# GENETIC ALGORITHMS FOR FINITE MIXTURE MODEL BASED TISSUE CLASSIFICATION IN BRAIN MRI

Jussi Tohka[*,**], Evgeny Krestyannikov[**], Ivo Dinov[*,***], David Shattuck[*], Ulla Ruotsalainen[**] and Arthur W. Toga[*]

[*] Laboratory of Neuro Imaging, Department of Neurology, UCLA Medical School, Los Angeles, CA, USA
[**] Institute of Signal Processing, Tampere University of Technology, Tampere, Finland
[***] Department of Statistics, University of California Los Angeles, CA, USA

jussi.tohka@tut.fi

**Abstract: Finite mixture models (FMMs) are an indispensable tool for unsupervised classification in brain imaging. Fitting a FMM to the data leads to a complex optimization problem. This optimization problem is difficult to solve with standard local optimization methods (e.g. by the expectation maximization (EM) algorithm) if a good initialization is not available. In this paper, we propose a new global optimization algorithm for the FMM parameter estimation that is based on the real coded genetic algorithms. Our specific contributions are two-fold: 1) We propose to use blended crossover in order to reduce the premature convergence problem to its minimum. 2) We introduce a completely new permutation operator specifically meant for the FMM parameter estimation. We demonstrate the good behavior of our algorithm compared to the EM-algorithm and a standard real coded genetic algorithm with the tissue classification task within the magnetic resonance brain imaging. Phantom images as well as real three dimensional image data with pathology are considered. The tissue classification results by our method are shown to be consistently more reliable and accurate than with the competing parameter estimation methods.**

## Introduction

Finite mixture models (FMMs) are weighted sums of relatively simple parametric probability density functions (pdf)s called *component densities*. A component density models the probability of the data from a certain class in an unsupervised classification problem. Then, the corresponding weighting factor in the FMM - called the *mixing parameter* - models the prior probability of that class. A sound, statistically based method for unsupervised classification involves minimizing the discrepancy between observed data and the FMM with respect to the unknown parameter values. This involves solving a complex optimization problem. The initializations for standard local algorithms have to be well selected for the parameter estimates to be good. In Fig. 1, the effect of a small change in the initialization for the expectation maximization (EM) algorithm [1] is demonstrated in the case of the tissue classification of magnetic resonance (MR) images. If the initialization for the EM algorithm was generated relying on stereotaxic registration and a brain atlas [2], such a small change could be caused by pathology or failed stereotaxic registration. To avoid the initialization problem, we propose a new method based on real coded genetic algorithms (RCGAs) to solve the optimization problem globally. We use tissue classification in MRI to demonstrate desirable features of our algorithm. We have chosen this application because it is important and well known. However, we emphasize that our algorithm is meant as a general tool for FMM fitting problems in brain imaging and it is not limited to the tissue classification application.

Previous approaches to the global FMM optimization within medical imaging include [3, 4]. In [3] tissue quantification within MRI was considered by optimizing the FMM parameters using the tree annealing algorithm [5]. However, the time-complexity of tree-annealing grows very fast with the increased number of variables in the problem. In [4], RCGAs were considered for tissue classification within T1 weighted MR brain images. However, as the authors applied the flat crossover operator that is known to cause premature convergence [6], their approach suffered local optima problems similar to the EM algorithm. The approach has been generalized to account for the partial volume effect [7], however, this generalization does not remove the local optimum problem associated with the flat crossover.

In this paper, we introduce a new RCGA for the fitting of FMMs within medical imaging. Our contributions are two-fold. Firstly, we apply a blended crossover (BLX) to avoid the premature convergence [8]. Secondly, we propose a new permutation operator designed to reduce the size of the search space by using *a priori* knowledge about the optimization problem. We compare our algorithm to the EM-algorithm and a RCGA similar to that in [4] using simulated and real MRI data. Our algorithm is shown to be much less sensitive to its initialization than those methods.


