Thyroid Segmentation and Volume Estimation in Ultrasound Images

Chuan-Yu Chang*, *Senior Member, IEEE*, Yue-Fong Lei, Chin-Hsiao Tseng, and Shyang-Rong Shih

Abstract—Physicians usually diagnose the pathology of the thyroid gland by its volume. However, even if the thyroid glands are found and the shapes are hand-marked from ultrasound (US) images, most physicians still depend on computed tomography (CT) images, which are expensive to obtain, for precise measurements of the volume of the thyroid gland. This approach relies heavily on the experience of the physicians and is very time consuming. Patients are exposed to high radiation when obtaining CT images. In contrast, US imaging does not require ionizing radiation and is relatively inexpensive. US imaging is thus one of the most commonly used auxiliary tools in clinical diagnosis. The present study proposes a complete solution to estimate the volume of the thyroid gland directly from US images. The radial basis function neural network is used to classify blocks of the thyroid gland. The integral region is acquired by applying a specific-region-growing method to potential points of interest. The parameters for evaluating the thyroid volume are estimated using a particle swarm optimization algorithm. Experimental results of the thyroid region segmentation and volume estimation in US images show that the proposed approach is very promising.

Index Terms—Particle swarm optimization (PSO), radial basis function (RBF) neural networks, region growing, thyroid segmentation.

I. INTRODUCTION

THE THYROID gland is a butterfly shaped organ and is composed of two cone-like lobes. Thyroid gland belongs to the endocrine system, and is located in the neck just in front of the larynx. It controls the secretion of the thyroid hormone, which regulates the temperature of the human body, and greatly affects childhood intelligence, growth, and adult metabolism. Too much or too little thyroid hormone (due to a thyroid that is too large or two small, respectively) causes pathological changes. Therefore, physicians often diagnose abnormal symptoms of the thyroid gland by its volume.

Ultrasound (US) imaging is currently the most popular diagnostic tool. It is inexpensive and easy to use; it can follow anatomical deformations in real time during biopsy and
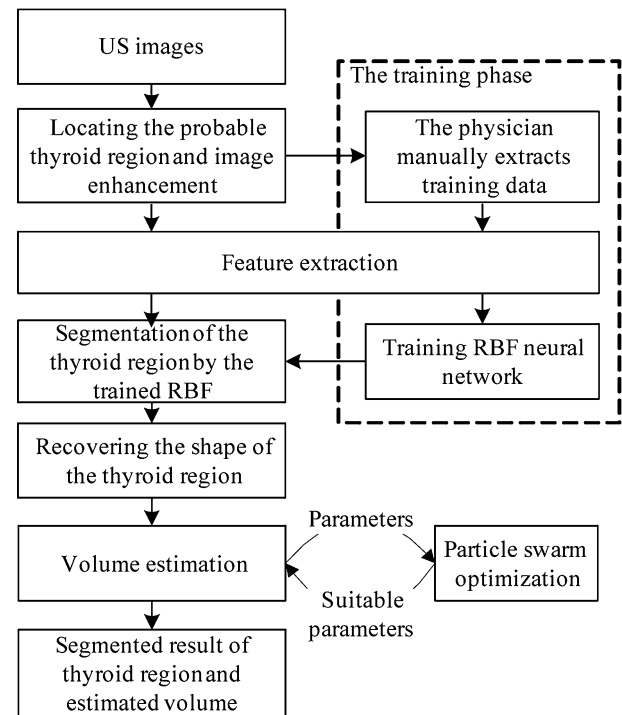
Fig. 1. Schematic steps of thyroid segmentation and volume estimation in US images.

treatment; and it is noninvasive and does not require ionizing radiation. However, US images contain echo perturbations and speckle noise, which can make diagnosis difficult.

Techniques to process US images are continuously being developed. Several methods for segmenting anatomical objects from US images have been presented, such as those for segmenting the prostate [2], [3], tumors in the breast [4], the carotid artery [5], [25], and the thyroid nodule [6], [23]. Among these segmentation methods, active contour models (ACMs) [7] have attracted attention due to their high performance. However, most active contour methods are sensitive to the gradient of the edge, and physicians are required to outline the rough contour of the thyroid gland. This is a time-consuming procedure, and an inaccurately outlined contour seriously affects the segmentation results.

The present study proposes a complete solution that uses a radial basis function (RBF) neural network [9] to automatically segment the thyroid gland. The particle swarm optimization (PSO) algorithm [10, 11] is then used to estimate the thyroid volume from US images. Fig. 1 shows a schematic diagram of the proposed method.
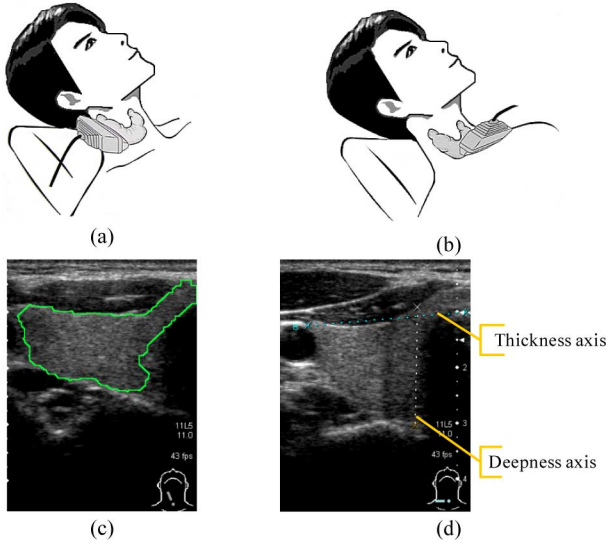
Fig. 2. Acquisition procedure. (a) Acquisition of longitudinal plane of right thyroid lobe. (b) Acquisition of transversal plane of right thyroid lobe. (c) Segmented result in the right longitudinal plane by our proposed method. (d) Thickness and deepness axes in the right transverse plane.

Fig. 3. (a) Original US image. (b) Horizontal projection of the US image. (c) Result of locating a probable thyroid region.

In the training phase, the physicians must manually outline the rectangular regions of interest (ROI) from the thyroid gland and nonthyroid tissues. Six textural features extracted from the ROIs are used to train the RBF neural network. The trained RBF neural network can then automatically roughly classify the thyroid regions from the US images. A specific-region-growing method is then applied to retrieve the complete thyroid region. Finally, based on the area of the segmented thyroid, the thickness, and the depth of thyroid gland, the volume is estimated using a PSO algorithm. Fig. 2 shows the acquisition procedure and the two views of the thyroid gland of a patient. The US probe is placed transversally and longitudinally to the left and right thyroid lobe, respectively. Fig. 2(a) and (b) shows the acquisition of longitudinal and transversal plane of right thyroid lobe, respectively. Fig. 2(c) is the segmentation results in the right longitudinal plane obtained using the proposed method, and Fig. 2(d) shows the thickness and deepness axes in the right transverse plane.

The rest of this paper is organized as follows. In Section II, the details of the adaptive weighted median filter (AWMF) [13], RBF neural network [9], reconstruction stages, and PSO algorithm [10], [11] are described. Section III presents the experiment results. Conclusions are given in Section IV.

