

● *Original Contribution*

AN EARLY VISION-BASED SNAKE MODEL FOR ULTRASOUND IMAGE SEGMENTATION

CHUNG-MING CHEN,^{*‡} HENRY HORNG-SHING LU,[†] and YU-CHEN LIN[†]

^{*}Institute of Biomedical Engineering, National Taiwan Univ., Taipei, Taiwan; and [†]Institute of Statistics, National Chiao-Tung University, Hsin-Chu, Taiwan

(Received 26 April 1999; in final form 13 September 1999)

Abstract—Due to the speckles and the ill-defined edges of the object of interest, the classic image-segmentation techniques are usually ineffective in segmenting ultrasound (US) images. In this paper, we present a new algorithm for segmenting general US images that is composed of two major techniques; namely, the early-vision model and the discrete-snake model. By simulating human early vision, the early-vision model can capture both grey-scale and textural edges while the speckle noise is suppressed. By performing deformation only on the peaks of the distance map, the discrete-snake model promises better noise immunity and more accurate convergence. Moreover, the constraint for most conventional snake models that the initial contour needs to be located very close to the actual boundary has been relaxed substantially. The performance of the proposed snake model has been shown to be comparable to manual delineation and superior to that of the gradient vector flow (GVF) snake model. © 2000 World Federation for Ultrasound in Medicine & Biology.

Key Words: Ultrasound, Image segmentation, Early-vision model, Snake, Discrete-snake model, Speckles, Texture.

INTRODUCTION

Image segmentation plays an essential role in both qualitative and quantitative ultrasound (US) image analyses. A correct segmentation promises a more accurate extraction of clinical information from US images for clinical applications. For example, in quantitative analysis, instead of only using the diameters to quantify the scale of a tumor, the area and the boundary of the tumor may be included for a more comprehensive characterization. In qualitative analysis, image segmentation may better visualize the surface of a three-dimensional (3-D) object of interest. One typical example is the 3-D reconstruction of the face of a fetus.

However, US image segmentation in practice is a very hard problem due to the complex nature of US images. Currently, the most widely used segmentation approach in a clinical US image system is based on thresholding. With this approach, one attempts to distinguish the object of interest from the background by selecting the proper ranges of grey scales. Although this

is a very convenient way to perform image segmentation, its result is usually unsatisfactory because of the speckles and textures in an US image. Another simple way for image segmentation provided in a clinical system is manual delineation. It allows a user to identify the desired boundary manually. Although the segmentation result of this approach is generally acceptable, the long delineation time and heavy working load are usually undesirable. For a volume image with hundreds of 2-D slices, it may take 0.5 h to complete drawing the boundary of the 3-D face of a fetus. Moreover, the results could be user-dependent and even time-dependent. That is, given a set of 2-D slices, different boundaries may be identified by different persons, and even by the same person at a different time.

To accomplish US image segmentation more efficiently and accurately, a computerized approach would be an ideal choice for clinical use. A computerized approach is expected to segment the object of interest automatically or semiautomatically with high reproducibility. In the classic image-processing area, image segmentation may be achieved by two categories of approaches, namely, the region-based and edge-based approaches. The region-based approaches *e.g.*, region growing, split-and-merge, etc.) aim to gather pixels with

[‡]Address correspondence to: Chung-Ming Chen, Institute of Biomedical Engineering, College of Medicine, National Taiwan University, #1, Sec. 1, Jen-Ai Road, Taipei, Taiwan. E-mail: chung@iotus.mc.ntu.edu.tw

similar properties. On the other hand, the edge-based methods try first to find the edges of the region-of-interest (ROI) as much as possible, then follow by an edge-linking operation to delineate the complete boundary of the ROI. For example, one may employ the Sobel operator to detect the edges and the graphic search algorithm to link the edge segments. Although the classic segmentation approaches have demonstrated their capability in various applications, their performance has been degraded severely in US image segmentation due to the noisy nature of US images.

The intrinsic noise of a US image mainly results from such factors as signal processing, image formation, interpolation, tissue property, interference and so on. Two major types of noises may be identified in an US image. The first one is the speckle, which appears as random mottles superimposed on a US image. The speckle is caused by the coherent interference of back-scattered echoes. The speckle occurs when the scatterer size is small compared to the underlying wavelength, and each sample volume consists of many scatterers. It degrades not only the perceivable resolution by a factor of 5 to 7 (Kozma and Christensen 1976), but also the discriminability of subtle difference in grey levels (George et al. 1976). The second one is the tissue-related textures, which may be composed of quasirepetitive patterns or sporadic spots. The pattern of the tissue-related textures varies with the tissue. For example, a liver with cirrhosis usually contains a coarser texture than a healthy liver. It should be noted that the tissue-related textures might be considered as either noise or useful information, depending on the type of image-processing techniques employed. For the nontexture image-processing approaches, the tissue-related textures basically behave like noise. On the other hand, for the texture image-processing approaches, the textural information of the tissue-related textures may be utilized in segmenting the object of interest. But, even with the texture image-processing approaches, the noisy nature may not be eliminated thoroughly due to the irregular pattern of the tissue-related textures.

With the speckle and the tissue-related textures, two classes of edges may be found in a US image; namely, the RG-edge and the RT-edge. The RG-edge is the edge formed by two adjacent areas with different mean grey levels. The RT-edge is the edge formed by two adjacent areas with different textures. As an example, Fig. 1 shows two snapshots of a liver that has the RT-edge at the lower boundary and the RG-edge at the upper boundary.

The difficulty of US image segmentation basically arises from the need to identify both types of edges at the same time. Conventionally, the denoising operations usually involve identifying the RG-edges. However, an

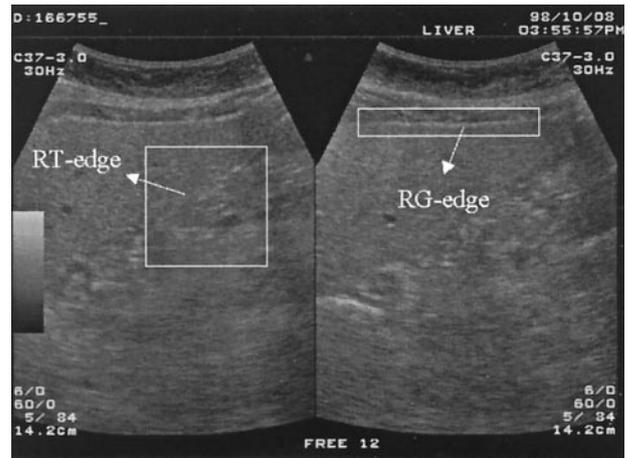


Fig. 1. An ultrasound image consisting of the RG-edge and the RT-edge.

improper denoising may either lead to a shifted edge or smear the desired edge. Furthermore, it not only cannot catch the RT-edges because the textures have been destroyed by the denoising operations, but also may generate the false edges from each texcel (texture element). On the other hand, an algorithm that may detect the RT-edge may not be able to detect the RG-edge effectively; also, the development of the RT-edge-detection algorithm is still in its infancy.

Another two serious problems in US image segmentation occur when the complete boundary of an ROI is to be formed. One difficulty is in distinguishing the desired edges from the false edges and the other is in composing the broken edges. On account of the speckle and the tissue-related textures, false edges may be inevitably generated while the desired RG-edges and RT-edges are sought. If the desired edges are not significant enough, the formation of a complete boundary may be easily trapped at the local optimum formed by the false edges. On the other hand, because of the noise and artefacts, the desired edges would most probably fragment into pieces, or a considerable portion of the desired boundary may be missing. In this case, the segmentation algorithm is required to find a reasonable approximation from the available pieces of edges.

