Segmentation in Ultrasonic *B*-Mode Images of Healthy Carotid Arteries Using Mixtures of Nakagami Distributions and Stochastic Optimization

François Destrempes, Jean Meunier, Marie-France Giroux, Gilles Soulez, and Guy Cloutier*

***Abstract*—The goal of this work is to perform a segmentation of the intimamedia thickness (IMT) of carotid arteries in view of computing various dynamical properties of that tissue, such as the elasticity distribution (elastogram). The echogenicity of a region of interest comprising the intima-media layers, the lumen, and the adventitia in an ultrasonic *B*-mode image is modeled by a mixture of three Nakagami distributions. In a first step, we compute the maximum a posteriori estimator of the proposed model, using the expectation maximization (EM) algorithm. We then compute the optimal segmentation based on the estimated distributions as well as a statistical prior for disease-free IMT using a variant of the exploration/selection (ES) algorithm. Convergence of the ES algorithm to the optimal solution is assured asymptotically and is independent of the initial solution. In particular, our method is well suited to a semi-automatic context that requires minimal manual initialization. Tests of the proposed method on 30 sequences of ultrasonic *B*-mode images of presumably disease-free control subjects are reported. They suggest that the semi-automatic segmentations obtained by the proposed method are within the variability of the manual segmentations of two experts.**

***Index Terms*—*B*-mode, Bayesian model, carotid artery, expectation maximization (EM) algorithm, exploration selection algorithm, mixtures of gamma distributions, mixtures of Nakagami distributions, segmentation, stochastic optimization, ultrasound.**

I. Introduction

THE intima-media thickness (IMT) is a double-line pattern visualized by echography on both walls of the common carotid arteries in a longitudinal image. It is formed by two parallel lines, which consist of the leading edges of two anatomical boundaries: the lumen-intima and media-adventitia interfaces [1]. Semi-automatic segmentation of the IMT on the far wall of carotid arteries in *B*-mode ultrasound images is a useful tool for clinical applications. For instance, one can use such segmentations as a preprocessing step in order to compute various dynamical properties of that anatomical region, such as the elastograms [2]. Rayleigh distributions have been used to model the local brightness of the speckle pattern in a *B*-mode image [3]–[9]. Rayleigh distributions correspond to the case of a high density of independent random scatterers within the range cell [10]. In [3], the image pixels are modeled by three Rayleigh distributions with means depending on their positions (inside the endocardium, between the endocardium and the pericardium, and outside the pericardium). In [4], the parameters of the Rayleigh distributions corresponding to two regions of the cross section of a coronary artery are estimated iteratively using the maximum likelihood estimator (MLE). Similarly, in [8], the parameters of the Rayleigh distributions corresponding to the blood and the arterial wall are estimated independently on each region with the MLE. In [5], the mixture parameters are preestimated using the iterative conditional estimation (ICE) procedure, upon considering a model with likelihood based on two Rayleigh distributions corresponding to the blood and the cardiac muscle. In [6], [7], and [9], the parameters are preestimated using the expectation-maximization (EM) algorithm [11], upon considering a model with no Markov prior, so that the intensities are viewed as independent variables.

Other distributions have been proposed for the local brightness of the speckle pattern in *B*-mode images. When unresolved coherent components (i.e., the spacing is smaller than the speckle size) are present within the tissue, the distribution becomes Rician [12]. In the case of a resolved coherent component (i.e., the spacing is larger than the speckle size), the distribution becomes a generalized Rician [12]. Finally, the case of few diffuse scatterers is modeled in [12] and [13] by a pre-Rician K-distribution [14], [15]. Now, except for Rayleigh distributions, none of these models can be easily estimated. In order to overcome this technical difficulty, Shankar [16] proposed to use Nakagami distributions [17] as an approximation of the true distributions. In [18], a compound model is proposed that modulates the scaling parameter of the Nakagami distribution with a gamma distribution. This is reminiscent of the compound decomposition of the homodyned K-distribution

Manuscript received March 17, 2008; revised July 07, 2008. First published August 8, 2008; current version published January 30, 2009. This work was jointly supported by grants from the Natural Sciences and Engineering Research Council of Canada (138570-06) and the Canadian Institutes of Health Research (PPP-78763). *Asterisk indicates corresponding author.*

F. Destrempes is with the Laboratoire de Biorhéologie et d'Ultrasonographie Médicale (LBUM), Centre de Recherche du Centre Hospitalier de l'Université de Montréal (CRCHUM), Montreal, QC, H2L 2W5 Canada.

J. Meunier is with the Département d'Informatique et de Recherche Opérationnelle (DIRO), Université de Montréal, Montreal, QC, H3T 1J4 Canada.

M.-F. Giroux and G. Soulez are with the Département de Radiologie, Centre Hospitalier de l'Université de Montréal (CHUM), Montréal, QC, H3T 1J4 Canada.

*G. Cloutier is with the Laboratoire de Biorhéologie et d'Ultrasonographie Médicale (LBUM), Centre de Recherche du Centre Hospitalier de l'Université de Montréal (CRCHUM), Montreal, QC, H2L 2W5 Canada, and with the Département de Radiologie, Radio-Oncologie et Médecine Nucléaire, Université de Montréal, Montréal, QC, H3T 1J4 Canada, and also with the Institut de Génie Biomédical, Université de Montréal, Montréal, QC H3T 1J4 Canada.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TMI.2008.929098

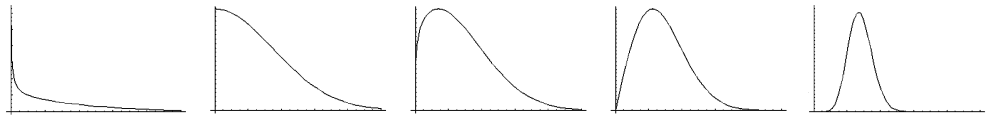

Fig. 1. Nakagami distributions for various values of the shape parameter m. From left to right: $0 < m < 1/2$ (pre-Rician distribution); $m = 1/2$ (generalized Rician distribution); $1/2 < m < 1$ (generalized Rician distribution); $m = 1$ (Rayleigh distribution); $m > 1$ (Rician distribution).

into a Rician distribution modulated with a gamma distribution [19]. In [20], the Nakagami distribution is modulated with an inverse Gaussian distribution instead of a gamma distribution. These two models can be viewed as infinite mixtures of Nakagami distributions. An alternative consists in finite mixtures of Nakagami distributions. In [21], four empirical models for the gray level of the *B*-mode image are tested, including the gamma distribution. In [22], the gamma distribution is adopted as a likelihood in a model for contour segmentation of ultrasound images. Let us mention at this point that a Nakagami distribution on the gray level of the *B*-mode image is equivalent to a gamma distribution on the square of the gray level, and this should not be confused with the gamma distribution on the gray level itself.

Note that when a log-compression operator or various filters are applied to the *B*-mode image, the model of Nakagami might not be valid anymore, since these operators affect the statistics of the data ([23, p. 705, second column, l.2-5]). In that case, a mixture of Gaussian distributions has been used in [24]. But then, a mixture of Nakagami distributions could also be used, except that each distribution will then lack the physical interpretation that stands when no filtering operators are applied. In the proposed method, it is assumed that the radio-frequency (RF) image is available and used to compute the *B*-mode image. In that case, mixtures of Nakagami distributions are used to model the brightness of the RF envelope. The parameters of the proposed model can be estimated using the EM algorithm [11]. Although there is no closed form to update parameters in the M-step of the EM algorithm in the case of mixtures of Nakagami distributions, an efficient iterative method is proposed here. To take into account the heterogeneity of tissue echogenicities in *B*-mode images, the distributions are estimated locally on nonoverlapping windows; then, at each pixel, the two distributions corresponding to the two closest adjacent (nonoverlapping) windows to that pixel are averaged in proportions corresponding to the distance from that pixel to each window, to ensure a smooth transition from one window to the next one. In [4], the Rayleigh distributions are also estimated locally, but on lines instead of regions, and no averaging of the distributions is considered. Furthermore, unlike [4]–[9], and [24], instead of associating each distribution to a single tissue, the echogenicity of the intima-media layers is modeled by a Bayesian model average of Nakagami distributions, in order to take into account the heterogeneity of the brightness in that anatomical region. The proposed distributions are used to model the likelihood of a Bayesian model. The prior of that model is based on a geometric constraint for rectilinear curves, an anatomical prior on the IMT for normal subjects, and a temporal constraint on the difference in the wall position between two subsequent frames. The solution sought is then formulated as the maximum a posteriori (MAP) of the Bayesian model. The MAP is computed using a stochastic optimization algorithm called the exploration/selection (ES) algorithm [25]. The adaptation of the ES algorithm in the context of the proposed method is novel, as well as the proposed model. The reader can consult [26] for a recent survey on ultrasound image segmentation.

II. Mixtures of Nakagami Distributions

A. Model

The Nakagami distribution [17] is defined by

$$\mathrm{N}(r|m,\Omega) = \frac{2m^m r^{2m-1}}{\Gamma(m)\Omega^m} e^{-mr^2/\Omega} \tag{1}$$

for $r \geq 0$, where Γ is the Euler gamma function. The real number $m > 0$ is called the *shape parameter*, and $\Omega > 0$ is called the *scaling parameter*. The gray level r (called *local brightness* in [26]) of the speckle pattern in a *B*-mode image was first modeled with that distribution by Shankar [16]. It reflects the local echogenicity of the underlying scatterers. When $m = 1$, the Nakagami distribution is equivalent to a Rayleigh distribution $\mathcal{R}(r|a) = (r/a^2)e^{-(r^2/2a^2)}$ with $2a^2 = \Omega$. When $m > 1$, the Nakagami distribution becomes Rician [16]; when $1/2 \leq m < 1$, the distribution is generalized Rician; and when $0 < m < 1/2$, the distribution is pre-Rician [16]. See Fig. 1 for an illustration of Nakagami distributions corresponding to various values of the shape parameter.

To model the brightness in a region encompassing more than one tissue, one is interested in mixtures of Nakagami distributions (MND)

$$f(r|p_1,\ldots,m_1,\ldots,\Omega_1,\ldots) = \sum_{i=1}^{\ell} p_i \mathrm{N}(r|m_i,\Omega_i) \tag{2}$$

where $p_i \geq 0$ and $\sum_{i=1}^{\ell} p_i = 1$. In this paper, the number of kernels ℓ is fixed.

The question of identifiability [27] of the MND model is now addressed. Namely, given the equality $f(r|p_1,\ldots,m_1,\ldots,\Omega_1,\ldots) = f(r|p'_1,\ldots,m'_1,\ldots,\Omega'_1,\ldots)$ for all $r \geq 0$, can one conclude that $p_i = p'_i$, $m_i = m'_i$, and $\Omega_i = \Omega'_i$ up to permutation of the kernels? The answer is yes, as follows from the known case [28, Proposition 2] of mixtures of gamma distributions. Indeed, a variable r follows a Nakagami distribution with parameters m, Ω if and only if the variable $I = r^2$ (called the *intensity* in [29] and [30]) follows a gamma distribution [16]:

$$\mathcal{G}(I|k,\theta) = \frac{1}{\theta^k \Gamma(k)} I^{k-1} e^{-I/\theta} \tag{3}$$

where $k = m$ and $\theta = \Omega/m$. Thus, for all practical purposes, the parameters of the MND model are uniquely determined by a large sample of data.

We found convenient to work directly with gamma distributions (in particular, gamma distributions are well known in the statistical literature). So, in practice, the change of variable $I = r^2$ on the data is first performed in order to obtain a mixture of gamma distributions (MGD)

$$f(I|p_1,\ldots,k_1,\ldots,\theta_1,\ldots) = \sum_{i=1}^{\ell} p_i \mathcal{G}(I|k_i,\theta_i) \qquad (4)$$

where the constraint $k_i\theta_i < k_{i+1}\theta_{i+1}$ is set in order to insure identifiability ($k\theta$ is the mean of a gamma distribution $\mathcal{G}(k,\theta)$). The transformed data is then normalized between 0 and 10^3. This range seemed convenient in our tests, as far as precision in the numerical methods is concerned.

