

Fast Algorithm for Probabilistic Bone Edge Detection (FAPBED)

Danilo Šćepanović^{*a}, Joshua Kirshtein^b, Ameet Kumar Jain^b, Russell H. Taylor^b

^a Department of Biomedical Engineering, Johns Hopkins University, Baltimore, MD – 21218

^b 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). FAPBED is designed to process CT volumes for registration to tracked US data. Tracked US is advantageous because it is real time, noninvasive, and non-ionizing, but it is also known to have inherent inaccuracies which create the need to develop a framework that is robust to various uncertainties, and can be useful in US-CT registration. Furthermore, conventional registration methods depend on accurate and absolute segmentation. Our proposed probabilistic framework addresses the segmentation-registration duality, wherein exact segmentation is not a prerequisite to achieve accurate registration.

In this paper, we develop a method for fast and automatic probabilistic bone surface (edge) detection in CT images. Various features that influence the likelihood of the surface at each spatial coordinate are combined using a simple probabilistic framework, which strikes a fair balance between a high-level understanding of features in an image and the low-level number crunching of standard image processing techniques. The algorithm evaluates different features for detecting the probability of a bone surface at each voxel, and compounds the results of these methods to yield a final, low-noise, probability map of bone surfaces in the volume. Such a probability map can then be used in conjunction with a similar map from tracked intra-operative US to achieve accurate registration. Eight sample pelvic CT scans were used to extract feature parameters and validate the final probability maps. An un-optimized fully automatic Matlab code runs in five minutes per CT volume on average, and was validated by comparison against hand-segmented gold standards. The mean probability assigned to nonzero surface points was 0.8, while nonzero non-surface points had a mean value of 0.38 indicating clear identification of surface points on average. The segmentation was also sufficiently crisp, with a full width at half maximum (FWHM) value of 1.51 voxels.

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

1 INTRODUCTION

The accuracy of Computer Assisted Orthopedic Surgery (CAOS) has shown improvement due to the development of new technologies for acquisition and analysis of patient data for complex procedures. The end goal described here is the registration of preoperative CT images to live patient anatomy in the OR. Many methods exist to register a patient to preoperative CT images [1, 3, 4, 7, 8]. The most popular registration algorithm, iterative closest point (ICP), iteratively transforms one set of spatial coordinates (points) to minimize the total distance between matching points in the second set [1]. ICP is desirable because it is monotonic and can be made globally optimal [1], but it requires a substantial point set. Such data sets can be obtained by physically sampling large areas of exposed bone, which is both invasive and time consuming. Alternative techniques can achieve the same results, but without the previously stated drawbacks. Two sufficiently large, unambiguous bone surface data sets – one from preoperative CT and the other from an intra-OR tracked US – should enable accurate registration of live patient to preoperative images using standard registration algorithms. The complexity lies within efficiently and accurately determining where the actual bone surface is located in these two image sets.

* Send correspondence to DaniloS@jhu.edu

Methods exist for using tracked US probes to sample patients and automatically process those images [6]. These methods can be used to create a 3D map of the surface, enabling localization of the patient's bone during the procedure [7, 8]. This data can then be used to register a patient to a preoperative image [2]. Currently, preoperative images, CT and MRI for example, must undergo manual or semi-automatic segmentation to create a bone surface data set suitable for registration. Accurate segmentation has become a prerequisite for accurate registration. Accurate segmentation is difficult because bone surfaces have inconsistent properties (orientation, density, shape, etc), unclear boundaries, and are often accompanied by artifacts and noise, especially in a modality like US [5, 6]. Thus a method that can either solve this problem or bypass it holds promise in image guided navigation systems. This paper builds on an automatic probabilistic segmentation procedure that can lead to a framework where accurate registration is made possible without absolutely accurate segmentation.

Localization of bone surfaces in US images has been studied [5], and an effective method exists for addressing the constraints of this modality and creating a low-noise bone surface probability map. US images are extremely prone to speckle noise, unusual artifacts and intensity inconsistencies which arise from individual machine settings. In order to obtain a segmentation that can be used with ICP, the images are filtered to boost the desired features and average out misleading artifacts. US images contain clues in the form of multiple reflections, shadow regions, etc. which help in deducing the presence of a surface. Probabilistic knowledge from such features is merged using Bayes rule, and an 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.

Localizing bone surfaces in a preoperative CT volume is currently time consuming, computationally intensive, and not fully automatic. Typically, a mesh is created following segmentation, and hence faulty segmentation could lead to inaccurate models. Factors such as noise, image striation, saturation effects (from metal in the images), and actual bone quality contribute to inaccuracies in segmentation algorithms. In addition, it is difficult to explicitly define bone edges in a CT volume. Figure 1 illustrates typical density (Hounsfield Units, HU) and gradient variety among surface points. Given such inconsistency, it is difficult for a single filtering method to capture all bone surfaces in a CT volume, and different edge detection methods achieve varying success in localizing edges: all can fail to detect some edges and can result in false positives. A probabilistic framework incorporating the results of independent methods into a coherent probability map is ideal because such a formulation minimizes the propagation of errors associated with any single method. The input from each method into the framework enhances the actual surface regions and suppresses the independent noise associated with each feature. The final image produced by such a framework is a probability map wherein each voxel has a value that denotes the likelihood that the actual bone surface lies within that voxel. As a result, a framework such as this does not output hard (binary) segmentation, but rather a segmentation that encompasses the

