

Multi-vector Segmentation of Breast MR Images via Hidden Markov Random Fields

STYLIANI PETROUDI, GEORGIOS KETSETZIS, MICHAEL BRADY
Medical Vision Laboratory, Department of Engineering Science
University of Oxford
Ewert House, Ewert Place, Summertown, Oxford, OX2 7BZ
UNITED KINGDOM

Abstract: In this paper we apply multi-vector Hidden Markov Random Fields to tissue segmentation of Magnetic Resonance (MR) breast images. Our proposed method performs segmentation using a stack of 3 MR breast slices $1mm$ apart. The approach takes into account neighborhood voxel information rather than merely neighborhood pixel information and the results are anatomically more plausible in comparison with standard two-dimensional segmentation techniques. The proposed algorithm incorporates an initial correction of the bias field, and automatic background removal. The k-means algorithm is used to provide an initial segmentation/classification. This classification allows for tissue parameter estimation, providing an initialization of probabilistic moments that are incorporated into a Gaussian probability model for each tissue class. The class labels follow a Gibbs distribution and the energy function is a sum of potentials taken from a multilevel logistic model for Markov Random Fields. The segmentation is obtained via maximization of the posterior probability distribution function and the solution is found by application of Besag's Iterated Conditional Modes (ICM) algorithm. After each ICM iteration, the tissue parameters are updated. The process continues iteratively until convergence. The segmentation results demonstrate anatomically plausible breast tissue segmentation and we expect the method to aid real time automatic segmentation of breast tissue, particularly in diagnosis of pathology.

Key-Words: *Magnetic Resonance Imaging, Breast Imaging, Hidden Markov Random Fields.*

1 Introduction

Segmentation of breast Magnetic Resonance (MR) images into different tissue classes, such as fat, healthy and malignant tissues is an important task during the diagnostic process. Breast MR images have a number of features. They are statistically complex as they are not piecewise constant and they possess a large number of classes. Moreover, MR breast images do not have high contrast between different tissues. By carefully choosing pulse sequence parameters and gradients, it is possible to highlight different components in the object being imaged and produce high-contrast images in order to facilitate

segmentation and classification. On the other hand, ideal imaging conditions are never realized in practice. Electronic noise, the bias field (intensity inhomogeneities in the RF field) and the partial volume effect (PVE) -multiple tissue class occupation within a voxel- cause classes to overlap in the image intensity histogram.

A wide variety of approaches have been proposed for MR image segmentation and they may be roughly divided into two categories: structural and statistical. Structural methods include various edge detection algorithms that have been applied to extract boundaries between different tissues [1]. However, such algorithms are vulnerable to artifacts and noise,

and fail in breast imaging because of the anatomical tissue complexity. Region growing [2] is another popular structural approach, in which an image is divided into small regions, which can be considered as “seeds” that grow under certain criteria; but it is not considered robust. Statistical methods approach segmentation/classification from a completely different perspective. Such methods label pixels according to probability values, usually determined by the intensity distribution of the image. Thresholding methods are simple statistical methods that are unlikely to produce reliable results for breast MR images since they are not robust under the presence of noise and have no way of incorporating discrepancies resulting from PVEs. Depending on whether a specific functional form for the density model is assumed or not, a statistical approach can either be parametric or non-parametric. Maximum *a posteriori* probability (MAP) or maximum likelihood (ML) are the most usually used principles for such attempts. In such methods the probability density functions (*pdfs*) of the different tissue classes need to be chosen very carefully.

In contrast to conventional approaches for breast MR segmentation, we perform tissue segmentation into a set number of classes by applying Hidden Markov Random Field - Maximum *a posteriori* (HMRF-MAP) labelling to stacks of 3 consecutive coronal axial MR breast slices that are 1mm apart. We seek the labelling of an image that is considered to be a realization of a MRF through minimizing the risk of misclassification, which is equivalent to maximizing the posterior probability. The minimal risk estimate is known as the MAP estimate. Under the MRF model the MAP estimate may be found by minimizing the posterior energy function [3]. This minimization is achieved by applying the Iterated Conditional Modes (ICM) algorithm proposed by Besag. ICM converges to a local minimum in only a few iterations and has been shown to outperform global methods when the image is of decent quality [3].

