Region-Of-Interest Selection and Statistical Analysis of Dynamic Breast Magnetic Resonance Imaging Data

Florence Forbes, Chris Fraley, Dianne Georgian-Smith, David M. Goldhaber, Nathalie Peyrard, Adrian Raftery

INRIA

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Region-Of-Interest Selection and Statistical Analysis of Dynamic Breast Magnetic Resonance Imaging Data

Florence Forbes — Chris Fraley — Dianne Georgian-Smith — David M. Goldhaber — Nathalie Peyrard — Adrian Raftery

N° 4249

Septembre 2001

THÈME 4

ISSN 0249-6399 ISRN INRIA/RR--4249--FR+ENG

INRIA
RHÔNE-ALPES

Region-Of-Interest Selection and Statistical Analysis of Dynamic Breast Magnetic Resonance Imaging Data

Florence Forbes , Chris Fraley , Dianne Georgian-Smith , David M. Goldhaber , Nathalie Peyrard , Adrian Raftery *

Thème 4 — Simulation et optimisation
de systèmes complexes
Projet IS2

Rapport de recherche n° 4249 — Septembre 2001 — 31 pages

Abstract: Magnetic Resonance Imaging (MRI) is emerging as a powerful tool for the diagnosis of breast abnormalities. Early investigation of contrast-enhanced MRI of the breast demonstrated that breast carcinoma consistently enhanced after the intravenous administration of contrast material (Gadolinium). However, many benign lesions also enhance reducing as a result MRI specificity. Further characterization is necessary to differentiate benign from malignant lesions and analysis of the temporal pattern of contrast uptake has been applied to improve specificity. Selecting a Region-Of-Interest (ROI) is an almost universal step in the process of examining the contrast uptake characteristics of a breast lesion. A limitation of most ROI analysis procedures is that only subjectively selected regions are examined. We propose a ROI selection method based on techniques from multivariate data analysis and statistical image segmentation. In addition to the enhancement rate of the contrast material, we use other characteristics of the signal intensity to group areas of the breast with similar characteristics and then select the group which corresponds to a fast and strong enhancement. We then investigate tools for subsequent analysis of signal intensity time course data in the selected region in order to determine if the enhancing lesion is likely to be benign or malignant. We tested our procedure on cases for which the diagnosis was available. The results are promising and are consistent with the diagnosed nature of the lesions.

Key-words: Magnetic resonance imaging, Region-Of-Interest, time-signal intensity curves, model-based clustering, morphology, Markov random fields.

* Florence Forbes and Nathalie Peyrard, Projet IS2, Inria Rhône-Alpes. Chris Fraley and Adrian E. Raftery, Department of Statistics, University of Washington. Doctor Dianne Georgian-Smith, Department of Radiology, Massachusetts General Hospital, Boston. David M. Goldhaber, International MRI Accreditation Resources, South San Francisco.

Unité de recherche INRIA Rhône-Alpes
655, avenue de l'Europe, 38330 Montbonnot-St-Martin (France)
Téléphone : +33 4 76 61 52 00 —Télécopie +33 4 76 61 52 52

Sélection de régions d'intérêt et analyse statistique de données d'imagerie à résonance magnétique du sein

Résumé : L'imagerie à résonance magnétique (IRM) apparaît comme un outil puissant d'aide au diagnostic des lésions du sein. C'est une technique sensible car les carcinomes réagissent fortement après injection d'un agent de contraste (Gadolinium), mais peu spécifique car des lésions bénignes peuvent également réagir fortement. Dans le but d'en améliorer la spécificité, nous proposons d'analyser les courbes d'évolution de l'absorption de l'agent de contraste. Une telle analyse nécessite la sélection préliminaire d'une région d'intérêt, souvent faite de manière subjective. Nous proposons ici une sélection automatique basée sur des techniques d'analyse de données multidimensionnelles et des méthodes statistiques de segmentation d'images. En plus du taux d'accroissement de l'agent de contraste, nous utilisons d'autres caractéristiques des courbes d'évolution afin de grouper les zones du sein de caractéristiques similaires et de sélectionner le groupe qui présente le plus rapide et le plus fort accroissement. Nous proposons ensuite des outils pour analyser les courbes dans la région sélectionnée afin de caractériser la nature de la lésion. Nous appliquons cette procédure sur quelques cas dont le diagnostic est connu et obtenons des résultats prometteurs en accord avec la nature diagnostiquée des lésions.

Mots-clés : Imagerie à résonance magnétique, région d'intérêt,courbes d'intensité, classification basée sur des modèles probabilistes, champs de Markov, morphologie bayésienne.

Contents

1 Introduction

Magnetic Resonance Imaging (MRI) is emerging as a powerful tool for the diagnosis of breast abnormalities. It can assist in the differential diagnosis of questionable lesions and is used as a complementary diagnostic modality in breast imaging. It has also proved useful in the evaluation of the extent of breast cancer as well as in treatment planning.

Because of the high reactivity of breast carcinomas after Gadolinium injection, this method has the potential to allow differentiation between malignant and benign tissues. However, some benign lesions also enhance, as a result reducing the specificity of MRI. Dynamic analysis of the temporal pattern of contrast uptake has been applied to improve specificity. For a recent review of the current knowledge base on detection and differential diagnosis of breast tumors see [13] and [15]. The diagnostic criteria that are in use for differential diagnosis can be divided into those related to lesion enhancement kinetics and those related to lesion morphology. Recent studies like [16] suggest that signal intensity time course data are useful for differentiating benign from malignant enhancing lesions. The authors conclude that the overall shape of the time-signal intensity curve is an important criterion, while a single attribute of the curve, such as the enhancement rate, may not be enough.

The evaluation of morphologic features and the extraction of architectural information is usually also based on post-contrast images of enhancing areas so that the above distinction is somewhat arbitrary. Integrating multiple diagnostic criteria (both qualitative and quantitative) is therefore recommended.

Selecting a region of interest (ROI) is an important first step in the process of examining the contrast uptake characteristics of a breast lesion from either a morphologic or kinetic point of view. However, no standard method for ROI selection and analysis of dynamic breast MR data has yet been established. A limitation of most ROI analysis procedures is that only subjectively selected regions are examined. In this paper, we propose an ROI selection method based on statistical techniques. We first use a multivariate classification method to produce color classification images in which parts of the breast with similar signal intensity time courses are assigned to a class represented by a color. This gives morphological information that can be used to select an ROI by focusing on the pixels with the strongest enhancement. We have also developed some tools for analyzing the enhancement kinetics for pixels in the selected region.

Section 3 gives a description of the classification method, and its use for ROI selection and enhancement kinetics analysis is discussed in Section 4. We then also propose, in Section 4, techniques for improving differential diagnosis based on the shapes of the curves in the selected region.

2 The Data

For this exploratory study, we considered sequences of images for three patients representing different cases (malignant and benign lesions). Several two-dimensional slices were available

for each patient. For each slice, 25 sequential magnetic resonance (MR) breast images were acquired (one image approximately every 10 seconds). See Figure 1 for examples of such images. Each image records the signal intensity at a given time after injection. Instead of working directly with these MR images, we summarized them in terms of five variables considered to be of significance for cancer diagnosis. These five variables are calculated from a curve fitting procedure using software developed at Toshiba America MRI, inc. ([14]). A fit analysis is carried out at each pixel location for the signal intensity curve. Figure 2 shows a signal intensity curve at a given pixel, after subtraction of the reference signal. The fitting model is assumed to be made of three successive sections: a zero signal, a second order polynomial curve and a flat line. We used the following five variables in our study:

- **Time to Peak:** the time at which the signal peaks.
- **Difference at peak:** absolute increase of intensity between the begining of the signal and the time at which the signal peaks.
- **Enhancement slope:** in units of intensity/time.
- **Maximum step:** maximum change between two adjacent dynamic samples.
- **Washout slope:** in units of intensity/time.

In addition to the images, diagnostic information is available. We have a set of 512×256 images corresponding to one slice for patient 05, containing a spherical lesion diagnosed as a carcinoma. Another set corresponds to 176×352 images for slice 6 of patient 08. The MR data was obtained less than a week after surgery and the radiologists concluded that there was no residual carcinoma, *i.e.* the margins of the surgery site were not suspicious. For patient 28, three slices are available (10, 11, 12) (192×192 images). A spherical carcinoma is known to be present in slice 10 and 11. After a tumor biopsy, the tumor size was estimated to be less than 16mm, which means that there should be no malignant tissues on slice 12 (a slice is 8mm).

To summarize, our starting point is thus five images for each case, made from the values of each variables at each pixel location, rather than the original 25 images. Although this preprocessing reduces the amount of data to be analyzed, the characterization of breast lesions based on these MR images remains a difficult task. In Section 3, we present the multivariate statistical methods for clustering and spatial segmentation we propose to synthesize the available information into a unique classification image.

