Clinical Feasibility of Quantitative Ultrasound Texture Analysis
A Robustness Study Using Fetal Lung Ultrasound Images

Alvaro Perez-Moreno, MS, Mara Dominguez, MS, Federico Migliorelli, MD, Eduard Gratacos, MD, PhD, Montse Palacio, MD, PhD, Elisenda Bonet-Carne, PhD

Objectives—To compare the robustness of several methods based on quantitative ultrasound (US) texture analysis to evaluate its feasibility for extracting features from US images to use as a clinical diagnostic tool.

Methods—We compared, ranked, and validated the robustness of 5 texture-based methods for extracting textural features from US images acquired under different conditions. For comparison and ranking purposes, we used 13,171 non-US images from widely known available databases (OUTEX [University of Oulu, Oulu, Finland] and PHOTEX [Texture Lab, Heriot-Watt University, Edinburgh, Scotland]), which were specifically acquired under different controlled parameters (illumination, resolution, and rotation) from 103 textures. The robustness of those methods with better results from the non-US images was validated by using 666 fetal lung US images acquired from singleton pregnancies. In this study, 2 similarity measurements (correlation and Chebyshev distances) were used to evaluate the repeatability of the features extracted from the same tissue images.

Results—Three of the 5 methods (gray-level co-occurrence matrix, local binary patterns, and rotation-invariant local phase quantization) had favorably robust performance when using the non-US database. In fact, these methods showed similarity values close to 0 for the acquisition variations and delineations. Results from the US database confirmed robustness for all of the evaluated methods (gray-level co-occurrence matrix, local binary patterns, and rotation-invariant local phase quantization) when comparing the same texture obtained from different regions of the image (proximal/distal lungs and US machine brand stratification).

Conclusions—Our results confirmed that texture analysis can be robust (high similarity for different condition acquisitions) with potential to be included as a clinical tool.

Key Words—image processing; quantitative ultrasound; robustness; texture analysis; ultrasound

Development of noninvasive and effective methods for reporting the pathophysiologic process status is still an elusive goal in modern medicine. Texture analysis methods have been extensively investigated on medical images, as they possess a vast amount of texture information relevant to clinical practice. This phenomenon occurs because medical images contain physical properties of tissues; the signal producing the image changes according to modifications of the tissue microstructure and...
composition. Texture analysis methods allow quantification of these subtle changes in the image.\(^1\)

Over the years, a large number of powerful texture-based methods have been developed thanks to improvements in computation capacity and image resolution.\(^2-4\) Specifically, texture analysis in ultrasound (US) images extracts information related to the speckle characteristics of the image. Oosterveld et al\(^5\) showed the close relationship between speckle and the “density” of the US scatter within a medium. In that study, Oosterveld et al\(^5\) suggested that US texture analysis could quantify the effective number density of tissues, as well as pathologic changes of this parameter. Thus, the principle goal of applying US texture analysis is to characterize the speckle variation between US images to distinguish those tissues altered as a consequence of the disease.

The ability of texture-based methods for extracting relevant texture features from medical US images and quantifying subtle changes in human tissues, which are nonvisible to the human eye, has been widely demonstrated.\(^6-8\) One of the first studies based on US texture analysis\(^7\) presented a perspective on tissue characterization features for extracting diagnostic information. Later, Tunis et al\(^8\) corroborated that textural information in US images is related to pathophysiologic processes. Thus, the potential clinical application of quantitative US texture analysis has been investigated in different medical fields.\(^9-12\)

Sujana et al\(^13\) used US texture analysis and classification methods for characterizing certain liver lesions, Chen et al\(^10\) for classification of breast tumors, and Vince et al\(^14\) for characterizing coronary plaques. In the fetal-maternal field, US texture analysis was introduced to evaluate associations of brain textures with neurobehavioral outcomes in preterm neonates.\(^12\)

Research in other quantitative US-based techniques reasserts a clinical trend in obtaining information related to the tissue microstructure, taking advantage of its acoustic properties. These techniques include elastography, flow estimation by Doppler imaging, shear wave imaging, spectral-based parameterization of US signals, and envelope statistics.\(^15,16\) Although some of these techniques have shown promising results for diagnostic purposes, most of them require specific devices and training for their integration into a clinical setting.\(^16\) We introduce quantitative US texture analysis as a technique that might be easily implemented into clinical practice, as it might provide valid information from standard US.

Up to the present, most of the studies have applied texture-based methods as part of a classification system, in which US texture features fed the classifier, evaluating its performance to predict the clinical outcome.\(^17,18\) There have been few application-oriented studies aimed at evaluating the relative powers of the texture-extractor methods before any classification or retrieval system. In fact, none of them have considered whether US texture features are robust enough (ie, repeatable regardless of different image acquisition parameters, such as illumination and resolution) to be used in a clinical setting. In particular, many have used a huge number of US images of the same tissue acquired under different conditions. It is worth considering the idea that speckle characteristics may be affected by different acquisition conditions, including but not restricted to those induced by operators, biological samples, and US system settings. Some quantitative US-based approaches have attempted to characterize pathologic tissues in a robust way,\(^19-22\) but these require complicated acquisition protocols to provide repeatable acquisition conditions for replicating the results. Furthermore, there are new texture-based methods that have not been widely applied for characterizing US texture in the literature,\(^23,24\) even though they might be useful because they compute local textural features related to local information.\(^25\) Finally, a fundamental step in the use of texture-based methods is the region of interest (ROI), which identifies the region of the image that corresponds to the piece of tissue that will be analyzed. Most studies overlook this step when evaluating texture analysis, although it is a fundamental step, as delineation (selection of the ROI) would be performed by different operators and, therefore, will be different each time. This factor might also affect the robustness of the specific textural features. For all of the above, a robustness assessment of variations in the US acquisition conditions and delineations of same type of tissue would represent a step forward in the exploration of the use of quantitative US texture analysis for clinical purposes.

We aimed to compare, rank, and validate the robustness of several texture-based methods to evaluate their feasibility as texture feature extractors in US images for use as clinical tools. Particularly, we
compared methods that compute local information. We included those methods most commonly found in the literature for US texture classification and newer methods as alternatives. To evaluate the methods, we acquired different US images of the same texture acquired under different conditions. Nevertheless, 2 main limitations were observed: (1) Not all parameters can be modified through the whole range when scanning real textures because of clinical limitations. For instance, different degrees of US wave absorption exist when crossing distinct tissues such as fat or bone, causing acoustic shadows; sometimes these artifacts cannot be avoided when the organ of interest is fixed and distant to the transducer (fetal evaluation). (2) It is not possible to change acquisition parameters in a precise and controlled way, especially because of operator variability when positioning the transducer. Thus, we decided to use an approach inspired by the image quality transfer process, which first selects and configures the methods by using images obtained from a different source but that are easier to acquire in a controlled setting, and later, the method is refined by using real images. Concretely, we used 2 sets of images for this study: (1) a controlled set of images, consisting of available non-US images acquired under controlled acquisition parameters (ie, illumination and rotation angle), emulating the acquisition conditions of a medical US setting and thus evaluating a huge number of images for each texture; and (2) a US image set comprising US images of fetal lungs acquired under similar conditions as those of a clinical setting. Hence, different texture-based methods were compared and ranked by using the controlled sample set, and the most robust methods were validated by using clinically acquired US images of fetal lungs.

