Draft. Full paper at: http://dx.doi.org/10.1109/ICCP.2010.5606444

Performance Evaluation of the Textural Analysis Algorithms in Liver Fibrosis Detection Using Ultrasound Software Phantoms

Cristian Vicas, Sergiu Nedevschi
Technical University of Cluj Napoca
Cluj Napoca, Romania
{cristian.vicas; sergiu.nedevschi}@cs.utcluj.ro

***Abstract*—Diagnosing liver fibrosis using non invasive procedures is challenging because the visual aspects in ultrasound imaging between healthy and fibrosis liver are very much alike. In this paper we use ultrasound simulation software to produce realistic liver tissue images. These images are produced by scanning several software liver phantoms. The phantoms proposed in this paper are generated based on relevant histological findings. The phantoms mimic the five METAVIR fibrosis stages from healthy liver to cirrhosis. A texture analysis based system, developed in our pervious studies is applied on these images. Accuracy in differentiating between healthy liver and various fibrosis stages is recorded. We show how various phantom properties affect the detection rates and present some explanations on why the liver fibrosis is not detectable to a level that is useful for physicians.**

Keywords-liver fibrosis; software phantom; textural analysis; ultrasound simulation

I. INTRODUCTION

The accurate diagnosis and staging of hepatic fibrosis is crucial for prognosis and treatment of liver diseases. There ar a number of pathologies that evolve with fibrosis. In this paper we focus only on fibrosis generated by chronic hepatitis C. To date, liver biopsy remains the gold standard for fibrosis assessment in hepatitis C. Liver biopsy suffers from several important drawbacks like morbidity, observer variability and sampling variation [1] hence the interest in developing a noninvasive fibrosis detection system. In literature there are a number of papers that propose the texture analysis methods applied on B-mode ultrasound images [2]. In our recent work [3] we have implemented such a system and applied it to 300 patients having various fibrosis stages. The fibrosis detection rates were relatively low compared to the data reported by other authors.

In a recent paper Bonekamp et al. [4] made a review of current imaging modalities that could detect and stage hepatic fibrosis and cirrhosis. In this review the B-mode ultrasound was mentioned but without any reference to computerized image analysis as being a reliable tool in fibrosis detection. This review confirms our findings that texture analysis does not provide a clinically useful tool for fibrosis staging.

Present paper is focused on fibrosis detection but from another perspective. Using relevant histological findings we propose software models that simulate the scatter patterns found in healthy and fibrotic liver. These models are used to build liver phantoms. Liver fibrosis is simulated according to METAVIR staging [5]. Ultrasound simulation software is applied on these phantoms and we generate ultrasound images. The images are then processed by means of texture analysis methods. We record the detection rates with respect to several phantom parameters and fibrosis stages.

The ultrasound simulation field is a rich one. The main applications are focused on cardiovascular imaging or simulation of ultrasound imaging from CT data, but there are approaches to ultrasound image segmentation, registration, etc.

In [6] the authors show that there are several tools available to simulate ultrasound images. The authors point out that the realism of the simulation cannot be achieved without computational effort. After analyzing several commercial and non commercial ultrasound simulation systems the authors in [7] claim that Field II package [8] gives the best results with carefully designed phantom, but this method has the disadvantage of being very slow. In a recent review [9] focused on ultrasound image segmentation the authors refer to Field II software package as an effective application for producing ground truth data.

There are other approaches that relates to Field II package. In [10] the authors propose a faster, convolution based system that simulates 2D/3D cardiac images. The authors show that the generated images have the same Rayleigh statistics and the same visual aspects with the Field II generated images.

In this paper the Field II software is used to generate US images from software phantoms. This software was developed by J.A. Jensen as a fast and accurate method for calculating the pulsed pressure field emitted from an arbitrary shaped, apodized and excited ultrasound transducer [8, 11-12].

The software phantoms described in literature are focused on the heart/cardiovascular system. Ultrasound liver imaging has received less attention. The authors in [13] proposed a software liver phantom that was focused mainly on other liver

diseases like steatosis. The features they use are built around the concept of attenuation. In [14] the authors proposed a liver phantom that is used to study the modifications of speckle distribution in cirrhosis.

In this paper we will present two software phantom models that simulate healthy, fibrosis and cirrhotic liver. The images are analyzed using a previously developed fibrosis detection system. This system is based on texture analysis and recognition.

Several textural attributes (features) are computed from each image. These features are used to train a classification schema. The classifier will then learn to discriminate between various diseases. In this paper the same detection system will be used to discriminate between synthesized images generated from different phantom models scores.

The paper is structured as following: In section 2 is described the ultrasound simulation setup and the phantom models. Section 3 presents the texture analysis system, section 4 presents the experimental results and section 5 presents the discussions and conclusions.

II. Ultrasound Image Simulation Setup

In this section we will present the simulation setup and the ultrasound liver models that we developed.

The ultrasound simulation procedure follows the guidelines presented in [11]. The probe used in this simulation is a 5.5Mhz linear probe, with 98 elements. Apodization is implemented using a Hanning window of 20 active elements. The excitation signal is two wavelengths long. Returned signal is sampled at 100Mhz. Speed of sound is set at 1540 m/s. In axial direction two scatters can be discriminated if they are situated at a distance larger than 1/2 of pulse width [15]. For the proposed simulation setup this axial resolution is 0.28 mm:

$$A_r = (N_p \lambda)/2 \qquad (1)$$

where A_r is the axial resolution, N_p is the number of wavelengths in excitation signal and λ is the wavelength of the ultrasound waves.