3 Producing Classification Images

We propose to use statistical segmentation methods to produce a few color images for each case, in which each color represents a group or class of pixels with similar time-signal intensity curves (summarized by the five selected variables). An important issue is the determination of the effective number K of components present in the data, *i.e.* the number of colors to

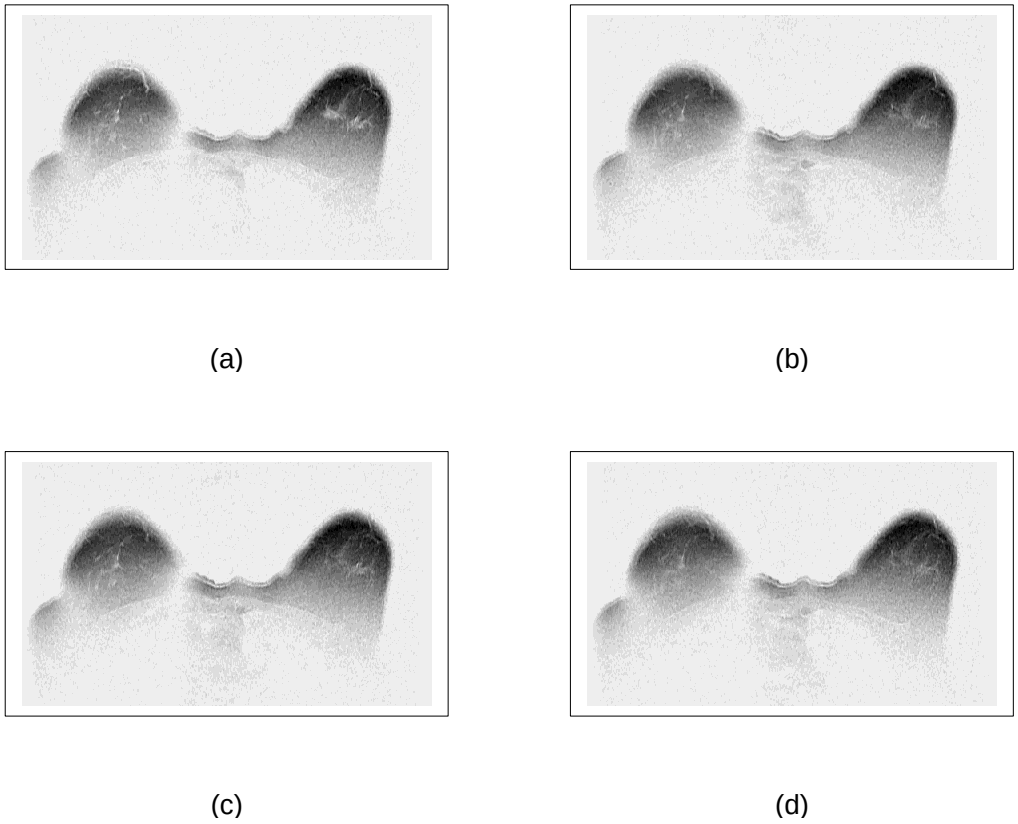


Figure 1: Patient 05, slice 09: dynamic MR images at t=10 (a), 70 (b), 150 (c), 250 (d) seconds.

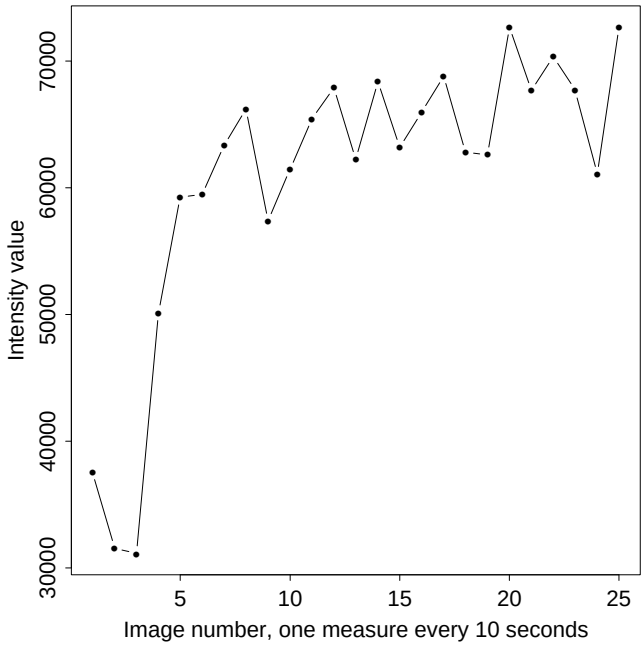


Figure 2: Signal intensity curve at a given pixel, for slice 09 of patient 05, 25 measures are acquired, one measure every 10 seconds.

use in the classification image. According to experts, the main components in the breast are blood vessels, air, fat, a possible tumor and other tissues of less interest. Because of the large number of pixels involved, those corresponding to air are eliminated prior to any further analysis. It is then reasonable to assume that the remaining pixels can be grouped into no fewer than 3 classes, and that while there may be more than 3 classes a large number of classes would not make sense in this context. Nevertheless, allowing somewhat more than 3 classes may provide better statistical performance in terms of identifying the main features of interest in the image. We chose to study three cases, $K = 3, 4$ and 10.

Model-based statistical methods for clustering multivariate observations are very flexible and have been applied successfully in many domains of practical interest (*e.g.* [1, 10, 11, 9])[1]. However for complex data such as those associated with tissue segmentation in medical imaging, these methods sometimes produce rather fragmented results that do not correspond directly to a meaningful classification, because they do not take into account spatial location and dependence. For this reason, we propose refining the results from nonspatial statistical clustering (Section 3.1) with spatial techniques (Section 3.2) ([7]). Small isolated regions are removed by automatically reassigning pixels located in them, reducing the spatial fragmentation of the classification.

3.1 Non Spatial Analysis

We propose to use *marginal mixture EM segmentation* as a first step in our analysis. The idea is to model the marginal distribution of (possibly multivariate) pixel intensities as a finite Gaussian mixture model, and use the EM (Expectation-Maximization) algorithm to estimate the model parameters (*e.g.* [20]; [4]). For multidimensional images, the relevant distributions are multivariate, and there has been much recent progress on estimation methods that combine agglomerative hierarchical clustering methods based on maximum classification likelihood ([1]; [8]) with the EM algorithm ([5]; [10]).

When the number of pixels is very large, the method cannot be used without modification because of its computational requirements. One possible approach is to take a sample of the pixels first, as was done for an MRI brain scan in [1]. Initialization via Minimum Spanning Trees ([18]), which does not require subsampling the pixels, is an alternative for making hierarchical clustering feasible for very large images. We used the MCLUST software for model-based clustering for our analysis, which is described in [10], [11], and [9]. It combines model-based Gaussian agglomerative hierarchical clustering methods ([1]; [8]) with the EM algorithm for Gaussian mixture models ([4]). The EM algorithm ([6], [17]) is an iterative method widely used in parameter estimation for incomplete data. For clustering applications via mixture models, the missing values are the cluster memberships of each pixel. To be effective, the EM algorithm generally requires a good initial estimate. Classifications produced by model-based Gaussian agglomerative hierarchical clustering, which are often

[1]These and all cited papers written by the authors of this report are available as technical reports at `http://www.stat.washington.edu/tech.reports`

good but rarely optimal, are a good way to initialize the EM algorithm for classification ([5]; [10]). The process is described in more detail below.

In what follows, observations (corresponding to image pixels) are denoted by y_i and are assumed to be 5-dimensional vectors, corresponding to the five more informative variables. The y_i are assumed to arise from a K-component Gaussian mixture model. More specifically, if y_i arises from class $k \in \{1, \ldots, K\}$, its distribution is multivariate normal (Gaussian) and depends on a mean vector μ_k and a covariance matrix Σ_k. The likelihood for our data is then

$$\mathcal{L}(\theta_1, \ldots, \theta_K; p_1, \ldots, p_K \mid y) = \prod_{i=1}^{n} \sum_{k=1}^{K} p_k f_k(y_i \mid \theta_k), \tag{1}$$

where $\theta_k = (\mu_k, \Sigma_k)$ and f_k is the multivariate normal density of the kth component in the mixture, parameterized by its mean μ_k and covariance matrix Σ_k:

$$f_k(y_i \mid \mu_k, \Sigma_k) \equiv \frac{\exp\left\{-\frac{1}{2}(y_i - \mu_k)^T \Sigma_k^{-1} (y_i - \mu_k)\right\}}{\sqrt{\det(2\pi\Sigma_k)}}. \tag{2}$$

Here p_k is the probability that an observation belongs to the kth component ($p_k \geq 0; \sum_{k=1}^{K} p_k = 1$).

Data generated by mixtures of mutivariate normal densities are characterized by groups or clusters centered at the means μ_k. The corresponding surfaces of constant density are ellipsoidal. Geometric features (shape, volume, orientation) of the clusters are determined by the covariances Σ_k, which may also be parameterized to impose cross-cluster constraints. There are a number of possible parameterizations of Σ_k ([1]; [4]), six of which have been implemented in MCLUST. Common instances include $\Sigma_k = \lambda I$, where all clusters are spherical and of the same size; uniform Σ_k, where all clusters have the same geometry but need not be spherical ([12]); and unrestricted Σ_k, where each cluster may have a different geometry ([19]). Banfield and Raftery (1993)[1] proposed a general framework for geometric cross-cluster constraints in multivariate normal mixtures by parameterizing covariance matrices through eigenvalue decomposition in the form

$$\Sigma_k = \lambda_k D_k A_k D_k^T, \tag{3}$$

where D_k is the orthogonal matrix of eigenvectors, A_k is a diagonal matrix whose elements are proportional to the eigenvectors, and λ_k is an associated constant of proportionality. Their idea was to treat λ_k, A_k and D_k as independent sets of parameters, and either constrain them to be the same for each cluster or allow them to vary among clusters.

Based on considerations discussed in earlier studies, we obtained a first segmentation via MCLUST with the constant-shape model $\Sigma_k = \lambda_k D_k A D_k^T$. More specifically, the algorithm provides an estimate of the conditional probability that each pixel belongs to each of the classes, given the observations. The segmentation derived from these conditional probabilities is the one which assigns each pixel to the class with the greatest probability. These probabilities are obtained using the EM algorithm. The EM algorithm is a general approach

to maximum-likelihood estimation for problems in which the data can be viewed as consisting of n multivariate observations x_i recoverable from (y_i, z_i), in which y_i is observed and z_i is unobserved. If the x_i are independent and identically distributed (iid) according to a probability distribution f with parameters θ, then the *complete-data likelihood* is

$$\mathcal{L}_c(\theta \mid x) = \prod_{i=1}^{n} f(x_i \mid \theta).$$

Further, if the probability that a particular variable is unobserved in x_i depends only on the observed data y and not on z, then the *observed data likelihood*, $\mathcal{L}_o(\theta \mid y)$, can be obtained by integrating z out of the complete data likelihood

$$\mathcal{L}_o(\theta \mid y) = \int \mathcal{L}_c(\theta \mid x)\, dz. \tag{4}$$

The Maximum Likelihood Estimate (MLE) for θ based on the observed data maximizes $\mathcal{L}_o(\theta \mid y)$.

