

An Information Fusion Method for the Automatic Delineation of the Bone-Soft Tissues Interface in Ultrasound Images

Vincent Daanen¹ Ph.D, Jerome Tonetti² M.D, Ph.D, and Jocelyne Troccaz¹Ph.D

¹ Joseph Fourier University,
TIMC Laboratory, GMCAO Department
Institut d'Ingénierie de l'Information de Santé (IN3S)
Faculty of Medecine - 38706 La Tronche cedex - France
² University Hospital
Orthopaedic Surgery Department
CHU A Michallon, BP217
38043 Grenoble

Abstract. We present a new method for delineating the osseous interface in ultrasound images. Automatic segmentation of the bone-soft tissues interface is achieved by mimicking the reasoning of the expert in charge of the manual segmentation. Information are modeled and fused by the use of fuzzy logic and the accurate delineation is then performed by using general a priori knowledge about osseous interface and ultrasound imaging physics. Results of the automatic segmentation are compared with the manual segmentation of an expert.

1 Introduction

In computer-aided orthopedic surgery (CAOS), the knowledge of the bone volume position and geometry in the operative room is essential. The usual way to acquire it is to register pre-operative data, which were accurately acquired, to intra-operative data. A recent way for acquire intra-operative data consists in the use of ultrasound imaging as intra-operative imaging [1, 2, 3] because this imaging modality is inexpensive, riskless and using a 6D localized ultrasound probe makes it possible to reconstruct the 3D shape of a structure after its delineation. The accurate delineation of structures in ultrasound images is still a very difficult task because of their very poor quality, i.e. low contrast, low signal-to-noise ration and speckle noise. Therefore, extraction of features can be :

- manual but manual segmentation is known to be operator-dependent and ultrasound images segmentation is difficult. This gives rise to uncertainty and inaccuracy of the segmentation and errors may reflect on registration results.
- semi-automatic : several methods based on the successive use of usual image processing methods such as contrast-enhancement, smoothing, edge-detection, morphological operations [4, 5] have been proposed but such methods often fail because of the poor quality of ultrasound images and therefore an user-interaction is required either to initialize the process or make the final choice.

- automatic: automatic methods are either based on active contours for the delineation of endocardial and epicardial boundaries of the heart ([6, 7]) or bone segmentation ([8, 9]) ; or are based on the multimodal registration of ultrasound datasets within more discriminant modalities such as CT [10] or MRI [11]. Compared with bone segmentation in ultrasound images,[2, 3, 12] belong to this last category of method. Moreover, in regards to orthopaedic CAS, the methods proposed in the literature are specific to a part of the human body : vertebrae [3], pelvis [12].

We propose a fully automated method designed for the delineation of the bone-soft tissues interface in ultrasound images based on information fusion. Data available in images are modeled and fused relatively to knowledge about the physics of ultrasound imaging. Expert's reasoning process is then mimicked in order to accurately delineate the osseous interface.

2 Material

Ultrasound imaging is achieved using a linear US probe, 25 mm large, working at a frequency of 7.5 MHz. The probe is localized in 3D space by an optical localizer. The US probe is calibrated according to the technique described in [13] (the pixel size is about 0.1mm/pixel). The position of an image pixel is known in 3D space with a precision in the range of the optical localizer (i.e. 1mm) (see [14] for details on the reconstruction of 3D shape from localized ultrasound images). Image size is 640×480 pixels and a subimage (214×422) is extracted before processing.

3 Method

In this section, we introduce the expert's reasoning and the way we mimic it in order to achieve the segmentation of the osseous interface.

3.1 Expert's Reasoning

Several information, based on the physics of the ultrasound imaging and on anatomy, can be used to delineate the osseous interface in ultrasound images :

1. bones appear to be hyper-echoic,
the amplitudes of the US echoes are proportional to the difference between acoustical impedances caused by successive tissue layers and in the case of bone imaging, the great difference of acoustical impedance between bones ($Z_{bone} \approx [3.65 - 7.09] \times 10^6 \text{ kg/m}^2/\text{s}$) and the surrounding soft tissues ($Z_{soft\ tissues} \approx 1.63 \times 10^6 \text{ kg/m}^2/\text{s}$) generates an important echo.
2. bones are said to 'stop' ultrasound waves,
this is due to the high absorption rate of bones which is about 10 dB/cm/MHz whereas the absorption rate for soft tissues is less than 2 dB/cm/MHz.
3. the reflection is almost completely specular,
only interfaces perpendicular to the direction of the ultrasound beam will reflect so features of interest appear to be composed of horizontal (or near horizontal) parts.

