A Multiparametric and Multiresolution Segmentation Algorithm of 3-D Ultrasonic Data

Djamal Boukerroui, Olivier Basset, Associate Member, IEEE, Atilla Baskurt, Member, IEEE, and Gérard Gimenez, Member, IEEE

Abstract—An algorithm devoted to the segmentation of 3-D ultrasonic data is proposed. The algorithm involves 3-D adaptive clustering based on multiparametric information: the gray-scale intensity of the echographic data, 3-D texture features calculated from the envelope data, and 3-D tissue characterization information calculated from the local frequency spectra of the radio-frequency signals. The segmentation problem is formulated as a Maximum A Posterior (MAP) estimation problem. A multi-resolution implementation of the algorithm is proposed. The approach is tested on simulated data and on in vivo echocardiographic 3-D data. The results presented in the paper illustrate the robustness and the accuracy of the proposed approach for the segmentation of ultrasonic data.

I. INTRODUCTION

THE STUDY of segmentation techniques devoted to 3-D L ultrasonic data presents several applications for medical imaging. The development of a 3-D acquisition scanner addresses the problem of visualization of the 3-D structures embedded in the volume. The segmentation allows extraction of the object of interest before a 3-D representation is constructed. The establishment of a diagnosis is facilitated when an accurate volume measurement of an organ or of a lesion is available, for example, when tracking the evolution of a tumor. Moreover, the shape of a lesion (regular or uneven) can provide valuable information about the benign or malignant status of a specific lesion. Taking into account 3-D data also improves the 2-D segmentation of an image. The continuity of data in the third dimension allows a more robust detection of object boundaries. This is true even when the volume of data is built from the temporal evolution of a dynamic image in a fixed plane (3-D M-mode imaging) and is particularly interesting for cardiac imaging.

In general, image segmentation is based on gray level values. However, ultrasonic images are of relatively poor quality, and segmentation is a difficult problem [1]. The image degradation includes primarily the speckle noise, the blurring of spatial information perpendicular to the propagation direction of ultrasonic waves, and the non-constant

A. Baskurt is with LIGIM, Lyon I University, Villeurbanne, France.

attenuation of ultrasound. When specular structures are imaged, the detected echo amplitude varies according to the orientation of the reflecting structure, and the contours can then appear discontinuous. Moreover, although dynamic focusing techniques are used, the lateral resolution is poor, and the structures are blurred in a direction perpendicular to the ultrasonic propagation.

In most cases, the structures to be detected, such as tumors, have acoustic characteristics similar to the surrounding tissues. Thus, the contrast between the various tissues is poor, which makes the determination of an accurate border difficult. The attenuation of ultrasound depends on the nature of the investigated tissues. Consequently, a homogeneous tissue does not appear quite homogeneous on the image. It may be visualized with a slight variation of intensity in the ultrasonic propagation direction, despite the correction of the time gain compensation, which is constant and independent of the nature of tissues.

The poor quality of conventional ultrasonic images and the slight differences between the various tissues make the automatic segmentation difficult. Techniques based on the thresholding approach or on edge detection from the pixels' intensities are generally not suited for ultrasonic data.

Several approaches have been proposed to segment ultrasonic images. Most of them are based on the pixel intensity and use a Bayesian framework to define an energy function that characterizes a homogeneous region or contour. Ashton and Parker [1] have proposed a modification of the adaptive clustering algorithm given by Pappas [2] to take into account the particularities of ultrasonic data. It is assumed that pixel intensities are given by a slowly varying class mean corrupted globally by uniform additive white Gaussian noise. Spatial smoothness constraints are incorporated in the algorithm by using a Markov random field (MRF) to model the region process. The Central Limit Theorem makes this model reasonably acceptable. It states that the distribution of the mean of a large number of independent random observations tends toward a Gaussian distribution centered on their collective mean. This is the case in a multi-resolution implementation of the algorithm for low-pass filtered and decimated ultrasonic images that are originally governed by non-Gaussian statistics. The segmentation process of this algorithm takes into account only the pixel gray level at a site and its surroundings to classify each pixel. Similar algorithms have been adapted to the segmentation of sequences of echocar-

Manuscript received July 19, 1999; accepted April 6, 2000.

D. Boukerroui, O. Basset, and G. Gimenez are with CREATIS, CNRS Research Unit (UMR 5515) and are affiliated with IN-SERM, INSA 502, 69621, Villeurbanne Cedex, France (e-mail: olivier.basset@creatis.univ-lyon1.fr).

diographic images based on boundary detection [3], [4]. Implementation of these algorithms involves a priori information about heart morphology and about the intensity distribution of the various tissues; it also takes into account the temporal continuity with the previous frames.

Other schemes for segmentation of ultrasonic images have been proposed. Mulet-Parada and Noble [5] suggest an intensity-amplitude invariant approach using a phasedbased feature detection method. Applied to an echocardiographic image sequence, the algorithm takes advantage of the temporal inconsistency of speckle to detect the acoustic boundaries. In [6], a linear combination of the pixel gray level and of the local entropy is used in the Minimum Cross Entropy thresholding technique to segment ultrasonic images containing fluid surrounded by soft tissue. This original method proposes generalization to a multivariate thresholding using several image parameters. However, the determination of the coefficients of the linear combination is still an open problem. Moreover, the method does not take into account spatial information of the segmentation map.