The EM iteration alternates between two steps, an 'E-step', in which the conditional expectation of the complete data loglikelihood given the observed data and the current parameter estimates is computed, and an 'M-step' in which parameters that maximize the expected loglikelihood from the E-step are determined.

In EM for multivariate normal mixture models, the 'complete data' are considered to be $x_i = (y_i, z_i)$, where $z_i = (z_{i1}, \ldots, z_{iK})$ with

$$z_{ik} = \begin{cases} 1 & \text{if } x_i \text{ belongs to group } k \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

Assuming that each z_i is independent and identically distributed according to a multinomial distribution of one draw on K categories with probabilities $p_1, \ldots, p_K$, and that the density of an observation y_i given z_i is given by $\prod_{k=1}^{K} f_k(y_i \mid \theta_k)^{z_{ik}}$ with f_k given by (2), the resulting complete-data loglikelihood is

$$l(\theta_1, \ldots, \theta_K, p_1, \ldots, p_k \mid x) = \sum_{i=1}^{n} \sum_{k=1}^{K} z_{ik} \log \left[p_k f_k(y_i \mid \theta_k) \right]. \tag{6}$$

The E-step of the EM algorithm for mixture models is given by

$$\hat{z}_{ik} \leftarrow \frac{\hat{p}_k f_k(y_i \mid \hat{\theta}_k)}{\sum_{j=1}^{K} \hat{p}_j f_j(y_i \mid \hat{\theta}_j)}, \tag{7}$$

while the M-step involves maximizing (6) in terms of p_k and θ_k with z_{ik} fixed at the values computed in the E-step. The value z^*_{ik} of $\hat{z}_{ik}$ at a maximum of (1) is the conditional probability that observation i belongs to group k. The maximum-likelihood classification of observation i is $\{m \mid z^*_{im} = \max_k z^*_{ik}\}$, so that $(1 - \max_k z^*_{ik})$ constitutes a measure of the uncertainty in the classification (*e.g.* [2]).

For the M-step, estimates of the means and probabilities have simple closed form expressions involving the data and $\hat{z}_{ik}$ from the E-step:

$$\hat{p}_k \leftarrow \frac{n_k}{n}; \quad \hat{\mu}_k \leftarrow \frac{\sum_{i=1}^{n} \hat{z}_{ik} y_i}{n_k}; \quad n_k \equiv \sum_{i=1}^{n} \hat{z}_{ik} \tag{8}$$

Computation of the covariance estimate $\hat{\Sigma}_k$ depends on its parameterization. For details of the M-step for Σ_k parameterized by eigenvalue decomposition (3), see [4].

EM estimation for mixture models has a number of limitations. First, the rate of convergence can be very slow. However, EM typically gives good results provided the data conforms reasonably well to the model and the iteration is started at reasonable values.

When applied to images, the algorithm can therefore provide a segmentation by converting the final $\hat{z}_{ik}$ estimates to a discrete classification: pixel i is assigned to class k which satisfies $\hat{z}_{ik} = \max_{j=1\ldots K} \hat{z}_{ij}$.

As an illustration, we focus on one of our data sets, corresponding to slice 09 for Patient 05. Figure 3 shows the segmentations, for 3, 4 and 10 clusters. These numbers do not include the background as a class so that the number of colors in the final segmented images is equal to the number of clusters plus one. In all three images, one can easily recognize the heart and tumor locations. The cluster corresponding to the heart and vessels is depicted in blue while the tumor is in red. The remaining colors indicate other tissues. We will refer to these three groups as to *heart, tumor* and *misc*. The latter group is composed of more than one cluster in the $K = 4$ and $K = 10$ cases.

In the segmentation obtained for $K = 3$ (Figure 3 (b)), many pixels in the skin area are classified as tumor, an indication that more classes are needed for useful image segmentation and tumor identification. This conclusion is further supported by the results for $K = 4$, in which the number of red pixels in the skin area decreases (see Figure 3 (c)), but the tumor remains solidly red.

If we compare the size (number of observations) of each cluster for $K = 3$ and $K = 4$ (Table 1), we can see that the second cluster of type *misc* and the cluster *tumor* are merged into a single one when K decreases from 4 to 3. The behavior of the five variables (see Figure 4) in these two clusters is actually very similar for the time to peak and the maximum step. The tumor, however, shows a greater range of values for difference at peak and enhancement slope and a smaller range of values for the washout.

For both $K = 3$ and $K = 4$, the main difference between the cluster *heart* and the cluster *tumor* resides also in the difference at peak, enhancement slope and washout variables. In this case the tumor shows a greater range of values for the three parameters.

The additional cluster produced for $K = 4$, referred to as *misc1* in the figures, seems to be mainly composed of outliers. The enhancement slope and the washout are equal to zero for most of the pixels in this cluster. (It should therefore be possible to further reduce the data set size by neglecting these pixels.)

Another important difference between the segmented images for $K = 3$ and $K = 4$ is the classification of the area to the left of the tumor, which is classified as tumor for $K = 3$ and non-tumor for $K = 4$. For a higher number of clusters (Figure 3 (d)), the tumor area

Table 1: Clusters volumes for $K = 3$ and $K = 4$.

	tumor	*heart*	$misc_1$	$misc_2$
K=3	1112	1597	707	
K=4	975	1312	751	378

is represented by more than one cluster. For instance, when $K = 10$, four colors can be distinguished in this area. Note that, as before, the method detects the presence of a cluster of pixels whose enhancement slope and washout variables are equal to zero.

In the case $K = 10$, it appears also that the vessels above the heart are classified as *tumor* instead of *heart*, which does not occur when $K = 3$ or $K = 4$. Also of note is that when $K = 4$ and $K = 10$, the tumor is surrounded by a thin border, composed of pixels from several clusters.

Similar analysis has been made for the other data sets. It appears that model-based clustering techniques produce very good and relatively inexpensive initial segmentations of the images that reproduce the important features derived by experts from the full set of 25 sequential images (see Figures 5, 6). However, the nature of the area adjacent to a potential tumor is of interest in the evaluation of the local extent of breast cancer. With this in mind, in the next section we present possible refinements of our initial classifications, taking into account spatial information. It can be also noted that in Figures 5 (c) and 6 (b), the tumor area is not always painted red as one may wish. This comes from the classification method which artificially assign colors (labels) to the different regions. In Section 4.1, we propose a way to detect the suspicious region and automatically assign it to a pre-determined color (*e.g.* red).

3.2 Spatial classification

Of the many classification/segmentation algorithms in existence that take spatial dependences into account, most proceed by transforming a given initial classification into an improved one. In this section we discuss refinements of our model-based classification to incorporate spatial information. For convenience, we transformed the original normalized data into grey-level images.

In section 3.2.1, we describe image smoothing via morphological filters, which treat each pixel in the initial classification only in the context of neighboring pixels. In Section 3.2.2, we discuss methods that make use of the original data in addition to the initial model-based classification.

3.2.1 Morphological filters

Morphological filters are procedures that successively apply a morphological rank operator (*e.g.* citeserra, citeheijmans, citeheijmans94) to each pixel and its neighbors until there are

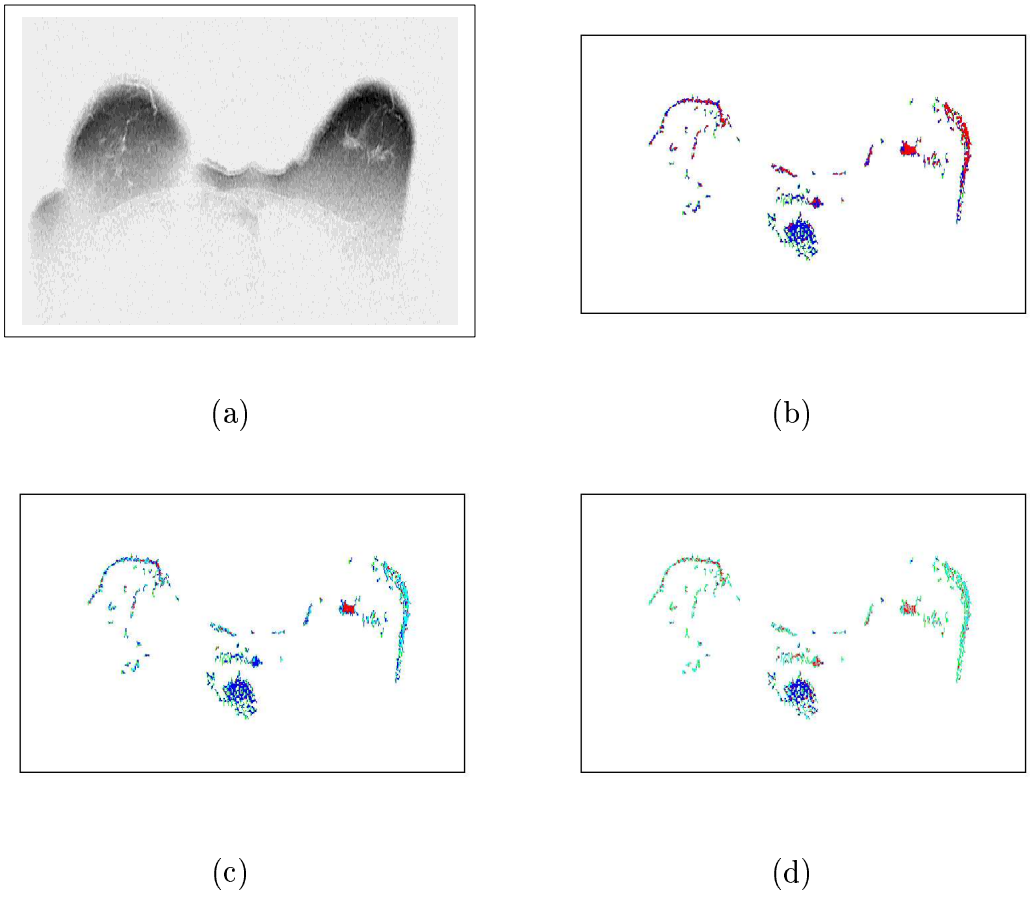
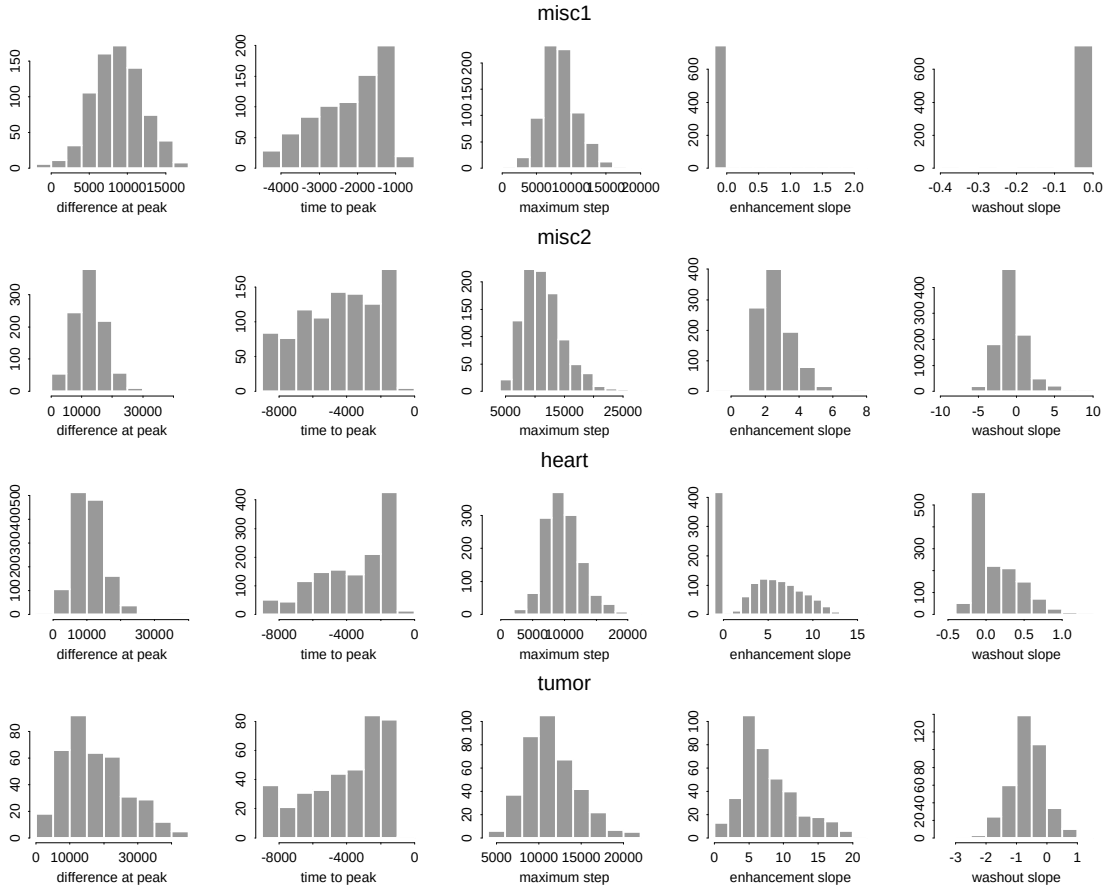