4. bone surfaces do not present major discontinuities
and therefore the found osseous interface should be as smooth as possible.
5. among an osseous interface, the contrast appears to be homogeneous.

The proposed method achieves the fusion of these information in 3 stages :

- first, an image processing step aims at modeling the information available in the images and then concentrate them into one image representing the membership of the pixel to a given property. This step models and fuses points 1,2 and 3 cited above.
- Then the computation of the continuousness cost function which purpose is to extract continuous osseous interfaces from the fuzzy image.
- Finally, we compute the optimal osseous interface from the candidates found at the previous step by choosing the one that ensures an osseous interface among which the contrast is maximum and homogeneous.

3.2 Image Processing Step

The image processing aims, in a first step, at modeling the information available in the image. Then the data fusion step concentrates the information in order to produce an image whom the value of a pixel represents the membership of the pixel to a property.

Fuzzy Intensity Image This stage attempts to model the first information listed above : bones appear to be hyper-echoic i.e. *bright* pixels constitute an indication of the location of the osseous interface, but is not an absolute criteria and consequently, the fuzzification function have to give an important (resp. low) membership value to *bright* (resp. *dark*) pixels.

In a previous development [15], we pointed out that binarizing the initial ultrasound image using the Otsu's threshold (T_{Otsu}) gives a good approximation of the echogenic area and so, of the position of the osseous interface.

We make use of this information to build the fuzzification function μ_{Int} : the criterion (we call V_{Otsu} : Fig.1-b, *solid curve*), needed to compute T_{Otsu} , is used as follows : first, V_{Otsu} is normalized and cumulated (Fig.1-b, *dotted curve*) and it is then shifted in order to force the membership function value : $\mu_{Int}(T_{Otsu}) = 0.5$.

Processing this way, the fuzzification function (Fig.1-b, *dashed curve*) is closed to the well-known S-function [16]. Finally, we apply the described fuzzification function over the gray-level image in order to construct the fuzzy intensity image¹ $FII(p)$ (Fig 1-c) which gives for a pixel p of the intensity image its membership degree to the echogenic area.

Fuzzy Gradient Image The transition from soft tissues to bone suggests to search for highly contrasted areas and so the fuzzy gradient image $FGI(p)$ is of great interest. The computation of the gradient image is a way to model informations 2 and 3 because it allows us :

¹Illustrating images have been cropped to half their length since the other half is a dark area and brings no information (the whole image is processed)

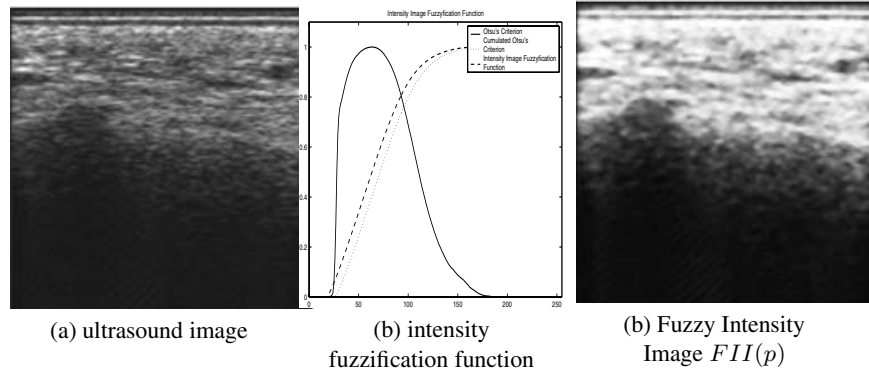


Fig. 1. Computation of Fuzzy Intensity Image (FII) from the original ultrasound image

- to detect features which have a particular direction by using a directional edge detector,
- and to get an information about the gray-level transition.