To cope with the influence of the noise in US images, varieties of approaches have been proposed previously for US image segmentation. For example, by using region-growing based on interpixel grey scale difference, component-labeling and dilation-erosion, and watershed segmentation, Brathwaite et al. (1996) proposed two automatic border-detection schemes to detect luminal borders in intravascular images. Czerwinski et al. (1995) suggested a directional median filtering approach to suppress speckle noise while retaining the

structure of the image. For accurate detection of the lumen-intima and media-adventitia boundaries, Fan *et al.* (1996) applied a nonlinear wavelet soft-thresholding algorithm to intracoronary US images. With a particular interest in horizontal image slices of the lower leg, Heckman (1996) developed a model-driven segmentation scheme to automatically find closed contours in a noisy image using a global search. Thomas *et al.* (1991) applied morphological operators to segment femurs in fetal US images. By using a linear combination of the grey level and the local entropy, Zimmer *et al.* (1996) proposed a bivariate extension of minimum cross-entropy thresholding to segment US images containing fluid surrounded by a soft tissue.

Although these works may have achieved satisfactory performance for the given problems, their approaches are either application-specific or segment out homogeneous regions with distinguishable mean grey levels from the background. As a consequence, they may suffer a limited extension to other types of US images. In contrast to these works, at least three classes of general approaches have been considered for US image segmentations. One is to use the speckle-reduction technique to alleviate the influence of the speckle. A good example is that Kotropoulos (1994) proposed a signal-adaptive maximum likelihood filter to remove speckle, and applied a self-organizing neural network to image segmentation. Another is to consider US images as texture images and apply texture image-segmentation approaches. For instance, two textural approaches have been suggested for US image segmentation (*i.e.*, a multiresolution framework incorporating simulated annealing) (Muzzolini *et al.* 1993) and a segmentation method based on a set of sampled textures (Muzzolini *et al.* 1994). Recently, Muzzolini *et al.* (1998) have proposed the Inck (incomplete knowledge) criterion function used in conjunction with a multiresolution classifier design that has been applied to US image segmentation. The other is to employ the deformable model to find the complete boundary of an ROI. For example, Chalana *et al.* (1996) proposed a multiple active-contour model for cardiac boundary detection on echocardiographic sequences. Theoretically, each class of approaches can only partially solve the aforementioned problems for US image segmentation. More specifically, the speckle reduction, the textural approaches and the deformable models are conceptually designed for finding the RG-edge, the RT-edge and the complete boundary, respectively. However, they might not be as effective for solving the other problems.

To find the complete boundary of an ROI composed of both RG-edges and RT-edges, in this paper we present a new snake model that combines the advantages of the textural approaches and the deformable models for segmentation of general US images. The key ideas of the

proposed snake model are to incorporate the distance map as a new image force and to move the snaxels (snake elements) discretely for better noise immunity and more accurate convergence. The distance map is an effective feature based on the early vision model proposed in one of our early works (Lin 1997). It is able to capture not only the RG-edges but also the RT-edges in US image. On the other hand, the discrete concept (*i.e.*, moving the snaxel discretely) only considers the local maxima in the distance map as the candidate positions for snaxel movement. As contrast, all points in the image are candidate positions in the conventional snake models. The discrete concept has allowed the proposed snake model to penetrate the energy barriers made by the noise more easily and guaranteed that all snaxels will converge to the edges defined by the distance map. Consequently, the constraint imposed by most conventional snake models, that the initial contour of a snake needs to be located very close to the actual boundary to avoid getting trapped in the local minima during deformation, has been relaxed greatly in our approach.

This paper is organized as follows. The biological and proposed early vision models are described next, followed by presentation of the proposed discrete snake model. Then, performance of the proposed segmentation algorithm is analyzed and conclusions are given.

THE EARLY VISION MODEL

Image force is one of the major forces that can guide a snake to converge to the desired boundary. However, the classic image-processing techniques frequently cannot provide an effective image force on US image due to the interference of the speckle and the tissue-related textures. On the other hand, the human vision has demonstrated its incredible capability in object recognition, even in a noisy environment. Although US image segmentation seems to be intractable to the classic techniques, the human vision shows no difficulty at all in identifying the boundary of the object of interest in a normal situation. Encouraged by the success of our recent effort on edge identification based on an early vision model (Lin *et al.* 1997) that may identify both of the RG-edges and RT-edges in a US image, we propose to incorporate the early vision model into the snake model for a better image force.

RT-edge identification based on an early vision model has been investigated extensively in the past for the generic texture images. Because the entire visual process still remains unclear to the scientists, these works have been mostly focused on simulating the simple cells of the V1 area, for which the scientists have the best idea compared to the other part of human vision. (Some

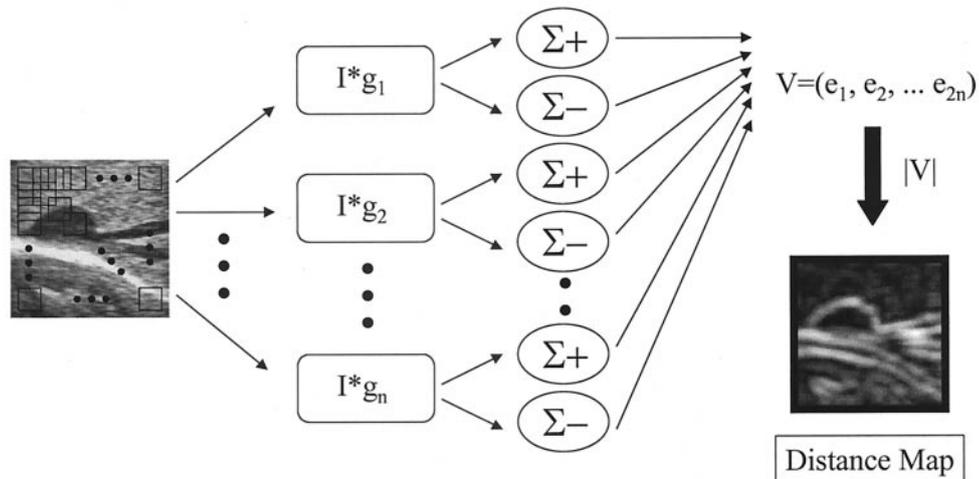


Fig. 2. Derivation of the distance map. The underlying liver US image is decomposed into overlapping blocks of subimage. Each block is filtered with a set of Gabor functions to derive its G-vector. The distance map is formed by the G-vector lengths of all blocks.

excellent works are Dunn et al. 1994; Jain and Farrokhnia 1991; Malik and Perona 1990; Tan 1995.

Based on the recent researches on psychophysics, psychophysiology and neurology, it has been widely adopted (Tan 1995) that simple cells are the fundamental processing elements for visual information. Each simple cell is tuned to a specific narrow frequency and orientation band. The receptive field profile of a simple cell may be modeled as an even or an odd Gabor function. The inputs of a simple cell are approximately summed and weighted by the corresponding Gabor coefficients. The output of a simple cell is half-way rectified.

Mathematically, a Gabor function may be expressed by a Gaussian function multiplied by a sinusoidal function,

$$g(x, y) = \exp\{-[(x - x_0)^2/a^2 + (y - y_0)^2/b^2]\pi\} \exp\{-2\pi i[u_0(x - x_0) + v_0(y - y_0)]\}. \quad (1)$$

In practice, the receptive field profile may also be modeled by several other functions closely *e.g.*, the DOOG functions). The common step of the vision model-based approaches is to generate the neuroimages by filtering the underlying image with a set of approximated receptive field profiles, such as Gabor functions, DOOG functions, and so on. After an image is filtered by the receptive field profiles of V1 cells, the scientists still have no thorough idea about how to compose the neuroimages. As a result, previous vision model-based approaches have either counted on the classic techniques (Jain and Farrokhnia 1991) or integrated the limited knowledge of

the visual processes (Malik and Perona 1990) to further process the neuroimages. For example, various techniques *e.g.*, the CLUSTER algorithm (Jain and Farrokhnia 1991), the neural network (Van Hulle and Tollenaere 1993), etc., have been suggested to utilize the textural information embedded in the neuroimages for the texture image segmentation.