Figure 3: MCLUST classifications for Patient 05, slice 09. (a) reference image, (b) 3 class segmentation, (c) 4 class segmentation, (d) 10 class segmentation.

Figure 4: Histograms for the five parameters in the different classes of Figure 3(c)



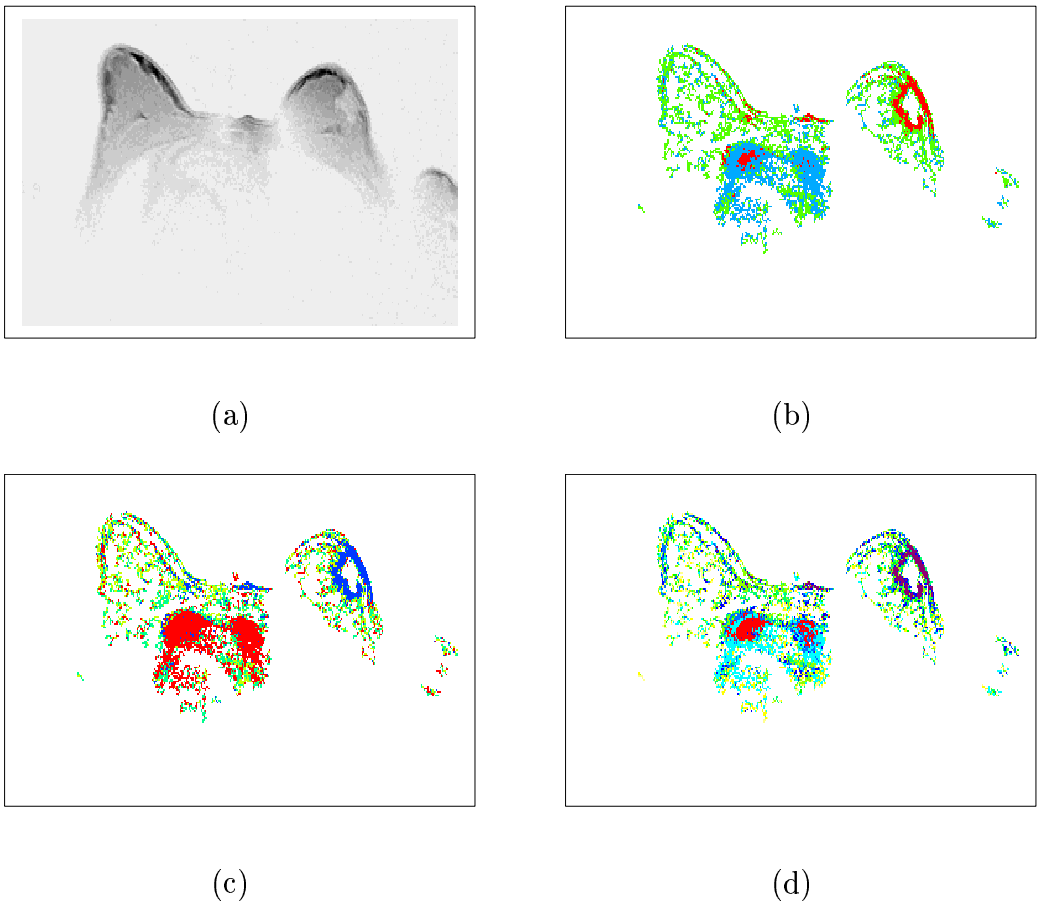


Figure 5: MCLUST classifications for Patient 08, slice 06. (a) reference image, (b) 3 class segmentation, (c) 4 class segmentation, (d) 10 class segmentation.

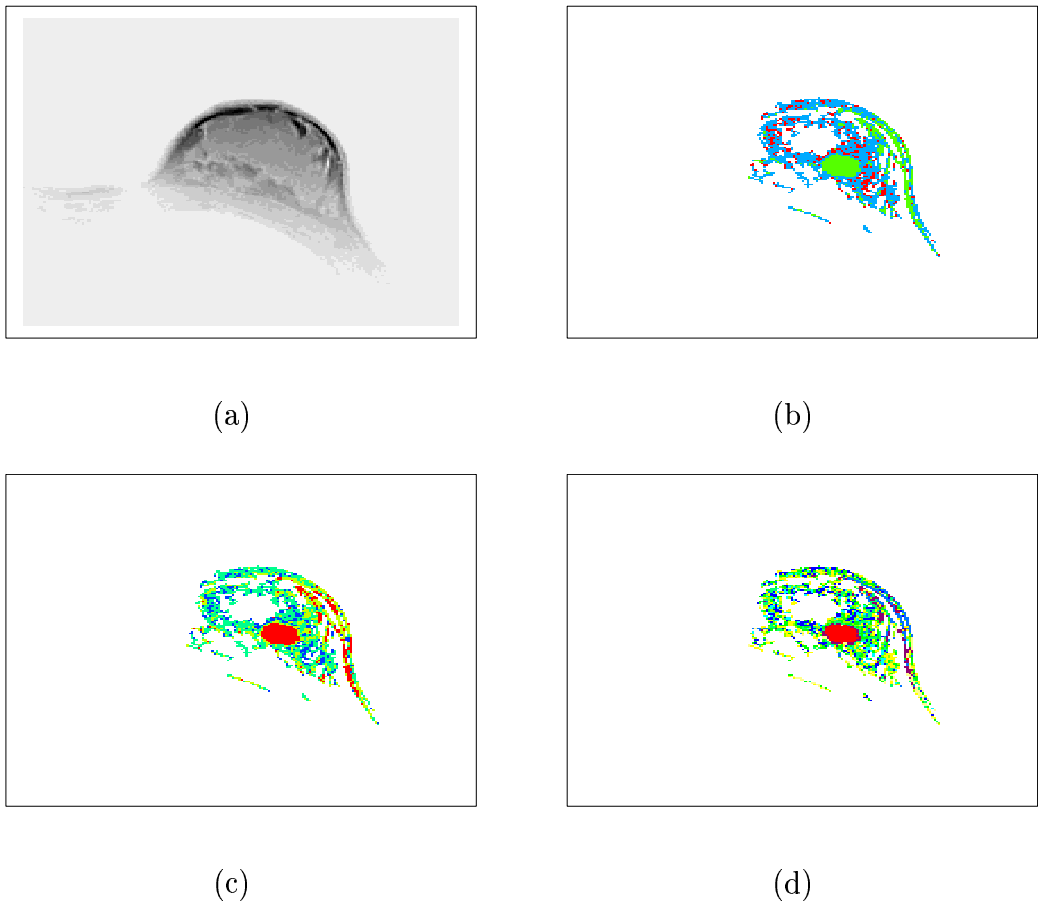


Figure 6: MCLUST classifications for Patient 28, slice 10. (a) reference image, (b) 3 class segmentation, (c) 4 class segmentation, (d) 10 class segmentation.

no further changes. Each pixel is considered in conjunction with its 8 surrounding neighbors, and the rank operator depends on a parameter r (the rank) as follows. Let c be the color which is taken most often by the neighbors of pixel i and let z_i be the current color of pixel i. We denote by $nb(k)$ the number of neighbors of pixel i which have color k . If $nb(c)$ is greater or equal to $nb(z_i) + r$, then the color of pixel i is changed to c, otherwise the color of pixel i remains equal to z_i.

