



Chest wall segmentation in automated 3D breast ultrasound scans



Tao Tan^{a,*}, Bram Platel^b, Ritse M. Mann^a, Henkjan Huisman^a, Nico Karssemeijer^a

^a Department of Radiology, Radboud University Nijmegen Medical Centre, The Netherlands

^b Fraunhofer MEVIS, Bremen, Germany

ARTICLE INFO

Article history:

Received 27 February 2012

Received in revised form 19 November 2012

Accepted 21 November 2012

Available online 5 December 2012

Keywords:

Chest wall segmentation
Automated 3D breast ultrasound
Cylinder fitting

ABSTRACT

In this paper, we present an automatic method to segment the chest wall in automated 3D breast ultrasound images. Determining the location of the chest wall in automated 3D breast ultrasound images is necessary in computer-aided detection systems to remove automatically detected cancer candidates beyond the chest wall and it can be of great help for inter- and intra-modal image registration. We show that the visible part of the chest wall in an automated 3D breast ultrasound image can be accurately modeled by a cylinder. We fit the surface of our cylinder model to a set of automatically detected rib-surface points. The detection of the rib-surface points is done by a classifier using features representing local image intensity patterns and presence of rib shadows. Due to attenuation of the ultrasound signal, a clear shadow is visible behind the ribs. Evaluation of our segmentation method is done by computing the distance of manually annotated rib points to the surface of the automatically detected chest wall. We examined the performance on images obtained with the two most common 3D breast ultrasound devices in the market. In a dataset of 142 images, the average mean distance of the annotated points to the segmented chest wall was 5.59 ± 3.08 mm.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

Breast cancer is the most common cancer that affects women, with 421,000 new cases diagnosed in the Europe each year and 129,000 women dying from the disease (Ferlay et al., 2010). While causes remain largely unknown, incidence is still increasing in most countries. If breast cancer is detected early, mortality is decreased due to more effective treatment. Furthermore, the quality of life of these women is maintained because early detection enables less radical treatment. Screening programs have therefore been introduced to detect early breast cancers in asymptomatic women.

The sensitivity of mammographic screening is seriously impaired in women with dense breasts. This because the fibroglandular and stromal tissues (dense tissues) have the same X-ray attenuation properties as tumors and thus both show equally bright on mammographic images. For extremely dense breasts, a sensitivity as low as 30% has been reported in the US (Mandelson et al., 2000).

As a complementary modality to mammography, breast ultrasound has a sensitivity which can surpass that of mammography for patients with dense breast tissue (Berg, 2004). In a recent study, Kelly et al. (2010) showed that detection sensitivity increased significantly using 2D image sequences covering the entire breast

when compared with mammography alone in dense breasts. Regular two dimensional breast ultrasound, however, suffers from operator dependence, a limited capability to visualize the entire breast and qualified radiologists have to perform this time-consuming task themselves.

These limitations have been technically resolved with the introduction of automated 3D breast ultrasound volume scanning systems which we referred to as ABUS in this paper. The system involves frontal compression of the breast by a dedicated membrane, allowing a wide transducer to automatically move over the breast surface. In the scanning device the transducer is translated with a constant speed across the breast to obtain a 3D volume of imaging data covering a large segment of the breast. Depending on the size of the breast, up to as many as five views are acquired for each breast. Positioning and compression are to some extent standardized and include anteroposterior, lateral, medial, superior, or inferior views. One main advantage of the 3D breast ultrasound system is to view reconstructed images in the coronal plane which offer a complete new way of looking at breast ultrasound and reveal signs of abnormality that could not be observed before, such as very specific spiculation patterns frequently occurring in the presence of cancer (Rotten et al., 1999; Watermann et al., 2005).

However, the reading of these 3D breast ultrasound images is time consuming and demanding for radiologists. Computer-aided detection (CADe) is expected to play an important role in the future to facilitate the reading of these images. CADe suffers from finding many false positives outside of the breast tissue. To remove false

* Corresponding author.

E-mail address: tao.tan911@gmail.com (T. Tan).

positives of a CAdE system beyond the chest wall, determining the location of the chest wall in ABUS images is necessary.

Furthermore, the accurate localization of the chest wall can be used as a landmark for determining spatial correlation between multi-modal images (Tanner et al., 2011). The spatial correlation is expected to be of great help for fusing information and correlating CAdE marks in multi-modal images.

To our knowledge, there is only one paper previously reporting on chest wall segmentation (Huisman and Karssemeijer, 2007), where a deformable volume model is used to determine the chest wall location in ABUS images. The method was only partly validated by determining how well presence of a chest wall in a ABUS scan could be detected, without determining accuracy of the segmentation itself. The performance of the method was limited.

In this paper we describe a novel method to accurately segment the chest wall.

2. Dataset

The automated 3D breast ultrasound images used in this study are a representative sample of diagnostic cases obtained in routine clinical care. We used two datasets obtained with two different automated 3D breast ultrasound devices. Dataset A includes 80 patients that were randomly selected from a series of cases obtained using the Somo-V automated 3D breast ultrasound system developed by U-systems (Sunnyvale, CA, USA). Dataset B also includes 80 patients that were randomly selected from cases obtained with the ABUS system developed by Siemens (Erlangen, Germany).

The device by U-systems generates an image with a maximal size of 14.6 cm by 16.8 cm on the coronal plane and a maximal depth of 4.86 cm while the device by Siemens generates an image with a maximal size of 15.4 cm by 16.8 cm on the coronal plane and a maximal depth of 6 cm.

The two devices differ in two aspects. Firstly, the frequency of the transducer by Siemens is between 5.0 and 14.0 MHz, which is adjustable to the breast size, while the frequency of the transducer by U-systems is fixed at 8.0 MHz or 10.0 MHz. Secondly, the compression of the membrane on the breast is manually applied using the device by U-systems, while the compression is applied automatically by the device by Siemens using a mechanical arm. For processing, both types of images were downsampled to obtain 0.6 mm cubic voxels after applying Gaussian smoothing on the original high-resolution image with a scale of 0.3 mm.

For each patient, one single volumetric view was randomly selected. The chest wall was completely invisible in 5 of 80 views in dataset A and in 13 of 80 views in dataset B. Those views in which the chest wall was completely invisible were excluded from this study resulting in 75 views and 67 views in dataset A and dataset B, respectively. The absence of the chest wall in an image may be caused by insufficient scanning depth set by the technician, extremely poor image quality or implants obstructing the chest wall in the image.