A multi-resolution texture segmentation approach was proposed by Muzzolini *et al.* [7]. Their approach generalized the conventional simulated annealing (SA) methods [8] to a multi-resolution framework and minimized an energy function, which is dependent on the resolution and the size of the texture blocks of the image. SA remains a computationally expensive method of minimization. The use of Besag's iterated conditional mode (ICM) [9], which corresponds to instantaneous freezing in SA for energy minimization, would improve the applicability of this algorithm. However, ICM requires an accurate initial parameters estimation, which is difficult under non-Gaussian statistics.

Because the pixel intensities are embedded in the speckle noise, an alternative for discrimination of the various tissues consists of using another kind of information, such as the measurements performed in tissue characterization experiments. They are generally based on textural measurements made on envelope echographic images or on acoustic measurements preformed on radio frequency (RF) signals. Some authors have proposed performing the segmentation of ultrasonic B-scan data from a texture analysis of the various kinds of tissues [6], [7], [10], [11]. The different constitution (size, distribution, type of reflectors) of the various tissues leads to different textural properties, which can help to distinguish the different structures. Textural features based on co-occurrence matrices have been widely used for the characterization of ultrasonic data. Originally proposed by Haralick [12], these features measure characteristics of the gray level spatial dependencies (second-order statistics). The most frequently used for tissue characterization are the entropy, contrast, correlation and angular second moment (ASM) features [13]-[15]. Nicholas et al. [13] have proposed a systematic approach to define a set of 93 textural features to characterize Bscan images, and the methods and criteria for selecting the optimal combination of the features are discussed. Their

approach has been applied to the discrimination between B-scan textures of normal human livers and spleens. Three of the most discriminating features are provided by the cooccurrence matrix method. The efficiency of this method for the characterization of echographic image texture (uncompressed data) has been studied [15]. It has been concluded that, in cases of diffuse scattering only, the more relevant parameter to differentiate between two textures is ASM. In cases in which there is a structural scattering component as well as diffuse scattering, correlation is the best parameter to detect the periodicity of this structural component.

However, on ultrasonic images, the texture of the various tissues depends on the imaging device. This drawback makes the choice of pertinent textural features that characterize the various tissues system dependent.

Acoustic parameters have also been extensively used for tissue characterization. Most of the applications concern liver and prostate tissues [16], [17]. Contrary to textural features, acoustical features are calculated from the unprocessed RF signals. Information on the attenuation; scattering; elastographic characteristics of tissues; and the size, distribution, and concentration of scatterers can be derived from the backscattered RF signal. These features are often associated with a classification procedure to identify the various tissues [18]–[22].

In this paper, we propose to combine 3-D textural and acoustical parameters in the segmentation process to improve the robustness and the accuracy of the detection of various tissues.

The segmentation of acoustic data needs to be robust to speckle noise, low contrast, and attenuation. The proposed approach favors robustness, taking advantage of 1) a multiresolution implementation. A rough initial segmentation is performed on the low frequency data to initialize the algorithm. Indeed, the aim of the segmentation of ultrasonic data is the definition of a rather large homogeneous area corresponding to the low frequency part of the spectrum. Then, a more rapid convergence and a more accurate segmentation are obtained when details are added at higher resolutions. 2) A multiparametric segmentation process. Information of different types, which may be either redundant or complementary, are involved in the optimization criteria. The fusion of various data increases the robustness of the algorithm. 3) Multidimensional data. The segmentation is performed on volumes. The volume is obtained by either 3-D data or by 2-D +T data. The spatial or temporal continuity in the third direction is used to improve the reliability of the segmentation.

This paper first presents the features involved in the segmentation process and describes the implemented algorithm. Then, segmentation results are illustrated on simulated and in vivo data.

II. Method

A. Parametric Volumes Calculation

1) Textural Features. The co-occurrence matrix method extended to 3-D data has been implemented for texture characterization. The method is based on the estimation of the second-order joint conditional probability density function, $p_{d\theta}(i, j)$. Each $p_{d\theta}(i, j)$ is the probability of going from a gray level, i, to a gray level, j, in a given direction, θ , at a given intersample spacing, d. In practice, the parameters d and θ are converted into the distances dx, dy, dz, which correspond to an integral number of pixels. The co-occurrence matrix $p_{dx,dy,dz}$ is a representation of the estimated values. It is a square matrix of dimension Ng (Ng is the number of gray level in the volume). Formally, the non-normalized matrix entries of a volume, V, are defined by $C_{dx,dy,dz}(i,j) =$ $\operatorname{card}\{(x, y, z), (x', y', z') \in V \times V / f(x, y, z) = j, f(x', y', z')$ =i, x-x'=dx, y-y'=dy, z-z'=dz, where card denotes the number of elements in the set, and f(x,y,z) gives the gray level of pixels (x,y,z).

In our application, as in [13], we chose to ignore the orientation dependency. Thus, the matrices were generated by pooling all of the frequencies calculated for a distance d = 1. Note that the intensity range is reduced to 64 gray levels for the matrices calculation.

To summarize the content of a co-occurrence matrix, a number of textural features have been proposed [12]. To build each parametric volume, a local estimation of each textural feature is performed for each voxel. It requires the calculation of a local co-occurrence matrix on a small volume centered on the considered voxel. The results presented in this study were obtained using the features: entropy (ENT), contrast (CON), correlation (COR), angular second moment (ASM), and sum average (SAV).

2) Acoustical Features. Two acoustical parameters are calculated from the RF signals. They are the mean central frequency (MCF) and the integrated backscatter (IBS). These parameters are often used for acoustical characterization. As these features depend on the structure of tissues, they can constitute a signature of the various tissues or of the pathological state. The MCF is related to the attenuation of the medium because of the dependence between the attenuation and the frequency. The attenuation increases with frequency and consequently modifies the frequency spectrum of a propagating acoustic wave. The IBS is an estimation of the backscattered energy, so it contains information about the number and the structure of the scatterers in the medium.