There are 25 lines acquired from the underlying medium. Care is taken that all 25 lines are uniformly distributed inside the phantom. The lateral resolution for the acquisition is 0.97 lines/mm. Envelope of the signal is detected using Hilbert transform. The envelope signal is sampled, logarithmically compressed and scaled to a 256 level graymap.

The images are synthesized at a resolution of 5 pixels/mm. At this resolution the physical distance between two pixels is 0.2mm. This distance is slightly smaller than the theoretical axial resolution for this simulation setup. This ensures that we do not loose information during scan conversion and interpolation.

Most of the measurements used for defining the phantom are based on the findings presented by Crawford et al. in [16]. Here, the authors performed a statistical analysis of 16 liver biopsies that were reported with no diagnostic abnormalities.

The basic morphological unit of the liver is liver lobule [17-19]. This lobule has a hexagonal shape and is 0.8-1.5mm in diameter. In the corners of the hexagon one can find the portal area (portal triad), a vascular system consisting of one large vein, an artery and one bile duct. In the center of the lobule is the central vein. The diameter of the portal triad is approx 0.2mm. The portal triad is not always present in all 6 corners of the liver lobule. In addition, not all the "triads" contain all three elements. Some portal spaces lacks the portal artery or the bile duct or both [16]. It is important to note that in healthy human liver the only region containing connective tissue is the portal area. The density of the portal spaces is 0.8 triads/mm^2 (with standard deviation SD=0.5). In this paper we do not distinguish between portal spaces with one, two or three elements. All portal spaces are simulated in the same way.

For the phantom synthesis we must account for the scatter interfaces. These interfaces appear when the ultrasound crosses the threshold between two regions with different impedances. We assume that most of the scattering is generated by the connective tissue. In the rest of the liver parenchyma we assume that there are randomly placed scatters having Gaussian amplitude. Central vein, found in the center of the lobule doesn't have connectivity tissue and it was ignored in the simulations.

In case of injury, the normal hepatic tissue is gradually replaced by fibrous connective tissue. Because this scarring tissue is similar to the connectivity tissue we assume that it has the same scattering properties as the regular connectivity tissue.

In this paper we are focusing on the lesions produced by chronic hepatitis C. In this disease METAVIR score classify the fibrosis in 5 stages according to the distribution of the scarring tissue [5]. METAVIR F0 score represents a healthy liver. METAVIR F1 assumes that there is minimal scarring in the portal spaces. F2 score assumes that the scarring has occurred and extends outside the areas in the liver that contains blood vessels. In F3 stage the scarring tissue create fibrosis bridges between portal areas. Fibrosis stage F4 or cirrhosis assumes that the regeneration process has begun and one can note advanced scarring and regeneration nodules surrounded by thick fibrosis walls.

Our software phantom models mimic the scattering aspect of the normal liver and of all 4 METAVIR stages. There are two types of models. One for normal liver and fibrosis stages 1, 2, 3 and one for fibrosis stage 4.

The phantom dimensions are fixed, 25.6 mm width (x axis), 25.6 mm height (z axis) and 5 mm thickness (y axis). The phantom is situated at 3 cm below the surface of the ultrasound probe. The medium is perfectly transparent for the ultrasound waves and it has zero attenuation for all the frequencies. There are no scatters in the medium, except for those defined in the phantom. From a 25.6x25.6 mm phantom results a 128x128 pixels image.

The phantom scatters are of two types. There are random scatters, distributed uniformly in the phantom, having normal distributed amplitude of standard deviation 1 and 0 mean. These scatters represent the interfaces produced by the hepatocytes and other non fibrotic tissues. We name this type A scatters. The other type of scatters is normally distributed along certain geometrical shapes specified by the model. They have also zero mean and one standard deviation for the amplitude. These scatters define the specific interfaces

Draft. Full paper at: http://dx.doi.org/10.1109/ICCP.2010.5606444

generated by connectivity tissue. We name them type B scatters.

The building block of the liver is the hexagonal liver lobule. In our phantom a hexagonal grid mimic the liver lobule and is used to position various elements. The size of hexagonal grid is chosen in such a way that one cell is circumscribed in a circle with a specified diameter of 1 mm. The position and rotation of the grid is randomly altered for each phantom that is generated. In Fig. 1.A is shown such a grid.

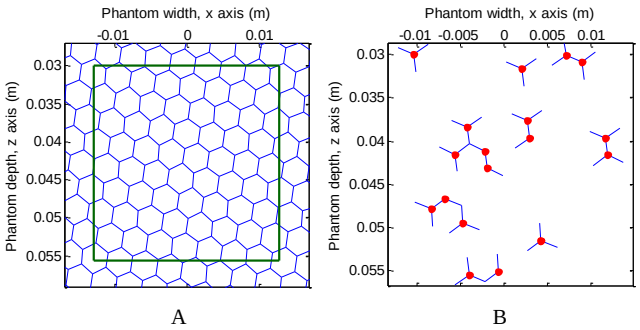
In the corners of the hexagon one can find the portal triads. These triads are simulated using a cylinder of diameter 0.15 mm centered on the corner point. The cylinder is aligned with y axis, along the thickness of the phantom. Scatters are positioned random along the walls of the cylinder. The distance between the scatter and the cylinder wall follows a normal distribution with SD=0.009mm and mean 0. The number of scatters around the cylinder is the same for all portal spaces. Care is taken that there are no type A scatters inside the cylinder. This empty space inside the cylinder simulates the large vein present in the triad. The fibrotic tissue normally present inside the portal space is simulated by these type B scatters situated around the cylinder.

One should note that not always the triads are present in the corners of the lobule. In our phantom we place a portal cylinder in a hexagon corner with a probability computed based on the desired density of portal spaces (Fig. 1.B).

