Computerized analysis of shadowing on breast ultrasound for improved lesion detection

Karen Drukker a) and Maryellen L. Giger
Department of Radiology MC2026, University of Chicago, 5841 South Maryland Avenue, Chicago, Illinois 60637

Ellen B. Mendelson
Lynn Sage Breast Center, Northwestern University, Chicago, Illinois 60611

(Received 1 April 2002; accepted for publication 17 April 2003; published 25 June 2003)

Sonography is being considered as the screening tool for breast cancer. We are developing computerized detection methods to aid in the localization of lesions on breast ultrasound images. The detection scheme presented here is based on the analysis of posterior acoustic shadowing for malignant solid lesions. Sonography is being considered for the screening of women at high risk for breast cancer. We are developing computerized detection methods to aid in the localization of lesions on breast ultrasound images. The detection scheme presented here is based on the analysis of posterior acoustic shadowing for malignant solid lesions. Sonography is being considered for the screening of women at high risk for breast cancer. We are developing computerized detection methods to aid in the localization of lesions on breast ultrasound images. The detection scheme presented here is based on the analysis of posterior acoustic shadowing for malignant solid lesions.

Key words: lesion detection, breast sonography, computer-aided diagnosis

I. INTRODUCTION

Breast cancer is currently the second leading cause of death for women in developed countries. The early detection of breast cancer may increase the success of treatment, and screening for breast cancer in women over 40 years of age is generally recommended, with mammography being the standard method used for periodic screening. Currently, other methods for detecting and diagnosing breast cancer include sonography (i.e., ultrasound), and magnetic resonance imaging (MRI). Here, we evaluate the possibility of using a computerized detection method in conjunction with sonography for the detection of breast cancer.

In the 1980s, breast sonography yielded limited success for localization and screening, with sonography being useful only for distinguishing simple cysts of the breast from solid lesions. The use of diagnostic and interventional sonography for breast cancer has grown rapidly over the last years, however. With a renewed interest in the use of non-ionizing radiation for the screening of high-risk women, various authors have recently shown that sonography can be used for the classification of solid benign and malignant masses.

The merits of sonography as an adjunct to mammography in screening have been studied by several investigators. Sonography was found to be helpful in the detection of otherwise occult malignancies in (young) women with dense breasts, in preoperative evaluation, and in the detection of masses associated with mammographically detected microcalcifications. Another study showed that the use of sonography as an adjunct to mammography results in an increase in the diagnostic accuracy.

The usefulness of sonography alone as a screening tool has not yet been established. In one study, sonography was found to be more effective than mammography for women under 35 years, since mammograms of younger women are often hard to interpret. Investigators in a large study on the effectiveness of ultrasound as a screening tool for women with dense breasts sonographically examined more than 3000 patients who had dense breasts and normal mammographic and physical examinations. The use of ultrasound in these patients increased overall cancer detection by 17%. It was shown that ultrasound was able to depict small, early-stage, otherwise occult malignancies, similar in size and stage as those detected by mammography, and smaller and lower in stage than palpable cancers in dense breasts.

These studies illustrate that sonography has potential as a screening tool. The added benefits are that sonographic equipment is relatively inexpensive and portable, and does not involve ionizing radiation. Sonography, however, is unable to image microcalcifications, making it unlikely it will...
Computer-aided diagnosis (CAD) methods in breast ultrasound are being explored by various researchers. To date, investigations have focused on characterizing sono-...
where $N$ is the number of data points in the region of interest $A$, $\langle \rangle$ is the arithmetic mean, and $\sigma_A$ is the standard deviation of the gray-value distribution in that area.

