Iterative Approximation of Empirical Grey-Level Distributions for Precise Segmentation of Multimodal Images

Ayman El-Baz

CVIP Laboratory, Department of Electrical and Computer Engineering, University of Louisville, Louisville, KY 40292, USA
Email: elbaz@cairo.spd.louisville.edu

Aly A. Farag

CVIP Laboratory, Department of Electrical and Computer Engineering, University of Louisville, Louisville, KY 40292, USA
Email: farag@cvip.uofl.edu

Georgy Gimel'farb

Centre for Image Technology and Robotics (CITR), Department of Computer Science, University of Auckland, Tamaki Campus, Auckland 1000, New Zealand
Email: g.gimelfarb@auckland.ac.nz

A new algorithm for segmenting a multimodal grey-scale image is proposed. The image is described as a sample of a joint Gibbs random field of region labels and grey levels. To initialize the model, a mixed multimodal empirical grey-level distribution is approximated with linear combinations of Gaussians, one combination per region. Bayesian decisions involving expectation maximization and genetic optimization techniques are used to sequentially estimate and refine parameters of the model, including the number of Gaussians for each region. The final estimates are more accurate than with conventional normal mixture models and result in more adequate region borders in the image. Experiments show that the proposed technique segments complex multimodal medical images of different types more accurately than several other known algorithms.

Keywords and phrases: density estimation, image segmentation, expectation maximization.

1. INTRODUCTION

A large number of image segmentation methods based on estimating marginal probability densities of signals and separating their dominant modes have been developed and tested for the last three decades (see [1, 2, 3, 4, 5] to cite a few). However many important applications such as medical image analysis or industrial vision still encounter difficulties in separating practically meaningful continuous or disjoint objects, even when signal densities are distinct to the point where their mixture becomes strongly multimodal. The basic issue is the accuracy of region borders, which are usually essential for correct interpretation of the objects. Typically, the fine separation of the signal modes to specify the region borders is obtained by intersecting tails of the signal distributions for the adjacent objects. Therefore, it is the tails that have to be precisely estimated in order to separate, for example, a darker object from a brighter background. One of the practical problems that inspired our approach is to accurately detect lungs in a spiral CT chest slice so that their borders closely match those outlined by a radiologist.

Because there always exists an overlap between the signal ranges for the different objects, the precise segmentation has to account for spatial distributions of the signals, too. Markov-Gibbs random field models show considerable promise in spatial image analysis [6, 7, 8, 9, 10, 11, 12, 13]. Thus we consider images to be segmented as samples of a two-level Gibbs random field of more or less continuous regions (the higher level) and grey levels in each region (the lower level) [7, 9, 10, 11]. We choose for each level the simplest probability models ensuring the necessary precision of the segmentation. At the lower level, the signals are described by a conditionally independent random field of grey levels having different marginal probability distributions in the regions. In practice, these latter distributions are quite intricate. Hence we approximate the probability density function for each region with a linear combination of both positive and negative Gaussians [14, 15, 16] which is more accurate

than a conventional normal mixture with only the positive components. Then a mixed empirical signal distribution over the image is approximated with a mixture of several linear combinations of Gaussians. Such a lower-level segmentation on the basis of the estimated probability densities is refined further at the higher level using the Bayesian maximum a posteriori (MAP) estimation of the segmentation map based on the joint Gibbs model of the maps and images.

We propose a new sequential EM-based algorithm that closely approximates a given multimodal empirical probability distribution of signals by estimating parameters of the linear combination of Gaussians for each region. Below for simplicity, we restrict our consideration to only a bimodal signal density describing a dark object and its bright background. But the extension of the proposed segmentation scheme onto the multimodal case is straightforward. In the bimodal case, the empirical mixed density is split first into two dominant positive Gaussians and a set of secondary alternating Gaussians. These latter describe deviations of the empirical distribution from the dominant components. The number of the dominant modes (i.e., the number of different types of regions to detect) is assumed to be known, but the number of the secondary terms is found on the basis of the approximation error.

The parameters of both the dominant and secondary terms are sequentially estimated using the EM-algorithm [17, 18, 19, 20, 21], or more specific, its early variant for the normal mixtures [22] (see also [21]). To initially segment the image, the approximated mixed density is divided further into two linear combinations of Gaussians (for the darker object and the brighter background, resp.) which yield the minimum classification error on the intersecting tails of the distributions.

At the higher level, the label intra- and interregion co-occurrences are specified by an auto-binomial Gibbs model with the nearest 8-neighborhood [11, 13]. The region map after an initial low-level classification is refined iteratively by repeating the following steps: (i) updating the higher-level map, given the low-level image model, and (ii) updating this latter, given the region map.

The paper is organized as follows. Section 2 introduces the Gibbs image model and describes the low-level and high-level stages of the proposed segmentation algorithm. Section 3 describes experiments with different types of medical images to show that the proposed technique produces region borders close to the "ground truth" given by the experts—radiologists—and is more accurate than several other segmentation algorithms.

2.   TWO-LEVEL GIBBS IMAGE MODEL

Let $\mathbf{R} = \{(x, y) : 1 \leq x \leq X, \ 1 \leq y \leq Y\}$ be a finite grid supporting images $g : \mathbf{R} \to \mathbf{Q}$ and their region maps $l : \mathbf{R} \to \mathbf{K}$. Here $\mathbf{Q} = \{0, \ldots, 255\}$ and $\mathbf{K} = \{1, \ldots, K\}$ are the sets of grey levels $q \in \mathbf{Q}$ and region labels $k \in \mathbf{K}$, respectively (for the bimodal case, $\mathbf{K} = \{1, 2\}$).

The two-level Gibbs image model is specified by a joint probability distribution $P(g, l) = P(g|l)P(l)$, where $P(l)$ is a prior higher-level Gibbs distribution of maps and $P(g|l)$ is a lower-level conditional distribution of images, given the map. Both the distributions are strictly positive. The MAP estimate $l^* = \arg\max_l L(g, l)$ of the map, given the image g, maximizes the joint log-likelihood function:

$$L(g, l) = \log P(g|l) + \log P(l). \tag{1}$$

To make the choice computationally feasible, a local maximum of the log likelihood of (1) is usually searched for iteratively by estimating or reestimating first the lower-level model and then using it to update the higher-level one. The process terminates when parameters of the current and pervious estimated models become equal to a value within a given accuracy range [7, 9, 10, 16].

Such an iterative maximization using the lower-level model in an explicit form and the higher-level model with numerically approximated components having no closed-form representation is implemented below.

2.1.   Low-level density model

To most accurately specify the lower-level model, we approximate the marginal grey-level probability density function in each region $k = 1, 2, \ldots, K$ with a linear combination of C_k Gaussians [14, 15]: for each $q \in \mathbf{Q}$,

$$p(q|k) = \sum_{i=1}^{C_k} w_{k,i}\varphi(q|\theta_{k,i}); \qquad \int_{-\infty}^{\infty} p(q|k)dq = 1. \tag{2}$$

Here, in contrast to the more conventional normal mixture models, the weights $w_{k,i}$ may be both positive and negative and have only one obvious restriction in line with (2): $\sum_{i=1}^{C_k} w_{k,i} = 1$. The weights now are not prior probabilities, and the combination in (2) is simply an approximation of the probability density function depending on parameters $w_{k,i}$, $\theta_{k,i}$ denoting the weight and the mean and variance of each component, respectively.

In the general case, the probability densities form only a proper subset of the set of all linear combinations in (2). The weights in the subset are limited as to keep nonnegative values of the combined densities to within the actual signal range, for example, $[0, 255]$ for the grey levels. This restriction leads to strongly interdependent parameters, but in our particular case, their interdependence may be ignored. We use the model to only approximate the intersecting tails of the empirical distributions, so its behavior outside the domain of close approximation (and the associated restrictions on the parameters) is of no concern. Moreover, the maximization of the likelihood function is also directed towards keeping the positive density values that approximate the empirical ones. The mixture of the K distributions $p(q|1), p(q|2), \ldots, p(q|K)$ in (2) has just the same form but with the larger number of the components

$(C = C_1 + C_2 + \cdots + C_K$ if all the values $\theta_{k,i}$ differ for both distributions):

$$p(q) = \sum_{c=1}^{C} w_c \varphi(q|\theta_c), \qquad (3)$$

where all the mixed elements do not relate to the regions and thus have only one subscript c. In principle, all the parameters $\mathbf{w} = \{w_c; c = 1, \ldots, C\}$ and $\theta = \{\theta_c : c = 1, \ldots, C\}$, given the number of the components C, can be sequentially found using an EM algorithm modified to account for the alternating signs of the weights. This modification will be considered elsewhere. Here an alternative EM-based approach for describing the multimodal density of (3) is used. The EM algorithm is quite natural in this MAP framework because it maximizes the first term, $\log P(g|l)$, in the log likelihood of (1).

We assume that the number of objects (types of regions) to be separated, or what is the same, the number of the dominant modes, is known for all the images to be segmented. For simplicity, let the empirical distribution have two separate dominant modes that represent the object and the background, respectively, and each can be roughly approximated with a single Gaussian. Deviations of the empirical distribution from the two-component dominant Gaussian mixture are described by other components of (3). Therefore there are two dominant positive weights, say, w_1 and w_2, such that $w_1 + w_2 = 1$, and the "subordinate" weights with much smaller absolute values such that $w_3 + \cdots + w_C = 0$.

