MR Image Segmentation and Data Fusion
Using a Statistical Approach

J. Kaufhold       M. Schneider       W. C. Karl       A. Willsky
BME Dept.          EECS Dept.          ECE/BME Dept.      EECS Dept.
Boston University  MIT                 Boston University  MIT
Boston, MA 02215   Cambridge, MA 02139 Boston, MA 02215 Cambridge, MA 02139

ABSTRACT
Magnetic resonance imaging (MRI) has become a widely used research and clinical tool in the study of the human brain. The ability to accurately segment the MRI data set into homogeneous regions such as gray matter, white matter, and cerebro spinal fluid aids in morphological quantitation of brain features. The large amount of data associated with typical MRI brain scans makes completely manual segmentation prohibitive on a large scale. In this paper we develop an estimation-theoretic interpretation of the segmentation problem which leads to a computationally efficient, statistically-based recursive technique for its solution. Being statistically based, the method also provides associated measures of uncertainty of the resulting estimates, which are useful both for evaluation of the estimates as well as their combination with other sources of information.

1. INTRODUCTION
Magnetic resonance imaging (MRI) has become a widely used research and clinical tool in the study of the human brain. The morphology of the various structures of the brain which can be obtained from such imagery correlate with various developmental differences, disease states, and injuries [1]. Accurate and reproducible quantification depends on accurate segmentation of the MRI brain data set into homogeneous regions such as gray matter, white matter, and cerebro spinal fluid (CSF). These regions and their boundaries have typically been delineated by specially trained technicians at great expense in time, money, and effort [4].

One popular approach to such segmentation problems is to formulate the problem in a variational framework [6]. In these approaches to segmentation, an energy functional captures desired properties of the resulting segmentation. The energy functional can enforce smoothness within homogeneous regions, preserve boundaries between homogeneous regions, and constrain edge properties. Unfortunately, finding the minimum of such a functional leads to large and computationally taxing optimization problems. Further, while these deterministic methods provide a rational approach to obtaining estimates of the underlying piecewise smooth field, they do not provide information about the uncertainty in these estimates, which are useful both for direct evaluation of the estimates as well as allowing these estimates to be rationally combined with other sources of information.

In this paper we present a statistically-based recursive approach to such segmentation problems which is computationally efficient and provides measures of uncertainty of the resulting estimates. Our approach is based on an interpretation of the segmentation problem as an equivalent recursive estimation problem.

2. VARIATIONAL SEGMENTATION METHOD
We start from a recently proposed variational formulation [6] of the segmentation problem based on a continuous edge process. This formulation obtains a segmentation as the solution of:

$$\arg \min_{f, s} \int_{\Omega} \mu (g-f)^2 + \lambda (1-s)^2 |\nabla f|^2 + \nu \left( \rho |\nabla s|^2 + \frac{1}{\rho} s^2 \right)$$

where $g$ is the observed MR imagery, $f$ is a smooth approximation to $g$ and the continuous edge process, $s$ is the estimate of the boundary locations and takes on values between 0 and 1 (edges exist where $s$ is approximately 1, and where $s$ is near zero, regions are homogeneous). Scalar weights, $\mu$, $\lambda$, $\rho$, and $\nu$ determine the properties of the overall segmentation (e.g. measured smoothness and edginess).
The minimization of (1) is equivalent to the joint minimization with respect to \( f \) and \( s \) of the following pair,

\[
E_s(f) = \int_{\Omega} \mu(g-f)^2 + \lambda(1-s)^2 |\nabla f|^2\, d\Omega \quad (2)
\]

\[
E_f(s) = \int_{\Omega} \left( \lambda |\nabla f|^2 + \nu \right) \left[ \frac{\lambda |\nabla f|^2}{\lambda |\nabla f|^2 + \nu} - s \right]^2 + \nu s^2 |\nabla s|^2\, d\Omega \quad (3)
\]

Discrete versions of \( E_s(f) \) and \( E_f(s) \) given in (2) and (3), respectively, may be written as:

\[
E_s(f) = \sum_{j=1}^n \|g(j) - f(j)\|_{\mu_I}^2 + \|D_r f(j)\|_{V(j)}^2 + \|f(j) - f(j-1)\|_{V(j)}^2 \quad (4)
\]

\[
E_f(s) = \sum_{j=1}^n \|h(j) - s(j)\|_{\nu_I}^2 + \|D_r s(j)\|_{\nu_I}^2 + \|s(j) - s(j-1)\|_{\nu_I}^2 \quad (5)
\]

Where \( D_r \) is a first difference operator over rows and the index \( j \) in the discretization above is the column number, so that \( f(j) \) corresponds to the samples of \( f \) in the \( j \)th column of the image.