The full skewness image is obtained by calculating the skewness for each possible mask position $(x,y)$ in the original image with a mask having the size and shape of the region of interest (ROI). The skewness values are assigned to the mask center points $(x,y)$ in the skewness image. Note that this procedure does not assign values to pixels in the skewness image closer to the edge than the full ROI size allows, and we have chosen to leave the borders of the skewness image blank. The pixel values in the skewness image are an estimate of the likelihood that a shadow is present, with skewness values theoretically ranging between $(-\infty, +\infty)$. The skewness image is subsequently thresholded in order to determine areas of interest. For normalization purposes, we automatically scale each skewness image to have zero mean and unit standard deviation, and employ a threshold value $t$ given in units of the standard deviation $\sigma_s$ of the calculated skewness image (excluding the undefined edge pixels),

$$t = m\sigma_s,$$  \hspace{1cm} (2)

where the choice for $m$ depends on the desired sensitivity and false-positive detection rate. Note that $\sigma_s$ equals one after the scaling procedure. The centers of areas of interest remaining after this thresholding are defined as detection points, i.e., as shadow candidates.

C. Implementation

A subsampling factor of 4 was used in the calculation of the skewness images. The images were cropped by 2 mm at all edges, since often artifacts were observed close to the image edge. The region of interest (ROI) was chosen as a rectangle since the shadow structures of interest tend to have a rectangular shape. Different ROI sizes were investigated. For a ROI height of 15 mm, widths of 1.25, 2.5, and 5 mm were used. For a ROI height of 10 mm, widths of 2.5 and 5 mm were investigated, and for a ROI height of 5 mm a width of 2.5 mm was employed.

The skewness values were calculated for each position of the ROI mask within the images, by calculating the skewness of the histogram of the pixel values in each ROI combined with a small number of noise pixels sampled from a uniform distribution. This “noise,” which contributed a uniform distribution to the histogram, was added in order to prevent undesirable computational problems upon encountering image regions with zero variation in pixel value. Note that image pixels were not changed in this procedure, and that noise was added only to the histogram. The number of added noise pixels in the histogram analysis was 10% of the number of pixels in each ROI. This number was determined empirically to avoid computational problems without significantly distorting the actual gray-value distribution. The values of the noise pixels were selected randomly between the minimum and maximum gray values of a given image. For a given image, the added histogram noise was the same throughout the analysis.

For the threshold value in the analysis of the skewness image, i.e., in the determination of areas of interest, a value of $t = 1.75\sigma$ was used unless otherwise specified, with the value being constant and equal for all images. In the performed FROC analysis the threshold value $t$ was used to sweep out the FROC curves and was increased in a stepwise fashion from 0.25 to 3.75 standard deviations [see Eq. (2)] to determine the effect of the threshold value on detection performance. The minimum size requirement for an area to be considered an area of interest was ten pixels.

D. Evaluating performance

The performance of the shadow detection scheme was obtained by designating detection output points as true-positive (TP) detections if they were located below the lesion in question. The image region in which detections were defined as true positive was determined from a vertical band below the lesion, with the width of the lesion being based on the radiologist outline of the lesion. All detection points located outside the vertical band were defined as false-positive (FP) detections. Since this shadow analysis method does not give a precise location of a lesion, but rather points to an area of special interest above the detection point, it was felt that this was a fair assessment of performance.

In the calculation of sensitivity (true-positive fraction per image or per case), only one true-positive detection was counted for a given lesion even if there were multiple detection points satisfying the true-positive criterion. Note that the analysis was performed in exactly the same way for all images and for all lesion types, independent of human-perceived shadowing characteristics. These shadowing characteristics are of interest, however, in assessing the performance of the detection method. Note that a lesion without a posterior acoustic shadow is likely to go undetected (a false-negative) using the proposed method. Since the sum of the true-positive and the false-negative fraction equals one by definition, the false negatives are not reported separately.

In Fig. 1, examples are given of images with different shadowing characteristics. The characterization of images was based on subjective visual inspection (KD) after consulting a radiologist and prior to the development and assessment of the computerized method. Images were classified as having substantial shadowing, having no—or at best very subtle—lesion shadowing, or as exhibiting artificial shadowing. This distinction is made here (see also Table I) only to help put in perspective the performance of the proposed shadow detection method. Results based on shadow characteristics are presented only in Table III, while for the other tables the perceived shadow characteristics were not taken into account.