To estimate parameters of this latter model (both the weight and the parameters of each individual Gaussian) and get the initial region map, we propose the following sequential algorithm (with obvious modifications, it can be used also to approximate either a unimodal empirical distribution or a distribution with three and more dominant modes).

(1) Form the empirical mixed density function from the relative frequencies $f(q|g)$ of grey levels in all the pixels $(x, y) \in \mathbf{R}$, its integral over the $-\infty \leq q \leq \infty$ range being equal to the sum of all the frequencies:

$$F(g) = \left\{ f(q|g) : q \in \mathbf{Q}; \sum_{q \in \mathbf{Q}} f(q|g) = 1 \right\}. \qquad (4)$$

(2) Use the conventional EM-algorithm to approximate $F(g)$ with a dominant mixture P_2 of two Gaussians:

$$p_2(q) = w_1 \varphi(q|\theta_1) + w_2 \varphi(q|\theta_2). \qquad (5)$$

(3) Find the deviation Δ of $F(g)$ from P_2:

$$\Delta = \{\delta(q) = f(q|g) - p_2(q) : q \in \mathbf{Q}\} \qquad (6)$$

and split it into the positive and negative parts:

$$\begin{aligned} \Delta_{\mathrm{p}} &= \{\delta_{\mathrm{p}}(q) : q \in \mathbf{Q}\}, \\ \Delta_{\mathrm{n}} &= \{\delta_{\mathrm{n}}(q) : q \in \mathbf{Q}\} \end{aligned} \qquad (7)$$

such that

$$\delta(q) = \delta_{\mathrm{p}}(q) - \delta_{\mathrm{n}}(q);$$

$$\delta_{\mathrm{p}}(q) = \begin{cases} \delta(q) & \text{if } \delta(q) > 0, \\ 0 & \text{otherwise}; \end{cases} \qquad (8)$$

$$\delta_{\mathrm{n}}(q) = \begin{cases} 0 & \text{if } \delta(q) < 0, \\ -\delta(q) & \text{otherwise}. \end{cases}$$

(4) Compute the scaling factor for the deviations:

$$s = \int_{-\infty}^{\infty} \delta_{\mathrm{p}}(q) dq \equiv \int_{-\infty}^{\infty} \delta_{\mathrm{n}}(q) dq. \qquad (9)$$

(5) If $s \leq \tau$ (a given accuracy threshold), terminate the process and output the model $P = P_2$.

(6) Otherwise, consider the scaled-up absolute deviation $\Delta_{\mathrm{abs}} = 1/2s(\Delta_{\mathrm{p}} + \Delta_{\mathrm{n}})$ as an "empirical density" and use iteratively the EM-algorithm to find the number C_{sec} of Gaussians such that their mixture approximates best the scaled-up absolute deviation; this search for the number of the secondary components is detailed below.

(7) Scale down the obtained mixture model $P_{3,\ldots,C}$, that is, scale down its weights: $w_i \to s \cdot w_c; c = 3, \ldots, C = C_{\mathrm{sec}} + 2$, and change the signs of the individual weights in accord with the corresponding deviations to obtain the alternating scaled-down model $sP_{3,\ldots,C}$, where each Gaussian is assigned with the sign of the deviation closest to the mean value for that component.

(8) Center the model $P_{3,\ldots,C}$ in order to guarantee zero sum of its weights and output the model $P = P_2 + sP_{3,\ldots,C}$.

Instead of the last three steps, the scaled-up absolute deviations $(1/s)\Delta_{\mathrm{p}}$ and $(1/s)\Delta_{\mathrm{n}}$ can be separately approximated with the two normal mixtures having C_{p} and C_{n} components, respectively. Then the scaled-down submodels $P_{C_{\mathrm{p}}}$ and $P_{C_{\mathrm{n}}}$ are added to and subtracted from the model P_2, respectively, in order to output the desired model P. Both variants have produced almost the same final models in our experiments. While the latter variant is more theoretically justified, the former one is twice as fast.

Since the EM algorithm converges to a local maximum of the likelihood function, it may be repeated several times with different initial parameter values as to choose the model P yielding the best approximation. In principle, the process can be continued iteratively in order to approximate more and more closely the residual absolute deviations between $F(g)$ and P. But because each Gaussian in the model P impacts all the density values $p(q)$, the iterations should be terminated when the approximation quality begins to decrease.

We specify the approximation quality by the Kullback-Leibler divergence between the empirical and estimated densities in the points $q \in \mathbf{Q}$:

$$D(F(g), P) = \sum_{q \in \mathbf{Q}} f(q|g) \log \left(\frac{f(q|g)}{p(q)} \right). \qquad (10)$$

The above divergence is zero if $f(q|g) = p(q)$ for all $q \in \mathbf{Q}$, that is, for the exact approximation, otherwise it specifies how far is the log likelihood of the current model from its maximum value for the exact one. Let ε_q denote the relative approximation error such that $p(q) = (1 + \varepsilon_q) f(q|g)$. Then

$$
\begin{aligned}
D(F(g), P) &= - \sum_{q \in \mathbf{Q}} f(q|g) \log(1 + \varepsilon_q) \\
&\approx \frac{1}{2} \sum_{q \in \mathbf{Q}} f(q|g) \varepsilon_q^2
\end{aligned}
\tag{11}
$$

so that $0 \leq D(F(g), P) \leq (1/2) \max_{q \in \mathbf{Q}} \varepsilon_q^2$. The sequential process has to be terminated when the divergence of (10) is closest to zero or, what is almost the same, when the weighted sum of the squared relative errors has the minimum value.

The search for a number c of the Gaussians in the model is based on the integral absolute error $E(\Delta_{\mathrm{abs}}, P_{3,\ldots,C})$ between the scaled-up absolute deviation Δ_{abs} acting as the density function and its mixture model $P_{3,\ldots,C}$. The number c of the components is increasing sequentially by unit step until the error $E(\cdot)$ begins to increase. Once again, due to the multiple local maxima, the search has to be repeated several times with different initial parameter values as to select the best approximation.

Now the final mixed model P has to be split into the two parts by relating the subordinate components to either the object or the background so as to minimize the expected error of classification. Let the two dominant Gaussians with the means μ_1 and μ_2, $0 < \mu_1 < \mu_2 < q_{\max}$, correspond to the background and the object, respectively. Then the subordinate components having the mean values greater than μ_2 and lesser than μ_1 belong to the object and the background, respectively. Other components having the means in the interval $[\mu_1, \mu_2]$ are compared to a threshold t in that interval as to get the minimum classification error

$$
e(t) = \int_{-\infty}^{t} p(q|2) dq + \int_{t}^{\infty} p(q|1) dq.
\tag{12}
$$

2.2. High-level region model

We describe the region maps as samples of a simple autobinomial Gibbs model [13] such that a region label $l(x, y)$ in the pixel $(x, y) \in \mathbf{R}$ depends on only its four nearest neighbors $l(x + \xi_i, y + \eta_i)$; $(\xi_i, \eta_i) \in \mathbf{N} = \{[1, 0]; [1, 1]; [0, 1]; [1, 1]\}$; $i = 1, \ldots, 4$. The model is specified with the Gibbs probability distribution:

$$
P(l) = \frac{1}{Z} \exp \left(U_1(l) + \sum_{i=1}^{4} U_{2,i}(l) \right),
\tag{13}
$$

where Z is a normalizing factor (the partition function) and $U_1(l)$ and $U_{2,i}(l)$ denote partial energies of pixelwise and pairwise interactions of region labels. The energies are the sums of Gibbs potentials $\{V_1(k|\alpha) : k \in \mathbf{K}\}$ and $\{V_{2,i}(j, k|\lambda_i) : (j, k) \in \mathbf{K} \times \mathbf{K}\}$ depending, respectively, on the labels $k = l(x, y)$ in the individual pixels and on the label

cooccurrences $(j = l(x, y), k = l(x + \xi_i, y + \eta_i))$, $(\xi_i, \eta_i) \in \mathbf{N}$, in the pairs of similarly oriented neighbors:

$$
\begin{aligned}
U_1(l) &= \sum_{(x,y) \in \mathbf{R}} V_1(l(x, y)), \\
U_{2,i}(l) &= \sum_{(x,y) \in \mathbf{R}} V_{2,i}(l(x, y), l(x + \xi_i, y + \eta_i)).
\end{aligned}
\tag{14}
$$

Each partial energy can be represented in terms of the potentials and relative frequency distributions of the labels $F_{\mathrm{map}}(l) = \{f(k|l) : k \in \mathbf{K}\}$ and their cooccurrences $F_{i,\mathrm{map}}(l) = \{f(j, k|l) : (j, k) \in \mathbf{K} \times \mathbf{K}\}$ over the region map l. These distributions are sufficient statistics of the model and allow for obtaining rough analytical first approximations of the potentials. The derivation scheme in [11] modified to better fit our model results in the following first approximations of the initial potential values:

$$
\begin{aligned}
V_1(k) &= -\log f(k|l), \\
V_{2,i}(j, k) &= \gamma \big(f_i(j, k|l) - f(j|l) \cdot f(k|l) \big).
\end{aligned}
\tag{15}
$$

The scaling factor γ can be also computed using the distributions $F_{\mathrm{map}}(l)$ and $F_{i,\mathrm{map}}(l)$, $i = 1, \ldots, 4$.

