Understanding Bone responses in B-mode Ultrasound Images and Automatic Bone Surface extraction using a Bayesian Probabilistic Framework

Ameet Kumar Jain, Russell H. Taylor
Computer Integrated Surgical Systems and Technology,
Department of Computer Science, Johns Hopkins University, Baltimore, MD - 21218

ABSTRACT

The registration of preoperative CT to intra-operative reality systems is a crucial step in Computer Assisted Orthopedic Surgery (CAOS). The intra-operative sensors include 3D digitizers, fiducials, X-rays and Ultrasound (US). Although US has many advantages over others, tracked US for Orthopedic Surgery has been researched by only a few authors. An important factor limiting the accuracy of tracked US to CT registration (1-3mm) has been the difficulty in determining the exact location of the bone surfaces in the US images (the response could range from 2-4mm). Thus it is crucial to localize the bone surface accurately from these images. Moreover conventional US imaging systems are known to have certain inherent inaccuracies, mainly due to the fact that the imaging model is assumed planar. This creates the need to develop a bone segmentation framework that can couple information from various post-processed spatially separated US images (of the bone) to enhance the localization of the bone surface.

In this paper we discuss the various reasons that cause inherent uncertainties in the bone surface localization (in B-mode US images) and suggest methods to account for these. We also develop a method for automatic bone surface detection. To do so, we account objectively for the high-level understanding of the various bone surface features visible in typical US images. A combination of these features would finally decide the surface position. We use a Bayesian probabilistic framework, which strikes a fair balance between high level understanding from features in an image and the low level number crunching of standard image processing techniques. It also provides us with a mathematical approach that facilitates combining multiple images to augment the bone surface estimate.

Keywords: Ultrasound, Bone, Segmentation, Automatic, Image Registration, CT

1 INTRODUCTION

Due to the rigid nature of the anatomy in Orthopedic Surgery, Computer Aided Orthopedic Surgery (CAOS) has seen many breakthroughs in the past two decades. Accurate execution of any preoperative surgical plan is becoming more and more reliable. One of the most crucial links in achieving this is the ability to accurately register the preoperative CT to the actual human anatomy in the OR [7, 11, 13, 19], the popular method being an Iterative Closest Point (ICP) based registration. A point cloud of the bone surface points was created by using a calibrated tracked pointer which digitized points on the surface of the exposed bone. This was then registered to the triangular mesh-model of the bone surface generated using the pre-operative CT. The registration accuracy depends on the extent of digitized surface. Hence to obtain a more accurate registration, additional surfaces of the bone have to be exposed. B-mode US acts as a non-invasive variable length bone digitizer, while improving registration by covering a larger surface.

Initial developments were made using A-mode US [8, 16], which can digitize a single surface point, a collection of which constitutes the point cloud. B-mode US on the other hand is capable of providing numerous surface points, hence requiring very few scans to register successfully [1, 17]. A significant time consuming part of the procedure is the construction of the bone surface mesh-model from the preoperative CT. To overcome this problem, a direct 3D-3D registration method can be employed [4, 15]. Coupling tracked US images [3], a 3D volume of the bone surface can be

Medical Imaging 2004: Ultrasonic Imaging and Signal Processing,
edited by William F. Walker, Stanislav Y. Emelianov, Proc. of SPIE Vol. 5373
(SPIE, Bellingham, WA, 2004) · 1605-7422/04/$15 · doi: 10.1117/12.535984

created. This volume can be directly registered to the preoperative CT, eliminating the need for the time-consuming semi-automatic segmentation process.

US images are extremely prone to speckle noise and unusual artifacts, making the 3D reconstruction inaccurate. Moreover US image intensity merely represents the distribution of the reflectivity coefficients and hence depends on the machine settings. In order to obtain a good registration we should not rely on image intensity, but filter the images to boost the desired features. A typical approach would be to accurately segment the feature in individual images before the 3D reconstruction. This offers good results provided the segmentation process is dependable, violation of which can cause significant ICP registration errors due to the presence of outliers. The most successful technique for 3D US reconstruction would be to filter the images prior to the 3D reconstruction, which in turn would suppress any non-consistent features [14].

US images are extremely noisy and exhibit various artifacts, yet the surface of the bone is fairly distinct in B-mode US images with many strong features like multiple reflections, shadow regions, etc; which help in deducing the presence of a surface. The accuracy of surface localization is still not very high due to the fact that the surface response is usually very thick, ranging anywhere from 2-4 mm. The main reason is that the imaging model is assumed planar, while the US beam thickness could reach 10mm. Moreover inaccurate calibration, envelope sampling of the RF data and a constant speed of sound assumption also add significantly to the final error. All these couple together, making it difficult to infer the bone surface precisely within an accuracy of 1mm. Consequently, to obtain the best possible registration, we need to accurately know the bone surface in US and then automatically segment it [10, 12, 18, 20], all within the limitations of US imaging.

We claim that since there are inherent uncertainties in resolving the surface location from a single image, we need a framework that can capture the likelihood of the surface occurrence[2, 6, 9]. We should be able to combine these 2D likelihoods with various other images and generate a superior 3D likelihood of the surface. This would give us a better quality estimate of the surface, when compared to blind techniques that create a 3D volume by simple averaging. A segmentation algorithm needs to accommodate for the likelihood of the surface based on image features like shadows, reflections, incidence angle, etc; a combination of which with high level human understanding should conclude the actual bone surface. The framework needs to accommodate for the occurrence of these features, their detection and the logical reasoning that follows [5]. A probabilistic framework accommodates for all the requirements. Since the final goal is registration our segmentation algorithm will not segment image pixels, rather it would give us the probability of pixels being on the surface. Probabilistic knowledge from independent sources is merged using Bayes rule. Using all the data, a more accurate three-dimensional surface likelihood volume can be created. This volume can either be used to segment out the bone directly (localized high probability connected points) or can be used as an input for registration. The framework also indirectly addresses the cyclic problem of segmentation and registration.