II. THYROID SEGMENTATION AND VOLUME-ESTIMATION APPROACH

In order to directly estimate the thyroid volume from US images, thyroid segmentation must be accurate. Keramids *et al.* extracted the local binary pattern (LBP) features from the ROI of the thyroid and applied a k-nearest neighbor (k-NN) algorithm to segment the thyroid gland. However, the segmented thyroid glands are messy [23] and their method cannot extract the complete thyroid gland. Therefore, in this paper, a com-
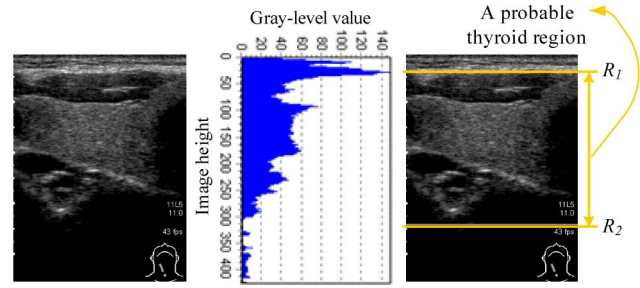
plete solution for segmenting thyroid glands is proposed. The methods for thyroid segmentation and volume estimation are presented in this section. There are five major steps, which are as follows: 1) locating the probable thyroid region and image enhancement; 2) feature extraction; 3) training RBF neural networks; 4) thyroid recovery; and 5) volume estimation. Details of these processes are described next.

A. Locating Probable Thyroid Region and Image Enhancement

In thyroid US images, low visual quality greatly affects the segmentation and the volume estimation results. A preprocessing step is thus required to enhance and locate the probable thyroid region. The preprocessing steps are as follows: 1) locating the probable thyroid region; 2) applying an AWMF [13] to reduce speckles; 3) applying two morphological operations [15] to enhance the filtering result; and 4) compensating for different US images according to the intensity template of the thyroid region.

1) Locating Probable Thyroid Region: In a thyroid US image, the thyroid gland is always in the middle, below the bright part and above the dark part of the image. Two reference values (R_1 and R_2) are defined to locate the probable thyroid region. R_1 is the row index with the largest average intensity in the horizontal projection of the US image. R_2 is the first row index with an average intensity of zero from the top to bottom in the horizontal projection of the US image. The probable thyroid region is located between the R_1th row and the R_2th row of the US thyroid image. An example of locating a probable thyroid region in an US thyroid image is shown in Fig. 3.

2) Adaptive Weighted Median Filter: The role of preprocessing is to remove the speckle noise and to reduce the influence of feature extraction. *AWMF* [13] is easy to use, has fewer free parameters, and preserves small details better than other nonlinear space-varying filters. Thus, an AWMF is applied to remove inevitable speckle noise and to enhance the probable thyroid region in the US images. AWMF is conducted on a fixed moving mask with the weights adjusted according to the local statistics. For a mask with a size of $M \times M$, the weight coefficient $w_{i,j}$ at position (i,j) is given by

$$w_{i,j} = \left[w_0 - \frac{gD\sigma_{x,y}^2}{\mu_{x,y}} \right] \tag{1}$$

where $\mu_{x,y}$ and $\sigma_{x,y}^2$ are the mean and variance of the $M \times M$ window centered in the (x,y) pixel, respectively. w_0 is the central weight, g is a scale factor, $[\cdot]$ is round-to-nearest function, and D is the Euclidean distance from the pixel point to the centre of the mask. If the weights are negative, they are set to zero. A 9×9 mask and parameters $w_0 = 10$ and $g = 0.25$ are used in this study [14]. The adaptive weighted median in an $M \times M$ region $\{I_{i,j}\}$ is defined as the pure median of the extended sequence formed by taking each term $I_{i,j}$, $w_{i,j}$ times, where $\{w_{i,j}\}$ are the corresponding weight coefficients. For example, if $w_{1,1} = 2$, $w_{1,2} = 3$, $w_{2,1} = 3$, and $w_{2,2} = 1$, the weighted median of the mask $\{I_{11}, I_{12}, I_{21}, I_{22}\}$ is given by $median\{I_{11}, I_{11}, I_{12}, I_{12}, I_{12}, I_{21}, I_{21}, I_{21}, I_{22}\} = I_{12}$.

3) Morphological Operation: A set of 3×3 closing and opening operators [15] were applied to remove the redundancy enhanced by AWMF.

4) Gray-Level Compensation: The variation of the gray level of the thyroid region in the US image greatly affects the segmentation results [16]. A gray-level compensation technique is thus applied to adjust the intensity of the probable thyroid region. In thyroid US images, the intensities of skin/fat are larger than those of other tissue. In general, the skin area occupied 20% of a thyroid US image. Hence, the normal-reference gray level GL_N is defined as the half gray level of an 8-bit image ($\text{GL}_N = 128$ in this study). The average value of the top 20% pixels (in intensity) in the test image $T(x,y)$ is regarded as the test-reference gray level GL_T. Accordingly, the intensity of the test image is adjusted by the value between the normal-reference gray level and the test-reference gray level using

$$T'(x,y) = \begin{cases} 255, & \text{if } T(x,y) - \Delta\text{GL} > 255 \\ 0, & \text{if } T(x,y) - \Delta\text{GL} < 0 \\ T(x,y) - \Delta\text{GL}, & \text{otherwise} \end{cases} \quad (2)$$

where $\Delta\text{GL} = \text{GL}_N - \text{GL}_T$, $x = 1, 2, \ldots, H$, and $y = 1, 2, \ldots, W$, where H and W denote the height and width of the probable thyroid region, respectively.

B. Feature Extraction

Textural features contain important information that can be used for the analysis and explanation of US images. In this paper, physicians manually extracted $2n$ ROIs with a size of $M \times M$ (n thyroid ROIs and n nonthyroid ROIs) from the probable thyroid region. Various feature-extraction methods have been implemented and analyzed. Six discriminative textural features were then extracted from the selected ROIs: the Haar wavelet [15], coefficient of local variation, histogram, block difference of inverse probabilities (BDIPs) [17], and normalized multiscale intensity difference (NMSID) [18]. These features are applied to the RBF neural network to classify the thyroid region, as described in Section II-C. These features are described in more detail as follows.

1) Haar Wavelet Features: The Haar wavelet features are significant features for segmentation in US images [19]. The mean and the variance of the low-low-frequency subband (LL

band) were computed as follows [15] :

$$\text{Mean of LL band}: \mu_{x,y} = \frac{1}{M^2} \sum_{(x,y)\in B} I(x,y) \quad (3)$$

$$\text{Variance of LL band}: \sigma_{x,y}^2 = \frac{1}{M^2} \sum_{(x,y)\in B} (I(x,y) - \mu_{x,y})^2 \quad (4)$$

where $I(x, y)$ denotes the intensity of a pixel (x,y) in the ROI block, which has passed through the Haar transformation, and B denotes a block size of $M \times M$.

2) Coefficient of Local Variation Feature: The coefficient of variation (CV) is a normalized measure of dispersion of a probability distribution. Because the texture of thyroid glands differs from those of other regions in the US image, CV is a useful index to represent it. The coefficient of local variation of a pixel located at (x, y) is defined as follows:

$$\text{LCV}_{x,y} = \frac{\sigma_{x,y}}{\mu_{x,y}} \quad (5)$$

where $\mu_{x,y}$ and $\sigma_{x,y}$ are the local mean and standard deviation of a pixel located at (x, y) with a block size of $M \times M$, respectively.

3) Histogram Feature: The histogram feature measures the texture characteristics of an $M \times M$ block. After the preprocessing, the thyroid gland occupies most of the area in the probable thyroid region. Thus, we extract the intensity of the largest area and add a tolerance of ± 10. The value of the histogram feature is defined as follows:

$$\text{HF} = \sum_{i=H-10, i\neq H}^{H+10} \text{histo}(i)$$

$$H = \arg\max_i (\text{histo}(i)) \quad (6)$$

where $\text{histo}(i)$ is the value of the histogram for an intensity equal to i of a block with a size of $M \times M$. ± 10 is a tolerance value determined from experiments.