These parameters can be easily estimated through a short-time Fourier analysis [20]. The estimation of the local power spectrum is performed by the Fast Fourier Transform (FFT) on a 64-point temporal window zeros padded to 512. A Hamming window is used to achieve local stationarity of the signal and to reduce the Gibbs phenomena. The MCF is calculated as the first spectral moment inside a -20 dB bandwidth and the IBS as the total power of the received signal in the reduced -20 dB bandwidth. For

each voxel of the parametric volume, the local estimation of the two features is performed from a window (64 points) of the RF signal, centered on the corresponding voxel.

B. Segmentation Method

1) Modeling. The following section describes the segmentation process considering, at first, a segmentation performed using a single parameter (for example, the envelope intensity volume).

We assume that the observed data Y is a random field defined on a 3-D rectangular grid S. Y_s denotes the value of Y at the site $s \in S$. A segmentation of the volume into regions will be noted by X, where $X_s = i$ means that the pixel at s belongs to region i. The number of different regions in X is k. Using Bayes' theorem, the a posteriori probability density function P(X = x|Y = y) has the form:

$$P(X = x | Y = y) \propto P(Y = y | X = x) P(X = x).$$
(1)

The conditional density function of Y given X is assumed to exist and to be strictly positive and is denoted by P(y|x). The probability P(X = x) is written as P(x). The volume may be segmented by estimating the voxel classification X given the observed volume Y using the MAP estimation of X expressed by

$$\hat{x}_{MAP} = \arg \max \{ P(X = x | Y = y) \} = x$$
arg min $\{ -\ln P(y|x) - \ln P(x) \}.$
(2)

Hence, once the distributions of P(y|x) and P(x) are defined, the problem of segmenting a volume will be reduced to that of minimizing an energy function.

To model the regional process X, we use an MRF because of its restriction to local interaction. So, according to the Hammersley-Clifford theorem [23], and for a given neighborhood system, the prior density P(x) can be written as a Gibbs density, which has the following form:

$$P(x) = \frac{1}{Z} \exp\left\{-\sum_{all \ cliques \ C} V_c(x)\right\}.$$
 (3)

Here, Z is a normalizing constant called the partition function. $V_c(x)$ are the clique potentials. A clique c is a subset of sites ($c \in S$) that are neighbors of each other. In this work, the first-order neighborhood system, with respect to the Euclidean distance, was used (Fig. 1), which considers only the two-site clique potentials defined as follows:

$$V_c(x) = \begin{cases} -\beta, \text{if } x_s = x_q \text{ and } s, q \in c \\ +\beta, \text{if } x_s \neq x_q \text{ and } s, q \in c \end{cases}, \quad \beta > 0.$$
(4)

The Gibbsian parameter β is positive so that two neighboring pixels are more likely to belong to the same class



Fig. 1. First-order 3-D neighborhood system with respect to the Euclidean distance (6-connexity).

than to different classes. Increasing β value increases regional size and leads to excessive smoothing of boundaries.

The conditional density distribution $P_s(y_s|x_s = i)$ of the observed gray intensity at a site s is assumed to be Gaussian with mean μ_s^i and variance $(\sigma_s^i)^2$. The local class mean μ_s^i is a slowly varying function of s. $(\sigma_s^i)^2$ is estimated independently for each class and is proportional to μ_s^i . Under these assumptions, $\ln P(y|x)$ may be written as

$$\ln P(y|x) \propto \sum_{s} \ln P_{s}(y_{s}|x_{s}) = -\sum_{s} \left(\ln (\sigma_{s}^{x_{s}}) + \frac{1}{2(\sigma_{s}^{x_{s}})^{2}} (y_{s} - \mu_{s}^{x_{s}})^{2} \right) .$$
(5)

Substituting P(x) from (3) and $\ln P(y|x)$ from (5) into (2) leads to the following energy function:

$$U(x|y) = \sum_{s} \left(\ln(\sigma_s^{x_s}) + \frac{1}{2(\sigma_s^{x_s})^2} (y_s - \mu_s^{x_s})^2 \right) + \sum_{all \ cliques \ C} V_c(x). \quad (6)$$

This function has two components. The first term constrains the regional intensity to be close to the data, and the second is a regularization term that imposes a smoothness constraint.

To improve the robustness of the algorithm, the energy function can be modified by adding other constraints based on parametric measurements that are representative of each region (Fig. 2). Let $\{y_1, ..., y_n\}$ be a set of features calculated on each site of the volume of data.

Gray-scale parametric volumes are modeled in the same way as gray-scale images in [2]. A parametric volume is assumed to be a collection of uniform or slowly varying intensities. The sharp transitions in gray levels may only occur at a region boundary. The feature value Y_i at a voxel location s is denoted $(y_j)_s$. We assume that the values given by a feature Y_j at the site s are modeled by a normal distribution of mean $(m_j)_s^{x_s}$ and variance $(\sigma_j^{x_s})^2$. Then, the complete energy function has the following form:

$$U(x|y, y_1...y_n) = \sum_{s} \left(\ln(\sigma_s^{x_s}) + \frac{1}{2(\sigma_s^{x_s})^2} (y_s - \mu_s^{x_s})^2 \right) + \sum_{all \ cliques \ C} V_c(x) + \sum_{j} \sum_{s} \left(\ln(\sigma_j^{x_s}) + \frac{1}{2(\sigma_j^{x_s})^2} \left((y_j)_s - (m_j)_s^{x_s} \right)^2 \right).$$
(7)

Finding the global minimum of this function requires a intensive computation. As an alternative to SA, the ICM algorithm has been used. Starting from an initial segmentation x_0 , the algorithm updates the label of sites in x to maximize the conditional density function at each site, knowing the label values at its neighborhood and the observation y. The algorithm converges to a local minimum, which is a reasonably acceptable solution under a good initialization. This algorithm was implemented in [1], [2], [24].