Materials and Methods

In this section, we briefly describe both the image data sets and their characteristics (image acquisition and image labeling) to determine which information related to acquisition conditions was evaluated. We also describe the ROIs to evaluate the robustness when different regions of the same tissue are delineated. Then, we introduce the texture-based methods and the metrics used to compare, rank, and validate the robustness of the methods for acquisitions and delineations. Finally, we describe the experiments’ designs used in this study.

Data Sets

Controlled Sample Set

Images with different textures were obtained from widely known available databases that previously have been used for testing classification methods: OUTEX (University of Oulu, Oulu, Finland) and PHOTEX (Texture Lab, Heriot-Watt University, Edinburgh, Scotland). These databases provide pictures of the same texture acquired under different conditions, varying illumination, spatial resolution, and rotation parameters and thus emulating the differences between US textures when acquired under different conditions in a controlled way. Three parameters whose changes might affect US speckle patterns and that might be indirectly adjusted by the radiologist when performing US scanning were used: (1) illumination, which is related to the gain parameter or image contrast and possible attenuation of the acoustic wave that has to cross different tissues until arriving at the desired tissue to be analyzed; we also used illumination for the US system’s color maps, which can be different for different systems, since it is inherent to the US system; (2) spatial resolution, which is related to the frequency, depth, zoom, and aperture of the transducer; and (3) rotation, which is determined by the unpredictable position of the organ and the transducer when performing a scan.

Clinical US Images

Fetal lung US images were acquired from patients with singleton pregnancies attending the Maternal-Fetal Medicine Department at the Hospital Clinic in Barcelona for routine pregnancy US scans. Multiple pregnancies and structural/chromosomal anomalies were excluded from the study. Ultrasound images of the same lung tissue acquired under different conditions were not available for all patients, since it was not feasible to acquire images with the whole range of acquisition parameters in a precise and controlled way. The study protocol was approved by the local Ethics Committee (approval number 3823-2007), and pregnant women provided written informed consent.
Image Acquisition and Labeling
Each data set was acquired and labeled as follows:

Controlled
The OUTEX and PHOTEX databases were downloaded from the links specified by Hossain and Serikawa.27 For the purpose of this study, only those textures that could be similar to the US patterns (ie, granulated, dotted, and flecked) were selected by visual inspection. An example of the selected textures is shown in Figure 1. Additionally, only those images that showed similar histograms as the ones computed from the real US textures were selected (see an example in online supplemental Figure S1). An analysis of variance was conducted to compare the mean, skewness, and kurtosis of the histograms computed from the controlled and clinical data sets. All images were digitally stored in portable network graphic and tagged image file formats and converted to grayscale values within a range between 0 and 255 values. Then, texture images were labeled according to controlled acquisition parameters.

A total of 69 textures were selected and labeled from the OUTEX database, in a total of 11,178 images. Specifically, OUTEX textures were labeled according to different illuminations (horizon, inca, and TL84) that emulate differences in the gain and US system’s color maps used for US image representation, resolution levels (100, 120, 300, 360, 500 and 600, dots per inch), and rotation degrees (0°, 50°, 100°, 150°, 300°, 450°, 600°, 750°, and 900°), obtaining 162 images per texture. Changes in resolution and rotation degrees emulate different acquisition conditions between US images due to frequency, depth, and organ position changes. Regarding the PHOTEX database, a total of 1,993 images were labeled from 34 tissues selected for the purpose of this study. The PHOTEX database images were labeled according to rotation degrees and tilt angles of illumination, since they emulate changes in US textures due to the transducer or organ position when insonating an organ. The acquisition parameters (rotation and tilt illumination) were controlled but differed for each texture.

Clinical
Ultrasound images of fetal lungs were acquired in an axial section of the fetal thorax at the level of the cardiac 4-chamber view. Acquisition settings such as gain, zoom, frequency, and time-gain compensation were not fixed and were adjusted according to clinical criteria. The depth and aperture of the transducer were adjusted to magnify the fetal thorax so that the thorax occupied about two-thirds of the screen. The aperture might change for each US system, each operator, and the unpredictable position of the fetus during acquisition. Changes in the aperture and frequency are related to changes in spatial resolution (see its distribution in online supplemental Figure S2). Scans were performed by certified...
radiologists using Sonoline Antares (Siemens Medical Solutions, Malvern, PA), Voluson 730 Pro and Voluson 780 Pro (GE Healthcare, Milwaukee, WI), ProSound α7 (Hitachi Aloka Medical, Ltd, Tokyo, Japan), and Aplio (Toshiba Medical Systems, Tokyo, Japan) US systems. All machines were equipped with curved linear transducers with a frequency range of 3 to 7.5 MHz. All images were collected digitally in the original Digital Imaging and Communications in Medicine format and then inspected by 2 investigators (A.P.-M. and E.B.-C) for image quality control. Images were considered noneligible if the fetal thorax occupied less than two-thirds of the screen or if color Doppler modes, calipers, or pointers were used. Furthermore, images were excluded if they had any of the following characteristics, as they can directly alter the values of the US features: presence of obvious acoustic shadows from the fetal ribs, saturation, and any type of postprocessing (such as smoothing). Image quality control was done manually, assisted by an ad-hoc graphical user interface that computed the proportion of the fetal thorax in the image by semiautomatically delineating an ellipse over the thorax, showed images to check the use of calipers, color Doppler imaging, or any type of postprocessing, and plotted acoustic shadows in green and saturated regions in red (pixel values close to 0 and 255, respectively).

A total of 713 US images were acquired from 385 fetuses. Forty-seven images were discarded, resulting in 666 useful images from a total of 355 patients after image quality control. Images were labeled according to the rotation angle, fetal spine position (left or right), and proximal lung (lung close to the transducer) as left or right. The same graphical user interface developed for image quality control was used to label the fetal lungs. By means of the graphical user interface, a clinical expert (F.M.) semiautomatically calculated the rotation angle, indicating the orientation of the fetal spine with respect to the atrioventricular bundle of the heart (see rotation angle distribution in online supplemental Figure S3). Additionally, the clinical expert also indicated the fetal spine position and the proximal lung as defined above. The same graphical user interface was used for delineation.

**Image Delineation**

Once images were labeled, different delineations were performed in each image for each data set:

![Nonoverlapped ROI divisions (dotted lines) of a texture image in the controlled setting (texture from the PHOTEX database). The original image is divided into 25 different ROIs.](image-url)
Controlled
An automatic delineation was performed for each texture image, considering 25 nonoverlapped and 28 overlapped but with different-size ROIs. In this manner, different regions of the same texture were evaluated, as shown in Figures 2 and 3, respectively.