For example, when $r = 1$, this is equivalent to apply the majority rule, according to which a pixel assumes the color which is taken most often by its neighbors. This corresponds to a *median filter* except that the pixels are updated successively rather than simultaneously.

Morphological filters tend to produce very smooth classifications, and more smoothing occurs as r decreases. We refer to morphological filters as *blind restoration*, since they do not make use of the orginal data. They have the clear disadvantage that useful information may be lost in the process, and it is not difficult to find instances in which they perform relatively poorly ([7]).

Nevertheless, morphological filters may be useful in our application to help isolate possible tumors from the initial MCLUST classification. The initial MCLUST classifications for the example image with $K = 3, 4, 10$ are shown in Figure 3. Applying blind restoration with $r = 1$, we obtained, for $K = 3, 4, 10$, the segmentations shown in Figure 7. For $r = 3$, the results are shown in Figure 8.

Blind restoration smooths the data considerably, and, as expected, the results are even smoother with $r = 1$ than with $r = 3$. This could be very useful because it eliminates a great deal of clutter and extraneous minor features in the image. For this image it clearly retains the tumor, and so the result is very satisfactory. However, in images where the tumor is less clearly identified, this level of smoothing runs the risk of eliminating the tumor altogether, and should probably be used only in conjunction with less smoothed images as well. Such images can be obtained using the procedure described in the next section. It has features similar to morphological filters but is statistically based and provided more sophisticated segmentations.

3.2.2 Bayesian Morphology

Bayesian image analysis is based on probability models. It includes a variety of tasks of which one of the most straightforward is image segmentation. The unknown classification, $z = \{z_i,\ i \in S\}$, where S is a set of pixels, is interpreted as a particular realization of a random vector Z. The observed data y is interpreted as a realization of a random vector Y. The vector Y depends on Z through a known conditional probability density function $\mathcal{L}(y|z)$ which incorporates the observed data formation model and the noise model.

Z is recovered by proposing an estimator $\hat{Z} = \hat{Z}(Y)$ of Z on the basis of Y. Bayesian image segmentation methods are based on the following principles. The unknown classification Z is assumed to be a realization of a random field with distribution $p(z)$. Then the classification $\hat{z}$ is based on the posterior density of z, namely

$$p(z|y) \propto \mathcal{L}(y|z)p(z) \ .$$

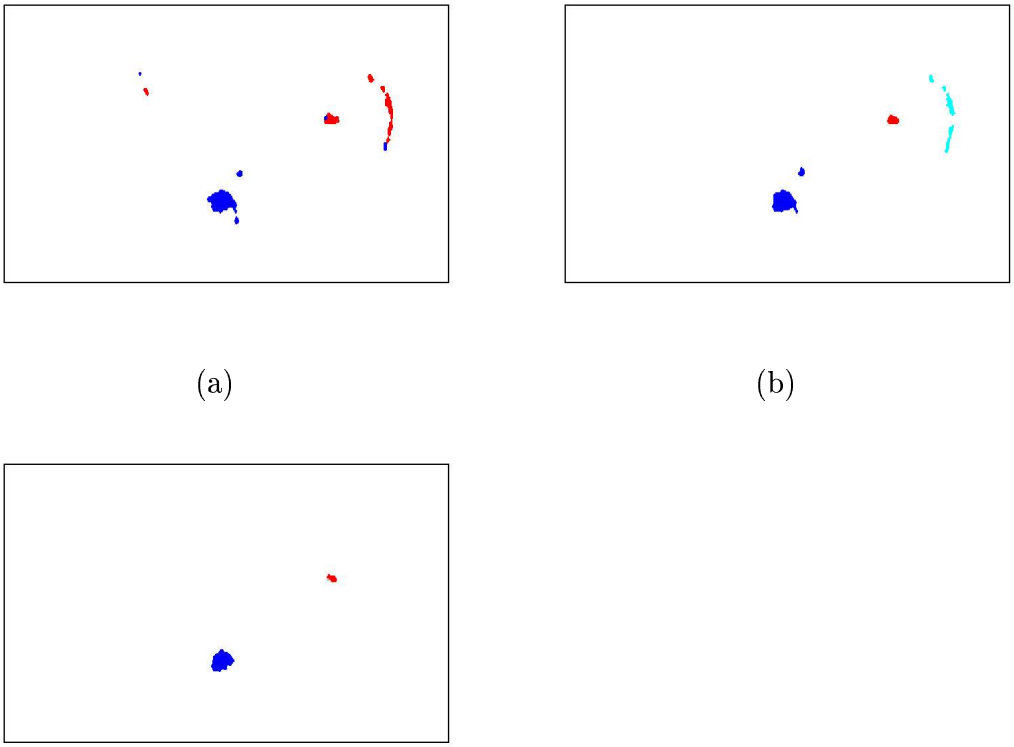


Figure 7: Blind restorations with $r = 1$ using images in Figure 3 as initial classifications. (a) K= 3, (b) K=4, (c) K= 10.

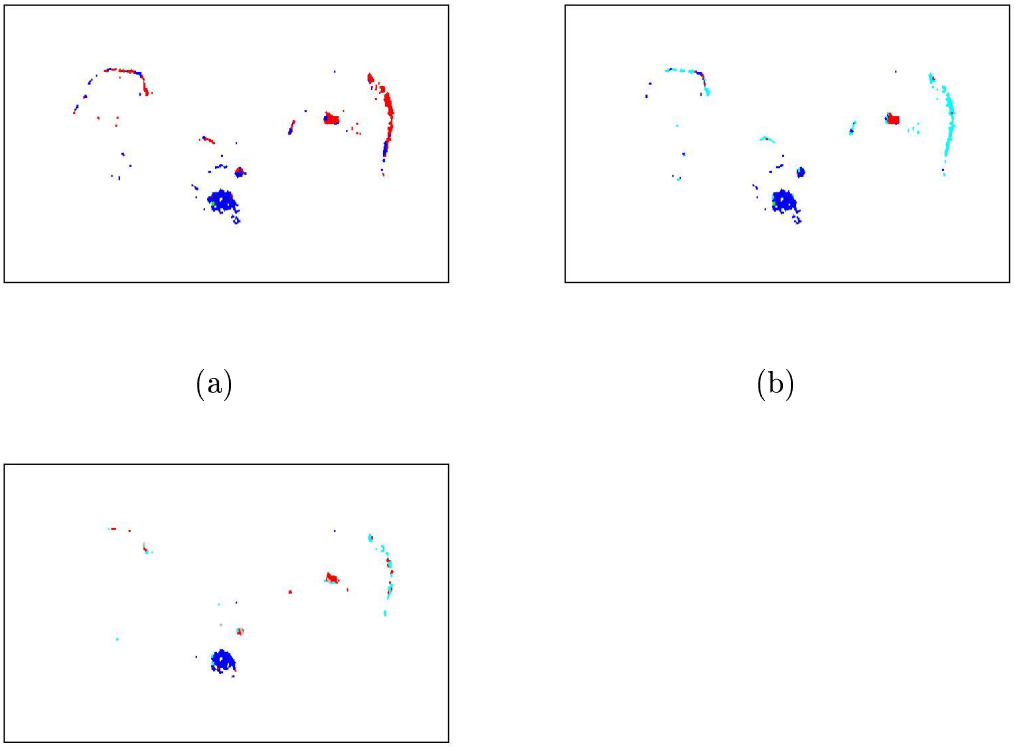


Figure 8: Blind restorations with $r = 3$ using images in Figure 3 as initial classifications. (a) K= 3, (b) K=4, (c) K= 10.

A standard restoration criterion consists of maximizing this density, leading to the maximum a posteriori (MAP) estimate of z .

One of the most popular modeling assumptions is to consider the image z as being a realization of a Markov random field. This means that $p(z)$ satisfies, for all pixels i in S,

$$p(z_i|z_{S\setminus\{i\}}) = p(z_i|z_{N(i)}),$$

i.e., the conditional distribution of $p(z)$ depends only on the values of pixels in a subset $N(i)$ of $S\setminus\{i\}$, called the *neighborhood* of pixel i. Another usual assumption is that, given $Z = z$, the Y_i are conditionally independent and have the same conditional density function $f(y_i|z_i)$ that depends only on z_i. Thus $\mathcal{L}(y|z)$ can be written as the product

$$\mathcal{L}(y|z) = \prod_{i\in S} f(y_i|z_i) \ .$$

Finding the MAP estimate under these assumptions can require heavy computation. A less computationally demanding method that provides a fast approximation to the MAP is the Iterated Conditional Modes (ICM) algorithm [3]. The ICM algorithm is iterative. Given a current estimate $\hat{z}$ of the image, a new one is computed by visiting each pixel in turn. When at pixel i, the current value there is replaced by the value that maximizes the conditional density

$$p(z_i|\hat{z}_{S\setminus\{i\}}, y), \tag{9}$$

given all other current pixel values $\hat{z}_{S\setminus\{i\}}$ and the fixed observation y. This choice is motivated by the following equality

$$p(z|y) = p(z_i|z_{S\setminus\{i\}}, y)\ p(z_{S\setminus\{i\}}|y) \ .$$

When pixels are updated sequentially, choosing values that maximize the conditional probability $p(z_i|z_{S\setminus\{i\}}, y)$ increases the posterior distribution and ensures the convergence to a local maximum of $p(z|y)$.

Under the previous modeling assumptions, maximizing the conditional density (9) is equivalent to maximizing

$$f(y_i|z_i)\ p(z_i|\hat{z}_{N(i)}), \tag{10}$$

since only the dependence on z_i is relevant for the maximization.

