Detection of Anatomical Structures on Ultrasound Liver Images Using Gabor Filters

C. Vicas¹, M. Lupșor², R. Badea², S. Nedevschi¹
¹Technical University of Cluj-Napoca, Romania, cristian.vicas@cs.utcluj.ro, sergiu.nedevschi@cs.utcluj.ro
²3rd Medical Clinic, Cluj-Napoca, Romania, mmlupsor@yahoo.com

Abstract - In this paper we present a novel method of detecting anatomical structures from liver ultrasound images. Local texture anisotropy is evaluated using a specially designed bank of Gabor filters. We show that the proposed method is well suited for processing liver ultrasound images and it is robust with respect to its parameters. A practical application of this method is then presented. Image features are computed on detected structures. These features are then used in conjunction with a classifier (Support Vector Machines) to discriminate between left and right lobe liver images. This classifier is used to label ultrasound images. The labeling algorithm obtained a very low error rate (<1%). It can successfully replace the human expert in image labeling.

I. INTRODUCTION

The aim of this paper is to introduce a novel technique that detects large anatomical features on ultrasound images. Liver tissue produces a speckled image that can be viewed as a relatively homogenous texture [1]. Large anatomical structures like liver capsule produce coherent and correlated speckles that give the local texture an orientation, hence the anisotropic property of the texture in that region.

Texture orientation plays a central role in edge or corner detection [2], image enhancement [3], optical flow [4], stereo disparity estimation [5], phase congruency [6] etc. There are many approaches [7] ranging from estimation of the image gradients [8], least-squares fitting of a plane in Fourier space [9], autocorrelation [10], steerable filters [11], Principal Component Analysis [12], structure tensors [13] etc.

Using Gabor filters [14] we propose a measure of the local texture anisotropy. Using this information one can decide if a texture region contains an anatomical structure or contains only regular tissue echoes generated by homogenous liver tissue.

This technique is then applied in automatic labeling of liver ultrasound images. The ultrasound machine embeds in the DICOM image some information like patient name, date, time, the probe frequency, gain, etc. Information like what organ is studied, or what is the angle of the visualization plane with the patient’s body must be written manually by the investigator. This process is relatively laborious and time consuming when a lot of images for each patient are acquired.

In this paper we also present an application that labels the images acquired following a certain protocol. This protocol establishes that the studied organ is the liver, establishes certain machine settings that should remain constant for all patients and establishes the sections that are observed. This protocol provides for each patient a series of images acquired from left and right lobe of liver. In Fig. 1 are shown two liver images.

None of these images are labeled by the physician in the moment of acquisition. Working with unlabeled images could be an impediment when we want to use these images to develop non invasive diagnosis tools for certain pathologies, like liver fibrosis. The golden standard procedure used to diagnose fibrosis is biopsy [15-16]. This biopsy is performed on the right liver lobe. Hence, it is important that the image processing is performed only on the right lobe images.

The paper is structured as following: in section 2 we give a short compendium on Gabor filters. In section 3 we present the proposed method and the application. Section 4 presents the experimental setup and the results on a large set of images. Discussions, conclusions and future research directions to this paper are presented in section 5.

II. GABOR FILTERS

In spatial domain a Gabor filter is a product between a complex sinusoidal carrier and a 2D Gaussian shaped envelope: \( g(x,y) = s(x,y)w(x,y) \) where \( s(x,y) \) is the complex sinusoid and \( w(x,y) \) is the Gaussian envelope [17-20].