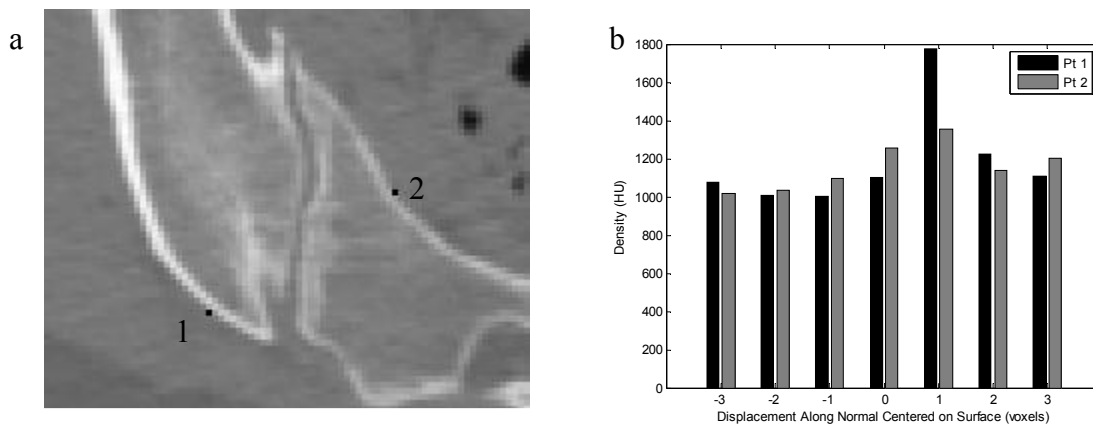


Figure 1: Bone surface points have varying properties. a) an enlarged part of a pelvic CT scan. Higher densities (HU) are shown by lighter colors. Two manually-segmented bone surface points 1 and 2 are shown in black. b) densities of voxels along a line that is perpendicular to bone surface and centered at the sample surface points. Black bars show the densities of voxels surrounding point 1, with 0 displacement showing the density of the surface point itself. Gray bars show the same information for point 2. Surface points do not have the greatest local density, and the densities themselves can vary greatly from point to point, the density of point 1 being about 150 HU less than that of point 2. Also, the relationship between surface and non-surface points is inconsistent, as evidenced by large local gradient differences for the two sample points.

actual bone edge with high probability.

Using a probabilistic framework to combine the outputs of a series of automatic bone edge detection algorithms has the benefit of needing no human interaction and produces a probability map which can be registered to a similar output from tracked intra-operative US data. We present a framework to automatically segment bone surface points from CT volumes for use in US-CT registration.

2 METHODS

Two volume-limiting and five processing methods were used to extract useful bone-surface data from the input CT volume. The volume-limiting methods systematically discard areas of the input volume which do not contain bone so that further processing may be concentrated on a reduced subset of the input. The output of each processing method was a probability volume, with the probability of bone surface recorded at each voxel. The following notational convention, similar to the one introduced in Jain et. al will be used [5]

$$P(Vol(i, j, k) = Surface) \equiv S(i, j, k) \quad \text{Eq. 1}$$

Thus, the probability that a surface is located at (i, j, k) in the given will be represented by S ; we will drop the voxel location (i, j, k) for syntactic clarity. Three of the processing methods use functions to transform raw feature values (density, gradient magnitude, energy absorption) that spanned the range [0,3000] to probabilities on the range [0,1]. These functions were calculated empirically using the actual bone surface points (known from manual segmentation of the 8 sample CT volumes) and the processed volumes as described by Equation 2:

$$\begin{aligned} P(S|F) &= \frac{P(F|S)P(S)}{P(F|S)P(S) + P(F|\bar{S})P(\bar{S})} \\ &= \frac{P(F|S)}{P(F|S) + P(F|\bar{S})} \end{aligned} \quad \text{Eq. 2}$$

In this formulation, $P(S|F)$ is the desired probability of bone surface given knowledge of a particular feature, and the simplification arises because the probability of a certain voxel being on a surface or not on a surface are equal given no other knowledge, so $P(S) = P(\bar{S}) = 0.5$. $P(F|S)$ is simply a normalized histogram of feature values for actual bone surface voxels, and $P(F|\bar{S})$ is a normalized histogram of feature values for voxels that are not bone surfaces.

The probability volumes generated by each method were compounded using Equation 3:

$$P(S|F_{1,2,\dots,i,j+1}) = \frac{P(S|F_{1,2,\dots,i})P(S|F_{i+1})}{P(S|F_{1,2,\dots,i})P(S|F_{i+1}) + (1 - P(S|F_{1,2,\dots,i}))(1 - P(S|F_{i+1}))} \quad \text{Eq. 3}$$

This equation has three nice properties: 1) combining conflicting estimates $(0.5+p)$ and $(0.5-p)$ gives 0.5 (unknown) as the result 2) combining 0.5 (unknown) with p results in p , and 3) it is biased away from 0.5 (unknown) so if p_1 and p_2 are both less than 0.5, the answer is strictly less than $\min(p_1, p_2)$, and if both p_1 and p_2 are greater than 0.5 the answer is strictly greater than $\max(p_1, p_2)$. These properties make intuitive sense in terms of how these probabilities should be compounded in order to produce maximal separation between bone surface and non-surface points.

The methods described below use information from *thresholding* and *volume-limiting* (2.1), voxel HU *density* (2.2), *gradient* (2.3), responses to *Canny* edge detection (2.4), and simulations of *energy absorption* (2.5) and *first impact* (2.6). Thresholding is a preprocessing step for all of the methods except 2.2. Volume-limiting creates a subset of data that contains all the bone, but excludes much of the soft tissue, and this limited subset is then processed by all the other

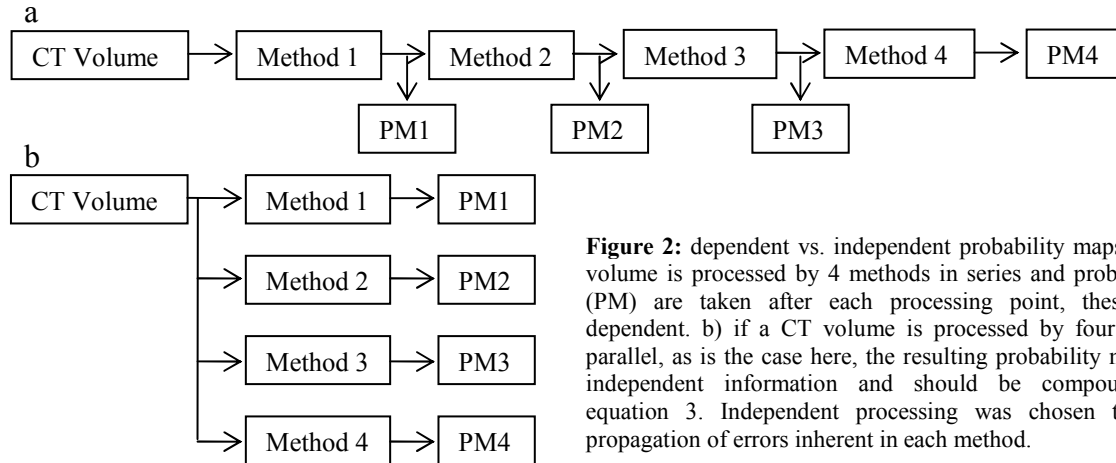


Figure 2: dependent vs. independent probability maps. a) if a CT volume is processed by 4 methods in series and probability maps (PM) are taken after each processing point, these PMs are dependent. b) if a CT volume is processed by four methods in parallel, as is the case here, the resulting probability maps contain independent information and should be compounded using equation 3. Independent processing was chosen to minimize propagation of errors inherent in each method.

methods. The gradient and energy absorption methods deduce the probability of bone surface given the values of these features, and Canny edge detection and first impact provide a more specific “template” describing the areas of highest bone surface probability. Below is more detail about how the parameters governing each method were estimated. The output of each method is an independent probability map where the values indicate the probability of a bone surface at the corresponding voxel. The probability maps are independent of each other because the input to each processing method is the same initial CT volume, as illustrated in Figure 2.