All images were processed and analyzed in an identical fashion. For “shadow-free” images, one expects a reduced number of detections. It should be noted, however, that the
computer analysis of the skewness of local gray-value histograms differs from visual perception, and hence may detect shadowing in images labeled as “shadow-free.” It would be a merit to detect shadowing on images where the shadowing is very subtle and not immediately apparent to a radiologist. For images exhibiting artifact shadows, one expects an increased number of false-positive detections. Artifact shadows in this context are defined as shadows not arising from the presence of a lesion [Fig. 3(d)]. Most artifact shadows were observed at the bottom corners of the ultrasound images. It should be noted that in this database these artifact shadows were not due to shadowing caused by normal structures in the breast, since those type of images were not saved during examinations. Since the shadowing characteristics of different lesion types differ substantially, detection results are reported by lesion type as well as for the entire database. While this is not a classification method, it is expected to work best for malignant lesions, since they exhibit posterior acoustic shadowing most frequently.

None of the images are considered “normal” in this study, but the performance (based on whether or not detection points are below a lesion) for images in this database with different shadowing characteristics serves as an indication of how performance depends on those shadowing characteristics. In order to fully evaluate the performance of a detection method, normal images—i.e., images without lesions—should be included in the analysis. This was not possible, however, due to the limitations of our current database. A concern for any computerized detection scheme is the false-positive detection rate, since a method resulting in too many false alarms has limited use. While we have assessed the false-positive rate for abnormal images, this may not truly

Fig. 1. Examples of various images of (a) a malignant lesion exhibiting substantial acoustic shadowing; (b) a malignant lesion with a lesser acoustic shadow still classified as substantial; (c) a malignant lesion that does not exhibit posterior shadowing; and (d) a malignant lesion with most of the shadowing due to an artifact.

<table>
<thead>
<tr>
<th>Lesion type</th>
<th>Number of images</th>
<th>Mean lesion size (mm)</th>
<th>Percent of images showing</th>
<th>Lesion shadow</th>
<th>Artifact shadow</th>
<th>No sign. shadow</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cyst</td>
<td>229</td>
<td>9</td>
<td>11.8</td>
<td>29.7</td>
<td>58.5</td>
<td></td>
</tr>
<tr>
<td>Benign solid</td>
<td>334</td>
<td>10</td>
<td>21.6</td>
<td>25.7</td>
<td>52.7</td>
<td></td>
</tr>
<tr>
<td>Malignant</td>
<td>194</td>
<td>11</td>
<td>37.6</td>
<td>19.6</td>
<td>42.8</td>
<td></td>
</tr>
<tr>
<td>Entire database</td>
<td>757</td>
<td>10</td>
<td>22.7</td>
<td>25.4</td>
<td>51.9</td>
<td></td>
</tr>
</tbody>
</table>

Table I. Subjective characterization of the database illustrating the presence of a posterior acoustic shadow, no significant shadow, and a significant artifact shadow (often occurring at the bottom corners of ultrasound images) by lesion type.
reflect the rate for patient examinations, including normal images. The performance of our method for the current database serves only as an indication of the false-positive rate expected for a more general population. It is not possible to make a reliable prediction about the false-positive rate for normal images, since ultrasound images of normal breast tissue may or may not have shadowing due to normal anatomical structures and interfaces. Moreover, the presence of lesions on all images of our database on the one hand limits the area a shadow may occur, but on the other hand allows lesions to be mistaken for shadows.

### III. RESULTS

An example of the skewness filtering procedure, using a ROI of 5 mm (width) by 15 mm (height), is shown in Fig. 2. The threshold value used in the skewness analysis is \( t = 1.75 \sigma \) for all results presented here, unless stated otherwise. The original image is shown and the analyzed region is marked as well as the ROI mask. In Fig. 2(b), the obtained skewness image is shown (with the area passing the threshold criterion outlined), and in Fig. 2(c) the final output of thresholding and locating the center is presented. The output format is aimed to visually aid the detection of lesion shadows. Note that the distance of a detection point to the lesion is not important in this analysis. A detection point indicates a need for further investigation in the vertical direction, and hence vertical arrows are used in the presentation of the computer detections. Figure 2(d) shows the radiologist outline of the (malignant) lesion and the detection arrow for comparison. An analysis of the shadowing of images is further illustrated in Fig. 3. The gray value histograms of the ROIs of interest and the obtained detections are shown for different lesion types.