We use a 5x5 'horizontal-direction' MDIF edge detector which is first-order derivative filter obtained by the convolution of the 4-connexity 3x3 mean lowpass filter with the Prewitt's derivative kernels. It allows us to select areas presenting horizontal (or near) contrast. Thresholding this image allows us to keep only areas where the transition occurs from bright-to-dark pixels (Fig. 2-a).

Finally, we use the S-shape function to perform the *fuzzification* of the gradient image and obtain the *Fuzzy Image Gradient $FGI(p)$* (Fig. 2-b). The parameters of the S-shape function are computed such that $S(x)$ is the closest s-shape function to the normalized cumulative histogram.

Data Fusion The data fusion step aims at concentrating all the information in order to produce a single membership value for each pixel of the analyzed image to the osseous interface. For our purpose, a pixel may belong to the osseous interface if both its gray-level and gradient are 'high'. This combination is naturally achieved by a 'conjunctive-type' combination operator *min* ; therefore the membership of a pixel to the osseous interface is given by :

$$FI(p) = \min(FII(p), FGI(p)) \tag{1}$$

$FI(p)$ denotes the global degree of membership of a pixel to an echogenic and highly contrasted area.

3.3 Determination of the Osseous Interface

According to the expert's reasoning, the optimal threshold described a continuous interface where the local contrast is maximum and homogeneous. For each membership

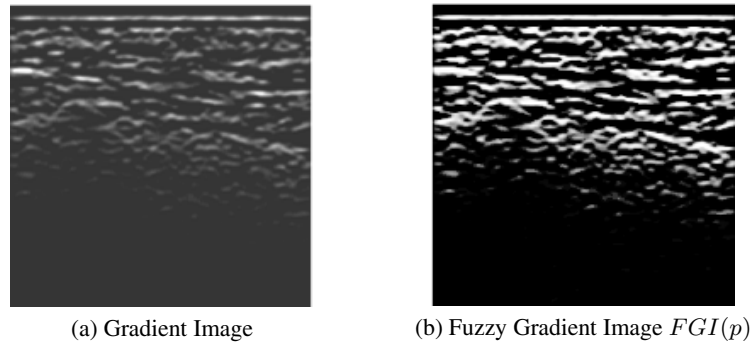


Fig. 2. Gradient image and Fuzzy Gradient Image of the ultrasound image shown in Fig 1-a

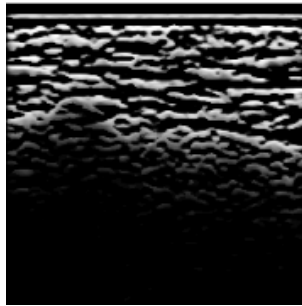


Fig. 3. Fusion Image

degree $0 < \mu < 1$ (μ space is discretized with a step $\delta_\mu = 0.005$), the defuzzification of $FI(p)$ is performed and the continuousness of the profile is evaluated. We then choose the membership degree which maximizes the local contrast and its homogeneity, and also ensures a local maximum continuity of the profile.

In this section, we present the defuzzification process which aims at extracting from the fuzzy image $FI(p)$ the osseous interface related to a membership degree μ_{ref} . We then explain the way we measure the continuity of a given osseous interface and finally we explain the way we compute an objective function which will allow us to determine the optimal membership.

Defuzzification Process To achieve this task, we make use of a priori knowledge about the physics of ultrasound imaging : as mentioned earlier, ultrasound imaging enhances the difference between acoustical impedances and because of the great difference of the acoustic impedances of the bone and its surrounding soft tissues, almost the entire ultrasound wave is reflected at the surface of the bone so that no imaging is possible

beyond it. Dealing with that in image processing, this means that, for a column of the image, the pixel of the osseous interface related to a membership value μ_{ref} is the last (from the top) pixel which has a membership equal or greater to μ_{ref} . At the end of this *defuzzification* process, at the most one pixel by column is highlighted. The 'curve' described by these pixels is called *profile* in the rest of the paper.

Evaluation of the Continuousness of the Profile As we mentioned earlier, actual osseous interfaces do not present discontinuities and therefore, the osseous interface we detect should be as smooth as possible. We use this property to determine the optimal defuzzification threshold by computing a function that reflects the continuousness of a computed osseous interface.