For fibrosis stage 1 we increase the thickness of the portal space by adding another cylinder of diameter 0.2 mm around each previously existing cylinder. The scatters are following the same distribution as for F0 cylinder. For fibrosis stage 2, we add the cylinders from stage one along with another cylinder having a diameter of 0.35 mm. For fibrosis stage 3 we have the cylinders from F2. In addition to those we add another one with the diameter of 0.4 mm. One should note that cylinders are concentric and they are added as the fibrosis stage increases. As a result, for fibrosis stage 3 we have 4 concentric cylinders with diameters of 0.15, 0.2, 0.35 and 0.4 mm, each one producing the same number of scatters. The total number of scatters per portal space increases as the fibrosis stage increases.

In addition to the increased fibrosis in portal space, in case of fibrosis stage 2 and 3 one can find fibrous septa that infiltrate the liver parenchyma starting from portal spaces [5]. These septa are simulated by positioning scatters along the hexagon edges. In case of F2 stage the scatters are concentrated around the end caps of the edges following a normal distribution. Care is taken that the scatters don't confluence because METAVIR F2 score implies that there are no portal to portal fibrous bridges.

For F3 stage simulation the scatters are uniformly distributed along the edges and along y axis (thickness). Each scatter's position is affected by a normal noise with SD=0.03mm and mean=0. These scatters simulate the fibrous bridges found in F3 stage. The number of scatters that are distributed along the edges for F2 and F3 cases is proportional with the number of scatters that are distributed along each cylinder. Only edges that have an end point inside of a portal triad are taken into consideration.



A B

Figure 1. A: The hexagonal grid. The diameter of the hexagonal cell is scaled in this image. The rectanlge in the middle of the image represents the phantom boundaries. B: Hexagonal grid with portal spaces (red dots). Note that the portal spaces are not present in all the hexagonal vertices.

The number of type B scatter points that are added for each cylinder and edge is a parameter of the algorithm. One should note that F2 phantom has three times more scatters in the portal area than F0 phantom.

In order to keep the phantom parameter complexity low we decided that the number of scatters that are simulated along edges is 0.8 times this parameter for F2 and 1.1 for F3 fibrosis. In Fig. 2 are shown some portal triads for each fibrosis stage.

For F4 stage we abandon the portal triads model and simulate the fibrous tissue that surrounds islands of normal liver hepatocytes. Along each edge of the grid we randomly set scatter points that are at a normally distributed distance from the edges. These scatters overlap the existing type A scatters. The phantom will consist of a hexagonal grid, each grid edge is a cylinder containing normally distributed scatters along its center. This grid is immersed in uniformly distributed type A scatters. In F4 stage one can find regenerating nodules having dimensions up to several centimeters. For F4 phantom the size of the grid is not fixed. The diameter of the circumscribed circle can take arbitrary values. Another parameter is the SD for the distance between the grid and the scatters (Fig. 3).

We also define a contrast ratio as being the ratio between the mean type B scatter density and type A density. The type A density is measured by dividing the total number of randomly placed scatters to the total phantom volume (expressed in mm). The type B density is measured by dividing the number of scatters placed along the grid to the total volume of the cylinders centered on the hexagonal grid and having the radius equal to 3 SD.

Given a certain ratio and a total number of scatters the algorithm automatically computes the necessary number of type A or B scatters.

The model used to simulate liver fibrosis stages 0, 1, 2 and 3 have the following parameters:

- Density of the portal spaces (measured in portal spaces/mm^2) – *PSD*
- Number of randomly spaced scatters (type A)
- Number of scatters per cylinder – *NSC*
- Fibrosis stage – *F*

Draft. Full paper at: http://dx.doi.org/10.1109/ICCP.2010.5606444

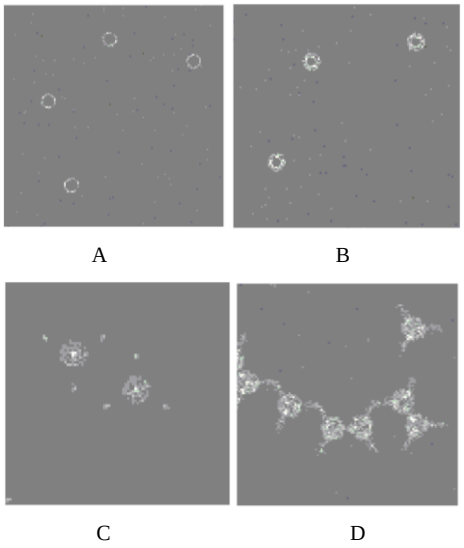


Figure 2. Aspects of software phantom. The type B scatters are highlighted. A) Normal liver. B) Fibrosis stage 1. C) Fibrosis stage 2. D) Fibrosis stage 3. The scale for each image vary.

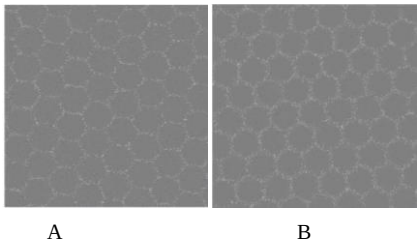


Figure 3. Fibrosis 4 pahntoms having different B scatter standard deviation. A) SD=0.064 mm, B) SD=0.1 mm

The phantom model used to simulate liver cirrhosis has the following parameters:

- Total number of scatters
- Dimension of the hexagon cell (measured in mm)
- The density ratio (contrast ratio)
- The B scatters standard deviation

III. TEXTURAL ANALYSIS

In literature there are several approaches in noninvasive detection of liver fibrosis based on ultrasound investigation.