Under the assumed continuity of and symmetric relationships between the two regions, all the potentials are bivalued:

$$
V_1(k) = \alpha \cdot k; \qquad V_{2,i}(j, k) =
\begin{cases}
\lambda_i, & j = k, \\
-\lambda_i, & j \neq k,
\end{cases}
\tag{16}
$$

with $\alpha \geq 0$ and $\lambda_i > 0$, $i = 1, \ldots, 4$.

To find the region map which maximizes the likelihood in (1), the potential estimates are refined using the genetic algorithm (GA) [23]. Because the partition function Z is unknown, it is approximated as proposed in [6]:

$$
Z \approx \sum_{(x,y) \in \mathbf{R}} \sum_{k \in \mathbf{K}} \exp \left(V_1(k) + \sum_{i=1}^{4} \sum_{j \in \mathbf{K}} V_{2,i}(k, j) \right).
\tag{17}
$$

To implement the GA optimization, the parameters α and λ_i, $i = 1, \ldots, 4$, in (16) are coded into a GA 20-bit "chromosome" using four bits per value. While the approximate log likelihood of (1) continues to increase, the following steps are repeated iteratively.

(1) Form a randomized population of $N = 30$ chromosomes from the top-ranked chromosomes for the previous iteration.

(2) Refine the initial region map using the Metropolis stochastic relaxation [13] with the potentials coded in each chromosome.

(3) Compute the log likelihood value for each chromosome using the refined map.

Therefore, the whole segmentation process is as follows.

Initialization: find an initial map by the Bayesian MAP classification of a given grey-scale image as shown in Section 2.1.
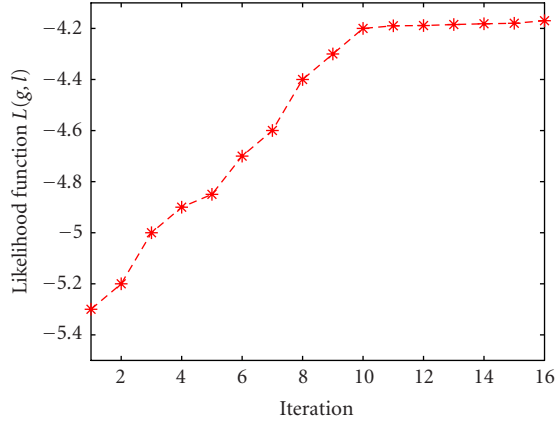
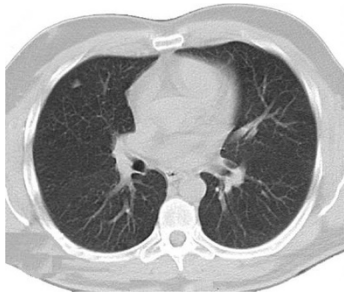Figure 1: Convergence of the proposed algorithm (for the experiment in Section 3, see Figure 2).

Figure 2: Typical slice from a chest spiral CT scan.

Iterative refinement: refine the initial map by repeating iteratively the following two steps:

(1) estimate the higher-level model which gives the maximum increase of the approximate log likelihood of the current region map, given the lower-level image model;

(2) recollect the empirical grey-level densities for the current regions, reapproximate these densities, and update the map using the pixelwise Bayesian classification.

Because at each step the approximate log-likelihood is greater than or equal to its previous value, the proposed algorithm converges to a locally optimum solution. Typical changes of the log-likelihood values in (1) at each iteration of the proposed algorithm are shown in Figure 1.

3. EXPERIMENTAL RESULTS

To assess robustness and computational performance, the proposed segmentation technique has been tested on three different medical imaging modalities. The images include axial human chest slices obtained by spiral-scan low-dose computer tomography (LDCT), axial human head slices obtained by time-of-flight magnetic resonance angiography (TOF-MRA), and axial human head slices obtained by magnetic resonance imaging (MRI). The two latter types were acquired with the Picker 1.5T Edge MRI scanner.
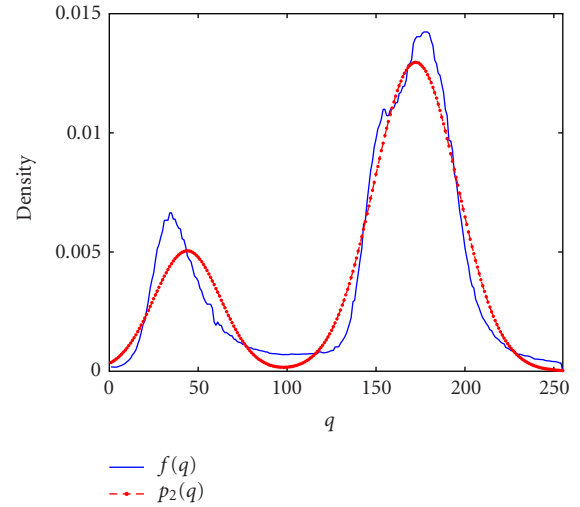

Figure 3: Empirical density $F(g) = (f(q) : q \in \mathbf{Q})$ versus the dominant mixture $P_2 = (p_2(q) : q \in \mathbf{Q})$.

Table 1: Initial (the upper row) and final (the bottom row) parameter estimates for the dominant Gaussians.

w_1	μ_1	σ_1^2	w_2	μ_2	σ_2^2
0.219	36.2	210.6	0.781	166.4	315.9
0.242	39.9	225.7	0.758	170.8	303.6

The TOF-MRA 512×512 and MRI 256×256 slices were 1 mm thick. The 8 mm thick LDCT slices were reconstructed every 4 mm with the scanning pitch of 1.5 mm.

In all the cases, the number of classes (i.e., dominant modes in the model) is specified by the user, and all other parameters are estimated by the proposed techniques. Section 3.1 discusses experiments with bimodal LDCT images. Experiments with three-modal MRA images and four-modal MRI images are described in Sections 3.2 and 3.3, respectively.

3.1. Lungs segmentation in LDCT images

We applied the proposed algorithm to a medical screening problem of separating lung tissues from the surrounding anatomical structures (e.g., chest, ribs, liver) in computer tomography (CT) images. The segmentation assumes that each CT slice has only two regions: the darker one (the lungs) and their brighter background. Because some lung tissues such as arteries, veins, bronchi, and bronchioles have grey levels close to those of the chest, the segmentation based on only the grey levels may lose some of these tissues. To obtain more accurate segmentation, our model accounts also for spatial relationships between the pixels.

Figure 2 shows a spiral CT slice for the chest, and its empirical grey-level density together with the dominant Gaussian mixture are presented in Figure 3. Parameters of the initial mixture are given in Table 1. The Levy distance [24] of
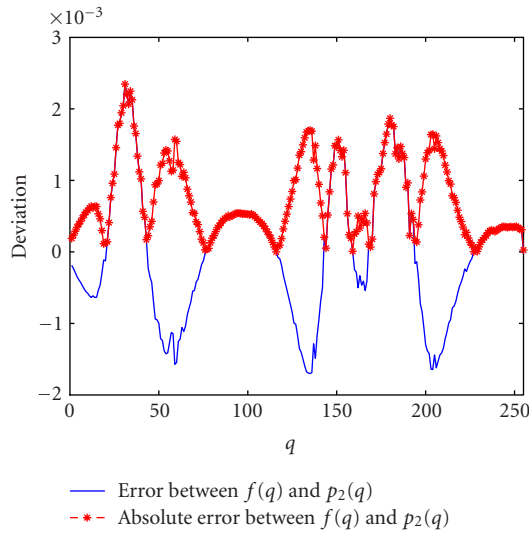
FIGURE 4: Deviations $\Delta(q) = f(q) - p_2(q)$ and the absolute deviations $|\Delta(q)|$ between the densities in Figure 3.
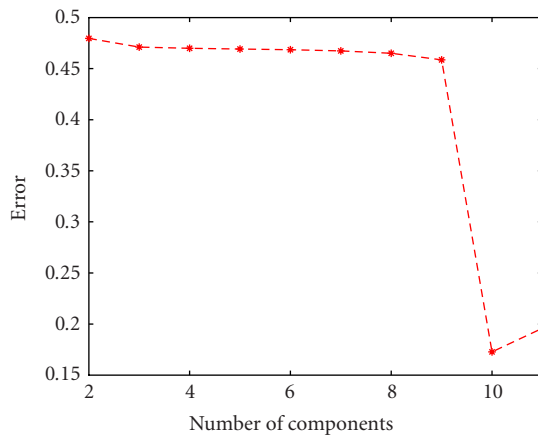

FIGURE 5: Error function $(E(\Delta_{abs}, P_{3,...,C}))$ versus the number C of the subordinate Gaussians approximating the scaled absolute deviation in Figure 4.
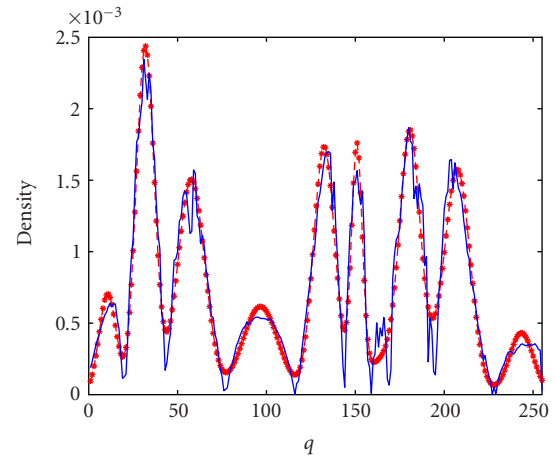