Clinical
Two operator-dependent delineations of both fetal lungs were considered, manual and semiautomatic ROIs, which were performed by a clinical expert (F.M.; Figure 4). Manual delineations included the largest possible homogeneous area of the fetal lung, avoiding the heart, gross vessels, and surrounding areas. Semiautomatic delineations were performed, indicating a fixed-size square region, following the same criteria as for manual delineations. After the operator-dependent delineations were performed, smaller ROIs were created automatically, eroding the manual and semiautomatic delineations repeatedly (Figure 5) until reaching the limit of 100 pixels for the smallest ROI.

Figure 3. Twenty-eight overlapped (with different sizes) ROIs of a texture image in the controlled setting (texture from the PHOTEX database).
Texture-Based Methods
The texture-based methods used for this study are expected to be able to extract grayscale, multiresolution, and rotation-invariant local features from US images, as robustness in relation to these characteristics will be required for their use in a clinical application. Additionally, the number of textural features obtained by each method should not be dependent on the ROI size or location within the same type of tissue. Textural image features were computed by several texture-based methods, which are widely known for texture classification in the computer vision field.\(^2\)\(^-\)\(^4\),\(^23\),\(^24\) For each texture-based method, different sets of textural features were extracted for each ROI and image. The texture-based methods used are detailed below (see a summary of the texture-based methods in online supplemental Table S1):

Gray-Level Co-occurrence Matrix
The gray-level co-occurrence matrix (GLCM) has been widely used to characterize textures in US images.\(^30\),\(^31\) This method counts pairs of horizontally adjacent pixels in a grayscale version of the image, as defined by Haralick and Shanmugam.\(^2\) Characteristics of the features extracted by this method are described in detail elsewhere.\(^2\) In our experiments, 1 adjacency direction (0°) and 8 gray levels when scaling the grayscale values in the image were used to compute the GLCM. Thus, there were 64 possible ordered combinations of values for each pair of pixels corresponding to the final 64 textural features.

Local Binary Patterns
The local binary patterns (LBP) method has been recently applied for texture characterization in US images.\(^32\),\(^33\) This method computes the distribution of binary patterns in the circular neighborhood of each pixel, which is characterized by a radius (R) and a number of neighbors (P). The principle is to threshold neighboring pixels compared with the central pixel. Thus, for each pixel a binary pattern is obtained. An LBP code at pixel \( p \) is computed by the scalar

Figure 4. Manual (A) and semiautomatic (B) delineations of the proximal (1) and distal (2) lungs. Ultrasound scans of fetal lungs in cardiac 4-chamber views at 37 weeks’ gestation.
product between the binary pattern and a vector of powers of 2:

$$LBP(p) = \sum_{i=0}^{P-1} 2^i \delta(f(q_i) - f(p)), \quad (1)$$

where $f(q_i)$ and $f(p)$ are gray levels of pixels $q_i$ and $p$, respectively, and $\delta$ is the Kronecker function. Then, the histogram of the LPB is used as texture features. The LPB method has some variants that have been widely used as texture features for medical images. In particular, we worked with the multiresolution grayscale and rotation-invariant approach based on recognizing those binary patterns that occur more often in a texture image than others. These frequent patterns are called uniform patterns and were explained in more detail by Ojala et al. In our study, uniform patterns were defined with $P = 16$ equally spaced pixels on a circle of radius $R = 1$, resulting in 18 specific texture features.

**Histogram of Oriented Gradients**

The histogram of oriented gradients (HOG) might obtain information about the anisotropy of a texture to determine the predominant directions of a texture. Recent studies have applied HOG to characterize textures in US images. Up to the present, however, the main purpose of applying this method on US images has been macrostructure detection such as nuchal translucency or motion estimation. We decided to include the HOG method in our study, since it may provide useful information related to tissue histologic characteristics. The HOG method counts frequencies of gradient orientation values in localized portions of an image. The gradient orientation is estimated at every pixel, and a histogram is computed to tell how often the respective gradient direction is present in the image. The specific textural features computed by this method were explained by Junior et al. For this study, each image to be analyzed (ROI) was divided in $3 \times 3$ cells (or portions) of the same size, and the number of histogram bins was $N_b = 9$, obtaining 81 textural features.

**Local Phase Quantization**

The Local phase quantization (LPQ) computes quantized phase information of the discrete Fourier transform, but it has not been extensively applied in texture classification for medical images and even less for characterizing US textures. It uses the local phase information extracted by a short-term Fourier transform computed over a rectangular $M \times M$ neighborhood $N_p$ at each pixel position $p$ of the image $f(p)$. The way of obtaining the features was explained in more detail by Ojansivu and Heikkilä. The same number of specific textural features is always computed, obtaining a total of 256 features for this study.

**Rotation-Invariant Local Phase Quantization**

The rotation-invariant local phase quantization (riLPQ) acronym corresponds to the rotation-invariant approach derived from the LPQ method. The riLPQ method compensates for the rotation of the image that has to be analyzed, considering the direction of the characteristics in the examination of the local phase. In this manner, the final textural features extracted should be the same regardless of the image rotation. For more detail, the specific features computed by this method were

---

**Figure 5.** Eroded ROIs from manual (A) and semiautomatic (B) delineations of the fetal lungs (clinical data set). Original ROIs from a distal/proximal lung at 37 weeks’ gestation.
described by Ojansivu et al. A total of 256 features are obtained by this method.

**Similarity Measurements/Metric Distances**

Robustness was evaluated and validated by measuring the similarity (or dissimilarity) between 2 sets of specific textural features extracted from 2 images of the same texture acquired under different conditions or the same conditions (the same image) with different ROIs. We used correlation and Chebyshev distances to compare the texture features because they provide different similarity information, which might be useful to construct a classification algorithm when developing a clinical application.

The correlation distance measures the similarity between the relative shapes of the 2 features sets. This distance is defined as a measure of statistical dependence between 2 random sets of features. In our study, the scale of correlation similarity values was inverted for comparison purposes. Consequently, a lower distance indicated more similarity (robustness); if the features were dependent, this measure was 0. Conversely, the features were independent when this measure was 1. The correlation distance ($D_{CR}$) used in this study can be expressed as

$$D_{CR}(X, Y) = 1 - \frac{\sum_{i=1}^{n} (X_i - X')(Y_i - Y')}{\sqrt{\sum_{i=1}^{n} (X_i - X')^2 \sum_{i=1}^{n} (Y_i - Y')^2}}$$

where $X = \{X_0, X_1, \ldots, X_{n-1}\}$ and $Y = \{Y_0, Y_1, \ldots, Y_{n-1}\}$ are the feature vectors extracted from images acquired under different conditions or different delineations considered statistically independent.

The Chebyshev distance ($D_{CH}$) measures the similarity between absolute values. In this study, we normalized distance between 0 and 1 for comparison purposes; in this manner, 2 sets of features were similar (robust) if the distance was close to 0 or not (distance close to 1). This similarity measurement can be expressed as

$$D_{CH}(X, Y) = \max_{0 \leq i \leq n} \left\{ \frac{|X_i - Y_i|}{\max\{D_{CH}(X, Y)\}} \right\}$$

where $X = \{X_0, X_1, \ldots, X_{n-1}\}$ and $Y = \{Y_0, Y_1, \ldots, Y_{n-1}\}$ are the feature vectors extracted from images acquired at different conditions or different delineations.