2.1 Thresholding and Volume Limiting

Two thresholding methods were used: Conservative Threshold (CTh), and Aggressive Threshold (ATh). Both of these functions have two parameters describing the minimum (Min) and maximum (Max) allowed values. Thresholding is applied so that any value outside this range is set to 0:

$$V(i, j, k) = \begin{cases} 0 & \text{if } V(i, j, k) < \text{Min or } V(i, j, k) > \text{Max} \\ V(i, j, k) & \text{if } \text{Min} \leq V(i, j, k) \leq \text{Max} \end{cases} \quad \text{Eq. 4}$$

Figure 3a summarizes the thresholding parameters relative to the number of actual bone surface points in each binned density range. The Max parameter was equal for both CTh and ATh, and was chosen as the maximum observed value of actual bone surface points ($\text{Max}_{\text{CTh}} = \text{Max}_{\text{ATh}} = 2491$ HU). The Min parameter for CTh was picked as the smallest value that consistently resulted in removal of extracorporeal noise while keeping the patient’s body outline mostly unaltered ($\text{Min}_{\text{CTh}} = 880$ HU). The Min parameter for ATh was a larger value than that for CTh, with the aim of removing as many non-surface points as possible, while maintaining virtually all bone surface points. This value was determined by plotting a threshold goodness ratio (TGR):

$$\text{TGR}(\text{Min}_{\text{ATh}}) = \frac{\% \text{Non-surface points removed using } \text{Min}_{\text{ATh}}}{\% \text{Surface points removed using } \text{Min}_{\text{ATh}}} \quad \text{Eq. 5}$$

Figure 3b shows the TGR for the 8 sample CT volumes as a function of Min_{ATh} and the vertical line indicates the actual value used ($\text{Min}_{\text{ATh}} = 1126$ HU). This value is the mean of the thresholds that followed the largest increase in the TGR for each CT volume.

Applying the ATh parameters as defined in Equation 4 removes 96% of non-surface voxels and only 10% of surface voxels on average. In some sets, a significant number of non-bone voxels are not removed and manifest as speckle-noise. To remove this noise after ATh, each slice was down-sampled by a factor of 5 in each dimension and these lower-resolution voxels were thresholded again, removing all voxels whose value was less than 15% of the maximum, but not altering the remaining voxels. This process removed speckle noise but did not eliminate a significant number of bone surface points. The final ATh method is the combination of aggressive thresholding and noise removal as described above.

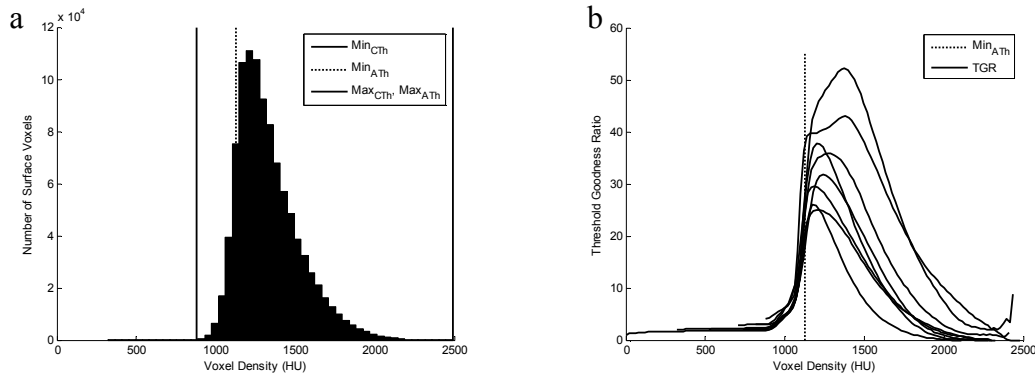


Figure 3: a) histogram of voxel densities for all actual bone surface points in 8 CT volumes. The vertical lines show the Min and Max parameters for CTh and ATh. All voxels below the Min parameter are discarded leading to negligible loss of surface points for CTh and a loss of 10% of the actual surface points for ATh. b) TGR as a function of Min_{ATh} for each CT volume. The dotted line shows the final Min_{ATh} , chosen as the mean of the densities following the greatest increase in TGR per CT volume.

CTh and ATh are useful for defining a significantly reduced CT volume subset that contains virtually all of the bone surface points. Any subsequent processing can be focused only on this subset and all voxels outside the subset can confidently be assigned a bone surface probability of 0.

Volume-limiting was performed following both CTh and ATh. To find the pertinent subset of a volume following CTh, the smallest possible rectangle encompassing the body outline was found for each slice. This was done by iteratively growing a rectangle from the middle of a slice toward the outside until each side only crossed zero-valued voxels. These rectangles contained 100% of the actual bone surface points and resulted in a volume reduction to 51% of the original volume on average. The volume-limiting procedure following ATh was similar, except that each distinct bone was placed within a rectangle. To find these smaller rectangles, a thresholded slice was downsampled by a factor of 5, groups of contiguous nonzero pixels were found, and the minimum and maximum x and y coordinates for each group were determined. These coordinates were then enlarged by 5 voxels to ensure that no bone was lost due to downsampling edge-effects, and they defined the outer bounds of a small rectangle. Volume-limiting following ATh retained 99.9% of bone surface points while reducing total volume to 15.5% of original on average across the 8 data sets. Due to this high-fidelity compression, all of the processing was done on the ATh volume-limited subsets. Table 1 summarizes the volume-limiting results per CT volume. Figure 4 shows the results of CTh, ATh, and the volume-limiting rectangles on a sample CT slice.

2.2 Feature 1: Density

The first of the processing methods uses a voxel's density in HU to calculate the probability of its being a bone surface point. The function describing the relationship between HU and bone surface probability was calculated by fitting an 11th order polynomial to empirically derived (according to equation 2) data points as shown in Figure 5a. The polynomial describes the nonzero portion of the function (between 960 and 2493 HU) and densities outside this range are assigned 0 probability. This polynomial is also scaled so that the maximum value is equal to 1 and the minimum

Table 1: Compression ratios obtained by volume-limiting the CT scans after CTh and ATh. Volume-limiting after CTh reduced the volumes to 51% of original size on average, while capturing 100% of bone surface points in each set. Volume-limiting after ATh reduced the volumes to 15.5% of original and captured 99.9% of bone surface points on average.

