involve non-parametric density estimation, such as the histogram binning method or kernel density estimators, to yield empirical estimates of $f_{0,i}(z_0, z_i)$ and its marginals which are subsequently substituted into an ensemble formula, e.g. (2) or (3). For example, when a non-parametric estimate of $\hat{f}_{0,i}$ is available the plug-in estimate of $\alpha$-information is $D_\alpha(\hat{f}_{0,i} \parallel \hat{f}_0 \hat{f}_i)$. The main difficulties with density plug-in methods are due to the infinite dimension of the spaces in which the unconstrained densities lie. Specifically: density estimator performance is poor without stringent smoothness conditions; no unbiased density estimators generally exist; density estimators have high variance and are sensitive to spurious outliers; for high dimensional feature spaces the integration in (2) might be difficult.

A more direct estimation method which applies to continuously valued features is based on constructing minimal graphs that span the feature point realizations. For example, to estimate the $\alpha$-entropy of the primary image's feature density $f_0$, the minimal spanning tree (MST) method constructs a graph that spans all of the feature realization points $\{Z_0(k)\}_{k=1}^{N^2}$ in such a way so as to minimize the total length of the graph's edges, examples are shown in Fig. 4. Specifically, if the MST puts an edge between $Z_0(i)$ and $Z_0(j)$ the edge has length $\|e_{ij}\| = \|Z_0(i) - Z_0(j)\|^\gamma$ where $\|\cdot\|$ is Euclidean distance and $\gamma = (1-\alpha)d$, where $d$ is the dimension of the feature vector $Z_0(k)$. This method can be applied to estimation of the $\alpha$-Jensen difference by constructing the MST over the merged sample $\{Z_0(k)\}_{k=1}^{N^2} \cup \{Z_i(k)\}_{k=1}^{N^2}$ and selecting $\beta = 1/2$ in (3). Minimal graph techniques can also be extended to estimation of $\alpha$-information but for lack of space we will not discuss this here (see [15]).

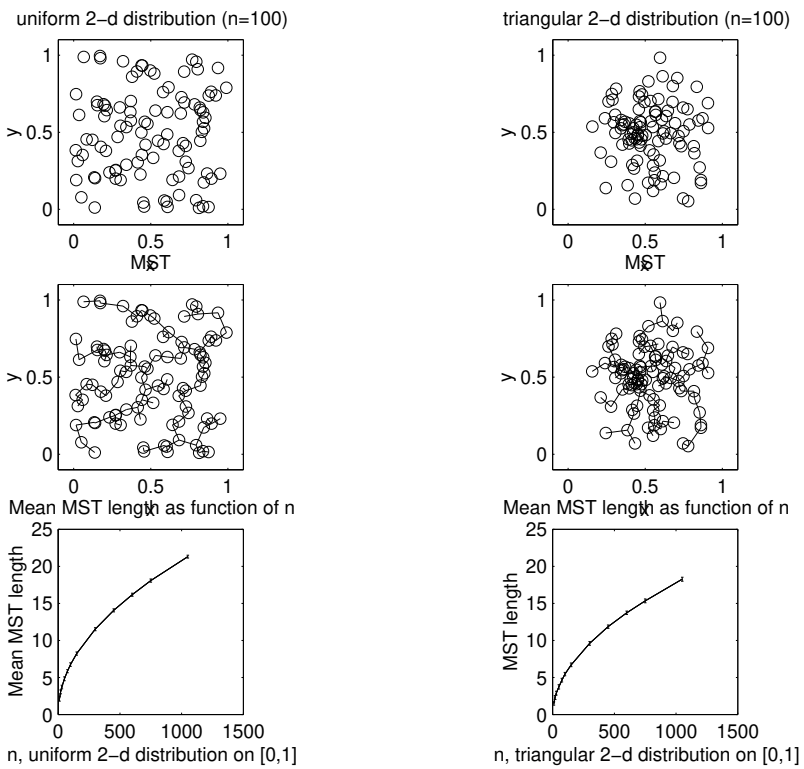


Figure 4: *Illustration of MST estimator of entropy for two dimensional features ($d = 2$) and edge weight exponent $\gamma = 1$. Left column: MST results for uniform feature density $f_0$ on the unit square. Right column: MST results for pyramidal feature density $f_0$ on the unit square. First row: $n = 100$ realizations of each feature density; Second row: MST for each case. Third row: total length of MST increases with rate $n^{\frac{d-\gamma}{d}} = \sqrt{n}$ and rate constants equal to the $\alpha$-entropy of $f_0$, $\alpha = \frac{d-\gamma}{d} = 1/2$.*