**Experiments**

Experiments were designed by following a similar approach as the image quality transfer method. First, the controlled sample set was used to determine reference values for comparison purposes when using correlation and Chebyshev distances. Concretely, the 3 best methods were selected, and then reference values for correlation and Chebyshev distances were determined. Once methods were selected, we evaluated the robustness of the selected methods using the clinical sample set by comparing the results with the measures previously obtained. A summary of the experiments, including the number of images for both sample sets, is displayed in Figure 6. The texture-based methods (GLCM, LBP, HOG, LPQ, and riLPQ) were ranked according to the robustness assessed with the controlled sample set. Then, only those methods that showed better robustness were validated with the clinically acquired US images. The experiments are explained in more detail below.

**Texture-Based Method Ranking Using the Controlled Sample Set**

For each texture and texture analysis method, the similarity measures (correlation and Chebyshev distances) were computed from the controlled databases (OUTEX and PHOTEX). The robustness in relation to each acquisition parameter was assessed; the parameter of interest was not fixed to any value, whereas the rest of the acquisition parameters were fixed, resulting in different acquisition scenarios. Then, both similarity measures were computed between the different textural features of the same texture acquired at different settings of the same parameter of interest. In this manner, the robustness of each acquisition parameter was isolated. This procedure was repeated for each parameter of interest until all of the acquisition parameters were unfixed once. Finally, to summarize the robustness of each acquisition parameter and texture, the mean and standard deviation were computed over fixed parameters (different scenarios) for each similarity measurement, resulting in a unique value (Mean ± Std). For instance, to assess illumination robustness using OUTEX database samples (illumination had horizon, inca, and TL84 labels), the resolution and rotation were fixed, resulting in a total of 54 scenarios (6 resolution levels and 9 rotation degrees) for each texture.
(Figure 7). Then, the mean and standard deviation were computed for each similarity measurement over the 54 scenarios. In this example, a total of 3 similarity values (Mean ± Std) from 2 similarity measures for 3 different labels were obtained for each texture. To compare the robustness of the texture-based methods for each acquisition parameter, for each similarity measure, the mean among similarity values was computed for each texture and then among all textures. In this manner, a unique value for each similarity measure, acquisition parameter, database (OUTEX and PHOTEX), and texture-based method was obtained.

The same approach was used to assess robustness regarding the different delineations; similarity measures were computed for the overlapped but different-size ROIs and the nonoverlapped ROIs delineated in the same texture image. The mean and standard deviation were computed over overlapped and nonoverlapped delineations for each similarity measure, resulting in a unique value for each texture image. Then, the robustness in relation to nonoverlapped and overlapped delineations was compared between the different texture-based methods, computing the mean among similarity values (Mean ± Std) for each similarity measure and each texture and then among all selected textures. A unique similarity value was obtained for each similarity measure, the nonoverlapped and overlapped delineations, each database, and each texture-based method.

Those texture-based methods that showed lower similarity values with regard to acquisition parameters and delineations were considered the most robust methods. Based on this criterion, methods were ranked from the most to the least robust in relation to acquisitions and delineations for each database (OUTEX and PHOTEX) first. Then, each texture-based method was globally ranked according to the
Figure 7. Flowchart of the robustness evaluation in relation to an acquisition parameter using a texture (from the OUTEX database) acquired under different illumination conditions as an example. For each similarity measurement, (Chebyshev and correlation), a mean similarity value (Mean ± Std) in relation to illumination is obtained for texture T and each texture-based method (z = 1 ... 5). Then, for each similarity measurement, the mean among all textures will be computed, obtaining a unique value for illumination and each method.

```
Meanz = mean( Meanp=1, Meanp=2, Meanp=3 )
Stdz = std( Stdp=1, Stdp=2, Stdp=3 )
```

"A total of Z unique values [meanz, stdz] are obtained for texture T and the unixed parameter"
number of times it ranked the best. The first 3 methods were selected for validation using clinical images.

*Validation of the Robust Methods Using the Clinically Acquired US Images*

The robustness of those methods that obtained better results when using the controlled sample set was validated by using fetal lung US images. Different experiments were performed as detailed below.

First, we assumed that the left and right lungs of the same patient have the same type of tissue, and in consequence, images of both lungs acquired under different conditions should show the same or similar textural features. Based on this assumption, the robustness in relation to illumination, resolution, and rotation was indirectly validated by computing similarity measurements between proximal and distal lungs that were at different depth positions. Different illumination and resolution conditions in the same tissue were indirectly achieved, since the lateral speckle size is strongly dependent on the depth within the tissue, and acoustic attenuation is dependent on depth.\(^5\),\(^40\) The robustness in relation to rotation was also assessed by using the fetal lung US images acquired with different fetal spine orientations. In this manner, the same US tissue at different rotation conditions (with respect to proximal and distal lungs) was achieved. For each texture-based method, the mean and standard deviation of the correlation and Chebyshev distances were computed among US fetal lung images for manual and semiautomatic delineations.

Second, to validate the robustness dependence of the selected texture-based methods on US systems, robustness results for illumination, resolution, and rotation were stratified for the different US system brands used in our clinical setting. No dependence on systems was considered when similar robustness was obtained between US systems of different brands. An analysis of variance was conducted over the stratified values (Siemens, GE, Toshiba, and Aloka).

Finally, the robustness in relation to different delineations was assessed for each texture-based method. Similarity measurements were computed between the eroded ROIs from the manual and semiautomatic delineations. The mean and standard deviation of the similarities were computed among all of the proximal and distal lungs for each method and the manual and semiautomatic delineations. All computations in this study were performed with MATLAB R2014b (version 8.4.0.150421; The MathWorks, Inc, Natick, MA).

**Results**

*Selection of Non-US Images (Controlled Data Set)*

No significant differences were shown between the mean, skewness, and kurtosis of the histograms computed from the selected non-US images and the histograms computed from the fetal lung US textures.

*Texture-Based Method Ranking*

Similarity results are presented as mean (standard deviation). Similarity results between features extracted from each texture acquired at different illumination, resolution, and rotation labels are given in Table 1. Regarding the OUTEX database, most methods showed high robustness when the illumination acquisition parameter was left free (horizon, inca, and TL84). For illumination in the PHOTEX database, the GLCM, LBP, and riLPQ texture-based methods showed more robustness in comparison with the rest of the methods (HOG and LPQ). Specifically, the HOG and LPQ methods resulted in correlation distances of 0.36 (0.15) and 0.29 (0.16), respectively. The GLCM, LBP, and riLPQ methods were the most robust methods for resolution and rotation parameters stratified in the OUTEX database, whereas the HOG and LPQ methods performed poorly for these parameters. The HOG and LPQ methods presented less robustness for rotation in the PHOTEX database than the other methods as well.