3. Method

The chest wall is the surface enclosing the ribs and touching the outer rib surfaces. These characteristic rib surfaces are used by human readers to locate the chest wall. Therefore, we propose to locate the chest wall by automatically detecting rib surfaces and subsequently fitting a surface through them.

We use multiple features and classifiers to automatically detect points on the surface of the ribs. Subsequently a cylinder is fitted to the resulting point distribution. This cylinder models the local shape of the chest wall in the view at hand.

To test our method, we manually marked 10–30 rib surface points in the each view of dataset A and B. These points are at the transition between breast tissue and rib shadows for each scan (Fig. 1). The annotation was verified with a radiologist.

This section is split into three subsections: the first subsection explains an intensity normalization step applied before chest wall segmentation; the second subsection shows that the partial chest wall visible in a single volumetric view can be modeled by the surface of a cylinder and the third subsection explains a method to automatically detect the rib surfaces.

3.1. Intensity normalization

Images in this study were obtained from three different institutes using devices by two different manufactures. Due to variations of scan settings, the intensity levels of the same tissue type (fatty tissue, dense tissue, etc.) varied from patient to patient. Therefore, before we perform chest wall segmentation, intensity normalization was performed on all images.

In ABUS imaging, the ultrasound signal can only be acquired if a good contact between the probe and the skin exists. Therefore only voxels in columns under the area with good contact hold valid information. The valid area on the skin is shown on coronal view in the resulting image. To be able to normalize intensities inside the breast in one image, a 2D breast mask was generated in the coronal plane. The same 2D mask is suitable for all coronal slices in the volume.

To this end, a coronal projection was obtained by averaging a series of 2D coronal slices. The breast mask was obtained by applying Otsu's thresholding method (Otsu, 1979) on the coronal projection. To avoid a negative effect of slices through the skin and thoracic volume on the mask extraction, only slices with a depth between 10% and 40% of the total depth of the image volume were used in the projection.

After thresholding, in the coronal projection, the boundary of the largest foreground connected component (breast) was refined by dilation and erosion operations. Holes inside the boundary caused by lesions or shadows were filled, obtaining the breast mask.

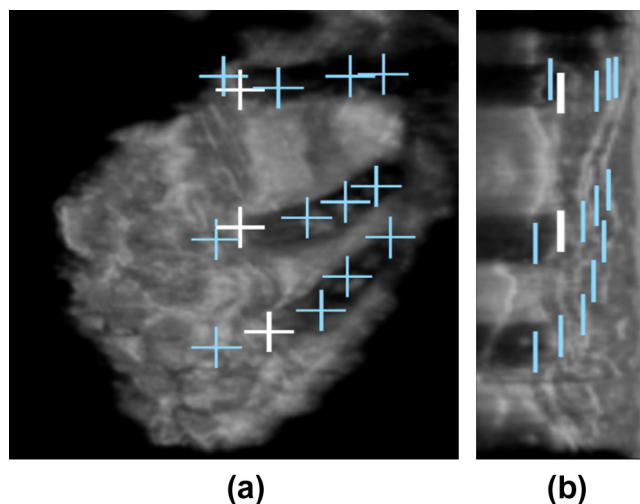


Fig. 1. Manually annotated points on the surface of the ribs on coronal view (a) and sagittal view (b). The crosses represent the manually annotated points on the rib surface, projected on the current slice. Annotated points on the current slice are represented by white crosses, and projected crosses are blue. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

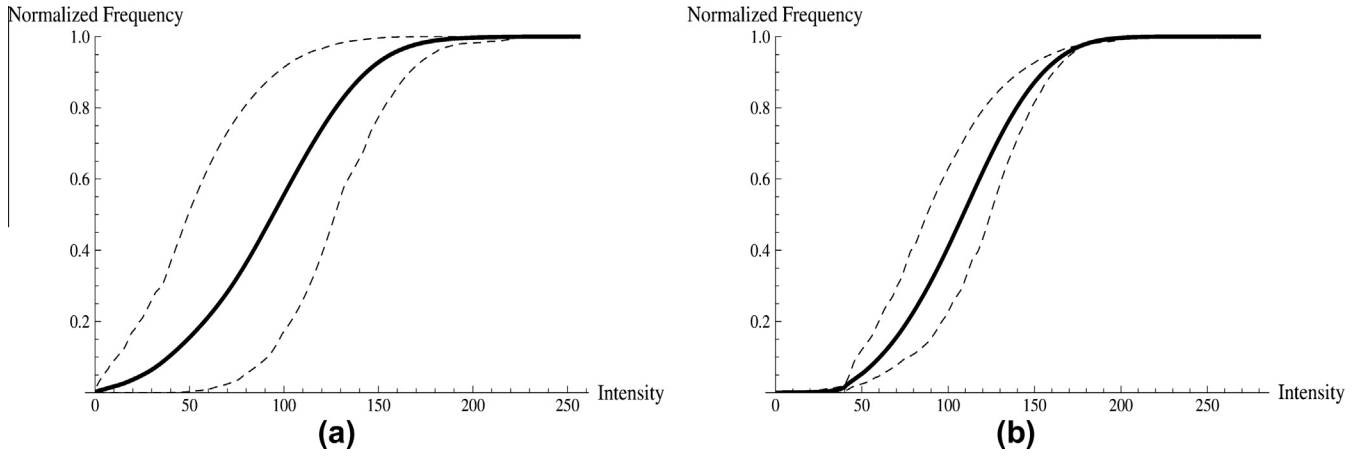


Fig. 2. The normalized cumulative histogram of intensities within the breast of all the images before normalization (a) and after normalization (b). The solid line stands for the mean of the normalized cumulative histograms of all images in our dataset and the dashed lines indicate the 95% confidence interval.

We use intensity normalization to scale the intensities within the breast mask across all images to a certain range. The standard-score normalization scales all intensities to have a zero mean and a unit standard deviation. However, the imaged breast tissue strongly depends on the composition of the breast, i.e. the proportion of fatty and dense tissue. Therefore standard-score normalization is not appropriate. We previously (Tan et al., 2012) introduced a normalization method in which the median intensity value of fatty tissue and dense tissue are used as reference values. However this method depends on the accuracy of the tissue segmentation and the intensities of dark shadows are not used. In our current method, within the breast mask, the 98th percentile of all intensity values, I'_{98} and the 2nd percentile of all intensity values, I'_2 were extracted. The normalized intensity I for each voxel is computed as:

$$I = \frac{(I_{98} - I_2)(I_o - I'_2)}{I'_{98} - I'_2} + I_2 \quad (1)$$

where I_o is the original voxel value and I_{98} and I_2 are constants representing reference values for the 98th percentile and the 2nd percentile of all intensity values within the breast mask.