Our breast MR image segmentation algorithm is based on the algorithm developed by Petroudi and Brady [4] where texture feature vectors are used to represent each pixel in an image. Multi-vector Gaussian Hidden Markov Random Fields (GHM-

RFs), based on feature vectors incorporate neighborhood information and achieve tissue segmentation as the result of applying iterations of Expectation-Maximization (EM) followed by ICM until a convergence criterion is met. The paper is organized in the following way. Section 2 provides background on HMRFs and multi-vector image segmentation, followed by formulation of the problem for breast MR image segmentation. Section 3 provides the complete algorithm for the proposed setup and the segmentation. Finally, Section 4 demonstrates results from application of the method to real datasets and discusses the performance of the proposed method as well as future research directions.

2 Segmentation using parametric and non-parametric MRFs

The objective of our work is segmentation/classification of MR breast images into different types of tissues. Automatic classification results, based only on the one-dimensional (1-D) intensity image, whether they take into account the pixel’s neighborhood or not, have been shown to be inferior to the segmentation of corresponding vector images [5]. Filter banks can be used to provide a multi-vector feature representation of an image. However, it has been shown that using non-parametric representation of MRFs achieves similar and even superior results in material classification [4].

A MRF is a collection of random variables which are defined on a finite lattice, and where each variable interacts with some subgroup of that lattice termed its neighborhood [6]. In modelling the interaction between data and model, certain parameters are required. Depending upon whether these are known or not, two paradigms result. MRF is the paradigm corresponding to *a priori* known statistical parameters. HMRF is the paradigm where statistical parameters are not known and need to be estimated. A HMRF is a stochastic process generated by a MRF whose state sequence cannot be observed directly but only through a field of observations. By imposing contextual constraints, we expect neighboring pixels to have the same class labels (in the case of piecewise

constant images) or similar intensities (in the case of piecewise continuous images). This is achieved through characterizing mutual influences among pixels using conditional MRF distributions [3].

MRI is a 3-D imaging modality with quite good plane spatial resolution, in the range 1 to 3 mm and with slice thickness in the range of 1 to 2 mm with no gap. These resolutions allow us to take advantage of the information available in the 3-D neighborhood corresponding to each point. To achieve this we employ MRFs in 2 directions (in the plane of the image and also in the direction of the slice acquisitions: perpendicular to the plane of the image) as well as in 2 settings: a parametric and a non-parametric. Each MRI slice is segmented into different classes using a statistical method based on HMRFs, that incorporate neighborhood information in the direction of the image plain. However, instead of using a scalar value to represent each point, we use a vector, consisting of the intensity at that location in the previous and following slice, hence incorporating a non-parametric local MRF in the direction of slice acquisition. Segmentation becomes a matter of multi-vector image classification into a set number of tissue classes with the use of HMRFs.

The procedure of estimating the unknown parameters for the HMRFs is known as model fitting. As we have introduced the idea of segmenting an image by segmenting a multi-vector representation of the image using HMRFs we need to introduce and make the assumption of a multivariable Gaussian emission function [3]. To fit the GHMRF at each iteration, we need an estimate of the means and covariances of the vector classes. For computational efficiency, we evaluate the means and covariances using their general statistical definition. Segmentation/classification of the image is the result of MAP estimation of the tissue labels by application of the ICM. We show that by incorporating, a multi-vector representation of a breast MR image using the non-parametric local MRF in the direction perpendicular to the plane of the image, a parametric GHMRF in the direction of the plane, an initial class estimation using k -means and iterative MAP estimation using ICM and mean and covariance updates, an accurate and robust segmentation of MR breast images is achieved.