ISSN: 1727-1983 

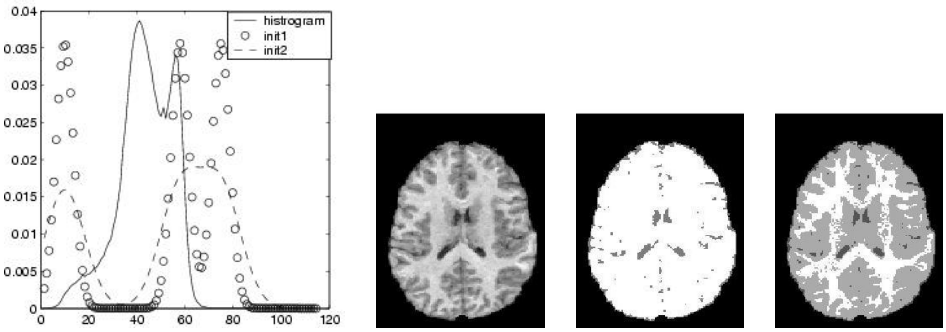


Figure 1: The local maxima problem with EM based parameter estimation. From left: two different initializations compared to the image histogram, a central axial slice of the skull stripped and non-intensity corrected [9] MR image, the tissue classification from the initialization 1, and the tissue classification from the initialization 2.

## Mixture models, maximum likelihood and classification

*General formulation*

Observed image intensities are denoted by $x_i \in \mathbb{R}, i = 1,\ldots,N$. All of these intensities come from one of the $K$ classes - here modeling intensities from different tissue types. Intensities drawn from the class $k$ follow the probability density function (pdf) $f_k(x,|\theta_k), k = 1,\ldots,K$. The parametric form of the pdf $f_k$ is known but the value of the parameter vector $\theta_k$ is unknown. The pdfs $f_k$ are called *component densities*. Each class has a prior probability $p_k \in [0,1]$ expressing the fraction of the intensity values following the density $f_k$. These *mixing parameters* satisfy $\sum_{k=1}^{K} p_k = 1$ and their values are initially unknown. We denote the set of all parameter values by $\Theta = \{p_i, \theta_i : i = 1,\ldots,K\}$. Combining the above facts, the pdf for whole data is

$$f(x|\Theta) = \sum_{k=1}^{K} p_k f_k(x|\theta_k). \quad (1)$$

The objective is to estimate the parameters $\Theta$ given the data $\{x_i : i = 1,\ldots,N\}$. Here, the estimation is based on the maximum likelihood (ML) principle. To find the ML estimate $\hat{\Theta}$, an optimization problem has to be solved:

$$\begin{aligned} \hat{\Theta} &= \arg\max l(\Theta) = \arg\max_{\Theta} \log \prod_{i=1}^{N} f(x_i|\Theta) \\ &= \arg\max_{\Theta} \sum_{i=1}^{N} \log f(x_i|\Theta). \end{aligned} \quad (2)$$

The log-likelihood $l(\Theta)$ will typically have several local maxima. Once we have the estimate $\hat{\Theta}$, the image voxels can be classified by the Bayes classifier based on their intensity values. That is, the class $\omega_i$ of the voxel $i$ is

$$\omega_i = \arg \max_{k \in \{1,\ldots,K\}} \hat{p}_k f_k(x_i|\hat{\theta}_k). \quad (3)$$

*Image models*

It is assumed that the brain has been extracted from the images before the parameter estimation. We also assume that the images have been corrected for the shading artifact. The images are assumed to be composed of two kinds of voxels: Those that contain only one type of tissue (pure voxels) and those that contain several types of tissues (partial volume (PV) voxels). Following the mixel model [10], it is assumed that the intensities of the pure voxels follow the normal density

$$f_k(x|\mu_k,\sigma_k^2) = [(2\pi)\sigma_k^2)]^{-1/2} \exp[-\frac{(x-\mu_k)^2}{2\sigma_k^2}] \quad (4)$$

where $\mu_k$ is the mean and $\sigma_k^2 > \varepsilon > 0$ is the variance of the class $k$. Constraining the variances to be larger than a small positive constant $\varepsilon$ is necessary because otherwise likelihood function in (2) could grow without a limit [11]. The pure voxel classes are white matter (WM), gray matter (GM), and cerebro spinal fluid (CSF).