FIGURE 6: Subordinate mixture estimated for the absolute deviation in Figure 4.
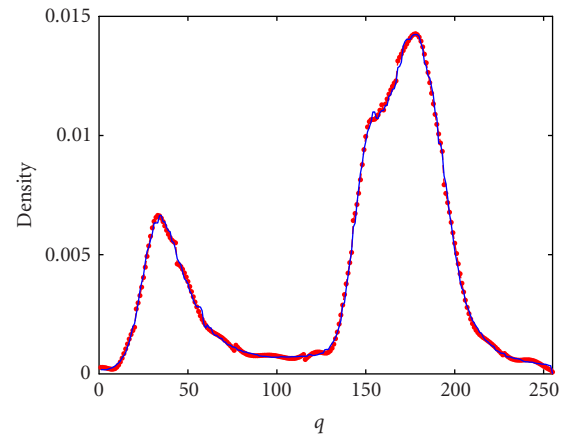

FIGURE 7: Empirical and estimated densities for the CT slice in Figure 2 (the linear combination of 12 Gaussians).
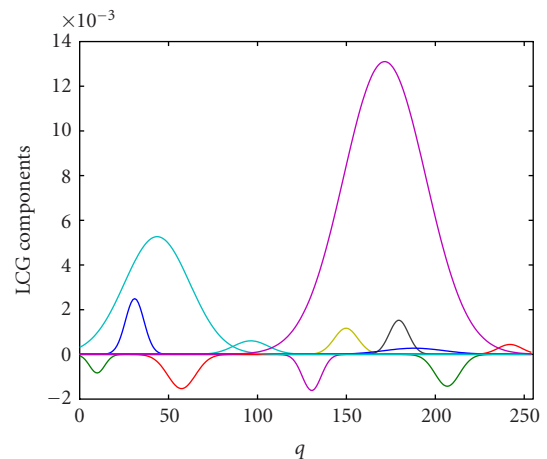
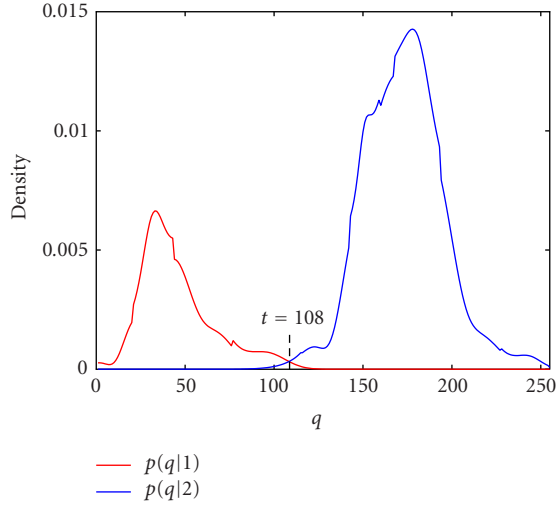

FIGURE 8: Components of the final density model P_{12}.

0.09 between these two distributions indicates a large mismatch. Figure 4 shows the scaled deviation of the dominant mixture from the empirical density. Our estimation of the number of components giving the minimum approximation error returns the ten Gaussians shown in Figure 5. Figure 6 presents the final result of this approximation.

Figure 7 shows the estimated density after the subordinate linear combination of Gaussians is added to the dominant mixture according to the deviation signs. The Levy distance between the empirical and final estimated distributions is now much smaller (0.02) than before (0.09).

Figure 8 shows the 12 components of the model. The Bayesian pixelwise classification is repeated for different partitions of the ten subordinate components. The minimum

FIGURE 9: Densities estimated for the lung and chest tissues in **Figure 2** (the minimum error 0.0045 for the threshold $t = 108$).

FIGURE 10: Approximation of the signal distribution for the CT slice in **Figure 2** by the conventional mixture of 12 Gaussians.
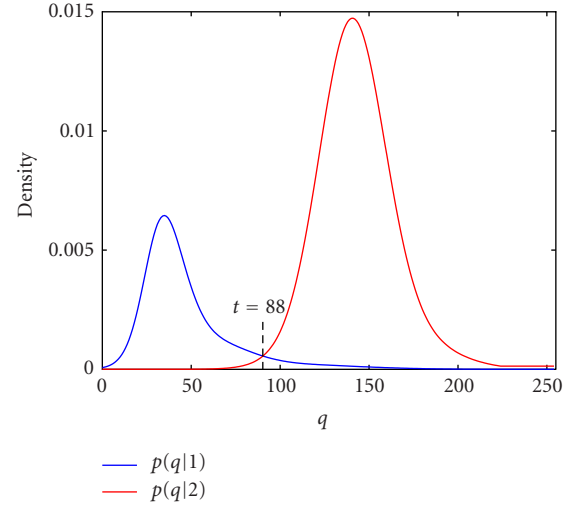

FIGURE 11: Densities for the lung and chest tissues in **Figure 2** estimated with the mixture of 12 Gaussians (the minimum error 0.013 for the threshold $t = 88$).
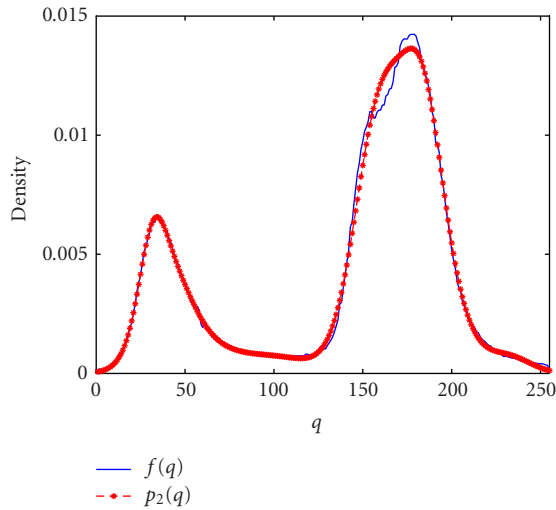
TABLE 2: The initial and final potentials for the map model for the LDCT image in **Figure 2**.

Parameters	α	λ_1	λ_2	λ_3	λ_4
Initial	1	0.81	0.71	0.49	0.51
Final	1	0.89	0.80	0.78	0.69

be related to only the object or the background. Therefore it comes as no surprise that our segmentation algorithm produces more accurate lung borders.

The potential estimates for the higher-level Gibbs model are obtained by maximizing the log likelihood in (1) using the genetic algorithm and the approximated partition function described in Section 2.2. The initial and final potentials are shown in Table 2. Table 1 shows the initial and final parameters for the dominant components. The final binary region map produced with these parameters is shown in Figure 12 together with the initial segmentation using the low-level model only, and the final region map obtained with both higher- and lower-level processing. For comparison, we show also the map refined by the Metropolis relaxation algorithm with the randomly selected parameters, the best segmentation of the same grey-level image obtained with the MRS algorithm [26] having the potential values 0.3 and three levels of resolution resulting in a notably larger error of 2.3%, the best segmentation obtained by the ICM algorithm proposed in [6] using the same potential values 0.3 resulting in even larger error of 2.9%, and the "ground truth," that is, the segmentation done by a radiologist. With respect to this latter binary map, the initial pixelwise classification using the densities approximated by the linear combinations of Gaussians is already sufficiently accurate so that the final refinement results in only a minor improvement.

classification error of 0.0045 on the intersecting distribution tails is obtained for the threshold $t = 108$ when the components 1–3 and 4–10 correspond to the first class (the lung tissues) and the second class (the chest tissues), respectively (see the estimates for each class in Figure 9).

To highlight advantages of using the linear combination of Gaussians, the same empirical density was approximated with a conventional mixture of 12 Gaussians. The results are shown in Figures 10 and 11. The classification error is almost three times higher (0.013) because one of the components combines the former tails of both classes and cannot

(a)

(b)

(c)

(d)

(e)

(f)

FIGURE 12: (a) Initial lower-level segmentation; (b) final lung regions (the error of 0.89%); (c) lung regions obtained with the randomly chosen higher-level parameters (the error of 1.86%); (d) segmentation by the MRS algorithm with the potential values 0.3 and three levels of resolution (the error of 2.3%); (e) segmentation by the ICM algorithm with the potential values 0.3 (the error of 2.9%); and (f) lung regions segmented by a radiologist. The errors with respect to this ground truth are highlighted by the yellow color.

Figure 13 compares our results to those obtained using the iterative thresholding proposed by Hu and Hoffman [25]. In contrast to their method, our segmentation does not lose abnormal lung tissues.

But the ground truth given by the radiologist may contain by itself errors due to hand positioning instabilities during the manual segmentation. In order to better measure the accuracy of our approach, we have created a phantom with the same grey-level distribution as the above CT slices. The phantom, its ground truth, and our segmentation are shown in Figure 14. The error 0.091% between the found regions and the ground truth confirms the high accuracy of the proposed approach.