In contrast to the hybrid approach adopted by the previous early vision models, our early vision model (Lin et al. 1997) has been designed based on the textural segregation model proposed by Chen (1994) with the support of psychophysical experiments. Moreover, the proposed vision model has attempted to exploit the local texture information rather than the global texture information in the filtering stage. As illustrated in Fig. 2, the underlying liver US image (*i.e.*, the image at the left side of Fig. 2) is decomposed into overlapped blocks of subimages. For each pixel in the underlying image, a block of $m \times m$ subimage centered at the pixel, which is illustrated by a square box, is extracted for feature computation. For each block of subimage, the G-vector is computed as follows. First of all, the image is filtered by n Gabor functions with difference orientations and central frequencies, which is represented by (I^*g_i) in Fig. 2. Then, the pixels in each neuroimage with positive values and negative values are summed up respectively, which is denoted by $(\Sigma+)$ and $(\Sigma-)$. This step gives two unique numbers for each neuroimage and is to simulate the half-way rectification phenomenon observed from the V1 cells. Finally, the G-vector of a subimage is the vector consisting of the $2n$ numbers computed from the n neuroimages. According to Chen (1994), the similarity

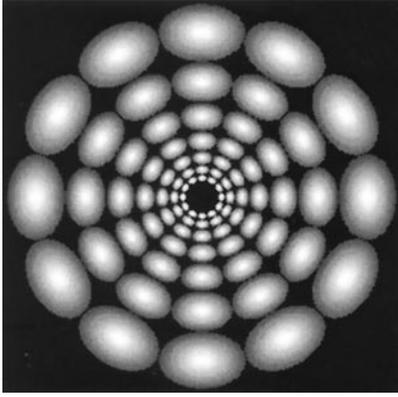


Fig. 3. A rosette map consisting of 42 Gabor filters. The half-peak magnitude frequency bandwidth is set to 1 octave for each Gabor filter. Only the half-peak supports of the filters are shown in the map.

between two textures may be modeled as the Euclidean distance of the G-vectors of the two textures.

Selection of the proper Gabor functions remains an open problem in the vision model-based approaches. In this study, we adopt the rosette map proposed by Jain and Farrokhnia (1991). As an example, Fig. 3 gives a rosette map containing 42 Gabor filters in the frequency domain. Only the half-peak supports of the filters are shown in the map. The DC component is at the center of the map. For a Gabor filter centered at the u axis, the half-peak magnitude bandwidth is given by:

$$B_r = \log_2 \left(\frac{u_0 + (2 \ln 2)^{1/2} \sigma_u}{u_0 - (2 \ln 2)^{1/2} \sigma_u} \right), \quad (2)$$

and

$$B_\theta = 2 \tan^{-1} \left(\frac{(2 \ln 2)^{1/2} \sigma_v}{u_0} \right), \quad (3)$$

where

- B_r = frequency bandwidth,
- B_θ = orientation bandwidth,
- u_0 = the center frequency,
- σ_u = the standard deviation of the Gaussian envelope in the u direction,
- σ_v = the standard deviation of the Gaussian envelope in the v direction.

Because it has been shown by several experiments that the frequency bandwidth of simple cells is approximately one octave (Pollen and Ronner 1983), B_r has been set to 1 in this study. For a better coverage of the

frequency domain, B_θ has been chosen to be 30° . After the G-vector has been derived for each block, the G-vector lengths of all blocks form a new characteristic image called distance map. The peaks of the distance map define the RT-edges and the RG-edges of the original US image. It is worthy of note that the block size determines the edges obtainable in different resolutions. For an object of interest with large texcels, a large block size would be required to catch the RT-edges. Empirically, for a US image with a good quality, block size 8 (*i.e.*, 8×8) or block size 16 (*i.e.*, 16×16) would be a reasonable choice, considering the computational efficiency. The more detailed discussion of our early vision model may be found in Lin *et al.* (1997) and Chen (1999).

THE DISCRETE SNAKE MODEL

The snake model was first proposed by Kass *et al.* (1987) to find a continuous closed contour that minimizes an energy function. A contour Γ is defined by $\mathbf{v}(s) = (x(s), y(s))$, where s is the arc length of the contour. Although various snake models have been proposed for different applications, the energy function to be minimized for most snake models may be generally defined as

$$E_{\text{snake}}(\Gamma(s)) = \oint [E_{\text{int}}(\mathbf{v}(s)) + E_{\text{image}}(\mathbf{v}(s)) + E_{\text{external}}(\mathbf{v}(s))] ds. \quad (4)$$

In eqn (4), E_{int} denoting the internal force, which controls the smoothness of the contour, is one of the major forces for the snake to shrink, and is usually given by:

$$E_{\text{int}}(\mathbf{v}(s)) = E_{\text{cont}}(\mathbf{v}(s)) + E_{\text{curv}}(\mathbf{v}(s)) = \frac{1}{2} (\alpha(s) |\Gamma_s(s)|^2 + \beta(s) |\Gamma_{ss}(s)|^2) \quad (5)$$

where $\Gamma(s) = (x(s), y(s))$ is the parametric representation of the contour. $\Gamma_s(s)$ and $\Gamma_{ss}(s)$ are continuity (*i.e.*, the absolute value of the first-order derivative) and curvature terms (*i.e.*, the absolute value of the second-order derivative), respectively. $\alpha(s) (\geq 0)$ controls the stretching property at the point s . $\beta(s) (\geq 0)$ controls the bending property at the point s .

E_{image} denotes the image force that stops the snake at the desired image features by minimizing the energy term. One common image force is $-\lambda |\nabla I|$, where $|\nabla I|$ is the length of the gradient vector of intensity and $\lambda (\geq 0)$ is a weighting factor. That is, the image force finds the boundary points that have maximum change of intensi-

ties. E_{external} represents the external constraint forces that pull the snake toward the desired image features or offer power for a snake to escape from the undesired local shape.

The basic idea of the snake model is to find a snake minimizing E_{snake} . That is,

$$E_{\text{snake}}(\mu) = \min_v E_{\text{snake}}(v), \quad (6)$$

where the contour function $v(s) = (x(s), y(s))$ is a function in $C^2 \times C^2$. That is, the second derivatives are assumed to be existent. Various approaches have been proposed to optimize eqn (4). For example, E_{snake} was minimized by the steepest descent approach in (Kass et al. 1987),

$$\frac{d\Gamma(s)}{dt} = -\frac{\delta E}{\delta \Gamma} = \alpha \Gamma_{ss}(s) - \beta \Gamma_{ssss}(s) + \lambda \vec{\nabla} |\vec{\nabla} I(v(s)) \cdot \vec{\nabla} I(v(s))|, \quad (7)$$

which may be derived from the calculus of variation. The snake model tends to deform centripetally when the internal force shrinks and there is no large image force to counterbalance it. Thus, centrifugal force may be required during deformation either to reverse the shrinking process or to ease deformation. To account for the centrifugal force, a snake-balloon model was proposed by Cohen (1991), which can be expressed as

$$E_{\text{snake}}(\Gamma(s)) = \oint [E_{\text{int}}(v(s)) + E_{\text{image}}(v(s)) + E_{\text{external}}(v(s))] ds - \gamma \iint_R dA, \quad (8)$$

where $\iint_R dA$ is the balloon force that is the area of the region R enclosed by the snake and $\gamma(\geq 0)$ is a weighting factor. By minimizing the balloon force, the area enclosed by the contour tends to be inflated; that is, the balloon is inclined to push the contour outward along its normal vector, $\vec{n}(s) = (-\dot{y}(s), \dot{x}(s))$. Because the snake model and the snake-balloon model share the same principle in deformation through energy minimization, the names of these two models will be used interchangeably here, even though they are historically different.