We chose the 98th percentile and the 2nd percentile of all intensity values as reference intensities for normalization instead of using the maximum and the minimum of all intensity values to avoid the influence of outliers on the normalization.

Fig. 2 shows the mean and the 95% confidence interval of all normalized cumulative histograms of intensities within the breast before intensity normalization and after. As we can see, before normalization, the distribution of the normalized cumulative histograms spans a much wider range compared to after intensity normalization.

3.2. The chest wall as a cylinder

Our method is based on the assumption that the partial chest wall visible in a single volumetric view can be modeled by the surface of a cylinder. To investigate if this conjecture is accurate, we fit the cylinder model to our manually annotated point sets for each scan.

We used a non-linear optimization method (Powell, 1964) to fit a cylinder to the annotated points in 3D space and evaluate the validity of our assumption by observing the average distance of the annotated points to the fitted cylinder surface for each scan.

In all equations in this paper, x and y axis are two orthogonal axis on coronal plane and the z axis (depth direction) is the ventral-dorsal axis. A cylinder in 3D space is fully defined by its radius R and its axis (center line). The axis in turn is defined by a point

(x_c, y_c, z_c) and an orientation vector (v_x, v_y, v_z) . We set $z_c = 0$ and $v_z = 1$ to remove redundancy. The normalized direction vector with length equal to 1 is denoted by (u_x, u_y, u_z) . Thus the parameter vector $\mathbf{p} = \{R, x_c, y_c, v_x, v_y\}$ fully defines the cylinder.

For any point (x, y, z) in 3D space, the closest point on the axis of the cylinder has coordinates (x_a, y_a, z_a) defined by

$$\begin{aligned} x_a &= x_c + D * u_x \\ y_a &= y_c + D * u_y \\ z_a &= z_c + D * u_z \end{aligned} \quad (2)$$

where $D = (x - x_c) * u_x + (y - y_c) * u_y + (z - z_c) * u_z$ which is the distance between the point (x_a, y_a, z_a) and (x_c, y_c, z_c) with $z_c = 0$. Therefore, given a set of manually marked rib-surface voxels (x_n, y_n, z_n) ($n = 1, 2, \dots, N$) the averaged squared distance to the cylinder surface $ASD(\mathbf{p})$ can be computed by

$$\frac{1}{N} \sum_{n=1}^N (\sqrt{(x_n - x_a)^2 + (y_n - y_a)^2 + (z_n - z_a)^2} - R)^2 \quad (3)$$

where N is the number of annotated points.

We constrain the fitted cylinder by introducing a penalty term $P(\mathbf{p})$

$$P(\mathbf{p}) = w_1 * \left(\frac{v_x^2 + \xi}{v_y^2 + \xi} \right) + w_2 * (R - R_{mean})^2 \quad (4)$$

in which the first term of $P(\mathbf{p})$ ensures that the orientation of the cylinder remains close to the body axis (y is directed towards to head). ξ is chosen to be a small positive number and is used to ensure a non-zero denominator. The second term penalizes the deviation of the radius of the fitted cylinder from a predefined average radius R_{mean} related to local curvature of the chest wall. The parameters w_1 and w_2 define the weights for each term.¹ The total cost function $C(\mathbf{p})$ is defined by

$$C(\mathbf{p}) = ASD(\mathbf{p}) + P(\mathbf{p}) \quad (5)$$

We used Powell's optimization method (Powell, 1964) to find the minimum of the cost function $C(\mathbf{p})$.

The cylinder fitting method was applied to the 75 images of dataset A and the 67 images of dataset B. The average of the mean distance of the annotated points in each image to the fitted cylinder surface for dataset A and dataset B was $2.19 \text{ mm} \pm 0.99 \text{ mm}$ and $2.43 \text{ mm} \pm 2.76 \text{ mm}$, respectively. Given this small error we

¹ In our work, we used $R_{mean} = 78 \text{ mm}$, which was an experimental estimate from our dataset. w_1 and w_2 were experimentally set to 100 and 150, respectively.

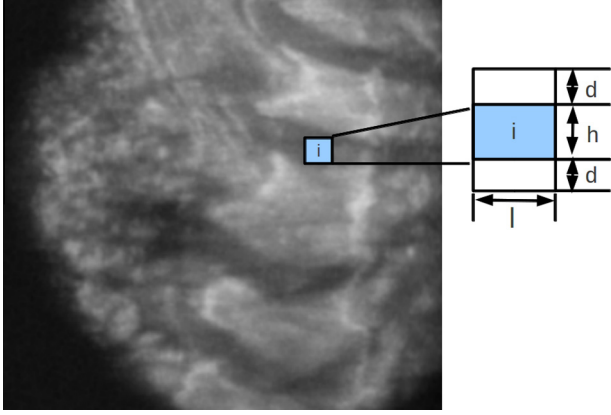


Fig. 3. For each voxel, a local rectangle is defined to extract local contrast on 2D coronal plane.

conclude that the surface of the chest wall in an ABUS scan can be approximated by the surface of a cylinder.

3.3. Chest wall segmentation

Our proposed method consists of two stages: first, rib-surface points are identified and second, the previously described fitting procedure is used to fit a cylinder to these points. In 3D breast ultrasound, rib-surface points are at the transition between breast tissue and rib shadows. Detecting rib shadows as landmarks is easier than detection of rib-surface points. Therefore, the extraction of rib-surface points involves two steps: the detection of rib shadows and subsequently the detection of the rib surface.

3.3.1. Detection of rib shadows

Due to strong attenuation of the ultrasound signal by the ribs, a shadow is cast behind them in the ABUS images (Fig. 1). The detection of these dark sheets aids us in localizing the rib-surface points. Here we describe a set of features we used in combination with a classifier to identify the rib shadows.