Similarity results for different delineations in the OUTEX and PHOTEX databases are given in Table 2. The HOG and LPQ methods were the worst in terms of robustness for different delineations from both databases. Maximum similarity values between textural features extracted by HOG and LPQ in different overlapped ROIs were 0.32 (0.13) and 0.42 (0.21), respectively, and they were 0.27 (0.13) and 0.42 (0.30) for the nonoverlapped ones. On the other hand, LBP and riLPQ performed better for the nonoverlapped delineations than the other methods.
Overall, the robustness performance for the GLCM, LPB, and riLPQ texture-based methods was favorable compared with HOG and LPQ. In fact, these methods showed similarity values close to 0 for the acquisition variations in almost all acquisition parameters and delineations from both controlled databases (OUTEX and PHOTEX). Table 3 shows the ranking of the robustness of the texture-based methods in relation to acquisition conditions and delineations for each data set.

**Validation of the Robust Methods**

Table 4 gives similarity results between proximal and distal lungs of all images. Overall results confirmed robustness for all of the evaluated methods (LBP, riLPQ, and GLCM) depending on the similarity measure and the 2 operator-dependent delineations (manual and semiautomatic). The highest similarity was shown for the riLPQ method using the manual delineation, but overall, the LBP method performed the best. The GLCM method was the worst in terms of robustness when using semiautomatic delineations and measuring the correlation distance, although the Chebyshev distance was close to 0.

Stratified results by US brand are shown in Table 4. A total of 198, 392, 56, and 20 fetal lung US images were acquired with the Siemens, GE, Toshiba, and Aloka US systems, respectively. Similar results were shown when comparing robustness stratified by US brands. Results demonstrated that variations in indirect illumination, resolution, and rotation were not dependent on the US system. No significant differences (P > .05) were found for the GLCM, LBP,

| Database | Method | Corr     | Cheb     | Corr     | Cheb     | Corr     | Cheb     |
|----------|--------|----------|----------|----------|----------|----------|----------|
| OUTEX    | GLCM   | 0.02 (0.02) | 0.17 (0.03) | 0.01 (0.01) | 0.06 (0.03) | 0.00 (0.00) | 0.02 (0.01) |
|          | LBP    | 0.02 (0.01) | 0.07 (0.02) | 0.06 (0.02) | 0.11 (0.02) | 0.00 (0.00) | 0.02 (0.01) |
|          | HOG    | 0.04 (0.03) | 0.05 (0.02) | 0.15 (0.06) | 0.11 (0.02) | 0.39 (0.11) | 0.17 (0.04) |
|          | LPQ    | 0.04 (0.03) | 0.02 (0.01) | 0.21 (0.09) | 0.04 (0.01) | 0.20 (0.11) | 0.04 (0.02) |
|          | riLPQ  | 0.01 (0.01) | 0.01 (0.00) | 0.17 (0.01) | 0.03 (0.01) | 0.01 (0.01) | 0.01 (0.01) |
| PHOTEX   | GLCM   | 0.01 (0.01) | 0.13 (0.10) | 0.03 (0.01) | 0.15 (0.05) | 0.02 (0.01) | 0.11 (0.02) |
|          | LBP    | 0.01 (0.01) | 0.02 (0.01) | 0.02 (0.02) | 0.06 (0.02) | 0.37 (0.11) | 0.36 (0.06) |
|          | HOG    | 0.36 (0.15) | 0.23 (0.11) | 0.43 (0.07) | 0.10 (0.02) | 0.17 (0.05) | 0.06 (0.02) |
|          | LPQ    | 0.29 (0.16) | 0.04 (0.02) | 0.20 (0.11) | 0.04 (0.02) | 0.20 (0.11) | 0.04 (0.02) |
|          | riLPQ  | 0.03 (0.02) | 0.03 (0.01) | 0.03 (0.04) | 0.02 (0.01) | 0.03 (0.02) | 0.03 (0.03) |

Data are given as mean (SD). Lower distances (values close to 0) indicate similarity (robustness), and higher distances (values close to 1) indicate dissimilarity. Cheb indicates Chebyshev distance; and Corr, correlation distance.

| Database | Method | Overlapped | Nonoverlapped |
|----------|--------|------------|---------------|
|          | Corr   | Cheb       | Corr           | Cheb           |
| OUTEX    | GLCM   | 0.02 (0.05) | 0.01 (0.00) | 0.05 (0.11) | 0.01 (0.01) |
|          | LBP    | 0.02 (0.03) | 0.06 (0.04) | 0.02 (0.03) | 0.06 (0.05) |
|          | HOG    | 0.18 (0.15) | 0.32 (0.13) | 0.19 (0.17) | 0.27 (0.13) |
|          | LPQ    | 0.25 (0.11) | 0.06 (0.03) | 0.17 (0.15) | 0.04 (0.04) |
|          | riLPQ  | 0.10 (0.04) | 0.04 (0.02) | 0.03 (0.04) | 0.02 (0.01) |
|          | GLCM   | 0.01 (0.04) | 0.01 (0.00) | 0.21 (0.30) | 0.03 (0.03) |
|          | LBP    | 0.01 (0.02) | 0.03 (0.02) | 0.02 (0.06) | 0.03 (0.03) |
|          | HOG    | 0.13 (0.17) | 0.31 (0.15) | 0.12 (0.13) | 0.23 (0.13) |
|          | LPQ    | 0.42 (0.21) | 0.04 (0.02) | 0.42 (0.30) | 0.03 (0.02) |
|          | riLPQ  | 0.09 (0.05) | 0.05 (0.03) | 0.04 (0.05) | 0.04 (0.02) |

Data are given as mean (SD). Lower distances (values close to 0) indicate similarity (robustness), and higher distances (values close to 1) indicate dissimilarity. Cheb indicates Chebyshev distance; and Corr, correlation distance.
...and riLPQ texture-based methods after stratifying by US brands.

Similarity results between textural features extracted from different ROIs are given in Table 5. Mean similarity values were computed among all proximal and distal lungs. Results confirmed the robustness in relation to delineations for all selected methods evaluated in the controlled setting (LBP, riLPQ, and GLCM).

**Discussion**

This study provides evidence that texture analysis can be used to extract robust information from US images acquired under different conditions. This finding supports the use of texture analysis to obtain valid features from US images, which is required for using those features for clinical purposes in classification or grading systems.

Different quantitative US-based techniques have been explored to extract information from the signals causing speckle that are associated with the underlying tissue microstructure. These techniques have shown promising results, such as transient elastography for the staging of liver fibrosis, spectral-based quantitative US parameters for characterizing breast cancer and detecting the response of breast cancer to therapy, and, most recently, shear wave elasticity imaging for the assessment of cervical softening. Some of these techniques are implemented on specific devices and have been shown to be invariant to different operators and systems. Despite this finding, some of them have not been capable of detecting...

