J Med Syst
DOI 10.1007/s10916-010-9588-7

ORIGINAL PAPER

TND: A Thyroid Nodule Detection System for Analysis of Ultrasound Images and Videos

Eystratios G. Keramidas · Dimitris Maroulis · Dimitris K. Iakovidis

Received: 21 April 2010 / Accepted: 1 September 2010

Abstract In this paper, we present a computer-aided-diagnosis (CAD) system prototype, named TND (Thyroid Nodule Detector), for the detection of nodular tissue in ultrasound (US) thyroid images and videos acquired during thyroid US examinations. The proposed system incorporates an original methodology that involves a novel algorithm for automatic definition of the boundaries of the thyroid gland, and a novel approach for the extraction of noise resilient image features effectively representing the textural and the echogenic properties of the thyroid tissue. Through extensive experimental evaluation on real thyroid US data, its accuracy in thyroid nodule detection has been estimated to exceed 95%. These results attest to the feasibility of the clinical application of TND, for the provision of a second more objective opinion to the radiologists by exploiting image evidences.

Keywords Computer-aided-diagnosis · Thyroid · Nodule · Ultrasound

Introduction

Among all radiological modalities, ultrasound (US) possesses a rare combination of advantages including portability, harmlessness, real-time data acquisition and affordability. Modern ultrasonographic systems can provide high resolution images allowing physicians to deduce useful information concerning the tissue characterization and structure. Especially for certain types of diagnoses including the prostate, breast, lung, and the thyroid gland, US is considered to be the dominant imaging modality.

The thyroid is a small gland which produces hormones that affect heart rate, cholesterol level, body weight, energy level, mental state and controls a host of other body functions. There are different types of thyroid cancer, but the most common ones (papillary carcinoma and follicular carcinoma) are highly curable if detected early. According to epidemiologic studies palpable thyroid nodules occur in 4%–7% of the population, but nodules found incidentally on US examinations show a frequency of 19%–67% [1]. Overall, a significant percentage of nodules detected by US examination, escape detection on clinical examination. The challenge is to utilize US imaging to detect thyroid nodules that are clinically occult due to their texture, size or shape.

Computerized analysis improves medical image interpretation, providing a reliable second opinion in detecting lesions, assessing disease severity, and leading to more accurate diagnostic decisions. Several software systems have been implemented aiming at the Computer Aided Diagnosis (CAD) on US images for different soft tissue internal organs, including those for liver [2–4], breast [5–7], prostate [8, 9]. Only a few US CAD systems have been proposed for the evaluation of the thyroid gland. Smutec et al. [10] proposed an approach for automatic differentiation between inflamed and healthy thyroid tissue. This approach was based on Muzzolini's spatial features [11] and Haralick's co-occurrence features (CM) [12]. Furthermore, a Support Vector Machines based system has been presented by Tsantis et al. [13] for assessing the malignancy risk of thyroid

E. G. Keramidas (✉) · D. Maroulis
Department of Informatics and Telecommunications,
University of Athens,
Panepistimiopolis,
GR-15784 Athens, Greece
e-mail: e.keramidas@gmail.com

D. K. Iakovidis
Department of Informatics and Computer Technology,
Technological Educational Institute of Lamia,
GR-35100 Lamia, Greece

Published online: 14 September 2010

nodules. This system combined statistical features computed from local grey-level histograms (GLH) and from Haralick's co-occurrence matrices [12]. More recently, a level-set active contour model has been proposed for automatic delineation of thyroid nodules [14], however this method requires that an initial contour is manually defined.

Such systems offer a less subjective means for the interpretation of thyroid US images, based on a variety of texture representation approaches summarized in Table 1. The use of GLH can be found in many approaches including the studies of Mailloux et al. [15, 16]for the discrimination of different types of thyroid tissue, the study of Morifuji [17] for malignancy determination of thyroid nodules, and the study of Hirning et al. [18], for the quantification and classification of echographic findings in the thyroid gland. These studies attest to effectiveness of GLH for the characterization of thyroid tissue. Nevertheless, GLH based features can not encode any information related to the spatial distribution of image pixels [19]. Thus, more recent approaches combine GLH features with statistical features for texture analysis of ultrasound thyroid images. Such statistical features include Haralick's co-occurrence features investigated in [13, 20], Muzzolini's spatial features proposed in [10], Radon Transform features presented in [21] and Local Binary Pattern features investigated in [22]. Most of these approaches use GLH and/or second or higher order statistical textural descriptors for the representation of thyroid ultrasound patterns, however, none of them takes into account the noise-originated uncertainty inherently present in all ultrasound images.

These patterns are categorized by sophisticated classification algorithms, capable of dealing with complex disease states requiring the integration of pre-diagnosed data and clinical information. Physicians involved in the process of diagnosis based on medical images are often challenged or discouraged by the complicated, hard to use, and expensive diagnosis software systems. Additionally most of these systems depend on the user to manually define regions of interest (ROI), within which the image processing is performed. This makes CAD systems susceptible to the physicians' experience and subjectivity.

Table 1 Literature review of feature extraction approaches for thyroid ultrasound image analysis

Feature extraction approach	Authors—Ref
Grey level histogram	Mailloux et al.—[13, 14]
	Morifuji et al.—[15]
	Hirning et al.—[16]
Muzzolini's features	Smutek et al.—[10]
Cooccurrence matrix	Tsantis et al.—[11]
	Skouroliakou et al.—[18]
	Smutek et al.—[10]
Radon transform	Savelonas et al.—[19]
Local binary patterns	Keramidas et al.—[20]

In this study, we propose an original methodological scheme for the detection of nodular tissue in US thyroid images and videos. This scheme involves a novel algorithm for automatic ROI definition through the unsupervised detection of the boundaries of the thyroid gland. We also propose a fuzzy logic-based approach to obtain an uncertainty-aware representation of thyroid ultrasound patterns. This approach involves the combination of fuzzy local binary pattern distributions (preliminarily investigated in [23]) with fuzzy grey-level histograms (FGLH). The proposed scheme has been implemented as a prototype exploratory analysis system, named TND (Thyroid Nodule Detector). TND offers a simple, practical, and user friendly interface providing a means to tune all the parameters that are relevant to the methods involved. A considerable series of experiments were conducted, for the evaluation of the best suited methodologies and parameter values involved in the proposed scheme. The ultimate purpose of the TND system, is the provision of a second opinion, for nodules that may not be diagnosed due to poor US image quality and/or for medical students, junior radiologists and others needing a second opinion.