B. Estimator of the Parameters of the Model

Let $\tilde{I} = I_1, I_2, \ldots, I_N$ be independent and identically-distributed (i.i.d.) samples of the MGD model. The goal is to estimate the vector of parameters $\chi = (p_i, k_i, \theta_i)$ of (4).

Under the Bayesian paradigm, the following priors on the parameters are set.

1) A Dirichlet prior [31] on the mixture parameters p_i

$$P(p_1,\ldots,p_\ell) = \mathcal{D}(p_1,\ldots,p_\ell|A_0;\alpha_1,\ldots,\alpha_\ell) \qquad (5)$$

where the Dirichlet distribution $\mathcal{D}$ is defined by $\left(\Gamma(A_0)/\prod_{i=1}^{\ell}\Gamma(A_0\alpha_i)\right)\prod_{i=1}^{\ell} p_i^{A_0\alpha_i-1}$, on the simplex $\sum_{i=1}^{\ell} p_i = 1$ and $p_i \geq 0$. Note that, under a Dirichlet prior, the mean value of p_i is α_i, whereas its variance is equal to $\alpha_i(1-\alpha_i)/(A_0+1)$. Thus, the larger the *confidence level* A_0, the smaller the variance of each proportion p_i. The special case $A_0 = \ell$ and $\alpha_i = 1/\ell$ corresponds to a uniform distribution. In this paper, the parameters of the Dirichlet distribution are adjusted based on prior anatomical information (see Sections III-D-2 and III-D-4).

2) A uniform prior on the parameters k_i and θ_i

$$P(k_1,\theta_1,\ldots,k_\ell,\theta_\ell) \propto \begin{cases} 1, & \text{on } K \\ 0, & \text{elsewhere} \end{cases} \qquad (6)$$

where K is a compact set of the form $([0,k_{\max}] \times [0,\theta_{\max}])^\ell$. In principle, one could take $k_{\max}$ and $\theta_{\max}$ as the upper limit of real numbers in a particular implementation. In practice, we restrict the search to a much smaller domain, upon taking $k_{\max} = 10$ and $\theta_{\max} = 10^3$ (the units are adjusted according to the normalization of the transformed data between 0 and 10^3). In our tests, this range appeared to be amply sufficient in order to contain the estimated values of k and θ.

We propose the MAP estimator

$$\hat{\chi} = \arg\max_{\chi} f(\tilde{I}|\chi)P(\chi). \qquad (7)$$

It can be shown from [32] that this estimator is strongly consistent[1]. The MAP can be computed using the EM algorithm [11] in the context of an arbitrary prior [11], [33]. The algorithm is detailed in Table I and its derivation is given in Appendix A. The stopping criterion adopted here is that the vector of parameters at the current iteration is at a Euclidean distance of the vector of parameters at the previous iteration that is smaller than 0.5% of its Euclidean norm (at the current iteration). In our tests, the EM algorithm is run with 10 random initializations, with a maximum of 10 000 iterations for each run. The solution with largest posterior value is then taken.

[1]This means that $\Pr(\lim_{N\to\infty}\hat{\chi}_N = \chi_0) = 1$, where χ_0 are the true parameters and $\hat{\chi}_N$ is the proposed estimator for a sample of size N.

III. Application to the Segmentation of Ultrasonic *B*-Mode Images of Normal Carotids

The goal is to perform a segmentation of the IMT of carotid arteries in sequences of ultrasonic *B*-mode images.

A. Anatomical Specifications

In Section III-C-4, anatomical information is used to construct a prior for the proposed segmentation model. Namely, we use the average IMT, as well as its spatial and temporal variance (in the sense of [34]). For that purpose, we use the recent study [34], in which 47 volunteers (29 women and 18 men) with a mean age of 49 ± 13 years underwent ultrasound examination, all of them presumed healthy. Setting a uniform prior on the four categories defined in [34, Table 2], one infers that the overall IMT in the common carotid artery is $\mu = 0.635$ mm with a spatial intrasubject variance of $\sigma_s^2 = (0.1411)^2$ mm^2 and a temporal intrasubject variance of $\sigma_t^2 = (0.1208)^2$ mm^2. We used that study in the implementation of the proposed method, since it covers a large span of age and includes only healthy subjects. Furthermore, in order to have prior information on the lumen diameter and the adventitia thickness, we used the study [35], in which 233 patients (113 women and 120 men) with a mean age of 61.6 ± 9.7 years underwent ultrasound examination. The study population consisted of 104 ischemic stroke patients and 129 patients without stroke. On this population (cf. [35, Table 1]), the IMT in the common carotid artery was 0.73 ± 0.28 mm, the lumen diameter was 5.67 ± 0.93 mm, the interadventitial diameter was 7.31 ± 1.10 mm, and the outer artery diameter was 9.11 ± 1.22 mm. One infers that the adventitia thickness was 0.90 ± 0.82 mm. This information is used in Section III-D.

B. Image Acquisition

The RF images were acquired with a Sonix RP echograph (Ultrasonix, Vancouver, BC, Canada) with a 10-MHz 38-mm linear array transducer. The frame rate was set by the radiologist and depended on the video sequence, but it was around 19 Hz. In the axial direction, 1 mm corresponds to 51.9 pixels, whereas in the longitudinal direction, 1 mm is equal to about 6.7 pixels (i.e., 256 scan lines for 38 mm).

C. Segmentation Model

1) Random Fields: Given a sequence of T RF images, let $r^{(t)}$ denote the brightness of the *B*-mode envelope of the RF signal in the tth frame, i.e., the norm of the Hilbert operator applied to the RF signal. This brightness reflects the local echogenicity of the underlying configuration of scatterers filtered by the point-spread function of the imaging system, but not directly the tissue structure.

TABLE I
EM ALGORITHM OF SECTION II-B AND APPENDIX A

ℓ	number of gamma kernels;
$I_1, ..., I_N$	independent and identically-distributed (i.i.d.) samples of the mixture of gamma distributions;
$c_1, ..., c_N$	the latent kernel variables;
p_i	proportion of the i-th kernel;
k_i, θ_i	parameters of the i-th kernel;
$A_0, \alpha_1, ..., \alpha_\ell$	parameters for the prior on $p_1, ..., p_\ell$.

Goal: to compute the Maximum A Posteriori estimator of equation (7) conditional to $I_1, ..., I_N$ (with at least two distinct elements).

for $j = 1, ..., N$ **do**
  Draw $c_j = i$ with probability $\omega_i = 1/\ell$.
end for
for $i = 1, ..., \ell$ **do**
  Set $p_i = \frac{N_i + A_0\alpha_i - 1}{N + A_0 - \ell}$, where N_i is the number of labels c_j equal to i.
  Let k_i, θ_i be the Maximum Likelihood estimator of the gamma distribution on the sample $\{I_j : c_j = i\}$.
end for
repeat
  E-step:
  for $j = 1, ..., N$ **do**
    Set $p(c_j = i) = \frac{p_i \mathcal{G}(I_j \mid k_i, \theta_i)}{\sum_{i=1}^{\ell} p_i \mathcal{G}(I_j \mid k_i, \theta_i)}$, for $i = 1, ..., \ell$.
  end for
  M-step:
  for $i = 1, ..., \ell$ **do**
    Update $p_i = \frac{P_i + A_0\alpha_i - 1}{N + A_0 - \ell}$, where $P_i = \sum_{j=1}^{N} p(c_j = i)$.
    Solve $\log x - \psi(x) = \log \frac{Q_i}{P_i} - \frac{R_i}{P_i}$, where $\psi(x)$ denotes the digamma function $\Gamma'(x)/\Gamma(x)$, $Q_i = \sum_{j=1}^{N} I_j p(c_j = i)$, and $R_i = \sum_{j=1}^{N} \log I_j p(c_j = i)$. Note that $\log x - \psi(x)$ is a decreasing function on $(0, \infty)$, $\lim_{x \to 0} \log x - \psi(x) = \infty$ and $\lim_{x \to \infty} \log x - \psi(x) = 0$. Since $\log \frac{Q_i}{P_i} - \frac{R_i}{P_i} > 0$, under the assumption that at least two sample elements are distinct, it follows that x can be found by a binary search. Update $k_i = x$.
    Update $\theta_i = \frac{Q_i}{k_i P_i}$.
  end for
until a stopping criterion is met

Let R be a region of interest that includes the inferior IMT on a longitudinal view, the adventitia, and at least 50% of the lumen, and that is delimited by two curves γ_- and γ_+, with γ_- in the lumen and γ_+ outside the carotid artery. The orientation of the axial axis is from the lumen toward the carotid artery wall of interest. In particular, $\gamma_- \leq \gamma_+$.

For each frame t, let $I^{(t)}$ be the observable random field $\left(I_s^{(t)}\right)$, where $I_s^{(t)}$ is the square of the brightness $r_s^{(t)}$ at pixel s of the corresponding B-mode image. Also, let $\gamma_1^{(t)}$ and $\gamma_2^{(t)}$ represent the interface between the IMT and the lumen or the adventitia, respectively, in the tth frame. In particular, $\gamma_1^{(t)}$ and $\gamma_2^{(t)}$ satisfy the point-wise relations $\gamma_- \leq \gamma_1^{(t)} \leq \gamma_2^{(t)} \leq \gamma_+$ (see Fig. 2). The pair of curves $\left(\gamma_1^{(t)}, \gamma_2^{(t)}\right)$ is viewed as a hidden discrete random field. The segmentation problem can then be formulated as the estimation of $\left(\gamma_1^{(t)}, \gamma_2^{(t)}\right)$ conditional to $I^{(t)}$.

2) Distributions: In order to take into account the longitudinal echogenicity variations of tissues, the region R is partitioned into B vertical strips of about 3 mm (20 pixels) wide each (see Fig. 2). The distribution of $I_s^{(t)}$ into each strip b is modeled by a mixture of three gamma distributions $\sum_{i=1}^{3} p_{b,i}^{(t)} \mathcal{G}\left(I_s^{(t)} | k_{b,i}^{(t)}, \theta_{b,i}^{(t)}\right)$. The distributions are ordered by increasing value of the means $k_{b,i}^{(t)} \theta_{b,i}^{(t)}$. See Fig. 3 for an example of an estimated distribution and its comparison with the empirical distribution.

The following assumptions on the tissues are made.

1) The lumen corresponds locally in the bth strip to the distribution with lowest mean

$$f_{b,1}^{(t)}\left(I_s^{(t)}\right) = \mathcal{G}\left(I_s^{(t)} | k_{b,1}^{(t)}, \theta_{b,1}^{(t)}\right) \tag{8}$$

since the lumen presents the lowest intensity.

2) The IMT corresponds locally to the mixture

$$f_{b,2}^{(t)}\left(I_s^{(t)}\right) = \sum_{i=1}^{3} q_{b,i}^{(t)} \mathcal{G}\left(I_s^{(t)} | k_{b,i}^{(t)}, \theta_{b,i}^{(t)}\right) \tag{9}$$

where $\sum_{i=1}^{3} q_{b,i}^{(t)} = 1$ and $q_{b,i}^{(t)} \geq 0$. One could have also considered only the gamma distribution $\mathcal{G}\left(I_s^{(t)} | k_{b,2}^{(t)}, \theta_{b,2}^{(t)}\right)$. But, in practice, the IMT presents the three types of intensity. For most images in our database, the media cannot be really distinguished from the intima tissue. Thus, note that the media does not correspond necessarily to one of the gamma distributions. A discussion on how the weights $q_{b,i}^{(t)}$ are adjusted will be presented in Sections III-D-2 and III-D-4. The proposed averaging model is more robust to the variability in intensity within the IMT due to various ultrasonographic phenomena (e.g., nonuniform acoustic gel distribution, acoustic reverberation, attenuation and shadowing, etc.).

3) The adventitia corresponds locally to the distribution with highest mean

$$f_{b,3}^{(t)}\left(I_s^{(t)}\right) = \mathcal{G}\left(I_s^{(t)}|k_{b,3}^{(t)}, \theta_{b,3}^{(t)}\right) \tag{10}$$

since the adventitia presents the highest intensity.