Since the work of Kass et al. (1987), many approaches have been proposed to augment the capability of the snake model. Williams and Shah (1992) gave a more accurate definition on curvature. Leymarie and Levine (1991) proposed a multistage approach attempt-

ing to make the snake resistant to the noise. Xu et al. (1994) have proposed a methodology for insensitive parameter selection by adding one normal force to counterbalance the internal force and another one to keep the snake moving. Yuen et al. (1996) have focused on the corner edge detection and proposed an enhanced snake algorithm by blowing the roughly converged snake to the concave part of an object segment by segment. To relax the constraint that the initial contour of a snake needs to be placed very close to the desired contour and to alleviate the difficulty in deforming into concave area, Xu and Prince (1998) have proposed a new external force called gradient vector flow (GVF) to guide snake deformation.

Even though the snake model has been studied extensively, it still suffers at least two fundamental problems. The first problem is that a snake can be easily trapped by the noise. Although the noise can be reduced or even eliminated by denoising techniques, either important fine detail may be lost or the snake may converge to an incorrect boundary. In addition to the noise, a snake can be easily trapped by the edges belonging to artefacts or other objects in the images. As a result, the initial contour of a snake is usually allocated very close to the real boundary in most snake models, which is an undesired constraint for clinical use.

The second problem arises from the difficulty in determination of weighting factors. The deformation behavior of a snake is basically controlled by the relative importance of different energy terms, which is defined by the weighting factors. For example, at the beginning of shrinking process, it may require a larger shrinking force to move the snake inward. In this case, a larger α would generally be expected. When the snake is close to the actual boundary, a larger image force would be desired to stop the deformation. Therefore, a larger λ would be anticipated. However, in practice, determination of a good parameter value is by no means trivial, not to say selecting parameter values adaptively and robustly. A well-known phenomenon caused by this problem is that the maximum curvature that may be captured by the snake is strongly dependent on the weighting factor of the curvature term. In fact, as a result of the compromise of different energy terms, a more general phenomenon is that a snake frequently cannot converge to the edge positions defined by the edge detectors. This phenomenon is called boundary deviation for convenience.

To cope with these two fundamental problems, in this paper we propose a new snake model, called the discrete snake model. The essential ideas of the discrete snake model are to carry out energy minimization on the distance map of the original image and to move the snake discretely. Performing snake deformation on the distance map rather than on the original image has two

advantages. One is that the noise effect due to the speckle and the tissue-related textures has been alleviated substantially. The other is that the desired edges have been highlighted, which provides a better image force to stop the snake deformation.

The discrete concept (*i.e.*, moving the snaxels discretely) suggests that, instead of searching the next position for a snaxel along its searching path pixel by pixel, it only consider the peak points of each path (*i.e.*, the local maxima) as the searching space. It is discrete in the sense that the candidate snaxel positions (*i.e.*, the local maxima) distribute discretely in contrast to continuously as in the conventional snake model. The advantages of the discrete concept are threefold. One is that the searching space for the optimal solution is greatly reduced. For example, it required less than 10 iterations for all the cases we tested to converge. Another is that it has a greater capability to go across the energy barrier made by the noise and, thereby, has better noise immunity. The other advantage is that, after the snake converges, it guarantees that all snaxels will reside on the edges defined by the distance map, which overcomes the boundary deviation problem due to the compromise among all energy terms.

In the proposed snake model, the searching path for the discrete snake model is predetermined to be the centripetal line from each initial snaxel to the center point defined by its adjacent snaxels. At every peak position on the path, the energy at the current position is compared with that at the next peak. The snaxel movement is performed if the energy decreases.

Determination of weighting factors

Although the discrete concept and the distance map have greatly alleviated the noise problem, a new approach is proposed in this paper for determination of the weighting factors adaptively. The image force employed in the discrete snake model is the grey level of the distance map. The ratio among the continuity force, the curvature force, the image force and the balloon force is set to be 1 : 1 : -3 : -1. The negative signs account for the counterbalance to the shrinking processes. This ratio has been chosen to balance the image force with the sum of the other three forces.

The weighting factors, α , β and γ , are changed at each peak position according to its relationship with the neighboring snaxels so that the force ratio (1:1:-3:-1) is maintained, given λ set to 3. Consider three consecutive peak points, $v^{k-1}(i)$, $v^k(i)$ and $v^{k+1}(i)$ on the i th searching path, which corresponds to the last, current and next snaxel positions, respectively. A searching path is an imaginary path from each snaxel on the initial contour to the center point defined by its adjacent snaxels. That is, each snaxel is moved along its searching path cen-

tripetally or centrifugally. Then, the energy function of the proposed snake model is:

$$E_{\text{snake}}^K(i) = (\alpha^K(i)|\Gamma_s^K(i)|^2 + \beta^K(i)|\Gamma_{ss}^K(s)|^2) - 3I(v^K(i)) - \gamma^K(i) \iint_{R^K} dA, \quad (9)$$

for $K = k - 1$, k , and $k + 1$. Note that the i th snaxel is moved from the k th position to the $(k + 1)$ st position if $E_{\text{snake}}^{k+1} < E_{\text{snake}}^k$.

To keep the image force in balance with the other forces, $\alpha^k(i)$, $\beta^k(i)$ and $\gamma^k(i)$ are chosen so that:

$$\alpha^k(i) = \frac{I(v^{k-1}(i))}{|\Gamma_s^{k-1}(i)|^2}, \beta^k(i) = \frac{I(v^{k-1}(i))}{|\Gamma_{ss}^{k-1}(i)|^2}, \gamma^k(i) = \frac{I(v^{k-1}(i))}{\iint_{R^{k-1}} dA}, \quad (10)$$

where $I(v^{k-1}(i))$, $|\Gamma_s^{k-1}(i)|^2$, $|\Gamma_{ss}^{k-1}(i)|^2$ and $\iint_{R^{k-1}} dA$ denote the image force, continuity force, curvature force and balloon force of the last snaxel on the i th path, respectively. Note that the continuity force for current peak position on the i th searching path is defined as

$$|\Gamma_s^k(i)|^2 = \min(|v^k(i) - v^k(i-1)|^2, |v^k(i+1) - v^k(i)|^2), \quad (11)$$

which is different from that in the conventional snake model. The conventional continuity force is defined by only the first term in the min function of eqn (11), which is the distance between the current snaxel position and the previous snaxel position on the boundary contour. Undesirable continuity force would occur with the conventional definition in case the previous snaxel position happens to be on an abnormal position, which may prevent the current snaxel from moving a further step. Hence, by considering the minimum of the two adjacent distances, more distance information is taken into account and the chance for current snaxel to jump to the next peak position is higher. In case of the first iteration, because last energy term information is not available, α , β and γ are all set to 1.