3 Methodology

The first step in any MR breast image segmentation consists of bias field and background pixel removal. The bias field correction follows the scheme proposed in [7], where intensity inhomogeneities are assumed to form a continuous field in the image and are approximated by Legendre polynomials. The associated parameters are computed via the solution of a non-linear energy minimization problem, computed from the distance of pixel values and class means. We thus remove the frequency artifact due to inhomogeneities in the radio-frequency field. Automatic background removal is achieved by Otsu’s thresholding method. Background intensities have a much lower value than breast intensities and the global threshold found by Otsu’s method minimizes the intra-class variance. Automatic background removal is performed to speed up further processing of the images.

In what follows, we establish the multi-vector image to be used for segmentation by stacking the 3 consecutive MR slices, with the breast MR slice to be classified, occupying the middle position. This is the representation that incorporates the non-parametric MRF in the direction of slice acquisition. For segmentation of the multi-vector image using GHMRFs and ICM we adopt the same notation as in [5]. Let X be a realization of a random field \mathbf{X} defined on a 2-D lattice \mathcal{S} , where \mathbf{X} is the set of class labels on the underlying image of a p -dimensional random field \mathbf{Y} on \mathcal{S} . We suppose that \mathbf{X} is comprised of pixels which belong to one of the m classes. For the purposes of the present paper we define a set number of tissue classes. The statistical parameters representing each class are unknown and need to be established at each iteration. Let c denote a clique and \mathcal{C} be the set of all cliques on \mathcal{S} . According to the Hammersley-Clifford Theorem, the joint *pdf* of X is a Gibbs distribution of the form

$$f(X) = \frac{1}{Z} e^{-U(X)} \tag{1}$$

where

$$Z = \sum_X e^{-U(X)} \tag{2}$$

is the partition function and $U(X)$ is an energy func-

tion defined by:

$$U(X) = \sum_{c \in \mathcal{C}} V_c(X) = \sum_{s \in \mathcal{S}} \sum_{j \in \mathcal{N}_s} \beta(1 - \delta(x_s - x_j)),$$

$$\forall s, s' \in \mathcal{C}, s \neq s' \quad (3)$$

with $V_c(X)$ being the potential function whose argument, X , is an element of the clique. \mathcal{N}_s denotes the neighborhood pixels of x_s in the associated clique. The MRF we consider here is a multilevel logistic model that has a second order neighborhood system with pairwise cliques: where x_s is the realization of the 1-D random field \mathbf{X} on \mathcal{S} . β can be interpreted as edge penalty. The observed image \mathbf{y}_s is obtained when the noise \mathbf{w}_s is superimposed on the signal $g(\mathbf{x}_s)$:

$$\mathbf{y}_s = g(\mathbf{x}_s) + \mathbf{w}_s \quad (4)$$

where $g(\mathbf{x}_s)$ is a function that maps the underlying label \mathbf{x}_s to its associated attribute $\underline{\mu}_{x_s}$ and Σ_{x_s} . The \mathbf{w}_s 's are independently distributed Gaussian random vectors with zero mean and unknown covariance matrix, which is class conditional. Therefore, the density of \mathbf{Y} , given the underlying true image $\mathbf{X} = X$, is

$$f(Y|X) = \prod_{s \in \mathcal{S}} f(\underline{y}_s | x_s). \quad (5)$$

Based on the observed image Y , the problem is to classify the observed random vector \underline{y}_s into one of the m different classes, subject to iteratively estimated parameters for the multivariable Gaussian distribution of the conditional *pdf* at every position. Based on the assumption that the pixel intensity y_s follows a Gaussian distribution with parameters $\theta_s = \{\mu_\ell, \sigma_\ell\}$, given the class label $x_s = \ell$,

$$p(y_s | x_s) = g(y_s; \theta_\ell) = \frac{1}{\sqrt{2\pi\sigma_\ell^2}} \exp\left(-\frac{(y_s - \mu_\ell)^2}{2\sigma_\ell^2}\right).$$