Ideally, the estimation of the parameters should be performed in a vertical strip centered at each pixel s in the region of interest R. Since this would have resulted in time consuming computations, we actually considered Bayesian model averaging [36]–[38] as follows. Let s be a pixel in the region of interest R, and assume that the vertical strip of width 3 mm centered at s has a proportion q_s of its area contained into the bth strip and the remaining part contained into the $(b+1)$th strip. Then, we set

$$f_i^{(t)}\left(I_s^{(t)}\right) = q_s f_{b,i}^{(t)}\left(I_s^{(t)}\right) + (1-q_s) f_{b+1,i}^{(t)}\left(I_s^{(t)}\right). \tag{11}$$

The proposed averaging operator yields a smoother transition of the distributions from one vertical strip to another.

3) Likelihood: Using the definition of the curves $\gamma_1^{(t)}$ and $\gamma_2^{(t)}$, the assumptions (8)–(10), and the averaging of (11), we define the likelihood as

$$f\left(I^{(t)}|\gamma_1^{(t)}, \gamma_2^{(t)}\right) = \prod_{\gamma_- \le s \le \gamma_1^{(t)}} f_1^{(t)}\left(I_s^{(t)}\right) \times \prod_{\gamma_1^{(t)} \le s \le \gamma_2^{(t)}} f_2^{(t)}\left(I_s^{(t)}\right) \times \prod_{\gamma_2^{(t)} \le s \le \gamma_+} f_3^{(t)}\left(I_s^{(t)}\right). \tag{12}$$

4) Prior: In the proposed model, various priors that set regularizing constraints on the segmentation process are considered.

For that purpose, let γ be a continuous piecewise linear curve of the form $(1, y_1), \ldots, (L, y_L)$, where the first coordinate indicates the longitudinal position and the second component indicates the axial position (these correspond to the horizontal and the vertical positions, respectively, in Fig. 2). Thus, the curve γ is parametrized by $\gamma(x) = (x, y(x))$, $x \in [1, L]$. Let $\bar{\dot{y}} = (1/L)\int_1^L \dot{y}(x)dx$, where $\dot{y} = dy/dx$. A smoothing spatial constraint is defined by the normalized geometrical energy (or action) of the curve γ

$$\Delta_{\rm sm}(\gamma) = \int_1^L |\dot{y}(x) - \bar{\dot{y}}|^2 dx. \tag{13}$$

This quantity is computed with the approximation

$$\Delta_{\rm sm}(\gamma) = \sum_{l=1}^{L-1} (s_l - \bar{s})^2 \tag{14}$$

where $s_l = y_{l+1} - y_l$, and $\bar{s}$ is the average value of $s_1, \ldots, s_{L-1}$.

Next, let γ and γ' be two curves of the form $(1, y_1), \ldots, (L, y_L)$ and $(1, y_1'), \ldots, (L, y_L')$, respectively. Let $d_l = y_l - y_l'$ (in mm), for $l = 1, \ldots, L$. A thickness spatial constraint is defined by the prior statistical energy

$$\Delta_{\rm th}(\gamma, \gamma') = -\sum_{l=1}^{L} \log \mathcal{N}(d_l|\mu, \sigma_s) = \frac{1}{2\sigma_s^2}\sum_{l=1}^{L}(d_l - \mu)^2 \tag{15}$$

where μ and σ_s are as in Section III-A, and $\mathcal{N}$ stands for the Gaussian distribution.

Then, a spatial prior for the first frame is defined by

$$P_1\left(\gamma_1^{(1)}, \gamma_2^{(1)}\right) \propto \exp\left(-\beta\Delta_{\rm sm}\left(\gamma_1^{(1)}\right) - \beta\Delta_{\rm sm}\left(\gamma_2^{(1)}\right) - \Delta_{\rm th}\left(\gamma_1^{(1)}, \gamma_2^{(1)}\right)\right) \tag{16}$$

where β is a positive real number. In the reported tests, the value of β was fixed throughout to 2. It appeared that the segmentations are affected by the value of β. So, ideally, the value of this parameter should be estimated in some fashion, but as of now, we have not found a way of doing so.

Finally, we define the temporal constraint by the prior statistical energy

$$\Delta_{\rm tm}(\gamma, \gamma') = -\sum_{l=1}^{L} \log \mathcal{N}(d_l|0, \sigma_t) = \frac{1}{2\sigma_t^2}\sum_{l=1}^{L} d_l^2 \tag{17}$$

with σ_t as in Section III-A.

Then, if $\gamma_1^{(t)}$ and $\gamma_2^{(t)}$ are the IMT boundaries in the tth frame, a spatio-temporal prior for the $(t+1)$th frame is defined by

$$\begin{aligned} P_{t+1}&\left(\gamma_1^{(t+1)}, \gamma_2^{(t+1)}|\gamma_1^{(t)}, \gamma_2^{(t)}\right) \\ &\propto \exp\left(-\beta\Delta_{\rm sm}\left(\gamma_1^{(t+1)}\right) - \beta\Delta_{\rm sm}\left(\gamma_2^{(t+1)}\right)\right) \\ &\quad \times \exp\left(-\Delta_{\rm th}\left(\gamma_1^{(t+1)}, \gamma_2^{(t+1)}\right)\right) \\ &\quad \times \exp\left(-\Delta_{\rm tm}\left(\gamma_1^{(t+1)}, \gamma_1^{(t)}\right) - \Delta_{\rm tm}\left(\gamma_2^{(t+1)}, \gamma_2^{(t)}\right)\right). \end{aligned} \tag{18}$$

5) Posterior Distributions: For the first frame, the posterior distribution is expressed as

$$P_1\left(\gamma_1^{(1)}, \gamma_2^{(1)}|I^{(1)}\right) \propto f\left(I^{(1)}|\gamma_1^{(1)}, \gamma_2^{(1)}\right) P_1\left(\gamma_1^{(1)}, \gamma_2^{(1)}\right) \tag{19}$$

[cf. (12) and (16)]. For the $(t+1)$th frame, one obtains from (12) and (18) the posterior distribution

$$\begin{aligned} P_{t+1}&\left(\gamma_1^{(t+1)}, \gamma_2^{(t+1)}|I^{(t+1)}, \gamma_1^{(t)}, \gamma_2^{(t)}\right) \\ &\propto f\left(I^{(t+1)}|\gamma_1^{(t+1)}, \gamma_2^{(t+1)}\right) \\ &\quad \times P_{t+1}\left(\gamma_1^{(t+1)}, \gamma_2^{(t+1)}|\gamma_1^{(t)}, \gamma_2^{(t)}\right). \end{aligned} \tag{20}$$

The segmentation of the IMT is then viewed as the following sequential MAP estimators:

$$\begin{aligned} \left(\gamma_1^{(1)}, \gamma_2^{(1)}\right) &= \arg\max P_1\left(\gamma_1, \gamma_2|I^{(1)}\right) \\ \left(\gamma_1^{(t+1)}, \gamma_2^{(t+1)}\right) &= \arg\max P_{t+1}\left(\gamma_1, \gamma_2|I^{(t+1)}, \gamma_1^{(t)}, \gamma_2^{(t)}\right) \end{aligned} \tag{21}$$

where $t = 1, \ldots, T-1$.

D. Segmentation Method

The proposed segmentation method is now presented in details.

1) Manual Initialization: A few pixels $s_1, \ldots, s_u$ (3–5 in our tests) located in the IMT in the first frame of the sequence are entered manually by an operator. The piecewise linear curve γ_0 defined by these pixels is computed.

From Section III-A, the IMT plus half the lumen diameter plus 1 SD of each is equal to 4.31 mm, whereas the IMT plus the adventitia thickness plus 1 SD of each equals $0.73 + 0.28 + 0.90 + 0.82 = 2.73$ mm. In our implementation, we actually consider (somewhat arbitrarily) the region within 5 mm toward the lumen and 3 mm toward the adventitia from the curve γ_0. Accordingly, the curves γ_- and γ_+ are obtained by translating γ_0 5 mm toward the lumen, and 3 mm toward the adventitia, respectively.

2) Estimation in the First Frame: In each vertical strip b, 800 points between the curves γ_- and γ_+ are chosen randomly according to a uniform distribution. A mixture of 3 gamma distributions $\sum_{i=1}^{3} p_{b,i}^{(1)} \mathcal{G}(k_{b,i}^{(1)}, \theta_{b,i}^{(1)})$ is estimated using the feature $I^{(1)}$ of the first frame at those 800 points, according to the EM algorithm of Table I. Let λ be the average axial distance between γ_+ and γ_-. Let $\mu = 0.635$ and $\sigma_s = 0.1411$ be the mean and the spatial standard deviation of the IMT as presented in Section III-A. Then, we adjust the hyper-parameters of the Dirichlet prior on the proportions p_i of the MGD by setting $\alpha_2 = \mu/\lambda$, thus imposing the prior average thickness μ to the region corresponding to the second gamma distribution. The two other gamma distributions are given an equal prior weight of $\alpha_1 = \alpha_3 = (\lambda - \mu)/(2\lambda)$, so that $\alpha_1 + \alpha_2 + \alpha_3 = 1$. Also, from the anatomical specifications, we want the variance of p_2 to be equal to $(\sigma_s/\lambda)^2$. Since the variance of the marginal variable p_2 of the Dirichlet distribution is equal to $\alpha_2(1-\alpha_2)/(A_0+1)$, one obtains $A_0 = \alpha_2(1-\alpha_2)(\lambda/\sigma_s)^2 - 1$, upon solving the equation $\alpha_2(1-\alpha_2)/(A_0+1) = (\sigma_s/\lambda)^2$.

The distributions f_i, for $i = 1, 2, 3$, are constructed as in Section III-C [cf. (8)–(11)]. For the first frame, the values $q_{b,1}^{(1)} = q_{b,2}^{(1)} = q_{b,3}^{(1)} = 1/3$ are taken in (9).

3) Segmentation in the First Frame: The continuous piecewise linear curves $\gamma_1^{(1)}$ and $\gamma_2^{(1)}$ located within 2 mm above and below γ_0, that maximize the posterior distribution of (19) are computed using the optimization algorithm of Section III-E. Note that from Section III-A, the IMT plus 3 SD is equal to 1.57 mm. We actually take 2 mm (somewhat arbitrarily) to cover an even wider span.

4) Estimation in the Subsequent Frames: Let $\gamma_1^{(t)}$ and $\gamma_2^{(t)}$ be the two solutions of the current frame t. Two curves γ_- and γ_+ are obtained by translating $\gamma_1^{(t)}$ 3 mm toward the lumen, and $\gamma_2^{(t)}$ 2 mm toward the adventitia, respectively. From Section III-A, half the lumen plus 1 SD is equal to 3.3 mm, whereas the adventitia thickness plus 1 SD is equal to 1.72 mm. Now, the estimation as in step 2 is performed, but with the following modifications.

Let λ_- and λ_+ be the average axial distances between γ_- and $\gamma_1^{(t)}$, and between $\gamma_2^{(t)}$ and γ_+, respectively. Let $\lambda = \lambda_- + \mu + \lambda_+$. We set $\alpha_1 = \lambda_-/\lambda$, $\alpha_2 = \mu/\lambda$, and $\alpha_3 = \lambda_+/\lambda$, in order to impose the prior average thickness λ_-, μ, or λ_+ to the region corresponding to the first, second, or third gamma distribution, respectively. Also, from the anatomical specifications, we want the variance of p_2 to be equal to $(\sigma_s/\lambda)^2$. We then deduce the value $A_0 = \alpha_2(1-\alpha_2)(\lambda/\sigma_s)^2 - 1$ as mentioned above.

In (9), $q_{b,i}^{(t)}$ is estimated as the proportion of points in the b-th strip between $\gamma_1^{(t)}$ and $\gamma_2^{(t)}$ for which the distribution $\mathcal{G}\left(k_{b,i}^{(t)}, \theta_{b,i}^{(t)}\right)$ *estimated at the current frame* is more likely than the two other distributions.