In Section 2 we discuss the response of a highly specular surface (bone) when imaged using US. This is followed by the mathematical probabilistic framework in Section 3. Section 4 elaborates on the *in vitro* experiments and preliminary results. We wrap up with future work and conclusions in Section 5.

2 BONE RESPONSE IN B MODE ULTRASOUND

US images give us an indication of high reflectivity points in the medium, *i.e.* they designate tissue boundaries. A high intensity pixel indicates a strong likelihood of the presence of a boundary, yet is not an absolute measurement of the reflectivity coefficient. Our ability to successfully use US images for image guided techniques depends on the accuracy in localizing the surface boundary, which is inherently limited by the uncertainty in the imaging physics. The bone surface is highly specular, creating a very high intensity feature in the image followed by a shadow. This high intensity feature looks like a line with a shape closely resembling the surface. However the thickness of this line can reach a value of 4mm in certain cases. We know that the actual surface is a unique thin line in this thick response and our problem is to accurately segment it. In this section we attempt to understand how any specular surface would be visible in US. More importantly given a thick line in the image that represents the bone, we try to understand where to expect the bone surface.

To understand this better, we performed *in vitro* experiments. We glued various fiducials with known dimensions on the surface of a pelvic bone immersed in a water tank. Localized high resolution US images were taken for each of the fiducials and appropriate measurements were made using the internal software, which are accurate to 0.1mm. By knowing the physical dimensions of the fiducials, we can estimate the position of the bone surface. Figure 1 shows some typical US images acquired with a CT fiducial, a plastic fiducial and an Optotrak marker attached. Note that to image the fiducial accurately, the US beam direction is orthogonal to the bone surface being imaged. The smaller dot-like response is the top of the fiducial surface, while the second thick response is that of the bone. Primarily there are 3 important measurements to be made – the dimensions of the first response, the thickness of the second response and the distance between the initial occurrences of both these responses. These along with the individual dimensions for each fiducial are given in Table 1.

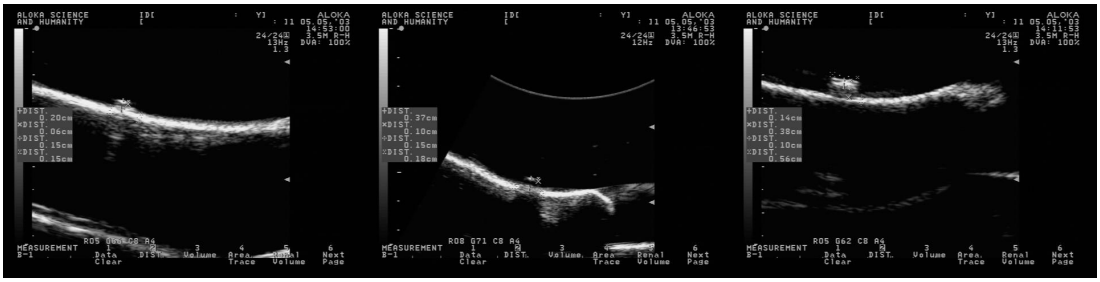

Figure 1: *High resolution US images with various fiducials attached on the bone surface – CT marker (left), Plastic (middle), Optotrak (right). Various measurements are made in each image, some of which are visible in the sample images.*

Fiducial Object	Fiducial Response Thickness (mm)	Distance to Bone (mm)	Actual Fiducial Thickness (mm)	Bone Response Thickness – left (mm)	Bone Response Thickness – right (mm)
CT Fiducial	0.6	2.0	2.0	1.5	1.5
Plastic Fiducial 1	1.0	3.7	3.7	1.5	1.8
Plastic Fiducial 2	0.8	4.0	3.7	-	-
Optotrak Marker	1.0	3.8	3.5	1.0	-
Flat Patch 1	1.6	1.6	2.2	1.2	1.6
Flat Patch 2	1.5	1.6	2.2	1.6	-
Wire1	0.6	2.5	1.3	1.2	1.1
Wire2 (axial)	1.0	1.5	1.5	-	-
Wire3 (axial)	0.8	1.4	1.4	1.0	1.0
Wire4	0.5	2.1	2.1	1.0	1.1
Wire5	0.5	3.2	2.5	1.6	1.0
Thread1	0.7	0.9	0.8	1.6	1.6
Thread2	0.6	0.7	0.7	1.7	-

Table 1: *Measurements from images taken with different fiducials attached to the bone surface. The physical thickness of the fiducial is measured using an accurate vernier calliper. All other measurements are made using the internal tool offered by the US machine.*

We notice that in most cases there is an agreement between the fiducial thickness (physical) and the distance of the first response of the fiducial to the first response of the bone surface. Fiducials with a point-like top surface (spherical CT fiducial, wires with the US image plane cutting them radially and threads) exhibit a response of about 0.6mm; the response is about 1.5mm for flat surfaces (flat patch 1&2); ranges from 0.6mm – 1.5mm depending on how flat the surface is. It is also noticed that bone surface response varies from 1.0mm – 1.8mm. It is to be noted that these images are magnified by the US machine and hence the envelope sampling of the US echo at this resolution produces a bone

response of about 1.5mm. Images created using the standard resolution produce a significantly thicker response. The measurements seem to indicate that the Point Spread Function (PSF) has a thickness of about 0.6mm and the Line Spread Function (LSF) has a thickness of about 1.5mm.

Depending on the position of the top surface of the fiducial inside the respons, we can predict the position of the bone surface. Looking at the data, it seems most probable that the fiducial surface (and hence the bone surface) lies in the top part of the response. The reasoning for this comes from the observation that independent of the fiducial being observed, the distance between the starting points of the two responses is equal to the physical dimensions of the fiducial. Although it is less likely, there still is a chance that the actual bone surface could be embedded inside the bone surface response to a (maximum) distance of 0.6mm, i.e. the minimum response from any fiducial.