General minimal graph techniques for entropy estimation were introduced by Hero and Michel [18, 19, 20]. We showed the remarkable fact that the length of a minimal spanning graph, such as a minimal spanning tree (MST) or $k$-nearest neighbors graph ($k$-NNG), forms a strongly consistent estimator of the $\alpha$-entropy. Furthermore, recently obtained theoretical results of Hero and Ma [16] assert that the minimal graph estimator of entropy can be significantly more accurate (lower mean-squared error) than plug-in estimators. Here $\alpha$ is related to the dimension

of the feature space and can be varied over 0 to 1 by using different length weighting exponents $\gamma$.

We have shown [20, 16] that minimal graph estimators have several advantages over standard plug-in estimators of entropy and mutual information: 1) they do not require specification of difficult auxiliary parameters such as bin size and kernel width; and 2) as shown by Hero and Ma [17], their estimation error can converge to zero much faster than that of density estimation methods; 3) fast implementations exist using MST and $k$-NNG constructions; 4) optimal pruning methods can be applied to these graphs to render the $\alpha$-entropy estimates robust to anomalous feature points, e.g. points due to speckle or tumors not present in both images. We believe that such pruning methods will be effective for registering two medical images where one of the images has small lesions. These minimal graph techniques have been applied with success to geo-registration problems by us [26]. A representative example is shown in Figs. 4 which shows a minimal spanning tree and its pruned version using a combinatorial optimization technique known as the $k$-point MST (see publications [26, 15] for explanation).

## 3.2 Higher Order Feature Selection

As an alternative to single-pixel features we will investigate higher order features which capture spatial dependencies such as edges, ridges, organ outlines, and other non-local structure. When added to single-pixels such features can produce much more robust features for registration. Two different types of features will be investigated: randomized feature trees which organize simple local image features into a hierarchy of subimages of increasing complexity, and randomized databases of local features selected by independent component analysis (ICA). These feature structures will allow us to sidestep the issue of rigidly defining non-local image configurations and clusters such as, image gradients, boundary detectors, distinct points, discriminating structures or other statistical parameters. For each of these methods we will investigate methods of bagging, boosting and randomization for feature selection from a representative image database used for training.

### 3.2.1 Randomized Feature Trees

Randomized feature trees were introduced by Amit and Geman [4] for shape recognition from binary transcriptions of handwriting and other computer vision applications. A set of primitive local features, called tags, are selected which provide a coarse description of the topography of the intensity surface in the vicinity of a voxel. Local image configurations, e.g. subimages associated with $4 \times 4$ pixel neighborhoods, are captured by coding each pixel with labels derived from the tags. Non-local spatial features are then captured by cataloging pairs of tags which are in particular relative spatial configurations.

The tags will be selected by the adaptive thresholding technique as described by Geman and Koloydenko [11] which was introduced to study the invariant characteristics of natural images. Let $\Delta$ be a positive granularity parameter. The quantized value assigned to a pixel within a $4 \times 4$ sub-image depends on the gray values of its neighbors. The darkest pixel(s) are assigned 0, the next brightest pixel(s) are assigned 0 if the difference is less than $\Delta$ and label 1 otherwise, the next brightest pixel(s) are assigned label 2 if the difference is less than $\Delta$, and so on. Using this scheme on our ultrasound image database tags associated with the relatively

uniform background areas (dark or bright) with small spatial variances are correctly classified as speckle and can be easily eliminated. Such tags are irrelevant to the image registration and our preliminary studies indicate that their elimination results in a reduction by almost 75% of the total number of possible tag types. The remaining 25% tag types were observed to distinguish straight and curved edges in the images.