CT Volume	1	2	3	4	5	6	7	8	Mean 1-8
Limited/Original, CTh	0.30	0.42	0.67	0.60	0.56	0.58	0.42	0.56	0.51
Limited/Original, ATh	0.10	0.15	0.17	0.18	0.15	0.22	0.14	0.13	0.155

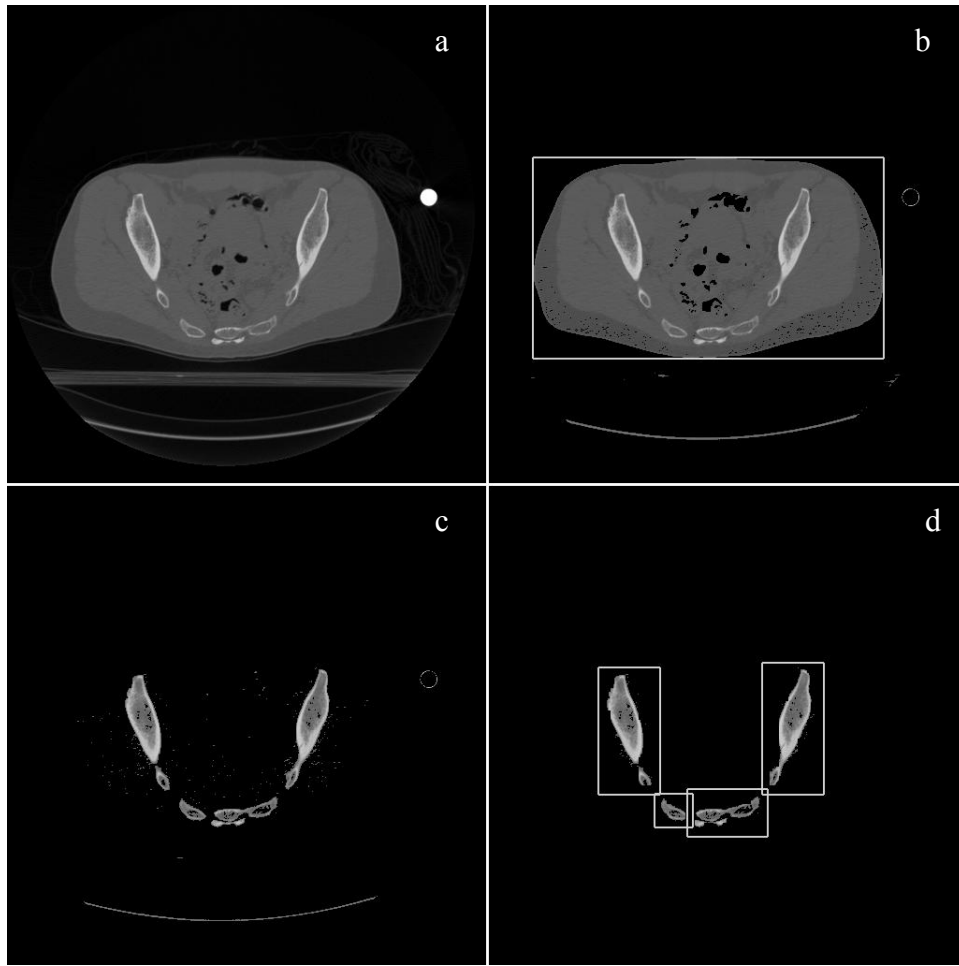


Figure 4: a) sample unprocessed slice b) original slice after CTh and volume-limiting c) original slice after ATH without noise-removal; speckle noise surrounds the bone regions d) original slice after ATH, speckle removal, and volume limiting. Overlapping rectangles were allowed to prevent eliminating bone.

value is 0, to promote clearer segregation of surface and non-surface points. Figure 3a shows the empirical data and final $P(S|Density)$ function. There is a fair amount of consistency between the 8 patients, justifying the use of constant-valued functions. Figure 6b shows a sample probability map resulting from the density feature.

2.3 Feature 2: XY Gradient

The gradient was computed for each slice independent of adjacent slices. The 2D gradient was chosen over the 3D gradient because the sampling between slices was sometimes coarse (as large as 3mm between slices) and introduced shadows of bones that were present in adjacent slices but not the one of interest. The magnitude of the XY gradient G at a location (i,j) on a slice was defined as:

$$X = \frac{V(i+1,j) - V(i-1,j)}{dist(i+1,i-1)}, Y = \frac{V(i,j+1) - V(i,j-1)}{dist(j+1,j-1)} \quad \text{Eq. 6}$$

$$G(i,j) = \sqrt{X^2 + Y^2}$$

Where $\text{dist}(\bullet)$ is the distance in millimeters between the two voxels. The processed volumes were then used to construct empirical $P(S|G)$ functions according to equation 2, and this data was fit with an 11th degree polynomial on the range [20,1134] and was set to zero outside this interval. Figure 5b shows the empirical data and $P(S|G)$ function. Again, the data between patients is consistent, supporting the idea that XY gradient is a conserved feature of bone surface points. Figure 6c shows the probability map derived from gradient information on a sample CT slice.

2.4 Feature 3: Canny Edges

The Canny edge detection algorithm (as implemented in the image processing toolbox of Matlab 6.5) was used as an additional feature. The Canny algorithm was run on an ATh processed volume and then smoothed with a moving average window of size 3. Given an image, the Canny algorithm uses information from a pixel's gradient and the values of adjoining pixels to trace the most likely connected lines around all possible edges. The algorithm's gap-filling property is desirable in the context of finding bone surface points because bones usually form closed contours within a CT slice. This closed contour feature is not explicitly implemented in any of the other methods (most of which operate on single voxels independently) but is important from a high-level consideration of the problem. As can be seen in Figure 6d, the Canny algorithm finds all the surface points, but also many extraneous edges within bones. The other methods work together to eliminate this extraneous conflicting information. The 1-voxel thick edges output by the Canny algorithm capture approximately 45% of actual bone edges on average, but after smoothing with the moving average, 96% of actual edge points are assigned nonzero values by this method. Since the output of the edge-finding algorithm is either a 0 or 1, no meaningful probability function for $P(S|Canny)$ can be estimated, so the probability volumes were scaled to span the range [0,1] after smoothing by the moving average.