In [7], we showed that the ICM algorithm could be formulated using a morphological terminology and proposed *Bayesian morphology*, a procedure that combines the speed of mathematical morphology with the principled statistical basis of ICM. To summarize, in Bayesian morphology, a succession of morphological rank operators is applied but their ranks are estimated at each iteration from the current classification and the original data instead of being predetermined or chosen arbitrary. These methods make more use of the available information than blind restoration, and as a result tend to produce classifications

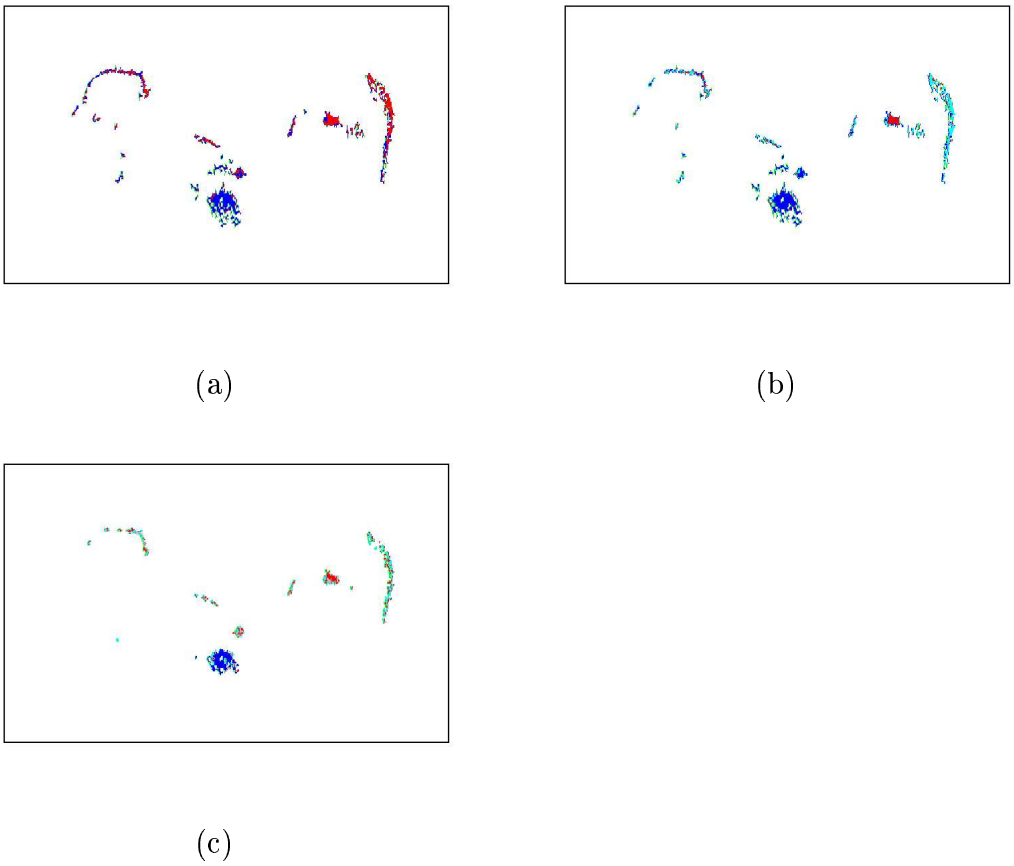


Figure 9: Bayesian morphology with $p = 0.99$ using images in Figure 3 as initial classifications. (a) K= 3, (b) K=4, (c) K= 10.

with more detail. In comparison to blind morphology, less noise is likely to be eliminated, while ambiguous features worthy of further consideration are more likely to be retained.

In addition to an initial classification, the method requires an initial value for a parameter p between 0 and 1 that controls the first operator applied. When p is close to 1, the resulting rank operator is a *median filter* ($r = 1$).

Figure 9 shows the classifications when p is set initially to 0.99. For this example, the classifications when p is set initially to 0.45 are similar for $K = 3$ and $K = 4$. For $K = 10$, the analysis with $p = 0.99$ is clearly better because it eliminates more noise, without removing any features of actual or potential interest.

3.3 Main results

1. Model-based clustering techniques provide informative initial segmentations but do not take into account spatial information.

2. Spatial analysis is useful in refining these initial classifications by:
   (a) giving a simultaneous (color) picture of all the (grey-level) bands;
   (b) eliminating noise and distracting features; and
   (c) enhancing features of potential interest.
3. Partitions into 4 colors/segments were adequate to reveal the tumor of interest. Three segments were too few because the resulting partition was not sufficient to distinguish the tumor from other tissue classes. Ten segments were too many, because the resulting partition divided the tumor pixels among several classes.
4. Spatial smoothing via Bayesian morphology (Figure 9) usefully enhances the initial segmentation, by eliminating some clear clutter, while retaining all the features of potential interest.
5. Moderate blind restoration (Figure 8) provides much more smoothing and a clearer picture, but at the possible cost of eliminating unclear features of potential interest. Strong blind restoration (Figure 7) smooths the image even further, so that there is even more potential loss of useful information.

   As regards the possibly visual use of our images, our tentative recommendation based on this limited study would be to provide radiologists with two different color synthetic images to consider, one in which statistical smoothing has been applied (*e.g.* Figure 9 (b)), and another based on a more drastic heuristic smoothing method (*e.g.* Figure 7 (b)). Note that there is a solid statistical basis for Bayesian morphology, but less so for the more drastic smoothing performed by blind restoration.

These classifications provide a first tool for guiding diagnosis. Often the lesion is easily identified and the radiologist can be asked to select regions that are suspicious or otherwise of interest to be further examined. In the next section, we show that additional information is present in the data that can be used for a more automatic detection and localisation of lesions of interest.

4 Region Of Interest Selection

One limitation of most ROI analysis procedures is that they are based on a subjective selection of the region. In this section, we propose a way to automatically select an ROI using the few color classification images produced in the first part of our study (Section 3). Our second step consists then of studying the values of each of the five variables within each class, on the basis of which we propose an ROI selection. Besides, the relevance of this selection can be visually assessed by looking at the whole classification images usually clearly showing the different features of interest. The analysis can then be carried out by looking at the shape of the ROI (morphologic feature analysis) or at the curves (kinetics analysis)

for pixels in a selected cluster or region in order to identify the nature of the lesion. We first tested our method using classifications obtained from an initial partition with 10 segments followed by a strong blind restoration (median filter). In the cases we analyzed, they did make the lesion appear very clearly with almost no clutter.

4.1 Deciding which of the segments is the ROI

In breast MRI, lesions are usually identified because they enhance after intravenous injection. Although the pathophysiologic basis of this feature has not been fully elucidated yet, some fundamental facts are known that should help in designing tools for lesion detection and differential diagnosis ([15] and [13]). In our study, we focus on rapidly and strongly enhancing lesions. The rationale for looking at enhancement rates is based on the observation that malignant lesions tend to have higher enhancement velocities than benign ones (see [15]), which, in our examples, proved to be a good indicator of suspicious regions. Good results are obtained when focusing on strong enhancements characterized by a large *difference at peak*, *i.e.* the absolute increase of intensity between the beginning of the signal and the time at which the signal peaks. For a given classification image, the mean value of *difference at peak* is computed for each class in the image. We then select the class with the largest mean value as containing the ROI. Further processing might be necessary to select then the ROI as the biggest connected component in the class (see Figure 11 (e)). Note that another criterion for rapid enhancement would be a small time to peak. This would correspond to a large value for the variable referred to as *time to peak* in our data, which is a linear transformation of the real time to peak. However, this criterion would sometimes select the heart area which enhances faster than the lesion of interest. Using the largest difference at peak instead, the class containing the suspicious region is always selected as desired. As an illustration, Table 2 shows the values for the classifications shown in Figure 10. The suspicious regions, in red, are the ones selected when using a maximum mean difference at peak criterion. The heart area is selected if a maximum mean time to peak criterion is used instead.

Note also that, on a practical point of view, the difference at peak parameter can be used to determine a meaningful color assignment. In most classification methods, images are produced using colors (or equivalently class labels) artificially assigned to the different classes. As shown on Table 2, the same region (*e.g.* the heart area) can be assigned to different class numbers depending on the classification (*e.g.* class 2 and 3 for the $K = 3$ and $K = 4$ classifications of Table 2). In our study, we propose to automatically display our results using the highest difference at peak criterion so that suspicious regions can be marked with a specific label and always displayed with a specific color (*e.g.* red).

4.2 Enhancement kinetics analysis

In diagnosis, an important point is to produce a *good* estimated pattern of uptake which should be representative of the lesion under study. A first idea is to use the ROI and compute the mean of all signals in the region. This should get rid of some noise but may be