$$(6)$$

and the conditional independence assumption of \mathbf{y} , equation (5), the joint likelihood probability used for segmentation becomes:

$$f(Y|X) = \frac{1}{(2\pi)^{Np/2}} \cdot e^{-\frac{1}{2} \sum_{s \in \mathcal{S}} [(y_s - \mu_{x_s})^T (\Sigma_{x_s})^{-1} (y_s - \mu_{x_s}) + \ln(\|\Sigma_{x_s}\|)]} \quad (7)$$

where \underline{y}_s and $\underline{\mu}_{x_s}$ are p -dimensional vectors while Σ_{x_s} is a $p \times p$ matrix and N is the total number of pixels in the image. $\underline{\mu}_{x_s}$, Σ_{x_s} are the mean and covariance associated with each class. Equation (7) can be written in the form $f(Y|X) = \frac{1}{Z'} e^{-U(Y|X)}$ with likelihood energy

$$U(Y|X) = \sum_{s \in \mathcal{S}} U(y_s | x_s)$$

$$= \frac{1}{2} \sum_{s \in \mathcal{S}} [\mathcal{H}(y_s, x_s) + \ln(\|\Sigma_{x_s}\|)] \quad (8)$$

where

$$\mathcal{H}(y_s, x_s) = (y_s - \mu_{x_s})^T (\Sigma_{x_s})^{-1} (y_s - \mu_{x_s}). \quad (9)$$

Since $\ln f(\mathbf{x}|\mathbf{y}) = -U(\mathbf{x}|\mathbf{y}) + C$ for some constant C , according to the MAP criterion the segmentation is given as:

$$\hat{x} = \arg \max_{\mathbf{x} \in \mathbf{X}} f(\mathbf{x}|\mathbf{y}) = \arg \min_{\mathbf{x} \in \mathbf{X}} (U(\mathbf{y}|\mathbf{x}) + U(\mathbf{x}))$$

$$= \arg \min_{\mathbf{x} \in \mathbf{X}} \sum_{s \in \mathcal{S}} \left[\frac{1}{2} \mathcal{H}(y_s, x_s) + \ln(\|\Sigma_{x_s}\|) + \sum_{j \in \mathcal{N}_i} \beta(1 - \delta(x_s - x_j)) \right]. \quad (10)$$

Segmentation/classification is the result of minimizing the solution of equation (10) iteratively. Our solution is given via the ICM method that requires an initial segmentation of the image, which we establish by application of the k -means algorithm. The ICM algorithm uses the ‘‘greedy’’ strategy in the iterative local minimization and convergence is guaranteed after only a few iterations. Given the data \mathbf{y} and the other labels $x_{\mathcal{S}-\{s\}}^{(k)}$, the algorithm sequentially updates each $x_s^{(k)}$ into $x_s^{(k+1)}$ by minimizing $U(x_s | \mathbf{y}, x_{\mathcal{S}-\{s\}})$, the conditional posterior probability, with respect to x_s . After the initial classification we evaluate the class means and covariances needed for the ICM optimization method using the following equations:

$$\hat{\underline{\mu}}_l = \frac{1}{r_l} \sum_{s \in \mathcal{S}} \underline{y}_s, \quad x_s = l \quad (11)$$

$$\hat{\underline{\Sigma}}_l = \frac{1}{r_l} \sum_{s \in \mathcal{S}} (\underline{y}_s - \underline{\mu}_l)(\underline{y}_s - \underline{\mu}_l)^T, \quad x_s = l \quad (12)$$

where $l = 1, \dots, m$ and r_l is the number of points in class l . The method continues by iterating the ICM algorithm followed by parameter updating using equations (11) and (12) until the segmentation converges.