### Table 3. Ranking of the Texture-Based Method Robustness

| Method | General Ranking | Acquisition Conditions | Delineations |
|--------|-----------------|------------------------|--------------|
|        |                 | OUTEX | PHOTEX | OUTEX | PHOTEX |
| LBP    | 1               | 2     | 1      | 2     | 1      |
| riLPQ  | 2               | 1     | 2      | 3     | 2      |
| GLCM   | 3               | 3     | 3      | 1     | 3      |
| LPQ    | 4               | 4     | 4      | 4     | 5      |
| HOG    | 5               | 5     | 5      | 5     | 4      |

Methods are ranked from the most (1) to the least (5) robust in relation to acquisitions and delineations for each database (OUTEX and PHOTEX).

### Table 4. Similarity Results for Proximal and Distal Fetal Lungs in the Clinical US Set

| Method | US Brand | Manual | Semiautomatic | Semiautomatic |
|--------|----------|--------|---------------|---------------|
|        |          | Corr   | Cheb          | Corr          | Cheb          | P     |
| LBP    | All      | 0.05 (0.04)* | 0.13 (0.10)* | 0.07 (0.06)* | 0.04 (0.04)* | .5971 |
|        | Siemens  | 0.04 (0.03) | 0.11 (0.09) | 0.05 (0.04) | 0.04 (0.04) |       |
|        | GE       | 0.06 (0.05) | 0.15 (0.11) | 0.08 (0.06) | 0.04 (0.04) |       |
|        | Toshiba  | 0.08 (0.05) | 0.07 (0.04) | 0.09 (0.05) | 0.02 (0.02) |       |
|        | Aloka    | 0.06 (0.05) | 0.04 (0.02) | 0.08 (0.07) | 0.01 (0.01) |       |
| riLPQ  | All      | 0.00 (0.00)* | 0.14 (0.11)* | 0.17 (0.23)* | 0.23 (0.18)* | .9956 |
|        | Siemens  | 0.00 (0.01) | 0.16 (0.12) | 0.12 (0.21) | 0.22 (0.16) |       |
|        | GE       | 0.00 (0.00) | 0.13 (0.11) | 0.20 (0.25) | 0.24 (0.19) |       |
|        | Toshiba  | 0.00 (0.00) | 0.11 (0.10) | 0.17 (0.22) | 0.23 (0.17) |       |
|        | Aloka    | 0.00 (0.00) | 0.22 (0.14) | 0.06 (0.14) | 0.27 (0.22) |       |
| GLCM   | All      | 0.16 (0.14)* | 0.14 (0.10)* | 0.51 (0.31)* | 0.04 (0.03)* | .9656 |
|        | Siemens  | 0.12 (0.12) | 0.14 (0.12) | 0.36 (0.27) | 0.04 (0.03) |       |
|        | GE       | 0.19 (0.14) | 0.15 (0.10) | 0.59 (0.30) | 0.04 (0.03) |       |
|        | Toshiba  | 0.13 (0.11) | 0.09 (0.04) | 0.44 (0.29) | 0.03 (0.01) |       |
|        | Aloka    | 0.19 (0.11) | 0.05 (0.04) | 0.48 (0.28) | 0.02 (0.01) |       |

Data are given as mean (SD). Lower distances (values close to 0) indicate similarity (robustness), and higher distances (values close to 1) indicate dissimilarity. Cheb indicates Chebyshev distance; and Corr, correlation distance.

*Results for all images.
specific diseases that are still prevalent in the general population, perhaps because their approaches are inadequate and are not able to obtain relevant information from any tissue. Quantitative US texture analysis might become a new clinical tool that might provide new insight for clinical diagnosis.

Several attempts have been made to obtain clinical information related to a pathophysiologic process by using quantitative US texture analysis in a robust way. Oosterveld et al analyzed the texture of B-mode images to differentiate diffuse liver diseases and evaluated its reproducibility over a 5-day period. In that study, the B-mode images were reconstructed by radiofrequency signals that were corrected by attenuation to remove the depth. Results showed the possibility of correcting the depth dependencies of the B-mode texture. Garra et al used quantitative analysis of US image texture to distinguish benign from malignant breast lesions, showing promising results. Nonetheless, Garra et al concluded that the method had US system dependence. Previous methods showed promising results but not feasibility for clinical practice. Other studies showed high diagnostic accuracy for detection of subtle changes in affected tissues that were nonvisible to the human eye. However, no perspective studies have been conducted to validate the robustness of the methods in a clinical setting.

To our knowledge, this work is the first study reporting accurate robustness of quantitative US texture analysis considering only the specific textural features and not the prediction rate for a clinical event, using machine-learning algorithms. The main difference between this study and the previous ones is that the robustness of US texture features was assessed by using a large number of controlled (nonmedical) images. The data sets used in this study emulated US acquisition conditions, which are usually present in a clinical setting. Additionally, several ROIs were used to assess robustness when delineating. Our study showed that the LBP, rILPQ, and GLCM methods were the 3 most robust methods for extracting information from images acquired under different conditions and different delineations in the controlled setting (Tables 1–3). It should be noted that the LBP and rILPQ methods were the most robust in both databases (OUTEX and PHOTEX). These methods have not been widely used for US texture classification in the literature. Thus, this finding opens the possibility of exploring new methods to develop US texture-based tools. Then, the most robust methods (LBP, rILPQ, and GLCM) were validated by using clinically acquired US images acquired by several US machines and operators. Our results validated robustness in relation to acquisition conditions using LBP, rILPQ, and GLCM and showed these methods to be invariant to US machines (Table 4). Concretely, LBP performed the best; the rILPQ and GLCM methods showed low similarity values in relation to acquisitions according to the delineation mode (manual or semiautomatic) and the similarity measure (correlation or Chebyshev). Robustness against multiple delineations was also validated by using clinically acquired US images. All methods resulted in low similarity values according to the delineation mode or the similarity measure (Table 5). These results confirm that a texture-based tool that integrates a classification system could be developed by using any of the tested methods.

Even though 3 of all of the texture-based methods (LBP, rILPQ, and GLCM) showed robustness when using clinically acquired US images, the use of these methods to develop a clinical tool needs to be demonstrated. Our results do not provide evidence of the suitability of these methods to assess pathophysiologic conditions involved in most tissues; that will depend on the intrinsic properties of textural

| Method | Manual | Semiautomatic |
|--------|--------|---------------|
|        | Corr   | Cheb          | Corr   | Cheb          |
| LBP    | 0.08 (0.05) | 0.05 (0.01) | 0.23 (0.09) | 0.02 (0.00) |
| rILPQ  | 0.00 (0.00) | 0.15 (0.08) | 0.02 (0.03) | 0.23 (0.11) |
| GLCM   | 0.03 (0.04) | 0.05 (0.01) | 0.12 (0.09) | 0.01 (0.00) |

Data are given as mean (SD). Lower distances (values close to 0) indicate similarity (robustness), and higher distances (values close to 1) indicate dissimilarity. Cheb indicates Chebyshev distance; and Corr, correlation distance.
features extracted by each texture analysis method. In fact, a method that always gives the same values will be the most robust method but completely useless. Additionally, robustness was assessed in the controlled setting over all acquisition conditions discretely and not considering specific ranges. In some cases, depending on the organ to be scanned (ie, carotid artery or fetal heart), acquisition protocols might include repeatable acquisitions with acquisition parameters fixed within particular ranges. Therefore, the discarded texture-based methods might obtain repeatable features within specific ranges and provide useful information related to the underlying pathophysiologic process. Moreover, it should be noted that robustness was validated by comparing proximal versus distal lungs. The robustness of the methods that showed higher similarities when comparing both fetal lungs would be improved by using a focal configuration and evaluating tissues within the same depth. Hence, when exploring texture US analysis to develop a clinical tool, an acquisition protocol should be designed to obtain the most repeatable acquisitions.