The pdfs for PV voxels are constructed by marginalization [12]. We assume that there are no more than two tissue types present in a voxel. The pdfs of PV voxels are dependent on the parameters of the appropriate pure voxel classes. For the mixture of tissue types $j$ and $i$, the pdf of the PV class is [13]

$$\begin{aligned} f_k(x) &= \int_0^1 [(2\pi)(w^2\sigma_j^2 + (1-w)^2\sigma_i^2)]^{-1/2} \\ &\times \exp[-\frac{(x - w\mu_j + (1-w)\mu_i)^2}{2(w^2\sigma_j^2 + (1-w)^2\sigma_i^2)}]dw. \end{aligned} \quad (5)$$

These pdfs have to be evaluated using numerical integration. The PV classes are numbered with indices $K+1,\ldots,K+K'$. Hence, the FMM (1) can be re-written as

$$f(x|\Theta) = \sum_{k=1}^{K} p_k f_k(x|\theta_k) + \sum_{k=K+1}^{K+K'} p_k f_k(x|\theta_i,\theta_j), \quad (6)$$

where $\sum_{k=1}^{K+K'} p_k = 1$, $i,j \leq K$ and $\Theta = \{p_k, \mu_i, \sigma_i^2 : k = 1,\ldots,K+K', i = 1,\ldots,K\}$. The PV classes are CSF/GM and WM/GM. The class background/CSF is not included due to its negligible effect to the FMM in most cases.

PV classes have a large influence on the pdf of the whole image. However, it is more desirable that all voxels


ISSN: 1727-1983 

are assigned to pure classes, because many further automatic image analysis procedures expect such input (e.g. [14]). Therefore, voxels initially classified to some PV class are re-classified to a pure voxel class as described in [13].

## Genetic algorithm

Genetic algorithms (GAs) mimic the genetic processes of biological organisms. Here, a GA is applied for maximizing the likelihood function with respect to the parameter vector $\Theta$. The basic structure of our GA is presented in Algorithm 1.

The populations consist of the parameter vectors for the mixture models to be fitted to the data. The parameter vectors are represented by a vector of real numbers, these kinds of GAs are said to be real coded. A survey about RCGAs can be found in [6]. The fitness of an individual (a parameter vector) is the likelihood of the data under the FMM model (6). The populations are initialized by a set of parameter vectors drawn randomly from the set of admissible parameter vectors. The applied population size is 100. The tournament selection with the tournament size of 2 is applied. The algorithm is elitistic, the individual with the best fitness score (the maximal likelihood) always survives to the next generation. For recombination, the BLX-0.5 operator is applied (subsection *Blended crossover*) with the crossover rate equal to 1. After recombination, a novel permutation operator is applied (subsection *Permutation operator*). To keep the algorithm simple and efficient, no mutation is applied. The algorithm is terminated when the difference in the likelihood score of the best individual and the mean likelihood score of the population drops below a certain threshold.

### *Blended crossover*

During the recombination, parameter vectors are combined, two at a time, to produce a new offspring via a crossover operator. We apply the BLX-$\alpha$ operator with $\alpha = 0.5$ [8]. If $\Theta^1 = [a_1, \ldots, a_M] \in \mathbb{R}^M$ and $\Theta^2 = [b_1, \ldots, b_M] \in \mathbb{R}^M$ are the parents then the offspring is $\Theta^{new} = [c_1, \ldots, c_M]$, where

$$c_i = r_i a_i + (1 - r_i) b_i. \tag{7}$$

Scalars $r_i, i = 1, \ldots, M$, are random numbers drawn uniformly from the interval $[-\alpha, 1 + \alpha]$. Each $c_i$ is a real scalar required to lie in the pre-specified interval $[c_i^{min} c_i^{max}]$. If after the recombination some $c_i \notin [c_i^{min} c_i^{max}]$, it is simply scaled to the nearest end point of that interval. The constraints of the type $\sum_{j=1}^{K+K'} c_j = 1$ are handled by normalizing the values $c_j, j = 1, \ldots, K + K'$ after the recombination by dividing them by the sum $\sum_{j=1}^{K+K'} c_j$. (Here, we assumed that the mixing parameters are the $K + K'$ first variates.)