Figure 4 shows the FROC curves for different lesion types obtained by varying the skewness threshold value using a given ROI size of 5 by 15 mm. Note that the FROC curves are not strictly monotonic because increasing the threshold values often results in the subdivision of detected regions, potentially increasing the number of false positives. Cyst images show limited shadowing, and hence—as expected—shadow detection results in a limited number of true-positive lesion detections. Images in the database of both benign solid lesions and malignant lesions show substantial shadowing, and hence shadow detection leads to a good performance in lesion detection. Table II illustrates the performance in terms of a true-positive fraction for a given false-positive rate per image. Note that throughout all analyses presented here, all images are considered to be “positive,” i.e., the true-positive fraction equals the number of true-positive detections divided by the total number of images (given that there was one lesion per image for all images). Only a single true-
positive detection is counted for each actual lesion, as explained previously.

Table III shows the performance for lesions with different shadowing characteristics (see Table I) and the fraction of images having no detection points. Here, a ROI size of 5 by 15 mm was employed. It is interesting to note that the detection scheme reached a true-positive fraction of 0.91 at 0.39 false positives per image for images with shadowing. Images without substantial shadowing were less likely to have detection points, as illustrated by the percentage of images without computer detections (37% and 31%, respectively, for the two threshold values). Note that there is no relationship between the true-positive fraction and the percentage of images without a detection, since an analysis of each image may or may not result in a true-positive detection with any number of false-positive detections, or may result in no detections at...
It should also be noted that for the employed threshold values, a non-negligible percentage of lesions classified by visual inspection as having no significant shadow were detected by the shadow analysis (44% and 54%, respectively), demonstrating an ability to detect a weak posterior acoustic change not immediately obvious to the human eye.

The effect of the ROI width and height on the true-positive detection rate is depicted in Fig. 5 for a false-positive rate of 0.25 FP/img, and summarized for malignant lesions in Table IV. In Fig. 5, the sensitivity by case is plotted as a function of the ROI width for different ROI heights, and a comparison is made between the performance for malignant lesions and the database as a whole. As expected, the sensitivity for detecting carcinoma is significantly higher than for the entire database. For a given false-positive per image level, performance is seen to improve for longer and wider ROIs (while maintaining a rectangular shape). One has to keep in mind, however, that the image size forms a physical limitation for the maximum ROI size that is reasonable.

IV. DISCUSSION

Lesion detection on ultrasound can be very difficult, due to the obscurity of lesions, numerous anatomic interfaces, and operator dependence. Here, we propose a computerized detection method based on analysis of posterior acoustic be-
behavior. In a previous paper, our detection method based on lesion shape was shown to perform best for cysts and was seen to have some difficulties identifying malignancies. The difficulties in detecting malignant lesions based on expected lesion shape characteristics arose because of the obscurity of lesions (leading to false negatives), “blending” of lesions and their acoustic shadows (leading to both false negatives and false positives), and the presence of extensive shadowing with a lesion-like appearance (leading to false-positive detections). For these reasons, we developed the methodology based on shadow analysis that was presented here.

The proposed shadow-detection method achieves the best performance for malignant lesions, since malignant lesions generally exhibit more posterior acoustic shadowing than benign masses. The performance for benign solid lesions is satisfactory as well, while cysts are not well detected based on shadow detection since they exhibit posterior acoustic enhancement rather than shadowing, and their edge (refraction) shadows are often too weak to be detected. Our focus was on detecting malignancies, however, especially those that were missed by our lesion-based method. In the future, we plan to combine the methods to enhance overall detection performance.