Most of the authors study the B-mode ultrasound image by means of texture analysis. In texture analysis there are two main steps. First step is the computation of several textural attributes that numerically describe the texture (using dedicated algorithms). The second step involves the training and evaluation of a classifier using the previously computed textural features.

A. Texture Features

In literature, there are many algorithms that describe the texture. Each algorithm treats the texture as being produced by a certain model [20]. There are algorithms that assume that the texture is a collection of pixels having random intensities, other algorithms treat the texture as a superposition of sinusoidal waves with different phase, amplitude and frequency, etc.

Each texture description algorithm has a certain number of parameters that control the feature extraction process. For each algorithm we used the proposed set of parameters found in corresponding paper. In the following we enumerate the algorithms and the papers that use these algorithms in fibrosis detection (when is possible).

The algorithms used in our previous research regarding the noninvasive diagnosis in diffuse liver diseases [2-3, 21-25] are: First order statistics [26-27], gray tone difference matrix [20], gray level co-occurrence matrix [26-29] , multiresolution fractal dimension [28], differential box counting [30-31], morphological fractal dimension estimators [32], Fourier power spectrum [28, 33], Gabor filters [34], Law's energy measures [28], texture edge co-occurrence matrix [30], phase congruency [35] and texture feature coding matrix [36].

There are 12 implemented algorithms that process the entire 128x128 image and compute 232 features per image. Each image histogram was equalized before entering the feature computation step. This ensures that the detection is not due some difference in overall gray level.

B. Learning Schema

We use logistic regression as a classifier [37]. There are other algorithms that are employed in fibrosis detection, algorithms varying from k-nearest neighbor, neuronal networks, decision trees, support vector machines, etc. [3, 26]. The reasons of this choice are that logistic regression produces simple to interpret model and doesn't overfit the data even if the number of instances is lower than the number of features [38]. Once the model was generated one can learn the relevance of the features by looking at the model coefficients. Another reason for this choice is that we already evaluated this algorithm in fibrosis detection field [25]. The performance criteria is accuracy, the ratio between correctly classified instances and the total number of instances in the dataset. To measure the accuracy for a certain dataset the following procedure is applied:

The available data is split in 10 disjoint folds. Iteratively, each fold is used for testing and the other 9 for training. At each step the predictions are collected. After 10 steps the accuracy is computed on all the predictions. This process is repeated 10 times with different random splits. The mean accuracy is reported as the performance of the algorithm for the current dataset. Before entering the classification step each feature is normalized to [0,1] interval. The test fold subset is also normalized using the same values that were used for corresponding training set.

In theory the accuracy can vary between 0% and 100%. In case of a balanced two class problem an accuracy of 50% means that the classifier randomly distributes the test data between the two classes. An accuracy of 100% means that the classifier perfectly predicts the class of each test instance. A balanced problem means that the volume of the tow classes is the same. This is true for both the testing and training data sets.

In this paper we also identify the relevant textural features for the problem addressed. In order to reduce the variability of

the result the following schema is used: The available dataset is spited in 10 folds. At each run, a new dataset is produced by eliminating one fold. For this dataset a logistic model is produced and the coefficients for each feature are recorded. After 10 runs, we report the mean coefficient value for each feature. Ordering the data in descending order of their absolute mean will produce a feature ranking for the specified dataset.

IV. Experimental Results

The experiments are focused on determining the influence of various phantom settings over the detection rates. We do not intend to discriminate between various fibrosis stages, only to differentiate between healthy liver and fibrotic liver.

The available data is grouped in two class problems. First class is normal liver and the second class is one of the fibrosis stages, F1, F2, F3 or F4. The class volumes are equal. We vary various parameters of the phantoms and record the detection accuracy.

A. *Fibrosis 1, 2 and 3 phantom model*

In the following experiment we identify the role of NSC parameter.

We generate phantoms with fibrosis stages 0, 1, 2 and 3 using the corresponding model. From each model we generate images having the *NSC* parameter 5, 10, 20, 30, 40, 50, 60, 80 and 90. For each combination of *NSC* and fibrosis stage we generate 100 images. The datasets are generated by combining each *NSC* parameter value for fibrosis 0 with each *NSC* parameter for the rest of fibrosis stages. The results for F0 and *NSC*=5, 20, 40 and 80 are summarized in Fig. 4. One can note that in Fig. 4.A there is no detection when all the images are generated using a low value for *NSC*. When *NSC* reaches higher values (20-60) we can detect higher fibrosis stages. Only at high *NSC* values (80-90) the detection rates for all fibrosis stages are acceptable. An interesting result can be observed in Fig. 4.B where detection rates for F0-F1 are raised when F0 images have *NSC*=80 and F1 images have NSC<20. We examined closer the datasets and noted that this phenomenon is due the fact that for these values the F1 images have lower contrast values than F0 images. In the rest of the datasets (where detection rates were high) the contrast values for F0 images are always lower than F,1,2,3 images. Relevant features and their significance will be discussed at the end of this section.

Another very important parameter is the average portal space density, *PSD*. If the distance between two portal spaces is below the theoretical discrimination resolution then the liver behaves like a homogenous medium because the ultrasounds cannot resolve different portal spaces. We want to know how the *PSD* parameter influences the detection rates.

In the following experiment we generate images having fibrosis stages 0, 1, 2 and 3. The *PSD* parameters take values in 0.1, 0.3, 0.6, 0.8, 1.3, 1.5 and 3 portal spaces/mm^2. From each combination of fibrosis stage and *PSD* values we generate 100 images. Accuracy is measured for each pair of F0 with F1, F2 and F3. In Fig. 5, Fig. 6, and Fig. 7 are shown the results. The detection rates are acceptable for a narrow range of *PSD* values. It is interesting to note that the *PSD* values for F0 stage don't affect much the detection rates.