The measure of the continuousness of a profile is achieved by applying the wavelet transform to it : the wavelet transform decomposes the profile with a multiresolution scale factor of two providing one low-resolution approximation (A_1) and one wavelet detail (D_1). We then apply the wavelet transform to A_1 and get a second order low-resolution approximation (A_2) and wavelet detail (D_2). The *Detail* signals are then used to quantify the discontinuities of the original profile. Experimentally, we choose the Daubechies-4 wavelet basis (several others basis have been tested and no dependence was pointed out at the exception of the Haar Basis). The wavelet decomposition of the profile is performed twice. To reject small interfaces (4/5 pixels) detected when the membership value used for the defuzzification is unsuitable (i.e. when it is too high), we add a penalization term related to the length of the profile (Pen). The 'amount' of discontinuities in the profile is computed as follows :

$$\varepsilon(\mu) = E(D_1) + E(D_2) + Pen \tag{2}$$

where $E(s)$ represents the energy of a signal $s(t)$ and is computed by :

$$E(s(t)) = \frac{1}{n} \sum_{i=0}^n s(t)^2 \tag{3}$$

Finally, $\varepsilon(\mu)$ is normalized (i.e. $\varepsilon(\mu)$ is linearly scaled from $[\varepsilon_{min} - \varepsilon_{max}]$ to [0-1]) and we compute the continuousness of the profile as :

$$C(\mu) = 1 - \varepsilon(\mu) \tag{4}$$

As one can see (Fig.4-a), the continuousness function $C(\mu)$ presents several local maxima. Each of them locates a membership degree μ where the associated profile is more continuous than the profiles of its neighbors and so each of them may be the optimal defuzzification threshold. We detect them by computing the watershed transform of $C(\mu)$. For each local maxima, the image is defuzzed to the corresponding membership degree μ and the local contrast is computed.

Local Contrast Computation For each pixel p belonging to a profile (i.e. the pixel p is highlighted after the defuzzification process related to a given μ), the local contrast

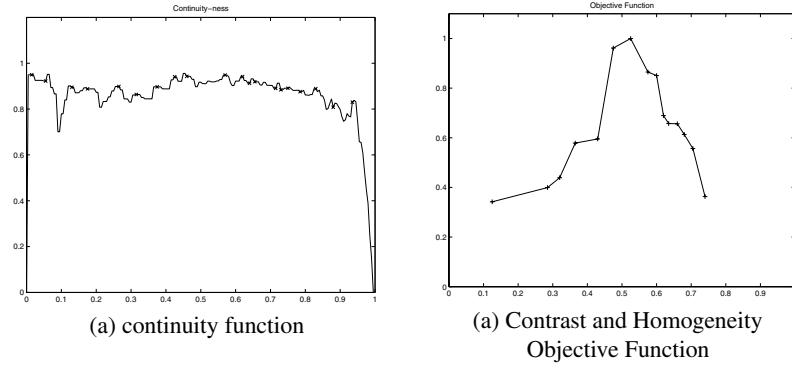


Fig. 4. Objective Functions

$LC(p)$ is computed between the above and the underneath areas of 10 pixels long. The local contrast associated to the pixel p is computed by :

$$LC(p) = \frac{\overline{Up} - \overline{Down}}{\overline{Up} + \overline{Down}} \quad (5)$$

where \overline{Up} (resp. \overline{Down}) is the mean value of the above (resp. underneath) area. This definition of the local contrast gives us a way to determine whenever the pixel p is in the vicinity of the osseous interface because the bone appears in the image as a 'light' area followed by a 'dark' area i.e. a positive local contrast. We then obtain a cost function $Contrast(\mu)$ related to the global contrast along the profile and defined by :

$$Contrast(\mu) = \sum_p LC(p) \quad (6)$$

Because the contrast along the profile is homogeneous, we also compute a measure of the homogeneity of the contrast along the profile. This is achieved by computing the standard deviation of the values of the contrast along the profile (this criterion was used as homogeneity measure by [17]) and gives us a function $StdDev(\mu)$

Optimal Defuzzification Threshold Determination The optimal membership degree $\mu_{Optimal}$ is chosen so that it maximized $Cost(\mu)$ (Fig.4-b) defined by :

$$Cost(\mu) = Contrast(\mu) + \frac{1}{StdDev(\mu)} \quad (7)$$

4 Results

The method was originally designed for sacral osseous interface delineation. Because the method tries to reproduce the expert's reasoning and because the a priori information we make use are not specific to a part of the human body, the method appears to

be usable on ultrasound images of different bones which are in relationship with various current research projects in CAOS. This section first presents the results of sacral images segmentation and then shows some results of vertebrae images segmentation.