4) BDIP Feature: The BDIPs [17] uses local probabilities in image blocks to measure local brightness variations of an image. BDIP is defined as the difference between the number of pixels in a block and the ratio of the sum of pixel intensities in the block to the maximum in the block, i.e.,

$$\text{BDIP} = \frac{M^2 - \sum_{(x,y)\in B} I(x,y)}{\max_{(x,y)\in B} I(x,y)} \quad (7)$$

where $I(x, y)$ denotes the intensity of a pixel (x, y) and B denotes a block with a size of $M \times M$.

5) NMSID Feature: NMSID [18] is defined as the differences between the pixel pairs with horizontal, vertical, diagonal,
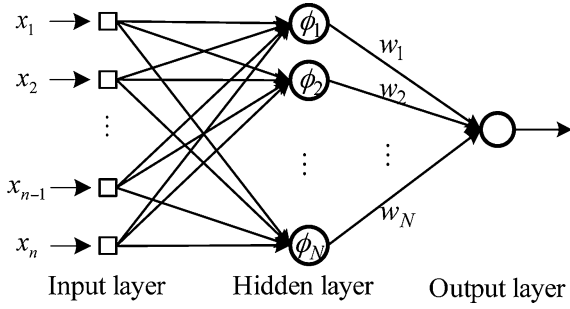
Fig. 4. Architecture of the RBF neural network.

and asymmetric-diagonal directions

$$
\text{NMSID} = \sum_{k=1}^{n}
$$

$$
\times \left[\begin{array}{l} \sum_{x=0}^{M-1} \sum_{y=0}^{M-k-1} |I(x,y) - I(x,y+k)| / M(M-k) \\ + \sum_{x=0}^{M-k-1} \sum_{y=0}^{M-1} |I(x,y) - I(x+k,y)| / M(M-k) \\ + \sum_{x=0}^{M-k-1} \sum_{y=0}^{M-k-1} |I(x,y) \\ \quad - I(x+k,y+k)| / (M-k)^2 \\ + \sum_{x=0}^{M-k-1} \sum_{y=0}^{M-k-1} |I(x,M-y) \\ \quad - I(x,M-(y+k))| / (M-k)^2 \end{array} \right] \Big/ 4
$$

(8)

where $I(x,y)$ denotes the intensity of a pixel (x,y) in a block with a size of $M \times M$ and n denotes the maximum horizontal or vertical distance.

C. RBF Neural Network

RBF neural networks [9] have attracted a lot of attention due to their good reliability in the field of image classification. A RBF neural network includes one hidden layer, which has high dimensionality. A nonlinear transformation is applied from the input layer to the hidden layer. The higher the dimension of the hidden space, the more accurate the approximation will be. The architecture of the proposed RBF neural network is presented in Fig. 4; it is applied to classify the thyroid gland region of the US image.

The input data of the RBF neural network can be written in vector form as follows:

$$
\mathbf{x}_i = [f_{i,1}, f_{i,2}, \ldots, f_{i,m}, \ldots, f_{i,6}]
$$

(9)

where $f_{i,m}$ is the mth feature of the ith block of size $M \times M$. All features are normalized before use by subtracting their mean value, and then, dividing the difference by their standard deviation. The mth normalized feature of the ith block is obtained by

$$
\bar{f}_{i,m} = \frac{f_{i,m} - \mu_m}{\sigma_m}
$$

(10)

where μ_m is the mean and σ_m is the standard deviation of feature m. These normalized feature vectors are then regarded as the training vectors of the RBF neural network.

In the proposed RBF neural network, which uses the stochastic gradient-based supervised learning algorithm [22], the error cost function is defined as follows:

$$
J(n) = \frac{1}{2} |e(n)|^2
$$

$$
= \frac{1}{2} \left[y_{\text{target}}(n) - \sum_{k=1}^{N} w_k(n) \phi_k (\|\mathbf{x}(n) - \mathbf{c}_k(n)\|) \right]^2
$$

$$
= \frac{1}{2} \left[y_{\text{target}}(n) - \sum_{k=1}^{N} w_k(n) \exp \left(-\frac{\|\mathbf{x}(n) - \mathbf{c}_k(n)\|_2^2}{\sigma_k^2(n)} \right) \right]^2
$$

(11)

where $\mathbf{x} \in R^{n \times 1}$ is an input feature vector, w_k is the weight vector between the output neuron and the kth hidden neuron, N is the number of neurons in the hidden layer, $\mathbf{c}_k \in R^{n \times 1}$ is the kth center node of the RBF, the symbol $\|\cdot\|$ denotes the Euclidean norm, and σ_k is the kth bandwidth of the Gaussian function defined as follows:

$$
\sigma_k = \frac{d_{\max}}{\sqrt{N}}
$$

(12)

where $d_{\max}$ is the maximum Euclidean distance between the selected center and N is the number of centers.

The updated equations for the network parameters are as follows:

$$
\mathbf{w}(n+1) = \mathbf{w}(n) - \mu_w \frac{\partial}{\partial \mathbf{w}} J(n) \Big|_{\mathbf{w}=\mathbf{w}(n)}
$$

$$
= \mathbf{w}(n) + \mu_w e(n) \Psi(n)
$$

(13)

$$
\mathbf{c}_k(n+1) = \mathbf{c}_k(n) - \mu_c \frac{\partial}{\partial \mathbf{c}_k} J(n) \Big|_{\mathbf{c}_k=\mathbf{c}_k(n)}
$$

$$
= \mathbf{c}_k(n) + \mu_c \frac{e(n) w_k(n)}{\sigma_k^2(n)}
$$

$$
\times \exp(-\|\mathbf{x}(n) - \mathbf{c}_k(n)\|^2 / \sigma_k^2(n)) [\mathbf{x}(n) - \mathbf{c}_k(n)]
$$

(14)

$$
\sigma_k(n+1) = \sigma_k(n) - \mu_\sigma \frac{\partial}{\partial \sigma_k} J(n) \Big|_{\sigma_i=\sigma_i(n)}
$$

$$
= \sigma_k(n) + \mu_\sigma \frac{e(n) w_k(n)}{\sigma_k^3(n)}
$$

$$
\exp(-\|\mathbf{x}(n) - \mathbf{c}_k(n)\|^2 / \sigma_k^2(n)) \|\mathbf{x}(n) - \mathbf{c}_k(n)\|^2
$$

(15)

where $\Psi(n) = [\phi_1(\mathbf{x}(n) - \mathbf{c}_1(n)), \ldots, \phi_N(\mathbf{x}(n) - \mathbf{c}_N(n))]^T$, $y_{\text{target}}(n) \in \{0,1\}$ is the corresponding desired output, and $\mu_w = 0.08$, $\mu_c = 0.05$, and $\mu_\sigma = 0.05$ are the appropriate learning rates. All the learning rates were determined empirically in this paper.