Convergence criteria

The convergence criterion for the discrete snake model can be as simple as all snaxels coming to a full stop. It is made possible by the following reasons. First

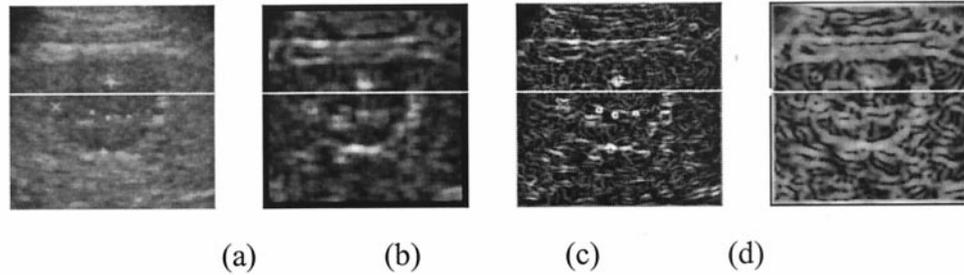


Fig. 4. (a) Original ultrasound image; (b) the distance map derived using the block size 8; (c) the filtered image by the Sobel operator; (d) the filtered image by the Gaussian function followed by the Sobel operator.

of all, in this model, the searching path has been explicitly defined as the centripetal line from each initial snaxel to the center point defined by its adjacent snaxels, which could prevent the snake from oscillating around the desired solution. Second, the image force from the distance map has been considered reliable and given a large weight, which could effectively stop the snake deformation. Thirdly, by considering only the peaks along each path, the internal force is more influential than that in the classic snake models. In short, these features have made the discrete snake model either easier jump to a better snaxel position or more accurately stop at a desired position.

EXPERIMENTS AND DISCUSSIONS

Validation of the proposed snake model has been carried out on 23 liver US images with various kinds of hepatic tumors, such as hepatocellular carcinoma (HCC), cavernous hemangiomas, metastatic liver cancer, hyperplastic nodule, and so on. Only one tumor per image was chosen for segmentation. The clinical US imaging system employed was Toshiba SSA-380A and the operating frequency was 3 MHz. All images were selected by medical doctors and captured through a frame grabber card. To evaluate the performance of the proposed snake model, the segmentation results have been compared with the tumor boundaries drawn by an experienced medical doctor (Dr. A). In comparison with the previous snake model, the GVF snake model (Xu and Prince 1998) has been chosen for the following two reasons. One is that, like the proposed snake model, the GVF snake model has been designed to allow a snake to deform from a distant initial contour. The other is that, by using the GVF program obtained from the Web pages of the original designers, we are able to faithfully reproduce the performance of the GVF snake model on the testing US images.

Distance map

The distance maps for all 23 US images are calculated using the block size 8. The even Gabor functions are selected according to the rosette map as suggested by Jain and Farrokhnia (1991). The center frequencies employed are $\sqrt{2}/2$, $\sqrt{2}$, and $2\sqrt{2}$. The angles used are 0, 30, 60, 90, 120, and 150°. As an example, the distance map of Fig. 4a is given in Fig. 4b. Generally speaking, the distance map is expected to provide better edge information than those classic approaches based on smoothing followed by edge detection. The reason is that the distance map can suppress the edges of the texcels and highlight the RT-edges, whereas the latter gives only the edges of texcels (*i.e.*, the RG-edges) in different resolutions. The consequence is that the distance map has more significant RT-edges, but less annoying false edges. To demonstrate these two properties, the image derived by the Sobel edge detector without smoothing beforehand is given in Fig. 4c. Because the Sobel operator is basically to capture the RG-edge, as expected, Fig. 4c shows a very noisy image with the RG-edges of all texcels in Fig. 4a. As a contrast, Fig. 4b highlights the RT-edges substantially.

With a closer observation, the cross-sectional profiles of the 56th row of Figs. 4a–c are plotted in Fig. 5. The two ellipses in the line profiles (Fig. 5) indicate the reasonable edge locations. One can see that the Sobel operator and the distance map both give a peak at the

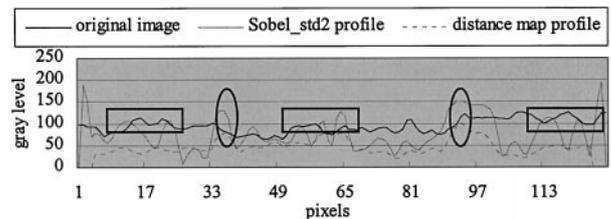


Fig. 5. Cross-sectional profiles of the 56th rows of Fig. 4a–c. The two ellipses indicate the reasonable edge locations.

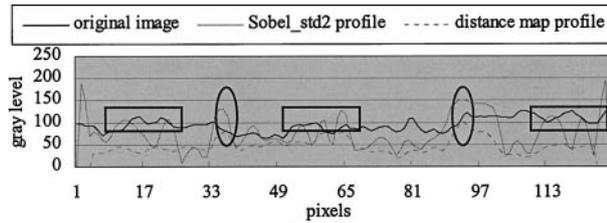


Fig. 6. Cross-sectional profiles of the 56th rows of Fig. 4a, b, d. The two ellipses indicate the reasonable edge locations. The rectangles point out the disturbing false edges in the “Sobel_std2 profile.”

edge locations, whereas the “Sobel profile” is much noisier than the “distance map profile.” For example, although the “Sobel profile” does give a peak in the right ellipse, another undesirable large peak also exists right next to the desired peak, which will interfere with the further determination of the true edges.

Although a less noisy profile may be anticipated by incorporating a smoothing operation before the Sobel operator, essentially, the attainable edge information is still the RG-edges of the texcels in the lower resolutions. For instance, Fig. 4d shows the edge information detected by using a Gaussian filter of standard deviation 2 for smoothing followed by the Sobel operator for edge detection. Similarly, the line profiles for the 56th rows of Fig. 4a, b and d are plotted in Fig. 6. As expected, Fig. 4d provides better edge information than Fig. 4c at the desired edge locations because of smoothing. However, the significant peaks indicating the RG-edges of the texcels in a lower resolution overwhelm the desired RT-edge information that we are looking for. For example, in the “Sobel_std2 profile” of Fig. 6, the strong peaks enclosed by the rectangles, the amplitudes of which are comparable to those peaks enclosed by the ellipses, define the undesirable RG-edges and may severely disturb the further identification of the desired boundary. In short, one may observe that the classic approaches basically give the more annoying false edges than the distance map because the classic approaches detect only the RG-edges in different resolutions.

Segmentation results and analyses

A self-developed Windows program, called MediaX, has been provided for the medical staff to perform US image segmentation using the proposed snake model. By using the MediaX, one can either choose predefined shapes, such as ellipses and rectangles, or draw any irregular shape as the initial contour. If necessary, the ratio among the weighting factors may be modified through the parameter setup dialogue. After the initial contour is given, the MediaX will automatically select the snaxels from the initial contour and generate the

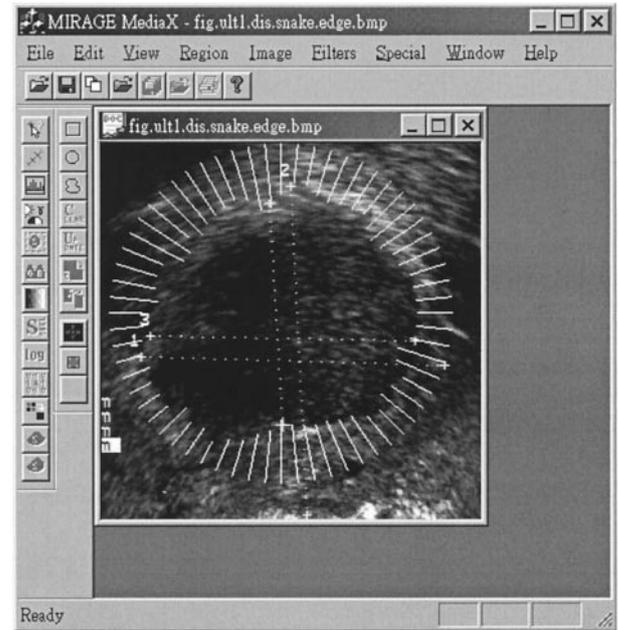


Fig. 7. The MediaX and a portion of an US image consisting of 64 centripetal searching paths. The outer and inner ending points of all searching paths form the initial contour and the final boundary found by the snake, respectively.

searching path for each snaxel. For example, Fig. 7 shows the MediaX and a portion of a captured US image consisting of 64 equally spaced snaxels, each with a centripetal searching path. The outer and inner ending points of all searching paths form the initial contour and the final boundary found by the snake, respectively. The boundary of the object of interest is marked with the “+”s by medical doctors.