2) Algorithm. In the modeling section, the conditional density distribution of the observed gray intensity at a site s is assumed to be Gaussian. Clearly, this assumption is not true in many cases, especially for ultrasonic data. However, the Central Limit Theorem makes it reasonably acceptable for low-pass filtered images. A multi-resolution implementation allows satisfaction of the hypothesis of Gaussian distribution.

Hence, the hierarchical structure proposed in [25] was used to implement our algorithm (Fig. 3). Starting from the highest resolution ultrasonic volume, a multiresolution pyramid is built using the discrete wavelet transform (DWT) approach [26]. Textural volumes at each resolution can be calculated in two ways: either by calculating the texture volumes only at the highest resolution and building the pyramid from the high resolution texture volumes or by calculating the texture volumes at each level of the gray level pyramid. The second approach is more interesting because textures are analyzed at each level. This allows a multi-scale analysis of textures from macro-texture to micro-texture. However, at low resolution levels, the texture information is very poor due to the successive low-pass filtering. In our implementation, the texture features at the two highest resolutions, 0 and -1, are calculated from the corresponding gray intensity data, and then the texture pyramids are completed by the DWT starting from the texture volumes at resolution -1 (Fig. 3).

The coarsest resolution is initially segmented, and the segmentation result is passed on to the immediate higher resolution and so on until the finest resolution is segmented. An initial solution of the minimization problem, at the coarsest resolution, is obtained with the K-means algorithm. Starting from this segmentation, the algorithm alternates between the estimation of region labels and



Fig. 2. Schematic representation of the data involved in the segmentation process.



Fig. 3. Decomposition scheme of the different volumes of data.

model parameters and is stopped when no further changes in the labels occur. A rectangular window is used for the estimation of the local class means and variances for all region types *i* and all pixels *s*. Given the region labels *x*, μ_s^i and σ_s^i are the average and the standard deviation of pixels of the region *i* inside the window of width *W* centered at *s*, respectively. Initially, for robust estimation of the model parameters, the window size *W* is equal to the whole volume, and, then, as the algorithm progresses, the segmentation becomes better, and smaller windows give more reliable and accurate estimations. Thus, the algorithm fits progressively to local characteristics of each region. In our implementation, as in [2], the window size is reduced by a factor of two until a final value W_{min} .

A major difficulty of this algorithm is the ad hoc choice of the Gibbsian parameter β and the number k of regions in the data. The authors in [1], [2] show experimentally on various kinds of images (natural scenes, ultrasonic cardiac images) that good results are obtained with k = 4. The problem of the estimation of β is studied in [25] in the special case of the segmentation of brain MRI images. The authors propose a prior model for the estimation problem. Some studies [2], [24] propose a constant value for parameter β , whereas Ashton and Parker [1] suggest adapting the



Fig. 4. Simulated ultrasonic image. Envelope (a) and IBS (b) images after logarithmic compression; MCF image (c). The histograms of the three images are given in (d).

TABLE I AVERAGE PERCENTAGE OF CORRECT PIXEL CLASSIFICATION OF SEGMENTATION RESULTS OBTAINED USING ENVELOPE DATA (A), ENVELOPE DATA WITH TEXTURE CHARACTERIZATION (B),

Envelope and IBS Images (c), and Envelope and IBS Images with Texture Characterization (d).

	а	b	с	d
Phantom 1	70.41%	87.81%	82.76%	89.10%
Phantom 2	96.08%	96.07%	95.90%	95.97%
Total	83.24%	91.94%	89.33%	92.53%

value of β to each resolution (β increases with resolution). In [27], the authors propose a mathematical approach to derive the Gibbsian parameter at each scale directly from the full resolution one. This method shows clearly that the Gibbsian parameter has a small value at the coarsest resolution and increases, but not linearly, within the resolution level until the final resolution. In our implementation, β is increased by $\Delta\beta > 0$ at each change of resolution level.

III. RESULTS

In this section, experimental results of the segmentation are presented on both 2-D simulated data and in vivo 2-D +T cardiac images. Real ultrasound images were used to evaluate visually the ability of the algorithm to segment echographic data. Synthesized images provide a controlled environment that allow the quantification of the performance of the proposed algorithm because a reference image is available. Generally, in our experiments, we take $W_{min} = 7$, the Gibbsian parameters $\beta = 0.5$ and $\Delta\beta = 0.2$, the number of decompositions n = 2 to 3, and the number of classes k = 2 to 4.

A. Simulated Data

1) Experiment 1. In this experiment, the 2-D simulated data are obtained using the approach proposed in [28]. A thick cylinder is generated with backscattering properties that differ from its surroundings. The resulting envelope image, after logarithmic compression of the form $D\ln(\cdot)$, with D = 56, is composed of three regions (Fig. 4) of different means and standard deviations: region A, $\mu_A = 160$,



Fig. 5. Segmentation results of the data presented in Fig. 4 (k = 3, n = 2, $\beta = 0.5$, $\Delta\beta = 0.2$) using the envelope image (without tissue characterization) (a); using envelope and two textural parametric images (ENT and ASM calculated from the co-occurrence matrices) (b); using envelope image, textural, and acoustical characterization, (ENT, SAV, MCF, and IBS) for an appropriate number of classes k = 3 (c), and for an overestimated number of classes in the segmentation process k = 4 (d).