The rest of this paper is organized as follows. Methodology section introduces a description of the main components of the proposed scheme. Then in section Performance Evaluation an experimental evaluation study on real US thyroid data is presented, demonstrating the effectiveness of the proposed scheme. Finally in the last section the conclusions of this study are summarized.

Methodology

The proposed scheme for thyroid nodule detection consists of five components (Fig. 1), namely: pre-processing, ROI definition, feature extraction, feature classification and post-processing. The pre-processing component incorporates all necessary modalities to convert input data, either images or videos, to a unified form suitable for further processing. The second component involves a novel algorithm for automatic detection of the thyroid boundaries and definition of a respective ROI that includes only the thyroid gland. The third component involves a set of feature extraction methods, where image texture and echogenity information is encoded via feature fusion. In the fourth component, classification algorithms categorize the patterns formed by the extracted features into predefined groups related to diagnosis. Finally, the output of the classification component is post-processed and visualized either as binary or as grayscale images or videos. Details on the components of TND system are described in detail in the following sub-sections.

Data preprocessing

In the preprocessing component an image normalization method has been utilized. In US images dynamic range of pixels intensities can be quite narrow, and their distributions may vary substantially. To deal with this phenomenon, pixel values are redistributed proportionately to cover the entire range of display brightness for each image or video frame [24].

Automatic ROI definition

The thyroid gland consists of two lobes located along either side of the trachea. Each lobe is surrounded by a thin fibrous capsule [25]. That capsule can be recognized in longitudinal thyroid US images as thin hyperechoic lines. Considering that thyroid nodules reside only within the thyroid parenchyma, the image analysis operations should be performed only within the thyroid boundaries.Although various approaches have been proposed for segmenting US images from thyroid, breast and prostate [26–29], even the most popular of them, such as active contour methods, present sensitivity to the gradient of the edge, and require physicians to roughly outline an initial contour of the area. To this end we propose a novel algorithm for the detection of the boundaries of thyroid gland, which will be referred to as TBD-2 in the rest of this paper. The proposed algorithm improves the Thyroid Boundaries Detection (TBD) algorithm presented in [22], and it is capable of tracking the boundaries of the thyroid gland (Fig. 2(c)) instead of providing a rough rectangular approximation of the gland as the original TBD (Fig. 2(b)). Restraining image analysis within the ROI defined by the gland's boundaries results in a consequent reduction of the computational cost as well as of the expected error rate in the subsequent classification task, since the irrelevant image regions, which would otherwise be considered as negative, are significantly reduced. TBD-2 is implemented in four steps: a) image pre-processing, b) analysis of the pre-processed image, c) identification of the areas corresponding to the thyroid boundaries, and d) refinement of the detected boundaries.

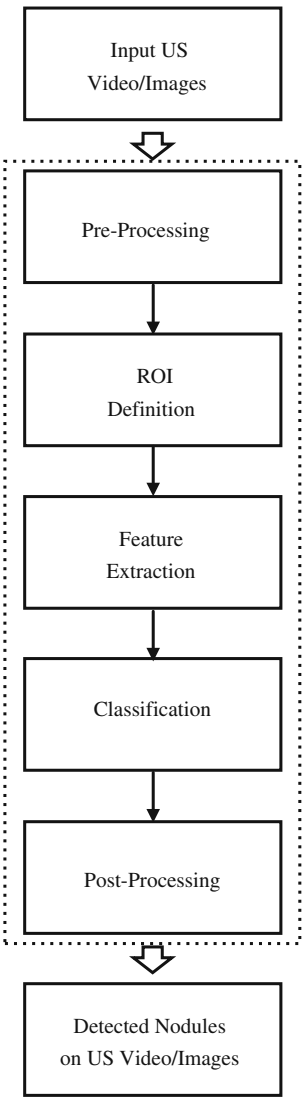

Fig. 1 Block diagram of the proposed CAD system

In the first step of the TBD-2 algorithm, a US image of $N \times M$ pixels and G grey levels is uniformly quantized into z discrete grey levels. By such a quantization process, a coarse segmentation of the US image is obtained, that accentuates the hyperechoic bounds of the thyroid gland. Let g_i be the original grey value of a pixel and g_i' the grey value of that pixel after quantization. Then g_i' can be computed as follows:

$$g_i' = \left\lfloor \frac{(G-1)\cdot\left(\left\lfloor\frac{2\cdot g_i\cdot(z-1)}{G-1}\right\rfloor + \left\lfloor\frac{2\cdot g_i\cdot(z-1)}{G-1}\right\rfloor \%2\right)}{2\cdot(z-1)} \right\rfloor \quad (1)$$

In the second phase the quantized image produced is sampled from top to bottom with K horizontal non-overlapping stripes of size $h \times M$ pixels.

For each stripe a weighted sum S_n is estimated by the equation:

$$S_n = \sum_{g=0}^{G} \left(w(g) \sum_{P_{g,n}} 1 \right) \quad (2)$$

where n is the stripe index incrementing from top to bottom, and $P_{g,n}$ represents the set of pixels of stripe n with grey value g.

It should be noted that $w(g)$, in Eq. 2, denotes a quadratic weight function which aims to amplify the contribution of higher grey levels, which clearly appear as hyperechoic lines in US images after the quantization process of the first stage. This weight function is defined as $w(g) = a \cdot g^2 + \beta \cdot g + \gamma$ where α, β, γ are constants. Constants β, γ are chosen so as

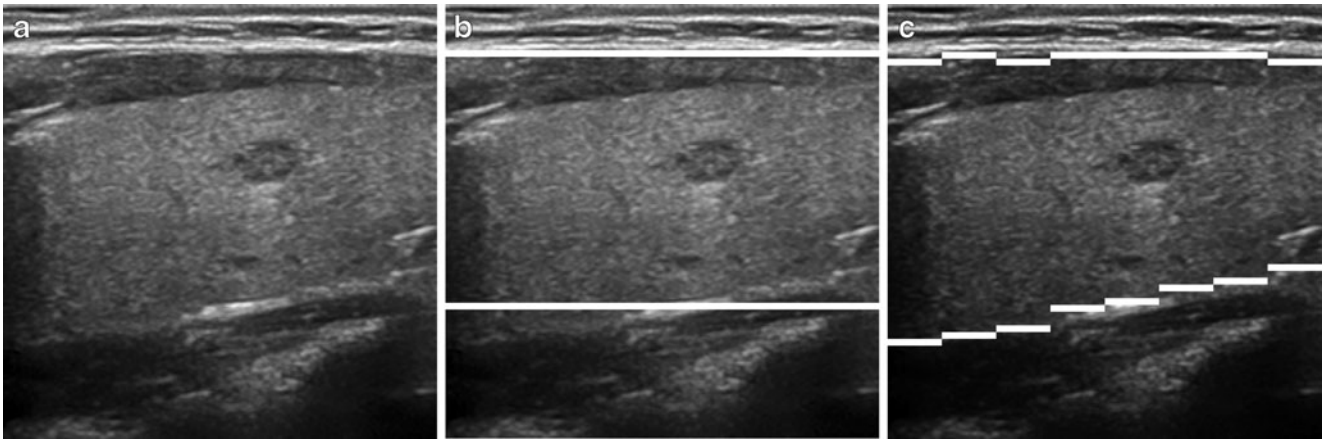

Fig. 2 **a** US image with thyroid nodule. **b** Detected thyroid boundaries superimposed to the original input image through TBD algorithm. **c** Refined detected thyroid boundaries superimposed to the original input image through TBD-2 algorithm