For all 23 liver US images, the initial contours have been placed reasonably far away from the tumors. The average and the maximum distances between the boundary derived by the proposed snake model and that drawn

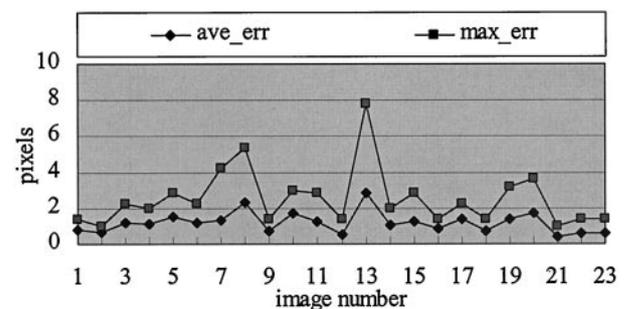


Fig. 8. The average and the maximum distances between the boundary derived by the proposed snake model and that drawn manually by Dr. A. The average and the maximum distances of each derived boundary are labeled as “ave_err” and “max_err.”

Table 1. The means and the standard deviations for the “ave_err” and “max_err” shown in Fig. 8 (unit: pixels)

	Mean	Standard deviation
ave_err	1.18	0.59
max_err	2.53	1.58

manually by Dr. A for each testing image are plotted in Fig. 8. The distance between each snaxel and the drawn boundary is defined as the length of the shortest path from the snaxel to the drawn boundary. The average and the maximum distances of all snaxels in each derived boundary are labeled as “ave_err” and “max_err”, respectively. Table 1 lists the means and the standard deviations for both “ave_err” and “max_err”. From Fig. 8 and Table 1, it is clear that the proposed snake model has successfully deformed from a distant initial contour to a boundary that is very close to the manually drawn boundary. For most testing images, the proposed snake model has achieved the average and the maximum distances less than 1.8 and 4.1 pixels, respectively. Pictorially, the segmentation results for 4 of the 23 liver US images are presented in Figs. 9–11 as examples. These 4 images are images 3, 7, 13 and 18, in which images 3 and 7 contain hypoechoic tumors and images 13 and 18

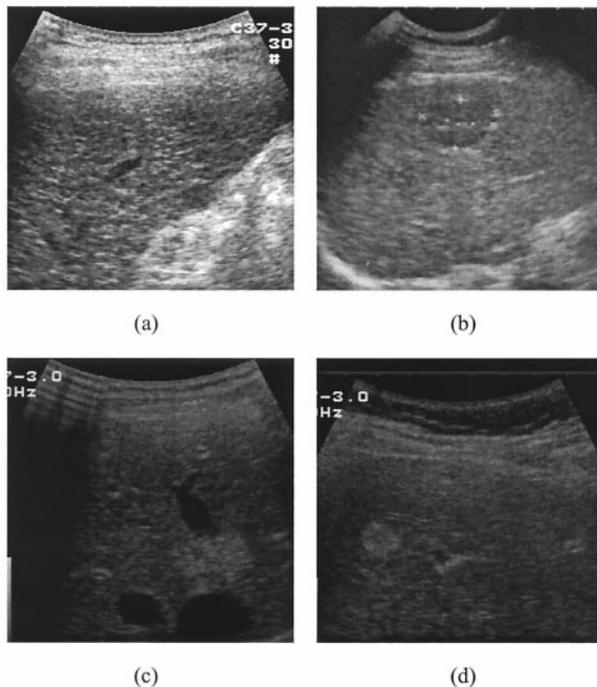


Fig. 9. Image clips of four testing image containing the tumors: (a) image 3; (b) image 7; (c) image 13; and (d) image 18.

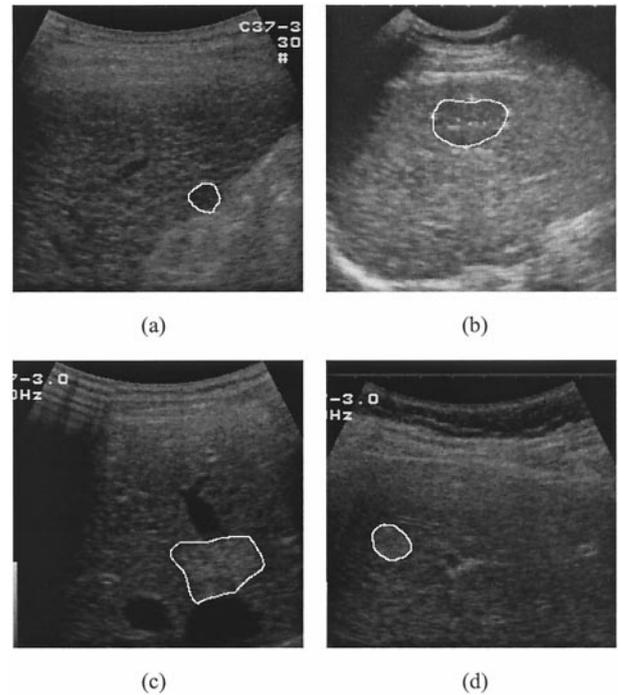


Fig. 10. The manually drawn boundaries by Dr. A for Fig. 9a–d, respectively: (a) image 3; (b) image 7; (c) image 13; and (d) image 18.

contain hyperechoic tumors. Figure 9a–d shows the parts of images consisting of the tumors in the original images 3, 7, 13 and 18, respectively. Figure 10a–d provides the manually drawn boundaries by Dr. A for Fig. 9a–d, respectively. Figure 11a–d gives the boundaries derived by the proposed snake model for Fig. 9a–d, respectively, in which the initial contours and the derived boundaries are as indicated. From Fig. 11, one can see that the proposed snake model has overcome the interference of various kinds of noise surrounding the objects of interest and attained acceptable boundaries. For instance, the snake in Fig. 11d has successfully penetrated the possible energy barriers made by the strips around the upper-right corner of the initial contour and reached the tumor.

Although further improvement of the proposed snake model is certainly required to optimize the snake deformation, the discrepancy between the derived boundary and the drawn boundary may be ascribed to two major error sources for the tested cases. One is the ill-defined boundary of a tumor. The other is the human-dependent judgement of a tumor boundary. An ill-defined boundary not only can misguide the snake deformation because of the false or weak edge information, but also can lead to quite different decisions of medical doctors on a tumor boundary. The human-dependent judgment is partly due to the ill-defined boundary and partly owing to the subjective opinions of the definition

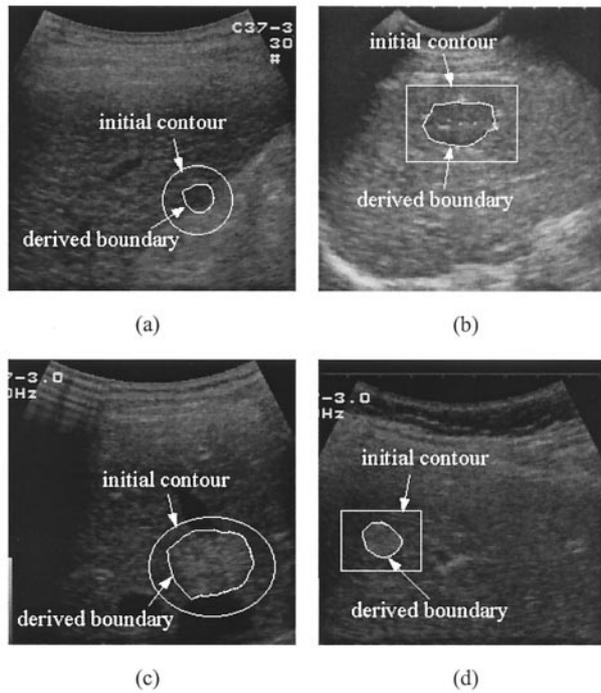


Fig. 11. The boundaries derived by the proposed snake model for Fig. 9a–d, respectively, in which the initial contours and the derived boundaries are as indicated: (a) image 3; (b) image 7; (c) image 13; and (d) image 18.