 $\sigma_A = 33$; region B, $\mu_B = 172$, $\sigma_B = 33$; and region C, $\mu_C = 116, \sigma_C = 30$. The three regions, before the log compression, are characterized by a Rayleigh distribution. We can note that the standard deviations of the three regions are very close. A theoretical formulation of the log compression statistics can be found in [29], [30]. It has been shown that the standard deviation of the log-compressed Rayleigh data depends only on the dynamic range parameter D, $\sigma = \frac{\pi}{\sqrt{24}}D$. In our case, the predicted standard deviation of the three regions is \approx 36. Fig. 4(a) shows the logarithmic envelope image obtained with the Hilbert transform and decimated in the ultrasound propagation direction to obtain isotropic data. Fig. 4(b) and (c) show the corresponding IBS and MCF images. Their normalized histograms are presented in Fig. 4(d). Clearly, the IBS image has a better contrast than the envelope image. Moreover, we can notice that, contrary to the envelope image histogram, the IBS histogram is bimodal.

The segmentation results of the simulated data are presented in Fig. 5. They are obtained with different combinations of the parametric images: a) using the envelope image only (without tissue characterization); b) using envelope and two textural parametric images (ENT and ASM calculated from the co-occurrence matrices); and c) and d) using envelope image, textural (ENT and SAV), and acoustical characterization (MCF and IBS).

It can be observed that texture characterization improves the accuracy of the segmentation, but some artifacts are still present at the borders of the image. These artifacts are eliminated when acoustical features are involved in the segmentation process. The borders between the regions, and particularly between regions A and B, are more accurately delineated. The segmentation results obtained when an overestimated number of classes is imposed are correct [Fig. 5(d)]. This is an illustration of the robustness of the algorithm.

The simulated data have been used to evaluate the interest of a multi-resolution and multi-parametric implementation of the segmentation algorithm. On a selected area of the synthesized image (a square area including the ring of the region B), the percentage of pixels classified in the correct region by the segmentation has been cal-



Fig. 6. Percentage of correctly classified pixels when only one kind of parametric data is used and when all of the features are involved in the segmentation process.

culated when only one kind of parametric data is used in the segmentation and when all of the data (envelope, textural, and acoustic data) are involved. The performance of the segmentation with various features has been calculated at different resolutions. Four values of correctly classified pixels are reported in Fig. 6. They correspond to the initial segmentation performed with the K-means and to the final segmentations (when the minimal value of the energy function is reached) at a resolution of -2 (2) decimations), -1 (one decimation), and at the full resolution. These experiments show that the correct classification rate is always larger when all of the features are used than when only one feature is considered. One can remark on this figure (Fig. 6) that the initial solution of the multi-parametric experiment is much better than the initial solution obtained with only one feature. As mentioned in the modeling section, a good initial solution is of major importance when a deterministic algorithm in the minimization process is used.

The acoustic features and the envelope, used individually, exhibit approximately the same rates. The textural feature SAV presents a lower rate of pixels correctly segmented. This can be explained by the inaccurate location of the border of the different regions with textural data because the parametric textural data are calculated on a rather large window (15×15 pixels). Using a large window to assess local information (at each site) blurs the tissue borders. When the feature ENT is used alone, only about 35% of the pixels are correctly classified in the three regions. This is due to the inability of this feature to discriminate the two regions A and C. However, this feature remains relevant in the global segmentation for identifying region B and separating it from the two other classes.

Globally, the segmentation is improved as the resolution increases. Moreover, the multi-resolution implementation of the algorithm decreases the processing time. Typically, the time required for the segmentation of the simulated image is about 6 s with a multi-resolution implementation, whereas 40 s are necessary to obtain a similar result without multi-resolution. 2) Experiment 2. The previous experiment shows the interest of a multi-parametric implementation of the algorithm. Other experiments have been carried out to verify these observations and to evaluate the performance of the proposed method on several images.

Two computer phantoms of $30 \times 30 \times 5$ mm are used. The synthesized images are generated using the FIELD program for the simulation of ultrasound imaging [31], [32]. The simulation approach is based on the calculation of spatial ultrasound field in a homogeneous attenuating medium and assumes a linear propagation. Any transducer and focussing schemes, including the transducer excitation, can be simulated with this approach. Therefore, the software produces realistic B-mode images.

The images are generated by specifying a number of independent scatters in a file where their position and amplitude are defined. Adjusting the number of scatters and their relative amplitude yields the proper image. In our experiments, a linear scan is done at 7.5 MHz with a 192-element transducer, using 64 active elements and the received RF signals sampled at 100 MHz. The scatters are positioned randomly in the phantoms. For the two phantoms presented hereafter, five images of 128 lines are generated with different attenuation of the media. Typically, one simulation requires about 13 h of simulation time on a 500-MHz PC. Fig. 7 shows the envelope images obtained without attenuation corresponding to the two phantoms.

In this experiment, the MCF parameter was not used in the segmentation process because the simulation model does not allow the simulation of several attenuating media in a single field. Thus, MFC image calculated on the RF signals cannot distinguish the different tissues of the phantom. All other control parameters are fixed at the same values as in Experiment 1.

Fig. 8 shows an example of segmentations of the image presented in Fig. 7(a), which were obtained with different combinations of the parametric images: a) using the envelope image only (without tissue characterization), b) using envelope and two textural parametric images (ENT and ASM calculated from the co-occurrence matrices), c) using the envelope and the IBS image, or d) using envelope image, textural (ENT and SAV) images, and the IBS image.

