The Finite Mixture Model (FMM) based approaches have been applied in Magnetic Resonance Imaging (MRI) to extract information about human anatomy. The idea is to model feature vector of a tissue using some known distribution (such as Gaussian, known as GMM). The performance of FMM deteriorates with increase in noise within data which may occur due to environment, patient movement, technician expertise level etc. The Spatially Variant Finite Mixture Model (SVFMM) is used as robust alternative to this. The student's t distribution has been explored, in place of Gaussian distribution, in context of FMM. Within the framework of SVFMM, Student's t distribution is used in this work to overcome the noise impact present in the data. The parameters of SVFMM are learned by 'sampling-resampling' based Bayesian Learning. The novelty of work lies in parameter estimation of mixture model in comparison to Frequency based estimations. This paper compares different mixture models, distribution functions and parameter estimation procedures. The column based sampling is used to train the model. The misclassification rate (MCR) is used as quantitative measure for performance evolution. The results of proposed method seem encouraging.
using Bayesian sampling-resampling approach. The performance of algorithm is compared with other approaches based on misclassification rate (MCR).

The paper is organized as follows: section 2 provides framework of SVFMM and the Bayesian sampling resampling for the parameter estimation. Section 3 presents the proposed algorithm for parameter estimation. Section 4 presents experimental results and comparison with other approaches based on MCR and computation time. Section 5 concludes the manuscript.

2. SVFMM AND BAYESIAN APPROACH

The finite mixture model (FMM) approach divides a density function into sub-density functions. The parameter estimation procedure assumes these sub density functions having classical parametric form such as Gaussian, Poisson etc. The algorithm estimates parameters of these functions, as well as the contribution of each class in the mixture. In general, all the component densities are assumed to follow same density function.

2.1 Spatially Variant Finite Mixture Model (SVFMM)

Let \( p_j^i \) denote the probability of the \( i^{th} \) pixel belonging to the \( j^{th} \) class; \( 0 \leq p_j^i \leq 1 \) and \( \sum_{j=1}^{L} p_j^i = 1 \forall \ i \). The SVFMM defines the density function of the observation \( x_i \) as [6]

\[
f_j(x_i | p^1, \ldots, p^n, \theta^1, \ldots, \theta^J) = \prod_{j=1}^{L} p_j^i f_j(x_i | \theta^j)
\]

If the observation \( x_i \) are modelled as statistically independent, the joint conditional density of observation can be formed as

\[
f(x_i | p^1, \ldots, p^n, \theta^1, \ldots, \theta^J) = \prod_{j=1}^{L} \sum_{i=1}^{L} p_j^i f_j(x_i | \theta^j)
\]

The mixing proportion \( p_j^i \) called as contextual mixing proportions in [21]. The SVFMM incorporates FMM as a special case and all label parameters converge to 0 or 1 which make labelling unambiguous. The numbers of parameters to be estimated in SVFMM are \( N \times L \) label probabilities and \( L \times N \) component density parameter vector. In SVFMM, each data element retains its pdf over the components which is used in the next iteration as a prior probability, where as in FMM, these are summed up and normalized to obtain single prior probability per component. Hence, labelling becomes spatially variant in case of SVFMM [9].

2.2 Bayesian Estimation

The purpose of Bayesian Inference is to make inferences from data using probability model for quantities observed and quantities about which we wish to learn [23]. The main concerning issue is how to model the priori information. The most convenient way is to assume that parameters have uniform distribution as a prior distribution. This makes all the value in parameter space to be equally likely; if nothing is known in specific. This sampling- resampling based method overcomes the drawback of conventional Bayesian approach which involves numerical integration or other approximation techniques to estimate posterior distribution. The Figure 1(a) shows the prior sample distribution of the mean as Uniform Distribution, where \( N \) is the maximum value that mean can have. The Figure 1(b) shows its transition into posterior distribution. The posterior distribution can also be assumed to have a particular distribution. This assumption helps in defining the property of posterior inference on mathematical framework. The Initialization of mean samples is done by Histogram Method during Simulation. By Histogram method, range of class mean was identified. This will make a more strong prior knowledge about the mean values.

![Figure 1: (a) Prior Inference, (b) Posterior Inference](image)

3. PROPOSED ALGORITHM

The following steps are to be performed:
1. Draw \( Z \) random samples each from all the prior distributions of the \( K_{\text{max}} \) means. Let us call the obtained samples as \( \{\mu_{11}, \mu_{22}, \ldots, \mu_{12}, \ldots, \mu_{K_{\text{max}}}, \mu_{K_{\text{max}}}, \ldots, \mu_{K_{\text{max}}} \} \).
2. When a pixel value \( x \) is observed, compute the sum of likelihoods for each Mean distribution, given that observation:

\[
L_j = \sum_{i=1}^{Z} l(\mu_{ji} | x), \quad j = 1, 2, \ldots, K_{\text{max}}
\]

The likelihood function is defined on the basis of the probability density function used.
3. The next step is to determine which cluster the pixel observation belongs to. The observation would belong to the cluster having the highest class probability.

\[
p_j = \frac{L_j}{\sum_{k=1}^{K_{\text{max}}} L_k}
\]