of a tumor boundary. For example, we have asked another experienced medical doctor (Dr. B) to delineate the boundaries of the tumors in Fig. 9 and the results are given in Fig. 12. For comparison, in Table 2, the row “DS vs. A” lists the average and the maximum distances between the boundary derived by the proposed snake model and that drawn by Dr. A for each image in Fig. 9. In addition, the row “B vs. A” lists these two distances between the boundaries drawn by Dr. A and Dr. B. The columns labeled as “average” and “maximum” contain the average and the maximum distances for each case, respectively.

From Fig. 10, Fig. 12 and Table 2, we can clearly see that Dr. B has made quite different edge definitions from those of Dr. A because of these two error sources. Table 2 also shows that, for each image, the distance between the derived boundary and the boundary drawn by Dr. A is comparable to that between the boundaries drawn by Dr. A and Dr. B. This phenomenon has two important implications. One is that, compared with manual delineation, the proposed snake model has achieved a reasonable performance. The other is that a better golden standard is required to evaluate the performance of a segmentation algorithm more accurately for US images with ill-defined boundaries.

In comparison with the GVF snake model, the same

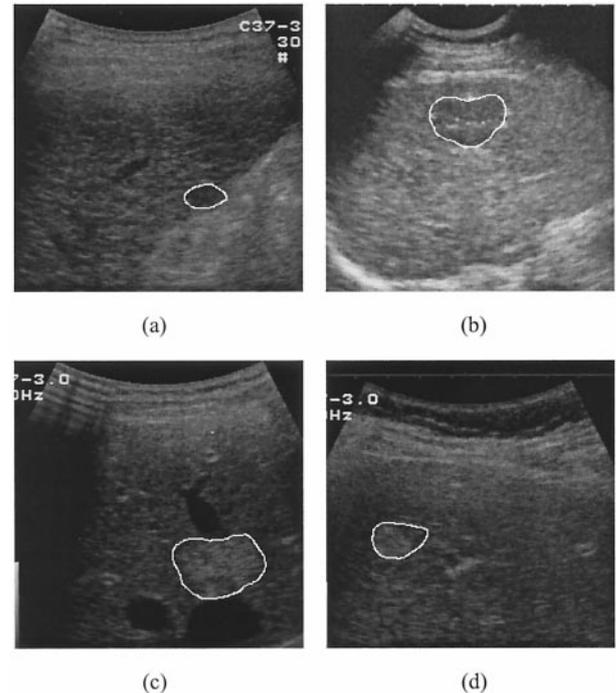


Fig. 12. The manually drawn boundaries by Dr. B for Fig. 9a–d, respectively: (a) image 3; (b) image 7; (c) image 13; and (d) image 18.

initial contour as that employed in the proposed snake model has been used in the GVF snake model for all 23 testing images. The essential idea of the GVF snake model is to extend the influence of the image force to a larger area by generating the GVF field exerted from the image force. The GVF field is computed from the edge map, which is derived using the Sobel edge detector on the Gaussian-smoothed version of the original image, in our implementation. Unfortunately, the GVF failed to give an acceptable result for any of these 23 images. As examples, Fig. 13 shows the final boundaries derived by the GVF snake model with distant initial contours for the 4 US images given in Fig. 9, which are typical segmentation results commonly found in the 23 tested cases. In Fig. 13a, the snake basically did not deform much and got stuck at the local minima made by the noise. The reason is that the GVF field did not spread the image force of the desired boundary far enough to attract the snake toward the object of interest. As a result, the attracting force propagated from the desired boundary is smaller than that of local noise. In Fig. 13b–d, one can see that some snaxels have been guided to the desired boundary, but others have been attracted toward undesired local minima by the GVF field. Because the segmentation results are far from satisfactory, their distances to the corresponding drawn boundaries are not provided in this paper.

Table 2. The average and the maximum distances between two pairs of boundaries for each image in Fig. 9

	Image 3		Image 7		Image 13		Image 18	
	Average	Maximum	Average	Maximum	Average	Maximum	Average	Maximum
DS vs. A	1.16	2.24	1.31	4.24	2.88	7.81	0.72	1.41
B vs. A	2.07	5.39	1.78	5.83	1.87	5.39	1.12	3.00

One is the boundary derived by the proposed snake model and that drawn by Dr. A, denoted as "DS vs. A." The other is the boundaries drawn by Dr. A and Dr. B, denoted as "B vs. A" (unit: pixels).

Even with much closer initial contours than those used in Fig. 13, the performance of the GVF snake model is still worse than the proposed snake model. Taking the 4 US images in Fig. 9 as examples, Fig. 14 shows the final boundaries obtained by the GVF snake model with close-by initial contours. Note that the same standard deviations have been employed for the corresponding images in both Figs. 13 and 14, which are 1, 2, 0.1 and 2, respectively. The average and the maximum distances between the boundary derived by the GVF snake model and that drawn by Dr. A for each image in Fig. 9 are also given in Table 3. By comparing Figs. 13–14 and Table 3 with Fig. 11 and Table 2, it is reasonable to conclude that the proposed snake model is superior to the GVF snake model for the segmentation of US images.

CONCLUSIONS

A new snake model has been proposed for US image segmentation in this paper. The key ideas of the proposed snake model are to incorporate the distance map computed from an early vision model as a new image force, and to move the snaxels (snake elements) discretely for better noise immunity and more accurate convergence. It has been shown that the distance map has effectively enhanced the edges of the object of interest and suppressed the speckle noises at the same time. It catches not only the RG-edges but also RT-edges. By performing deformation only on the peaks of the distance map, the discrete snake model has greatly conquered the difficulty encountered in the conventional snake models

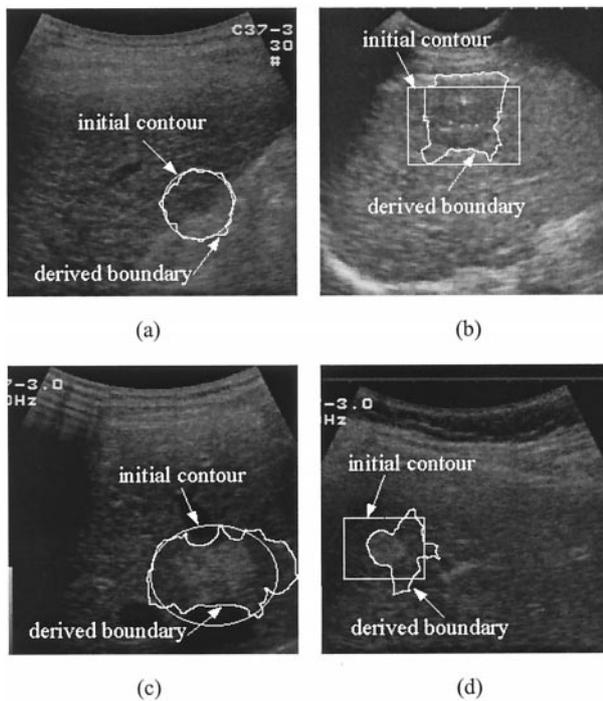


Fig. 13. The boundaries derived by the GVF snake model with a distant initial contour for Fig. 9a–d, respectively, in which the initial contours and the derived boundaries are as indicated: (a) image 3; (b) image 7; (c) image 13; and (d) image 18.