Shadow contrast: we observed that on 2D coronal plane, rib shadows appear as embedded dark horizontal bands. To characterize local contrast for voxel i on coronal plane, we defined a rectangle at center i with a fixed length l (5 mm) and a width h as shown in Fig. 3. A shadow contrast feature is defined by:

$$SC(h) = \bar{I}(h + 2d) - \bar{I}(h) \quad (6)$$

where $\bar{I}(h)$ represents the average intensity of a rectangle at center i with a length of l and a width of h and d is chosen to be 3 mm. SC is computed as a function of h ($5 \text{ mm} < h < 16 \text{ mm}$) and the maximum of SC is chosen as the feature value. This feature is computed in 2D in the coronal planes. Fig. 4 shows this feature as overlay.

Dark shadow enhancement: in 3D space, a rib shadow appears as a dark sheet and therefore we apply a sheet detector to enhance the rib shadows in ABUS images. The sheet detector utilizes the eigenvalues and eigenvectors of the Hessian matrix.

We compute the elements of the Hessian matrix by convolving image $I(\mathbf{x})$ with Gaussian kernel

$$L(\mathbf{x}, \sigma) = I(\mathbf{x}) \otimes G(\mathbf{x}, \sigma) \quad (7)$$

where $G(x, \sigma)$ is the Gaussian filter and $L(x, \sigma)$ is the result of the convolution. The normalized Hessian matrix is defined as

$$H(\mathbf{x}) = \sigma^2 \begin{bmatrix} \frac{\partial^2 L(\mathbf{x}, \sigma)}{\partial x^2} & \frac{\partial^2 L(\mathbf{x}, \sigma)}{\partial x \partial y} & \frac{\partial^2 L(\mathbf{x}, \sigma)}{\partial x \partial z} \\ \frac{\partial^2 L(\mathbf{x}, \sigma)}{\partial y \partial x} & \frac{\partial^2 L(\mathbf{x}, \sigma)}{\partial y^2} & \frac{\partial^2 L(\mathbf{x}, \sigma)}{\partial y \partial z} \\ \frac{\partial^2 L(\mathbf{x}, \sigma)}{\partial z \partial x} & \frac{\partial^2 L(\mathbf{x}, \sigma)}{\partial z \partial y} & \frac{\partial^2 L(\mathbf{x}, \sigma)}{\partial z^2} \end{bmatrix} \quad (8)$$

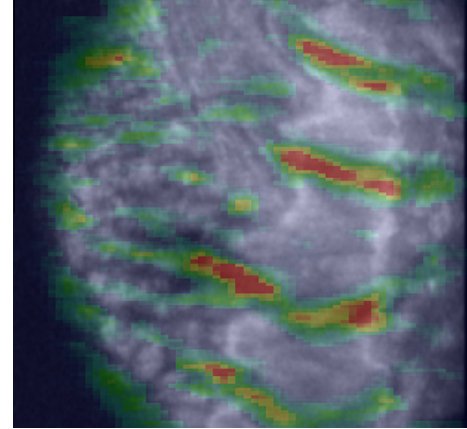


Fig. 4. A 2D coronal plane with rib shadow contrast (SC) overlay.

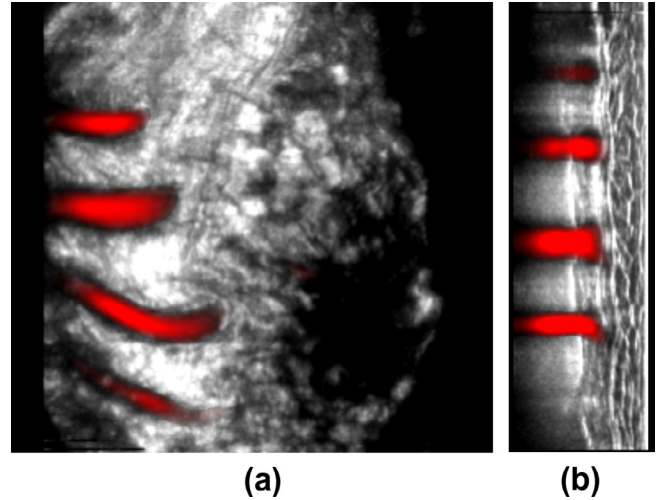


Fig. 5. A scan with dark shadow enhancement overlay (red). A coronal slice (a) and a sagittal slice (b) are shown. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

where σ is the scale.² The eigenvalues of $H(\mathbf{x})$ are λ_1, λ_2 and λ_3 which correspond eigenvectors V_1, V_2, V_3 where $|\lambda_3| > |\lambda_2| > |\lambda_1|$. The dark sheet enhancement filter (Sato et al., 2000) is defined as:

$$DS(\lambda_1, \lambda_2, \lambda_3) = \begin{cases} \lambda_3 * e^{-(\lambda_1^2 + \lambda_2^2)} & \text{if } \lambda_3 > 0 \\ 0 & \text{if } \lambda_3 < 0 \end{cases} \quad (9)$$

The first eigenvector V_3 gives the direction of the greatest curvature. To make the operator more sensitive to sheets perpendicular to the body axis or vertical direction, a weight factor F is calculated as the inner product between V_3 and body-axis vector $(0, 1, 0)$.

Then, the directional dark shadow enhancement image (Fig. 5) is computed as

$$SH(\mathbf{x}) = DS(\mathbf{x}) * F \quad (10)$$

Additional shadow features: the shadow contrast and shadow enhancement alone, however, are not specific enough. Therefore we used a number of additional features to improve the detection of rib shadows including:

- I , image intensity

² σ is set to 5 mm which is estimated from the rib-shadow width.

- $|G_z|$, the absolute values of the first order Gaussian derivative of the intensity image in the depth direction at a scale of 0.6 mm (one voxel) to characterize the fine tissue structure and 2.4 mm (four voxels) to characterize to coarse structures but at a scale small enough to have a low response within the rib shadows
- z , the depth of the voxel

To compute the likelihood that a voxel belongs to a rib shadow, we apply a two-class soft classifier using the features described above as input and using the manual annotations for training.

For the class of rib shadows we used the column of voxels below the annotated rib surface points. For the other class, we used a set of regularly spaced samples (2.4 mm in each direction) that are at least 6 mm above the chest wall (as defined by a fitted cylinder on the manually annotated rib-surface points).

We used an ensemble of five neural networks to classify this data. Each network consists of an input layer of n nodes representing the features, where n is the number of features, and a hidden layer of eight hidden nodes and an output layer with one node. All networks were initiated with random weights to reduce the influence of the initial weights on the final output of the trained network. The back-propagation algorithm was used for training with a learning rate of 0.005. The final classifier output was computed by averaging the five network outputs. This output, $LS(\mathbf{x})$, is considered to represent the likelihood that a voxel is part of a rib shadow (Fig. 6). This likelihood is used as input to the rib-surface detector described below.