4 Results Discussion and Conclusion

We have applied the presented local MRF multi-vector HMRF breast MR segmentation algorithm to 12 sets of images with excellent results. Two such segmentation examples are shown in Figures 1 and 2. As it can be seen, choosing 6 tissue classes results in good correspondence with breast anatomy. In our examples these tissue classes are subsets of the more general categories of fibroglandular tissue and fat. A larger number of classes may be useful in better classification of PVE voxels, in the direction perpendicular to the image acquisition. As the set number of tissue types increases it is more probable that such points will create their own class, leading us to overcome any difficulties arising due to PVE. The segmentation results will be even better when the algorithm is applied to Contrast Enhanced MR images of the breast as tissue differentiation will be superior.

This paper presents a new approach in segmentation of breast MR images using multi-slice intensity information, utilizing local non-parametric MRFs combined with GHMRFs to impose contextual constraints on image voxels. Moreover, it incorporates information from neighboring slices to assign a tissue class. Segmentation is the result of a MAP estimation applied on the multi-vector image and obtained using ICM. Such a segmentation approach achieves a more physiologically plausible characterization of contextual constraints between neighboring voxels and as a result it is more reliable in real time segmentation than conventional 2-D segmentation methods. The illustrated results, obtained after applying the method to 4 patients, capture reliably the anatomical information contained in the original images and demonstrate that the method works in practice. The algorithm is fast and the segmentation steps are completely automated. We are now evaluating a version of this algorithm where the number of tissue classes is

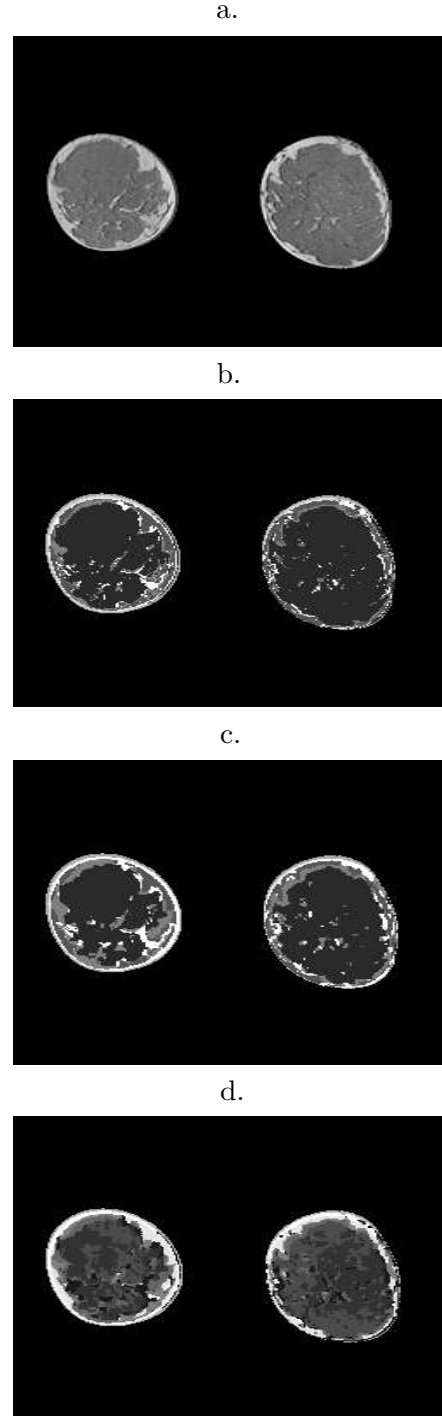


Figure 1: (a) A breast MR slice. (b) The segmentation result using 6 tissue classes and the k-means clustering algorithm. (c) The segmentation result of the new method, after 10 iterations using 6 tissue classes. (d) The segmentation result of the new method, after 10 iterations using 10 tissue classes.