Once selected, the tag features are efficiently organized on a tree structured database, called the feature tree, as illustrated in Figures 5 and 6 for $4 \times 4$ tag subimages. For the purposes of image registration, the primary and deformed secondary images $X_0$ and $X_i$ will each be dropped down the feature tree and incidences and coincidences of features at all of the nodes of the two trees are counted. The counter is incremented for every coincidence of a particular feature pair occurring at a common position within each of the two images. This results in a feature coincidence histogram $\hat{f}(z_0, z_i)$. The histogram marginals $\hat{f}(z_0)$ and $\hat{f}(z_i)$ of the coincidence histogram are extracted by summing over one of its arguments. These will be then used in the mutual $\alpha$-information formula (2) to come up with a registration score for the images.

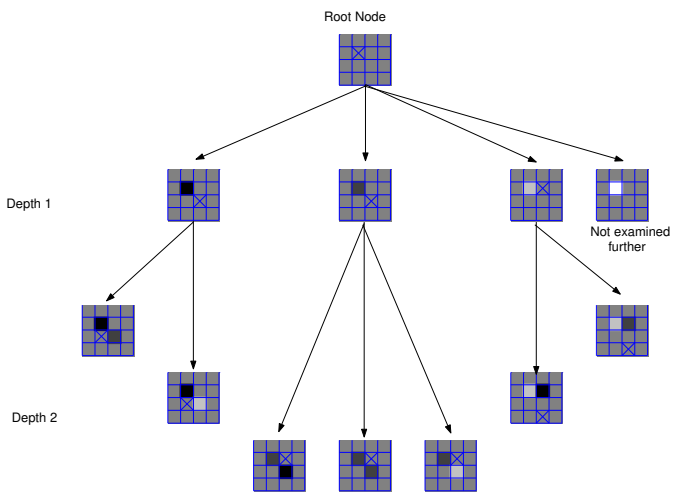


Figure 5: *Part of feature tree data structure.*

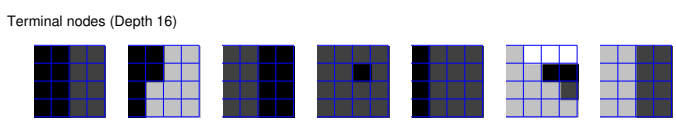


Figure 6: *Leaves of feature tree data structure.*

Feature trees have had the most success in pattern recognition when a large number of features, tags, and spatial relationships are used. The resultant high computation load is reduced by exploiting the recursive partitioning property of the tree hierarchy: for binary trees the search complexity is reduced from $K$ to $\log 2(K)$ and for quad trees further search complexity reductions are obtained, where $K$ is the number of features considered. Additional reductions in average search complexity are obtained by putting the most common and discriminating feature tags near the top of the tree and pruning a large number of the remaining tags. Discovery of the most discriminating spatial relationships and the most discriminatory local features is essential for finding the most convenient feature organization in the tree. We plan to investigate

various methods for organizing features and pruning including tree balancing techniques, Huffman coding methods, and inductive learning using randomization [4], bagging [6] and boosting [10]. These latter methods were developed for classification and are based on a forest of different feature trees. Each of these feature trees is constructed from randomly sampling from the database of images and the classifier is obtained by aggregation of decisions of the trees, e.g. using majority voting or averaging of the posterior distributions, to produce the most stable and informative features. We will extend these methods to image registration in two ways: 1) substitution of more appropriate feature selection criteria than the standard miss-classification error; and 2) aggregating trees by averaging the computed $\alpha$-information or $\alpha$-Jensen values. Examples of feature selection criteria to be investigated are: the curvature of the empirical registration function and the number of coincident features for co-registered images compared to the average number of coincident features for misregistered images.

#### 3.2.2 ICA Features

Independent components analysis (ICA) is an iterative method which is closely related to the projection pursuit technique of non-linear regression and was applied to images by Olshausen, Hyvärinen and others [24, 31, 21]. ICA is an extension of the better known singular value decomposition (SVD) methods and have been shown to yield sparse low dimensional feature spaces for natural images. We apply the method of ICA to the ultrasound registration problem as follows. Based on a random sample of a large number of $M \times M$ subimages from the representative training database of ultrasound images we construct a basis of a small number $p$ of $M \times M$ subimages which decomposes the database of subimages into

$$X_l = \sum_{k=1}^{p} a_{lk} S_k$$

where $\{S_k\}$ are statistically independent components ($M \times M$ images) and $a_{lk}$ are coefficients which approximate each subimage $X_l$ in the database. The selection of of independent components is performed using an independence criterion such as kurtosis, likelihood, or minimum description length (MDL).