The weighting matrices \( V(j) \) are diagonal and capture the spatially varying, edge-dependent weights on \( \lambda(1-s)^2 \) in (2). The weighting matrices \( W(j) \) are diagonal and capture the spatially varying weights, \( \lambda |\nabla f|^2 + \frac{\nu}{\rho} \) in (3). In practice, the joint minimum of (4) and (5) are solved for in a coordinate descent framework. First, the \( f \) corresponding to the minimum energy of (4) is found with \( s \) fixed, then the \( s \) corresponding to the minimum energy of (5) is found with \( f \) fixed. This process is iterated to some practical convergence.

Each of the minimizations, (4) and (5) is a large, computationally intensive, optimization problem, which must be performed repeatedly to obtain the segmentation. Further, only the estimates of the image field \( \hat{f} \) and edge process \( \hat{s} \) are provided by this process, with no measure of the reliability of those estimates.

3. STOCHASTIC INTERPRETATION

To avoid these difficulties we interpret each of the minimization problems corresponding to (4) and (5) in an estimation theoretic context. Such an interpretation of this segmentation problem allows a variety of statistically-motivated approaches to be taken [7].

In particular we interpret each minimization problem as an equivalent dynamic estimation problem. Based on this interpretation we are able to formulate an efficient, statistically-based recursive solution, which provides not only estimates of the fields themselves, but also corresponding measures of uncertainty. For simplicity we focus here on the problem of processing two-dimensional MRI slices – i.e. images, but this method can be generalized to arbitrarily high dimensionality.

To proceed, consider equation (4). It is straightforward to show [2] that the value \( \tilde{f}(j) \) for each \( j \) that minimizes the energy in (4) is the same as the maximum-likelihood estimate of each \( f(j) \) given the following set of constraints and observations:

\[
\begin{align*}
\textbf{f}(j) &= \textbf{f}(j-1) + \textbf{q}(j-1), \quad \textbf{q}(j) \sim \mathcal{N}(0, V^{-1}(j)) \\
\textbf{y}(j) &= C(j) \textbf{f}(j) + \textbf{r}(j), \quad \textbf{r}(j) \sim \mathcal{N}(0, R(j))
\end{align*}
\]

for \( j = 1, \ldots, n \), where \( \textbf{x} \sim \mathcal{N}(\textbf{m}, \textbf{P}) \) denotes a Gaussian random vector with mean \( \textbf{m} \) and covariance \( \textbf{P} \), and the variables \( \textbf{g}(j), C(j) \) and \( R(j) \) are defined as:

\[
\begin{align*}
\textbf{y}(j) &= \begin{bmatrix} \textbf{g}(j) \\ 0 \end{bmatrix}, \quad C(j) = \begin{bmatrix} I & 0 \\ 0 & V^{-1}(j) \end{bmatrix}, \\
R(j) &= \begin{bmatrix} 1/\mu_I & 0 \\ 0 & 1/\nu_I \end{bmatrix}
\end{align*}
\]

Note that the last term in (4) has been deliberately rewritten as a dynamic equation in (6). Given the form we have written (6), (7) in, we can see that another, equivalent interpretation of these equations specifies a dynamic estimation problem for the \( \tilde{f}(j) \). In particular, the components \( \tilde{f}(j) \) of the minimizer \( \tilde{f} \) of (4) are precisely the same as the series of smoothed estimates of \( f(j) \) based on the dynamic equation (6) and the observation (7) for \( 1 \leq j \leq n \). The advantage of the formulation (6) and (7) is that we may use efficient, recursive Kalman filtering based techniques for its solution. Specifically, we use the Mayne-Fraser form of the discrete-time Kalman smoother to arrive at the estimates of \( f(j) \). The preceding development shows how (4) can be interpreted as a stochastic estimation problem; a similar argument holds for (5).