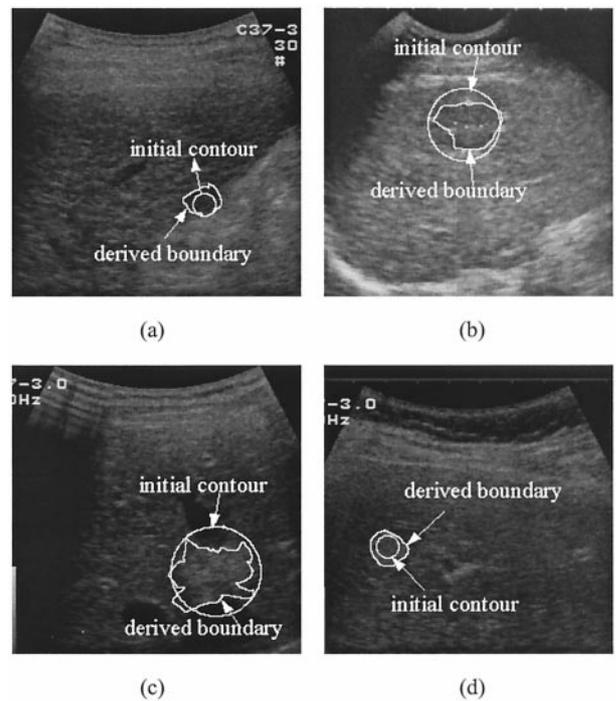


Fig. 14. The boundaries derived by the GVF snake model with a distant initial contour for Fig. 9a–d, respectively, in which the initial contours and the derived boundaries are as indicated: (a) image 3; (b) image 7; (c) image 13; and (d) image 18.

Table 3. The average and the maximum distances between the boundary derived by the GVF snake model and that drawn by Dr. A for each image in Fig. 9 (unit: pixels)

	Image 3		Image 7		Image 13		Image 18	
	Average	Maximum	Average	Maximum	Average	Maximum	Average	Maximum
GVF vs. A	1.58	4.0	2.85	14.9	4.82	13.3	0.91	2.24

(*i.e.*, easily trapped by the local minima resulted from the speckle noises and failing to catch the insignificant edge features and RT-edges). As a consequence, the constraint for most conventional snake models that the initial contour needs to be located very close to the actual boundary has been relaxed substantially.

The performance of the proposed snake model has been compared with manual delineation and the performance of the GVF snake model. From the experimental results, the mean values of the average and the maximum distances between the boundaries derived by the proposed snake model and the manually drawn boundaries are as small as 1.18 and 2.53 pixels, which are comparable to the discrepancy between manually drawn boundaries. On the other hand, the GVF snake model, though also designed with the capability of using a distant initial contour, has been shown to be inferior to the proposed snake model.

Acknowledgements—This work was supported in part by National Science Council and National Health Research Institute, Taiwan, R.O.C.

REFERENCES

- Brathwaite PA, Chandran KB, McPherson DD, Dove EL. Lumen detection in human IVUS images using region-growing. Proceedings of the 1996 23rd Annual Meeting on Computers in Cardiology 1996:37–40.
- Chalana V, Linker DT, Haynor DR, Kim Y. A multiple active contour model for cardiac boundary detection on echocardiographic sequences. IEEE Trans Med Imaging 1996;15(3):290–298.
- Chen CM. An edge identification framework for ultrasound image segmentation based on an early vision model and wavelet analysis. Technical Report NTUBME-MI-027. Taiwan: Institute of Biomedical Engineering, National Taiwan University, 1999.
- Chen IP. Texture perception: a linear system approach. PhD Dissertation. University of California at Berkeley, Berkeley, CA, 1994.
- Cohen LD. On active contour models and balloons. Comput Vision Graphics Image Processing: Image Understanding 1991;53(2):211–218.
- Czerwinski RN, Jones DL, O'Brien WD Jr. Ultrasound speckle reduction by directional median filtering. Proceedings of the 1995 IEEE International Conference on Image Processing 1995;1:358–361.
- Dunn D, Higgins WE, Wakeley J. Texture segmentation using 2-D Gabor elementary functions. IEEE Trans PAMI 1994;16(2):130–149.
- Fan L, Braden GA, Herrington DM. Nonlinear wavelet filter for intracoronary ultrasound images. Proceedings of the 1996 23rd Annual Meeting on Computers in Cardiology 1996:41–44.
- George N, Christensen CR, Bennerr JS, Guenther BD. Speckle noise in displays. J Opt Soc Am 1976;66:1282–1290.
- Heckman T. Searching for contours. Proc SPIE 1996;2666:223–232.
- Jain AK, Farrokhnia F. Unsupervised texture segmentation using Gabor filters. Pattern Recognition 1991;24(12):1167–1186.
- Kass M, Witkin A, Terzoulou D. Snake: Actour contour models. Int J Comput Vision 1987;1:321–331.
- Kotropoulos C. Nonlinear ultrasonic image processing based on signal-adaptive filters and self-organizing neural networks. IEEE Trans Image Processing 1994;3(1):65–77.
- Kozma A, Christensen CR. Effect of speckle on resolution. J Opt Soc Am 1976;66:1257–1260.
- Leymarie F, Levine MD. Tracking deformable objects in the plane using an active contour model. IEEE Trans Pattern Anal Machine Intel 1993;15(6):617–634.
- Lin CB, Chen CM, Su S. A computational early vision model for segmentation of clinical ultrasound images. Proceedings of the 19th Annual International Conference of the IEEE EMBS 1997:605–608.
- Malik J, Perona P. Preattentive texture discrimination with early vision mechanisms. J Opt Soc Am 1990;A7:923–932.
- Muzzolini R, Yang YH, Pierson R. Multiresolution texture segmentation with application to diagnostic ultrasound images. IEEE Trans Med Imaging 1993;12(1):108–123.
- Muzzolini R, Yang YH, Pierson R. Texture characterization using robust statistics. Pattern Recognition 1994;27(1):119–134.
- Muzzolini R, Yang YH, Pierson R. Classifier design with incomplete knowledge. Pattern Recognition 1998;31(4):345–369.
- Pollen DA, Ronner SF. Visual cortical neurons as localized spatial frequency filters. IEEE Trans Syst Man Cyber 1983;13:907–916.
- Tan TN. Texture edge detection by modeling visual cortical channels. Pattern Recognition 1995;28(9):1283–1298.
- Thomas JG, Peters RA, Jeanty P. Automatic segmentation of ultrasound images using morphological operators. IEEE Trans Med Imaging 1991;10(2):180–186.
- Van Hulle MM, Tollenaere T. A modular artificial neural network for texture processing. Neural Networks 1993;6:7–32.
- Williams D, Shah M. A fast algorithm for active contours and curvature estimation. Image Understanding 1992;55(1):14–26.
- Xu C, Prince JL. Snakes, shapes, and gradient vector flow. IEEE Trans Image Processing 1998;7(3):359–369.
- Xu G, Segawa E, Tsuji S. Robust active contours with insensitive parameter. Pattern Recognition 1994;27(7):879–894.
- Yuen PC, Wong YY, Tong CS. Contour detection using enhanced snakes algorithm. Electron Letters 1996;32(3):202–204.
- Zimmer Y, Tepper R, Akselrod S. A two dimensional extension of minimum cross entropy thresholding for segmentation of ultrasound images. Ultrasound Med Biol 1996;22(9):1183–1190.