The prior distribution of the Mean of this cluster is updates to obtain a posterior distribution using step 4. The distributions of the means of the other clusters are left unchanged.
4. The steps to update a prior distribution to posterior one are:
   a. If the \( j^{th} \) distribution is to be updated, compute weights \( q_i \) for each samples \( \mu_{ji} \) of the prior distribution as follows:

\[
q_i = \frac{l(\mu_{ji} | x)}{L_j}, \quad i = 1, \ldots, Z
\]
   b. \( \{\mu_{j1}, \mu_{j2}, \ldots, \mu_{jZ} \} \) are then resampled using the weighted bootstrap method with weights \( \{q_1, q_2, \ldots, q_j \} \) to obtain samples for the posterior distribution of \( \mu_j \) which are \( \{\mu^*_1, \mu^*_2, \ldots, \mu^*_Z \} \).
5. When the next pixel observation is made, the posterior samples \( \{\mu^*_1, \mu^*_2, \ldots, \mu^*_Z \} \) become the prior samples of the \( j^{th} \) cluster mean. Steps 2 to 5 are repeated for every observation which is under consideration.
In comparison to algorithm presented in [17], [18], the proposed algorithm provides the labelling of each pixel according to its likelihood value. A comparison has been presented in the next section.

4. EXPERIMENTAL RESULTS

The work is carried out on 3D MRI phantom brain images. The data and its ground truth are downloaded from McConnell Brain Imaging Centre, Montreal Neurological Institute, McGill University, Canada [27]. The dimensions data used are 181x217x181. The specification of the data is as follows: modality: T1, protocol: ICBM, Noise: 3%, RF: 40%. Throughout the experiments, the number of classes kept fixed to four: (i) CSF, (ii) GM, (iii) WM and (iv) Background. To train the mixture model, few samples are used for estimation of parameters. The samples are obtained using column sampling as used in [30]. All the experiments are performed in MATLAB (R2008b version) on Intel core(TM)2 Duo CPU E8400 @3.00GHz, 1.99GB RAM. The classification accuracy is measured by misclassification rate (MCR) as follows:

$$\text{MCR} = \frac{\text{Number of misclassified pixels}}{\text{Total number of pixels}} \times 100$$

### 4.2 Sampling Percentage

This section describes about one of the most crucial parameter of any sampling method, i.e., Number of samples requires. The small number of samples results in poor estimate of the parameters and may lead to high error whereas large number of samples makes it more time consuming method and small improvement in the results. So, every application requires sufficient number of samples that can give better estimate within some time constraints. The Table 2 presents the misclassification rate measure using different sampling percentage with ML Estimation technique. The samples taken from whole volume using Column Sampling method and Random Sampling method (represented as Col and Rad in the table respectively). In most of the cases, Column Sampling gives better results or minimum MCR as compared to Random Sampling. Similarly, Table 3 presents misclassification rate using Bayesian Learning method. The number of prior samples, N, is adjusted experimentally. Figure 1 and Figure 2 show misclassification rate of slice 110 after segmenting whole volume data.

### 4.3 Time Comparison

Table 4 show the computational time of all the methods when different sampling percentage is used. All the figures are in seconds. These tables reveal that ML estimation is time consuming as sampling percentage increases whereas Bayesian Learning estimation takes less time in estimating the parameters.

---

**Table 1: Misclassification Rate of volume data when 1% and 3% samples taken from each direction and estimated using ML estimation. The figures in bold show minimum MCR.**

<table>
<thead>
<tr>
<th>Direction</th>
<th>1% samples</th>
<th>3% samples</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>FMM_N</td>
<td>FMM_St</td>
</tr>
</tbody>
</table>

**Table 2: Misclassification Rate of the volume data at different sampling percentages.**

<table>
<thead>
<tr>
<th>Sampling percentage</th>
<th>Maximum Likelihood Estimation</th>
<th>Bayesian Learning Estimation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>FMM_N</td>
<td>FMM_St</td>
</tr>
<tr>
<td></td>
<td>Col</td>
<td>Rad</td>
</tr>
<tr>
<td>5</td>
<td>4.7983</td>
<td>4.8093</td>
</tr>
</tbody>
</table>

**Table 3: Misclassification Rate of the volume data at different sampling percentages.**

<table>
<thead>
<tr>
<th>Sampling percentage</th>
<th>Bayesian Learning Estimation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>FMM_N</td>
</tr>
<tr>
<td></td>
<td>Col</td>
</tr>
<tr>
<td>1</td>
<td>5.0935</td>
</tr>
<tr>
<td>2</td>
<td>5.0793</td>
</tr>
<tr>
<td>3</td>
<td>5.7916</td>
</tr>
<tr>
<td>5</td>
<td>4.6562</td>
</tr>
<tr>
<td>7</td>
<td>6.0401</td>
</tr>
<tr>
<td>10</td>
<td>5.6276</td>
</tr>
</tbody>
</table>

---
CONCLUSION

The work discusses platform for segmentation of 3D volume images using the Bayesian estimation of SVFMM using Student’s t distribution function. The results of proposed method are incrementally better than existing one. The work compared FMM [17], [18] and SVFMM of Bayesian estimation. The comparison also includes effect of change of distribution from Gaussian to Student’s t distribution. A suitable feature vector can be used based on prior information of boundary, shape structure information to improve segmentation accuracy. The real time implementation of this requires sampling percentage to be fixed and its computational time in that environment. The simulations are done on normal human brain MRI phantom image so it can extend to real data and tumor detection also.

REFERENCES