When considering BLX-$\alpha$ operators, $\alpha = 0.5$ is the optimal choice in that it balances the relationship between exploration (finding completely new solutions) and exploitation (improving already found solutions) [8, 6]. The flat crossover, used in [4], is equal to the BLX-0 crossover that overdoes the exploitation.

**Algorithm 1** Genetic Algorithm

```
t ← 0, initialize a population P(t) of mixture models
evaluate P(t) by computing the likelihood of each individual in it
while NOT termination condition do
  t ← t + 1
  select P(t) from P(t − 1)
  recombine P(t)
  permute Θ ∈ P(t) when necessary
  evaluate P(t)
end while
```

### *Permutation operator*

We ensure that $\mu_1 \leq \mu_2 \leq \ldots \leq \mu_K$ after the recombination. This is done by performing a certain permutation of variables:

For all index pairs $i, j$ where $i < j \leq K$ do the following: If $\mu_j < \mu_i$ then swap $\mu_i$ and $\mu_j$, $\sigma_i^2$ and $\sigma_j^2$, $p_i$ and $p_j$.

This is the permutation operator. The operator increases the efficiency of the GA based mixture modeling due to the reduction of the size of the search space by the operator. As can be seen from Eqs. (1) and (6), if a permutation is applied to the classes of the FMM, the likelihood of the data remains the same whereas the interpretation of the FMM changes dramatically. This kind of identifiability problem is encountered because the intensities of the pure voxel classes are all modeled with the same parametric distribution. Therefore, without extra information, the FMM does not include a possibility to interpret the classes correctly. For example in T1-weighted MR data, WM intensities is greater than GM intensities on average. This kind of information can be used to interpret the classes in a FMM and to reduce the search space size via the permutation operator.

### *Fast implementation*

The likelihood function in Eq. (2) needs to evaluated numerous times during our GA. Hence, a notable speed up can be achieved by using a well-known connection between maximum-likelihood and the Kullback-Leibler divergence [15, 16]. The EM algorithm can be sped up using a similar strategy [17].

The Kullback-Leibler divergence between the true pdf $g(\cdot)$ and the parametric density $f(\cdot|\Theta)$ is

$$D(g(\cdot), f(\cdot|\Theta)) = \int g(x) \log \frac{g(x)}{f(x|\Theta)} dx. \tag{8}$$

Minimizing this divergence with respect to $\Theta$ is equivalent to the maximizing the likelihood (2) when $N$ tends to


ISSN: 1727-1983 

Table 1: The misclassification rates with 50 random initializations and BrainWeb data. Left: The minimum, mean, median, and maximum misclassification rates, resulting from random initializations, with T1-weighted image with 5 % of noise are shown (the Bayes error is 6.0 %). Right: The mean misclassification rates from random initializations with other types of data.

| | T1 with 5 % noise | | | |
|---|---|---|---|---|
| | min | mean | med | max |
| BLX-GA | 6.1 | 6.2 | 6.2 | 6.5 |
| FLAT-GA | 6.0 | 6.4 | 6.3 | 7.7 |
| EM | 6.8 | 29.2 | 8.5 | 80.7 |

| pulse seq. | T1 | | T2 | | | PD | | |
|---|---|---|---|---|---|---|---|---|
| noise % | 3 | 7 | 3 | 5 | 7 | 3 | 5 | 7 |
| BLX-GA | 4.1 | 9.6 | 8.8 | 15.2 | 20.8 | 16.1 | 28.3 | 36.1 |
| FLAT-GA | 4.2 | 9.8 | 12.5 | 16.2 | 20.4 | 19.0 | 29.8 | 37.2 |
| EM | 18.9 | 23.9 | 41.3 | 42.5 | 42.5 | 60.9 | 63.6 | 60.1 |

infinity [15]. By replacing the integral in (8) by its trapezoidal approximation and $g(\cdot)$ by its Parzen estimate $\hat{g}(\cdot)$ [11], we get a faster way to obtain the ML estimate:

$$\hat{\Theta} = \arg\min_{\Theta} \hat{D}(\hat{g}(\cdot), f(\cdot|\Theta)); \quad (9)$$

$$\hat{D}(\hat{g}(\cdot), f(\cdot|\Theta)) = \sum_{i=1}^{M-1} (z_{i+1} - z_i)\hat{g}(z_i) \log \frac{\hat{g}(z_i)}{f(z_i|\Theta)}. \quad (10)$$