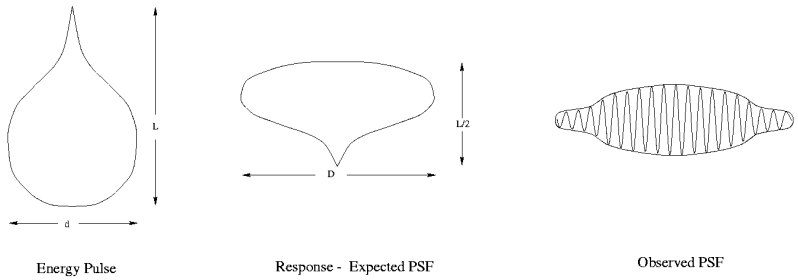

Figure 2: *Approximate shape of an incident acoustic energy pulse (left); expected theoretical response from a ideal point like reflector (middle); the observed response in the high resolution images is in agreement with the expected theoretical response (right)*

In what follows, a model is being suggested for the response of any strongly specular surface when imaged using US. It is known from the physics of US imaging that the acoustic energy propagated in the medium is shaped in the form of a tear drop. With simple calculations, it can be seen that the expected shape of the response (reflected energy) will be an inverted tear drop. Since the imaging device calculates the distance of the object using the time at which the echo is received, the length of the return signal is halved. The thickness of the tear drop is not a constant and changes with the time of travel (increasing after it reaches the plane of focus). The typical length (L) of the tear drop is about 3λ, where λ is the wavelength of the wave which remains a constant. Coupling the frequency of the acoustic wave (3.5MHz US probe) with the speed of sound in water at room temperature (1540m/s), we obtain the wavelength (λ) to be 0.44mm. Thus the length of the expected PSF turns out to be 0.66mm ($3\lambda/2$), which agrees closely to our measurements. As shown in Figure 2, the shape of the observed response agrees well with that of the expected PSF.

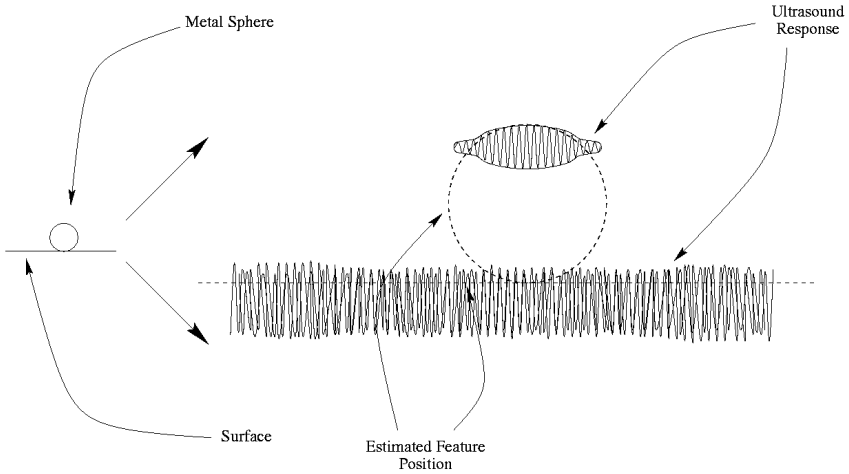

Figure 3: *The expected model for the image formation with the spherical CT fiducial attached to the bone surface. The top of the fiducial sphere gives a response similar to the expected PSF and the flat bone surface gives a response similar to an expected LSF.*

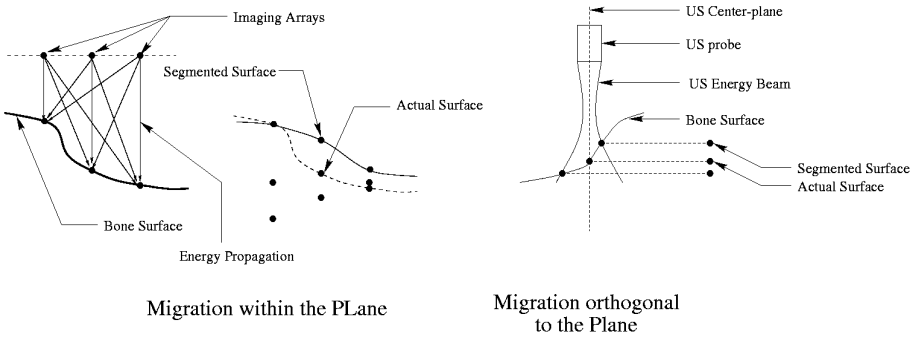

Figure 4: *Exact location of bone surface may migrate within the response, the amount of migration depending on local 3D geometry. Each imaging array may receive reflections from surfaces outside its direct line of sight and hence records a thicker response along its own scanline (left). Moreover a tilted surface may convolve with the 3D US energy beam and produce a thick response when projected to the 2D imaging plane (right).*

Figure 3 describes the image formation process for a CT fiducial attached to the bone surface where the reflecting surface is normal to the direction of incident energy. The top of the fiducial behaves like a point reflector, while the surface acts like a line. This is in close agreement with what is observed in Table 1. There are two main reasons why any specular surface produces an extended (thick) response. These are illustrated in Figure 4.