to satisfy $w(0)=0$ and $w(1)=G$. Therefore $w(g)$ is finally derived by the following equation:

$$w(g) = \left(a + \frac{1 - a \cdot G^2}{G \cdot g}\right) \cdot g^2 \tag{3}$$

A measure directly proportional to the rate of change of S_n, between two successive stripes is computed as follows:

$$D_n = \left.\frac{d(S_i)}{di}\right|_{i=n} \cdot S_n\,, \quad n = 1, 2, \ldots K \tag{4}$$

Then stripes that contain the outer and the inner thyroid boundaries are selected. If n_{outer} and n_{inner} are the stripe indices that correspond to a rough estimation of the outer and the inner boundaries respectively, then n_{outer} and n_{inner} should satisfy the following conditions:

$$\begin{cases} n_{outer} = \arg\min\limits_{n}[D_n \cdot \omega_1(n)] \\ n_{inner} = \arg\max\limits_{n}[D_n \cdot \omega_2(n)] \\ n_{inner} - n_{outer} > \Delta, \quad \Delta > 0 \end{cases} \tag{5}$$

where $\omega_1(n)$ and $\omega_2(n)$ are defined by the following equations:

$$\omega_1(n) = \begin{cases} \log\left(\frac{\lambda \cdot K - n}{\lambda \cdot K} + 1\right) & n < \lambda \cdot K \\ 0 & n \geq \lambda \cdot K \end{cases} \tag{6}$$

$$\omega_2(n) = \begin{cases} \log\left(\frac{n - (1-\lambda) \cdot K}{\lambda \cdot K} + 1\right) & n < (1-\lambda) \cdot K \\ 0 & n \geq (1-\lambda) \cdot K \end{cases} \tag{7}$$

$$\lambda = 1 - \frac{\Delta}{N} \tag{8}$$

Parameter K is the total number of stripe samples per image, Δ represents a minimum anteroposterior diameter of the thyroid gland, and the logarithmic weight functions $\omega_1(n)$ and $\omega_2(n)$ bias n_{outer} and n_{inner} towards the upper and lower image regions, respectively.

Because of round shape of the thyroid lobes, their boundaries on longitudinal US images are not always perfect horizontal lines. Thus a refinement phase for a more detailed detection of these boundaries appears to be necessary. In this phase the image area around the n_{outer} stripe of height $3h$ is vertically sampled from top to bottom with non overlapping rectangular windows of dimensions $l' = M/v$ and $h' = 2h/v$, where v is a parameter controlling the boundary detection detail. For each window a weighted sum $S_{n'}$ is computed according to Eq. 2. Then the window with the maximum sum value on each column is chosen as the one over the hyperechoic lines of the outer boundaries of each lobe. The same steps are applied around n_{inner} stripe for the detailed detection of the inner boundaries of each lobe (Fig. 2c).

A pseudo code of the TDB-2 algorithm described above is depicted in Fig. 3.

Feature extraction

Feature extraction is performed on sub-images sampled in a raster scanning fashion with a sliding window of user-defined size and sliding step, which allows an overlap between the consecutive samples. The novel combination of fuzzy distributions that is implemented in the proposed scheme as well as the calculations involved are described in the following subsections.

1) Textural Features

Texture patterns appearing in US images can be represented by a fuzzy distribution of Local Binary Patterns, referred to as Fuzzy Local Binary Patterns (FLBP) features [23]. Although the original approach of Local

```
Step 1:   Quantize_Image()

Step 2:   For n = 1 to Number_of_Stripes do
              I_n = Get_Next_Stripe()
              S_n= Compute_Weighted_Sum(I_n)
          EndFor
          For n = 1 to Number_of_Stripes -1 do
              D_n = Compute_Derivative_Times_Sn(S_n)
          EndFor

Step 3:   n_outer =Find_Next_Min( [D_n*ω1(n)])
          n_inner  = Find_Next_Max( [D_n*ω2(n)])
          While (TRUE)
              If( n_inner - n_outer > ( T/h ) ) then
                 n_inner is the stripe index of the inner boundary
                 n_outer is the stripe index of the outer boundary
                 Exit
              Else
                  k  = Find_Next_Min([D_n*ω1(n)])
                  l  = Find_Next_Max([D_n*ω2(n)])
                  If( Choose_Next_To_Alter()== n_outer )
                     n_outer = k
                  Else
                     n_inner = l
                  EndIf
                     Continue;
              EndIf
          EndWhile
Step 4: For i = 1 to Number_of_Columns do
          n_refined_outer = argmax_n'(S_n')
          n_refined_inner = argmax_n'(S_n')
          EndFor
```

Fig. 3 Pseudo code for ROI definition schema

Binary Pattern (LBP) [30] has been used successfully in many studies [31], it has also been proven to be sensitive to small variations of the pixel intensities usually caused by noise. The FLBP is an enhanced extension of the LBP approach, capable of better coping with speckle noise [23], a common characteristic of all US images [32].

Contrary to the original LBP operator, where a single LBP code characterizes a 3×3 neighbourhood, in FLBP approach, a neighbourhood can be characterized by more than one LBP codes. Figure 4 illustrates FLBP feature extraction scheme, where multiple LBP codes are involved in the characterization of a 3×3 neighbourhood.