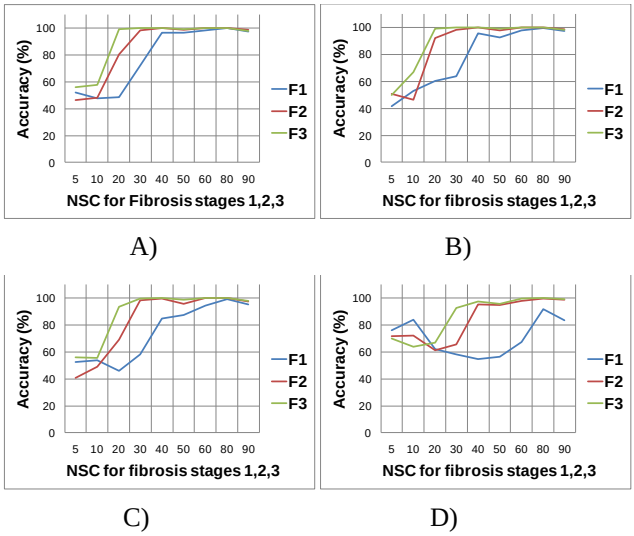


Figure 4. Detection rates in detecting fibrosis sgates 1, 2 and 3. A) *NSC* for F0 is 5. B) *NSC* for F0 is 20. C) *NSC* for F0 is 40. D) *NSC* for F0 is 80.

This behavior could be because the F0 phantoms appear as homogenous medium. The slightly descendent slope of detections recorded when the PSD values for F0 increases tend to endorse this theory. The *PSD* value range that gives a good detection increases as the fibrosis stage increases.

In all three fibrosis stages one can note a significant drop in detection rates when the values exceed the boundaries. Good detection rates are observed in narrower density ranges than the portal density range observed in biopsies. Authors in [16] noted 0.8 portal spaces/mm2 with a standard deviation of 0.5. Only F0 vs. F3 dataset have good detection in the range that extends ±1 standard deviation for portal space density found in human liver.

B. *Fibrosis 4 phantom model*

The METAVIR F4 phantom has the number of scatters fixed to 30000 and the B scatter SD to 0.064 mm. This standard deviation yields a cylinder of approx 0.3 mm diameter full with scatters. The other two parameters, the grid dimension and the density ratio are studied below.

We compare the F4 phantom with a phantom containing no type B scatters. First, we establish a large value for grid dimension (10 mm diameter) and iterate the density ratios, from 1 through 4 in 0.1 increments. Of course, at density ratio of 1 there are no B scatters in the phantom. As expected, the detection rates increase as the contrast ratio increases. In Table I are summarized the results.

TABLE I. Detection rates of pure type A phantoms vs F4 phantoms having various density ratio values

Density ratio	Accuracy
≤1.8	≤55%
1.9	73%
2	74%
2.1	84%
2.2	82%
2.3	90%
2.4	93%
≥2.5	100%

TABLE II. DETECTION RATES OF PURE TYPE A PHANTOMS VS F4 PHANTOMS HAVING VARIOUS GRID CELL DIAMETERS

Grid scale (diameter)	Accuracy
<2 mm	<57%
3 mm	60.1%
4 mm	70%
6 mm	80%
8 mm	92%
>10 mm	≥99%

The detection rates reach 100% for a density ratio greater than 2.5. In the next experiment we set the density ratio to 2.5 and iterate through several scales. From equation (1) the minimal distance between two scatters that can be differentiate is 0.28 mm. In this experiment the following diameters were chosen: 0.2, 0.4, 0.8, 1.2, 1.6, 3, 4, 6, 8, 10 and 12 mm. From Table II one can see that the accuracy reaches an acceptable level only at a cell diameter of 6 mm which is far greater than the theoretical resolution limit. This can be due the fact that the type B scatters have the same amplitude as the rest of the scatters. The only difference is that Type B scatters are correlated and not uniformly distributed. We do not alter the amplitude of the scatters because in the in vivo ultrasound liver images there are no strong scatters, especially in homogenous areas.

In Fig. 8 are shown the accuracies when differentiating between F0 phantom and F4 phantom with different hexagonal grid dimensions. The descendent slope is a surprising result. Its significance is that the detection rates decreases as the regeneration nodules grow in dimensions. This result is somehow conflicting with the results shown in Table II where a bigger grid means greater detection performance. However, when looking at the features one can note an inversion for the relevant textural attributes. The Gray Tone Difference Matrix coarseness is greater for F0 case than for F4, in opposition with the result found in Table II where coarseness was greater for F4 cases.

The significance of these findings is that the discrimination between healthy and cirrhosis liver doesn't vary monotonically with the regeneration nodule size. It is possible that a nodule could become so large that the ROI will fit inside it. In this case the textural features will characterize a healthy, homogenous liver tissue. Even if the fibrosis wall is captured inside the ROI is possible that it will not give enough "coarseness" to the texture. In this case, a multirezolution approach is indicated, but this approach should address the entire liver tissue not only a small ROI.

C. *Relevant features*

Another important aspect of texture analysis is the identification of relevant texture features. From each experiment we take the datasets that produces detection rates greater than 50-60% and study the relevant features according to the procedure described above. The results were similar in terms of the relevant algorithms. In the following we will only enumerate the relevant algorithms.