Similar to the training phase, the testing US image is split up into overlapping blocks of size $M \times M$. The overlap is 50%.
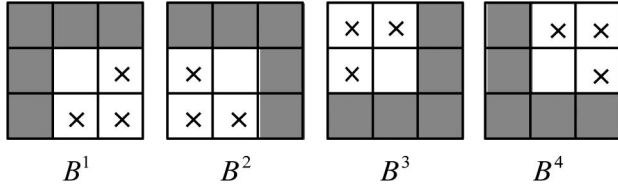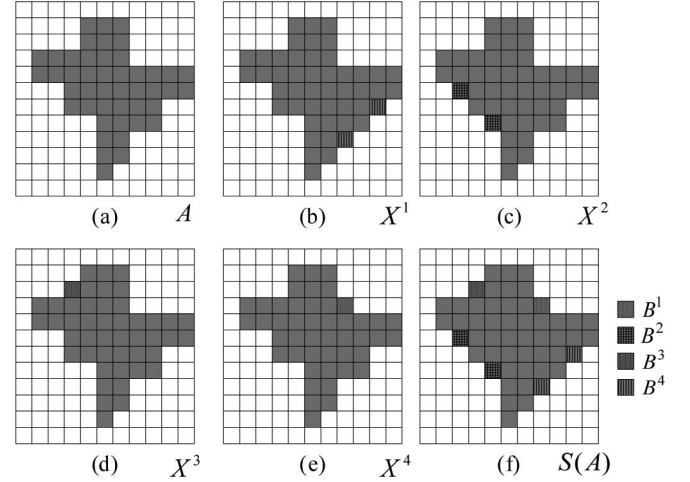
Fig. 5.   Structuring elements.

Fig. 6.   (a) Set A. (b)–(e) Results with the structuring elements. (f) Final result of filtering out the blocking shape of the segmented thyroid region showing the contribution of each structuring element.

The normalized feature vectors of the overlapping block are considered as input vectors in the trained RBF neural network. The trained RBF neural network classifies the block into the thyroid gland and the nonthyroid gland. For each thyroid block, the number of thyroid blocks was calculated in its 8-nearest neighbors. If the number is smaller than four, the block is reassigned to nonthyroid glands. Finally, the largest connected component [20] is extracted from the classified US image. The region of the largest connected component is considered as part of the thyroid gland region.

D. Recovering Shape of Thyroid Region

Using the aforementioned procedures, a pure region of the thyroid gland can be extracted (i.e., no pixels belong to the nonthyroid gland region). However, the shape of the segmented thyroid region is serrated, and thus, a refinement procedure is required to recover the complete thyroid gland region. Consequently, three specific reconstruction stages are applied to recover the complete shape of the thyroid gland.

The first reconstruction stage is filtering out the blocking shape of the segmented thyroid region. Let B^i, $i = 1, 2, 3, 4$ represent the four structuring elements, as shown in Fig. 5. Entries marked by "×" indicate the "do not care" condition.

The procedure consists of implementing the following equation:

$$X^i = (A * B^i) \qquad i = 1, 2, 3, 4 \qquad (16)$$

where A is the segmented thyroid region and $*$ denotes a match operator [15]. The match operator is a morphological operator designed to locate simple shapes within an image. A structuring element is said to have found a match in the thyroid gland region if the 3×3 region under the structuring element mask at that location matches the pattern of the mask. For a particular mask, a pattern match occurs when the center of the 3×3 region in the thyroid gland region is 0, and the five pixels under the shaded mask elements are 1. The result of filtering out the blocking shape of the segmented thyroid region A, denoted by $S(A)$, is thus

$$S(A) = \left(\overset{4}{\underset{i=1}{\cup}} X^i \right) \cup A. \qquad (17)$$

Fig. 6 shows the procedure given in (16) and (17). The origin of each structuring element is at its center.

The second reconstruction stage is based on the convex-hull concept [15]. Fig. 7 shows the four structuring elements of a convex hull. The general convex-hull algorithm does not take the local statistical information of the structuring element into account; the algorithm was thus improved here to consider this
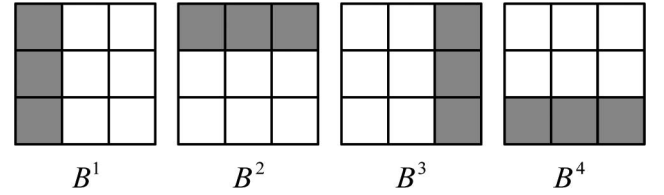

Fig. 7.   Structuring elements of the convex hull.

information. If a pattern match occurs, the criterion of convex-hull-based region growing is examined. The criterion is defined as follows:

$$|m_0 - m_1| < T \qquad (18)$$

where m_0 denotes the mean gray-level value of 3 pixels under the shaded mask elements, m_1 denotes the mean gray level value of 6 pixels under the bright mask elements, and T is a threshold. T is set to four from trial and errors experiments. If a 3×3 mask centered on pixel (x, y) conforms to this criterion, then pixel (x, y) is included in the thyroid gland region.

The thyroid gland usually contains cysts and blood vessels, making messy holes in the obtained image. A region filling operation and a 7×7 closing morphological operation are applied to fill these holes [15]. Using the aforementioned processes, the complete thyroid gland can be obtained, as shown in Fig. 2(c). The thickness axis of the segmented thyroid is acquired using the longest horizontal axis of the segmented thyroid. Let D_L and D_R denote the thickness axis of segmented thyroid glands in the left and right longitudinal plans of US images, respectively.

E. Volume Estimation

Since computed tomography (CT) imaging is expensive and involves hazardous radiation, US imaging is the most commonly used auxiliary tool currently utilized in clinical diagnosis. Hence, this study proposes a complete solution to estimate the volume of the thyroid gland directly from US images. PSO is a population-based stochastic optimization technique developed

by Eberhart and Kennedy [10], [11], inspired by the social behavior of bird flocking or fish schooling. The PSO algorithm is usually used to obtain a set of potential solutions that evolves to approach a convenient solution (or set of solutions) for a problem. The PSO algorithm has been reported to have strengths of fast convergence and robust stability over other evolutionary optimization mechanisms, such as genetic algorithms or ant colony algorithms [26]. Therefore, the PSO algorithm is here applied to estimate the parameters of the thyroid volume equation, which is defined by

$$\text{Volume}_{\text{US}} = a \times (\text{Area}_L \times D_L + \text{Area}_R \times D_R) + b \quad (19)$$

where a and b are scale and bias parameters that are estimated by the PSO algorithm, respectively. Area_L and Area_R represent the areas of the left and the right lobes of segmented thyroid regions in the left and right longitudinal planes of US images, respectively. With this optimization scheme, the scale and bias parameters for volume estimation can be directly estimated from the US images. The fitness function of the PSO algorithm is defined by

$$\text{Fitness} = |[a \times (\text{Area}_L \times D_L + \text{Area}_R \times D_R) + b]$$
$$- \text{Volume}_{\text{CT}}| \quad (20)$$

where $\text{Volume}_{\text{CT}}$ denotes the volume calculated from the CT image. To calculate the actual thyroid volume, the thyroid regions were manually outlined and verified by two experienced radiologists from the CT images. The areas of all slices and the interslice distance were then integrated to calculate the actual volume of thyroid gland.

The particle velocity and position can be mathematically modeled according to the following equations:

$$V_{id} = w \times V_{id} + c_1 \times \text{rand}() \times (p_{id} - x_{id})$$
$$+ c_2 \times \text{rand}() \times (p_{gd} - x_{id}) \quad (21)$$
$$x_{id} = x_{id} + V_{id} \quad (22)$$

where V_{id} denotes velocity of the dth dimension of the ith particle, w denotes the inertia weight, c_1 and c_2 are learning constants, $\text{rand}(\cdot)$ is a uniformly distributed random number between 0 and 1, x_{id} denotes the current position of the ith particle, p_{id} denotes the ith particle (p_{id} is used to keep track of the particle coordinates in the problem space which are associated with the best solution (fitness) it has achieved so far), and p_{gd} denotes the best solution (fitness) of the entire population. Fig. 8 shows the flow chart of the general PSO algorithm. There are five major steps. The particle positions and velocities are randomly initialized. The fitness function for each particle is calculated and the best solution for each particle is updated.