As an example, using the ICA algorithm a set of $8 \times 8$ basis vectors were learned from 10 consecutive image slices extracted from a single ultrasound volume scan of the breast (Case151) shown in Fig. 1. Here ICA was implemented using Olshausen's `SPARSENET` code (available from `http://redwood.ucdavis.edu/bruno/`). 64 of the ICA basis vectors are shown in Fig. 7. Only basis elements that corresponded to distinct edge or texture information in the associated basis were retained, resulting in a total of approximately 256 different types. Two different methods of constructing the feature vectors $Z_0$ and $Z_i$ will be investigated: 1) a tag assigned to a particular pixel is the feature type that has least Euclidean distance from the $8 \times 8$ sub-image centered there; 2) sub-images of the reference and secondary images are projected onto the ICA basis and either the coefficients or residuals are used to construct the joint histogram. In method 1 standard plug-in estimates of the $\alpha$-information or $\alpha$-Jensen difference will be used. In the latter case the joint histogram is the realization of a continuous density $f_{0,i}$ and the minimal graph methods of Hero *etal* will give more accurate estimates of the $\alpha$-information or $\alpha$-Jensen difference.

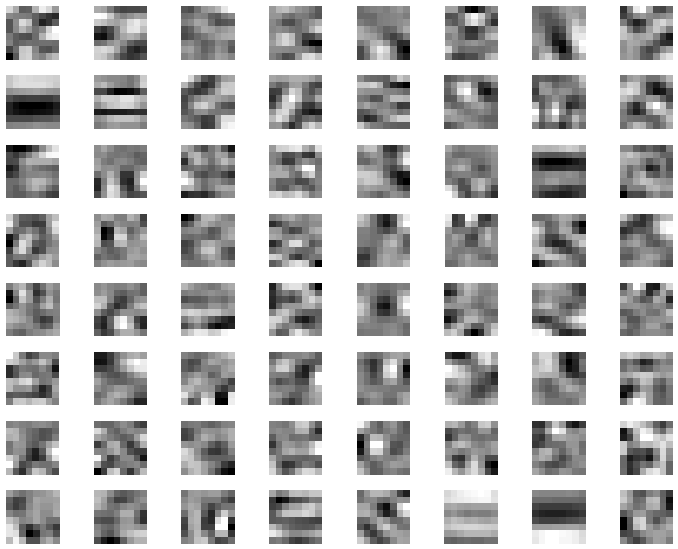

Figure 7: *Estimated ICA basis set for ultrasound breast image database*

#### 3.2.3 Spatially Organized Features

By including selected spatial relationships between local tag and ICA features as part of the feature set, one can enhance registration performance. We capture these relationships by superimposing a disk over each pixel and detecting the simultaneous occurrence in the disk of a tag or feature of given type at the center of the sector and another tag of given type in some outlying spatial sector, e.g. in sectors NW, NE, SE, SW. This additional spatial information can be used to discriminate and eliminate undesired regions, e.g. shadow regions, from each image and also to introduce useful local geometric invariances into the registration process. Local invariance is an attractive property in non-rigid image registration, since it allows pixel neighborhoods (tags) to undergo relative displacement without affecting the overall registration. For example local invariance can be built into the features to allow the registration algorithm to be insensitive to small relative offsets between boundaries due to small compression deformations in one of the two images.

## 4 Preliminary Results

For illustration we used the database of ultrasound breast images represented by the three images in Figure 1. We simulated the decorrelating effect of speckle in two different scans of the same slice of breast by registering a slice to a rotated version of a proximal slice approximately 2mm (up to four correlation lengths) away along the depth of the scan. Feature tree and ICA feature selection were performed without randomization, bagging or boosting. Randomization would be expected to provide even better results than those obtained here. The $\alpha$-information was estimated from the features using plug-in histogram binning methods. For each of the feature sets investigated we searched over the choice of $\alpha$ in $\alpha$-information registration criterion which maximized the curvature, i.e. yielded the most highly resolved peak.