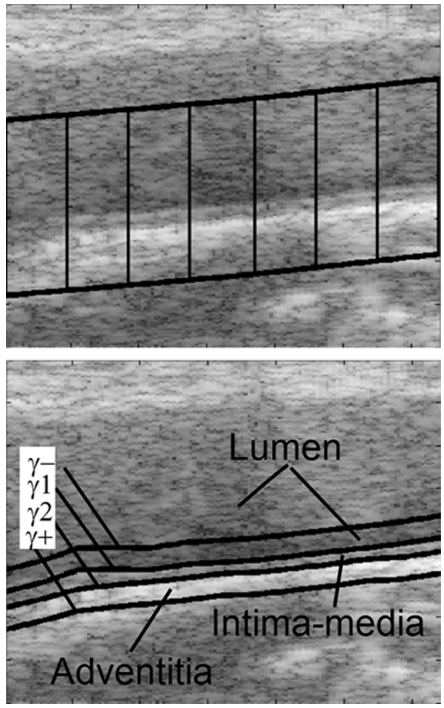

Fig. 2. Top: subdivision of a region of interest into nonoverlapping vertical strips (see Section III-C-2). Bottom: the curves γ_1 and γ_2 delimiting the IMT are constrained within the region between γ_- and γ_+. The solution (γ_1, γ_2) of the proposed method is viewed as the MAP of a Bayesian model (see Section III-C-1).

5) Segmentation in the Subsequent Frames: The continuous piecewise linear curves $\gamma_1^{(t+1)}$ and $\gamma_2^{(t+1)}$ within 1 mm toward the lumen from $\gamma_1^{(t)}$ and 1 mm toward the adventitia from $\gamma_2^{(t)}$ that maximize the posterior distribution of (20) are computed using the optimization algorithm of Section III-E. The choice of 1 mm is arbitrary, but it seems amply sufficient to allow tracking the movement of the wall from one frame to the next one, provided the transducer is held fixed.

6) Final Segmentation of the First 20 Frames: Once the whole video sequence has been segmented, the first 20 frames are estimated and segmented again using the solution found at the 21st frame and working backwards. This procedure helps correcting the errors found in the first few frames due to an error in the initialization. The choice of 20 frames was somewhat arbitrary but it seemed sufficient for the correction purpose.

E. Optimization Algorithm

In order to perform the segmentation steps of Sections III-D-3 and III-D-5, an optimization algorithm is needed. In the context of segmentation of ultrasonic images, various optimization algorithms have been used, such as the simulated annealing algorithm in [39], the simulated annealing algorithm with a modified Gibbs sampler [40], the iterated conditional mode (ICM) algorithm [41] in [42], iterative multigrid dynamic programming in [3], and a genetic algorithm in [5].

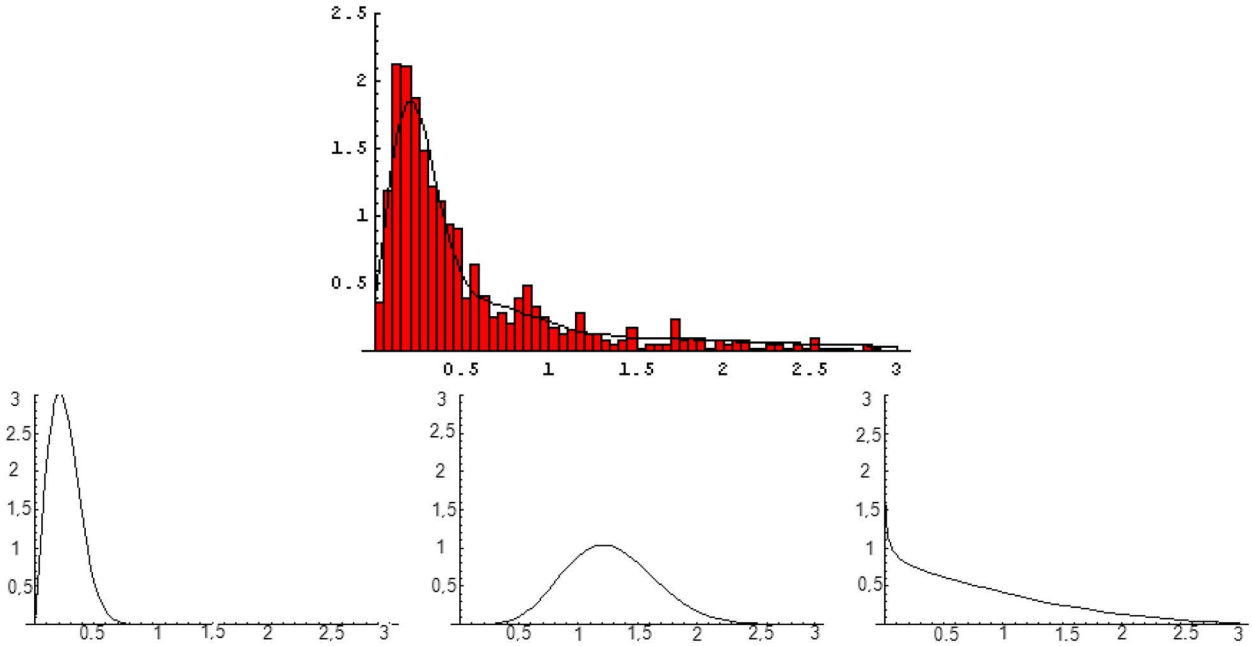

Fig. 3. Example of the distribution of the *B*-mode echogenicity in a vertical strip as a mixture of three Nakagami distributions, as estimated by the EM algorithm. Top: The estimated mixture of Nakagami distributions is compared with the empirical distribution (normalized histogram). Bottom, from left to right, the three estimated Nakagami distributions in the mixture and their proportions: lumen (almost Rayleigh) distribution ($k_1 = 1.04, k_1\theta_1 = 0.08, p_1 = 0.53$); intima-media Rician distribution ($k_2 = 2.92, k_2\theta_2 = 0.61, p_2 = 0.10$); adventitia pre-Rician distribution ($k_3 = 0.41, k_3\theta_3 = 2.93, p_3 = 0.37$).

In our case, we resort to the ES global minimization algorithm of François [25], [43]. The ES algorithm is a particular generalized simulated annealing algorithm that finds asymptotically a global minimum of a function F defined on a finite search space E. A population of n solutions (or particles) is initialized randomly. Then, at each iteration and for each solution independently of the others, two operators are available: the random exploration of a solution within a neighbourhood of the search space endowed with a connected graph structure (exploration step), or the replacement of the solution by the best solution of the previous iteration (selection step). The exploration is chosen with probability p_i, called the exploration probability that depends on the iteration i. This probability decreases to 0 as the number of iterations increases, at a known rate that ensures convergence to the optimal solution (see Table II). Namely, let $\tilde{D}$ be the diameter of the search space endowed with its connected graph structure. In order that the ES algorithm converge to a global minimum, it is sufficient to have $n > \tilde{D}$ (number of particles) and $p_i = i^{-1/\tau}$, where $\tau \geq \tilde{D}$. See for instance [25] or [44, Ch. 2], for detailed explanations. The general form stated in [45, p. 40, line 1], for which the convergence is proved in [44], [46], [47], is actually used in the proposed method; namely, the exploration distribution can be any positive distribution on the neighborhood of a solution, and not just the uniform distribution considered in [25], [43]. This yields a flexible form of the ES algorithm.

Here, the function F is as in Table II, and the search space is the set of pairs of curves (γ_1, γ_2). Since each piecewise linear curve is described by L control points, we consider the search space $E = \{(\gamma_1, \gamma_2) : \gamma_- \leq \gamma_1 \leq \gamma_2 \leq \gamma_+\}$. That set is finite upon considering a finite precision for real numbers on the computer. The graph structure is defined by the symmetric relation $(\gamma_1, \gamma_2) \sim_\rho (\gamma_1', \gamma_2')$ if and only if the axial coordinate of each control point of γ_1 (γ_2) is within a distance ρ of the axial coordinate of the corresponding point of γ_1' (γ_2', respectively). For each control point, the distance is normalized so that the axial coordinates of γ_- and γ_+ are at a distance equal to 1. With that graph structure, the search space has diameter equal to $1/\rho$ (in our tests, we take systematically $\rho = 1/8$). Next, if we set $D = 1/(2\rho)$, and we let the exploration step consist of up to D basic explorations within a distance ρ (the number of steps being drawn from a binomial distribution), then the diameter of the new implicit underlying graph structure is equal to $\tilde{D} = 2$.[2] Thus, it is enough to take $n > \tilde{D} = 2$ and $\tau \geq \tilde{D} = 2$ in order to have the asymptotic convergence property. In our tests, we took $n = 30$ (in practice, it is preferable to have more particles than needed in order to accelerate the convergence of the algorithm) and $\tau = 15$ (we found emprically that $\tau = n/2$ is in general a good choice). See [47] and [48] for similar choices of the internal parameters of the ES algorithm.[3]

The application of the ES algorithm in the context of this paper can be viewed as a variant of the localization and reconstruction methods of [48]–[51]. Strictly speaking, (18) is equivalent to a Gaussian statistical shape prior on the IMT of the carotid artery (in a longitudinal view). However, contrary to the above references, no reduction of dimensionality has been used. Instead, the variant presented in Table II is based on local moves of the control points (of the piecewise linear curves). In particular, we think that the proposed variant could be well adapted to free curves (i.e., without shape prior).

[2]Incidentally, if one wished to have $\tilde{D} = 1$, one should take $D = 1/\rho$.

[3]In [47] and [48], the radius of exploration ρ is denoted r. Here, we do not want to confuse with the speckle brightness r.

TABLE II
ES ALGORITHM OF SECTION III-E

(γ_1, γ_2)	pair of intima-media piecewise linear curves;
L	number of control points on each curve;
$F\big((\gamma_1, \gamma_2)\big)$	$-\log P_1(\gamma_1, \gamma_2 \mid I^{(1)})$ or $-\log P_{t+1}(\gamma_1, \gamma_2 \mid I^{(t+1)}, \gamma_1^{(t)}, \gamma_2^{(t)})$;
n	size of the population ($n = 30$);
i	iteration number;
$\big((\gamma_1, \gamma_2)_l^{[i]}\big)$	population of the pairs of intima-media curves at iteration i, with $l = 1, ..., n$;
$p_i = i^{-1/\tau}$	probability of exploration at iteration i ($\tau = 15$ in our tests);
ρ	radius of exploration ($\rho = 1/8$ in our tests);
$D = 1/(2\rho)$	diameter of exploration;

Parameter initialization: Initialize randomly $(\gamma_1, \gamma_2)_l^{[0]}$, for $l = 1, ..., n$. Set $i = 1$.
repeat
  Update $i \leftarrow i + 1$.
  Determine the best current solution $\alpha\Big(\big((\gamma_1, \gamma_2)_l^{[i]}\big)\Big) = (\gamma_1, \gamma_2)_l^{[i]}$, where $F\big((\gamma_1, \gamma_2)_k^{[i]}\big) > F\big((\gamma_1, \gamma_2)_l^{[i]}\big)$ for $k < l$, and $F\big((\gamma_1, \gamma_2)_k^{[i]}\big) \geq F\big((\gamma_1, \gamma_2)_l^{[i]}\big)$ for $k > l$.
  for $l = 1, 2, ..., n$ **do**
    Let u be a random number between 0 and 1.
    if $u \leq p_i = i^{-\frac{1}{\tau}}$ **then**
      Exploration: Draw j from a binomial distribution $B(1/(D-1), D-1)$. Set $(\gamma_1, \gamma_2) = (\gamma_1, \gamma_2)_l^{[i]}$.
      for $j + 1$ times **do**
        Replace (γ_1, γ_2) by (γ_1', γ_2') drawn as follows.
        for $j = 1, ..., L$ **do**
          With probability $1/L$, set randomly the j-th control point (c. p.) of γ_1' within a radius ρ of the j-th c. p. of γ_1 and between γ_- and γ_+. Then, set randomly the j-th c. p. of γ_2' within a radius ρ of the j-th c. p. of γ_2, with the constraint that it is located between the j-th c. p. of γ_1' and γ_+.
          Else, set the j-th c. p. of γ_1' and γ_2' equal to the corresponding point of γ_1 and γ_2, respectively.
        end for
      end for
      Set $(\gamma_1, \gamma_2)_l^{[i+1]} = (\gamma_1, \gamma_2)$.
    else
      Selection: Set $(\gamma_1, \gamma_2)_l^{[i+1]} = \alpha\Big(\big((\gamma_1, \gamma_2)_l^{[i]}\big)\Big)$.
    end if
  end for
until a stopping criterion is met

IV. EVALUATION METHODS

A. Evaluation of the Estimation Method

Let $\sum_{i=1}^{3} p_{b,i}^{(t)} \mathcal{G}\left(I_s^{(t)} | k_{b,i}^{(t)}, \theta_{b,i}^{(t)}\right)$ be the estimated mixture of gamma distributions in the bth strip of the tth frame, as in Sections III-D-2 and III-D-4. We want to test the goodness-of-fit of that distribution on the corresponding sample set $\tilde{I}$. For that purpose, we propose the following Bayesian test in the context of [52, Sec. 5.2.4], that uses quantiles of distributions.