The above experiments as well as additional experiments with 120 different bimodal LDCT slices have shown that our

(a)

The error: 0.98%   The error: 1.09%

(b)

The error: 3.01%   The error: 13.1%

(c)

(d)

FIGURE 13: (a) Original CT slices; (b) our lung regions (the segmentation errors are only around the outer edge); (c) segmentation by the method in [25]; and (d) regions given by a radiologist. The errors are highlighted by the yellow color.

segmentation yields much better results than several more conventional algorithms. As indicated in Table 3, the most accurate algorithm among these latter algorithms, namely, the MRS [26], has the larger error range of 1.9–9.8% and the mean error of 5.1% with respect to the ground truth. Our segmentation has the notably smallest error range of 0.21–3.25% and its mean error of 0.72% is more than ten times less.
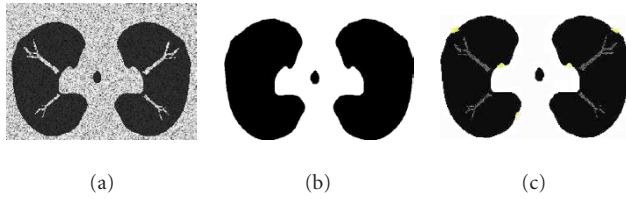
FIGURE 14: (a) Generated phantom; (b) ground-truth image (black lung and grey chest regions); (c) our segmentation (the error of 0.091% around the outer edge is highlighted by the yellow color).

TABLE 3: Accuracy and time performance of our segmentation in comparison to the three conventional algorithms (IT [25], MRS [7], and ICM [6]).

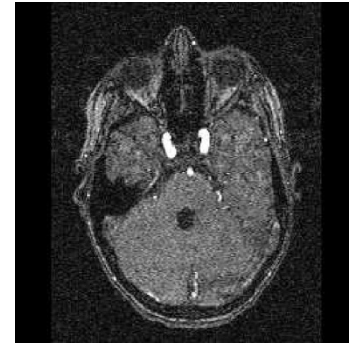| Characteristics	Segmentation algorithm			
	Ours	IT	MRS	ICM
Minimum error (%)	**0.21**	2.81	1.9	2.03
Maximum error (%)	**3.25**	21.9	9.8	17.1
Mean error (%)	**0.72**	10.9	5.1	9.8
Standard deviation (%)	**0.81**	6.04	3.31	5.11
Average time (s)	**117**	197	91	125

TABLE 4: Initial parameter estimates for the dominant Gaussians for the MRA image in Figure 15a.

Parameter	Class		
	Bones and fat	Brain tissues	Blood vessels
Mean (μ)	24.7	105.7	210.7
Variance (σ^2)	126	318	250
Weight (w)	0.518	0.456	0.026

(a)

(b)

FIGURE 15: (a) Typical TOF-MRA scan slice and (b) the deviations between the empirical distribution and the dominant mixture.
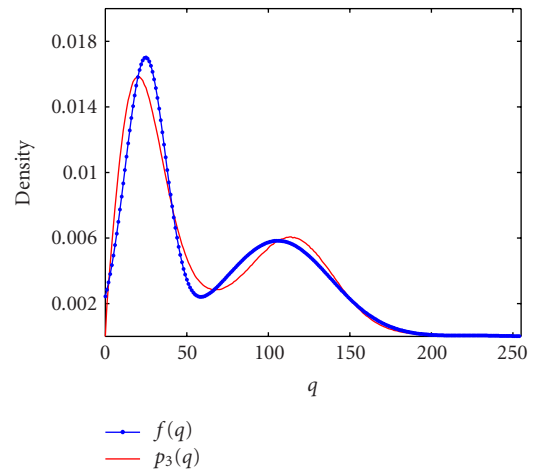
3.2. Segmenting MRA images: blood vessels

Precise segmentation with three dominant modes is obtained in a similar way. Figure 15 shows an MRA image and its three-modal empirical grey-level distribution approximated with the dominant three-component normal mixture. The three classes represent dark bones and fat, brain tissues, and bright blood vessels, respectively. The goal is to separate the latter class in spite of its large intersection with the second class and the very low prior probability. Initial parameters of the dominant mixture are given in Table 4, and the Levy distance of 0.08 indicates a big mismatch between the mixture and the empirical distribution. Figure 16 shows the scaled deviations between these two distributions as well as the six estimated subordinate Gaussians giving the minimum approximation error.

Figure 17 shows the approximated absolute deviation and the whole model obtained after the subordinate parts are combined with the dominant mixture. The resulting Levy distance between the empirical distribution and the model decreases to 0.01 from the original value of 0.08. The minimum classification error of 0.01 on the intersecting distribution tails is obtained for the separation thresholds $t_1 = 57$ and $t_2 = 190$. In this case, the subordinate Gaussians 1–3, 4–5, and 6 correspond to the first (bones and fat), second (brain tissues), and third (blood vessels) classes, respectively. The nine components of the whole model are shown in Figure 17c, and the estimated individual models for each class are shown in Figure 17d.

Figure 18 presents the initial and the final region maps for the third class. The first eight iterations of the map refinement converge to the final region map shown in Figure 18b having the error about 0.51% with respect to the expert's "ground truth" in Figure 18c. The initial and final potentials for the map model for the MRA image in Figure 15a are shown in Table 5. The final estimated parameters for the dominant Gaussians of the model are shown in Table 6.
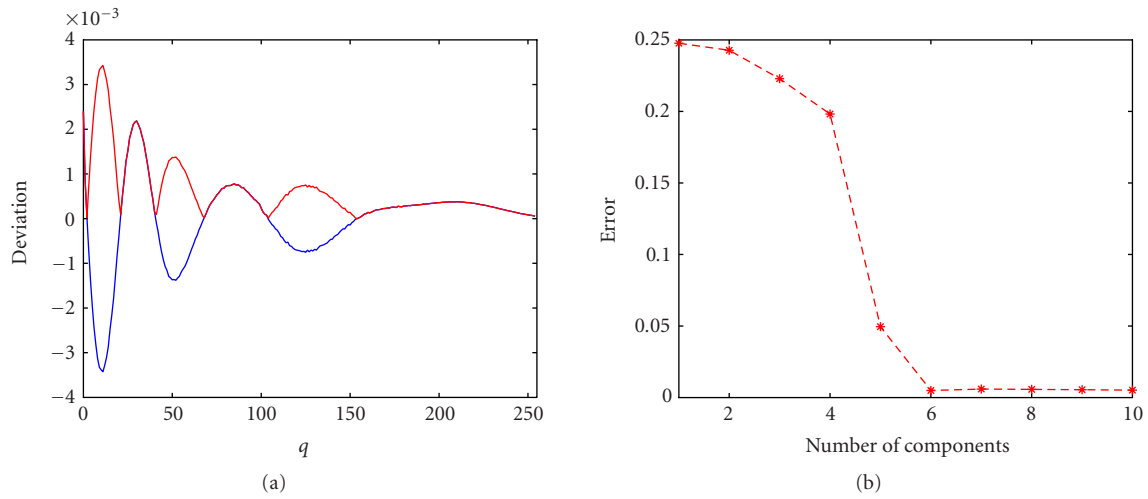
(a)

(b)

FIGURE 16: (a) Estimated subordinate components of the absolute deviation and (b) the absolute error as a function of the number of Gaussians approximating the scaled absolute deviation in Figure 16a.
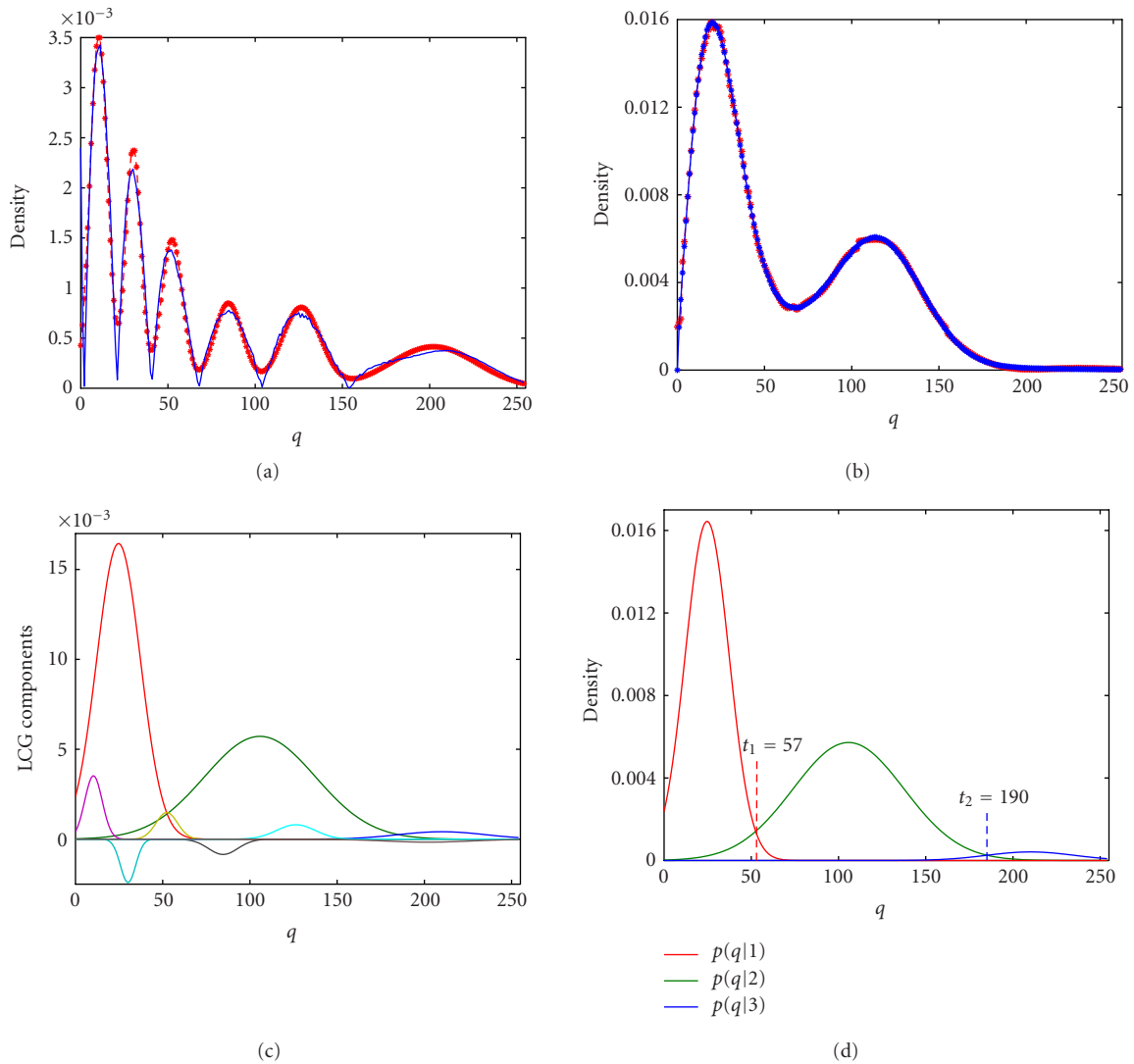