2.5 Feature 4: Energy Absorption

Energy Absorption (EA) is an iterative algorithm which considers an ATh processed CT slice from the periphery toward the center. A theoretical *energy packet* travels down a row or column of voxels, and each voxel *absorbs* a fraction of the remaining energy relative to its density in HU. More specifically, a voxel with density D will have an EA of:

$$EA(i, j) = \begin{cases} 0.005(\text{remainingEnergy}) & \text{if } D(i, j) < 0.39(MaxD) \\ (D(i, j) / MaxD) & \text{if } D(i, j) \geq 0.39(MaxD) \end{cases} \quad \text{Eq. 7}$$

where $MaxD$ is close to the largest observed voxel value ($MaxD=2400$). At the start of a line of voxels, the total energy equals 3, and as it moves across voxels, the amount of energy absorbed by each voxel is subtracted from the total until there is no energy left. The given image is considered from every side (left, right, top, bottom) to attempt capturing closed contours. The EA images are then smoothed with a moving average window of size 3, as was done for the Canny processed slices.

This method tends to give the highest values for voxels that are just within the actual bone surface boundary, and it does not highlight internal edges and parts of complex shapes due to its iterative nature and limited energy per line. The EA volumes are transformed to probability volumes using an 8th degree $P(S|EA)$ polynomial on the range (0.06,0.9), calculated as for Density and Gradient features. The data points and fitted polynomial are shown in Figure 5c.

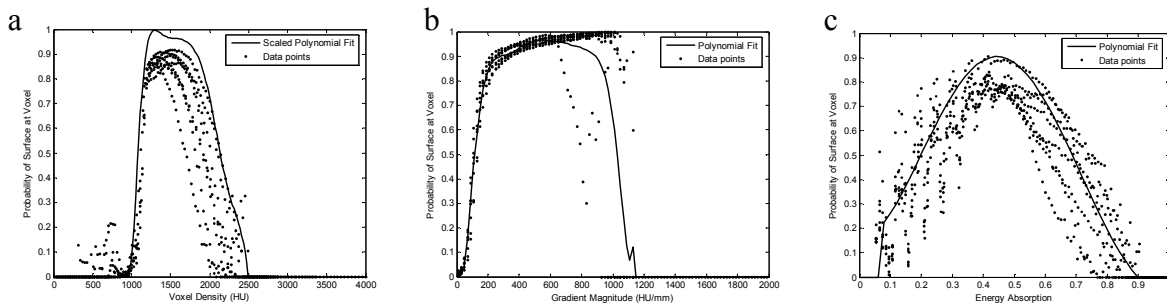


Figure 5: a) surface given density function b) surface given gradient function c) surface given energy absorption function

2.6 Feature 5: First Impact

First Impact (FI) is similar to EA, but only the first encountered nonzero voxel gets a value of 1 and all other voxels on the same scanline become 0. First impact is calculated on each slice of a volume-limited ATh processed CT volume from all 4 sides and it produces a binary volume with ones delineating the supposed bone surfaces and zeros everywhere else. After thickening this 1-voxel thick line by using a moving average of size 3 and setting all nonzero voxels to 1, this method captures about 60% of actual bone surface points on average and is generally biased toward the outside of actual bone surface contours. When compounded, the voxels in the overlapping region between EA and FI probability volumes have the highest probability of being actual surface points.

2.7 Compounding

To compute the final probability volume, Equation 3 was used to compound the probability volumes resulting from three processing methods (D, G, EA). The Canny and FI probabilities were used as templates (multiplied point-by-point instead of using Equation 3) since they did not have meaningful $P(S|F)$ functions. Following compounding, the final probabilities were multiplied by a constant so that the mean of the distribution of nonzero values was 0.5.

The initial probability volume was all 0.5-valued (unknown) and information given by individual features was compounded in the order described above. The combined effect of all the methods yields a final probability map with consistently high-valued surface points and few high-valued non-surface points, a result that captures more actual surface points and excludes more non-surface points than any single method. Figure 7 shows the progression of a sample slice through the compounding process.