In this paper, we select $M = 100$ and $z_j = \min_i x_i + \frac{(j-0.5)}{M}(\max_i x_i - \min_i x_i)$. The Parzen estimates $\hat{g}(z_i)$ are computed using Gaussian window functions with the standard deviation of $(\max_i x_i - \min_i x_i)/M$. This way a speed up of the factor $N/M$ can be obtained. For a typical MR image this factor can be 10000.

**Experiments and results**

In this section, we compare the performance of our algorithm (BLX-GA) to those of the EM algorithm and a GA similar to [4] (FLAT-GA). With the EM-algorithm, the PV classes were ignored. The FLAT-GA algorithm was otherwise exactly the same as our BLX-GA algorithm except that it used the flat crossover and it did not contain the permutation operator. The classes were ordered after the parameter estimation with FLAT-GA and EM, so that the improvements by BLX-GA were not due to the identifiability problem.

In the first experiment, the algorithms were compared using the simulated brain MRI images from the Brainweb database by the Montreal Neurological Institute `http://www.bic.mni.mcgill.ca/brainweb` [18, 19]. We applied the images with no intensity non-uniformity, with the resolution of 1 mm × 1mm × 1mm and with the image size of 181 x 217 x 181. The images studied were simulated with the T1 and T2 weighted as well as proton density (PD) pulse sequences. The brain volume was extracted based on the ground-truth. The quantitative results were computed using the misclassification rate as the criterion of success. Each algorithm was run 50 times with a random initialization with each image and the misclassification rates resulting each of these runs were calculated. The results are presented in Table 1. As can be seen, both GAs were clearly more reliable than the EM-algorithm, which typically achieved the average performance of the GAs with only the best of the 50 random initializations. The EM-algorithm worked well only when the initialization was well chosen, usually the results by EM were completely useless. BLX-GA was clearly more reliable than FLAT-GA as can be observed by the worst performance with a T1-weighted scan (see the 'max' column in Table 1). Also, the mean performance of BLX-GA was consistently better than that of FLAT-GA, see the right panel in Table 1. This difference was especially clear with the T2-weighted and proton density images with the lowest noise level.

In the second experiment, the repeatability of BLX-GA using 3-D T1-weighted MR brain scans of healthy and diseased (Alzheimer's disease (AD), schizophrenia and childhood schizophrenia) subjects were studied (Fig. ). Since the algorithm involves random components, the repeatability of it is an important question. The images were corrected from the intensity non-uniformity using the N3-method [9]. The brain was extracted using the Brain Surface Extractor [20] (healthy,schizophrenia), using the Brain Extraction Tool (childhood schizophrenia) [21] or manually (AD). We estimated the FMM parameters of the image data starting from a random initialization and tissue classified the image. This was repeated 50 times. Based on the 50 tissue classified images, the average classification was constructed by deciding the label of each voxel by a majority vote, see Fig. . The percentages of brain voxels classified differently compared to the average classification were 2.8 % (healthy subject), 5.3 % (AD subject), 1.2 % (schizophrenic subject) and 2.9 % (schizophrenic child). These percentages can be considered to be very low and they indicate good reproducibility of the algorithm. With FLAT-GA, the percentages were higher in most cases 6.0 % (normal), 3.4 % (AD), 3.4 % (schizophrenic) and 3.8 % (schizophrenic child).

In addition to the reproducibility study, we computed the resulting Kullback-Leibler divergences (10) for both BLX-GA and FLAT-GA. These may be considered as quantitative indicators of the success of the algorithms in the optimization problem. The mean values of these are given in Table 2. The Kullback-Leibler divergences produced by FLAT-GA were usually about ten times larger than with BLX-GA. This indicates that the results by BLX-GA were more correct than with FLAT-GA. Particularly with the AD subject, where the reproducibility of FLAT-GA was better, BLX-GA achieved more accurate parameter estimates as can be seen from Table 2.