Given a fixed number of bins M, and a distribution $f(I)$, we consider the M equiprobable bins $B_1, \ldots, B_M$ corresponding to the M quantiles of f. Next, given a sample $\tilde{I}$ of size N, we let N_i be the number of samples falling in the ith bin, for $i = 1, \ldots, M$. We assume that the numbers $N_1, \ldots, N_M$ follow a multinomial distribution $\mathcal{M}(N_1, \ldots, N_M | p_1, \ldots, p_M) = \left(N! / \prod_{i=1}^{M} N_i!\right) \prod_{i=1}^{M} p_i^{N_i}$, with unknown parameters $\vec{p} = (p_1, \ldots, p_M)$. We test the null hypothesis $H_0 : \vec{p} = \vec{p}_0 = (1/M, \ldots, 1/M)$, for if f were the true distribution for the sample $\tilde{I}$, each bin would have probability $1/M$ (by construction). The alternative hypothesis is $H_1 : \vec{p} \in \Theta_1 = \left\{(p_1, \ldots, p_M) : p_i \geq 0, \sum_{i=1}^{M} p_i = 1\right\} \setminus \{\vec{p}_0\}$.

The likelihood $f(\tilde{I}|\vec{p}_0) = \mathcal{M}(N_1, \ldots, N_M | 1/M, \ldots, 1/M)$ under the null hypothesis is equal to

$$\frac{N!}{\prod_{i=1}^{M} N_i!} \prod_{i=1}^{M} M^{-N_i} = \frac{N!}{\prod_{i=1}^{M} N_i!} M^{-N}. \tag{22}$$

We choose the Dirichlet distribution $g_1(\vec{p}) = \mathcal{D}(p_1, \ldots, p_M | M/2; 1/M, \ldots, 1/M) = \Big(\Gamma(M/2) / \prod_{i=1}^{\ell} \Gamma(1/2)\Big) \prod_{i=1}^{\ell} p_i^{-1/2}$ as prior for $\vec{p}$. This is actually Jeffrey's prior [53], [54] for the multinomial model. The corresponding marginal $m_1(\tilde{I}) = \int_{\Theta_1} f(\tilde{I}|\vec{p}) g_1(\vec{p}) d\vec{p}$ is equal to

$$\int_{\Theta_1} \frac{N!}{\prod_{i=1}^{M} N_i!} \frac{\Gamma\left(\frac{M}{2}\right)}{\Gamma\left(\frac{1}{2}\right)^M} \prod_{i=1}^{M} p_i^{N_i - 1/2} d\vec{p}$$
$$= \frac{N!}{\prod_{i=1}^{M} N_i!} \frac{\Gamma\left(\frac{M}{2}\right)}{\Gamma\left(\frac{1}{2}\right)^M} \frac{\prod_{i=1}^{M} \Gamma\left(\frac{N_i+1}{2}\right)}{\Gamma\left(\frac{N+M}{2}\right)}. \tag{23}$$

We obtain the Bayes factor

$$B_{0,1} = \frac{f(\tilde{I}|\vec{p}_0)}{m_1(\tilde{I})} = \frac{\Gamma\left(\frac{1}{2}\right)^M}{\Gamma\left(\frac{M}{2}\right)} \frac{\Gamma\left(\frac{N+M}{2}\right)}{\prod_{i=1}^{M} \Gamma\left(\frac{N_i+1}{2}\right)} M^{-N}. \tag{24}$$

The null hypothesis H_0 is accepted if and only if $B_{0,1} \geq 1$.

TABLE III
MEAN AND STANDARD DEVIATION OF THE SHAPE PARAMETER k AND THE AVERAGE $k\theta$ OF THE THREE GAMMA DISTRIBUTIONS, AS ESTIMATED OVER ALL VERTICAL STRIPS AND ALL FRAMES OF THE 30 SEQUENCES WITH THE EM ALGORITHM OF TABLE I

distribution	common carotid artery ($N = 7283$)	proximal internal carotid artery ($N = 9200$)
lumen	$k = 0.94 \pm 0.21$ and $k\theta = 0.31 \pm 1.34$	$k = 0.99 \pm 0.19$ and $k\theta = 0.42 \pm 1.13$
intima-media distr.	$k = 1.45 \pm 1.54$ and $k\theta = 2.86 \pm 6.44$	$k = 1.74 \pm 1.67$ and $k\theta = 3.98 \pm 11.35$
adventitia	$k = 0.42 \pm 0.14$ and $k\theta = 22.11 \pm 26.77$	$k = 0.54 \pm 0.26$ and $k\theta = 22.06 \pm 28.22$

B. Evaluation of the Segmentation Method

Let N video sequences of *B*-mode images be given. For each sequence, one expert chooses a region of interest and a temporal interval of two to three cardiac cycles. Then, two independent experts segmented manually the IMT for each frame of these N sequences. This yields the curves $\gamma_{i,j,1}^{(t)}$ and $\gamma_{i,j,2}^{(t)}$ for the ith sequence $(i = 1, \ldots, N)$, obtained by the jth expert $(j = 1, 2)$, on the tth frame (the maximal value $T(i)$ for t varies with the sequence and corresponds to two to three cycles).

Also, applying the proposed segmentation method of Section III-D to the same sequences, one obtains the curves $\gamma_{i,0,1}^{(t)}$ and $\gamma_{i,0,2}^{(t)}$ for the i-th sequence on the t-th frame.

Now, let $D(\gamma, \gamma')$ be a distance between two curves. In the reported tests, the average point-to-point distance and the Hausdorff point-to-point distance [55] between the two curves were considered. One then defines the following distance between two segmentations $j \neq j' \in \{0, 1, 2\}$ of a same video sequence i

$$D_{i,j,j'} = \max \left\{ \frac{1}{T(i)} \sum_{t=1}^{T(i)} D\left(\gamma_{i,j,1}^{(t)}, \gamma_{i,j',1}^{(t)}\right) \right. \\ \left. \frac{1}{T(i)} \sum_{t=1}^{T(i)} D\left(\gamma_{i,j,2}^{(t)}, \gamma_{i,j',2}^{(t)}\right) \right\}. \quad (25)$$

For each pair of indices $j \neq j' \in \{1, 2\}$ and for each index $k \in \{1, 2\}$, we compare the population of distances between the segmentation of the experts j and j', i.e., $D_{1,j,j'}, \ldots, D_{N,j,j'}$, with the population of distances between the semi-automatic segmentation and the segmentation of expert k, i.e., $D_{1,0,k}, \ldots, D_{N,0,k}$. This leads to the one-sided p-value of Welch's approximate t-test that the random variable $D_{\cdot,0,k}$ is no more than $D_{\cdot,j,j'}$ (null hypothesis). With a confidence level of $\alpha = 0.05$, the test succeeds whenever $p \geq 0.05$. Note that $(1/N)\sum_{i=1}^{N} D_{i,0,k} \leq (1/N)\sum_{i=1}^{N} D_{i,j,j'}$ if and only if $p \geq 1/2$.

V. EXPERIMENTAL RESULTS

A. Data

$N = 30$ video sequences of *B*-mode images from 15 presumably disease-free control subjects were considered. For each subject, longitudinal views of the right distal common carotid artery and right proximal internal carotid were acquired by one expert radiologist, for a total of two video sequences per patient. The number of expert technicians for the manual segmentations was 2.

B. Performance of the Estimation Method

Two versions of the proposed method were tested: 1) the MGD model (which is equivalent to a mixture of Nakagami distributions after the change of variable $I = r^2$) estimated by the EM algorithm (Section II-B); and 2) the mixture of exponential distributions (MED) model (which is equivalent to the mixture of Rayleigh distributions after the same change of variable) estimated by the EM algorithm.[4] Using the evaluation method of Section IV-A, we compared the goodness-of-fit of the estimated mixtures of distributions performed on the 30 video sequences. For the distal common carotid artery, the percentage of success of the goodness-of-fit test was 83.2% with the MGD model, but only 73.4% with the MED model. For the internal proximal carotid artery, the percentage of success was 91.5% with the MGD model, but it was reduced to 87.7% with the MED model. The mean value and standard deviation of the estimated parameters of the gamma distributions (using the EM algorithm) are presented in Table III. For the MGD model, there were only 224 runs of the EM algorithm out of 16 484 for which the convergence was not reached within the maximum of 10 000 iterations.[5]

C. Performance of the Segmentation Method

In our tests, the first author chose the initial points for the semi-automatic segmentations in the 30 video sequences. A clinician (the fourth author) confirmed that there were no atherosclerotic plaque appearing in the video sequences of the internal proximal carotid and distal common carotid arteries. This prior information was used only for the semi-automatic computerized segmentations. The manual segmentations were done independently by the two technical experts. In particular, they had to decide by themselves if there were a plaque or not. Furthermore, the clinician and the first author decided to restrict the region of interest R in the case of 3 of these video sequences

[4]In our implementation used for the reported tests, a distribution is dropped out whenever its estimated proportion becomes less than 5% (i.e., 40 points out of 800) within the EM algorithm. This exception never occured for the 16 484 estimations in the case of the model of a mixture of Nakagami distributions. In the case of the model of a mixture of Rayleigh distributions, it occured only 3 times out of the 16 484 estimations (corresponding to all vertical strips of the frames of the 30 sequences).

[5]When convergence was not achieved, the parameters were taken as in the last iteration. We also implemented a stochastic algorithm, called the exploration/selection/estimation (ESE) procedure [46] in the context of a mixture of gamma distributions in order to initialize the EM algorithm. In that case, the EM algorithm always reached convergence within 10 000 iterations. However, the goodness-of-fit tests and most importantly, the segmentation results as in Tables IV and V were practically the same as with the simpler use of 10 runs of the EM algorithm, as proposed here. For simplicity sake, we decided not to include the details of the ESE procedure in this paper.