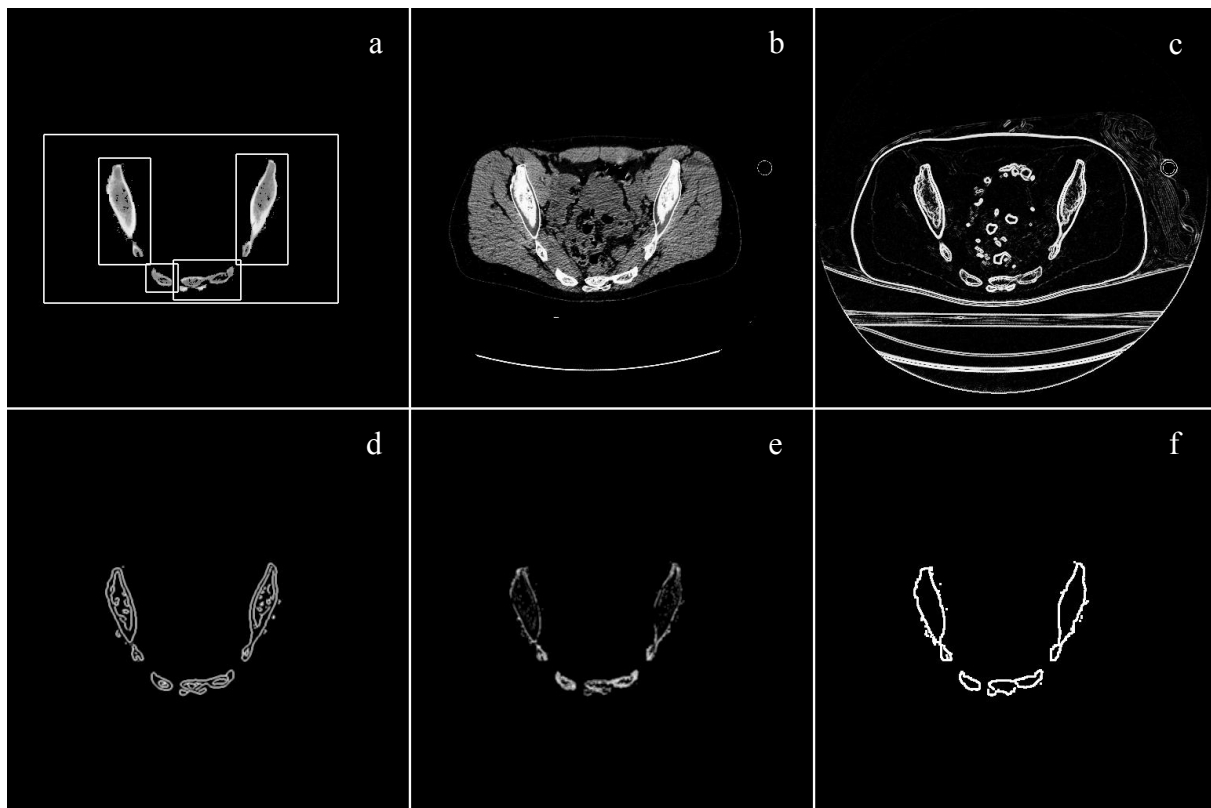


Figure 6: Summary of processing methods. a) the large rectangle shows the volume-limited area following CTh, and the four small boxes within it show the volume limited area following ATh b) sample slice showing $P(S|D)$. All bone areas have high values without significant surface/interior separation c) sample slice showing $P(S|G)$. Gradient methods find double edges and soft-tissue boundaries d) sample slice after Canny processing. Most bone surface points are found, but areas inside the bones are also labeled with high probability e) sample slice showing $P(S|EA)$. This is similar to the Canny output, but with less intra-bone noise and generally tighter contours f) sample slice showing FI. Compared with the $P(S|EA)$ result, these contours are larger, consistent with the idea that they highlight the area immediately outside of surface contours.



Figure 7: Probability map evolution as a result of compounding. a) the initial probability map assumes no knowledge about the location of surfaces b) probability map after compounding with $P(S|D)$. Only ATh volume-limited areas were processed, effectively removing noise from soft-tissues. c) probability map after adding information from $P(S|G)$. At this point, the majority of noise is localized inside the bones. d) probability map after compounding with *Canny* results. The outside edges of the bone are intensified and some of the internal noise is removed e) probability map after compounding with $P(S|EA)$. This method reduces line thickness and internal-bone noise f) final probability map after compounding with *FI*. The surface boundaries remain thin although somewhat smeared, and all internal-bone noise has been removed.

3 RESULTS AND DISCUSSION

Data was gathered and validated using eight female CT volumes, which were manually segmented using Analyze 6.1 software (Mayo Clinic). Some features (density, gradient, energy absorption) required the use of empirical data in order to construct $P(S|Feature)$ functions. This data was collected from the segmented volumes as described in the methods. All eight CT volumes were used for empirical data-gathering, and the algorithm was validated by processing the same eight volumes and comparing the results to the manually-segmented gold standards. Enough consistency was observed between individual volumes; therefore, the results are not expected to be vastly different if a completely novel CT volume is processed using parameters determined from the eight available volumes.

FAPBED was implemented in Matlab 7.01, and run on a 1.8GHz Pentium M processor with 1GB RAM. The unoptimized code takes between 2.9 and 7.8 minutes (mean 5.0 min) to process a single volume. The algorithm was executed on each of the 8 available CT volumes and statistics are shown in Table 2. Figure 7 shows the results of compounding the methods sequentially on a sample CT slice. Lighter color represents higher probability of edge. On average, FAPBED assigned nonzero probabilities to 59% of actual surface voxels, while 92% of actual surface points were within 5 voxels of a $P(S)>0.5$ value. The algorithm does not find internal bone surface points, surface points which are occluded by another surface point along a vertical or horizontal row, and surface points between two close bones as illustrated in Figure 8. This is not a serious drawback in the context of US-CT registration, since these types of surfaces cannot be imaged by US.

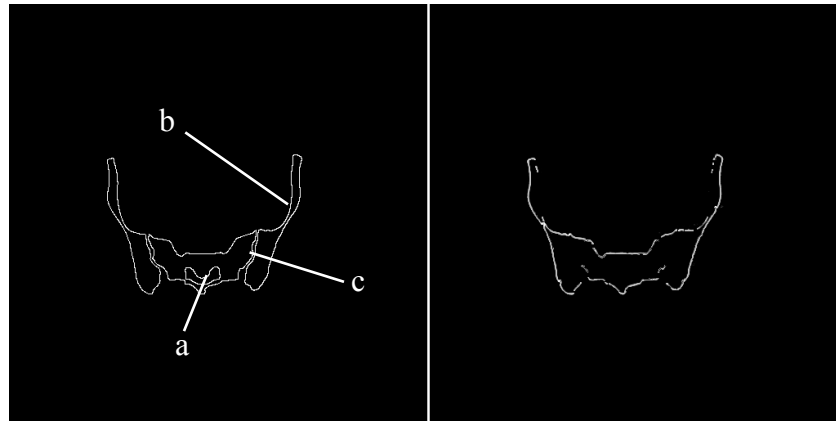


Figure 8: The image on the left is the manually-segmented gold standard. The FAPBED-generated probability map on the right fails to capture a) internal bone edge points b) bone surfaces occluded by a surface along vertical and horizontal lines and c) bone surfaces between two closely-spaced bones. These omissions are not expected to interfere with US-CT registration because they cannot be imaged by US, so their presence or absence in the FAPBED-processed image is immaterial.

To evaluate the noise present in the processed volume, the number of nonzero-valued voxels that are not actual surface points were found and expressed relative to the number of actual surface points in the volume. Dividing this value by the ratio discussed above gives an average number of non-surface points associated with a found surface point. This value was slightly over 3 on average meaning that a typical surface point is accompanied by about 3 non-surface points in the processed volume. This value varied from 1.69 for the best set to 6.3 for the worst. Performance was far below average for dataset 6 because the soft tissues surrounding bones were more dense than average, so ATh, a precursor method to 3 of the 5 processing methods, did not perform as usual and generally failed to produce clean separation of bone from tissue. This caused the EA and FI methods to assign high probabilities to voxels far from the actual bones.

The above analysis is an underestimate of the final accuracy, since all nonzero probabilities are considered as equal indications of a bone surface. Calculating the mean values of the actual bone surface points versus the mean values of the non-surface points, a clearer classification is found, with bone surface points having a mean of around 0.8 while non-surface points had a mean of about 0.38. On average, actual surface points had $P(S)$ values equal to more than twice those of non-surface points.

Table 2: Statistics concerning noise and surface point retention in FAPBED processed volumes. Column 2 shows the ratio (#nonzero surface points)/(total # surface points). Column 3 shows the ratio (#nonzero non-surface points)/(total # surface points). Column 4 is the ratio of values in columns 3 and 2, and is a measure of the average number of non-surface points per actual surface point found. Column 5 contains the mean $P(S)$ of all nonzero surface points. Column 6 has the mean $P(S)$ of all nonzero non-surface points. Column 7 is the ratio of columns 4 and 5, and measures the mean separation between surface points and non-surface points.