Our analysis presented here did not contain images completely devoid of a lesion, hampering the evaluation of expected false-positive rates for a more general database. The merit of our method, however, is illustrated by its high sensitivity for detecting shadowing in cancer images. Since tissue interfaces on normal images may also exhibit posterior acoustic shadowing (potentially resulting in false-positive detections), further testing of our method is warranted. It may be beneficial to classify shadows as being caused by a lesion or another entity (e.g., a normal anatomical structure such as a rib) based on calculated image features of detected shadow candidates.

With improving sonographic technology, such as spatial compound scanning, less posterior acoustic shadowing due to normal structures is observed, and the image quality is superior to that of the images used in this study. The performance of our CAD method for images collected with this new technology is expected to improve because of less artifact shadowing. Detection thresholds of the CAD scheme may need to be modified, however, since also actual lesions exhibit less posterior shadowing using these new scanning techniques. Presently, the equipment used in collecting the database for the analysis presented here remains the most widely used.

The skewness analysis presented here was based on a predetermined level of added noise to the histograms in order to avoid computational difficulties. While it was determined that this noise level did not obscure the gray value distributions for the employed region of interest (ROI) masks for this database, it may deserve further study to determine the minimal (optimal) amount of noise pixels necessary to avoid computational problems. In practice, it may be also beneficial to automatically couple the ROI size with the image size. We found that for a given ROI height, the wider the ROI the better the performance, and, in general, that the larger the (rectangular) ROI, the better the performance. The image size is a limiting factor to the maximum reasonable ROI size, however, while shadowing characteristics are dependent on the lesion size and type. Moreover, the larger the ROI size with respect to the image size, the more the shadow search is limited to the central region of the image. The effect of ROI size on detection performance needs to be re-

<table>
<thead>
<tr>
<th>Image characteristic</th>
<th>Number of images</th>
<th>Threshold value ( t )</th>
<th>TPF by image</th>
<th>FP/img</th>
<th>% of images with no detections</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lesion shadow</td>
<td>172</td>
<td>1.75</td>
<td>0.78</td>
<td>0.15</td>
<td>17</td>
</tr>
<tr>
<td>Artifact shadow</td>
<td>192</td>
<td>1.75</td>
<td>0.48</td>
<td>0.48</td>
<td>19</td>
</tr>
<tr>
<td>No significant shadow</td>
<td>393</td>
<td>1.75</td>
<td>0.44</td>
<td>0.36</td>
<td>37</td>
</tr>
<tr>
<td>Entire database</td>
<td>757</td>
<td>1.75</td>
<td>0.53</td>
<td>0.34</td>
<td>28</td>
</tr>
<tr>
<td>Lesion shadow</td>
<td>172</td>
<td>1.25</td>
<td>0.91</td>
<td>0.39</td>
<td>2</td>
</tr>
<tr>
<td>Artifact shadow</td>
<td>192</td>
<td>1.25</td>
<td>0.65</td>
<td>0.61</td>
<td>10</td>
</tr>
<tr>
<td>No significant shadow</td>
<td>393</td>
<td>1.25</td>
<td>0.54</td>
<td>0.53</td>
<td>31</td>
</tr>
<tr>
<td>Entire database</td>
<td>757</td>
<td>1.25</td>
<td>0.65</td>
<td>0.52</td>
<td>19</td>
</tr>
</tbody>
</table>
searched further, ideally in a more realistic real-time scanning environment.

In summary, the computerized detection method presented here based on an analysis of posterior acoustic shadowing shows promise as an aid for the detection of lesions on breast ultrasound images.

**ACKNOWLEDGMENTS**

This work was supported in part by USPHS Grants No. CA89452 and No. T32 CA09649, and U.S. Army Medical
Research and Materiel Command 97-2445. The authors would like to thank Luz Venta, M.D. for her help in collecting the sonographic database, and Carl J. Vyborny M.D., Ph.D. for outlining the sonographic lesions in the database. M.L.G. is a shareholder in R2 Technology, Inc. (Los Altos, CA). It is the University of Chicago Conflict of Interest Policy that investigators disclose publicly actual or potential significant financial interest that would reasonably appear to be directly and significantly affected by the research activities.

Electronic mail: kdrukker@hotmail.com