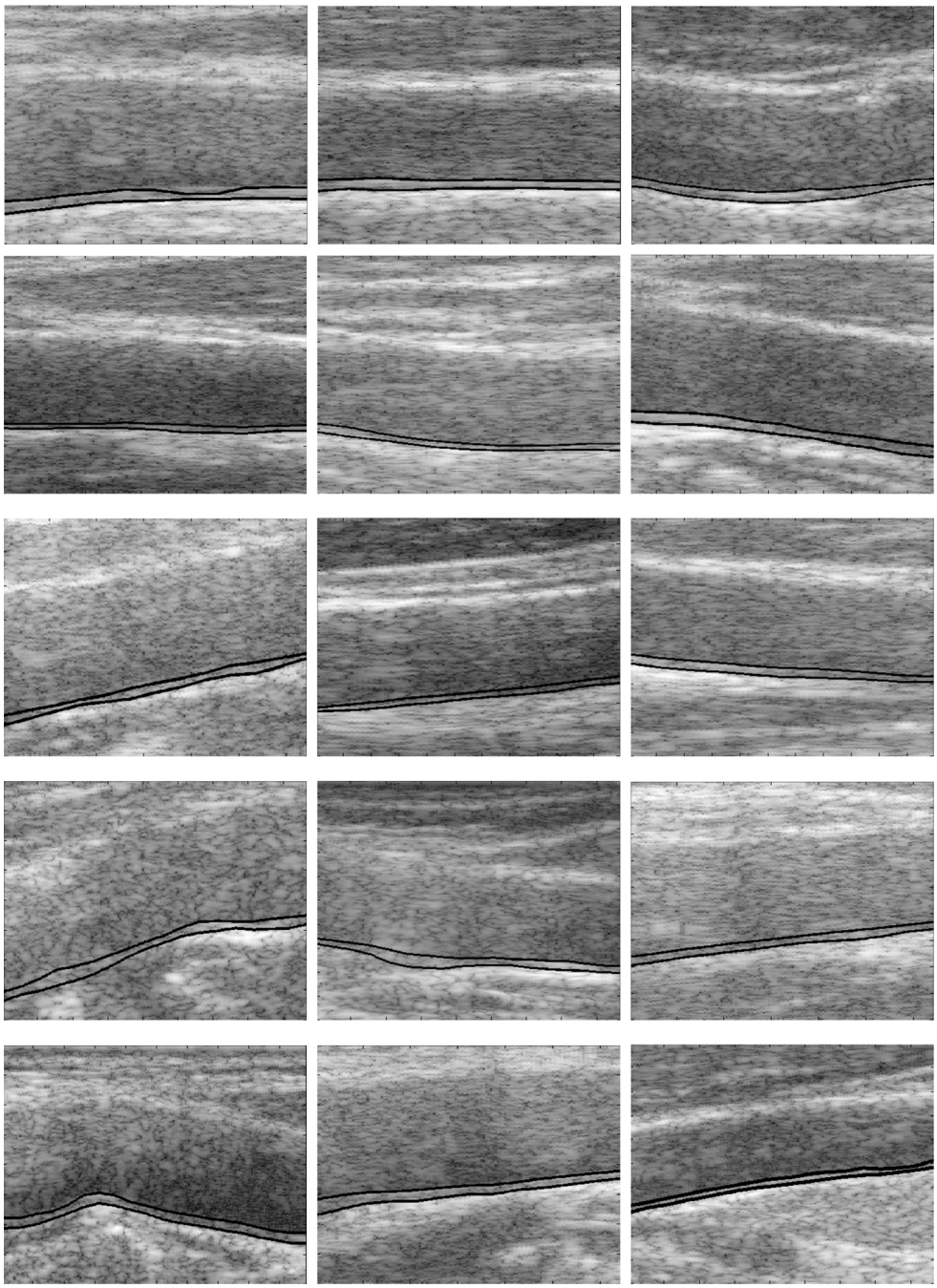

Fig. 4. Segmentations of the first frame of the 15 video sequences of B-mode images of common carotid arteries defined as the MAP of a Bayesian model that is computed using the ES optimization algorithm. The mixtures of gamma distributions that are used to define the likelihood of that Bayesian model, are estimated using the EM algorithm.

of the internal proximal carotids, due to a great uncertainty in the location of the IMT (see Fig. 5).

Again, the two versions of the proposed method mentioned in Section V-B were tested on the 30 video sequences. Figs. 4 and 5 show the segmentations of the IMT on the first frame for the 30 video sequences, using the MND model as estimated by the EM algorithm. Using the evaluation method of Section IV-B, a detailed analysis of the empirical results is presented in Tables IV and V.

Namely, Table IV concerns the 15 video sequences of the common carotid artery, whereas Table V presents a similar analysis for the internal proximal carotid artery. Row 1 compares the two experts upon presenting the distance $D_{i,1,2}$ and its standard deviation. Rows 2 and 3 present an analysis of the results for the MGD model estimated by the EM algorithm; more precisely, row 2 compares the segmentation method with the first expert (distance $D_{i,0,1}$ and standard deviation); the p-value of the one-sided difference t-test of Section IV-B, for the comparison of the segmentations obtained by the proposed method with the manual segmentations of expert 1, is indicated in parenthesis. Row 3 concerns the second expert (the distance is now denoted $D_{i,0,2}$). Rows 4 and 5 concern the MED model, as es-

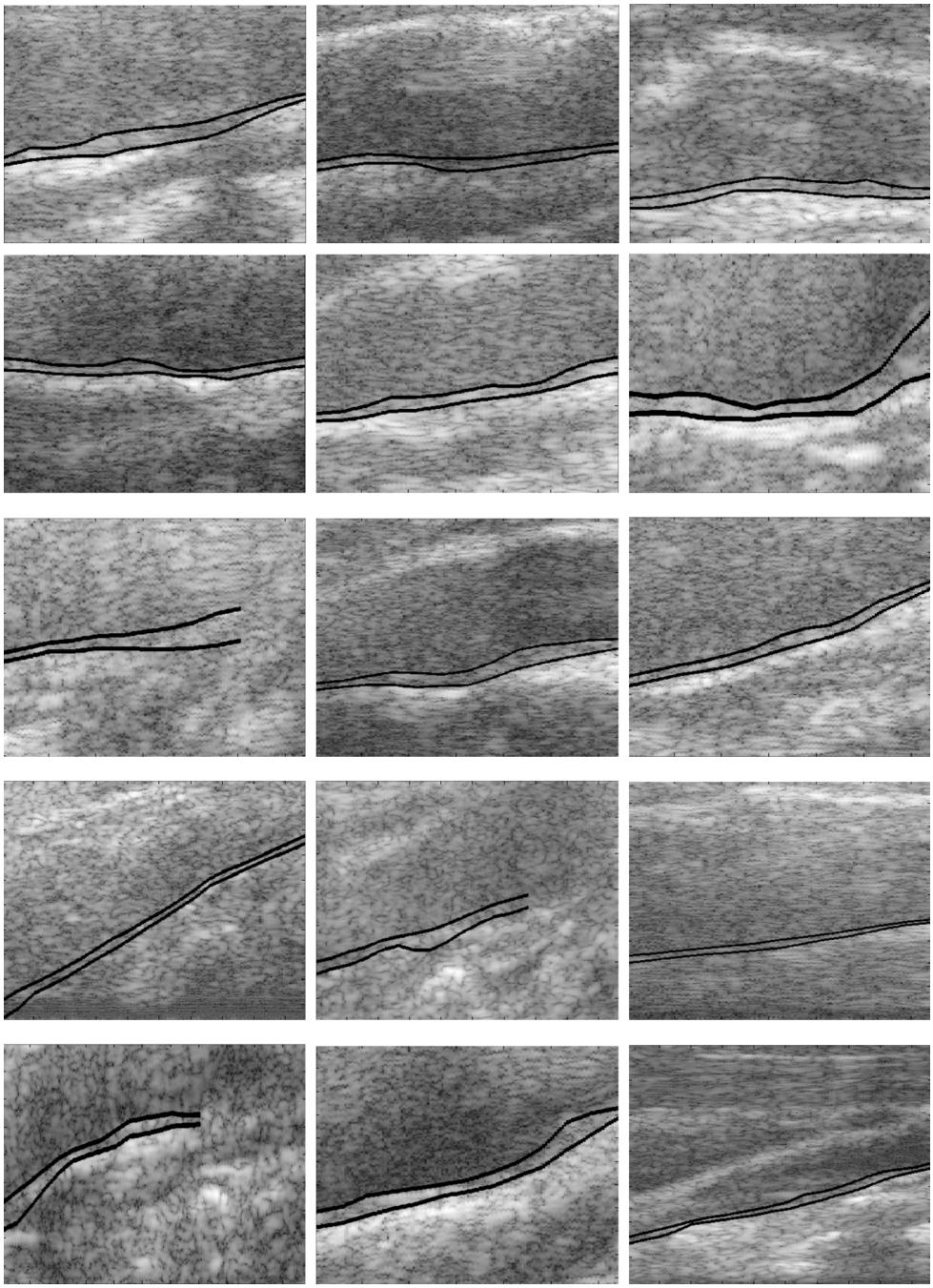

Fig. 5. Segmentations of the first frame of the 15 video sequences of *B*-mode images of internal proximal carotid arteries defined as the MAP of a Bayesian model that is computed using the ES optimization algorithm. The mixtures of gamma distributions that are used to define the likelihood of that Bayesian model, are estimated using the EM algorithm.

timated by the EM algorithm. Finally, columns 1–4 use, respectively, the average distance for the interface lumen-intima, the Hausdorff distance for the same interface, the average distance for the interface media-adventitia, and the Hausdorff distance for that same interface.

The EM algorithm combined with the segmentation method took 14 h and 41 min (this represents 16 483 estimations and 2227 segmentations) for the 30 video sequences. The implementation was in C++ and the tests were run on a 3 GHz Pentium 4 CPU.

VI. Discussion

A. Statistical Model

The results of the goodness-of-fit tests reported in Section V-B indicate that the simpler MED model is less adequate than the proposed MGD model. This is explained by the fact that the Rayleigh distribution covers only the case of a high density of random scatterers. In particular, Table III indicates that the lumen distribution is practically Rayleigh (high density of random scatterers, i.e., $k = 1$), the adventitia distribution is

TABLE IV
COMPARISON OF THE SEGMENTATIONS OBTAINED BY THE PROPOSED METHOD USING MGDS OR MEDS, ON ONE HAND, WITH THE MANUAL SEGMENTATIONS PERFORMED BY TWO EXPERTS, ON THE OTHER HAND, FOR THE 15 VIDEO SEQUENCES OF THE COMMON CAROTID ARTERY. HERE, THE SAMPLE SIZE IS $N = 15$. THE p-VALUE OF THE STATISTICAL TEST THAT THE DISTANCE BETWEEN THE ALGORITHM AND AN EXPERT IS NO MORE THAN THE DISTANCE BETWEEN THE TWO EXPERTS IS INDICATED WITHIN PARENTHESIS (THE HIGHER THE p-VALUE, THE BETTER THIS TEST SUCCEEDS). TWO TYPES OF DISTANCES ARE USED: THE MEAN DISTANCE AND THE HAUSDORFF DISTANCE; DISTANCES ARE AVERAGED OVER ALL FRAMES OF A SEQUENCE. SEE SECTION IV-B FOR FURTHER DETAILS

Common carotid	lum.-int. (mean dist.)	lum.-int. (Hausd. dist.)	med.-adv. (mean dist.)	med.-adv. (Hausd. dist.)
Exp. 1 vs Exp. 2	0.22 ± 0.11mm	0.45 ± 0.23mm	0.15 ± 0.05mm	0.38 ± 0.18mm
Exp. 1 vs MGD	0.21 ± 0.13mm (0.57)	0.46 ± 0.21mm (0.45)	0.16 ± 0.07mm (0.27)	0.41 ± 0.18mm (0.28)
Exp. 2 vs MGD	0.18 ± 0.11mm (0.84)	0.46 ± 0.31mm (0.47)	0.15 ± 0.10mm (0.54)	0.40 ± 0.31mm (0.39)
Exp. 1 vs MED	0.21 ± 0.14mm (0.61)	0.46 ± 0.25mm (0.46)	0.31 ± 0.42mm (0.075)	0.76 ± 0.90mm (0.062)
Exp. 2 vs MED	0.20 ± 0.16mm (0.62)	0.49 ± 0.35mm (0.39)	0.31 ± 0.47mm (0.10)	0.73 ± 0.96mm (0.090)

TABLE V
COMPARISON OF THE SEGMENTATIONS OBTAINED BY THE PROPOSED METHOD FOR THE MGDS OR THE MEDS, ON ONE HAND, WITH THE MANUAL SEGMENTATIONS PERFORMED BY TWO EXPERTS, ON THE OTHER HAND, FOR THE 15 VIDEO SEQUENCES OF THE INTERNAL PROXIMAL CAROTID ARTERY. HERE, THE SAMPLE SIZE IS $N = 15$. THE p-VALUE OF THE STATISTICAL TEST THAT THE DISTANCE BETWEEN THE ALGORITHM AND AN EXPERT IS NO MORE THAN THE DISTANCE BETWEEN THE TWO EXPERTS IS INDICATED WITHIN PARENTHESIS (THE HIGHER THE p-VALUE, THE BETTER THIS TEST SUCCEEDS). TWO TYPES OF DISTANCES ARE USED: THE MEAN DISTANCE AND THE HAUSDORFF DISTANCE; DISTANCES ARE AVERAGED OVER ALL FRAMES OF A SEQUENCE. SEE SECTION IV-B FOR FURTHER DETAILS