4.1 Sacral Images Segmentation

The proposed method has been tested on ultrasound images of sacrum coming from cadaver datasets or patient datasets : about 250 images have been processed. For each image, the manual segmentation of the expert is available and constitutes our bronze-standard.

For each image within a dataset, we compute the differences between the manual segmentation and the segmentation computed by our method per image column. We then compute the mean error for each image (Table 1-column 1) and the Hausdorff distance and mean absolute distance (average of all the maximum errors within a subset) (Table 1-column 2). In order to evaluate the ability of the proposed method to delineate the osseous interface in strongly corrupted images, we also compute the Signal-to-MSE ratio (Table 1-column 3), which corresponds to the classical Signal-To-Noise ratio computed in the case of additive noise, and is defined as [18] :

$$S/mse = 10 * \log_{10} \left(\frac{\sum_{i=1}^K S_i^2}{\sum_{i=1}^K (\hat{S}_i - S_i)^2} \right) \tag{8}$$

where

S is the original image

\hat{S} represents the denoised image

and K is the number of pixels (i.e. 90308 for a 214×422 image).

\hat{S} is obtained by filtering S with a 5x5 median filter. As one can easily see from Equ. 8, the less S/mse ratio, the noisier the image.

Dataset	Segmentation Error	Max Errors	S/mse
	mean/SD (pixel)	mean/max (pixel)	mean/SD (dB)
Patient 1 (51 images)	7.808 / 1.995	12.137 / 22	5.052 / 0.185
Patient 2 (49 images)	8.807 / 3.177	16.905 / 25	5.206 / 0.428
Patient 3 (69 images)	4.545 / 3.874	17.0789 / 35	8.905 / 0.283
Cadaver 1 (37 images)	3.495 / 1.931	9.830 / 36	8.786 / 0.340
Cadaver 2 (41 images)	2.679 / 1.456	7.294 / 19	9.019 / 0.259
Cadaver 3 (39 images)	4.056 / 3.213	12.14 / 38	7.984 / 0.177

pixel size is $0.112mm \times 0.109mm$

Table 1. Segmentation errors

As one can see (table 1), as compared to the manual delineation of the expert :

- the mean error of segmentation is always less than 10 pixels (i.e. 1mm) even on highly corrupted images. However, it is clear that the accuracy of the delineation is correlated within the amount of noise and therefore, we think that taking into account the noise (measured by the S/sme ratio by example) during the fusion and/or delineation process may be a way to improve the delineation. This could be done by weighting the *Fuzzy Intensity Image* values by a factor depending on the noise (as described in [19]).
- The maximum error still remains substantial but, according to us, it is not the error we should focus on : we point out that these errors occur at more or less one pixel on complex shapes (such as medial sacral crest or sacral hiatus) giving thus an important maximum error relatively to the manual delineation but the overall error on the global shape still remains negligible and has very limited impact on the registration step which follows.
- The proposed method is also sufficiently fast to be used during the intra-operative stage of a CAOS : the time needed to delineate one image is less than 4 s. The processing of large datasets such as *Patient 3* takes about 4 minutes (on a standard PC, Pentium III-800Mhz) whereas it took more than 30 minutes in the case of a manual delineation (according to [20]).

4.2 Spine Image Segmentation

42 images of the L4-vertebra were acquired on a healthy volunteer and processed. In order to see a large part of the spinous process, we use a linear US-probe of 80 mm large working at a frequency of 7.5 MHz. The image size is 768×576 pixels and a subimage (346×467) is extracted before processing and one image is processed in less than 6 s.