ISSN: 1727-1983 

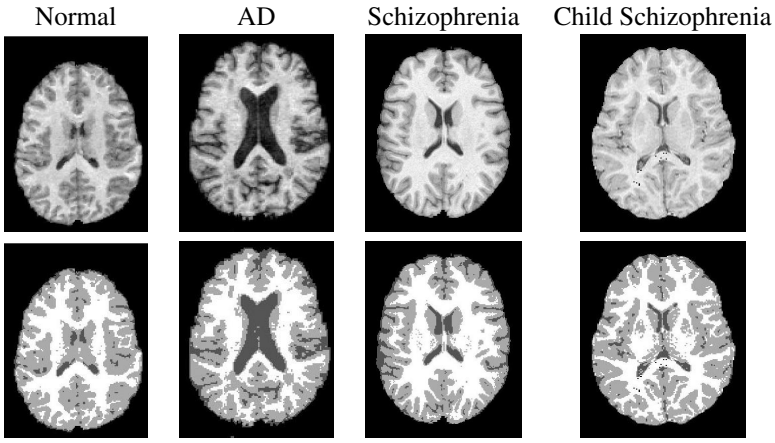


Figure 2: Top row: Axial cross-sections of the images applied in the reproducibility test. Bottom row: Axial cross-sections of the average voxel classification that were decided by the majority vote from 50 random initializations for BLX-GA parameter estimation. The 'normal' image is the same as in our example in Fig. 1. The images are in the native space.

Table 2: The average Kullback-Leibler divergences from 50 random initializations. Lower value means better result. The abbreviation 'SZ' stands for schizophrenia.

| | Normal | AD | SZ | Child SZ |
|---|---|---|---|---|
| BLX-GA | 0.0057 | 0.0077 | 0.0015 | 0.0063 |
| FLAT-GA | 0.0879 | 0.0252 | 0.0512 | 0.0774 |

## Discussion

We have presented a global algorithm for the FMM parameter estimation. The algorithm is based on the optimization of the ML criterion using a novel GA. The algorithm is not sensitive to its initialization which makes it particularly useful in cases where a well-principled initialization for a local algorithm, such as the EM algorithm, is not available. Moreover, the parameters for arbitrary FMMs can be estimated, since the algorithm does not require any assumptions about the parametric models for the component densities. This is important in many brain imaging applications where e.g. the influence of the partial volume effect is large. The running time of the algorithm varied from tens of seconds to five minutes for the tissue classification of a 3-D MR image with a 3.0 Ghz PC.

Our GA contained two novel features compared to the previous GA-based approaches for FMM optimization. 1) We applied the blended crossover which reduces the problem of the premature convergence to its minimum. 2) A completely new permutation operator was introduced. These two contributions were demonstrated to improve voxel classification results with 3-D MR brain images.

*Acknowledgments*

This work was supported by the NIH/NIMH research grant R01 MH071940 and the NIH/NCRR resource grant P41 RR013642. The work of J.T. was supported by the Academy of Finland under grants 204782 and 104834. E.K. was supported by the Tampere Graduate School of Information Science and Technology (TISE) and by the Academy of Finland under the grant 104834. The work of I.D. was supported by the UCLA grant OID IIP (IIP0318) and the NSF grant CCLI-EMD (044299). Additional support was provided by the National Institutes of Health through the NIH Roadmap for Medical Research, Grant U54 RR021813 entitled Center for Computational Biology (CCB). Information on the National Centers for Biomedical Computing can be obtained from `http://nihroadmap.nih.gov/bioinformatics`.

## References

[1] A.P. DEMPSTER, N.M. LAIRD, and D.B. RUBIN. Maximum likelihood from incomplete data via the EM algorithm. *J R Stat Soc Ser B*, 39(1):1 – 39, 1977.

[2] J. MAZZIOTTA, A. TOGA, A. EVANS, and ET AL. A four-dimensional probabilistic atlas of the human brain. *J Am Med Inform Assoc*, 8:401 – 430, 2001.