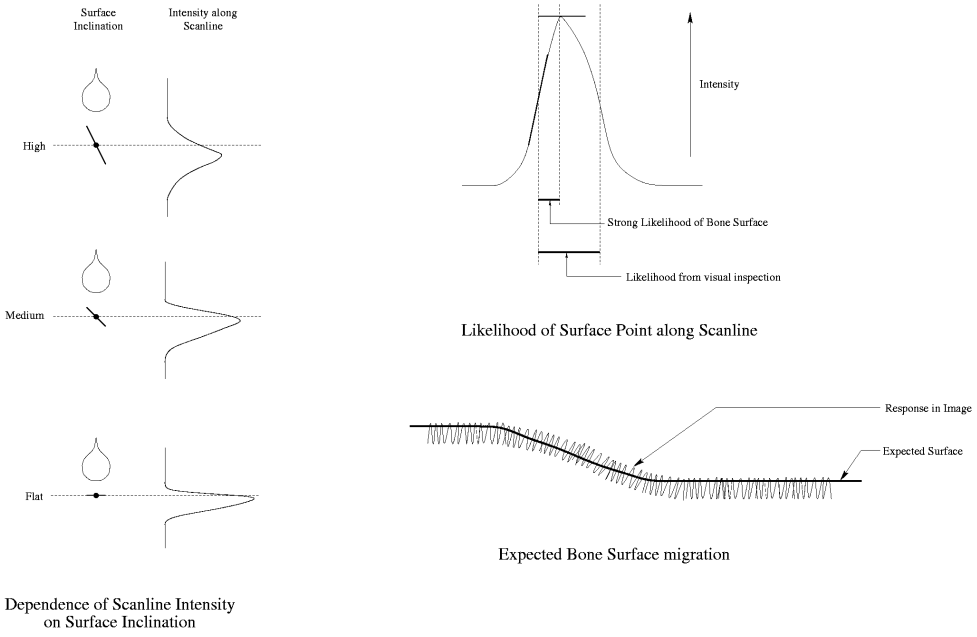

Figure 5: *Effect of the local bone surface geometry on the shape of the intensity profile along each scanline. Greater the inclination, thicker is the response (left). Given intensity profile alone a scanline or a fuzzy line response in the image, the high likelihood expected localization of the surface of the bone (right). Actual position of the surface could vary depending on the local geometry and cannot be deduced by looking at a single image. Similar responses from different images could have the surface at different positions.*

The basic imaging model for ultrasound is assumed planar, i.e. the received echo is projected to a plane. On the other hand, the true beam profile is 3D. Moreover the scanline for each imaging array element is assumed single dimensional, while it is infact a tear drop with a depth varying thickness. In other words, the final US image is a planar projection of, a 3D convolution of the depth varying energy pulse with the reflectivity coefficients in the medium. Figure 5 illustrates how the local direction of the bone surface would influence the echo intensity structure along any scanline. More importantly it indicates the presence of the bone surface within the intensity response arising in the different cases. Thus given an intensity profile we expect the bone surface to have a strong likelihood between the point of highest gradient and highest intensity. This whole region stands equi-probable to the possibility of containing a bone surface, i.e. the actual intensity value does not have any strong bearing in the bone surface.

The highest gradient signifies the maximum resistance being offered to the incident energy, while the highest intensity signifies the maximum energy being reflected. The actual surface is close to these two points and cannot be localized precisely from the image. Knowledge of the true 3D geometry and its relationship to the present image-slice orientation is required to make any further accurate inference. Without this information it would be inaccurate to try and segment out the bone surface definitively. Any algorithm which does that, is inherently adding an error (from lack of sufficient information) to its result. However, the amount of added error and its acceptability depends completely on the end-application. One possible way to overcome this problem is to be able to take multiple scans of the same surface from different orientations and couple them together to create a 3D volume. The 3D volume from multiple scans would boost consistent data, while it would suppress inconsistent artifacts. Thus a more accurate information boosting process is developed. Section 3 develops the mathematical framework which would be capable of accommodating the above concepts and couples it with standard image processing techniques.

3 A PROBABILISTIC FRAMEWORK FOR BONE SEGMENTATION

We design a probabilistic framework to mathematically incorporate the inherent uncertainties involved in US bone imaging. The output of the segmentation process will be a probability value at each pixel signifying the likelihood of a surface. In other words the image acts like an occupancy grid, with each pixel position being a binary random variable *S(i,j)*. We omit the use of pixel coordinates *(i, j)* for readability. The occupancy grid is composed of cells; our aim being to evaluate the expected state of each cell. In the framework that follows, the occupancy grid is modeled as a Markov random field (MRF) of order 0, i.e. individual cell states can be estimated independent of each other. Though this is not completely true, it is a fair assumption since in general the occupancy state of one cell is independent of any other cell that is far away. Note that a surface never occurs in isolation, and hence the states of the immediate neighbors of any cell are not completely independent of each other. For higher order MRFs, computationally more expensive models can be incorporated. In what follows *P* represents the probability of an event.

$$S(i,j)=\begin{cases}1 & if \quad pixel \quad (i,j) \quad is \quad on \quad the \quad surface \\ 0 & if \quad pixel \quad (i,j) \quad is \quad not \quad on \quad the \quad surface\end{cases}$$

$$P[S(i,j)=1] \quad + \quad P[S(i,j)=0] \quad = \quad 1$$
$$P[S(i,j)=1] \equiv P[S(i,j)] \equiv P(S)$$
$$P[S(i,j)=0] \equiv P[\bar{S}(i,j)] \equiv P(\bar{S})$$

Equation 1: *Each pixel has a binary random variable, the value of the RV indicating the presence of a surface or the absence of it at that pixel. Each RV can only take a value of either 0 or 1. Abbreviations are used for easy readability.*

An image *(I)* consists of various features $\{F\}_i = \{F_1, F_2, F_3 \ldots F_i\} \equiv I$. Any inference from the image relies on our ability to extract and utilize these features. The probability of interest is the probability of the random variable at each pixel,

given the image. One method to implement this would be by extracting different features in succession and then applying them to the present probabilistic estimate. Bayes formula is an excellent way to do the mathematical integration.

$$P(Surface|Image) = \frac{P(Image|Surface)P(Surface)}{P(Image)}$$

Equation 2: *An application of Bayes Rule to infer the surface probability. This is also used in Maximum Likelihood Estimate based segmentation techniques.*

$$P(S|\{F\}_{i+1}) = \frac{P(F_{i+1}|S)P(S|\{F\}_i)}{P(F_{i+1}|S)P(S|\{F\}_i) + P(F_{i+1}|\overline{S})P(\overline{S}|\{F\}_i)}$$