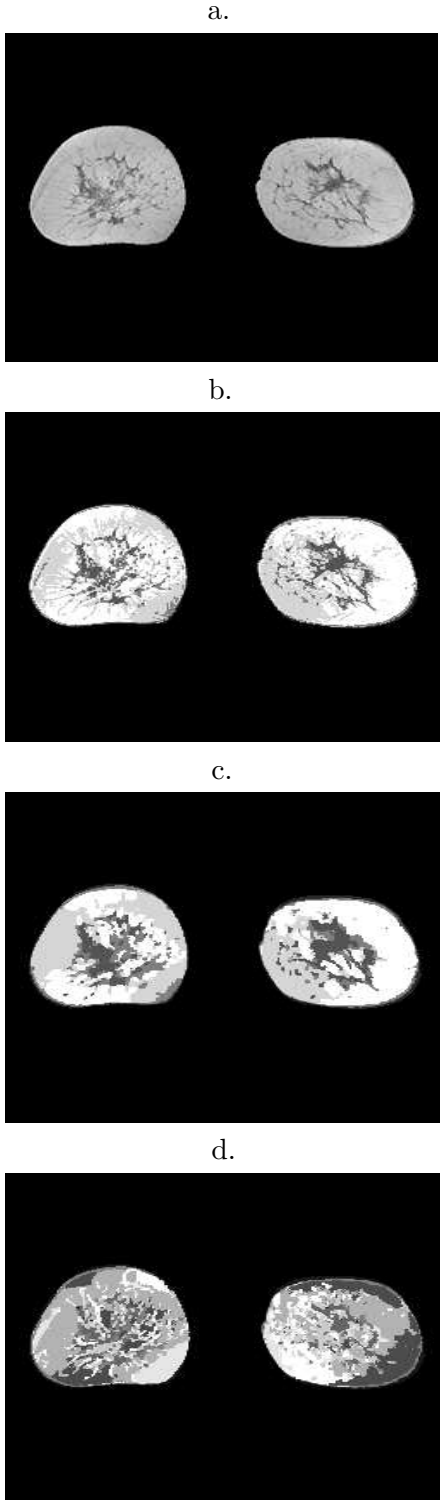


Figure 2: (a) A breast MR slice. (b) The segmentation result using 6 tissue classes and the k-means clustering algorithm. (c) The segmentation result of the new method, after 10 iterations using 6 tissue classes. d) The segmentation result of the new method, after 10 iterations using 10 tissue classes.

allowed to change and where the GHMRF parameters are updated using an EM method. Our approach will be used in 3-D registration and reconstruction of the breast. In the future, we plan to evaluate a larger set of features for each point in our segmentation. We also intend to incorporate the corresponding T_1 value, as we aim to create statistical models of T_1 for healthy, malignant and fat tissues.

References

- [1] Clarke L.P., Velthuizen R.P., Camacho M.A., Heine J.J., Vaidyanathan M., Hall L.O., Thatcher R.W., and Silbiger M.L. MRI segmentation: methods and applications. *MRI*, vol. 13(3), 1995, pp. 343–368.
- [2] Zhu S.C. and Yuille A. Region competition: Unifying snakes, region growing, and Bayes/MDL for multiband image segmentation. *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 18(9), 1996, pp. 884–900.
- [3] Zhang Y., Brady M., and Smtih S. Segmentation of Brain MR Images Through a Hidden Markov Random Field Model and the Expectation Maximization Algorithm. *IEEE Trans. Med. Imag.*, vol. 20(1), 2001, pp. 45–57.
- [4] Petroudi S. and Brady M. Classification of Mammographic Texture Patterns. In *Digital Mammography*. Medical Physics Publishing, 2004.
- [5] Fwu J.K. and Djuric P.M. Unsupervised vector image segmentation by a tree structure-ICM algorithm. *IEEE Trans. Med. Imag.*, vol. 15(6), 1996, pp. 871–880.
- [6] Brault P. and Mohammad-Djafari A. Bayesian Segmentation and Motion Estimation in Video Sequences using a Markov-Potts Model. In *Proceedings of WSEAS Conference on Applied Mathematics 2004*, 2004.
- [7] Styner M., Gerig G., Brechbhlher C., and Szekeley G. Parametric estimate of intensity inhomogeneities applied to MRI. *IEEE Trans. Med. Imag.*, vol. 19(3), 2000, pp. 153–165.