Fig. 8(a) shows that when the envelope image is used alone, the algorithm is more sensitive to the speckle pattern. This leads to several misclassified regions in the background tissue. The classification error is reduced when other parametric images are introduced in the segmentation process. Texture characterization improves the segmentation results globally [Fig. 8(c)]. However, as noticed in Experiment 1, some artifacts are still present because of the low frequency of this kind of information. One can remark, in this example, a small region at the border of the inner tissue that is classified as a tissue of the background. Because the textural information is an average measurement on a window, such artifacts can occur at the border of two regions. In that case, the window used to calculate the textural feature is averaging an hypoechoic



Fig. 7. Simulated ultrasonic images using FIELD II program. Phantom 1 (a); phantom 2 (b) (without attenuation).



Fig. 8. Segmentation results of phantom 1 (k = 3, n = 2, $\beta = 0.5$, $\Delta\beta = 0.2$) using the envelope [Fig. 7(a)] (a); using envelope and two textural parametric images (ENT and ASM calculated from the co-occurrence matrices) (b); using envelope and IBS images (c); and using envelope image, textural, and acoustical characterization (ENT, SAV, and IBS) (d). The real contours are superimposed on the segmented images.



Fig. 9. Envelope slice after logarithmic compression calculated from the RF data.



region (inner) and an hyperechoic region (ring). The average value obtained is close to the characteristics of the outer region. It can be observed from the inspection of the four segmentations, that when the texture characterization is involved in the segmentation process, the detected boundaries are less accurately located. This shift is more important for the boundaries parallel to the ultrasound propagation direction because of their fuzziness aspect in ultrasound B-mode images.

For the two phantoms and for the four previous combinations of the parametric images, the average percentage of correctly classified pixels in the segmentation results is given in Table I. The segmentation accuracy of phantom 1 with the multi-parametric algorithm is substantially better than the segmentation using the envelope image alone. However, the segmentation accuracy of the second phantom is approximately constant for the four parameter combinations (around $96 \pm 0.08\%$). This result was expected because the second phantom is relatively easy to segment and most of the misclassified pixels are localized at the boundaries parallel to the ultrasound beam.

In summary, taking into account texture and acoustical characterization in the segmentation process improves the accuracy and the robustness of the segmentation algorithm. The improvement is the result of the complementarity of the parametric measurements. The texture characterization leads to a robust global segmentation but can produce some artifacts, and the contours are not accurately localized. The acoustic characterization involves images with a higher resolution than textural images and contributes to the reduction the artifacts and the improvement of the boundary location.

B. In Vivo Data

The segmentation algorithm has been tested on cardiac RF images acquired from a dog's heart at the University





Fig. 10. Parametric images calculated from the RF data: IBS image after logarithmic compression (a), MCF image (b), and derived from the envelope data (ENT image calculated on the 3-D co-occurrence matrices) (c).



Fig. 11. Segmentation results of the slice shown in Fig. 9 (k = 3, n = 3, $\beta = 0.5$, $\Delta\beta = 0.2$) with different combinations of features: envelope volume only (a); envelope volume and two textural features ENT and SAV (b); envelope volume, textural features ENT, SAV, and acoustic features MCF and IBS (c).

of Leuven with a dedicated system [33]. A temporal sequence of 50 sector images of 120 scan lines was acquired at 5 MHz and sampled at 40 MHz, 8-bit. The frame rate is 30 images/s. From the sequence of images, we construct a 3-D volume (dimensions: $352 \times 265 \times 99$). Fig. 9 shows an envelope slice after logarithmic compression and sectorial reconstruction. In Fig. 10, we present the IBS and the MCF images calculated from the corresponding RF data of the envelope images given in Fig. 9, as well as the "Entropy" slice calculated from the 3-D co-occurrence matrix of the envelope volume. We notice, similar to the simulated data, that the IBS image presents a better contrast than the envelope image.

The segmentation results of the slice shown in Fig. 9 are illustrated in Fig. 11. Three combinations of the different feature classes involved in the segmentation process are presented. All other control parameters of the algorithm are constant (three classes k = 3, three levels of decomposition n = 3, and the Gibbsian parameters are fixed to $\beta = 0.5$ and $\Delta\beta = 0.2$, where $\Delta\beta$ is the increment of β at each resolution).

The results show that, when acoustical parameters are included in the segmentation process [Fig. 11(c)], the contours are more accurately determined than when only texture characterization is used [Fig. 11(b)]. This is particularly true for the definition of the ventricular cavity. The algorithm is also efficient for segmentation of the heart wall. The outer contour is almost continuous. Notice that the valve appears clearly on the segmented image. On Fig. 12, the contours of the segmented regions are superimposed on the original envelope slices, which allows visual evaluation of the segmentation. The time necessary for the processing of a single image (352×265 pixels) on a digital personal workstation (433MHz, RAM: 128 Mo) is typically 6 s and increases up to 2 h for the 3-D processing of each volume of data.

Fig. 13 shows the evolution of the location of the inner left ventricle border versus time at the two axial points indicated by arrows in Fig. 11(c). The sequence of images corresponds to more than one heartbeat cycle (about 1.5 s), and the wall motion can be observed. The relative positions of the curves indicate the diastole and systole. We can notice the temporal continuity of the two curves caused by the 3-D processing of the data.