III. EXPERIMENTAL RESULTS

Four experiments were performed to show the capability of the proposed method. The images used for the experiments were taken from the Division of Endocrinology and Metabolism at National Taiwan University Hospital, Yunlin Branch. All US images were captured with a Toshiba Xario SSA-660A instrument with the following parameters: thyroid echo = 4–6 cm,
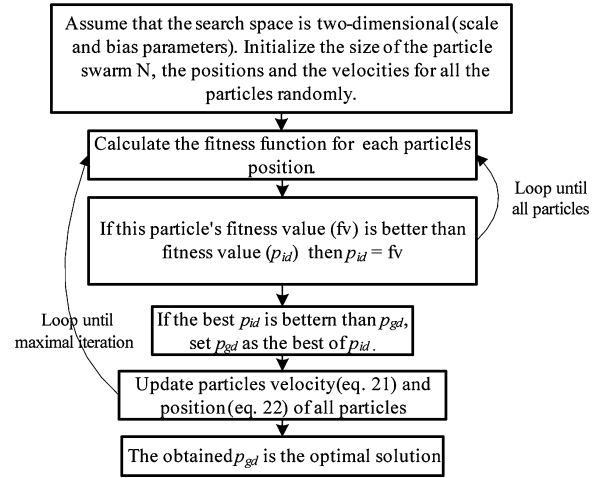

Fig. 8. Flowchart describing the general PSO algorithm.

linear transducer = 5–7.5 cm. CT images were captured with a General Electric LightSpeed-16 scanner with the following parameters: 120 kVp and 248–396 mA.

Since CT imaging is expensive and involves hazardous radiation, only five patients underwent both the US and CT inspections. Accordingly, the testing dataset contains a total of 20 US images and five CT image series from five patients; each CT series contained 80 slices. An additional 20 US images taken of five patients were used to train the RBFNN for segmenting thyroid regions from US images. In the 20 training US images, a total of 60 training patterns, including 30 thyroid tissues and 30 nonthyroid tissues extracted by an experienced physician, were used to train the RBF neural network. The size of the extracted ROI blocks was 16×16.

For the classification task, the RBF neural network was implemented with one output node and 90 hidden neurons. Two convergence conditions of the RBF neural network were defined as: 1) the maximum iteration was set to 10000 and 2) the correction value of synaptic weights was less than 0.00001. When one of the two conditions was satisfied, the training procedure was stopped. Fig. 9 shows the segmentation results obtained using the proposed method. Fig. 9(a) shows the original US image, which has serious speckle noise. The result of locating the probable thyroid region is shown in Fig. 9(b). Fig. 9(c) and (d) shows the results after image enhancement. Fig. 9(e) shows the classification result of the RBF neural network classifier. The segmented thyroid region after shape recovery is shown in Fig. 9(f). Fig. 9(g) shows the contour of the thyroid gland outlined manually by the physician. Fig. 9(g) and (f) are very similar.

In order to illustrate the segmentation performance of the proposed method, five standardized measurements were adopted, which are as follows: accuracy, sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV). The five measurements are defined as follows:

$$\text{Accuracy} = \frac{A_{\text{TP}} + A_{\text{TN}}}{A_P + A_N} \quad (23)$$

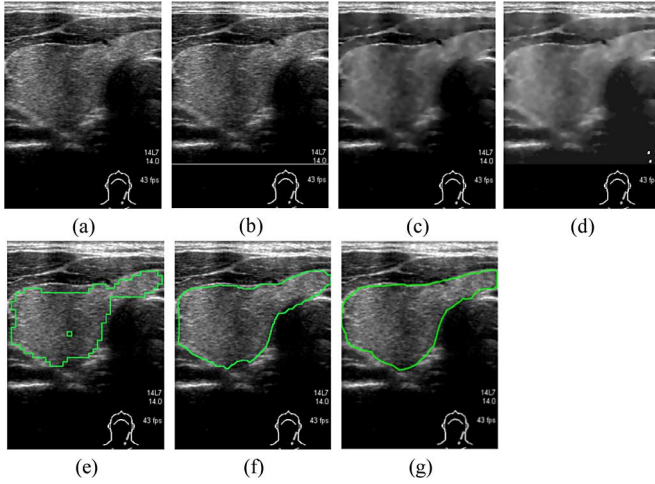$$\text{Sensitivity} = \frac{A_{\text{TP}}}{A_P} \quad (24)$$

Fig. 9.   (a) Original US image. (b) Result of locating the probable thyroid region. (c) Result of the AWMF. (d) Result of the morphological operation and gray-level compensation. (e) Result of the segmented thyroid region produced by the RBF neural network classifier. (f) Result after shape recovering. (g) Contour of the thyroid gland manually outlined by a physician.

$$\text{Specificity} = \frac{A_{\text{TN}}}{A_N} \qquad (25)$$

$$\text{PPV} = \frac{A_{\text{TP}}}{A_{\text{TP}} + A_{\text{FP}}} \qquad (26)$$

$$\text{NPV} = \frac{A_{\text{TN}}}{A_{\text{TN}} + A_{\text{FN}}} \qquad (27)$$

where A_P is the total number of actual positive pixels and A_N denotes the total number of actual negative pixels. A_{TP} denotes the number of pixels in the actual thyroid gland region segmented using the proposed method and A_{FP} denotes the number of pixels of nonthyroid regions, which are falsely segmented as pixels of thyroid gland region. Hence, the true negative pixels A_{TN} and false negative pixels A_{FN} can be defined by $A_{\text{TN}} = A_N - A_{\text{FP}}$ and $A_{\text{FN}} = A_P - A_{\text{TP}}$, respectively.

A. Gray-Level Compensation

Gray-level compensation is a procedure of image enhancement steps. The process is applied to adjust the intensity of the thyroid gland regions in different US images to decrease variation. Table I shows the performance of the proposed method with/without gray-level compensation. It shows that gray-level compensation significantly improves the segmentation results.

B. Threshold of Convex-Hull-Based Region Growing

Convex-hull-based region growing is a shape recovery procedure. To determine an appropriate threshold value, the performance of convex-hull-based region growing was evaluated for various threshold values. The results summarized in Table II. Threshold values of 4 and 8 achieved the highest performance. The performance is gradually decayed when T is larger than 8. Thus, the appropriate threshold range is between 4 and 8. $T = 4$ was used in the present study.