(a)

(b)

(c)

(d)

FIGURE 17: (a) Subordinate mixture estimated for the absolute deviation in Figure 16a; (b) the empirical and estimated densities for the MRA image in Figure 15a; (c) the model components; and (d) the individual models of the classes "bones," "brain tissues," and "blood vessels."

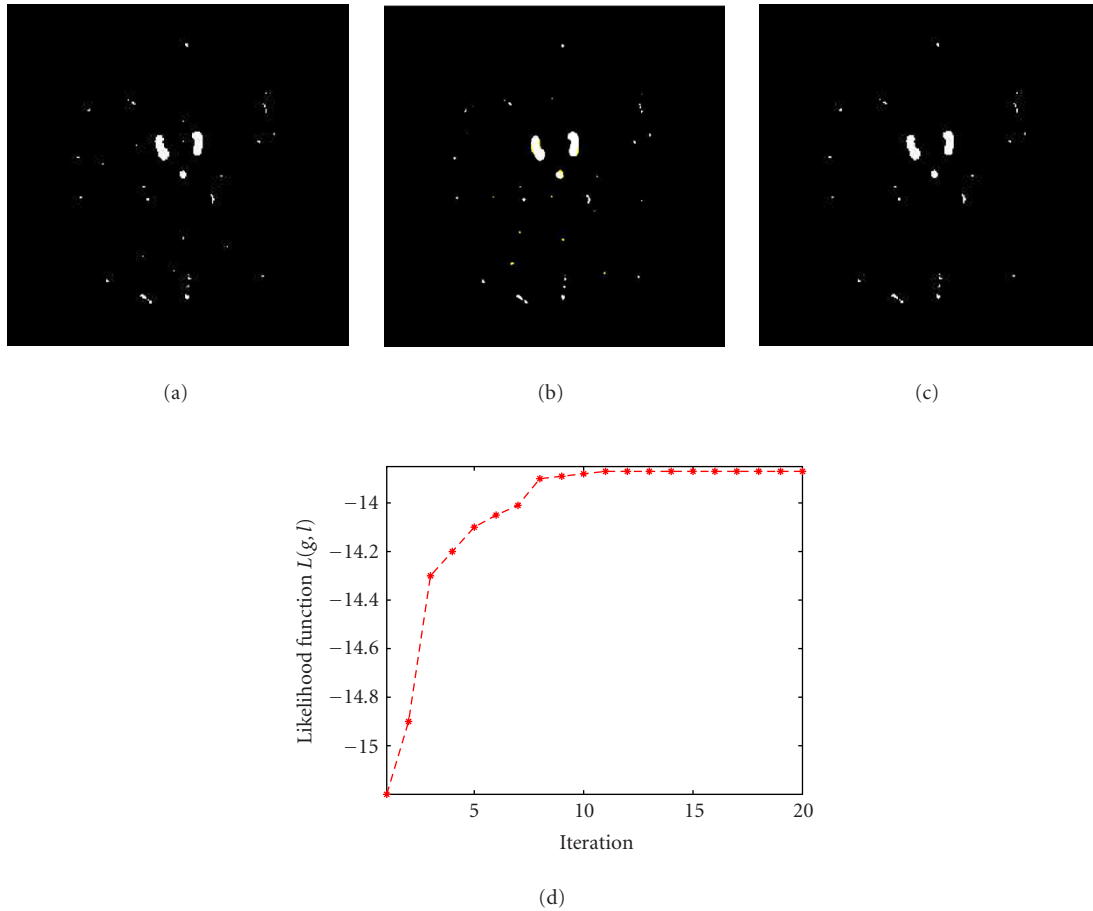(a)   (b)   (c)

(d)

FIGURE 18: (a) Initial and (b) final segmentation of blood vessels by our approach (the final error of 0.51%); (c) the ground truth (blood vessels outlined by a radiologist); and (d) the convergence of the proposed algorithm for the experiment in Figure 15a. The errors are highlighted by the yellow color.

TABLE 5: Initial and final potentials of the map model for the MRA image in Figure 15a.

Parameters	α	λ_1	λ_2	λ_3	λ_4
Initial	1	0.08	0.05	0.09	0.09
Final	1	0.16	0.11	0.17	0.15

TABLE 6: Final parameter estimates for the dominant Gaussians for the MRA image in Figure 15a.

Parameter	Class		
	Bones and fat	Brain tissues	Blood vessels
Mean (μ)	23.2	101.5	208.9
Variance (σ^2)	99	329	270
Weight (w)	0.52	0.451	0.028

3.3. Segmentation of brain tissues

Figure 19 shows a weighted T2 MRI image and its four-modal empirical grey-level distribution approximated with the dominant four-component normal mixture. The four classes represent dark bones and fat, gray matter, white matter, and cerebrospinal fluid (CSF), respectively. The goal is to separate these classes with the minimum classification error. Initial parameters of the dominant mixture are given in Table 7. The Levy distance of 0.11 indicates a large mismatch between the empirical distribution and the dominant

mixture. Figure 20 shows the scaled deviations of the dominant mixture from the empirical distribution and the changes of the approximation error for the increasing number of components. The minimum error is obtained for the 16 subordinate Gaussians. Figure 21a presents this initial estimate.

Figure 21b shows the estimated model after the subordinate components are combined with the dominant mixture. The resulting Levy distance between the empirical
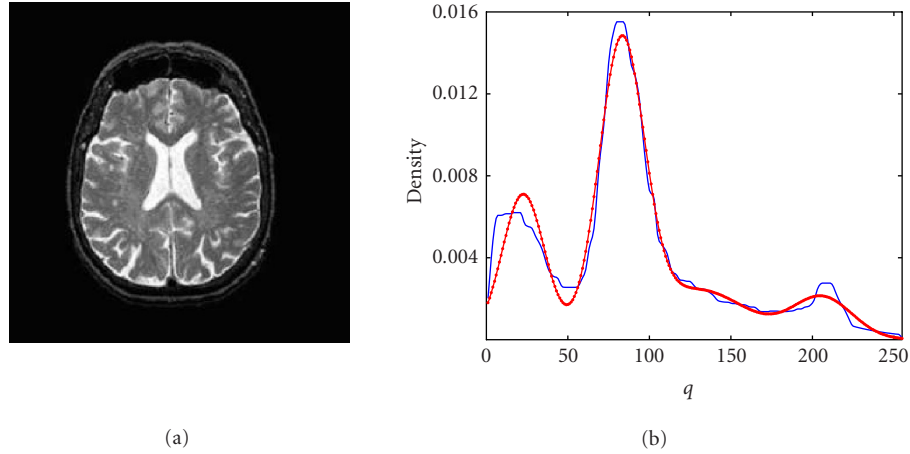
(a)   (b)

FIGURE 19: (a) Typical MRI T2 weighted image and (b) the deviations between the empirical distribution and the dominant mixture.