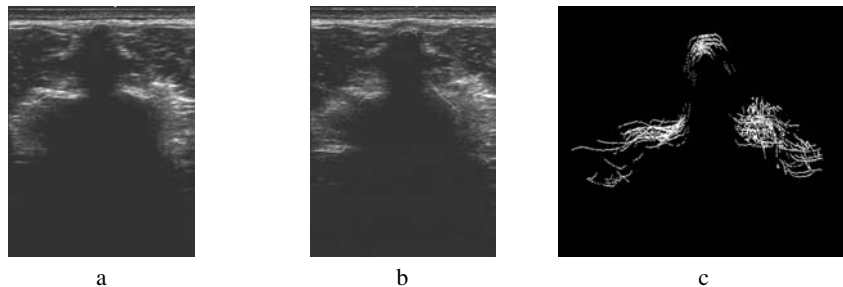


Fig. 5. vertebra image segmentation

One can clearly recognize the spinous process on Fig. 5-a and Fig.5-b. The overall shape of the vertebra is also clearly visible on Fig. 5-c which is a 3D point cloud computed from the delineation of the vertebra surface in the ultrasound images.

5 Discussion

Recently, lots of methods dedicated to the *indirect* delineation of the bone surface in ultrasound images have been proposed in the literature [3, 12, 21] but these methods have not been tested on real patients' datasets yet. Moreover, the ultrasound imaging is constrained by the use of a mechanical system [3, 21] ; and a good initial estimation of the rigid registration matrix between the CT and US datasets is often required to achieve the bone surface segmentation in the ultrasound images [2, 12].

The method described in this paper does not require neither a dedicated ultrasound images acquisition system nor an estimation of the rigid registration matrix between the CT and US datasets to perform the delineation of the osseous interface in ultrasound images. Moreover, it has been extensively tested on images acquired on cadavers (about 120 images) and on real patients (about 170 images).

Although the method is sensible to noise, the mean errors are still acceptable : we measure a maximum mean error of 8.8 pixels (i.e. 0.8 mm) with a S/mse ratio of 5.206 dB which corresponds to a highly corrupted image (according to [18]).

We think that an important point have also to be made clear : the validation, based on the comparison to a single expert segmentation, may appear limited. However, segmenting bones on ultrasound images is very unusual for physicians and it is difficult to find several expert users. Moreover, gold-standard does not exist and tests on phantoms or isolated bones would not allow to draw conclusion and we then consider that this evaluation is just a first step.

Finally, we did not notice any dependence of the accuracy to the visualization parameters tuning, i.e. the different gains (proximal, distal) the physician can play with to enhance the contrast of the image on the ultrasound scanner. The only condition is that the osseous interface should not get bogged down in noise and according to us, it is an acceptable condition since the physician has to validate the images during the acquisition stage and this validation can only be done if he is able to localize approximatively the osseous interface.

6 Conclusion

In this paper, we present a method for automatic delineation of the osseous interface in ultrasound image. The method is based on the fusion of the pixels intensity and gradient properties in a first step and then on the fusion of information extracted from the physics of ultrasound imaging and a priori knowledge.

Tests were performed with images coming from patients or cadaver studies. The method has been used to delineate osseous interface in ultrasound images of the sacrum which may present several shapes ; we also use it to delineate the osseous interface in vertebrae images and good results were obtained in all cases so that it's independent of the shape to be recovered and that's why we think that the described method is a first step toward robust delineation of the osseous interface in ultrasound images.