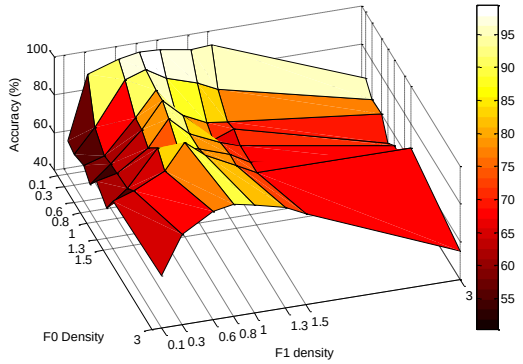


Figure 5. Detection rates between F0 and F1 at different *PSD* values

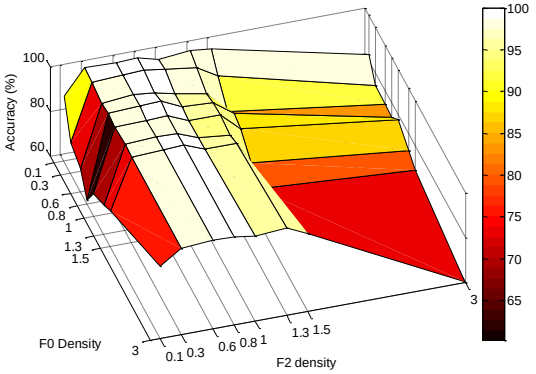


Figure 6. Detection rates between F0 and F2 at different *PSD* values

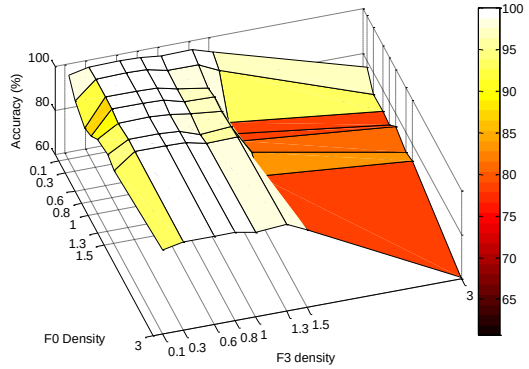


Figure 7. Detection rates between F0 and F3 at different *PSD* values

In vast majority of the cases the relevant features were produced by Gray Tone Difference Matrix. Attributes like coarseness, busyness and contrast are among the most relevant features. Other relevant algorithms found in top positions are entropy from First Order statistics, entropy and correlation from Gray Level Co-occurrence Matrix and Fourier Power Spectrum.

Other features like Laws energy measures [39-40] or Multiresolution Fractal Dimension [28] are found in higher positions only when the detection rates are lower.

We can draw the conclusion that contrast related features are relevant when the classes can be easily differentiated. Other features are not to be discarded because they are relevant when the contrast features are not capable of good class discrimination.

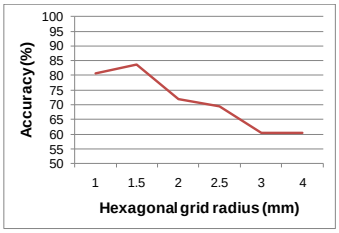


Figure 8. Detection rates between F0 and F4 at different grid sizes

Other classification schemas could be able to use the information provided by these features and increase the detection rates.

Synthesized images have the advantage that the "ground truth" is always known, the quantity and the quality of the fibrosis can be controlled and the number of images is limited only by the computation costs. Moreover, simulated US images are artifact free and are operator independent.

Real ultrasound images are corrupted by various artifacts, the number of images varies greatly from one fibrosis stage to another and there is operator variability in US acquisition and in ROI establishment. We also have a significant variability in establishing the "ground truth" even when the liver biopsy is performed [1].

V. Discussions and Conclusions

There are some disadvantages of the simulated ultrasound images and especially to the phantoms proposed here. Our phantoms tend to underestimate the variability of the real tissue. The phantoms assume that the section is perpendicular to the hepatic lobules. The models assume only 1 scale. The "real" liver might exhibit more details at higher scales. Maybe a fractal approach to the simulation could provide a more realistic model. Crawford et al. [16] mentioned that the vascular tree could exhibit such a behavior. The portal artery and the biliary duct are not modeled. In the same time, the portal triad is modeled as a cylinder. This might not be an accurate representation. A star like shape, where the rays of the stars follow the inter lobule divisions might be more appropriate.

In our phantoms there are no other anatomical elements like larger vein branches, arteries, etc. These elements can easily be found in any liver. For fibrosis detection most of the authors suggest that the ROI should be placed in a homogenous area, avoiding artifacts and anatomical features. In this paper, the phantoms were generated in such a way that they mimic an ideal artifact free region of the liver. Adding certain anatomical features to the phantom, even with low probability might increase the realism of the simulation. The simulation setup assumes that the phantom is suspended into a perfectly transparent medium. This mediun does not produce scatters or attenuation. The ROI is fixed to the size of the phantom. In *in vivo*, the ROI corresponds to a tissue sample that is few cm deep into the liver. Ultrasounds has to pass through various structures like more liver tissue, liver capsule, fat tissue, muscles, skin and skin-probe interface. These structures might degrade the signal quality.

Present simulation setup used a linear array for imaging. In abdominal imaging the physicians use frequently convex probes. Simulation of these probes along with phased array imaging will remain an important research direction, especially because polar scan conversion alters both axial and lateral resolution in a position dependent manner. In the same time is important to know how various acquisition settings influence the detection rates. These settings vary from ultrasound frequency to compression and reconstruction procedures. This paper establishes the detection rates in case of a given phantom and set of parameters. Starting from this one can investigate the effects of various probe parameters.