In spatial domain \( s(x,y) = e^{j(2\pi(xu_0+yu_0)+\phi)} \) where \((u_0,v_0)\) define the spatial frequency and \( P \) define the phase. If \((u_0,v_0)\) represents a point in frequency space then it can be specified in polar coordinates in terms of \( F_0 \) and \( \gamma_0 \):

\[
F_0 = \sqrt{u_0^2 + v_0^2}, \quad \gamma_0 = \arctan \left( \frac{u_0}{v_0} \right).
\]

Then \( s(x,y) \) becomes

\[
s(x,y) = e^{j(2\pi F_0(x\cos \gamma_0 + y\sin \gamma_0)+\phi)}, \quad \text{where} \quad j = \sqrt{-1}.
\]

The Gaussian envelope is a simple 2D Gaussian function:

\[
w(x,y) = Ke^{-\frac{(x-x_0)^2}{A^2} - \frac{(y-y_0)^2}{B^2}},
\]

where

\[
(x-x_0) = (x-x_0)\cos \theta + (y-y_0)\sin \theta,
\]

\[
(y-y_0) = -(x-x_0)\sin \theta + (y-y_0)\cos \theta,
\]

and \((x_0,y_0)\) are the coordinates for the peak of the Gaussian and \( \theta \) is the rotation angle. \( A \) and \( B \) are standard deviations. In this paper the following conventions are assumed: \((x_0,y_0) = (0,0); \theta = \gamma_0; \quad P = 0; \quad F_0 = 1/\text{Period}; \quad K = 2\pi/AB, \quad \text{where} \quad \text{Period} \quad \text{is the wavelength of the sinusoidal carrier. The definition of the Gabor kernel in spatial domain becomes:}

\[
g(x,y) = \frac{2\pi}{AB} e^{-\frac{(x\cos \theta + y\sin \theta)^2}{A^2} - \frac{(-x\sin \theta + y\cos \theta)^2}{B^2}} \times e^{\frac{2\pi}{\text{Period}}(x\cos \theta + y\sin \theta)}.
\]

Knowing that

\[e^{ix} = \cos x + j\sin x,\]
s(x,y) can be decomposed in two parts, real and imaginary part. As a result, in spatial domain, the complex Gabor function can be decomposed in two convolution kernels, \( \tilde{g}_r(x,y) \) and \( \tilde{g}_i(x,y) \).

By performing the Fourier transform on the Gabor function and expressing in polar coordinates one can compute the magnitude:

\[
G(u,v) = 2\pi e^{-\pi(\Delta\theta)^2} \text{ sinc}(\Delta F) \cdot \sqrt{\frac{2\pi}{\Delta F}}
\]

where \( G(u,v) \) is the Fourier transform of \( g(x,y) \) and \( G(u,v) \) represents the magnitude of complex number \( G(u,v) \).

As one can note, the magnitude of the response depends on frequency and an orientation bandwidth. The bandwidth is computed in terms of half magnitude response.

The half-response is defined as the location in the frequency domain where the \( |G(u,v)| = 0.5 |G(u_0,v_0)| \). Usually the frequency bandwidth is measured in octaves. Let \( F_{\min} < F_0 < F_{\max} \) be the maximum and minimum frequencies such that

\[
|G(u,v)| = 0.5 |G(u_0,v_0)|
\]

and

\[
F_{\min} \leq \sqrt{u^2 + v^2} \leq F_{\max}.
\]

The frequency bandwidth is defined as

\[
\Delta F = \log_2\frac{F_{\max}}{F_{\min}},
\]

where \( \Delta F \) is expressed in octaves.

In a similar way is defined the angle bandwidth

\[
\Delta \theta = \theta_{\max} - \theta_{\min}
\]

where \( \Delta \theta \) is expressed in radians.

From equation (5)-(9), one can compute that

\[
\Delta F = \log_2\frac{F_0}{F_0 - (1/A)} \sqrt{2\pi/\ln 2}.
\]

and

\[
\tan\left(\frac{1}{2} \Delta \theta\right) = (1/2B\tan(\theta_0)) \sqrt{2\pi/\ln 2}.
\]

From the equations (10),(11) the terms \( A \) and \( B \) can be determined with respect to central angle, frequency and bandwidths:

\[
A = \frac{1}{F_0} \sqrt{\frac{\ln 2}{\pi}} \left(\frac{2\Delta F + 1}{2\Delta F - 1}\right)
\]

and

\[
B = \frac{1}{F_0} \sqrt{\frac{\ln 2}{\pi}} \left(1/\tan\frac{1}{2} \Delta \theta\right).
\]

By looking at equations (1),(2),(3),(12),(13) one can express a Gabor filter in terms of:

- \( 1/\text{Period} \) – central frequency response
- \( \theta_0 \) – central orientation response
- \( \Delta F \) – frequency bandwidth
- \( \Delta \theta \) – angle bandwidth

In actual applications a Gabor filter bank is used. Usually this filter bank is constructed in such a way that it will cover all frequency spectra by ensuring that the angle and frequency half-magnitude points of adjacent filters touch each other.