<i>CT volume</i>	<i>Nonzero / total surface points</i>	<i>Non-surface / total surface points</i>	<i>Mean # nonsurface points per surface point</i>	<i>Mean value of surface points</i>	<i>Mean value of non-surface points</i>	<i>Ratio of mean values for surface vs. non-surface points</i>
1	0.7025	1.1883	1.6916	0.6789	0.3936	1.7249
2	0.6255	1.2384	1.9799	0.7460	0.3743	1.9928
3	0.5960	1.4934	2.5059	0.7229	0.4099	1.7636
4	0.5848	1.4675	2.5095	0.7994	0.3780	2.1146
5	0.5787	1.5374	2.6568	0.7757	0.3938	1.9696
6	0.4260	2.6878	6.3096	0.9384	0.3406	2.7551
7	0.5847	2.2465	3.8420	0.8785	0.3773	2.3282
8	0.6225	1.7343	2.7860	0.8402	0.3726	2.2549
Mean 1-8	0.5901	1.6992	3.0352	0.7975	0.3800	2.1130

Table 3 shows more detailed information regarding average distance between actual bone surface points and the closest nonzero or *bone surface* points (*bone surface* defined as $P(S) > 0.5$). Three subsets of all segmented surface points were analyzed: all surface points, surface points within 5 voxels of a nonzero processed point, and surface points that have a nonzero processed value themselves. On average, an actual surface point was within 0.89 voxels of a nonzero FAPBED processed point. This value is an underestimate because surface points that were surrounded by 5 zero-valued voxels in each direction were excluded from the analysis. Approximately 8% of actual surface points fell into this category. The average distance to the nearest *bone surface* voxel was 1.00, with about 98% of actual surface points being within 5 voxels of a bone point. For the subset of surface points that were within 5 voxels of a nonzero point, an average 99.4% were also in the neighborhood of a *bone surface* point. Finally, for the population of actual surface points that were assigned nonzero $P(S)$ values, the average distance to a *bone surface* point was 0.08 voxels, and all of these points were in the 5-voxel neighborhood of *bone surfaces*. Full width at half maximum values were also calculated for areas of adjacent voxels centered at and perpendicular to the actual surface for this set of points. This value, 1.51 voxels on average, implies very sharply defined bone surfaces as can be seen in Figure 9d.

The most meaningful validation for registration purposes is computed by looking in the neighborhood of actual surface points, since accurate local registration involves relatively small rotations and translations. To be useful for registration, the probabilities should be highest at the location of the actual bone edges and should significantly drop in the immediate vicinity. Such functions were calculated by finding lines that are perpendicular to the bone surface and pass

Table 3: Detailed information about three subsets of all actual surface points. The first four data columns show information constructed from all of the segmented data points. Data columns 5 and 6 contain values for the subset of surface points that were at most 5 voxels away from a nonzero value in the processed image. The last three columns contain data for the subset of surface points that were assigned a nonzero value in the processed volume. Three types of data are shown: distances, probabilities, and full width half max (FWHM) values. The distance columns show the mean and standard deviation for distances between each surface point within the named subset and the nearest nonzero or *bone* point. Bone points were defined as processed voxels with a $P(S) > 0.5$. The probability columns show the proportion of the corresponding surface point subset that have the listed characteristic (proximity to bone or nonzero voxels). FWHM was only calculated for the third surface point subset, and it quantifies the average sharpness of $P(S)$ along bone surface normals centered at the surface (as in Figure 9). Perfect segmentation would yield a FWHM value of 1.

Volume	All surface points				Surface points with a nonzero processed value within 5 voxels		Surface points that have a nonzero value themselves		
	Distance to nearest nonzero voxel	Probability of nonzero point within 5 voxels	Distance to nearest bone point	Probability of bone point within 5 voxels	Distance to nearest bone point	Probability of bone point within 5 voxels	Distance to nearest bone point	Probability of bone point within 5 voxels	Full Width Half Max
1	0.70 ± 1.49	0.96	0.81 ± 1.54	0.95	0.81 ± 1.54	0.994	0.10 ± 0.30	1.0	1.62 ± 0.60
2	0.80 ± 1.51	0.93	0.91 ± 1.59	0.92	0.91 ± 1.59	0.995	0.08 ± 0.28	1.0	1.45 ± 0.55
3	0.82 ± 1.51	0.92	0.93 ± 1.56	0.91	0.93 ± 1.56	0.995	0.10 ± 0.30	1.0	1.43 ± 0.56
4	0.86 ± 1.59	0.90	0.96 ± 1.64	0.89	0.96 ± 1.64	0.994	0.09 ± 0.30	1.0	1.40 ± 0.55
5	0.94 ± 1.65	0.92	1.04 ± 1.70	0.91	1.04 ± 1.70	0.996	0.09 ± 0.28	1.0	1.44 ± 0.58
6	1.35 ± 1.76	0.91	1.44 ± 1.83	0.90	1.44 ± 1.83	0.991	0.06 ± 0.24	1.0	1.52 ± 0.59
7	0.87 ± 1.52	0.93	0.97 ± 1.59	0.92	0.97 ± 1.59	0.994	0.08 ± 0.28	1.0	1.66 ± 0.65
8	0.82 ± 1.56	0.93	0.93 ± 1.62	0.93	0.93 ± 1.62	0.995	0.08 ± 0.28	1.0	1.59 ± 0.63
Mean	0.89 ± 1.58	0.92	1.00 ± 1.63	0.92	1.00 ± 1.63	0.994	0.08 ± 0.28	1.0	1.51 ± 0.59

through each segmented surface point. Looking at the unprocessed and processed volumes along these normals gives the graphs in Figure 9. The unprocessed volumes do not show a clear difference between bone surface and soft tissue/bone interior. This difference becomes more clear after FAPBED processing (Figure 9b), with a significant decrease in the probability of bone surface points for displacements of even 1 voxel from the actual bone surface (the 'displacement along normal' plots are similar for individual CT volumes, only mean data across volumes is shown here). The graphs in Figure 9b-d show the data for the three actual surface point subsets defined above. Panel b) shows an average trend toward higher values at surface points and lower values for the surrounding voxels, but the variation is also significant (due to the inclusion of zero-valued surface points), implying that it would be hard to produce an accurate absolute segmentation. Panel c) shows the results for the subset of actual points that are within 5 voxels of nonzero processed points. This result is similar to b) because about 92% of points are common to both sets. Panel d) shows the clearest surface versus non-surface point segregation. Standard deviation is significantly smaller, and actual surface points are well contained above $P(S)=0.5$ and non-surface points in the vicinity are mostly below this value.