Internal proximal carotid	lum.-int. (mean dist.)	lum.-int. (Hausd. dist.)	med.-adv. (mean dist.)	med.-adv. (Hausd. dist.)
Exp. 1 vs Exp. 2	0.33 ± 0.17mm	0.73 ± 0.33mm	0.40 ± 0.29mm	0.88 ± 0.67mm
Exp. 1 vs MGD	0.32 ± 0.14mm (0.60)	0.83 ± 0.46mm (0.25)	0.33 ± 0.23mm (0.75)	0.89 ± 0.65mm (0.48)
Exp. 2 vs MGD	0.32 ± 0.18mm (0.56)	0.85 ± 0.48mm (0.22)	0.33 ± 0.22mm (0.76)	0.82 ± 0.54mm (0.61)
Exp. 1 vs MED	0.33 ± 0.16mm (0.49)	0.86 ± 0.46mm (0.19)	0.44 ± 0.32mm (0.34)	1.07 ± 0.74mm (0.23)
Exp. 2 vs MED	0.32 ± 0.18mm (0.57)	0.86 ± 0.49mm (0.20)	0.37 ± 0.30mm (0.59)	0.90 ± 0.65mm (0.46)

most often pre-Rician (low density of scatterers) and sometimes generalized Rician (coherent components with spacing larger than the pulse width), and the intima-media distribution is on average Rician (coherent components with spacing smaller than the pulse width) but presents the greatest variability in shape among the three distributions. The results concerning the lumen and the adventitia distributions seem to be in agreement with anatomical information. Namely, in the carotid, the erythrocytes can be viewed as random scatterers in high density; the adventitia is primarily fibrocellular [56, p. 254], and is composed of loose connective tissue [56, p. 143], whose cells can be viewed as random scatterers in low density together with a resolved coherent component. As for the intima-media distribution, the intima consists of a monolayer of endothelial cells [56, p. 253], that can be viewed as random scatterers in high density together with an unresolved coherent component, whereas the media is composed of smooth muscular cells, elastic fibers, and collagen [56, p. 253]. Thus, the IMT offers a greater variability in the configuration of its underlying scatterers, hence perhaps explaining the greater variability in its echogenic statistical properties. Finally, the results of Tables IV and V indicate that the MED model performs less than the MGD, in particular at the interface between the media and the adventitia. For this interpretation, we adopt the practitioner's point of view that a higher p-value in a hypothesis test gives a higher confidence level in the null hypothesis tested (but see [52, Sec. 5.3], for further discussion).

B. Validation of the Segmentations

Using a significance level of $\alpha = 0.05$, one concludes from the experimental results reported in Tables IV and V that the mean distance between the semi-automatic segmentations obtained by the proposed method and the manual segmentations performed by either expert, is not significantly more than the mean distance between the manual segmentations of the two experts, in the case of disease-free common and internal proximal carotid arteries.

These conclusions hold for the model of MGDs (or equivalently, Nakagami distributions, after a change of variable). Tables IV and V indicate that the p-values are somewhat lower when using the simpler model of mixtures of exponential distributions (MED) (or equivalently, Rayleigh distributions after the same change of variable), especially for the media-adventitia interface.

The same conclusions hold for the Hausdorff distance, but the p-values are lower than for the mean distance. Nevertheless, for the common carotid arteries, the distance between the semi-automatic segmentation and either manual segmentation is at most

0.01 mm more than the distance (0.45 mm) between the two manual segmentations, on average for the lumen-intima interface, and the distance between the semi-automatic segmentation and either manual segmentation is at most 0.03 mm more than the distance (0.38 mm) between the two manual segmentations, on average for the media-adventitia interface. For the internal proximal carotid arteries, the distance between the semi-automatic segmentation and either manual segmentation is at most 0.1 mm more than the distance (0.73 mm) between the two manual segmentations, on average for the lumen-intima interface and the distance between the semi-automatic segmentation and either manual segmentation is at most 0.01 mm more than the distance (0.88 mm) between the two manual segmentations, on average for the the media-adventitia interface. For the simpler MED model, the results obtained are worse, as indicated in Tables IV and V.

It appears that there is a greater variability between the two experts in the case of the internal proximal carotid than in the case of the common carotid. In particular, there was a disagreement on the presence of a plaque or not. But then, the semi-automatic segmentations appear to be within that range of variability, especially when using the mean distance and the MGD model. Thus, we conclude that the semi-automatic segmentation method is overall no less reliable than the manual segmentations performed by the expert technicians. We also conclude that the proposed MGD model is significantly better than the simpler MED model.[6]

Assumptions On the Tissue Echogenicity

In the examples where the proposed method performed well, the assumptions made in Section III-C-II on the echogenicity of the tissues were for the most part verified. On the other hand, in the examples where the proposed method did not perform well, the main cause of difficulty, in our opinion, might be the fact that these assumptions were not quite satisfied. On the other hand, the analysis presented in Tables IV and V indicates that the proposed method is nevertheless robust to the estimation procedure, in the sense of Section VI-B.

C. Computation Time

Although the average computation time is 24 s per frame (14 h and 41 min for 2227 frames), further improvements can be made on a commercial version, provided one is given a few CPUs. We think that 1–3 s might be enough to treat one frame, with a multithread implementation. In fact, the 10 runs of the EM algorithm are applied independently on each vertical strip. Also, in the ES algorithm, each of the 30 particles explores independently a new solution in its neighborhood. So, the clinician might carry on the patient examination, while the segmentation of the whole sequence is performed within 1–4 min. Thus, there is a reasonable hope to see a clinically applicable version in the future.

[6]This statistical test is one of the few ones for which the Bayes factor can be computed directly from the p-value: $B_{0,1} = p/(1-p)$ (cf. [57]). Thus, in a Bayesian framework, one accepts the null hypothesis if and only if $B_{0,1} \geq 1$, i.e., $p \geq 1/2$, when both kinds of decision errors are equally weighted. So in the case of this test, the Bayesian approach (that amounts to using $\alpha = 1/2$) is more severe than the frequentist approach (that uses $\alpha = 0.05$ [58] for historical reasons). If one prefers to use $\alpha = 1/2$, then our conclusion holds only for the mean distance and the MGD model. In any case, we have reported the p-values.

VII. Conclusion

In the sequences of *B*-mode images tested, the model of mixtures of Nakagami distributions is more conform to the true statistical distributions than the simpler model of mixture of Rayleigh distributions. The parameters of the proposed model can be efficiently and reliably estimated using the EM algorithm. The test reported here suggest that the semi-automatic segmentations obtained by the proposed method are within the variability of the manual segmentations of two experts in the case of disease-free carotids. We are currently developing an adaptation of our method to carotids presenting pathologies.

Appendix

In this appendix, the EM algorithm is explained in the case of a mixture of gamma distributions.

As usual, it is convenient to introduce a latent variable $c \in \{1, 2 \ldots, \ell\}$, that indicates which kernel is chosen. The MGD model is then equivalent to the hierarchical model

$$\begin{aligned} c &\sim \mathcal{M}(p_1, \ldots, p_\ell) \\ I|c = i &\sim \mathcal{G}(k_i, \theta_i) \end{aligned} \tag{26}$$

where $\mathcal{M}$ denotes the multinomial distribution.

Let $\chi^{(t)}$ be the current estimation of the parameters $(p_1, \ldots, k_1, \ldots, \theta_1, \ldots)$. Let $\tilde{c}$ be the configuration of latent labels (c_j). Let $\tilde{I} = (I_1, \ldots, I_N)$ be i.i.d. samples of a mixture of gamma distributions. In the EM algorithm, one considers the expectation

$$Q(\chi|\chi^{(t)}, \tilde{I}) = \sum_{\tilde{c}} \log \Big\{ f(\tilde{I}, \tilde{c}|\chi)\pi(\chi) \Big\} p(\tilde{c}|\chi^{(t)}, \tilde{I}). \tag{27}$$

Up to a constant, this is equal to

$$\sum_{j=1}^{N} \sum_{i=1}^{\ell} \Big(\log \mathcal{G}(I_j|k_i, \theta_i) + \log p_i \Big) p(c_j = i) + \sum_{i=1}^{\ell} (A_0\alpha_i - 1) \log p_i \tag{28}$$

where $p(c_j = i)$ is a short-hand for $p(c_j = i|\chi^{(t)}, \tilde{I})$.

The E-step consists in computing $p(c_j = i)$. Using Bayes' Theorem, one obtains $p(c_j = i) = \Big(p_i \mathcal{G}(I_j|k_i, \theta_i) / \sum_{i=1}^{\ell} p_i \mathcal{G}(I_j|k_i, \theta_i) \Big)$.

For the M-step, the term $\lambda \left(1 - \sum_{i=1}^{\ell} p_i \right)$ corresponding to the constraint $\sum_{i=1}^{\ell} p_i = 1$ is added to (28). Setting the partial derivative with respect to p_i equal to 0, one obtains

$$\frac{1}{p_i} \left(\sum_{j=1}^{N} p(c_j = i) + A_0\alpha_i - 1 \right) - \lambda = 0. \tag{29}$$

Setting $P_i = \sum_{j=1}^{N} p(c_j = i)$ yields $P_i + A_0\alpha_i - 1 = p_i\lambda$. In particular, $\lambda = N + A_0 - \ell$. This gives the expression $p_i = (P_i + A_0\alpha_i - 1)/(N + A_0 - \ell)$. Moreover, setting the partial derivatives with respect to k_i and θ_i equal to 0, one obtains, respectively, the equations

$$\sum_{j=1}^{N} (-\log \theta_i - \psi(k_i) + \log I_j)\, p(c_j = i) = 0; \tag{30}$$

$$\sum_{j=1}^{N}\left(-\frac{k_i}{\theta_i}+\frac{I_j}{\theta_i^2}\right)p(c_j=i)=0. \qquad (31)$$

Setting $Q_i=\sum_{j=1}^{N} I_j p(c_j=i)$, (31) implies the identity $\theta_i=(Q_i/k_iP_i)$. Substituting back into (30), and setting $R_i=\sum_{j=1}^{N}\log I_j p(c_j=i)$, one obtains the updating expression for k_i: $\log k_i-\psi(k_i)=\log(Q_i/P_i)-(R_i/P_i)$, where $\psi(x)$ denotes the digamma function $\Gamma'(x)/\Gamma(x)$. Now, the function $\log$ being concave on $(0,\infty)$, it follows from Jensen's inequality [59] that $\log(Q_i/P_i)-(R_i/P_i)\geq 0$, where equality holds if and only if all I_j are mutually equal. This latter case happening with probability 0, it follows that $\log(Q_i/P_i)-(R_i/P_i)>0$ with probability 1. Since $\log x-\psi(x)$ is a decreasing function on $(0,\infty)$, $\lim_{x\to 0}\log x-\psi(x)=\infty$ and $\lim_{x\to\infty}\log x-\psi(x)=0$, one concludes that $k_i>0$ can be found by a binary search, unless all data elements are mutually equal.

ACKNOWLEDGMENT

The authors would like to thank C. Schmitt, J. Fromageau, and R. L. Maurice for useful discussions on ultrasonic imagery, as well to Z. Qin and I. Renaud for technical assistance in performing the manual segmentations. The authors would like to thank the anonymous reviewers for their helpful comments on the presentation of this work.

REFERENCES

[1] P.-J. Touboul, M. G. Hennerici, S. Meairs, H. Adams, P. Amarenco, N. Bornstein, L. Csiba, M. Desvarieux, S. Ebrahim, M. Fatar, R. H. Hernandez, M. Jaff, S. Kownator, P. Prati, T. Rundek, M. Sitzer, U. Schminke, J.-C. Tardif, A. Taylor, E. Vicaut, K. K. S. Woo, F. Zannad, and M. Zureik, "Mannheim carotid intima-media thickness consensus (2004–2006)," *Cerebrovascular Diseases*, vol. 23, no. 1, pp. 75–80, 2007.

[2] C. Pellot-Barakat, F. Frouin, M. F. Insana, and A. Herment, "Ultrasound elastography based on multiscale estimations of regularized displacement fields," *IEEE Trans. Med. Imag.*, vol. 23, no. 2, pp. 153–163, Feb. 2004.

[3] J. Dias and J. Leitao, "Wall position and thickness estimation from sequences of echocardiograms images," *IEEE Trans. Med. Imag.*, vol. 15, no. 1, pp. 25–38, Feb. 1996.

[4] C. Haas, H. Ermert, S. Holt, P. Grewe, A. Machraoui, and J. Barmeyer, "Segmentation of 3-D intravascular ultrasonic images based on a random field model," *Ultrasound Med. Biol.*, vol. 26, no. 2, pp. 297–306, 2000.