One major drawback for Gabor filters is that they have a DC response dependent on the filter frequency bandwidth:

\[
DC_g = e^{-\frac{\ln 2\Delta F + 1}{2\Delta F - 1}}
\]

Some DC corrections were investigated in [21]. The authors note that for small bandwidths (\( \Delta F < 0.7 \)) the compensation is not needed.

Based on equation (4) the complex Gabor filter can be decomposed in two convolution kernels. By applying these two kernels on an image we obtain the real and the imaginary response of the filter. In most papers the only significant value is the magnitude of the response:

\[
|g(\Pi) \cdot I| = \sqrt{|g_{re}(\Pi) \cdot I|^2 + |g_{im}(\Pi) \cdot I|^2}
\]

where \( \Pi = (F_0, \theta_0, \Delta F, \Delta \theta) \) represents the parameters for the Gabor filter. The \( \cdot \) operation represents the convolution.

When generating the convolution in spatial domain the kernel width should be \( W=k \cdot \max(A,B) \) where \( A \) and \( B \) are the standard deviations defined in equations (12),(13) and \( k \) should take values as 2.5 or 3 depending on the desired precision.

For large bandwidths and high frequencies the values for \( A \) and \( B \) are rather small thus the kernels are small. For small bandwidths and/or lower frequencies the spatial kernels tend to become large. In these cases one should consider performing the convolution in Fourier space. Care should be taken when choosing the \( F_0 \) frequency in order to avoid aliasing [21].

### III. Proposed Method

#### A. Detection of Anatomical Structures

The main idea behind the proposed method is that in a region without any edges, the texture is relatively homogenous. In a region that contains an edge or other linear feature the texture presents an orientation, hence, anisotropy. Liver ultrasound images are characterized by a fine grained texture. This texture is not directly related to the liver tissue but to the interference patterns of adjacent structures that have the dimension close to the ultrasound wavelength [1, 22].

The interfaces between larger structures and liver tissue have a different reflectivity and generate a larger echo. Because these structures have their own microstructures the interference pattern is still present. As a result, these structures generate fuzzy edges. Sometimes these edges are marked only by a subtle texture difference or by a soft difference in luminosity.

In addition to that, some authors claim that the grey level difference pattern follows a Rayleigh distribution [23]. Present method assumes that at anatomical interface locations
the texture has a degree of anisotropy and tries to quantify this anisotropy.

With the help of a bank of oriented filters one can estimate the magnitude response of some frequency band as a function of orientation for each image location. For an isotropic texture the distribution of the response is fairly uniform. For an oriented texture there is a peak in the response at texture orientation. Using a proposed formula we compute the “energy” of the filter responses and determine if there are peaks in some orientations.

The filter bank is implemented using Gabor filters. A central frequency along with a number of grade divisions are established. Each filter in the bank has the same central frequency response $F_0$, the same $\Delta F$ and $\Delta \theta$ bandwidths. The only difference is the central orientation response, $\theta_0$. The angle bandwidth and the individual $\theta_0$ are computed knowing the desired number of filters in the bank.

Let $Mag_o(x,y)$ be the magnitude response of filter $o$ at location $(x,y)$. Let,

$$Mag_o(x_w,y_w) = \sum_{x,y} \frac{Mag_o(x + x_w,y + y_w)}{W}$$ \hspace{1cm} (16)

be the average response of the filter over a window. In equation (16) $W$ denotes the width of the square smoothing window. $W$ is approx. $A/2$ where $A$ is the frequency standard deviation, as defined in equation (12) and $(x_w,y_w)$ denote the window coordinates. These windows overlap by a factor of 0.25.