3.3.2. Detection of rib surface

Based on the likelihood map for rib shadows, the extracted dark sheet enhancement feature, and the original intensity image, a number of features are extracted to characterize rib-surface points.

Neighboring rib shadows: a 2D coronal projection (Fig. 7a) is obtained by averaging all 2D coronal slices of the likelihood map. A 2D binary projection (Fig. 7b) is generated by thresholding the 2D coronal projection. The threshold was defined as the 85th percentile of all likelihood values within the breast mask in each image, as this roughly corresponds to the relative area of the projected rib shadows, $\pm 15\%$.

To be able to utilize the contextual information of neighboring rib shadows, for each voxel, a template is defined as Fig. 8. The template consists of three rib-shadow segments $S0, S1$, and $S2$ with length l (3 mm) and width a ($5 \text{ mm} < a < 15 \text{ mm}$) and two non-shadow segments NS with length l and width b ($5 \text{ mm} < b < 27 \text{ mm}$). Using this template located at each voxel i , the width a and b are locally estimated by:

$$\arg \max_{a,b} \bar{I}_{S0 \cup S1 \cup S2_a} - \bar{I}_{NS_b} \quad (11)$$

where $\bar{I}_{S0 \cup S1 \cup S2_a}$ is the average value of three rib-shadow segments $S0, S1$, and $S2$ on the 2D binary projection and \bar{I}_{NS_b} is the average value of non-shadow area NS .

Using the estimated rib shadow width a and non-shadow width b , a contextual feature (Fig. 7c) for voxel i indicating the neighboring rib shadows (CNR) is defined:

$$\text{CNR} = \bar{I}_{S1 \cup S2_a} - \bar{I}_{NS_b} \quad (12)$$

where $\bar{I}_{S1 \cup S2_a}$ is the average value of the two neighboring shadow segments $S1$ and $S2$ of the segment $S0$ on the 2D binary projection. Since this feature is computed from the 2D binary coronal projection, it has the same value in the z -direction, for each (x, y) location.

Coronal distance to the nipple: the distance of voxel (x, y, z) to the nipple in coronal view d_{npl} is extracted. This maybe helpful to reduce false positives, since strong shadowing is caused by the nipple and ducts connecting the nipple. On coronal slices close to the transducer, the nipple is visible as a dark circle. Therefore we applied a Hough circle transform (Duda and Hart, 1972) on the coronal slices closer than 4.2 mm to the transducer. The detection results for these slices were fused to obtain the position of the nipple by voting.

Additional rib-surface features: additional features are incorporated characterizing rib-surface. These features are based on the intensity image, the previously described rib shadow likelihood image $LS(\mathbf{x})$ and dark sheet enhancement feature (DS):

- I , image intensity (rib surface appears as bright bands)
- z , the depth of the voxel: it is incorporated since the depth of the rib surface cannot be too close to the skin or too deep
- $Gz_{DS}(x, y, z)$, the maximal value of the first order Gaussian derivative in the depth direction on the original feature image $DS(x, y, z)$ across a scale of 0.6 mm, 2.4 mm and 4.8 mm.
- $Gz_{LS}(x, y, z)$, the maximal value of the first order Gaussian derivative in the depth direction on the rib shadow likelihood image $LS(\mathbf{x})$ across a scale of 0.6 mm, 2.4 mm and 4.8 mm.
- $\frac{\sum_{k < z} LS(x, y, k)}{z}$, the average feature value (LS) of voxels above the voxel.
- $\frac{\sum_{k > z} LS(x, y, k)}{N_z - z - 1}$, the average feature value (LS) of voxels below the voxel.
- $\frac{\sum_{k < z} LS(x, y, k)}{z} - \frac{\sum_{k > z} LS(x, y, k)}{N_z - z - 1}$.

We used a classifier with the same settings as the classifier used in the previous step except that the classifier in this step has a input layer of nine nodes based on the features computed above. For the class of rib-surface points we use the manually annotated points. For the non-rib voxels we used a set of regularly spaced samples that are at least 6 mm away from the chest wall (as defined by a fitted cylinder on the manually annotated rib-surface points).

This classification yields a likelihood for each voxel $Li(x, y, z)$, representing the probability that a voxel belongs to the rib surface (see Fig. 9).

The rib-surface points are determined by thresholding image $Li(x, y, z)$ with threshold defined as the 98th percentile of all likelihood values within the breast mask in each image.

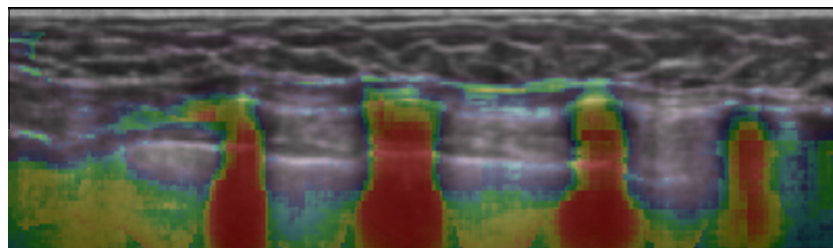


Fig. 6. A sagittal view image overlaid with the rib shadow likelihood LS . The probability ranges from blue (low) to red (high). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