IV. CONCLUSION

When ultrasonic data are involved, a segmentation based only on intensity information is a difficult task because of the speckle noise. In this work, both acoustic and texture characterizations are introduced in the segmentation process of ultrasonic data. The algorithm is applied to simulated ultrasonic data and cardiac RF images. The various tissues are outlined more accurately when textural and acoustical features are combined. These features reinforce and complete the information provided by standard echographic images. For example, the algorithm profits from the contrast enhancement given by the IBS im-



(a4)

(c4)

Fig. 12. The region contours detected are superimposed on the original envelope slices at four instants of the cardiac cycle: (a) segmentation without tissue characterization, (b) segmentation with texture characterization, and (c) segmentation with texture and acoustical characterization.

(b4)



Fig. 13. Evolution of the location of the inner border of the left ventricle versus time at the two axial points shown in Fig. 11(c). The vertical axis indicates a number of pixels.

ages. The use of different kinds of data, which are complementary and/or redundant, and the fusion of these data in the algorithm lead to a more robust segmentation.

Acknowledgments

We gratefully acknowledge Bart Bijnens and Jan D'hooge (Medical Image Computing, University of Leuven) for the acquisition of the cardiac RF data.

References

- E. A. Ashton and K. J. Parker, "Multiple resolution Bayesian segmentation of ultrasound images," *Ultrason. Imag.*, vol. 17, pp. 291–304, 1995.
- [2] T. N. Pappas, "An adaptive clustering algorithm for image segmentation," *IEEE Trans. Signal Proc.*, vol. 40, no. 4, pp. 901– 914, 1992.
- [3] N. Friedland and D. Adam, "Automatic ventribular cavity boundary detection from sequential ultrasound images using simulated annealing," *IEEE Trans. Med. Imag.*, vol. 8, no. 4, pp. 344–353, 1989.
- [4] J.M.B. Dias and J.M.N. Leitao, "Wall position and thickness estimation from sequences of echocardiographic images," *IEEE Trans. Med. Imag.*, vol. 15, no. 1, pp. 25–38, 1996.
- [5] M. Mulet-Parada and J. A. Noble, "2D+T boundary detection in echocardiography," Med. Image Anal., vol. 4, pp. 21–30, 2000.
- [6] Y. Zimmer, R. Tepper, and S. Akselrod, "A two-dimensional extension of minimum cross entropy thresholding for the segmentation of ultrasonic images," *Ultrasound Med. Biol.*, vol. 22, no. 9, pp. 1183–1190, 1996.
- [7] R. Muzzolini, Y.-H. Yang, and R. Pierson, "Multiresolution texture segmentation with applicaton to diagnostic ultrasound images," *IEEE Trans. Med. Imag.*, vol. 12, no. 1, pp. 108–123, 1993.
- [8] S. Kirkpatrick, C. Gelatt, and M. Vecchi, "Optimization by simulated annealing," *Science*, vol. 220, pp. 671–680, 1983.
- [9] J. Besag, "On the statistical analysis of dirty pictures," J. R. Stat. Soc. B, vol. 48, no. 3, pp. 259–302, 1986.

- [10] D. Boukerroui, O. Basset, A. Baskurt, A. Hernandez, N. Guérin, and G. Gimenez, "Texture based adaptive clustering algorithm for 3D breast lesion segmentation," in *Proc. IEEE Ultrason. Symp.*, Toronto, Canada, 1997, pp. 1389–1392.
- [11] D. Boukerroui, O. Basset, N. Guérin, and A. Baskurt, "Multiresolution texture based adaptive clustering algorithm for breast lesion segmentation," *Eur. J. Ultrasound*, vol. 8, no. 2, pp. 135– 144, 1998.
- [12] R. M. Haralik, K. Shanmugam, and I. Dinstein, "Textural features for image classification," *IEEE Trans. Syst.*, Man, Cybern., vol. 3, no. 6, pp. 610–621, 1973.
- [13] D. Nicholas, D. K. Nassiri, P. Garbutt, and C. R. Hill, "Tissue characterization from ultrasound B-scan data," *Ultrasound Med. Biol.*, vol. 12, no. 2, pp. 135–143, 1986.
- [14] O. Basset, Z. Sun, J. L. Mestas, and G. Gimenez, "Texture analysis of ultrasonic images of the prostate by means of cooccurrence matrices," *Ultrason. Imag.*, vol. 15, no. 3, pp. 218– 237, 1993.
- [15] F.M.J. Valckx and J. M. Thijssen, "Characterization of echographic image texture by cooccurrence matrix parameters," *Ultrasound Med. Biol.*, vol. 23, no. 4, pp. 559–571, 1997.
- [16] F. L. Lizzi, E. J. Feleppa, M. Astor, and A. Kalisz, "Statistics of ultrasonic spectral parameters for prostate and liver examinations," *IEEE Trans. Ultrason., Ferroelect., Freq. Contr.*, vol. 44, no. 4, pp. 935–942, 1997.
- [17] G. Schmitz, H. Ermert, and T. Senge, "Tissue-characterization of the prostate using Radio Frequency ultrasonic signals," *IEEE Trans. Ultrason., Ferroelect., Freq. Contr.*, vol. 46, no. 1, pp. 126–138, 1998.
- [18] A. Lorenz, M. Blüm, H. Ermert, and T. Seng, "Comparison of different Neuro-Fuzzy classification systems for detection of prostate cancer in ultrasonic images," in *Proc. IEEE Ultrason. Symp.*, Toronto, Canada, 1997, pp. 1201–1204.
- [19] J. M. Thijssen, B. J. Oosterveld, P. C. Hartman, and G.J.E. Rosenbusch, "Correlations between acoustic and texture parameters from RF and B-mode liver echograms," *Ultrasound Med. Biol.*, vol. 19, no. 1, pp. 13–20, 1993.
- [20] H. Rijsterborgh, F. Mastik, C. T. Lancee, P. Verdouw, J. Roelandt, and N. Bom, "Ultrasound myocardial integrated backscatter signal processing: Frequency domain versus time domain," *Ultrasound Med. Biol.*, vol. 19, no. 3, pp. 211–219, 1993.
- [21] K. A. Wear, R. F. Wagner, and B. S. Garra, "A comparison of autoregressive spectral estimation algorithms and order determination methods in ultrasonic tissue characterization," *IEEE Trans. Ultrason., Ferroelect., Freq. Contr.*, vol. 42, no. 4, pp. 709–716, 1995.
- [22] G. Georgiou and F. Cohen, "Unsupervised segmentation of RF echo into regions with different scattering characteristics," *IEEE Trans. Ultrason., Ferroelect., Freq. Contr.*, vol. 45, no. 3, pp. 779–787, 1998.
- [23] B. J. Besag, "Spatial interaction and the statistical analysis of lattice systems," J. R. Stat. Soc. B, vol. 26, no. 2, pp. 192–236, 1974.
- [24] C. Bouman and B. Liu, "Multiple resolution segmentation of textured images," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 13, no. 2, pp. 99–113, 1991.
- [25] B. C. Vemuri, S. Rahman, and J. Li, "Multiresolution adaptive k-means algorithm for Segmentation of brain MRI," in *Proc. Int. Comp. Sci. Conf. Image Analysis Computer Graphics*, Hong Kong, 1995, pp. 347–354.
- [26] S. G. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 11, no. 7, pp. 674–693, 1989.
- [27] F. Heitz, P. Perez, and P. Bouthemy, "Multiscale minimization of global energy functions in some visual recovery problems," *CVGIP: Image Understanding*, vol. 59, no. 1, pp. 125– 134, 1934.
- [28] J. Meunier and M. Bertrand, "Echographic image mean gray level changes with tissue dynamics: A system-based model study," *IEEE Trans. Biomed. Eng.*, vol. 42, no. 4, pp. 403–410, 1995.
- [29] D. Kaplan and Q. Ma, "On the statistical characteristics of logcompressed Rayleigh signals: Theoretical formulation and experimental results," J. Acoust. Soc. Amer., vol. 95, no. 3, pp. 1396–1400, 1994.