The main strength of our study was that the feasibility of texture analysis to obtain US features in a robust way was tested by using non-US images acquired under controlled conditions similar to US and clinically acquired fetal lung US images. On the one hand, the non-US set provides different images of the same tissue acquired in a very precise way, in contrast to whichever US setting that depends on the ability of the radiologist. This approach opens the possibility of evaluating a higher number of images of the same texture acquired under different conditions than in the theoretical case of evaluating real US images. Furthermore, images were acquired by combining parameters over the whole range, thus emulating possible acquisition conditions of whichever US setting at which textures are scanned from any organ is used. On the other hand, testing the robustness of US texture-based methods using fetal lung US images expands opportunities to explore the same methods for quantifying textural changes in other organs, even in adult scans, in which acquisition conditions might be more repeatable. Another strength of our study was the use of the fetal lung US images to compare the same lung tissue at different depths (proximal and distal fetal lungs). Our results represent a forward step in relation to a study published by Thijssen, which suggested that texture analysis based on second-order statistics should be used in the axial direction exclusively, since the speckle size changes strongly according to depth and attenuation. Finally, several US systems were used to acquire our clinical images. Speckle patterns might be related to the system, since wave propagation fundamentals, such as wavelength and gain, are postprocessed in the system. In our study, we demonstrated that it is possible to configure similar settings in different US systems without affecting the robustness of the selected methods (LBP, rILPQ, and GLCM).

This study had some limitations that should be acknowledged. First, noncontrolled resolution images in the PHOTEX database might have affected the robustness evaluation between nonoverlapped delineations. We believe that nonoverlapped ROIs (of the same image) show different textural content between them when the resolution is high. For instance, the GLCM method resulted in high dissimilarity (correlation distance) only for nonoverlapped delineations in the PHOTEX database (Table 2), in which the resolution was not controlled. Second, we used clinically acquired US images of fetal lungs to validate the robustness of the selected texture-based methods, but only robustness for different lungs (proximal versus distal) and delineations of the same tissue were assessed. In fact, for this study, we assumed that proximal and distal lungs of the same patient would have the same tissue, although, to our knowledge, this assumption has not been previously demonstrated in the literature. Ideally, the robustness evaluation should be performed by using different controlled acquisitions of the same organ and patient. Although different US images of the same patient were acquired in some cases, acquisition conditions were similar, since the images were acquired for clinical purposes with the use of a similar setting. To evaluate robustness for US images acquired under different conditions in a controlled way, a robustness study using different US images of the same tissue (ie, carotid artery or liver in adults) should be performed. Third, this study evaluated the repeatability of specific textural features obtained from images acquired under different conditions and different delineations without demonstrating their ability to detect differences against a clinical outcome of interest. We acknowledge that an additional study to compare the
prediction of a clinical outcome with the same US tissue acquired under different conditions should be performed. Nonetheless, the use of texture analysis to develop a robust clinical tool was recently demonstrated by Palacio et al.\(^4\) In that work, a prospective multicenter study in 20 centers worldwide was undertaken, including a total of 730 samples for the final analysis, different operators, and different US systems. The results showed that quantitative US of fetal lung texture predicted neonatal respiratory morbidity with a sensitivity, specificity, positive predictive value, and negative predictive value of 74.3%, 88.6%, 51.6%, and 95.5%, respectively. These promising results support our findings, suggesting that texture analysis may provide robust and relevant information that could be useful for clinical diagnosis.

In summary, this study provides evidence that US tissues can be characterized by quantitative texture analysis in a robust way, allowing its use for diagnostic purposes in clinical practice. These results should be confirmed in larger clinical images of the same tissue acquired under different controlled conditions and validated by using this information to examine the ability to detect differences against a clinical outcome in an effective manner.

References

1. Castellano G, Bonilha L, Li LM, Cendes F. Textures analysis of medical images. Clin Radiol 2004; 59:1061–1069.
2. Haralick RM, Shanmugam K. Textural features for image classification. IEEE Trans Syst Man Cybern Syst 1973; 3:610–621.
3. Junior OL, Delgado D, Gonçalves V, Nunes U. Trainable classifier-fusion schemes: an application to pedestrian detection. In: Proceedings of the 12th International IEEE Conference on Intelligent Transportation Systems. Piscataway, NJ: Institute of Electrical and Electronics Engineers; 2009:1–6. IEEE.
4. Ojala T, Pietikainen M, Maenpaa T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Trans Pattern Anal Mach Intell 2002; 24:971–987.
5. Oosterveld BJ, Thijssen JM, Verhoef WA. Texture of B-mode echograms: 3-D simulations and experiments of the effects of diffraction and scatterer density. Ultrasound Imaging 1985; 7:142–160.
6. Bergen JR, Adelson EH. Theories of visual texture perception. Spat Vision 1991; 10:114–134.
7. Insana MF, Garra BS, Rosenthal SJ, Hall TJ. Quantitative ultrasonography. Med Prog Technol 1988; 15:141–153.
8. Tunis AS, Czarnota GJ, Giles A, Sherar MD, Hunt JW, Kolios MC. Monitoring structural changes in cells with high-frequency ultrasound signal statistics. Ultrasound Med Biol 2005; 31:1041–1049.
9. Bonet-Carne E, Palacio M, Cobo T, et al. Quantitative ultrasound texture analysis of fetal lungs to predict neonatal respiratory morbidity. Ultrasound Obstet Gynecol 2015; 45:427–433.
10. Chen DR, Chang RF, Kuo WJ, Chen MC, Huang YL. Diagnosis of breast tumors with sonographic texture analysis using wavelet transform and neural networks. Ultrasound Med Biol 2002; 28:1301–1310.
11. Hartman PC, Oosterveld BJ, Thijssen JM, Rosenbusch GJ. Variability of quantitative echographic parameters of the liver: intra- and interindividual spread, temporal- and age-related effects. Ultrasound Med Biol 1991; 17:857–867.
12. Tenorio V, Bonet-Carne E, Figueras F, et al. Correlation of quantitative texture analysis of cranial ultrasound with later neurobehavior in preterm infants. Ultrasound Med Biol 2014; 40:2285–2294.
13. Sujana H, Swarnamani S, Suresh S. Application of artificial neural networks for the classification of liver lesions by image texture parameters. Ultrasound Med Biol 1996; 22:1177–1181.
14. Vince DG, Dixon KJ, Cothren RM, Cornhill JF. Comparison of texture analysis methods for the characterization of coronary plaques in intravascular ultrasound images. Comput Med Imaging Graph 2000; 24:221–229.
15. Oelze ML. Emerging quantitative ultrasound applications: from diagnosing disease to monitoring of therapy. J Acoust Soc Am 2016; 140:S136.
16. Oelze ML, Mammou J. Review of quantitative ultrasound: envelope statistics and backscatter coefficient imaging and contributions to diagnostic ultrasound. IEEE Trans Ultrasound Ferroelectr Freq Control 2016; 63:336–351.
17. Christodoulou CI, Pattichis CS, Pantziaris M, Nicolaides A. Texture-based classification of atherosclerotic carotid plaques. IEEE Trans Med Imaging 2003; 22:902–912.
18. El-Gayar MM, Soliman H. A comparative study of image low level feature extraction algorithms. Egyptian Informatics J 2013; 14:175–181.
19. Garra BS, Krasner BH, Hori SC, Ascher S, Mun SK, Zeman RK. Improving the distinction between benign and malignant breast lesions: the value of sonographic texture analysis. Ultrasound Imaging 1993; 15:267–285.
20. Oosterveld BJ, Thijssen JM, Hartman PC, Romijn RL, Rosenbusch GJ. Ultrasound attenuation and texture analysis of diffuse liver disease: methods and preliminary results. Phys Med Biol 1991; 36:1039–1064.
21. Schmitz G, Ermert H, Senge T. Tissue-characterization of the prostate using radio frequency ultrasonic signals. IEEE Trans Ultrason Ferroelectr Freq Control 1999; 46:126–138.
22. Serizawa M, Maeda K. Noninvasive fetal lung maturity prediction based on gray level histogram width. Ultrasound Med Biol 2010; 36:1998–2003.
23. Ojansivu V, Heikila J. Blur insensitive texture classification using local phase quantization. In: Proceedings of the International Conference on Image and Signal Processing. Berlin, Germany: Springer; 2008:236–243.