Equation 3: *A simple iterative rule for updating the probability values by adding another feature. The formula is an extension of the Bayes rule. Given a cumulative probability estimate for the surface based on i features, this formula gives us the new cumulative estimate based on (i+1) features. It requires a model for the occurrence of the (i+1)th feature given information about the surface.*

In this recursive formulation [5], the cell probability estimates can be updated based on any new feature discovered in the image. The present probability estimate is stored in the occupancy grid, which stores the probability for each cell based on all the features discovered so far. The algorithm needs the initial priors to start the iterations and a model for $P(F_i|S)$ and $P(F_i|\overline{S})$. The initial priors can be initialized to the maximum entropy setting of P(S) = $P(\overline{S})$ = 0.5 for each cell in the grid. The models for $P(F_i|S)$ and $P(F_i|\overline{S})$ need to be developed individually for each feature. Equation 4 provides some basic rules that aid in developing these models.

$$\begin{aligned} P(\overline{S}|F_i) &= 1 - P(S|F_i) \\ P(\overline{F_i}|S) &= 1 - P(F_i|S) \\ P(F_i|\overline{S}) &\neq 1 - P(F_i|S) \end{aligned}$$

Equation 4: *Some simple rules that help in computing the probability estimates based on various factors.*

In some cases, it becomes extremely difficult to estimate a reliable model for $P(F_i|S)$, but it may be easier to directly compute $P(S|F_i)$. Thus if we have two different estimates from two different sources, we need a way to couple the probabilistic information together. Let the two computed estimates be $P(S|\{F_i\})$ and $P(S|\{F_j\})$ for two different features $\{F_i\}$ and $\{F_j\}$. By attempting to integrate these probabilities together, we are evaluating $P(S|\{F_i, F_j\})$. Two standard methods to do this are by using either a *Linear Opinion Pool* or an *Independent Opinion Pool*, as described in equation 5.

$$\begin{aligned} P(S|\{F_i, F_j\}) &= w_1 P(S|\{F_i\}) + w_2 P(S|\{F_j\}) \\ P(S|\{F_i, F_j\}) &= \left(\frac{1}{\kappa}\right) P(S|\{F_i\})P(S|\{F_j\}) \end{aligned}$$

Equation 5: *Rules for combining probabilities from different sources, choice of which depends on their dependence on each other.*

In the linear opinion pool method w_1, w_2 reflect the confidence in the respective information sources and $w_1 + w_2 = 1$. A deficiency of this method is that it does not allow adequate reinforcement of similar opinions. When the information sources are independent, the *independent opinion pool* method can be used, where the constant κ is for normalization. When applicable we prefer the *independent opinion pool* to combine the probabilities, since it does not rely on any weight parameters to be chosen by a global decision-making agent. The combination of the two methods is easily extended to cases with many features and cumulative estimates can be constructed as a combination of the above.

Moreover, this technique can combine relevant information from sources other than the image, for example from a prior registration.

$$P = \frac{P_1 P_2}{P_1 P_2 + (1-P_1)(1-P_2)}$$

Equation 6: *Rule for combining probability estimates from two completely independent sources, to obtain an updated estimate.*

For example, let P_1, P_2 be two such probability estimates from independent sources. Equation 6 describes the global estimate for the combined probability using the independent pool method, which has extremely desirable properties. It is commutative and associative, making the order of feature incorporation immaterial. The combination of a probability , with 0.5 (unknown), yields *p* as the final answer. Combining conflicting estimates *(0.5-p)* with *(0.5+p)* gives 0.5 (unknown) as the answer.

After constructing the necessary mathematical framework for combining probabilities, we need to identify individual features that are associated with a bone surface response. Some of the features typical in US images of the bone are high intensity, high gradient, specific shape of the intensity-profile along a scanline, multiple reflections, shadow regions, angle between the incident energy and the surface normal, etc. with very simple filters to recognize them. For most of these features either a simple feature model $P(F \mid S)$ can be constructed, or a surface prediction $P(S \mid F)$ can be made. In both cases, the inferred probabilities can be coupled together to create a consolidated final expectation map of the surface of the bone.

The probability map $P(S \mid F)$ can be very easily constructed for the intensity feature. In general higher the intensity, higher is the likelihood of a surface at that pixel. Hence, a simple linear scaling of the intensity values is sufficient. A more sophisticated model can also be used but is unwarranted since this is just a beginning estimate. The second feature is the approximate shape of the intensity profile near the bone, which is not required to be extremely precise but is an estimation of the observed shape. We looked at various intensity profiles from various images and observed that the response is a curve that is usually about 7 pixels thick with a very rapidly rising front edge and a slow decaying falling edge. We modeled it by joining 2 Gaussians having a standard deviation of 1 pixel for the rising edge, a standard deviation of 2 pixels for the falling edge, and a maximum value of 1.

Once we have an approximate model for a bone surface response, we need to construct the probability map of $P(F \mid S)$ and $P(F \mid \overline{S})$. It should be noted that S here is a conditional random variable, i.e. $S \equiv \langle [S(i,j)=1] \mid \{F\}_i \rangle$ where $\{F\}_i$ is the list of features considered so far. Thus the requirement is to compute the probability map $P(F \mid S)$ with the present estimate of the cell states. The intensity model does not directly correspond to $P(F \mid S)$ since the former relates to the US response from a single surface, while in reality two surfaces that are close to each other may influence the other's response. Equation 7 gives us the necessary structure to compute this, where $G_{S(i,j)}$ stands for a specific local grid configuration that includes all cells except *S(i,j).* Similar computations estimate $P(F \mid \overline{S})$.

$$\begin{aligned} P(F \mid S(i,j)) &= P\left(F \mid \left[S(i,j) \wedge S(u,v) = {}^{0}\!/_{1} \quad \forall (u,v) \in Neighbourhood(i,j)\right]\right) \\ &= \sum_{\{G_{S(i,j)}\}} P\left(F \mid \left[S(i,j) \wedge G_{S(i,j)}\right]\right) P\left(G_{S(i,j)} \mid S(i,j)\right) \\ \sum_{\{G_{S(i,j)}\}} P\left(G_{S(i,j)} \mid S(i,j)\right) &= 1 \end{aligned}$$