TABLE I
PERFORMANCE OF PROPOSED METHOD WITH AND WITHOUT
GRAY-LEVEL COMPENSATION

	With Gray-level compensation	Without Non-gray-level compensation
Accuracy	96.52	93.52
Sensitive	91.58	67.85
Specificity	97.61	96.39
PPV	89.14	87.18
NPV	98.07	94.29

TABLE II
PERFORMANCE OF DIFFERENT THRESHOLD VALUES

	T=4	T=8	T=10	T=12	T=14	T=16	T=18	T=20
Accuracy	96.52	96.52	96.12	96.01	95.46	95.06	94.83	93.34
Sensitive	91.58	91.58	93.82	94.07	94.92	95.26	95.42	95.55
Specificity	97.61	97.61	96.60	96.39	95.49	94.90	94.59	93.94
PPV	89.14	89.14	85.72	85.04	82.42	80.83	79.94	78.24
NPV	98.07	98.07	98.58	98.64	98.86	98.94	98.87	98.99
Average	94.58	94.58	94.17	94.03	93.43	93.00	92.75	92.21

TABLE III
SEGMENTATION PERFORMANCE OF PROPOSED METHOD

US images	Measurement Indices [%]				
	Accuracy	Sensitivity	Specificity	PPV	NPV
Case 1	96.15	94.46	96.79	89.99	98.07
Case 2	96.46	91.36	97.79	92.13	97.72
Case 3	96.98	91.85	98.09	90.62	98.23
Case 4	96.56	90.23	98.11	91.94	97.62
Case 5	96.47	89.97	97.30	81.01	98.69
Average	96.52	91.58	97.61	89.14	98.04

TABLE IV
AVERAGE PERFORMANCE OF FOUR CASES WITH DIFFERENT
SEGMENTATION METHODS

Segmentation methods	Measurement Indices [%]				
	Accuracy	Sensitivity	Specificity	PPV	NPV
The proposed method	96.54	91.98	97.69	91.17	97.91
AWMF+ACM	94.56	85.66	96.90	88.42	92.27
AWMF+Watershed	88.27	78.80	90.78	70.19	94.14

C. Segmentation

Table III shows the quantitative measurements of the proposed method. All the average values are higher than 89%, thus demonstrating the effectiveness of the proposed method.

The proposed method was compared with the AWMF + ACM [7] and AWMF + watershed model [8]. The ACM and watershed model were coded using the OpenCV library. In the evaluations, the optimum parameters for each method were obtained by trial and error. Table IV shows that the proposed

(a) (b) (c)

(d) (e) (f)
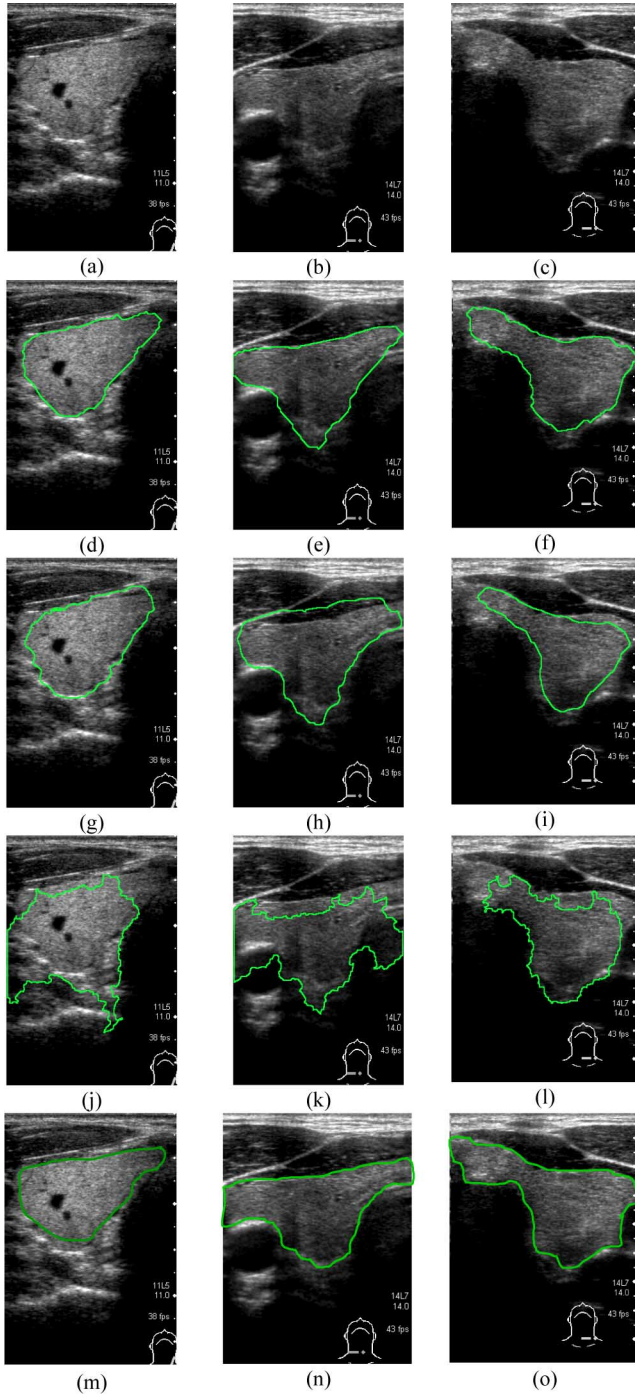
(g) (h) (i)

(j) (k) (l)

(m) (n) (o)

Fig. 10. (a)–(c) Original US images. (d)–(f) Segmentation results of Fig. 11(a)–(c) using the proposed method shown with the overlapped contour, respectively. (g)–(i) Segmentation results using AWMF + ACM. (j)–(l) Segmentation results using AWMF + Watershed. (m)–(o) Contour of the thyroid gland manually outlined by a physician.

Fig. 11. Segmentation power of thyroid, where *Az* is equal to 0.891.

shows the segmented results obtained using AWMF + watershed. Finally, Fig. 10(m)–(o) shows the contours of the thyroid glands manually outlined by a physician. The receiver–operating characteristic (ROC) curves of thyroid segmentation are shown in Fig. 11 [24]. The value of true-positive fraction almost reaches 1 with a tiny loss of the false-positive fraction and the value of *Az* equals 0.891. The results show that the proposed method correctly segments the thyroid gland region.

D. Volume Estimation

The detailed parameter settings for the PSO algorithm are as follows: the maximum number of iterations was 200, the population size was 12, the dimension of the search space was set to 2, $w = 1/(2 \times \ln(2))$, and c_1 and c_2 were both set to $0.5 + \ln(2)$ [12]. To evaluate the accuracy of the estimated volume, the difference (*Diff*) and the *MSE* between the estimated volume and the volume calculated from CT was applied. *Diff* and *MSE* are defined as follows:

$$\text{Diff} = \left| G_i - \hat{G}_i \right| \tag{28}$$

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^{N} (\text{Diff})^2 \tag{29}$$

where N is the case number, G_i is the volume estimation obtained using the proposed method, and $\hat{G}_i$ is the volume obtained from the CT images.

As mentioned earlier, only five CT series images were taken of five patients. To demonstrate the accuracy and efficiency of the proposed volume estimation method, two cases were selected randomly as training data and the others were used as testing data. The experiment was performed ten times. The performance of the proposed volume estimation method is shown in Table V. The MSE values are below 0.9. In Table V, the training data are in gray.