The degree, to which each LBP code characterizes a neighbourhood, depends on the membership functions $m_0()$ and $m_1()$ computed for each peripheral pixel $i \in [0, 7]$. Membership functions $m_0()$ and $m_1()$ can be defined as follows:

$$m_0(g_i) = \begin{cases} 0 & \text{if } \Delta g_i \geq T \\ \dfrac{T - \Delta g_i}{2 \cdot T} & \text{if } -T \leq \Delta g_i < T \\ 1 & \text{if } \Delta g_i \leq -T \end{cases} \tag{9}$$

$$m_1(g_i) = \begin{cases} 1 & \text{if } \Delta g_i \geq T \\ \dfrac{T + \Delta g_i}{2 \cdot T} & \text{if } -T \leq \Delta g_i < T \\ 0 & \text{if } \Delta g_i \leq -T \end{cases} \tag{10}$$

where g_i is the grey level value of pixel i (Fig. 4a), and $\Delta g_i = g_i - g_{center}$. For both $m_0()$ and $m_1()$, $T \in [0, G]$ represents a parameter that controls the degree of fuzziness.

For a 3×3 neighbourhood, LBP codes can be obtained from the following equation,

$$LBP = \sum_{i=0}^{7} d_i \cdot 2^i \tag{11}$$

and the corresponding contribution C_{LBP} of each LBP code in the FLBP histogram can be defined as:

$$C_{LBP} = \prod_{i=0}^{7} m_{di}(g_i)$$

where $d_i \in \{0, 1\}$, and thus $m_{di}()$ can be either $m_0()$ or m_1.

Thus each neighbourhood is characterized by a set of ordered pairs of LBP codes and contribution values (Fig. 4d). In other words, each 3×3 neighbourhood contributes to more than one bins of the FLBP histogram. This histogram forms a feature vector, representing the underlying texture.

2) Intensity Features

In US images a substantial amount of information concerning the pathology of the examined tissue is contained in image echogenity [33]. Several studies on US medical images have been using echogenity features based on grey-level histograms (GLH) [10, 13]. In the proposed system, fuzzy grey-level histograms (FGLH) have been utilized for intensity representation, given that they are well known for their insensitivity to noise [34].

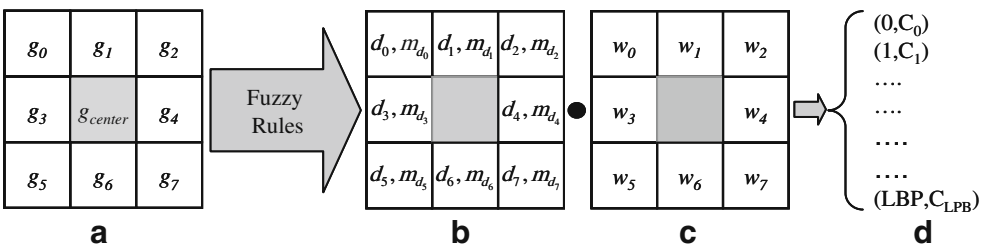

Fig. 4 FLBP computation scheme for a 3×3 neighbourhood. **a** 3×3 pixels neighbourhood. **b** Fuzzy thresholded values along with corresponding membership values. **c** Binomial weights. **d** LBP codes and corresponding contribution values

The normalized fuzzy histogram of an image region, with dimensions $N \times M$ pixels, can be defined as:

$$\mathrm{H}(g) = \frac{1}{N \cdot M} \sum_{i=1}^{N \cdot M} m_{\mathrm{g}}(g_i) \tag{13}$$

where m_g is a membership function that defines the degree of membership of pixel i with grey value g_i to histogram bin g. A commonly used triangular function is the following:

$$m_g(g_i) = \begin{cases} \dfrac{T - |g_i - g|}{T^2} & |g_i - g| < T \\ 0 & otherwise \end{cases} \tag{14}$$

where $T \in [0, \mathrm{G})$ determines the degree of fuzziness.

Classification

The fusion of FLBP and FGLH feature vectors extracted from sub-images sampled from the parenchyma of the thyroid gland, are subsequently classified into a predefined set of classes. The output of the classification phase is a class label representing either normal or nodular tissue. For the classification phase of the proposed scheme two widely used approaches have been evaluated, the Support Vectors Machines (SVM) and the *k*- Nearest Neighbours (*k*-NN). Both methodologies have advantages and disadvantages which are briefly described in the following paragraphs.

1) Support Vector Machines

The Support Vector Machine (SVM) [35] is a widely accepted classifier, considered very effective for pattern recognition, machine learning and data mining. It is based on the structural risk minimization principle described by Vapnik [36]. An optimal hyperplane decides the separation between individual classes of patterns. The optimality is based on the logic that the average distance between the hyperplane and the closest training points on both sides should be maximal. This aids in avoiding overfitting training data and maximizing the classification performance without being affected by the magnitude of the features-to-samples ratio (a phenomenon known as "curse of dimensionality") [37].

2) *k*-Nearest Neighbourhood

A *k*-Nearest Neighbour (*k*-NN) classifier [38], is also integrated in the system proposed, offering a good alternative when simplicity andease of the training phase are the predominant issues. The *k*-NN method is non parametric and generally effective classification approach. In addition it allows easy and fast incorporation of new data into an existing trained system. However, the price to pay with the *k*-NN is that for large datasets and multidimensional feature spaces, it requires large memory size and heavy computational load.

Post-processing

Misclassified pixels or small groups of misclassified pixels may appear as spots or small oblong formations within the output binary images. In order to reduce this effect, the post-processing component introduces the majority voting decision criterion proposed in [39]. This criterion takes advantage of the overlap between the consecutive sub-images sampled by the raster scanning approach and considers that a pixel can belong to more than one of the sampled sub-images. Therefore, a class label for each pixel can be aggregated by the majority of the class labels of the sub-images that this pixel belongs to, and the sub-images its neighboring pixels belong to, within a pixel radius R.

Performance evaluation

The proposed scheme presented in the previous section has been implemented as a prototype exploratory analysis system, named TND (Thyroid Nodule Detector). The functionality of TND is considered in two phases, a training phase and an application phase. In the first training phase, TND can be trained with a dataset annotated by experts, where different types of tissue have been identified. Once the TND system has been trained, it can be used to evaluate new US images or videos of the thyroid gland. Indicatively, two of the main windows of the graphical user interface (GUI) of the TND system concerning the application phase are illustrated in Fig. 5.

In the evaluation process a set of real US thyroid images have been used. These images have been provided by the Euromedica Medical Center of Athens in Greece with the approval of its ethics committee. Examinations were performed using a General Electric VOLUSON 730 sonographic imaging system [40]. A broadband probe with a frequency range of 6.0 to 12.0 MHz was used. The settings of the US scanner that affect image attributes (dynamic range 60 dB, grey-map linear, frame rate high, persistence low) were kept constant throughout the entire study. Ultrasound images were digitized at 8-bit grey-levels and their effective resolution was 480×440 pixels.