Equation 7: *The rule for estimating the probability of a feature based on the information from neighboring pixels. It uses as input a model for the feature based on information from a single cell.*

Configuration Probability	Configuration	Surface Inference
$(1-p_1)p_2p_3$	(0, 1, 1)	(0, 1, 0)
$p_1p_2(1-p_3)$	(1, 1, 0)	(1, 0 ,0)
$p_1p_2p_3$	(1, 1, 1)	(1, 0, 0)
$1 - p_1p_2 - p_2p_3 + p_1p_2p_3$	5 possibilities	(0.5, 0.5, 0.5)

Table 2: *When observing multiple reflections, there are 8 possible combinations possible, each having a certain probability. Since they are far away from each other, the state of each cell is considered independent of the other; allowing us to compute the probability of the combination. Out of the eight configurations, only 3 indicate multiple reflections. The subsequent surface inference is shown in the last column.*

Another strong feature that can be observed in the images is the presence of multiple reflections of the bone. When the bone surface is normal to the incident energy beam, it reflects most of the acoustic energy. This reflected energy in some cases is reflected again from the crystal creating another energy pulse. These subsequent energy pulses can give rise to artifacts in the image which appears as several occurrences of the bone surface, spaced at some multiple of the first occurrence. These reflections can be high intensity and can be very easily mistaken for the true surface. To obtain an overall model for $P(F \mid S)$ would be extremely difficult since secondary reflections can be very difficult to predict, without an accurate knowledge of the local 3D geometry. On the other hand it is not too difficult to infer information about the likelihood of a surface based on the reflections it exhibits. Hence this is an excellent case where it is easier to compute $P(S \mid F)$ than $P(F \mid S)$. Given a reflection number (secondary, tertiary,) any cell indicating a high probability for the surface can either be a reflection of some surface or be a surface point. Moreover if it is a surface point, then it can produce a reflection. We need to identify these cases when they occur and appropriately compute the probability map $P(S \mid F)$. Let *n* be the reflection number that is being considered. If *i* is any scanline and *j* the depth for some cell *(i,j)* in the grid; p_2 the current probability of *(i,j)* being a surface point, p_1 the current probability of *(i/n, j)* being a surface point and p_3 the current probability of *(in, j)* being the surface point. So either *(i, j)* can be a reflection of *(i/n, j)* or *(in, j)* can be a reflection of *(i, j).* Since we consider each of these cells states to be independent of each other, we can compute the probability of every possible configuration of the 3 cells. Moreover for each configuration, we can compute the probability $P(S \mid F)$ for the cell *(i, j).* Table 2 gives us the details for each configuration. A value of *0.5* indicates that no additional inference can be made about the surface.

$$P(S \mid F) = \begin{cases} 0.5 + 0.5(-p_1p_2 + p_2p_3 - p_1p_2p_3) & if \quad (in, j) \in IMAGE \\ 0.5 + 0.5(-p_1p_2 - p_2p_3 + p_1p_2p_3) & if \quad (in, j) \notin IMAGE \end{cases}$$

Equation 8: *The probability map for surface inference based on multiple reflections. The probability values are computed using the information from Table 2.*

We have looked at a framework to convert high level understanding of features into a probabilistic platform. The platform is capable of merging relevant information from other sources too. The application we consider does not require us to hard segment the surface from each image. We will use the final probability maps to couple information together from various images and create a 3D volume. Many simple heuristics can hard segment the surface in each image, an example being the maximum probability cell from each scanline.

4 RESULTS AND DISCUSSION

Experiments were conducted to acquire *in vitro* data to test the proposed framework. A female pelvic bone was immersed in a water tank and fixed rigidly. The system described in [3] was used to capture 3D tracked images of the bone. A set of 100 spatially tracked images constituted a single data set. These images were supplied as input to the image segmentation program which produces the final bone probability map. The algorithm was implemented using the Matlab 6.5 environment installed on a PC with a Windows 2000 operating system (2.4GHz Pentium4 with 512MB RAM). The algorithm ran for 5s for each image before it produced the final probability map, with 80% of the time being

used to remove multiple reflections, scan-convert the image to produce scanlines and to back-scan-convert the scan-lines to form the final image.

Figure 6: *The performance of the method for a typical image. Most of the artifacts/noise has been successfully removed by reducing its probability of being on the bone surface. The original image (left), the final probability map (middle right), probability of pixels more likely to be bone than not bone (right), the highest probability point along each scanline (middle left).*

Figure 6 displays the segmentation performance of the algorithm for a typical image. It shows the original image, a segmentation of the highest probability value for each scanline, the final probability map (PM) and the pixels with a probability greater than 0.5 in the final PM (more likely to be bone than not bone). A precise pixel wide segmentation of the surface is not suggested as input to a registration algorithm. It is shown to give the reader an estimate on the general performance of the algorithm. A numerical accuracy analysis of the segmentation algorithm was not performed for two reasons: a) an accurate segmentation attempt from a single US image is inherently flawed (Section 2); b) The combination of US calibration error and a tracker error would produce a mean error of about 1mm which would be insufficient to reliably measure the segmentation to an accuracy of 1 pixel (0.5mm).

Figure 7 shows some of the intermediate probability maps generated for the same image. Image1 is the original image scan-converted to create scan-lines, 2 is the image with simple noise removal, 3 is the PM based on intensity, 4 is the expected response probability given the presence of a surface at that pixel, 5 is the expected response probability given the absence of a surface at that pixel, 6 is the Bayesian PM based on 4&5, 7 is the PM based on multiple reflections and 8 is the final PM combining all the above. The different rows in Figure 8 show the performance of the algorithm for different images. The first column is the input image, the second column the maximum probability point as the segmented surface, the third column is the final probability map and the last column is the probability map more likely to be bone than not bone (probability > 0.5).