24. Ojansivu V, Rahtu E; Heikila J. Rotation invariant local phase quantization for blur insensitive texture analysis. In: Proceedings of the 19th International IEEE Conference on Pattern Recognition. Piscataway, NJ: Institute of Electrical and Electronics Engineers; 2008:1–4.

25. Thijssen JM. Ultrasonic speckle formation, analysis and processing applied to tissue characterization. Pattern Recognit Lett 2003; 24:659–675.

26. Alexander DC, Zikic D, Zhang J, Zhang H, Criminisi A. Image quality transfer via random forest regression: applications in diffusion MRI. In: Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention. Cham, Switzerland: Springer; 2014:225–232.

27. Hossain S, Serikawa S. Texture databases: a comprehensive survey. Pattern Recognit Lett 2013; 34:2007–2022.

28. Ojala T, Maenpaa T, Pietikainen M, Viertola J, Kyllo nen J, Huovinen S. OUTEX: new framework for empirical evaluation of texture analysis algorithms. In: Proceedings of the 16th International IEEE Conference on Pattern Recognition. Vol 1. Piscataway, NJ: Institute of Electrical and Electronics Engineers; 2002:701–706.

29. Targhi AT, Geusebroek JM, Zisserman A. Texture classification with minimal training images. In: Proceedings of the 19th International IEEE Conference on Pattern Recognition. Piscataway, NJ: Institute of Electrical and Electronics Engineers; 2008:1–4.

30. Giesen RJ, Huynen AL, Aarnink RG, de la Rosette JJ, Debruyne FM, Wijkstra H. Construction and application of hierarchical decision tree for classification of ultrasonographic prostate images. Med Biol Eng Comput 1996; 34:105–109.

31. Poonguzhali S, Ravindran G. Automatic classification of focal lesions in ultrasound liver images using combined texture features. Inform Tech J 2008; 7:205–209.

32. Barrios N, Perez-Moreno A, Migliorelli F, et al. Quantitative analysis of the cervical texture by ultrasound and correlation with gestational age. Fetal Diagn Ther 2017; 41:265–272.

33. Keramidas EG, Iakovidis DK, Maroulis D, Dimitropoulos N. Thyroid texture representation via noise resistant image features. In: Proceedings of the 21st International IEEE Symposium on Computer-Based Medical Systems. Piscataway, NJ: Institute of Electrical and Electronics Engineers; 2008:560–565.

34. Nanni L, Lumini A, Brahnam S. Local binary patterns variants as texture descriptors for medical image analysis. Artif Intell Med 2010; 49:117–125.

35. Beyerer J, León FP, Frese C. Machine Vision: Automated Visual Inspection—Theory, Practice and Applications. Berlin, Germany: Springer; 2015.

36. Pazinato DV, Stein BV, de Almeida WR, et al. Pixel-level tissue classification for ultrasound images. IEEE J Biomed Health Inform 2016; 20:256–267.

37. Ravishankar H, Annangi P, Washburn M, Lanning J. Automated kidney morphology measurements from ultrasound images using texture and edge analysis. In: SPIE Medical Imaging. Bellingham, WA: International Society for Optics and Photonics; 2016:97901A.

38. Deng Y, Wang Y, Chen P, Yu J. A hierarchical model for automatic nuchal translucency detection from ultrasound images. Comput Biol Med 2012; 42:706–713.

39. Tenbrinck D, Schmid S, Jiang X, Schäfers K, Stypmann J. Histogram-based optical flow for motion estimation in ultrasound imaging. J Math Imaging Vision 2013; 47:138–150.

40. Huisman HJ, Thijssen JH. An in vivo ultrasonic model of liver parenchyma. IEEE Trans Ultrason Ferroelectr Freq Control 1998; 45:739–750.

41. Friedrich-Rust M, Ong MF, Martens S, et al. Performance of transient elastography for the staging of liver fibrosis: a meta-analysis. Gastroenterology 2008; 134:960–974.

42. Sadeghi-Naini A, Papanicolaou N, Falou O, et al. Quantitative ultrasound evaluation of tumor cell death response in locally advanced breast cancer patients receiving chemotherapy. Clin Cancer Res 2013; 19:2163–2174.

43. Sannachi L, Tadayyon H, Sadeghi-Naini A, et al. Non-invasive evaluation of breast cancer response to chemotherapy using quantitative ultrasonic backscatter parameters. Med Image Anal 2015; 20:224–236.

44. Rosado-Mendez IM, Carlson LC, Woo KM, et al. Quantitative assessment of cervical softening during pregnancy in the Rhesus macaque with shear wave elasticity imaging. Phys Med Biol 2018; 63:085016.

45. Palacio M, Bonet-Carne E, Cobo T, et al. Prediction of neonatal respiratory morbidity by quantitative ultrasound lung texture analysis: a multicenter study. Am J Obstet Gynecol 2017; 217:196.e1–196.e14.