Shown in Figure 8 are typical trajectories (left figure) of the $\alpha$-information registration

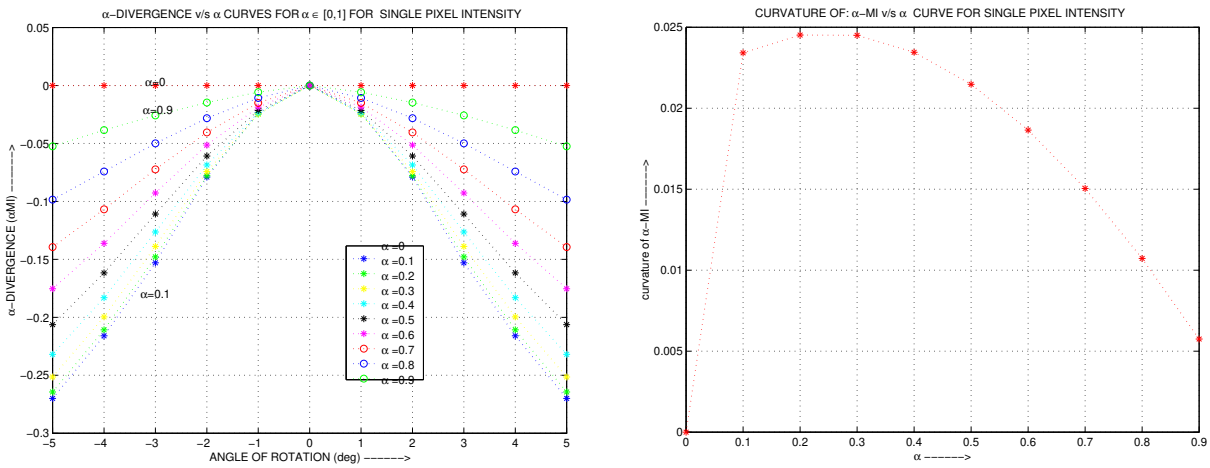


Figure 8: *Resolution of $\alpha$-information as function of alpha for case 151*

| | 151 | 142 | 162 | 151/8 | 151/16 | 151/32 |
|---|---|---|---|---|---|---|
| pixel | 0.3/0.9 | 0.6/0.3 | 0.6/0.3 | | | |
| tag | 0.5/3.6 | 0.5/3.8 | 0.4/1.4 | | | |
| spatial-tag | 0.99/14.6 | 0.99/8.4 | 0.6/8.3 | | | |
| ICA | | | | 0.7/4.1 | 0.7/3.9 | 0.99/7.7 |

Table 1: *Numerator =optimal values of $\alpha$ and Denominator = maximum resolution of mutual $\alpha$-information for registering various images (Cases 151, 142, 162) using various features (pixel, tag, spatial-tag, ICA). 151/8, 151/16, 151/32 correspond to ICA algorithm with 8, 16 and 32 basis elements run on case 151.*

function over rotation angles applied to the secondary image slice for the standard single-pixel features and various values of $\alpha$. Right panel of Figure 8 is a plot of the local (magnitude) curvature computed by parabolic interpolation of each trajectory about its peak. A high curvature value is better than a low value since it implies higher sensitivity to small misregistration angle errors. Note that the highest peak curvature does not occur for $\alpha = 1$ meaning that that the standard Shannon MI criterion is sub-optimal. Shown in Table I are resolution-optimizing $\alpha$ values and resultant optimal peak curvature values for each of the three breast image cases and different sets of features. Note that the resolution optimizing value of $\alpha$ are close to 1 only in 3 out of the 12 cases studied. Note also that all of the proposed feature based methods have significantly higher resolution than the standard single-pixel (column labeled "pixel" in Table I) registration method. Systematic methods for determining the optimal value of $\alpha$ is one of the aims of future work as is the issue of optimally combining the single-pixel, ICA and tag features.

These preliminary results are promising and there remain several important research issues that we will investigate during this grant period. These issues include: selection of features, feature organization in the feature tree, selection of alpha parameter, numerical optimization of the mutual alpha-information, validation and comparison by simulation and experiment.