[3] P. SANTAGO and H. D. GAGE. Quantification of MR brain images by mixture density and partial volume modeling. *IEEE Trans Med Imag*, 12(3):566 – 574, September 1993.

[4] P. SCHROETER, J.-M. VESIN, T. LANGENBERGER, and R. MEULI. Robust parameter estimation of intensity distribution for brain magenetic resonance images. *IEEE Trans Med Imag*, 17(2):172 – 186, 1998.


ISSN: 1727-1983 

[5] G.L. BILBRO and W. E. SNYDER. Optimization of functions with many minima. *IEEE Trans Syst Man Cybern*, 21:840 – 849, July/August 1991.

[6] F. HERRERA, M. LOZANO, and J. L. VERDEGAY. Tackling real-coded genetic algorithms: Operators and tools for behaviorial analysis. *Artif Intell Rev*, 12(4):265 – 319, August 1998.

[7] M. BACH CUADRA, B. PLATEL, E. SOLANAS, T. BUTZ, and J. THIRAN. Validation of tissue modelization and classification techniques in T1-weighted MR brain images. In *Proc. of MICCAI 2002, Lecture Notes In Computer Science 2488*, pages 290 – 297, 2002.

[8] L. J. ESHELMAN and J. D. SCHAFFER. Real-coded genetic algorithms and interval schemata. In L.D. Whitley, editor, *Foundations of Genetic Algorithms 2*, pages 185 – 202. Morgan Kaufmann Publishers, 1993.

[9] J. G. SLED, A. P. ZIJDENBOS, and A. C. EVANS. A non-parametric method for automatic correction of intensity non-uniformity in MRI data. *IEEE Trans Med Imag*, 17(1):87 – 97, 1998.

[10] H. S. CHOI, D. R. HAYNOR, and Y. KIM. Partial volume tissue classification of multichannel magnetic resonance images - a mixel model. *IEEE Trans Med Imag*, 10(3):395 – 407, 1991.

[11] R. O. DUDA, P. E. HART, and D. G. STORK. *Pattern Classification*. Wiley-Interscience, New York, 2nd edition, 2001.

[12] P. SANTAGO and H. D. GAGE. Statistical models of partial volume effect. *IEEE Trans Image Process*, 4(11):1531 – 40, November 1995.

[13] J. TOHKA, A. ZIJDENBOS, and A.C. EVANS. Fast and robust parameter estimation for statistical partial volume models in brain MRI. *NeuroImage*, 23(1):84 – 97, 2004.

[14] D. MACDONALD, N. KABANI, D. AVIS, and A.C EVANS. Automated 3-D extraction of inner and outer surfaces of cerebral cortex from MRI. *Neuroimage*, 12(3):340 – 356, 2000.

[15] G. GRAY. Bias in mispecified mixtures. *Biometrics*, 50(2):457 – 470, 1994.

[16] R.I JENNRICH. Asymtotic properties of non-linear least squares estimators. *Ann Math Stat*, 40(2):633 – 643, 1969.

[17] Y. WANG, T. ADAH, J. XUAN, and Z. SZABO. Magnetic resonance image analysis by information theoretic criteria and stochastic site models. *IEEE Trans Inform Tech Biomed*, 5(2):150 – 158, June 2001.

[18] L. COLLINS, A.P. ZIJDENBOS, V. KOLLOKIAN, J.G. SLED, N.J. KABANI, C.J. HOLMES, and A.C. EVANS. Design and construction of a realistic digital brain phantom. *IEEE Trans Med Imag*, 17(3):463 – 468, 1998.

[19] R.K.-S. KWAN, A.C. EVANS, and G.B. PIKE. MRI simulation-based evaluation of image-processing and classification methods. *IEEE Trans Med Imag*, 18(11):1085–97, 1999.

[20] D. W. SHATTUCK, S. R. SANDOR-LEAHY, K.A. SCHAPER, D.A. ROTTENBERG, and R.M. LEAHY. Magnetic resonance image tissue classification using a partial volume model. *Neuroimage*, 13(5):856 – 876, 2001.

[21] S. SMITH. Fast robust automated brain extraction. *Human Brain Mapping*, 17:143 – 155, 2002.


ISSN: 1727-1983 