A low MSE indicates that the volume directly estimated from US images is similar to that estimated from the CT images. The proposed volume estimation method was compared with Brunn *et al.*'s method [1], Wael *et al.*'s method [21], and standard genetic algorithm [27], which are still used in clinical volume estimation in most hospitals. Table VI shows the performance of the volume estimation methods. The detailed parameter settings for the genetic algorithm are summarized as follows: maximum

method outperforms the other segmentation methods. Because of space limitations, only three of the 20 US image segmentation results that were obtained are shown in Fig. 10. Fig. 10(a)–(c) shows the original US images. The contours of the segmented thyroid glands from the proposed method are superposed onto the images in Fig. 10(d)–(f). Fig. 10(g)–(i) shows the segmentation results obtained using AWMF + ACM, while Fig. 10(j)–(l) Seg-
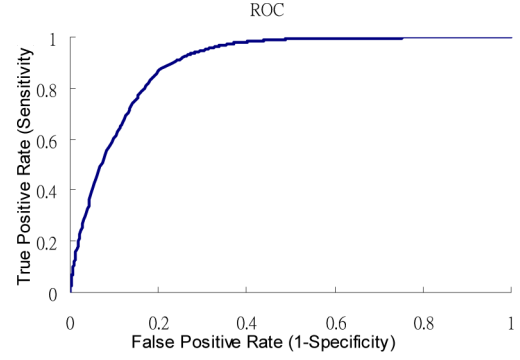
TABLE V
VOLUME ESTIMATION OF PROPOSED METHOD

Training numbers	Patients	Case1	Case2	Case3	Case4	Case5	MSE
	CT (ml)	22.172	17.987	22.246	18.969	13.190	
1 (a=0.8967,b=0.6962)	US image	22.578	17.577	21.502	20.268	12.609	0.582
	Diff	0.406	0.410	0.744	1.299	0.581	
2 (a=0.9729,b=-0.9499)	US image	22.792	17.366	21.624	20.285	11.976	0.873
	Diff	0.620	0.621	0.622	1.316	1.214	
3 (a=0.9162,b=-0.0205)	US image	22.338	17.228	21.238	19.976	12.152	0.743
	Diff	0.166	0.759	1.008	1.008	1.038	
4 (a=0.8165,b=1.6956)	US image	21.621	17.067	20.641	19.517	12.543	0.889
	Diff	0.551	0.920	1.605	0.548	0.647	
5 (a=0.901,b=0.3217)	US image	22.309	17.284	21.228	19.987	12.292	0.679
	Diff	0.137	0.703	1.018	1.018	0.898	
6 (a=0.9231,b=-0.1284)	US image	22.398	17.250	21.290	20.019	12.135	0.745
	Diff	0.226	0.737	0.956	1.051	1.055	
7 (a=0.8874,b=1.0853)	US image	22.741	17.791	21.676	20.454	12.875	0.598
	Diff	0.569	0.196	0.570	1.485	0.315	
8 (a=0.9083,b=0.4464)	US image	22.612	17.546	21.522	20.271	12.514	0.613
	Diff	0.440	0.441	0.724	1.302	0.676	
9 (a=0.862,b=2)	US image	23.035	18.228	22.001	20.814	13.452	0.868
	Diff	0.863	0.241	0.245	1.846	0.262	
10 (a=0.8143,b=1.7823)	US image	21.654	17.112	20.677	19.555	12.601	0.838
	Diff	0.518	0.875	1.569	0.587	0.589	

TABLE VI
PERFORMANCE WITH DIFFERENT VOLUME ESTIMATION METHODS

Patients	Case1	Case2	Case3	Case4	Case5	MSE
CT (ml)	22.172	17.987	22.246	18.969	13.190	
Brunn *et al.*	15.879	13.896	16.879	14.461	8.471	31.932
Diff	6.293	4.091	5.367	4.508	4.719	
Wael *et al.*	17.536	15.346	18.636	15.970	9.355	13.040
Diff	4.636	2.641	3.61	2.999	3.835	
Genetic algorithm (a=0.9251,b=0.3932)	22.968	17.809	21.858	20.584	12.684	0.737
Diff	0.797	0.178	0.388	1.616	0.506	
The proposed method (a=0.8967,b=0.6962)	22.578	17.577	21.502	20.268	12.609	0.582
Diff	0.406	0.410	0.744	1.299	0.581	

number of iterations: 500, population size: 200, crossover rate: 0.7, mutation rate: 0.0001, and two-point crossover: roulette wheel selection. The fitness function was the same as that for the PSO algorithm. The MSE value of the proposed method is 0.582, which indicates that it outperforms Brunn *et al.*'s, Wael *et al.*'s, and standard GA methods.

IV. CONCLUSION

US images are a widely used tool for clinical diagnosis, although it is time consuming for physicians to manually segment the thyroid gland region. The alternative to estimate the volume of a thyroid gland using CT imaging is expensive and involves hazardous radiation. Thus, a convenient system for thyroid segmentation and volume estimation in US images is of interest. The proposed method includes image enhancement processing to remove speckle noise, which greatly affects the segmentation results of the thyroid gland region obtained from US images.

The probable thyroid gland region is located in the US image, and then, an RBF neural network is used to classify the region into thyroid and nonthyroid gland areas. Finally, a region growing method is applied to recover an accurate shape of the thyroid gland region. The experiment results show that the proposed method can be used to segment the thyroid gland region and to estimate thyroid volume directly from US images. The proposed method offers two significant improvements: 1) it can automatically segment the thyroid gland region from US images and 2) it can accurately estimate the volume of the thyroid from US images.

REFERENCES

[1] J. Brunn, U. Block, G. Ruf, I. Bos, W. P. Kunze, and P. C. Scriba, "Volumetric analysis of thyroid lobes by real-time ultrasound," *Dtsch Med Wochenschr*, vol. 106, pp. 1338–1340, Oct. 1981.

[2] N. Hu, D. B. Downey, A. Fenster, and H. M. Ladak, "Prostate boundary segmentation from 3d ultrasound images," *Medical Physics*, vol. 30, no. 7, pp. 1648–1659, Jul. 2003.

[3] B. Chiu, G. H. Freeman, M. M. A. Salama, and A. Fenster, "Prostate segmentation algorithm using dyadic wavelet transform and discrete dynamic contour," *Phys. Med. Biol.*, vol. 49, no. 21, pp. 4943–4960, Nov. 2004.

[4] D. R. Chen, R. F. Chang, W. J. Wu, W. K. Moon, and W. L. Wu, "3-D breast ultrasound segmentation using active contour model," *Ultrasound Med. Biol.*, vol. 29, no. 7, pp. 1017–1026, Jul. 2003.

[5] C. Baillard C. Barillot, P. Bouthemy, "Robust adaptive segmentation of 3d medical images with level sets," Institut National de Recherche en Informatique et en Automatique (INRIA), Le Chesnay Cedex, France, Tech. Rep. 4071, Nov. 2000.

[6] D. E. Maroulis, M. A. Savelonas, D. K. Iakovidis, S. A. Karkanis, and N. Dimitropoulos, "Variable background active contour model for computer-aided delineation of nodules in thyroid ultrasound images," *IEEE Trans. Inf. Technol. Biomed.*, vol. 11, no. 5, pp. 537–543, 2007.

[7] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *Int. J. Comput. Vision*, vol. 1, no. 4, pp. 321–331, 1987.

[8] L. Vincent and P. Soille, "Watersheds in digital spaces: An efficient algorithm based on immersion simulation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 13, no. 6, pp. 583–598, 1991.

[9] F. M. Ham and I. Kostanic, *Principles of Neurocomputing for Science and Engineering*. New York: McGraw-Hill, 2001.

[10] R. C. Eberhart and J. Kennedy, "A new optimizer using particle swarm theory," in *Proc. 6th IEEE Int. Symp. Micro Mach. Human Science*, pp. 39–43, Oct. 1995.

[11] J. Kennedy and R. C. Eberhart, "Particle swarm optimization," in *Proc. IEEE Int. Conf. Neural Netw.*, vol. 4, pp. 1942–1948, 1995.

[12] M. Clerc, *Standard PSO version 2006*. http://www.particleswarm.info/Standard_PSO_2006.c, 2006.