# 5 Research Plan

We will develop and test our new approaches to registration to increase accuracy and robustness for 3D ultrasound image registration. Our research plan will be comprised of several components: 1) investigation of different types of features for enhancing the capability of the feature tree to discriminate misregistered images in the presence of speckle noise, shadowing and distortions; 2) investigation of the new registration functions, and in particular optimum choices of alpha to sharpen the registration peak; 3) development of fast minimal graph implementations using hierarchies of minimal spanning trees; 4) exploration and compensation for failure modes of the registration algorithm by controlled studies of robustness of the algorithm to increasing amounts of distortion; 5) validation of results with a database of real medical ultrasound images. Initially all of these tasks will be guided by a database of unannotated ultrasound breast images that have been previously acquired from 80 subjects. Later in the project we validate these techniques on muscolo-skeletal scans, e.g. shoulder and achilles tendon, when similar 3D volumetric ultrasound data becomes available. Throughout this work we will perform numerical optimization and grid deformation using, respectively, the Carver-Mead simplex algorithm and the control-point driven thin plate spline software which is implemented in the MIAMI Fuse©.

**1. Feature Selection**

Initially we will concentrate on registration of 2D slices for optimizing the proposed feature selection procedure before moving to 3D extensions. Spatial relationships between tags will be explored that can detect anomolous regions or shadows, and these will be included into the feature tree for additional robustness. Combinations of single-pixels, tags, and ICA features will also be investigated. We will implement randomization, bagging and boosting methods using a subset of 24 of the 80 subjects and evaluate improvements in registration performance on 24 image volumes randomly selected from the rest of the 80 subjects. The 24 subjects chosen for training set will be representative of the eight different classes of breast tissue complexity. Initially, to gauge pre-screened registration performance of our new method we will generate and apply separate sets of features for each class. However, we will also investigate universal sets of features which can automatically screen out false breast classes during the registration process. We will extend the feature selection algorithms to 3D volumes using ICA and feature tree tags in order to eliminate out of plane distortion effects inherent to 2D registration methods. We will also extend these techniques to include 3D color power mode Doppler image information as well as gray scale imaging. Computational issues in 3D feature selection and processing will be addressed (see item 3 below).

**2. Study of New Registration Functions**

We have shown [15] that there exists a theoretically optimal value of $\alpha$ for detecting the peak of the $\alpha$-information trajectory over the set of transformations $T_i$. This theory also suggests that the optimal value of $\alpha$ is small when $f_{0,i}$ is close to $f_0 f_i$, i.e. when the registration is very coarse, while $\alpha$ tends towards 1 for nearly perfectly registered images. Thus we will investigate the idea of varying $\alpha$ over the course of registration: a coarse registration regime for which $\alpha$ is small in the initial iterations and increases to a fine regime in the final iterations. In between these extremes $\alpha$ might be made to increase continuously or step through only a few values over the iterations.

These issues and others will be studied for the simpler-to-compute $\alpha$ Jensen difference. We will also undertake a theoretical and experimental comparison between the registration

performance of the $\alpha$ Jensen difference and the $\alpha$-information. When the primary image volume is exactly equal to a deformation of the secondary image volume then, based on previous study [15], we know that the $\alpha$ Jensen difference and the $\alpha$-information are both optimized at the true deformation. When there are gray level differences in addition to deformations between the two images we strongly suspect that each of these optimizations will yield different registrations. We will to characterize these differences through a combination of parameteric approximations, e.g. a multivariate Gaussian density, to the feature distributions, asymptotic expansions about the optimum deformation, and simulation studies.

The choice of alpha also plays a role in controlling the number of maxima in the registration function. Maintaining a small number of local maxima will be important as exhaustive search is impractical for fully 3D registration having many degrees of freedom in the transformation sequence $\{T_i\}$ applied to the secondary image. The presence of many local maxima in the registration function makes it necessary to reinitialize the maximization several times to make sure that convergence to a global maximum has occurred. The selection of $\alpha$ will be investigated through a combination of analysis, simulation, and experiment with real scans to try and identify what schedules of $\alpha$ yield the fastest possible convergence of the registration algorithm.

**3. Algorithm Acceleration**

As explained above, the implementation of the randomized feature trees and randomized ICA feature classifiers requires two computationally intensive steps: 1) the learning or training of features from a representative database of images; and 2) the implementation of the classifiers to collect coincidence data from the primary and transformed secondary images. In both steps the incidence (tag) or projection (ICA) of every feature (or feature candidate) must be computed at every voxel in the two images, a very time consuming task for 3D image volumes. We will implement simple pre-processing algorithms that can reduce the number of features and feature candidates to be explored. In the learning phase the feature filters will be used which *a priori* eliminate many irrelevant feature candidates that are not present in both images. Such features include speckle and shadowing which can be identified by local intensity discontinuities, consistently dark (low intensity) tags or ICA basis images, and discriminating spatial relationships between local features. In both the learning and implementation phases we will apply 2D and 3D fast algorithms based on matched filtering and mathematical morphology [12] to identify and eliminate from consideration voxels associated with these irrelevant features.