References

- [1] J. Tonetti, L. Carrat, S. Blendea, Ph. Merloz, J. Troccaz, S. Lavallée, and JP Chirossel. Clinical Results of Percutaneous Pelvic Surgery. Computer Assisted Surgery using Ultrasound Compared to Standard Fluoroscopy. *Computer Aided Surgery*, 6(4):204–211, 2001.
- [2] G. Ionescu, S.Lavallée, and J. Demongeot. Automated Registration of Ultrasound with CT Images : Application to Computer Assisted Prostate Radiotherapy and Orthopedics. In *CS Serie, Springer Verlag, MICCAI*, volume 1679, pages 768–777, 1999.
- [3] 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(3):146–155, 2002.
- [4] A. Krivanek and M. Sonka. Ovarian Ultrasound Image Analysis : Follicle Segmentation. *IEEE Transactions on Medical Imaging*, 17(6):935–944, 1998.
- [5] SD. Pathak, DR. Haynor, and Y. Kim. Edge-Guided Boundary Delineation in Prostate Ultrasound Images. *IEEE Transactions on Medical Imaging*, 19(12):1211–1219, 2000.
- [6] I. Mikic, S. Krucinski, and JD. Thomas. Segmentation and Tracking in Echocardiographic Sequences : Active Contours Guided by Optical Flow Estimates. *IEEE Transactions on Medical Imaging*, 17(2):274 –284, 1998.
- [7] JG. Bosch, SC. Mitchell, BPF. Lelieveldt, F. Nijland, O. Kamp, M. Sonka, and JHC. Reiber. Automatic Segmentation of Echocardiographic Sequences by Active Appearance Motion Models. *IEEE Transactions on Medical Imaging*, 21(11):1374–1383, 2002.
- [8] F. Lefebvre, G. Berger, and P. Laugier. Automatic Detection of the Boundary of the Calcaneus from Ultrasound Parametric Images using an Active Contour Model ; Clinical Assessment. *IEEE Transactions on Medical Imaging*, 17(1):45–52, 1998.
- [9] P. He and J. Zheng. Segmentation of Tibia Bone in Ultrasound Images using Active Shape Models. In *Proceedings of the 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2001.
- [10] BH Sollie. Automatic Segmentation and Registration of CT and US images of Abdominal Aortic Aneurysm using ITK. Master’s thesis, Norwegian University of Science and Technology. Faculty of Information Technology, Mathematics and Electrical Engineering, 2002.
- [11] A. Roche, X. Pennec, G. Malandain, and N. Ayache. Rigid Registration of 3-D Uultrasound with MR Images : a New Approach Combining Intensity and Gradient Information. *IEEE Transactions on Medical Imaging*, 20(10):1038–1049, 2001.
- [12] D.V. Amin, T Kanade, A.M. DiGioia III, and B. Jaramaz. Ultrasound Registration of the Bone Surface for Surgical Navigation. *Computer Aided Surgery*, (8):1–16, 2003.
- [13] T. Langø. *Ultrasound Guided Surgery : Image Processing and Navigation*. PhD thesis, Norwegian University of Science and Technology, Trondheim, Norway, 2000.
- [14] C. Barbe, J. Troccaz, B. Mazier, and S. Lavallée. Using 2.5D Echography in Computer Assisted Spine Surgery. In *Proceedings of the 15th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 160–161, 1993.
- [15] V.Daanen, J.Tonetti, J. Troccaz, and Ph. Merloz. Automatic Determination of the Bone-Soft Tissues Interface in Ultrasound Images. First Results in Iliosacral Screwing Surgery. In *Proceedings of Surgetica-CAMI 2002*, pages 144–151, 2002.
- [16] Z.Chi, H.Yan, and T.D. Pham. *Fuzzy Algorithms : with Applications to Image Processing and Pattern Recognition*, volume 10 of *Advances in Fuzzy Systems - Applications and Theory*, section 2. World Scientific, 1996.
- [17] M. Garza, P. Meer, and V. Medina. Robust Retrieval of 3D Structures from Image Stacks. *Medical Image Analysis*, 3(1):21–35, 1999.

- [18] A. Achim, A. Bezerianos, and P. Tsakalides. Novel Bayesian Multiscale Method for Speckle Removal in Medical Ultrasound Images. *IEEE Transactions on Medical Imaging*, 20(8):772–783, 2001.
- [19] S.Vial, D. Gibon, C. Vasseur, and J. Rousseau. Volume Delineation by Fusion of Fuzzy Sets Obtained from Multiplanar Tomographic Images. *IEEE Transactions on Medical Imaging*, 20(12):1362–1372, 2001.
- [20] L. Carrat, J.Tonetti, S. Lavallée, Ph. Merloz, L. Pittet, and JP. Chirosset. Treatment of Pelvic Ring Fractures : Percutaneous Computer Assisted Iliosacral Screwing. In *CS Serie, Springer Verlag, MICCAI*, volume 1496, pages 84–91, 1998.
- [21] O. Schorr and H. Wörn. A New Concept for Intraoperative Matching of 3D Ultrasound and CT. In *Proceedings of MMVR : Medecine Meets Virtual Reality*, pages 446–452, 2001.