On the whole, 64 patients who had ultrasonographic examinations of their thyroid gland were enrolled in this study. A total of 118 longitudinal thyroid US images with nodules classified as Grade 3 or Grade 4 [41] were acquired accompanied with ground truth information. This experimental evaluation focuses on the detection of these

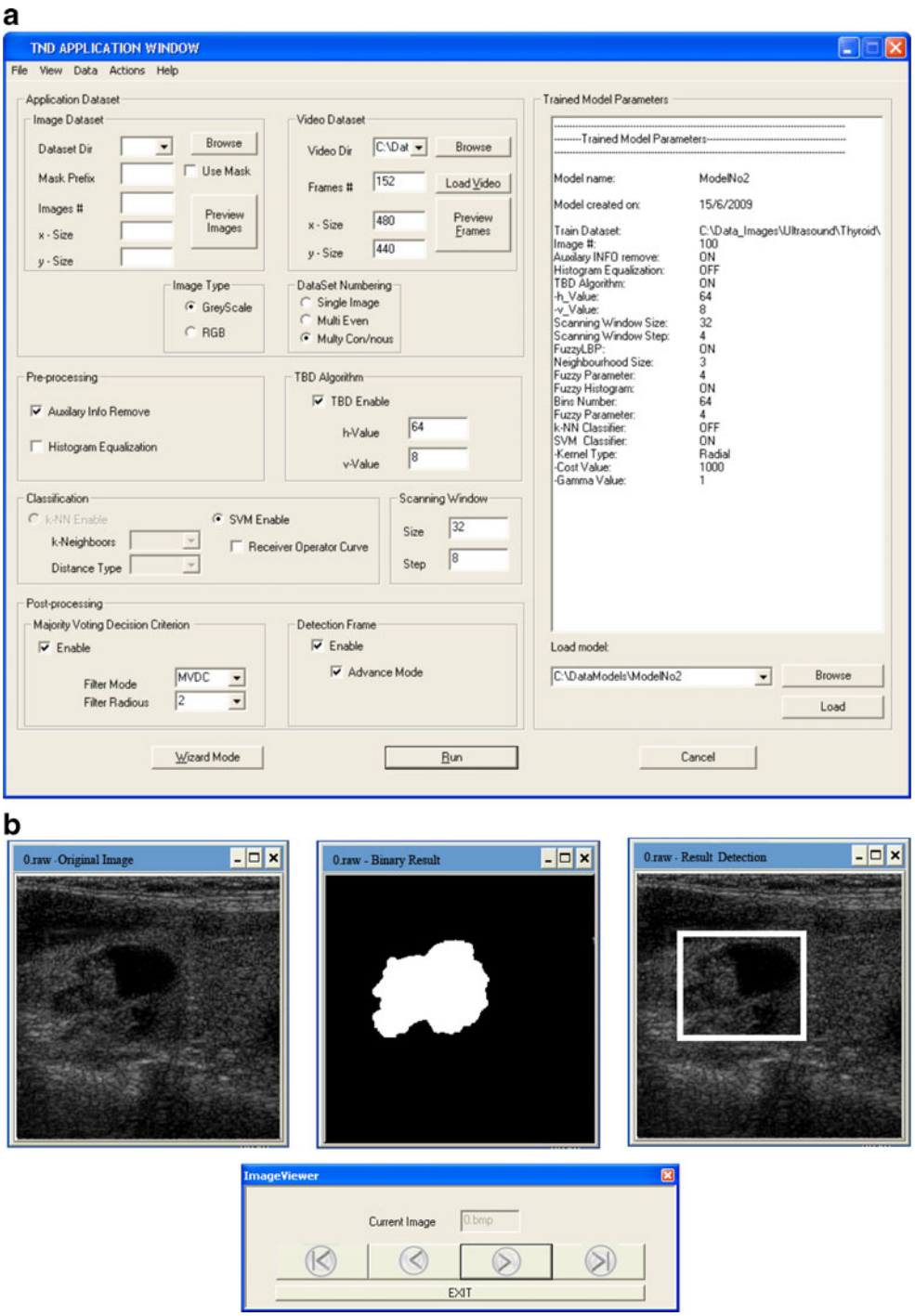

Fig. 5 **a** Example of TND application phase window. **b** Example of TND input/output windows. From left to right: original input image, the binary output image, and the greyscale output image with the nodule framed by a rectangular box

types of nodules because they are associated with a significantly higher malignancy risk compared to other types of nodules [41].

Three expert radiologists marked up the boundaries of the thyroid lobes and of the existing nodules. The ground truth concerning the boundaries of the lobes was obtained by following the rule that a pixel belongs to the thyroid gland when it is included in at least two out of the three delineations drawn by the experts [42]. The same rule has been applied for the ground truth concerning the nodules. Classification percentage actually quantifies the matching between classified pixels of the US images and the reference ground truth images.

The experimental evaluation that follows aims to present the effectiveness of the proposed TBD-2 algorithm and the nodule detection performance obtained by the proposed system.

Table 2 Optimal system configurations and corresponding classification accuracies

Setup no	ROI		Feature extraction					Classification				Post-processing	NDA(%)	APP(%)
	TBD-2		SW		FGLH		FLBP					MVDC		
	h	*v*	WD	WS	Bins	*T*	*T*	Type	Kernel	*g*	*c*	*R*		
1	-	-	32	8	32	13	13	SVM	RBF	1	2^3	3	90.1	82.2
2	32	16	32	8	32	13	13	SVM	Sigmoid	1	2^2	2	95.2	91.3

TBD-2 Thyroid Boundaries Detection—2, *h* Stripe Height, *v* Horizontal Refinement Parameter, *SW* Scanning Window, *WD* Window Dimension, *WS* Window Step, *FGLH* :Fuzzy Grey Level Histograms, *Bins* Quantized Histograms Bins, FLBP *Fuzzy Local Binary Patterns, T* Fuzzification Parameter, *SVM* Support Vector Machines, *RBF* Radial Basis Function, *g* Gamma Parameter for SVM, *c* Cost Parameter for SVM, *MVDC* Majority Voting Decision Criterion, *R* Radius Parameter for MVDC, *NDA* Nodule Detection Accuracy, *APP* Accuracy Per Pixel

Evaluation of the TBD-2 algorithm

Exhaustive experiments have been performed with the TBD-2 algorithm to determine the optimal set of parameters that minimizes the error in detecting the boundaries of the thyroid lobes. There are two parameters that the user can tune in the TBD-2 algorithm; *h* that controls the vertical resolution, and *v* that controls the horizontal resolution. The values tested for *h* varied from 8 to 128, and for *v* from 2 to 64, where $v<2h$. For large *h* and small *v* the proposed method leads to gross detection of the boundaries, resulting in a diminution of the methods accuracy, whereas higher accuracy can be obtained for smaller values of *h* and larger *v*. The optimal values of the investigated parameters are $h=32$, and $v=16$, for which the mean accuracy in boundaries detection reached a maximum of 91.6±3.2%.