[5] M. Mignotte, J. Meunier, and J.-F. Tardif, "Endocardial boundary estimation and tracking in echocardiographic images using deformable templates and Markov random fields," *Pattern Anal. Appl.*, vol. 4, pp. 256–271, 2001.

[6] M. -H. R. Cardinal, J. Meunier, G. Soulez, E. Thérasse, and G. Cloutier, "Intravascular ultrasound image segmentation: A fast-marching method," in *Medical Image Computing and Computer Assisted Intervention*, ser. Lecture Note Computer Science. Berlin, Germany: Springer-Verlag, 2003, pp. 432–439.

[7] M.-H. R. Cardinal, J. Meunier, G. Soulez, R. L. Maurice, E. Thérasse, and G. Cloutier, "Intravascular ultrasound image segmentation: A three-dimensional fast-marching method based on gray level distributions," *IEEE Trans. Med. Imag.*, vol. 25, no. 5, pp. 590–601, May 2006.

[8] E. Brusseau, C. L. de Korte, F. Mastik, J. Schaar, and A. F. W. van der Steen, "Fully automatic luminal contour segmentation in intracoronary ultrasound imaging—A statistical approach," *IEEE Trans. Med. Imag.*, vol. 23, no. 5, pp. 554–566, May 2004.

[9] M.-H. R. Cardinal, J. Meunier, G. Soulez, R. L. Maurice, and G. Cloutier, "Automatic 3-D segmentation of intravascular ultrasound images using region and contour information," in *Medical Image Computing and Computer Assisted Intervention*, ser. Lecture Note Computer Science, J. S. Duncan and G. Gerig, Eds. New York: Springer-Verlag, 2005, vol. 3749, pp. 319–326.

[10] R. F. Wagner, S. W. Smith, J. M. Sandrick, and H. Lopez, "Statistics of speckle in ultrasound b-scans," *IEEE Trans. Sonics Ultrason.*, vol. 30, no. 3, pp. 156–163, May 1983.

[11] A. Dempster, N. Laird, and D. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. R. Stat. Soc. (Ser. B)*, vol. 39, pp. 1–38, 1977.

[12] M. F. Insana, R. F. Wagner, B. S. Garra, D. G. Brown, and T. H. Shawker, "Analysis of ultrasound image texture via generalized Rician statistics," *Opt. Eng.*, vol. 25, no. 6, pp. 743–748, 1986.

[13] P. M. Shankar, J. M. Reid, H. Ortega, C. W. Piccoli, and B. B. Goldberg, "Use of non-Rayleigh statistics for the identification of tumors in ultrasonic B-scans of the breast," *IEEE Trans. Med. Imag.*, vol. 12, no. 4, pp. 687–692, Dec. 1993.

[14] R. D. Lord, "The use of the Hankel transform in statistics. I," *Biometrika*, vol. 41, pp. 44–55, 1954.

[15] E. Jakeman and P. N. Pusey, "A model for non-Rayleigh sea echo," *IEEE Trans. Antennas Propag.*, vol. 24, no. 6, pp. 806–814, Nov. 1976.

[16] P. M. Shankar, "A general statistical model for ultrasonic scattering from tissues," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 47, no. 3, pp. 727–736, May 2000.

[17] M. Nakagami, "The m distribution—A general formula of intensity distribution in rapid fading," in *Statistical Methods on Radio Wave Propagation*, W. C. Hoffman, Ed. New York: Pergamon, 1960, pp. 3–36.

[18] P. M. Shankar, "A compound scattering pdf for the ultrasonic echo envelope and its relationship to K and Nakagami distributions," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 50, no. 3, pp. 339–343, Mar. 2003.

[19] E. Jakeman and R. J. A. Tough, "Generalized k distribution: a statistical model for a weak scattering," *J. Opt. Soc. Am. A*, vol. 4, pp. 1764–1772, 1987.

[20] Karmeshu and R. Agrawal, "Study of ultrasound echo envelope based on Nakagami-inverse Gaussian distribution," *Ultrasound Med. Biol.*, vol. 32, no. 3, pp. 371–376, 2006.

[21] Z. Tao, H. D. Tagare, and J. D. Beaty, "Evaluation of four probability distribution models for speckle in clinical cardiac ultrasound images," *IEEE Trans. Med. Imag.*, vol. 25, no. 11, pp. 1483–1491, Nov. 2006.

[22] Z. Tao, C. C. Jaffe, and H. D. Tagare, "Tunneling descent: A new algorithm for active contour segmentation of ultrasound images," in *Information Processing in Medical Imaging 2003*, ser. Lecture Notes in Computer Science. New York: Springer, 2003, vol. 2732, pp. 246–257.

[23] R. W. Prager, A. H. Gee, G. M. Treece, and L. H. Berman, "Decompression and speckle detection for ultrasound images using the homodyned k-distribution," *Pattern Recognit. Lett.*, vol. 24, no. 4–5, pp. 705–713, 2003.

[24] M. B. Hansen, J. Moller, and F. A. Togersen, "Bayesian contour detection in a time series of ultrasound images through dynamic deformable template models," *Biostatistics*, vol. 3, no. 2, pp. 213–228, 2002.

[25] O. François, "Global optimization with exploration/selection algorithms and simulated annealing," *Ann. Appl. Probab.*, vol. 12, no. 1, pp. 248–271, 2002.

[26] J. A. Noble and D. Boukerroui, "Ultrasound image segmentation: A survey," *IEEE Trans. Med. Imag.*, vol. 25, no. 8, pp. 987–1010, Aug. 2006.

[27] D. M. Titterington, A. F. M. Smith, and U. E. Makov, *Statistical Analysis of Finite Mixture Distributions*. New York: Wiley, 1985.

[28] H. Teicher, "Identifiability of finite mixtures," *Ann. Math. Stat.*, vol. 34, pp. 1265–1269, 1963.

[29] E. Jakeman, "k -distributed noise," *J. Opt. A: Pure Appl. Opt.*, vol. 1, pp. 784–789, 1999.

[30] R. F. Wagner, M. F. Insana, and D. G. Brown, "Statistical properties of radio-frequency and envelope-detected signals with applications to medical ultrasound," *J. Opt. Soc. Am. A*, vol. 4, no. 5, pp. 910–922, 1987.

[31] M. Escobar, "Estimating normal means with a Dirichlet process prior," *J. Am. Stat. Assoc.*, vol. 89, no. 425, pp. 268–277, 1993.

[32] R. Redner, "Note on the consistency of the maximum likelihood estimate for nonidentifiable distributions," *Ann. Stat.*, vol. 9, no. 1, pp. 225–228, 1981.

[33] M. Tanner, *Tools for Statistical Inference: Observed Data and Data Augmentation Methods*, ser. Lecture Notes in Statistics. New York: Springer-Verlag, 1991, vol. 67.

[34] J. M. Meinders, L. Kornet, and A. P. G. Hoeks, "Assessment of spatial inhomogeneities in intima media thickness along an arterial segment using its dynamic behavior," *Am. J. Physiol.—Heart Circulatory Physiol.*, vol. 285, pp. H384–H391, 2003.

[35] R. Kazmierski, C. Watala, M. Lukasik, and W. Kozubski, "Common carotid artery remodeling studied by sonomorphological criteria," *J. Neuroimag.*, vol. 14, no. 3, pp. 258–264, 2004.

[36] D. Madigan, A. E. Raftery, C. T. Volinsky, and J. A. Hoeting, "Bayesian model averaging," in *Integrating Multiple Learned Models, (IMLM-96)*, P. Chan, S. Stolfo, and D. Wolpert, Eds. : , 1996, pp. 77–83.

[37] J. A. Hoeting, D. Madigan, A. E. Raftery, and C. T. Volinsky, "Bayesian model averaging: A tutorial (with discussion)," *Stat. Sci.*, vol. 14, pp. 382–401, 1999.

[38] J. A. Hoeting, D. Madigan, A. E. Raftery, and C. T. Volinsky, "Correction to: Bayesian model averaging: A tutorial (with discussion)," *Stat. Sci.*, vol. 15, pp. 193–195, 1999.

[39] R. Muzzolini, Y. H. Yang, and R. Pierson, "Multiresolution texture segmentation with application to diagnistic ultrasound images," *IEEE Trans. Med. Imag.*, vol. 12, no. 1, pp. 108–123, Mar. 1993.

[40] N. Friedland and D. Adam, "Automatic ventricular cavity boundary detection from sequential ultrasound images using simulated anneal," *IEEE Trans. Med. Imag.*, vol. 8, no. 4, pp. 344–353, Dec. 1989.

[41] J. Besag, "On the statistical analysis of dirty pictures," *J. R. Statist. Soc. B*, vol. 48, no. 3, pp. 259–302, 1986.

[42] D. Boukerroui, A. Baskurt, J. A. Noble, and O. Basset, "Segmentation of ultrasound images—multiresolution 2-d and 3-D algorithm based on global and local statistics," *Pattern Recognit. Lett.*, vol. 24, no. 4–5, pp. 779–790, 2004.

[43] O. François, "An evolutionary strategy for global minimization and its Markov chain analysis," *IEEE Trans. Evol. Comput.*, vol. 2, pp. 77–90, 1998.

[44] F. Destrempes, "Estimation de paramètres de champs Markoviens cachés avec applicationsà la segmentation d'images et la localisation de formes," Ph.D. dissertation, Univ. Montréal, Montreal, QC, Canada, 2006.

[45] F. Destrempes, "Détection non-supervisée de contours et localisation de formesà l'aide de modèles statistiques," M.S. thesis, Univ. Montréal, Montreal, QC, Canada, 2002.

[46] F. Destrempes, M. Mignotte, and J.-F. Angers, "A stochastic method for Bayesian estimation of hidden Markov random field models with application to a color model," *IEEE Trans. Image Process.*, vol. 14, no. 8, pp. 1096–1124, Aug. 2005.

[47] F. Destrempes, J.-F. Angers, and M. Mignotte, "Fusion of hidden Markov random field models and its Bayesian estimation," *IEEE Trans. Image Process.*, vol. 15, no. 10, pp. 2920–2935, Oct. 2006.

[48] F. Destrempes, M. Mignotte, and J.-F. Angers, "Localization of shapes using statistical models and stochastic optimization," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 9, pp. 1603–1615, Sep. 2007.

[49] F. Destrempes and M. Mignotte, "Unsupervised localization of shapes using statistical models," in *4th IASTED Int. Conf. Signal Image Process.*, Kaua'i , HI, Aug. 2002, pp. 66–71.

[50] S. Benameur, M. Mignotte, H. Labelle, and J. A. de Guise, "A hierarchical statistical modeling approach for the unsupervised 3D biplanar reconstruction of the scoliotic spine," *IEEE Trans. Biomed. Eng.*, vol. 52, no. 12, pp. 2041–2057, Dec. 2005.

[51] S. Benameur, M. Mignotte, F. Destrempes, and J. A. de Guise, "3D biplanar reconstruction of scoliotic rib cage using the estimation of a mixture of probabilistic prior models," *IEEE Trans. Biomed. Eng.*, vol. 52, no. 10, pp. 1713–1728, Oct. 2005.

[52] C. P. Robert, *The Bayesian Choice*, 2nd ed. New York: Springer Verlag, 2001.

[53] H. Jeffreys, "An invariant form for the prior probability in estimation problems," *Proc. R. Soc. London (Ser. A)*, vol. 186, pp. 453–461, 1946.

[54] H. Jeffreys, *Theory of Probabiliy*, 3rd ed. Oxford, U.K.: Oxford Univ., 1961.

[55] D. P. Huttenlocher, G. A. Klanderman, and W. J. Rucklidge, "Comparing images using the Hausdorff distance," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 15, no. 9, pp. 850–863, Sep. 1993.

[56] W. J. Zwiebel and J. S. Pellerito, *Introduction to Vascular Ultrasonography*, 5th ed. New York: Elsevier Saunders, 2005.

[57] J. O. Berger, *Statistical Decision Theory and Bayesian Analysis*, 2nd ed. New York: Springer-Verlag, 1985.

[58] R. A. Fisher, *Statistical Methods and Scientific Inference*. Edinburgh, U.K.: Oliver Boyd, 1956.

[59] W. Rudin, *Real and Complex Analysis*. New York: McGraw-Hill, 1987.