- [30] V. Dutt and J. Greenleaf, "Statistics of the log-compressed echo envelope," J. Acoust. Soc. Amer., vol. 99, no. 6, pp. 3817–3825, 1996.
- [31] J. A. Jensen, "Field: A program for simulating ultrasound systems," in Proc. Med. Biol. Eng. Comp., 10th Nordic-Baltic Conf. Biomed. Imag., pt. 1, pp. 351–353, 1996, Available at: http://www.it.dtu.dk/~jaj/field/papers.html.
- [32] J. A. Jensen and P. Munk, "Computer phantoms for simulation ultrasound B-mode and CFM images," in Proc. 23rd Acoust. Imag. Symp., Boston, MA, 1997, Available at: http://www.it.dtu.dk/~jaj/field/papers.html.
- [33] B. Bijnens, M.-C. Herregods, J. Nuyts, G. Vandeweghe, P. Suetens, and F.V.D. Werf, "Acquisition and processing of the Radio-Frequency signal in echocardiography: A new global approach," *Ultrasound Med. Biol.*, vol. 20, no. 2, pp. 167–176, 1994.



Atilla Baskurt (M'98) was born in Ankara, Turkey in 1960. He received the B.S. degree in 1984, the M.S. in 1985, and the Ph.D. in 1989, all in electrical engineering from INSA of Lyon, France. From 1989 to 1998, he was "Matre de Conférences" at INSA of Lyon. Since 1998, he has been with the University Claude Bernard of Lyon, where he is a professor in electrical and computer engineering. His research interests include digital signal and image processing, image compression, and image segmentation, especially for biomedical

and multimedia applications. He is a member of IEEE and EURASIP.



Djamal Boukerroui was born in Bejaa, Algeria in 1972. He received the B.S. degree in 1990 and the M.S. degree in electronics in July 1995 from ENP of Algiers (Algeria). He received the Ph.D. degree in image processing at CREATIS Laboratory of INSA in Lyon (France) in 2000. Since March 2000, he has been with the Department of Engineering Science of the University of Oxford as a postdoctoral research assistant. His research interests include image segmentation, ultrasound imaging, markov random field, multiresolution ap-

proaches, and texture characterization.



Olivier Basset (M'96–A'96) was born in Bourg en Bresse (France) in 1963. He received, in 1991, the Ph.D. degree in acoustics from the National Institute of Applied Sciences (INSA) in Lyon, France. He is currently Associate Professor in the University Claude Bernard-Lyon I.

He is a member of CREATIS [Research Unit associated with CNRS (#5515) and affiliated with INSERM]. His research interests concern ultrasonic image processing, segmentation, data fusion, and texture analysis. Dr.

Basset is an associate member of IEEE.

Gérard Gimenez (M'92) received the M.S. degree from the University Claude Bernard (Lyon, 1970), the Ph.D. degree from the University of Provence (Marseille, 1972), and the Doctorate of Science degree from the University of Orsay (Paris XI, 1978).

He is currently Professor at the National Institute of Applied Sciences (INSA Lyon) and Director of CREATIS, a research unit associated with CNRS, which gathers engineers, scientists, and medical doctors working together on medical imaging. He is also head of the Doctoral Studies Center in electrical engineering for the Lyon area.

His research interests are in the interaction between ultrasonic interrogating waves and investigated media. He also promotes transfer of knowledge in image processing and computer engineering toward medical doctors dealing with specific pathologies.