For comparison reason the TBD algorithm proposed in [22] has also been tested for the detection of the boundaries of the thyroid lobes on the same image dataset. In this case, after exhaustive experimentation for the determination of the optimal set of parameters, the maximum mean accuracy in boundaries detection was 85.5±2.1%.

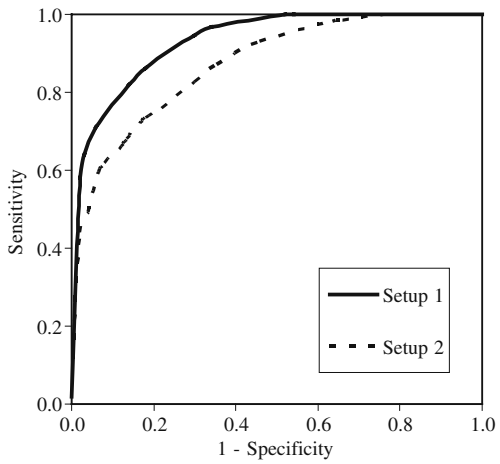

Fig. 6 ROC curves obtained by the two optimal system configurations presented in Table 2

Evaluation of the thyroid nodule detection scheme

The TND system has been applied on real longitudinal US thyroid images and videos for the experimental evaluation of the proposed nodule detection scheme. For this purpose a leave-one-out cross validation scheme has been applied [43] that uses the training data very efficiently and presents relatively small bias especially for small datasets [44]. Multiple training and testing sessions were carried out, where different values for system parameters were applied. An exhaustive search of combinations of parameters values have been based on the minimum classification error criterion. The system has been evaluated in terms of pixelwise classification accuracy and nodule detection accuracy. The nodule detection accuracy is the percentage of all existing nodules that have been correctly detected and framed by a rectangular box.

Initially the performance of the proposed scheme is evaluated without the automatic ROI definition, provided by the TBD-2 algorithm, neither for the training nor for the application phases. In this case the combination of FLBP and FGLH features has been extracted from the whole US image, including the skin, fat, and tissue surrounding the gland. The best classification accuracy has been obtained with the SVM classifier and the set of parameters values are apposed in the first row of Table 2. This setup resulted in 90.1% accuracy in nodules detection, and 82.2% in pixelwise classification between nodular and normal tissue. The receiver operating characteristics (ROC) curve for the

Table 3 Confusion matrix for the best system setup

	Predicted normal	Predicted nodular
Normal	87.8%	12.2%
Nodular	5.4%	94.6%

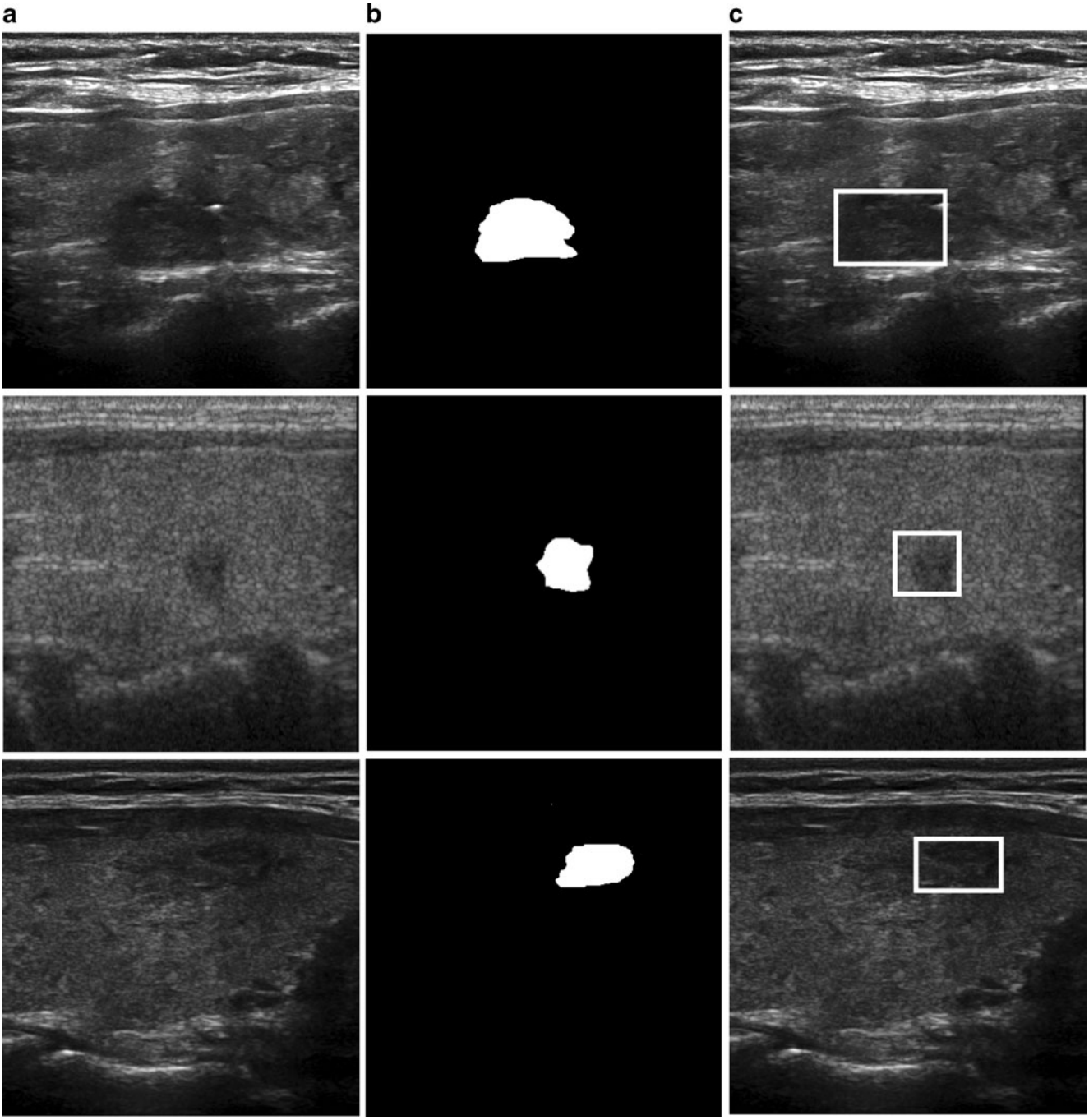

Fig. 7 Indicative result images **a** Input US images with thyroid nodule. **b** output images where tissue classified as nodular is represented by white pixels and normal tissue as black, **c** output images with the nodule framed by a rectangular box

pixelwise classification with this set up is depicted in Fig. 6 (Setup 1).