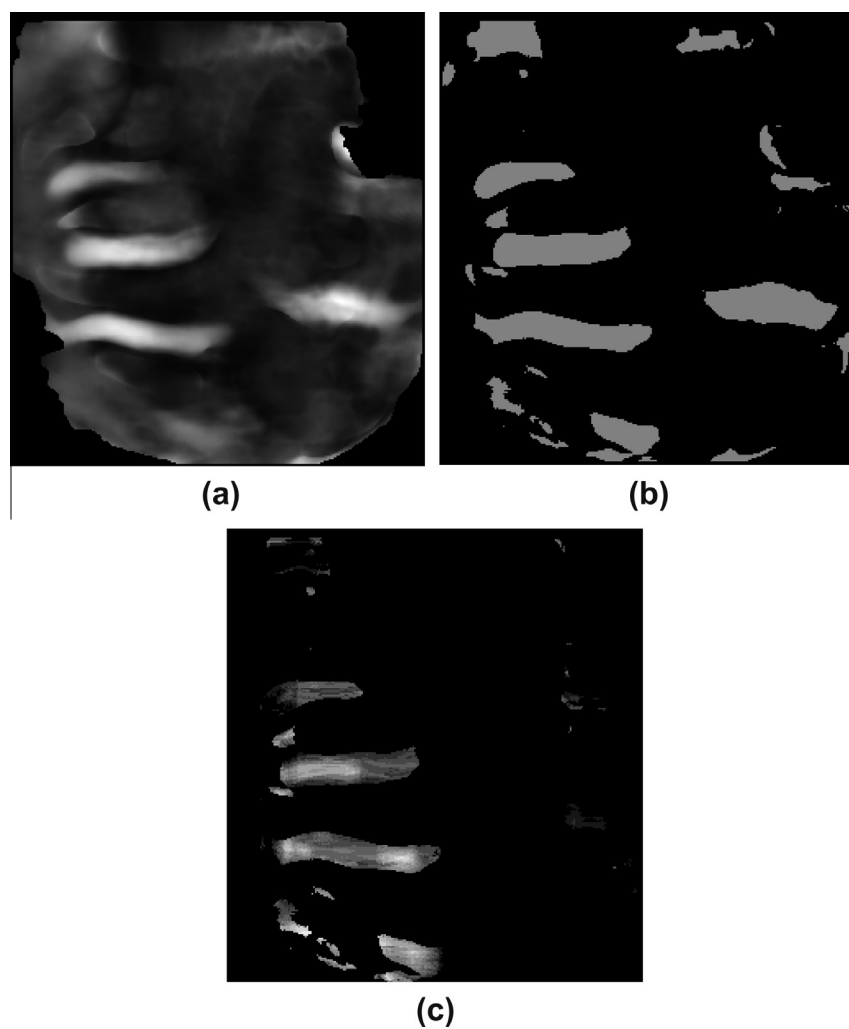


Fig. 7. (a) A 2D projection obtained by averaging all 2D coronal slices of probability map. (b) A binary 2D projection obtained by thresholding the image (a). (c) Computed contextual feature map.

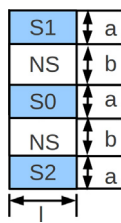


Fig. 8. A template for extracting contextual information. Blue segments stand for rib shadow regions with a fixed length l and a variable width a and white segments stand for non-rib shadow regions in between with a fixed length l and a variable width b . This template is used to compute contextual features indicating the existence of neighboring rib shadows at each pixel in the 2D binary projection.

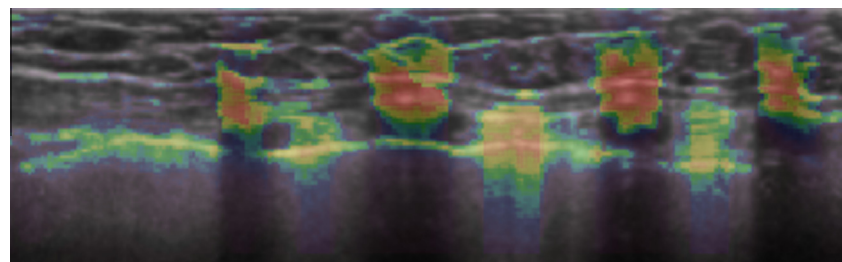


Fig. 9. A sagittal view image overlaid with the rib-surface likelihood L_i . The probability ranges from blue (low) to red (high). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

3.3.3. Removal of false positive due to the nipple

Due to strong shadowing of the nipple and shadowing of ducts connected to the nipple, there were still a number of false positives were left after thresholding. Therefore among detected rib-surface points, points whose distance to the nipple in 2D coronal projection are within 9 mm are removed.

3.3.4. Cylinder fitting to detected rib-surface points

In this step, we fit our cylinder model to the detected rib-surface points in a similar manner as described in Section 3.2. However, this time we weight the contribution of each rib-surface point by its likelihood LI :

$$\frac{1}{N} \sum_{n=1}^N w_n (\sqrt{(x_n - x_a)^2 + (y_n - y_a)^2 + (z_n - z_a)^2} - R)^2 \quad (13)$$

where w_n is defined as

$$w_n = LI(x_n, y_n, z_n)^{p1} * \left(\frac{d_{npl}(x_n, y_n)}{d_{npl_{max}}} \right)^{p2} \quad (14)$$

$p1$ is chosen to enhance the weight difference of rib-surface points with different likelihoods and $p2$ is chosen to enhance the weight difference of rib-surface points with different distance to the nipple and $d_{npl_{max}}$ is the maximal distance of the rib-surface points to the nipple ($p1$ and $p2$ were both experimentally set to 3).

3.3.5. Classifier training

Given a dataset, we split it into two folds (fold 1 and fold 2) to perform twofold cross-validation. To avoid bias, the two classifiers (rib-shadow classifier and rib-surface classifier) are both trained on data from fold 1 and tested on data from fold 2. To obtain the result for fold 1, we repeat the above procedure by swapping the two folds.

3.3.6. Implementation and computation time

Our segmentation method was implemented in C++. Using a computer with a 2660 MHz CPU and 8 GB RAM, the total computation time for the chest wall segmentation in an image is about 6 min and 30 s in which the cylinder fitting procedure takes 1 min and 30 s. We used a workstation developed by our group for visualization.

4. Results

We compute an error measure for each chest wall segmentation. This error measure is defined as the average distance of all annotated rib surface points to the automatically segmented chest wall. We applied our method separately on dataset A and dataset B and we also trained and tested our method on a combined dataset in which images of dataset A and dataset B were mixed (Table 1).

For the 75 scans of dataset A, the mean error was 5.73 mm with standard deviation 3.11 mm. For the 63 scans of dataset B, the mean error was 5.43 mm with standard deviation 3.07 mm. For all 142 images, the mean error was 5.59 mm with standard deviation 3.08 mm. A boxplot and a cumulative histogram plot summarize the result (Figs. 10 and 11).

Table 1
Segmentation results.