In our previous work we have shown that the detection rates on liver ultrasound images is relatively low. A major finding of this paper is that there are limitations to direct fibrosis detection even in ideal cases. These limitations are due to the spatial resolution of the ultrasound imaging. If the major scatters are too dense the ultrasound image appears as the underlying medium is a homogenous tissue. In the same time, if the scatters are too spare or too low as intensity they are not resolved in a detectable fashion.

The detection depends heavily on the physical parameters of the phantom. The detection rates of mild fibrosis (F1,F2) are acceptable only in a narrow band of portal space density values. The actual density values that can be found in human liver have a boarder range. The simulated datasets are homogenous with respect to these parameters. This means that all the F0 images have the *PSD, NSC*, etc. parameters fixed.

In *in vivo*, the portal density, the lobule dimensions, etc. vary for each patient. On actual patient data the classifiers trained to detect fibrosis performed better than a random classifier. This means that exist some information about the fibrosis, information that is detected by textural analysis. Another possibility is that the texture analysis detects only patients having liver properties fitted for detection. Feature ranking revealed an interesting fact. The features that are sensitive to contrast change are relevant. Gray Tone Difference Matrix is very powerful in describing the statistical relations between a pixel and its neighbors. Another statistical algorithm Gray Level Co-occurrence Matrix algorithm is not so sensitive because is orientation and scale sensitive.

The power of other features is shown in comparison cases where the detection rates are lower. No further investigation was done in this direction, but the preliminary data indicates that one cannot ignore these features. Their power might be evidenced only by a classifier capable to learn a more sophisticated model. In our previous experiments algorithms like Gabor filters or fractal dimensions were proved to give relevant features also.

In clinical practice is important to discriminate between different fibrosis stages. In the experiments presented here we did not address this topic. Although is not difficult to group the data in 5 class problem and modify the classification schema to handle multiple classes we choose not to pursue this research direction leaving it for a further paper. Another interesting research direction is to evaluate the discrimination power of the texture analysis versus the discrimination power of a trained physician.

REFERENCES

[1] P. Bedossa, D. Dargere, and V. Paradis, "Sampling variability of liver fibrosis in chronic hepatitis C," *Hepatology*, vol. 38, pp. 1449-1457, 2003.

[2] S. Nedevschi, C. Vicas, M. Lupsor, R. Badea, and M. Grigorescu, "The employment of textural and non textural image analysis algorithms in assessing the diffuse liver diseases," *Automat. Comput, Appl. Math.* , vol. 17, pp. 12-17, 2008.

[3] C. Vicas, S. Nedevschi, M. Lupsor, and R. Badea, "Fibrosis detection from ultrasound imaging. The influence of necro-inflammatory activity and steatosis over the detection rates," *Automat. Comput, Appl. Math.*, vol. 16, pp. 26-32, 2007.

[4] S. Bonekamp, I. Kamel, S. Solga, and J. Clark, "Can imaging modalities diagnose and stage hepatic fibrosis and cirrhosis accurately?," *Journal of Hepatology*, vol. 50, pp. 17-35, Jan 2009.

[5] P. Bedossa and T. Poynard, "An algorithm for the grading of activity in chronic hepatitis C. The METAVIR Cooperative Study Group," *Hepatology*, vol. 24, pp. 289-293, Aug 1996.

[6] T. Reichl, J. Passenger, O. Acosta, and O. Salvado, "Ultrasound goes GPU: real-time simulation using CUDA," *Procs SPIE*, 2009.

[7] O. Kutter, R. Shams, and N. Navab, "Visualization and GPU-accelerated simulation of medical ultrasound from CT images," *Comput Methods Programs Biomed*, vol. 94, pp. 250-266, Jun 2009.

[8] J. Jensen and P. Munk, "Computer phantoms for simulating ultrasound B-mode and cfm images," *Acoustical Imaging*, vol. 23, pp. 75-80, 1997.

[9] J. A. Noble and D. Boukerroui, "Ultrasound image segmentation: A survey," *Ieee Transactions on Medical Imaging*, vol. 25, pp. 987-1010, Aug 2006.

[10] H. Gao, H. Choi, P. Claus, S. Boonen, S. Jaecques, G. Van Lenthe, G. Van Der Perre, W. Lauriks, and J. D'hooge, "A fast convolution-based methodology to simulate 2D/3D cardiac ultrasound images," *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, vol. 56, pp. 404-409, 2009.

[11] J. Jensen, "Field: A program for simulating ultrasound systems," *Medical and Biological Engineering and Computing*, vol. 34, pp. 351-352, 1996.

[12] J. Jensen and N. Svendsen, "Calculation of pressure fields from arbitrarily shaped, apodized, and excited ultrasound transducers," *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, vol. 39, pp. 262-267, 1992.

[13] D. Zatari, N. Botros, and F. Dunn, "A simulation algorithm for ultrasound liver backscattered signals," *Ultrasonics*, vol. 33, pp. 469-474, Nov 1995.

[14] T. Yamaguchi and H. Hachiya, "Modeling of the Cirrhotic Liver Considering the Liver Lobule Structure," *Jpn J Appl Phys*, vol. 38, pp. 3388-3392, 1999.

[15] J. Jan, *Medical image processing, reconstruction, and restoration: concepts and methods*. Boca Raton,FL: Taylor & Francis Group, 2006.

[16] A. R. Crawford, X. Z. Lin, and J. M. Crawford, "The normal adult human liver biopsy: a quantitative reference standard," *Hepatology*, vol. 28, pp. 323-331, Aug 1998.

[17] R. I. Badea, M. Dudea, P. A. Mircea, and F. Stamatian, *[Clinical Ultrasonography]* vol. I. Bucuresti: Editura Medicala, 2000.