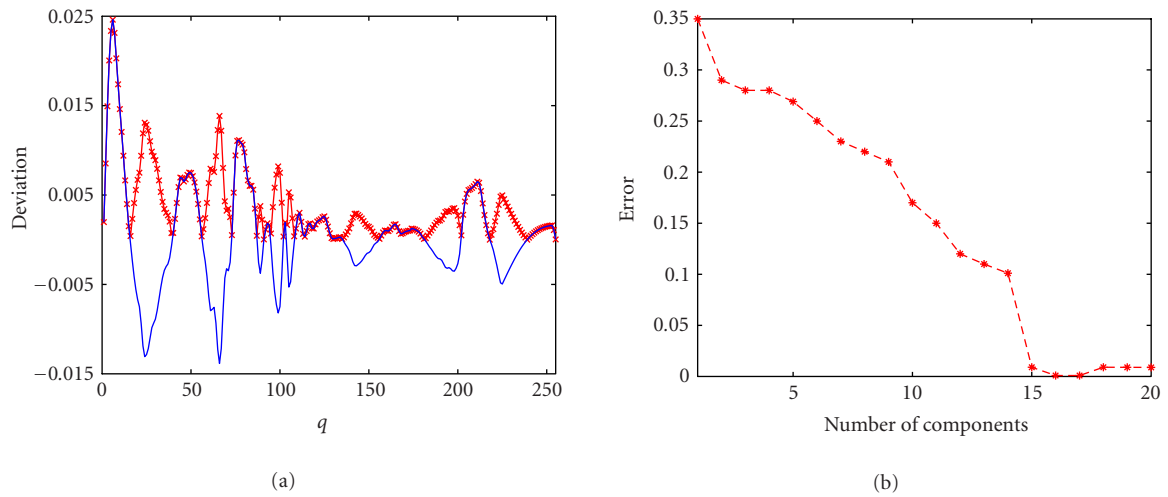(a)   (b)

FIGURE 20: (a) Estimated subordinate components of the absolute deviation and (b) dynamics of the error as a function of the number of Gaussians approximating the scaled-up absolute deviation in Figure 20a.

TABLE 7: Initial parameter estimates for the dominant Gaussians for the MRI image in Figure 19a.

Parameter	Class			
	Bones and fat	White matter	Gray matter	CSF
Mean (μ)	21.7	82.0	128.6	205.0
Variance (σ^2)	171.1	192.6	892.7	337.4
Weight (w)	0.23	0.49	0.19	0.09

distribution and the model decreases to 0.02 from the original value of 0.11. The minimum classification error of 0.01 on the intersecting distribution tails is obtained for the thresholds $t_1 = 53$, $t_2 = 110$, and $t_3 = 179$ when the components 1–3, 4–8, 9–12, and 13–16 are assigned to the first (bones and fat), second (white matter), third (gray matter), and fourth (CSF) classes, respectively. The twenty components of the model are shown in Figure 21c, and the estimated individual models of each class are shown in Figure 21d.

The map refinement process converges after 15 iterations as shown in Figure 22d. The initial and final potentials are shown in Table 8. Table 9 shows the final estimated parameters for the dominant Gaussians. Figure 22 presents the initial and final region maps. The final error is about 3% with respect to the expert's map in Figure 22c.

4. CONCLUDING REMARKS

We introduced a novel sequential EM-based algorithm for accurate segmentation of intrinsically multimodal grey-scale
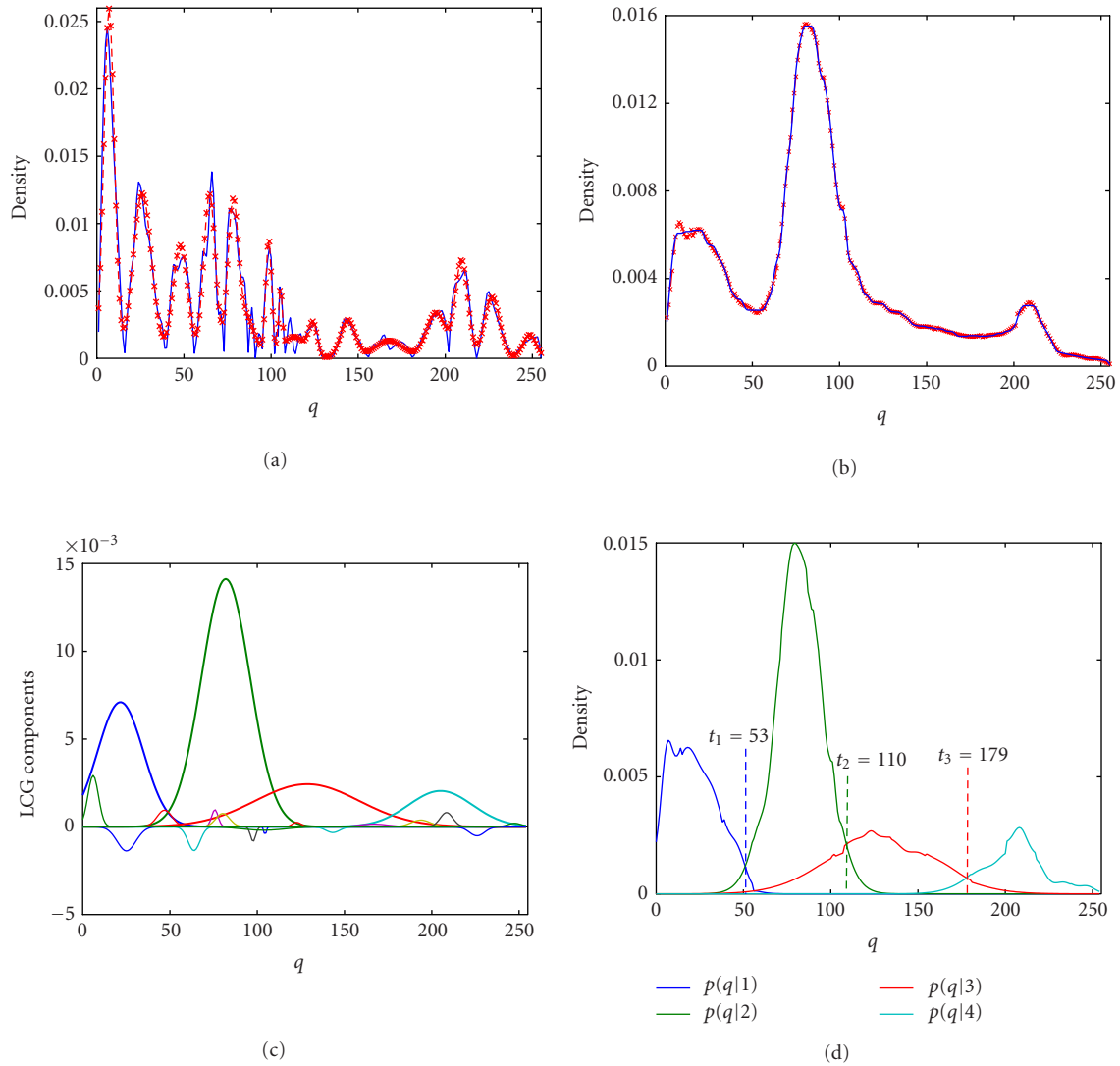
FIGURE 21: (a) Subordinate mixture estimated for the absolute deviation in Figure 20; (b) empirical and estimated densities for the MRI image in Figure 19a; (c) the model components; and (d) the estimated models of the classes "bones," "gray matter," "white matter," and "CSF."

TABLE 8: Initial and final potentials of the map model for the MRI image in Figure 19a.

Parameters	α	λ_1	λ_2	λ_3	λ_4
Initial	1	0.25	0.15	0.37	0.35
Final	1	0.37	0.29	0.41	0.49

TABLE 9: Final parameter estimates for the dominant Gaussians for the MRI image in Figure 19a.

Parameter	Class			
	Bones and fat	White matter	Gray matter	CSF
Mean (μ)	22.9	83.8	125.5	207.0
Variance (σ^2)	190.1	201.6	780.6	299.4
Weight (w)	0.22	0.5	0.195	0.085

images. The empirical probability distribution of grey levels is approximated by a mixture of linear combinations of Gaussians more precisely than by a conventional mixture of Gaussians. This results in a high-quality initial segmentation map that needs only a small refinement by further processing based on the joint Gibbs random field model of region

maps and grey-scale images. The model parameters are estimated analytically and then refined using a genetic optimization technique.
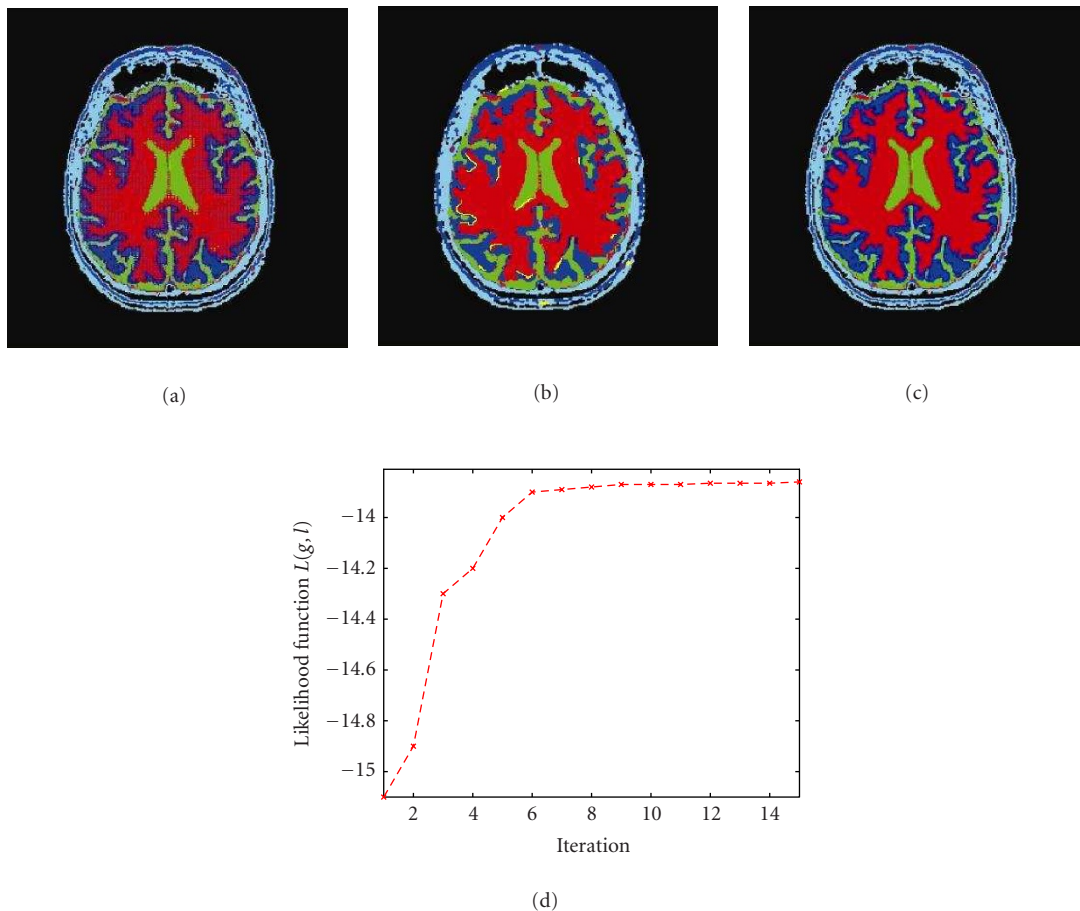
(a)