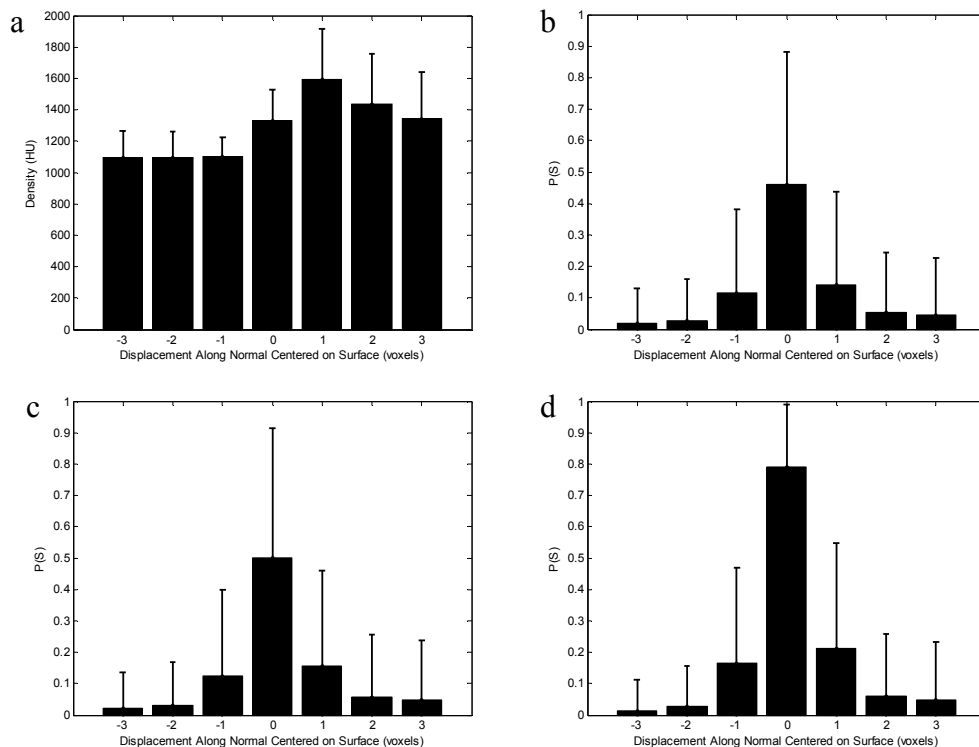


Figure 9: Mean values along lines centered at actual surface points and perpendicular to bone surface for the unprocessed CT volume and three populations of surface points in the processed CT volumes; error bars show standard deviation. a) mean density along bone normals for unprocessed CT image. The actual bone edge (displacement of 0) is generally adjacent to the highest average density, and there is little difference between bone surface and surrounding tissue. b) $P(S)$ values along bone surface normals for the processed volumes, including all actual bone surface points. The standard deviations are high and the mean $P(S)$ at 0 displacement is lower than expected because this set also includes edges that were assigned 0 probability, as described in Figure 8. The peak of the distribution is at 0 displacement however, indicating that on average, surface points are assigned higher probabilities than surrounding tissues. c) $P(S)$ values along bone surface normals for the processed volumes including only those actual surface points that had a nonzero processed point within 5 voxels. The probabilities are slightly larger than in b) because of the exclusion of some zero-values, but the results are generally the same as in b) because c) contains about 92% of the points in b). d) $P(S)$ values along bone surface normals for the subset of actual edge points that had a nonzero value assigned to them. Here, the standard deviations are significantly smaller than in populations b) and c), and there is a larger difference between actual edge (0 displacement) and surrounding tissue (absolute displacement >0).

4 CONCLUSIONS AND FUTURE WORK

FAPBED provides a quick, reliable and fully automatic means for segmenting CT volumes for use in US-CT registration for CAOS. The algorithm uses empirically-gathered data from 8 female CT volumes to extract bone surface features and then converts these values to probabilities of bone surface. Individual probability maps are compounded using a simple probabilistic framework (Equation 3). With an average runtime of five minutes, the algorithm assigns high probabilities to actual surface points and low probabilities to non-surface points in the immediate vicinity. Internal and otherwise occluded bone surface points are not captured by the algorithm, but these points cannot be imaged by US, and are not used for US-CT registration. Registration of a processed, tracked-US dataset to a FAPBED processed CT volume is expected to yield good results if implemented as a gray-level maximization [2] or ICP [1] for small rotations and translations. Below average performance was observed for one of the eight sample CT volumes because the soft tissue density of this volume was unusually high.

Desirable improvements to FAPBED are: generalization to all CT volumes regardless of patient gender (the current FAPBED parameters are derived from female CT volumes and are expected to perform differently on male CT volume inputs), generation of algorithm parameters in real-time (based on dynamic properties of the input volume itself) as opposed to relying on consistency between patients, and improvements to detect more of the actual surface points (including occluded and internal surfaces). With such improvements, FAPBED could reliably and automatically be used as a precursor to automatic 3D mesh generating algorithms.

5 ACKNOWLEDGEMENTS

The authors gratefully acknowledge the support and help offered by Mr. Ofri Sadowski (JHU CISST ERC) for obtaining CT volumes and helpful comments and advice. We also thank Dr. Branislav Jaramaz (Carnegie-Mellon University) for supplying the CT volumes. This work is supported in part by the National Science Foundation under grant EEC9731478.

6 REFERENCES

- [1] P. J. Besl and N. D. McKay, "A method for registration of 3-d shapes," IEEE Trans. Pat. Anal. and Mach. Intel. 14(2), 239-256, Feb 1992
- [2] B. Brendel, S. Winter, A. Rick, M. Stockheim and H. Ermert, "Registration of 3D CT and Ultrasound Datasets of the Spike using Bone Structures," Computer Aided Surgery, 7, 146-155, 2002.
- [3] E. A. Firle, S. Wesarg, G. Karengelis and C. Dold, "Validation of 3D Ultrasound – CT Registration of Prostate Images," SPIE, San Diego, Feb 2003.
- [4] 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.
- [5] A. K. Jain and R. H. Taylor, "Understanding Bone Responses in B-mode Ultrasound Images and Automatic Bone Surface Extraction Using a Bayesian Probabilistic Framework," Proceedings of SPIE, 5373, 131-142, April 2004.
- [6] 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.
- [7] 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.
- [8] Y. Zhang, R. Rohling and D. K. Pai, "Direct Surface Extraction from 3D Freehand Ultrasound Images," IEEE Visualization, 45-52, Boston, Oct 2002.