The experiments have been repeated with the feature sets proposed by Smutec et al. [10] and Tsantis et al. [13]. In the former case the best nodule detection accuracy obtained with the SVM classifier was 87.6%. The respective pixelwise classification accuracy reached 77.8%. The feature set proposed by Tsantis et al. [13] resulted in a nodule detection accuracy of 83.3% and a pixelwise classification accuracy of 74.2%. This comparison makes evident that the proposed combination of FLBP and FGLH features is more suitable for the discrimination of nodular from normal thyroid tissues.

In the second set of experiments the TBD-2 algorithm determined the boundaries of the lobes of the thyroid gland. Then, the feature extraction process has been applied only to the regions that belong to the thyroid gland, excluding any other tissue surrounding the gland. For the TBD-2 algorithm, the optimal parameters values obtained in the section of the

evaluation of the TBD-2 algorithm have been used. In that case the best overall detection performance has been accomplished again, with the SVM classifier and the second set of parameters values presented in Table 2. Through this set of parameters 95.2% of the existing nodules have been detected, and 91.3% of pixels representing thyroid tissue have been correctly classified as nodular or normal. The confusion matrix concerning the pixelwise classification accuracy is presented in Table 3. The ROC curve of the proposed methodology is depicted in Fig. 6 (Setup 2), where a substantial improvement can be clearly seen due to the utilization of the TBD-2 algorithm. Additionally, indicative output images are illustrated in Fig. 7.

Discussion and conclusions

In this study an original scheme has been presented for the detection of nodular tissue in US thyroid images and videos. The proposed scheme involves four components in which novel contributions have been considered. These include:

- the TBD-2 algorithm for automatic definition of the ROI including only the thyroid parenchyma, and
- the combination of Fuzzy Local Binary Patterns (FLBP) and Fuzzy Grey Level Histogram (FGLH) features

A prototype CAD system, named TND, implementing the proposed methodology has been developed and applied for exploratory analysis on real US images and videos. Through a simple and practical GUI, TND offers to the user the potential of tuning a variety of parameters relevant to the algorithms incorporated.

The feasibility of the proposed methodology has been investigated through an extensive experimental application of TND system on real US data. Experimental results showed that the proposed FLBP-FGLH fuzzy feature combination is more effective than the most relevant ones proposed in the literature [10, 13]. Overall the experimental evaluation showed that TND system can be clinically applied to provide physicians with a second opinion on the problem of nodule detection.

Future prospects for the enhancement of TND system include incorporation of sophisticated automatic methods for selection of parameters values, such as genetic algorithms, and integration of a suitable contour approach such as the one proposed in [14] for delineation of the detected nodules on the thyroid US images.

Availability of the software

A demonstration version of the presented software is available for downloading from our web site: http//rtsimage.di.uoa.gr/TND.htm

Acknowledgment We would like to thank EUROMEDICA S.A., Greece for the provision of the medical images. We would also like to thank N. Dimitropoulos, M.D. and G. Legakis, M.D. for their continuous support and advice. This work was supported by the Greek General Secretariat of Research and Technology (25%), the European Social Fund (75%), and the private sector, under the framework of Measure 8.3 of E.P. Antagonistikotita—3rd European Support Framework—PENED 2003 (grant no. 03-ED-662).

References

1. Welker, M., Orlov, D., Thyroid Nodules. *American Family Physician*, 67, 2003.
2. Raeth, U., Schlaps, D., Limberg, B., Zuna, I., Lorenz, A., Kaick, G., Lorenz, W. J., and Kommerell, B., Diagnostic accuracy of computerized B-scan texture analysis and conventional ultrasonography in diffuse parenchymal and malignant liver disease. *J Clin Ultrasound* 13:87–99, 1985.
3. Abe, C., Kahn, C. E., Doi, K., and Katsuragawa, S., Computer-aided detection of diffuse liver disease in ultrasound images. *Investig Radiol* 27:71–77, 1992.
4. Kadah, Y. M., Farag, A. A., Zurada, J. M., Badawi, A. M., and Youssef, A. M., Classification algorithms for quantitative tissue characterization of diffuse liver disease from ultrasound images. *IEEE Trans Med Imaging* 15:466–478, 1996.
5. Horsch, K., Giger, M. L., Vyborny, C. J., and Venta, L. A., Performance of computer-aided diagnosis in the interpretation of lesions on breast sonography. *Acad Radiol* 11:272–280, 2004.
6. Joo, S., Yang, Y. S., Moon, W. K., and Kim, H. C., Computer-aided diagnosis of solid breast nodules: use of an artificial neural network based on multiple sonographic features. *IEEE Trans Med Imaging* 23:1292–1300, 2004.
7. Kuo, W. J., Chang, R. F., Moon, W. K., Lee, C. C., and Chen, D. R., Computer-aided diagnosis of breast tumors with different US systems. *Acad. Radiology* 9:793–799, 2002.
8. Huynen, A., Giesen, R., De La Rosette, J., Aarnink, R., Debruyne, F., and Wijkstra, H., Analysis of ultrasonographic prostate images for the detection of prostatic carcinoma: the automated urologic diagnostic expert system. *Ultrasound Med Biol* 20:1–10, 1994.
9. Rosette, J., Computerized analysis of transrectal ultrasonography images in the detection of prostate carcinoma. *Br J Urol* 75:485–491, 1995.
10. Smutek, D., Šara, R., Sucharda, P., Tjahjadi, T., and Švec, M., Image texture analysis of sonograms in chronic inflammations of thyroid gland. *Ultrasound Med Biol* 29:1531–1543, 2003.
11. Muzzolini, R., Yang, Y. H., and Pierson, R., Texture characterization using robust statistics. *Pattern Recognit* 27(1):119–134, 1994.
12. Haralick, R. M., Dinstein, I., and Shanmugamm, K., Textural features for image classification. *IEEE Trans. On Systems, Man and Cybernetics* 3(6):610–621, 1973.
13. Tsantis, S., Cavouras, D., Kalatzis, I., Piliouras, N., Dimitropoulos, N., and Nikiforidis, G., Development of a support vector machine-based image analysis system for assessing the thyroid nodule malignancy risk. *Ultrasound Med Biology* 31:1451–1459, 2005.
14. Maroulis, D. E., Savelonas, M., Karkanis, S. A., Iakovidis, D. K., Dimitropoulos, N., Computer-Aided Thyroid Nodule Detection in Ultrasound Images, *IEEE International Symposium on Computer-Based Medical Systems—CBMS*: 271–276, 2005.
15. Mailloux, G., Bertrand, M., Stampfler, R., and Ethier, S., Local histogram information content of ultrasound B-mode echographic texture. *Ultrasound Med Biol* 11:743–750, 1985.