We will also investigate fast minimal spanning tree (MST) approximations for approximating the $\alpha$-information and $\alpha$-Jensen difference registration functions. The plain vanilla MST has polynomial runtime complexity in the number of feature points $n$, e.g. it has $O(n^2)$ complexity in the plane. Recently approximate MST algorithms with nearly $O(n \log n)$ complexity have been implemented [9]. Even faster algorithms can be implemented by partitioning heuristics, like the heuristic we recently introduced in [16]. Initial results are quite promising on a hierarchical scheme where the MST is only run on the centroids of a $k$-means algorithm, yielding average complexity of $O(k \log k) + O(n \log(n/k))$. We will investigate optimal choice of $k$ to minimize runtime yet maintain adequate registration performance. In general it is unclear to what extent approximation of the MST will affect registration accuracy. This important tradeoff between runtime and approximation error will be studied both experimentally and via analysis.

**4. Sensitivity Studies and Robust Algorithms**

We will investigate robust density estimation and the $k$-point MST for rejection of irrelevant

spurrious features that degrade registration performance. As mentioned above, in the case of ICA features (projection coefficients of the image onto ICA basis) the feature densities are continuous, but with unknown degree of smoothness. For this case both adaptive kernel density estimation and edgelet-curvelet based shrinkage estimators will be investigated as plug-in estimators of the $\alpha$-information registration function. These types of density estimators are known to attain minimax error performance over smooth classes of densities [8] and we have obtained similar minimax results [16] for the associated plug-in estimates of the registration function. The pruned MST estimator of $\alpha$-Jensen difference will also be evaluated, where pruning will be implemented to reject spurrious features using the $k$-point MST framework proposed by Hero *etal* [20, 15]. Such MST estimators of the $\alpha$-Jensen registration function can be shown to attain minimax error when the feature densities are not smooth, e.g they contain step or ridge discontinuities [16]. We will resolve the relative advantages of plug-in vs. MST estimators by studying the smoothness properties of the ICA feature density.

To improve robustness, we will also investigate feature types that can indicate the presence and location of anomalies which can be subsequently pruned from the feature tree to enhance registration accuracy. This study will entail identifying the failure modes of the registration algorithm when presented with images having varying degrees of anatomical differences, e,g, due to tumors. We will insert such anomalous regions into speckle contaminated replicates of a single image volume. Performance will be evaluated by both registration accuracy, i.e. distance of registration function peak from true rotation/translation/defoemation, and by post-registration constrast of the anomolous region determined from the difference between the registered primary ad secondary images.

**5. Validation and comparison by simulation and experiment**

We will develop techniques to modify a copy of an image enough to make registration nontrivial while allowing rigorous truth data on registration success. The developed tags, feature sets and tree structures will be evaluated in comparison with the existing single-pixel techniques for accuracy and robustness on 2D data from 24 subjects not used for training of the registration methods. This comparison will be done with the new methods alone and in combination with the single pixel techniques. Use of the best combination of features plus single voxel techniques at the optimized value of $\alpha$ will be compared with single voxel alone at $\alpha = 1$ to establish superiority of our methods on a real database. After integration of the new methods into a module of MIAMI Fuse© and development of 3D tags, features and tree structures, this comparative validation on the 24 subjects will be repeated with registration of nonlinearly-warped 3D gray scale and color flow volumes. Initially, successful registration will be declared if the algorithm produces a difference image with mean-squared-error less than 1% of maximum constrast. It is too early to estimate the variances to calculate the statistical power of this test but we will evaluate experimental $p$-values to establish statistical significance. We will then modify the registration criterion by accounting for spatially localized displacement errors which can be calculated for thousands of points in each scan volume. In the event that high computational complexity prevented whole volume registration we will reduce the study to registering half of the total image volume or less, either through truncation or downsampling of the original whole volume image. Final registration displacement error as a function of the eight classes of breast tissue complexity will be plotted as a function of breast class and we will evaluate the slope of the fit by confidence limits on that slope.