Testing set	Training set	Mean error
A	A	5.73 mm \pm 3.11 mm
B	B	5.43 mm \pm 3.07 mm
A + B	A and B separately	5.59 mm \pm 3.08 mm
A + B	A + B pooled	5.95 mm \pm 3.54 mm

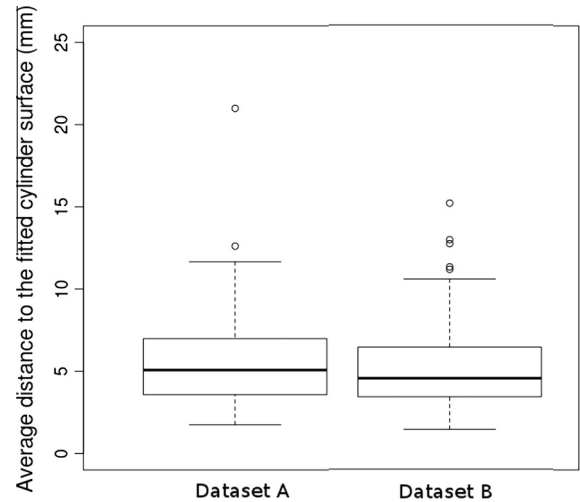


Fig. 10. Box plot of the average distance of the annotated points (reference standard) in each image to the fitted cylinder surface using detected rib-surface points in dataset A and dataset B.

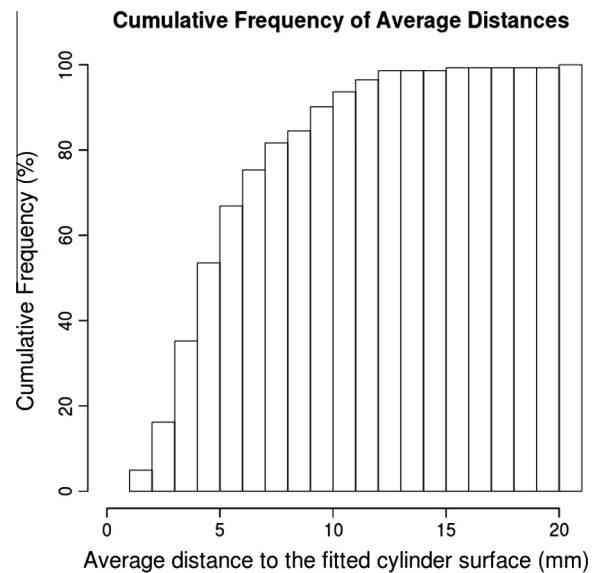


Fig. 11. Cumulative histogram for the averaged distance of annotated rib-surface points to the segmented surface for all images in dataset A and dataset B with separated training.

Three segmentation examples are shown in Fig. 12 with an average distance of 4.64 mm, 9.03 mm, and 12.61 mm, respectively. By visually inspecting our segmentation results we determined that an average distance of 8 mm or less generally results in an acceptable segmentation for the development of CADE systems and advanced visualization tools for ABUS reading. As seen from Fig. 11, for all 142 images, about 82% of scans had an average distance less or equal to 8 mm.

By pooling all 142 images together for training and testing, the mean error was 5.95 mm with standard deviation 3.54 mm which is significantly different to the resulting with separated training (paired t -test, $p = 0.008$).

5. Discussion and conclusion

Our result in Section 3.2 demonstrate that the surface of the chest wall in an ABUS scan can be well approximated by the

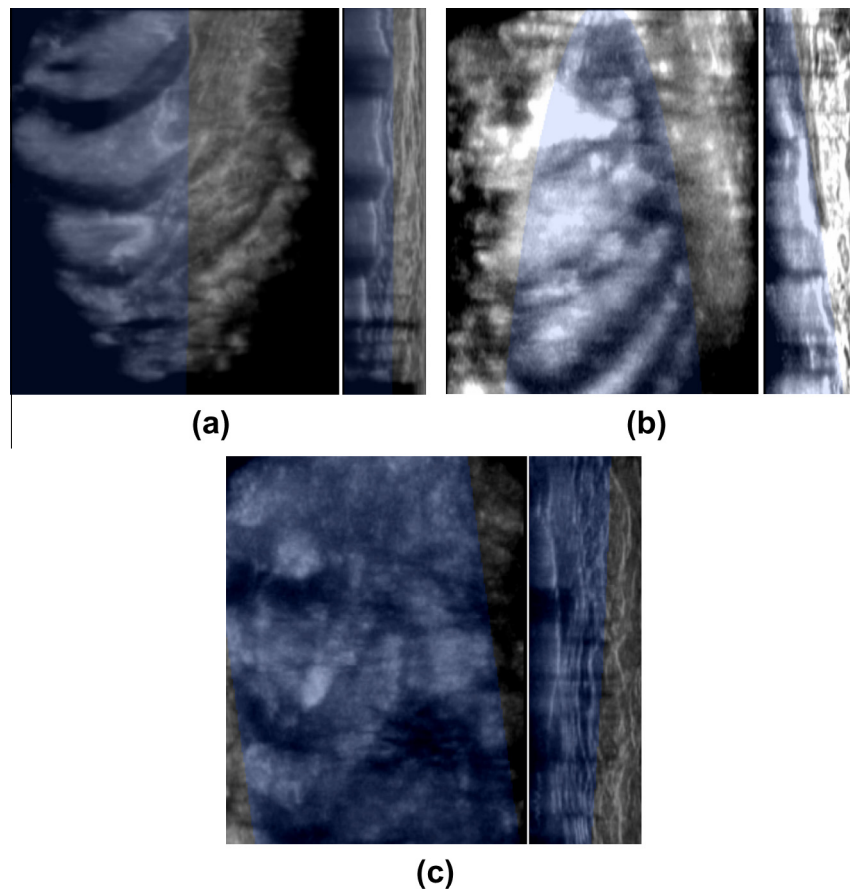


Fig. 12. Segmentation examples with chest wall segmentation as overlay. Three segmentation examples are shown in coronal and sagittal planes with an average distance between the automatically detected chest wall and the manually annotated rib surface points of 4.64 mm (a), 9.03 mm (b) and 12.61 mm (c) in ABUS, respectively. As we can see for case (c), our method over segmented the chest wall with a large error.