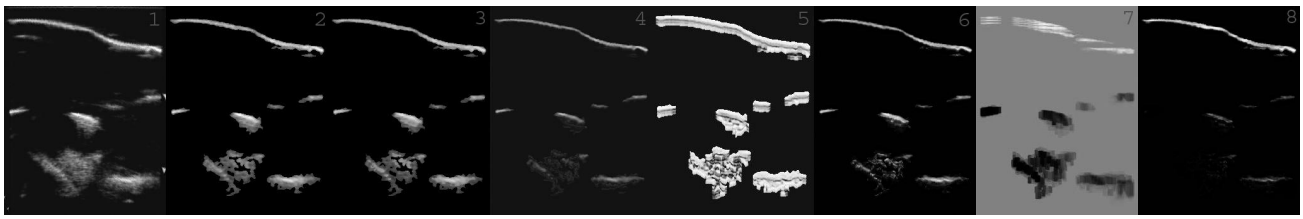

Figure 7: *Various intermediate Probability Maps (PM). (1) Scanlines from the original image (2) Simple noise reduced image (3) Intensity based PM (4) PM showing the shape correlation at each pixel given that it is probably a surface point (5) PM showing the shape correlation at each pixel given that it is not a surface point (6) Shape based PM obtained by coming 5&6 (7) PM indicating multiple reflections (8) Final PM obtained by the combination of all the above*

We can see that the algorithm works well and is able to filter most of the noise. Pixels with probabilities greater than 0.5 indicate the presence of a bone surface in close proximity. When similar information from other US images is coupled together, positive reinforcement of the bone surface probabilities will to boost the actual surface. Artifacts/noise from any image will not get reinforced and hence would quickly die to a near-zero probability value. Thus the final 3D probability map would reveal voxels having high probability for bone surface, which can be directly registered to the preoperative 3D CT volume. Exact segmentation of the bone surface, though unnecessary, can be done on the 3D PM to extract the surface. This 3D surface would be more accurate than any segmentation from a single US image, since the 3D probability map encodes the 3D local geometry.

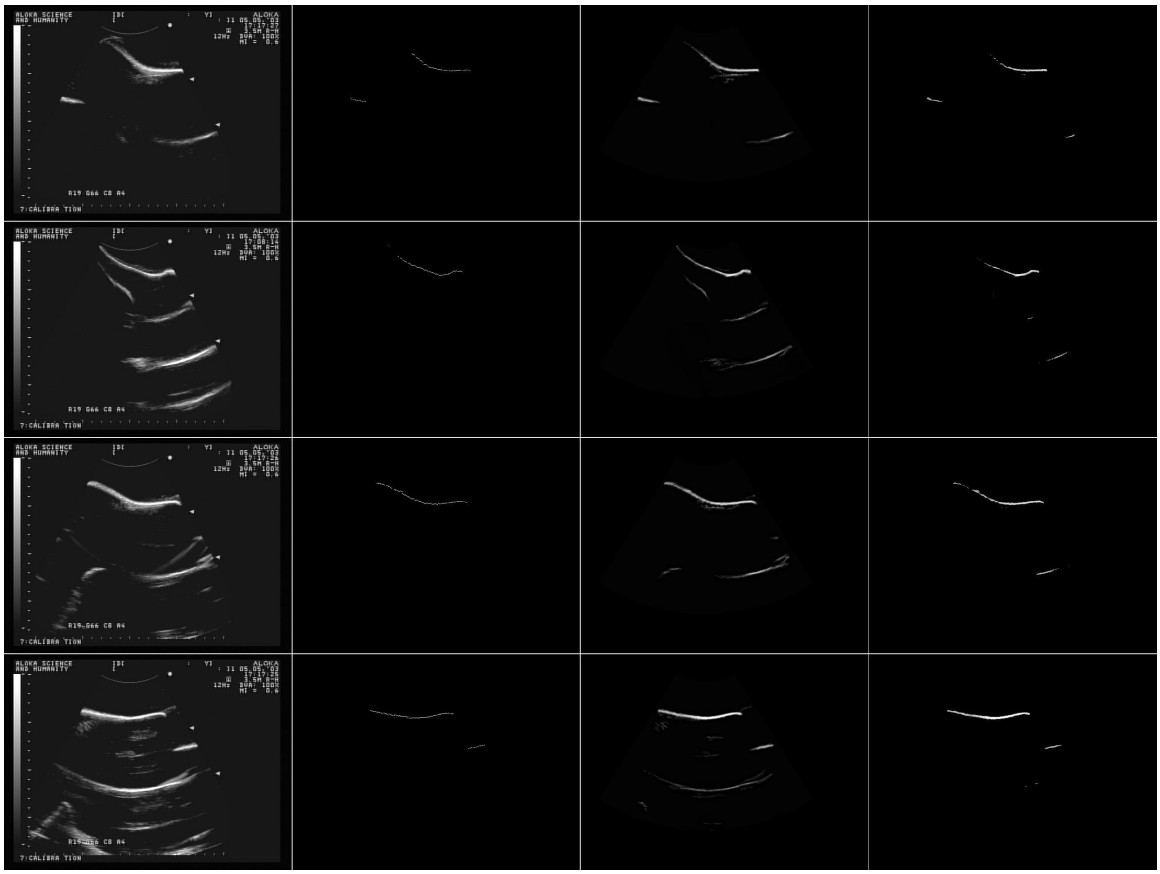

Figure 8: *Performance of the proposed framework on various images. Original images (column1), Final Probability Maps (column 3), Probability of pixels more likely to be bone than not bone (column 4), Highest probability pixels along each scanline (column 2).*

5 CONCLUSION AND FUTURE WORK

We have studied the uncertainties associated with the response of a highly specular surface when imaged using US. Our experiences with the proposed probabilistic framework indicate that the framework is feasible for segmentation and registration algorithms. It is generic and extendable to other anatomies or imaging modalities. It is broad enough to accommodate for the various features observed in any image, and their low level computations, which are based on high level understanding of how the features interact to produce a response. The framework can model an uncertainty in any feature, which would be resolved later by the addition of further specific information. Thus it is able to couple exact knowledge with inexact information. This knowledge can be from a variety of sources and not limited to a single imaging modality. Thus our framework offers a platform to combine information and uncertainties in our understanding of the environment. The final blending of information would boost the desired knowledge, while suppressing the undesired noise.