The “energy” for the window centered in $(x_w,y_w)$ is then estimated using the following proposed equation:

$$E(x_w,y_w) = \frac{\sum_o (Mag_o(x_w,y_w))^2}{\left(\sum_o Mag_o(x_w,y_w)\right)^2}$$ \hspace{1cm} (17)

where $o$ denotes the orientation of the filter used to compute the response.

For each pixel, the energy value is a linear combination between the energies of the windows that overlap the pixel. Each pixel energy value is weighted according to the distance between current pixel and the center of the corresponding window. A threshold is applied to isolate the texture regions with high energy. Each $Mag_o$ term contains a DC response as shown in equation (14). The DC response depends only on filter bandwidth. Each filter that is employed here has the same bandwidth and is applied on the same region of the image. The DC response for each $Mag_o$ term is identical for each $o$. One can easily see that the DC terms cancel themselves out in equation (17). As a result, no DC compensation is required for the proposed solution, regardless of the frequency bandwidth value.

The method depends on some parameters. First we have to decide what frequency band (or bands) we investigate. Second, a threshold must be established for the energy. If the energy is lower than selected threshold then the surrounding texture can be considered isotropic. The frequency band selects the dimensions of the features that we want to detect, and the threshold indicates the sensibility of the method.

Although several frequency bands can be used this is not recommended because at most of the natural textures the orientation is dependent of scale. The method is limited to linear or quasi-linear features because for a circular feature the proposed method will detect an isotropic texture. However, if the size of the feature is larger than the filter selection size, the feature’s edges will be detected. Because of the band pass filter the smaller features will be ignored. As a result, it is possible that the texture has a strong orientation at finer scales and the proposed method will correctly detect larger features ignoring the finer scale anisotropy.

This is a very useful behavior because ultrasound images are characterized by this fine scale anisotropy. The resolution in axial direction is greater than the resolution in lateral direction [1, 22-23]. Moreover, these properties vary with the relative distance from the probe.

### B. Ultrasound Image Labeling

The proposed method is used to discriminate between left and right lobe images. Empirically one can note that the right liver lobe produces a larger area in the ultrasound image than left lobe. The liver capsule appears lower in the right images than in the left. Another aspect is that the surrounding area of the liver contains many artifacts, most of them produced by the bowel, stomach, etc. The proposed method will detect many features in these regions.

The difference between left and right images is given by the localization and quantity of these responses. In order to characterize the spatial localization of the detected artifacts the image is divided in several non overlapping regions. Each region is then described using the amount of detected artifacts. One should expect for example, that for right liver images the regions in the middle of US image contain fewer artifacts than for the images acquired from left lobe.

First step of the labeling algorithm is a preprocessing step. The ultrasound image suffers an inverse polar projection as in Fig. 2. The center of the projection overlaps the virtual source of ultrasound waves.

This transformation ensures that the resulting image has the same physical properties on any horizontal or vertical direction. Pixels belonging to any horizontal direction are obtained after the same period of time form the returning echo signal. Moreover, the pixels on any vertical direction belong to the same A line. These properties ensure that the properties of the image are translation invariant. These sentences are not valid for original ultrasound image.

The proposed method is applied to this polar projected image. We employ a filter bank having 6 orientations and a single frequency band as in Fig. 3.

The binary image generated by the proposed method is then divided in 32 areas using a 4x8 grid as in Fig 4. The grid dimensions are a compromise between the need of describing the spatial localization of the artifacts and the need to keep the feature dimensionality low.

From each cell grid we measure the area of the detected anatomical features.

We also measure the total area of detected features. Each cell in the grid is described using two numbers: the absolute area and the relative area ratio. Relative area is computed dividing the area in the current cell with the detected area on the entire.
image. Using the values from all grid cells, each cell providing two features, we build a feature vector. A series of images are manually labeled. A feature vector is computed from each image, creating a dataset. This dataset has two classes; each vector belongs to the left or to the right class, depending on the image from where it was computed.