[18] L. P. Gartner and J. L. Hiatt, *Color textbook of histology*, 2nd ed. Philadelphia: W.B. Saunders, 2001.

[19] D. W. Fawcett, W. Bloom, and E. Raviola, *A textbook of histology*, 12th ed. New York: Chapman & Hall, 1994.

[20] A. Materka and M. Strzelecki, "Texture analysis methods a review," Technical University of Lodz, Institute of Electronics,1998.

[21] M. Lupsor, R. Badea, S. Nedevschi, C. Vicas, S. Tripon, H. Stefanescu, C. Radu, and M. Grigorescu, "The assessment of liver fibrosis using the computerized analysis of ultrasonographic images. Is the virtual biopsy appearing as an option?," *Acta Electrotehnica*, vol. 48, pp. 245-250, 2007.

[22] M. Lupsor, R. Badea, C. Vicas, S. Nedevschi, H. Stefănescu, M. Grigorescu, M. Radu, D. Crisan, Z. Sparchez, H. Branda, A. Serban, and A. Maniu, "The diagnosis performance of the attenuation coefficient computed on the ultrasonic image in comparison with the classical ultrasonographic examination, to quantify the hepatic fat content in nonalcoholic steatohepatitis patients.," *Proceedings of EASL Special Conference on NAFLD/NASH and related metabolic disease*, p. 169, 2009.

[23] M. Lupsor, S. Nedevschi, C. Vicas, R. Badea, M. Grigorescu, S. Tripon, H. Stefanescu, C. Radu, A. Serban, and T. Suteu, "Estimating the fibrosis stage in the human liver tissue using image processing methods on ultrasonographic images. Preliminary results," *Proc. EMMIT, 3rd Ed.*, pp. 311-318, 2007.

[24] C. Vicas, S. Nedevschi, M. Lupsor, and R. Badea, "Automatic detection of liver capsule using Gabor filters. Applications in steatosis quantification," *IEEE Proc.Intelligent Computer Comm. Process.*, pp. 133-140, Aug. 2009.

[25] C. Vicas, S. Nedevschi, M. Lupsor, and R. Badea, "Fibrosis detection using ultrasound and serological markers as features for additive logistic regression.," *2009 IEEE International Conference on Intelligent Computer Communication and Processing*, 2009.

[26] Y. M. Kadah, A. A. Farag, J. M. Zurada, A. M. Badawi, and A. B. M. Youssef, "Classification algorithms for quantitative tissue characterization of diffuse liver disease from ultrasound images," *Ieee Transactions on Medical Imaging*, vol. 15, pp. 466-478, Aug 1996.

[27] A. M. Badawi, A. S. Derbala, and A. M. Youssef, "Fuzzy logic algorithm for quantitative tissue characterization of diffuse liver diseases from ultrasound images," *Int J Med Inform*, vol. 55, pp. 135-147, Aug 1999.

[28] C. M. Wu, Y. C. Chen, and K. S. Hsieh, "Texture Features for Classification of Ultrasonic Liver Images," *Ieee Transactions on Medical Imaging*, vol. 11, pp. 141-152, Jun 1992.

[29] W. C. Yeh, S. W. Huang, and P. C. Li, "Liver fibrosis grade classification with B-mode ultrasound," *Ultrasound in Medicine and Biology*, vol. 29, pp. 1229-1235, Sep 2003.

[30] G. T. Cao, P. F. Shi, and B. Hu, "Liver fibrosis identification based on ultrasound images captured under varied imaging protocols," *J Zhejiang Univ Sci B*, vol. 6, pp. 1107-1114, Nov 2005.

[31] W. Lee, K. Hsien, and Y. Chen, "A study of ultrasonic liver images classification with artificial neural networks based on fractal geometry and multiresolution analysis," *Biomedical Engineering - Applications, Basis&Communications*, vol. 16, pp. 17-25, 2004.

[32] Y. Xia, D. G. Feng, and R. C. Zhao, "Morphology-based multifractal estimation for texture segmentation," *Ieee Transactions on Image Processing*, vol. 15, pp. 614-623, Mar. 2006.

[33] C. Abe, C. E. Kahn, Jr., K. Doi, and S. Katsuragawa, "Computer-aided detection of diffuse liver disease in ultrasound images," *Invest Radiol*, vol. 27, pp. 71-77, Jan 1992.

[34] A. Ahmadian, A. Mostafa, M. D. Abolhassani, and Y. Salimpour, "A texture classification method for diffused liver diseases using Gabor wavelets," *2005 27th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, vol. 1-7, pp. 1567-1570, 2005.

[35] G. Cao, P. Shi, and B. Hu, "Ultrasonic Liver Characterization Using Phase Congruency," *Proc. 27th Ann. Int. Conf. IEEE-EMBS* pp. 6356-6359, Jan. 2005.

[36] M.-H. Horng, Y. N. Sun, and X. Z. Lin, "Texture feature coding method for classification of liver sonography," *Computerized Medical Imaging and Graphics*, vol. 26, pp. 33-42, Jan-Feb 2002.

[37] I. Witten and E. Frank, *Data Mining: Practical machine learning tools and techniques*: Morgan Kaufmann Pub, 2005.

[38] V. Popovici, "Assessment of classification models for medical applications," presented at the Workshop on Computers in Medical Diagnoses, Cluj-Napoca, Romania, 2009.

[39] K. Laws, "Texture Energy Measures," *Proc. Image Understanding Workshop*, pp. 47-51, 1979.

[40] K. Laws, *Textured image segmentation*: University of Southern California, Los Angeles Image Processing Institute, 1980.