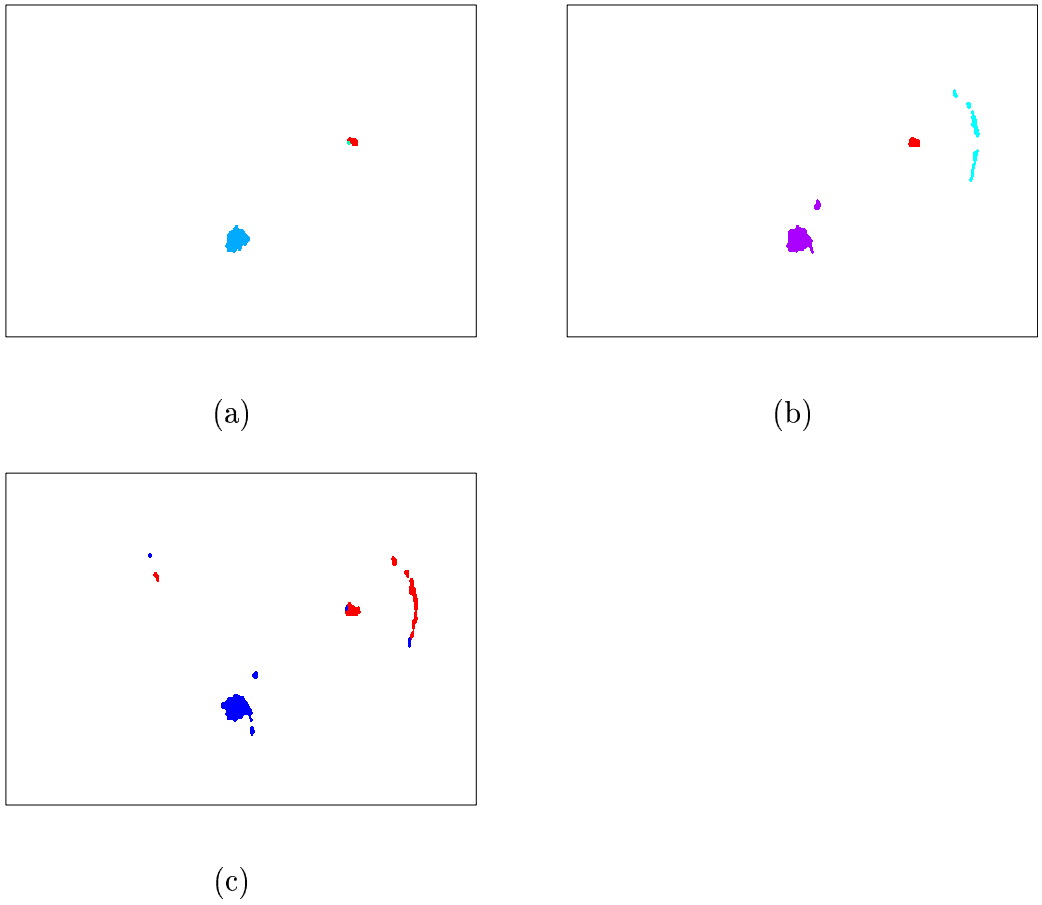


Figure 10: Median filter ($r = 1$) on MCLUST classifications for Patient 05, slice 09. (a) K= 10, (b) K=4, (c) K= 3. The regions in red are the ones with the largest mean difference at peak.

Patient 05 slice 09			
K=3	mean difference at peak	mean time to peak	interpretation
class 2	10363	**-2061**	heart
class 3	**17179**	-4578	lesion and skin
K=4			
class 2	13006	- 5468	skin
class 3	10519	**-1834**	heart
class 4	**27068**	-2819	lesion
K=10			
class 8	**29840**	-2222	lesion (main)
class 9	25512	-3063	lesion (border)
class 10	9736	**-1582**	heart

Table 2: Patient 05, slice 09. The classifications in Figure 10 into $K = 3, 4, 10$ classes are used to compute mean values for variable *difference at peak* in each non-background component. The largest value corresponds to the lesion of interest while taking the largest mean *time to peak* would select the heart area.

biased if the region is too big. We could also select one or a few pixels in the ROI with the highest probability.

Figures 11 (g)-(i) are mean curves computed using classifications in Figures 11 (d)-(f). These curves are obtained by averaging all signals in the ROI's defined as the largest connected component in the red region of Figures 11 (d)-(f). We also investigated other ways to compute such *mean* curves using weights. The idea is to give more weight to pixels in the ROI which are typical of the lesion and less weight to pixels for which we are more uncertain. The question is then to find reasonably good weights as automatically as possible.

Let S_{roi} denote the set of pixels in the ROI and $w = \{w_i, i \in S_{roi}\}$ a vector of weights associated to each pixels. A mean curve can then be computed by multiplying each signal in the ROI by $w_i / \sum_{i \in S_{roi}} w_i$ before summing all the signals. For w_i in $\{0, 1\}$, using weights is equivalent to select some of the pixels. The question is which pixels to keep ($w_i = 1$) and which to disgarde ($w_i = 0$). As an illustration, we first kept the 33% pixels in the ROIs with the highest difference at peak values. Figure 12 shows the selected pixels and the corresponding mean curves in three cases. We then used the conditional probabilities estimates provided by MCLUST (see section 3.1) and kept the pixels for which the probability to belong to the lesion class was very close to one (we used all pixels within 10^{-7} of 1). Results are shown in Figure 13 (upper curves). More generally, the conditional probabilities estimates can be used as weights. This gives the results in Figure 13 (middle curves). The curves are very similar to the mean curves in Figure 11 (g)-(i) because in the lesions the conditional probabilities estimates are close to one.

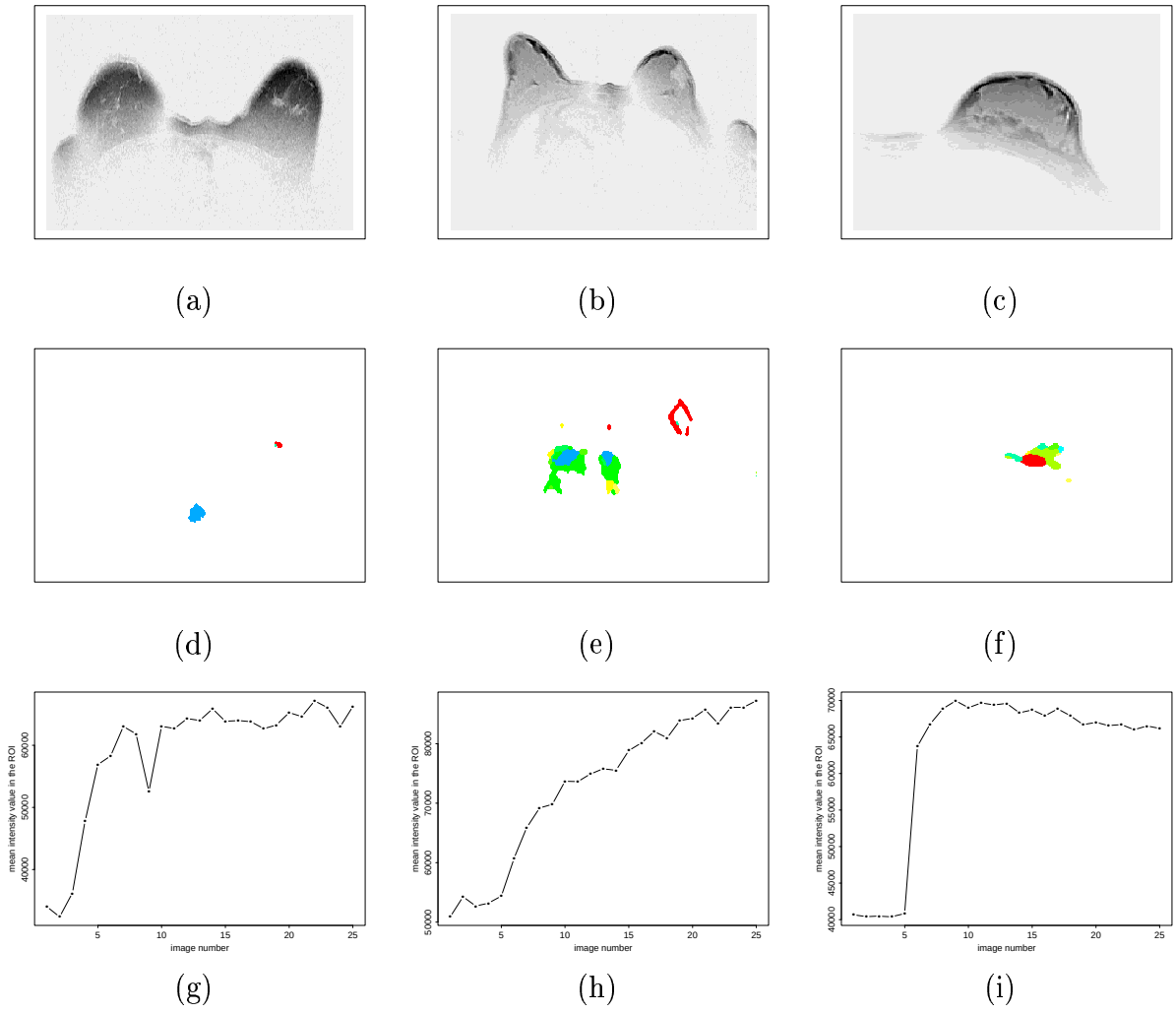


Figure 11: ROIs and associated mean curves in three cases. (a), (b), (c) dynamic image at $t = 1$ for patient 05, slice 09, patient 08, slice 06 and patient 28, slice 10. (d), (e), (f) ROIs selections (largest connected component in the red regions). (g), (h), (i) mean time-intensity signals in the ROIs.

For our cases, it seems that selecting only few pixels in the ROIs, those with the highest difference at peak values or with the highest conditional probabilities, provides mean curves where features (slope enhancement, washout, etc.) are more clearly marked.

4.3 Time-intensity signal analysis in the ROI

Assuming that we have assigned a representative curve to the lesion under study, our third step is then to focus on the shape of this curve. We take into account information from