The features incorporated presently are the intensity values in an image, the shape of a bone response and multiple reflections. Other important features observed in US images of the bone that have yet to be included are angle of incidence, shadow regions below the bone, etc. The suggested new 3D PM reconstruction methodology has to be implemented, followed by a final registration step, the accuracy of which would validate the success of the framework. Since the framework is extremely broad, it can be used to construct a more intelligent 3D PM for the bone surface by incorporating the tracker uncertainties. *In Vivo* experiments need to be performed to conduct validation in real scenarios.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the support and help offered by Mr. Emad Boctor for supplying the US images whenever needed. We also thank Ming Li for the Pelvic bone and the CT scan; Prof. Gregory D. Hager, Prof. Randy E. Ellis and Prof. Jerry L. Prince for their advice. The work is supported by the National Science Foundation under various Engineering Research Center grants.

REFERENCES

[1] D. Amin, *"US Registration for Surgical Navigation"*, PhD Thesis, CMU, 2001.
[2] E. Ashton and K. Parker, "Multiple Resolution Bayesian Segmentation of Ultrasound Images",Ultrasonic Imaging,**17,** 291 - 304, 1995.
[3] E. M. Boctor, A. Jain, M. Choti, R. H.Taylor and G. Fichtinger, "Rapid calibration method for registration and 3D tracking of ultrasound images using spatial localizer", *SPIE Medical Imaging: Ultrasonic Imaging and Signal Processing*, 5035, 521-532, San Diego, 2003.
[4] B. Brendel, S. Winter, A. Rick, M. Stockheim and H. Ermert, "Registration of 3D CT and Ultrasound Datasets of the Spine using Bone Structures",Computer Aided Surgery,**7,** 146 - 155, 2002.
[5] A. Elfes, *"Occupancy Grids: A Probabilistic Framework for Robot Perception and Navigation"*, PhD Thesis, CMU, 1989.
[6] M. T. D. Figueiredo and J. M. N. Leitao, "Bayesian Estimation of Ventricular Contours in Angiographic Images",IEEE Transactions on Medical Imaging,**11,** 416 - 429, 1992.
[7] E. A. Firle, S. Wesarg, G. Karangelis and C. Dold, "Validation of 3D Ultrasound - CT Registration of Prostate Images", *SPIE*, San Diego, Feb 2003.
[8] D. R. Gauldie, *"Calbration and Registration with 3D A-mode Ultrasound"*, Master's Thesis, Queen's University, 2002.
[9] C. A. Glasbey, "Ultrasound Image Segmentation using a Point Distribution Model in a Bayesian Framework", *British Machine Vision Conference*, 113–122, 1996.
[10] P. He and J. Zheng, "Segmentation of Tibia Bone in Ultrasound Images using Active Shape Models", *23rd Annual Conference - IEEE/EMBS*, 1-4, Istanbul, Turkey, Oct 2001.
[11] G. Ionesco, S. Lavallee and J. Demongeot, "Automated Registration of Ultrasound with CT Images: Application to Computer Assisted Prostate Radiotherapy and Orthopedics", *MICCAI*, 768 - 777, Cambridge, Sept 1999.
[12] J. Kowal, C. Amstutz, J. Ioppolo, M. Styner and L.-P. Nolte, "Fast Automatic Bone Contour Extraction in Ultrasound Images",IEEE Transactions on Medical Imaging,**Conditionally Accepted,** 2003.
[13] J. Kowal, R. Hamdan, P. Heini, R. Kothe, L. P. Nolte and M. Styner, "Ultrasound Based Registration for Minimally Invasive Interventions", *CAOS*, abstract, Marbella, Spain, 2003.
[14] D. F. Leotta and R. W. Martin, "Three-Dimensional Spatial Compounding of Ultrasound Scans with Weighting by Incidence Angle",Ultrasonic Imaging,**22,** 1-19, 2000.
[15] K. E. Lunn, A. Hartov, E. W. Hansen, H. Sun, D. W. Roberts and K. D. Paulsen, "A Quantitative Comparision of Edges in 3D Intraoperative Ultrasound and Preoperative MR Images of the Brain", *MICCAI*, 1081 - 1090, Utrecht, The Netherlands, 2001.
[16] C. R. Maurer, R. P. Gaston, D. L. G. Hill, M. J. Gleeson, M. G. Taylor, M. R. Fenlon, P. J. Edwards and D. J. Hawkes, "AcouStick: A Tracked A-Mode Ultrasonography System for Registration in Image-Guided Surg", *MICCAI*, 953 - 962, Cambridge, Sept 1999.
[17] D. M. Muratore, J. H. Russ, B. M. Dawant and R. L. Galloway, "Three-Dimensional Image Registration of Phantom Vertebrae for Image-Guided Surgery: A Preliminary Study",Computer Aided Surgery,**7,** 342-352, 2002.
[18] J. G. Thomas, R. A. Peters and P. Jeanty, "Automatic Segmentation of Ultrasound Images Using Morphological Operators",IEEE Transactions on Medical Imaging,**10,** 180 - 186, Sept 1991.
[19] J. Tonetti, L. Carrat, S. Lavallee, P. Cinquin, P. Merloz and L. Pittet, "Ultrasound based registration for percutaneous computer assisted pelvis surgery : Application to Iliosacral screwing of pelvis ring fractures", *CARS*, Berlin, 1997.
[20] Y. Zhang, R. Rohling and D. K. Pai, "Direct Surface Extraction from 3D Freehand Ultrasound Images", *IEEE Visualization*, 45 - 52, Boston, Oct 2002.