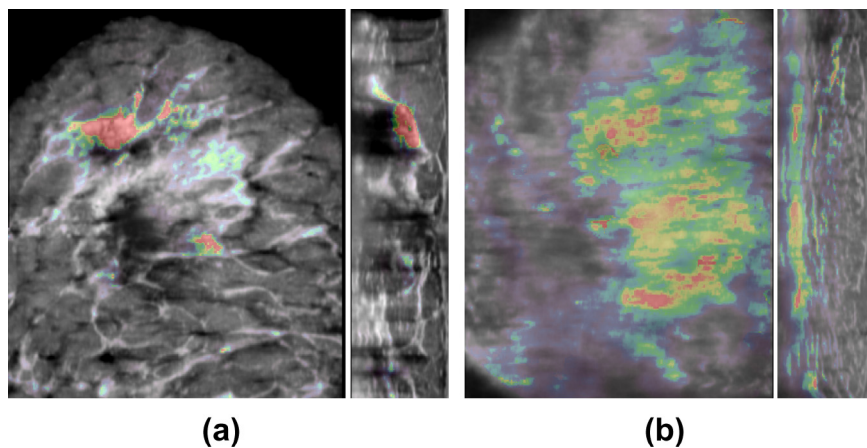


Fig. 13. Two examples of cases with poor result in coronal and sagittal views with the likelihood map for rib surface as overlay. False positives are caused by dense tissue in case (a) and rib surface in case (b) are assigned low likelihood due to insufficient scanning depth.

surface of a cylinder. Therefore we investigated an automated method to locate chest wall surface using this model. Using this the average of the mean distance of the annotated points to the segmented chest wall surface was 5.73 mm (dataset A) and 5.43 mm (dataset B), which is promising for follow-up processing, such as computer-aided detection and inter- and intra-modal image registration.

From the fitting error using manual annotations (2.19 mm for dataset A and 2.43 mm for dataset B), we can see there is a limitation of using the cylinder model. The limitation may be due to our

cylinder model. To improve the method, in the future we can include additional parameters that allow bending of the cylinder as well as an elliptical cross-section. We can also use our result as a starting point for further refinement, for instance, using graph search methods to optimize a defined layer or surface (Garvin et al., 2008).

In our work, training and testing is involved for the evaluation. We investigated if it was necessary to use the images from the same manufactures for training and testing as opposed to mixing

the images from two different manufactures. We found that by using images from the same manufactures for training and testing, significantly better results were obtained than by pooling images of different manufactures.

The segmentation method had poor results for a few cases. One reason for poor performance is that for extremely dense breasts, dense tissue also causes shadows behind it. Our rib-surface detection algorithm in those cases assigns high probabilities to the dense tissue, resulting in a poor segmentation of the chest wall. An example is shown in Fig. 13a. Another reason is that for some images, rib shadows are not obvious enough due to insufficient scanning depth. For example in Fig. 13b, the scanning depth is not sufficient to include rib shadows (the rib surface is too close to the bottom of image). For further improvement, false positives of the rib-surface detector must be reduced, for instance by incorporating more features characterizing dense tissue and ducts.

Acknowledgments

This work was presented at the MICCAI (the International Conference on Medical Image Computing and Computer Assisted Intervention) 2011 Workshop on Breast Image Analysis. This work has been funded by the HAMAM Project (IST-2007-224538) within the Seventh Framework Programme (FP7) of the EU. The authors thank André Grivegnée from Cancer Prevention and Screening Clinic, Jules Bordet Institute, Brussels, Belgium, László Tabár from Department of Mammography, Falun Central Hospital, Sweden and Matthieu Rutten from Department of Radiology, Jeroen Bosch Ziekenhuis, Den Bosch, the Netherlands for providing the data.

References

- Berg, W., 2004. Supplemental screening sonography in dense breasts. *Radiologic Clinics of North America* 42, 845–851.
- Duda, R.O., Hart, P.E., 1972. Use of the Hough transformation to detect lines and curves in pictures. *Communications of the ACM* 15, 107–111.
- Ferlay, J., Shin, H.R., Bray, F., Forman, D., Mathers, C., Parkin, D.M., 2010. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *International Journal of Cancer* 127, 2893–2917.
- Garvin, M.K., Abramoff, M.D., Kardon, R., Russell, S.R., Wu, X., Sonka, M., 2008. Intraretinal layer segmentation of macular optical coherence tomography images using optimal 3-D graph search. *IEEE Transactions on Medical Imaging* 27, 1495–1505.
- Huisman, H., Karssemeijer, N., 2007. Chestwall segmentation in 3D breast ultrasound using a deformable volume model. In: *Information Processing in Medical Imaging Lecture Notes in Computer Science*, pp. 245–256.
- Kelly, K.M., Dean, J., Comulada, W.S., Lee, S.-J., 2010. Breast cancer detection using automated whole breast ultrasound and mammography in radiographically dense breasts. *European Radiology* 20, 734–742.
- Mandelson, M.T., Oestreicher, N., Porter, P.L., White, D., Finder, C.A., Taplin, S.H., White, E., 2000. Breast density as a predictor of mammographic detection: comparison of interval- and screen-detected cancers. *Journal of the National Cancer Institute* 92, 1081–1087.
- Otsu, N., 1979. A threshold selection method from gray level histograms. *IEEE Transactions on Systems, Man, and Cybernetics* 9, 62–66.
- Powell, M., 1964. An efficient method for finding the minimum of a function of several variables without calculating derivatives. *The Computer Journal* 7, 152–162.
- Rotten, D., Levaillant, J.M., Zerat, L., 1999. Analysis of normal breast tissue and of solid breast masses using three-dimensional ultrasound mammography. *Ultrasound in Obstetrics and Gynecology* 14, 114–124.
- Sato, Y., Westin, C., Bhalerao, A., Nakajima, S., Shiraga, N., Tamura, S., Kikinis, R., 2000. Tissue classification based on 3D local intensity structures for volume rendering. *IEEE Transactions on Visualization and Computer Graphics* 6, 160–180.
- Tan, T., Platel, B., Huisman, H., Sánchez, C.I., Mus, R., Karssemeijer, N., 2012. Computer aided lesion diagnosis in automated 3D breast ultrasound using coronal spiculation. *IEEE Transactions on Medical Imaging* 31, 1034–1042.
- Tanner, C., Karssemeijer, N., Szekely, G., 2011. Deformation models for registering MR and 3D ultrasound breast images. In: *Proc. IEEE Int Biomedical Imaging: From Nano to Macro Symp.*, pp. 582–585.
- Watermann, D.O., Földi, M., Hanjalic-Beck, A., Hasenburger, A., Lüghausen, A., Prömpeler, H., Gitsch, G., Stickeler, E., 2005. Three-dimensional ultrasound for the assessment of breast lesions. *Ultrasound in Obstetrics and Gynecology* 25, 592–598.