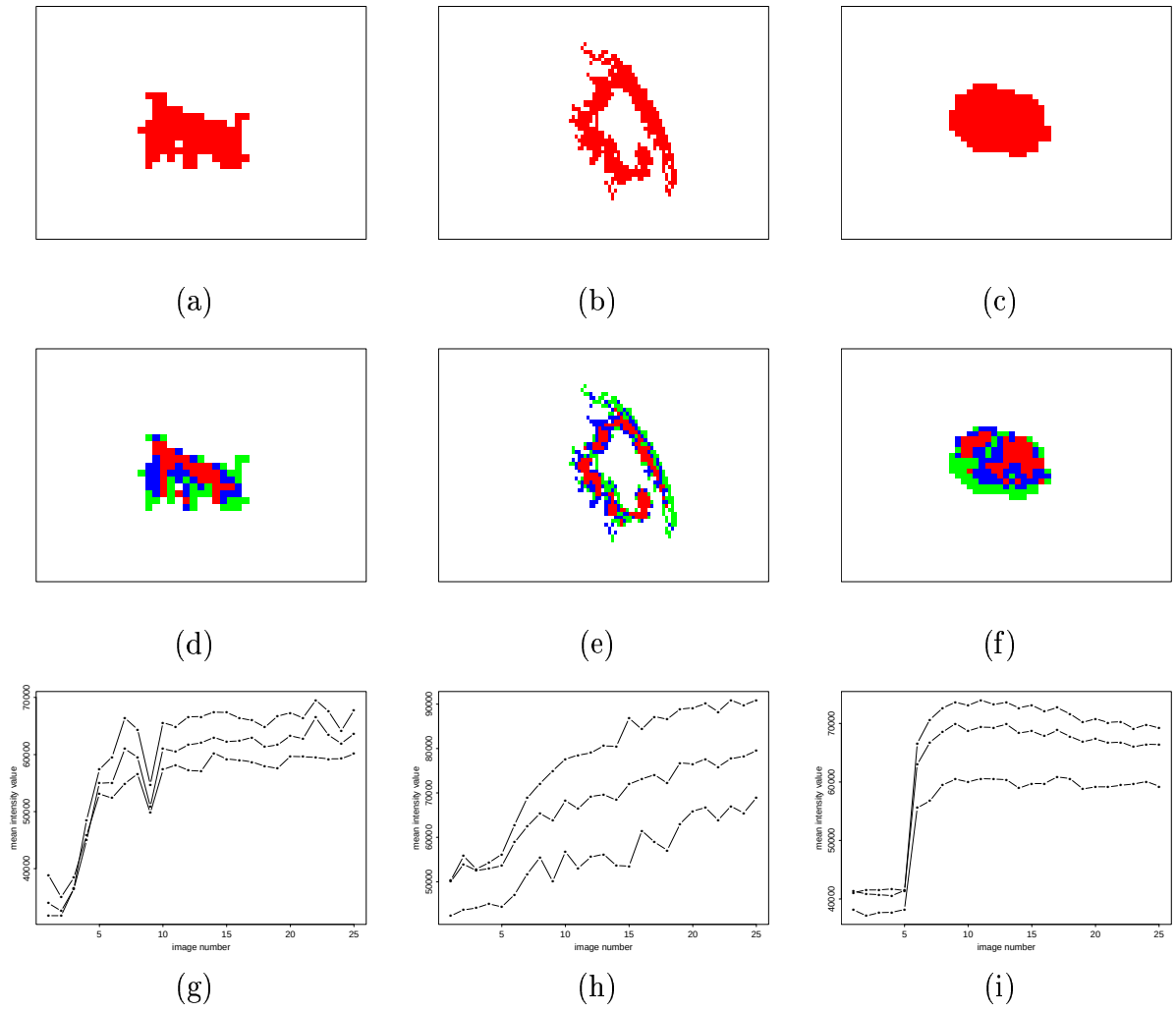
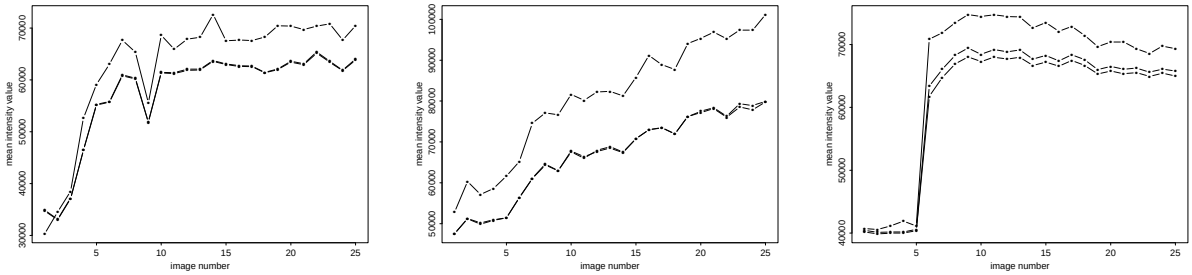


Figure 12: ROIs using 33% and 67% difference at peak quantiles, and associated mean curves in three cases. (a), (b), (c) zoomed ROIs from MCLUST classifications with $K = 4$, for patient 05, slice 09, patient 08, slice 06 and patient 28, slice 10. (d), (e), (f) ROIs segmentations using difference at peak quantiles: highest 33% values in red, lowest 33% values in green. (g), (h), (i) mean time-intensity signals in each class (upper curve for the red class, lower curve for the green class).



(a) (b) (c)

Figure 13: Mean curves using conditional probabilities estimates from MCLUST. (a) patient 05, slice 09, (b) patient 08, slice 06 (c) patient 28, slice 10. Upper curves: mean curves when selecting pixels with conditional probability very close to 1. Middle curves: weighted mean curves when using conditional probabilities as weights. Lower curves: mean curves using all pixels in the ROIs.

other sources, [16], [15] and [13]. They distinguish three patterns of signal intensity curves on the basis of three characteristics, the *enhancement rate*, the presence of a *plateau* and that of a *washout slope*. Type I shows a mono-phasic enhancement that persists until the late post-contrast period (linear time course). This type is indicative of a benign lesion. Type II is a bi-phasic enhancement where signal intensity reaches the maximum approximately 2-3 minutes after injection and stays at this level (plateau curve). This type has been observed in both benign and malignant lesions. Type III is characterized by a washout enhancement. As in type II, peak enhancement is already reached in the early post-contrast period but then it is followed by an intensity loss. A type III pattern strongly supports the diagnosis of a malignant breast tumor.

The curves computed for the analysis of patients 05, 08 and 28 are easily assigned to type II, I and III respectively. This is consistent with the known diagnostics which correspond respectively to a carcinoma, a benign lesion and a carcinoma and confirms that our procedures are of interest for the differentiation between malignant and benign lesions.

5 Discussion

We have applied model-based clustering followed by spatial smoothing techniques to data derived from an MR breast image with the object of producing one or more classifications useful to experts for breast cancer diagnosis. In particular, the classifications obtained after morphological filtering (Figures 8 (b) and 7 (b)) clearly indicate the tumor in the image that we analyzed.

However, this particular image may not represent a typical situation since the tumor is relatively easy to distinguish by eye. While the more conservative segmentations (*e.g.* Figure 9 (b)) lack smoothness to some extent, they may well retain features of potential

interest when applied to images in which tumors are less apparent. This trade-off between smoothness and resolution needs to be assessed by further empirical research on other images. The trained human eye can often discern features that are not completely delineated, but it cannot reconstruct features that have been removed. However the ideal for radiologists would presumably be to have very definite, ideally black and white, images showing the tumor and non-tumor areas.

The present report is quite limited in that it is largely exploratory and has been based on the analysis of only a few sets of MR images. However, it does suggest that the combination of model-based clustering and statistical methods for spatial segmentation we have investigated is quite promising for tumor identification.

Acknowledgments

This research was supported by the Office of Naval Research under grants N00014-96-1-0192 and N00014-96-1-0330 and Toshiba America MRI, inc. The authors are grateful to Leon Kaufman, David Haynor, Doug Ortendahl for useful comments and discussions and to Andrew Jianhua Li for providing the fitting procedure software.

References

[1] J. D. Banfield and A. E. Raftery. Model-based Gaussian and Non-Gaussian Clustering. *Biometrics*, 49:803–821, 1993.

[2] H. Bensmail, G. Celeux, A.E. Raftery, and C.P. Robert. Inference in model-based cluster analysis. *Statistics and Computing*, 7:1–10, 1997.

[3] J. Besag. On the statistical analysis of dirty pictures. *Journal of the Royal Statistical Society, series B*, 48:259–302, 1986.

[4] G. Celeux and G. Govaert. Gaussian Parsimonious Clustering Models. *Pattern Recognition*, 28:781–793, 1995.

[5] A. Dasgupta and A. E. Raftery. Detecting features in spatial point processes with clutter via model-based clustering. *Journal of the American Statistical Association*, 93:294–302, 1998.

[6] A. P. Dempster, N. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via the EM algorithm (with discussion). *Journal of the Royal Statistical Society, series B*, 39:1–38, 1977.

[7] F. Forbes and A. E. Raftery. Bayesian Morphology: Fast Unsupervised Bayesian Image Analysis. *Journal of the American Statistical Association*, 94:555–568, 1999.

[8] C. Fraley. Algorithms for model-based gaussian hierarchical clustering. *SIAM Journal of Scientific Computing*, 20(1):270–281, 1998.

[9] C. Fraley and A. E. Raftery. Model-based clustering, discriminant analysis and density estimation. Technical Report 380, University of Washington, Department of Statistics, October 2000.

[10] C. Fraley and A.E. Raftery. How many clusters? which clustering methods? - answers via model-based cluster analysis. *The Computer Journal*, 41(8):578–588, 1998.

[11] C. Fraley and A.E. Raftery. MCLUST: Software for model-based cluster analysis. *Journal of Classification*, 16:297–306, 1999.

[12] H.P. Friedman and J. Rubin. On some invariant criteria for grouping data. *Journal of the American Statistical Association*, 62:1159–1178, 1967.

[13] T. H. Helbich. Contrast-enhanced magnetic resonance imaging of the breast. *European Journal of Radiology*, 34:208–219, 2000.

[14] A. Jianhua Li. Streaming Data Analysis Algorithms and Implementations. Technical report, TAMI RIL report, 1999.

[15] C.K. Kuhl. MRI of breast tumors. *European Radiology*, 10:46–58, 2000.

[16] C.K. Kuhl, P. Mielcareck, S. Klaschik, C. Leutner, E. Wardelmann, J. Gieseke, and H. H. Schild. Dynamic Breast MR Imaging: Are Signal Intensity Time Course Data Useful for Differential Diagnosis of Enhancing Lesions? *Radiology*, 211:101–110, 1999.

[17] G. J. McLachlan and T. Krishnan. *The EM Algorithm and Extensions*. Wiley, 1997.

[18] C. Posse. Hierarchical Model-Based Clustering for Large Data Sets. *To appear in Journal of Computational and Graphical Statistics*, 2001.

[19] A.J. Scott and M.J. Symons. Clustering methods based on likelihood ratio criteria. *Biometrics*, 27:387–397, 1971.

[20] D. M. Titterington, A. F.M. Smith, and U. E. Makov. *Statistical Analysis of Finite Mixture Distributions*. John Wiley, Chichester, U.K., 1985.

Unité de recherche INRIA Rhône-Alpes
655, avenue de l'Europe - 38330 Montbonnot-St-Martin (France)

Unité de recherche INRIA Lorraine : LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex (France)
Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)
Unité de recherche INRIA Rocquencourt : Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex (France)
Unité de recherche INRIA Sophia Antipolis : 2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399