[13] T. Loupas, W. N. McDicken, and P. L. Allan, "An adaptive weighted median filter for speckle suppression in medical ultrasonic images," *IEEE Trans. Circuits Syst.*, vol. 36, pp. 129–135, Jan. 1989.

[14] Y. Chen, R. Yin, P. Flynn, and S. Broschat, "Aggressive region growing for speckle reduction in ultrasound images," *Pattern Recognit. Lett.*, vol. 24, pp. 677–691, Feb. 2003.

[15] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 2nd ed. Englewood Cliffs, NJ: *Prentice-Hall International Edition*, 2002.

[16] R. L. Hsu, Mohamed Abdel-Mottaleb, and Anil K. Jain, "Face detection in color image," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 5, May 2002.

[17] Y. D. Chum and S. Y. Seo, "Image retrieval using BDIP and BVLC moments," *IEEE Trans. Circuits Syst. Video Technol*, vol. 13, no. 9, pp. 951–957, Sep. 2003.

[18] E. L. Chen, P. C. Chung, C. L. Chen, H. M. Tsai, and C. I. Chang, "An automatic diagnostic system for CT liver image classification," *IEEE Trans. Biomed. Eng.*, vol. 45, no. 6, pp. 783–794, 1998.

[19] C. M. Wu, Y. C. Chen, and K. S. Hsieh, "Texture features for classification of ultrasonic liver images," *IEEE Trans. Med. Imag.*, vol. 11, no. 2, pp. 141–152, Jun. 1992.

[20] L. D. Stefano and A. Bulgarelli, "A simple and efficient connected components labeling algorithm," in *Proc. 10th Int. Conf. Image Anal.*, pp. 322–327, 1999.

[21] W. Shabana, E. Peeters, and M. D. Maeseneer, "Measuring thyroid gland volume: Should we change the correction factor," *Amer. J. Roentgenology*, vol. 186, pp. 234–236, 2006.

[22] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *Nature*, vol. 323, pp. 533–536, 1986.

[23] E. G. Keramidas, D. K. Iakovidis, D. Maroulis, and S. Karkanis, "Efficient and effective ultrasound image analysis scheme for thyroid nodule detection," in *Lecture Notes in Computer Science*, vol. 4633. Heidelberg: Springer, pp. 1052–1060, 2007.

[24] S. Yu and L. Guan, "A CAD system for the automatic detection of clustered microcalcifications in digitized mammogram films," *IEEE Trans. Med. Imag.*, vol. 19, no. 2, pp. 115–126, Feb. 2000.

[25] J. C. R. Seabra, L. M. Pedro, J. F. e Fernandes, and J. M. Sanches, "A 3-D ultrasound-based framework to characterize the echo morphology of carotid plaques," *IEEE Trans. Biomed. Eng.*, vol. 56, pp. 1442–1453, 2009.

[26] A. Abraham, H. Guo, and H. Liu, "Swarm intelligence: Foundations, perspectives and applications," in *Swarm Intelligent Systems, Studies in Computational Intelligence*, vol. 26, pp. 3–25, 2006.

[27] D. B. Fogel, *Evolutionary Computation: Toward a New Philosophy of Machine Intelligence*, 3rd ed. Piscataway, NJ: IEEE Press, 2006.

Chuan-Yu Chang (M'02–SM'07) received the B.S. degree in nautical technology and M.S. degree in electrical engineering from National Taiwan Ocean University, Keelung, Taiwan, in 1993 and 1995, respectively, and the Ph.D. degree in electrical engineering from National Cheng Kung University, Tainan, Taiwan, in 2000.

From 2001 to 2002, he was in the Department of Computer Science and Information Engineering, Shu-Te University, Kaohsiung, Taiwan. From 2002 to 2006, he was in the Department of Electronic Engineering, National Yunlin University of Science and Technology, Yunlin, Taiwan. Since 2007, he has been with the Department of Computer Science and Information Engineering, National Yunlin University of Science and Technology, where he is currently an Associate Professor. He has authored or coauthored more than 100 publications in journals and conference proceedings. His research interests include neural networks and their application to medical image processing, wafer defect inspection, digital watermarking, and pattern recognition.

Dr. Chang is a Lifetime Member of Chinese Image Processing and Pattern Recognition Society and Taiwanese Association for Artificial Intelligence. He was the recipient of the Excellent Paper Award of the Image Processing and Pattern Recognition Society of Taiwan in 1999, 2001, and 2009, the Best Paper Award in International Computer Symposium in 1998, the Best Paper Award in the Conference on Artificial Intelligence and Applications in 2001, 2006, 2007, and 2008, and the Best Paper Award in National Computer Symposium in 2005. He has also been listed in *Who's Who in Science and Engineering*, *Who's Who in the World, Who's Who in Asia, and Who's Who of Emerging Leaders*.

Yue-Fong Lei received the B.S. degree from the Department of Information Management, Aletheia University, Taipei, Taiwan, in 2006, and the M.S. degree from the Department of Computer Science and Information Engineering, National Yunlin University of Science and Technology, Yunlin, Taiwan, in 2008.

He is currently an Engineer at Coretronic Corporation, Hsinchu, Taiwan. His research interests include neural networks and their applications to image processing.

Chin-Hsiao Tseng received the M.D. degree from the Medical School, College of Medicine, National Taiwan University, Taipei, Taiwan, in 1986, and the Ph.D. degree from the School of Public Health, National Taiwan University, in 1996.

From August 2006 to July 2008, he was the Director of the Division of Endocrinology and Metabolism and the Department of Medical Research and Development, National Taiwan University Hospital, Yunlin Branch. He was with the National Health Research Institute, Taiwan, the School of Public Health, Taipei Medical University, School of Public Health, National Taiwan University, National Yunlin University of Science and Technology, and the Armed Forces Institute of Pathology, Washington, DC. He is currently a Senior Attending Physician with the Department of Internal Medicine, National Taiwan University Hospital and a Professor at the College of Medicine, National Taiwan University and the School of Public Health, Taipei Medical University, Taipei. His research interests include clinical and epidemiological studies associated with endocrine diseases. He is currently a member of the Editorial Advisory Board of the following medical journals: the *Current Diabetes Reviews*, the *Taiwan Geriatrics and Gerontology*, the *Open Environmental Science*; the *Open Environmental Letters*, the *Open Environmental Reviews*, the *Open Toxicology Journal*, the *Open Heart Failure Journal*, the *Open Diabetes Journal*, the *Open Translational Medicine Journal*, the *International Journal of Chemical, Environmental, and Pharmaceutical Research*, the *World Journal of Hepatology*, the *World Journal of Cardiology*, and the *RASAYAN Journal of Chemistry*. He has authored or coauthored more than 100 refereed papers and more than ten book chapters in either English or Chinese on arsenic-related health problems and diabetes mellitus, especially focusing on epidemiologic aspects. He has been invited to give lectures and presentations for more than 200 times on diabetes and arsenic-related health hazards locally or internationally. Also in recent years, he has been invited by more than 50 international medical journals as Referee for scientific papers submitted for publication for more than 100 times.

Shyang-Rong Shih received the M.D. degree from the College of Medicine, National Taiwan University, Taipei, Taiwan, in 2000.

From 2000 to 2005, she was in the Department of Internal Medicine, National Taiwan University Hospital. From 2005 to 2008, she was in the Division of Endocrinology and Metabolism, Department of Internal Medicine, National Taiwan University Hospital, Yunlin Branch, Taiwan. Since 2008, she has been with the Division of Endocrinology and Metabolism, Department of Internal Medicine, National Taiwan University Hospital, where she is currently an Attending Physician.