We thus cast the solution to the original segmentation problem as a sequence of recursive estimation problems. We then apply efficient near-optimal filtering methods for their solution. As a result, rather than solving a single large optimization problem at each iteration, we instead solve a sequence of much smaller, near-optimal recursive estimation problems. The solution efficiently provides estimates as well as error variance information.
4. INFORMATION-FORM KALMAN FILTER

The Mayne-Fraser method for obtaining the solution to the smoothing problem requires a pair of Kalman filtering operations [3]. To perform each of these Kalman filtering steps based on (6), (7) we use an efficient version of the information form of the Kalman filter developed in [2]. This work shows how to perform efficient near optimal filtering via reduced order Markov Random Field Modelling.

There are two computationally taxing parts of the optimal filter as it is defined in [2]. These bottlenecks are a) the matrix inversion in the prediction step for the information matrix and b) the implicit definition of the updated estimate as a solution of a system of equations in the update step of the filter. Our approach, as in [2], is to make efficient approximations to expedite the steps without noticeably compromising the accuracy of the estimates. While we do not develop it here, the MRF-based interpretation of the update step in [2] coupled with [5] further allows us to pose the calculation in the update step of each Kalman filter as another dynamic estimation problem which has the same form as (6), (7). This allows us to again apply the same estimation theoretic approach we develop for the overall problem to efficiently and accurately solve its subproblems.

To summarize, the maximum-likelihood estimate of (6), (7) provided by the use of the optimal Mayne-Fraser smoother is exactly the same as the minimizer of (4). However, we implement a suboptimal Mayne-Fraser smoother to ameliorate the bottlenecks a) and b) described above. We estimate $f$ and $s$ 20% faster than comparable techniques used to solve for the same estimates based directly on (4) and (5). The approximation errors for such approximations perturb the estimates less than 2% of full scale for more than 90% of the pixels (for a given choice of parameters, $\mu$, $\lambda$, $\rho$, and $\nu$). However, in addition to the estimates, the smoother also produces error variance information along with the estimates. Solving for those error measures with traditional methods requires on the order of $n$ times as much computation as our method (where $n$ is the number of image columns).

5. RESULTS

A typical MRI proton density (PD) weighted brain scan is shown on the top of Figure 1. Note that the brain has already been segmented from the rest of the scan. A segmentation of this PD weighted scan using our statistically-based recursive approach is shown on the top middle of Figure 1. This is the smooth approximating field, $f$, in (1). The corresponding continuous edge field estimate, $s$, is shown on the bottom middle of Figure 1. Our approach also provides a measure of uncertainty for each pixel estimate. The estimation error standard deviation for each value of the edge estimate is shown in Figure 1 on the bottom. One way such uncertainty information can be used is to fuse the edge information obtained from different sources (e.g. edge estimates from registered T1-weighted imagery or CT data). In Figure 2 we demonstrate such edge fusion. We combine the edge estimates obtained from T1, T2 and PD weighted images of the same brain using a maximum likelihood framework. The fused edge image contains edge information that does not exist in the edge image of any single channel. For instance, in the fused edge estimate, the boundary indicated by the arrow is complete, indicating incorporation of collateral edge information, since the same boundary in the PD edge estimate is incomplete. Note also how the estimation error decreases as more information is added to the fused estimate of the edge field.

6. CONCLUSIONS

In this paper we have presented a recursive approach to the segmentation of MRI imagery. This approach was based on an estimation theoretic interpretation of the segmentation problem and while efficient, provides solutions indistinguishable from direct solution of an equivalent variational problem. In addition to its efficiency, being statistically based the method also provides associated measures of uncertainty of the resulting estimates. Such measures are useful not only for evaluation of the segmentation but also for subsequent stages of the processing. An illustration of the use of such information was demonstrated through the fusion of edge information obtained from registered multichannel (PD, T1, and T2) imagery.

7. REFERENCES


Figure 1: Original PD weighted image, smoothed image, edge estimate, and edge estimate standard deviation of error

Figure 2: Fusion of edge estimates and standard deviation of estimation error

Massachusetts General Hospital, Boston, MA, USA Using Smart Source Parsing Mar English.