(b)

(c)

(d)

FIGURE 22: (a) Initial and (b) final segmentation by our approach (the final error of 3%); (c) the ground truth by an expert; and (d) the convergence of the proposed algorithm for the experiment in Figure 19a. The errors are highlighted by the yellow color.

Our experiments confirm that the proposed algorithm segments complex images in some medical applications more precisely than other known techniques. The algorithm is also suitable for approximating different empirical data curves with linear combinations of Gaussians. Future extensions will include the like approximation of multivariate densities, other Gibbs models for the higher level, and the use of stochastic optimization techniques for estimating the models' parameters.

REFERENCES

[1] C. A. Glasbey, "An analysis of histogram-based thresholding algorithms," *CVGIP: Graphical Models and Image Processing*, vol. 55, no. 6, pp. 532–537, 1993.

[2] J. Kittler and J. Illingworth, "Minimum error thresholding," *Pattern Recognition*, vol. 19, no. 1, pp. 41–47, 1986.

[3] N. Otsu, "A threshold selection method from gray level histograms," *IEEE Trans. Syst., Man, Cybern.*, vol. 9, no. 1, pp. 62–66, 1979.

[4] N. R. Pal and S. K. Pal, "A review on image segmentation techniques," *Pattern Recognition*, vol. 26, no. 9, pp. 1277–1294, 1993.

[5] O. D. Trier and A. K. Jain, "Goal-directed evaluation of binarization methods," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 17, no. 12, pp. 1191–1201, 1995.

[6] J. E. Besag, "On the statistical analysis of dirty pictures," *Journal of the Royal Statistical Society*, vol. B48, no. 3, pp. 259–302, 1986.

[7] C. A. Bouman and M. Shapiro, "A multiscale random field model for Bayesian image segmentation," *IEEE Trans. Image Processing*, vol. 3, no. 2, pp. 162–177, 1994.

[8] R. Chellappa and R. L. Kashyap, "Digital image restoration using spatial interaction models," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 30, no. 3, pp. 461–472, 1982.

[9] A. El-Baz and A. A. Farag, "Image segmentation using GMRF models: parameters estimation and applications," in *Proc. IEEE International Conference on Image Processing (ICIP '03)*, vol. 2, pp. 177–180, Barcelona, Spain, September 2003.

[10] A. A. Farag and E. J. Delp, "Image segmentation based on composite random field models," *Optical Engineering*, vol. 31, no. 12, pp. 2594–2607, 1992.

[11] G. L. Gimel'farb, *Image Textures and Gibbs Random Fields*, Kluwer Academic, Dordrecht, the Netherlands, 1999.

[12] A. K. Jain, "Advances in mathematical models for image processing," *Proc. IEEE*, vol. 69, no. 5, pp. 502–528, 1981.

[13] R. C. Dubes and A. K. Jain, "Random field models in image analysis," *Journal of Applied Statistics*, vol. 16, no. 2, pp. 131–164, 1989.

[14] A. Goshtasby and W. D. Oneill, "Curve fitting by a sum of Gaussians," *CVGIP: Graphical Models and Image Processing*, vol. 56, no. 4, pp. 281–288, 1994.

[15] T. Poggio and F. Girosi, "Networks for approximation and learning," *Proc. IEEE*, vol. 78, no. 9, pp. 1481–1497, 1990.

[16] H. W. Sorenson and D. L. Alspach, "Recursive Bayesian estimation using Gaussian sums," *Automatica*, vol. 7, no. 4, pp. 465–479, 1971.

[17] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistical Society*, vol. 39B, no. 1, pp. 1–38, 1977.

[18] T. K. Moon, "The expectation-maximization algorithm," *IEEE Signal Processing Mag.*, vol. 13, no. 6, pp. 47–60, 1996.

[19] R. A. Redner and H. F. Walker, "Mixture densities, maximum likelihood and the EM algorithm (review)," *SIAM Review*, vol. 26, no. 2, pp. 195–239, 1984.

[20] M. I. Schlesinger, "Relation between learning and self-learning in pattern recognition," *Kibernetika*, no. 2, pp. 81–88, 1968 (Russian).

[21] M. I. Schlesinger and V. Hlavac, *Ten Lectures on Statistical and Structural Pattern Recognition*, Kluwer Academic, Dordrecht, the Netherlands, 2002.

[22] N. E. Day, "Estimating the components of a mixture of normal distributions," *Biometrika*, vol. 56, pp. 463–474, 1969.

[23] D. E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison-Wesley, Boston, Mass, USA, 1989.

[24] J. W. Lamperti, *Probability*, John Wiley & Sons, New York, NY, USA, 1996.

[25] S. Hu, E. A. Hoffman, and J. M. Reinhardt, "Automatic lung segmentation for accurate quantitation of volumetric X-Ray CT images," *IEEE Trans. Med. Imag.*, vol. 20, no. 6, pp. 490–498, 2001.

[26] C. A. Bouman and B. Liu, "Multiple resolution segmentation of textured images," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 13, no. 2, pp. 99–113, 1991.

Ayman El-Baz received the B.S. and M.S. degrees in electrical engineering from Mansoura University, Egypt, in 1997 and 2000, respectively. He joined the Computer Vision and Image Processing (CVIP) Lab, the University of Louisville, Kentucky, in May 2001 as a Ph.D. student. During his stay at the University of Louisville, he has been involved in the applications of image processing and computer vision for medical image analysis. His current research includes image modeling, image segmentation, 2D and 3D registration, visualization, and surgical simulation including finite element analysis, where he has authored or coauthored more than 25 technical articles. Mr. El-Baz is a regular reviewer for a number of technical journals and conferences including the IEEE Transactions on Image Processing, and the International Conference on Medical Image Computing and Computer-Assisted Intervention. He was the recipient of the Research Louisville Award in 2002 and the Speed School Award for Outstanding Student, Research and Creative Activity in 2003. He is a Member of the IEEE and a Member of Eta Kappa Nu.

Aly A. Farag received his B.S. degree in electrical engineering from Cairo University, his M.S. degree in biomedical engineering from Ohio State University, Columbus, another M.S. degree in bioengineering from the University of Michigan, Ann Arbor, and his Ph.D. degree in electrical engineering from the Purdue University, Indiana. Dr. Farag joined the University of Louisville, Kentucky, in August 1990, where he is currently a Professor of electrical and computer engineering. His research interests are concentrated in the fields of computer vision and medical imaging. Dr. Farag is the founder and Director of the Computer Vision and Image Processing Laboratory (CVIP Lab), the University of Louisville, which supports a group of more 20 graduate students and postdocs. His contribution has been mainly in the areas of active vision system design, and volume registration, segmentation, and visualization. He has authored or coauthored more than 80 technical articles in leading journals and international meetings in the fields of computer vision and medical imaging. Dr. Farag is an Associate Editor of the IEEE Transactions on Image Processing. He is a regular reviewer for a number of technical journals and to national agencies including the US National Science Foundation and the National Institute of Health. He is a Senior Member of the IEEE and SME, and a Member of Sigma Xi and Phi Kappa Phi.

Georgy Gimel'farb graduated from Kiev Polytechnic Institute, Ukraine, and received a Ph.D. degree in engineering cybernetics from the Institute of Cybernetics, Academy of Sciences of Ukraine, and a D.S. (Eng.) degree in control engineering from the Higher Certifying Commission of the USSR, Moscow, Russia. After working in the Institute of Cybernetics for a long time, Dr. Gimel'farb joined the University of Auckland, New Zealand, in July 1997, where he is currently an Associate Professor of computer science. His research is focused on image analysis, computer vision, and statistical pattern recognition with the main contributions in the areas of computational stereovision and probabilistic texture modeling and analysis. He has authored or coauthored two books and more than 200 technical articles in professional journals and international meetings in the fields of pattern recognition and computer vision. He is a reviewer for several technical journals and a member of program committees for a number of international conferences and workshops.