A support vector machine (SVM) classifier is trained using this labeled dataset. SVM uses a Gaussian (RBF) kernel [24-26]. The two main meta parameters of the classifier (cost $C$ and exponent $\gamma$ for the RBF kernel) are optimized using an exhaustive search schema (grid search) [24]. A value set is proposed for each parameter and the algorithm search all possible combinations. Each combination is evaluated using 2-fold cross validation.

Best found pair, along with surrounding pairs is then evaluated using 5-fold cross validation and the best pair is selected. The search schema returns a classifier trained on the entire input dataset. The performance of this classifier is evaluated on a new, distinct dataset [27].

IV. EXPERIMENTAL RESULTS

In this section the proposed method is used to label ultrasound images.

The images are acquired from patients suffering of various diffuse liver diseases (hepatitis B,C, steatohepatitis, etc). Over 200 patients were scanned using a 5.5 MHz phased array probe. Almost all the machine settings were kept constant for all patients. The physician was allowed to alter the focal depth in such a way that the best image quality is just below the liver capsule. Left and right lobe images were acquired.

First step in applying proposed method is to establish the parameters like the Gabor frequency response, bandwidth and energy threshold. These three parameters were established on a set of 80 liver images using a human expert. These images were processed using various parameter sets. These parameter sets were chosen consecutively, in a greedy like search. The initial set was: central frequency 1/16, bandwidth 1.5 octaves and the threshold was set to 0.1. First, the central frequency was optimized, then bandwidth and in the end, the threshold.

Another iteration was performed, trying to further optimize the found parameter set. No major modifications were made. The final parameter set was: frequency 1/24, bandwidth 2 and threshold 0.17. There were tested only 18 parameter sets until the physician was satisfied with the results. In Fig. 5 are shown several ultrasound images processed using this parameter set.

The next step was to train the classifier. In present experiment a number of 1000 images were manually labeled by the same human expert. We ensured that in this lot 500 images were from right lobe and 500 from left lobe. These images were processed using the proposed method with the found parameter set. The resulting dataset consisting of 1000 labeled feature vectors was then used to train the SVM classifier. Meta parameters that control the SVM training were searched using the grid search technique explained earlier. Cost $C$ meta parameter takes values $2^{-3}$, $2^{-2}$, $2^{-1}$ … $2^8$. Exponent $\gamma$ takes values $2^{-15}$, $2^{-14}$ … $2^4$. 

DRAFT. Full paper at: http://dx.doi.org/10.1109/AQTR.2010.5520792
Another unlabeled dataset was built by processing 5000 images. The classification algorithm trained at the previous step was used to label these images. The human expert investigated the dataset and noted the wrongfully labeled images.

Error rate was reported, as the ratio between the number of incorrect labeled images and total number of images.

The recorded error rates were 0.92%. We noted that the misclassified images are distributed similar between the two classes.

V. DISCUSSIONS AND CONCLUSIONS

The anatomical features detection method does not have a very good spatial localization. Depending on the threshold value the detected regions might shift. However, the value of threshold does not influence greatly the number or the shapes of detected features.

Few parameter sets were tried before a satisfying detection was achieved. We conclude that the method is robust with respect to its main parameters.

The detection of anatomical features can be applied to other areas of research. Detected areas could act as a seed for a dedicated segmentation algorithm, like watershed [28-30].

Further development in the direction of isolating the US image features (including major blood vessels, shadowing artifacts, etc) is achievable by tuning the frequency bands and creating complex rules for detecting each type of artifact. This will allow classification of the detected features (differentiation between upper and lower liver capsule, blood vessel types, cysts, etc.). Another line of research could be the employment of the structure tensor [31] to detect the anisotropy. Structure tensors are known to detect that there is an orientation even in the presence of circular features.

Ultrasound image labeling is a very useful tool because it can be used to replace a human expert. Human experts, especially physicians are a scarce resource. The high performance of this tool allows immediate integration with the existing image handling tools.

REFERENCES