16. Mailloux, G., Bertrand, M., Stampfler, R., and Ethier, S., Computer analysis of echographic textures in hashimoto disease of the thyroid. *J Clin Ultrasound* 14:521–527, 1986.
17. Morifuji, H., Analysis of ultrasound B-mode histogram in thyroid tumors. *Nippon Geka Gakkai Zasshi* 90(2):210–221, 1989.
18. Hirning, T., Zuna, I., and Schlaps, D., Quantification and classification of echographic findings the thyroid gland by computerized b-mode texture analysis. *Eur J Radiol* 9:244–247, 1989.
19. Julesz, B., Textons, the elements of texture perception, and their interactions. *Nature* 290:91, 1981.
20. Skouroliakou, C., Lyra, M., Antoniou, A., and Vlahos, L., Quantitative image analysis in sonograms of the thyroid gland. *Nucl Instrum Meth Phys* 569:606–609, 2006.
21. Savelonas, M. A., Iakovidis, D. K., Dimitropoulos, N., and Maroulis, D., Computational Characterization of Thyroid Tissue in the Radon Domain, *IEEE Internationa Symposium on Computer-Based Medical Systems* 189-192, 2007.
22. Keramidas, E. G., Iakovidis, D., Maroulis, D., and Karkanis, S. A., Efficient and effective ultrasound image analysis scheme for thyroid nodule detection. *Lect Notes Comput Sci* 4633:1052–1060, 2007.
23. Iakovidis, D. K., Keramidas, E., and Maroulis, D., Fuzzy local binary patterns for ultrasound texture characterization, image analysis and recognition. *International Conference (ICIAR 2008) Springer LNCS* 5112:750–759, 2008.
24. Wilhjelm, J., Gronholdt, M. L., Wiebe, B., Jespersen, S. K., Hansen, L. K., and Sillesen, H., Quantitative analysis of ultrasound B-mode images of carotid atherosclerotic plaque: correlation with visual classifcation and histological examination. *IEEE Trans Med Imaging 17:910–922, 1998.*
25. Rumack, C. M., Wilson, S. R., Charboneau, J. W., Johnson, J. A., *Diagnostic Ultrasound, Mosby*, ISBN 0323020232, 2004.
26. Maroulis, D. E., Savelonas, M. A., Iakovidis, D. K., Karkanis, S. A., and Dimitropoulos, N., Variable background active contour model for computer-aided delineation of nodules in thyroid ultrasound images. *IEEE Trans Inf Technol Biomed* 11(5):537–543, 2007.
27. Chen, D.-R., Chang, R.-F., Wu, W.-J., Moon, W. K., and Wu, W.-L., 3-D breast ultrasound segmentation using active contour model. *Ultrasound Med Biol* 29(7):1017–1026, 2003.
28. Hu, N., Downey, D. B., Fenster, A., and Ladak, H. M., Prostate boundary segmentation from 3d ultrasound images. *Med Phys* 30 (7):1648–1659, 2003.
29. Chiu, B., Freeman, G. H., Salama, M. M. A., and Fenster, A., Prostate segmentation algorithm using dyadic wavelet transform and discrete dynamic contour. *Phys Med Biol* 49(21):4943–4960, 2004.
30. Ojala, T., Pietikainen, M., Harwood, D., A comparative study of texture measures with classification based on featured distribution. *Pattern Recognition* 29, 1996.
31. Petrou, M., and Sevilla, P. G., Image Processing: Dealing With Texture, *John Wiley and Sons Ltd*, 2006.
32. Jasjit, S., Wilson, D., and Laxminarayan, S., (Eds.) Handbook of Biomedical Image Analysis, 2005, ISBN: 978-0-306-48550-3.
33. Simeone, F. J., Daniel, G. H., and MuLler, P. R., High-resolution real-time sonography. *Radiology* 155:431–439, 1985.
34. Jawahar, C. V., and Ray, A. K., Fuzzy statistics of digital images. *IEEE Signal Process Lett* 3:225–227, 1996.
35. Chapelle, O., Haffner, P., and Vapnik, V. N., Support vector machines for histogram-based image classification. *IEEE Trans Neural Netw. IEEE Trans* 10:1055–1064, 1999.
36. Vapnik, V. N., *The nature of statistical learning theory.* Springer-Verlag, New York, 1995.
37. Burges, C., A tutorial on support vector machines for pattern recognition, *Kluwer Academic Publishers*, 1998.
38. Theodoridis, S., and Koutroumbas, K., *Pattern Recognition, Academic Press*, 2008.
39. Maroulis, D., Iakovidis, D., Karkanis, S., and Karras, D., COLD: a versatile system for detection of colorectal lesions in endoscopic images. *Comput Meth Programs Biomed* 70:151–166, 2003.
40. General Electric Healthcare. Ultrasound Imaging System, Voluson 730 Pro. Retrieved August 6 2010, http://www.gehealthcare.com/usen/ultrasound/genimg/products/voluson730/index.html.
41. Tomimori, E. K., Camaro, C. Y., Bisi, H., and Medeiros-Neto, G., Combined ultrasongraphic and cytological studies in the diagnosis of thyroid nodules. *Biochimie* 81:447–452, 1999.
42. Kaus, M. R., Warfield, S. K., Jolesz, F. A., and Kikinis, R., Segmentation of Meningiomas and Low Grade Gliomas in MRI, *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 1-10, 1999.
43. Samarasinghe, S., Neural Networks for Applied Sciences and Engineering, *Auerbach Publications Boston, USA* ISBN:084933375, 2006.
44. Leary, R. H., Rosen, J. B., and Jambeckz, P., An optimal structure-discriminative amino acid index for protein fold recognition. *Biophys J* 86:411–419, 2004.