Automatic segmentation of magnetic resonance images using a decision tree with spatial information

Wen-Hung Chao\textsuperscript{a,b}, You-Yin Chen\textsuperscript{a,*}, Sheng-Huang Lin\textsuperscript{c}, Yen-Yu I. Shih\textsuperscript{d}, Siny Tsang\textsuperscript{e}

\textsuperscript{a} Department of Electrical and Control Engineering, National Chiao Tung University, No. 1001, Ta-Hueh Rd., Hsinchu 300, Taiwan, ROC
\textsuperscript{b} Department of Biomedical Engineering, Yuanpei University, No. 306, Yuanpei St., Hsinchu 300, Taiwan, ROC
\textsuperscript{c} Department of Neurology, Buddhist Tzu Chi General Hospital, No. 707, Sec. 3, Chung Yang Rd., Hualien 970, Taiwan, ROC
\textsuperscript{d} Institute of Biomedical Sciences, Academia Sinica, No. 128, Sec. 2, Academia Rd., Taipei 115, Taiwan, ROC
\textsuperscript{e} College of Criminal Justice, Sam Houston State University, Huntsville, TX 77341-2296, USA

\textbf{Abstract}

Here we proposed an automatic segmentation method based on a decision tree to classify the brain tissues in magnetic resonance (MR) images. Two types of data – phantom MR images obtained from IBSR (http://www.cma.mgh.harvard.edu/ibsr) and simulated brain MR images obtained from BrainWeb (http://www.bic.mni.mcgill.ca/brainweb) – were segmented using an automatic decision tree algorithm to obtain images with improved visual rendition. Spatial information on the general gray level ($G$), spatial gray level ($S$), and two-dimensional wavelet transform ($W$) was combined in-plane in two coordinate systems (Euclidean coordinates $(x, y)$ or polar coordinates $(r, \theta)$). The decision tree was constructed based on a binary tree with nodes created by splitting the distribution of input features of the tree. The spatial information obtained from MR images with different noise levels and inhomogeneities were segmented to compare whether the use of a decision tree improved the identification of human anatomical structures in a neuroimage. The average accuracy rates of segmentation for phantom images with a noise variation of 15 gray levels were 0.9999 and 0.9973 with spatial information ($G_x$, $x$, $y$, $r$, $\theta$) and ($S_x$, $x$, $y$, $r$, $\theta$), respectively, and 0.9999 and 0.9819 with spatial information ($G_x$, $x$, $y$, $S$, $r$, $\theta$) and ($W_x$, $x$, $y$, $G$, $r$, $\theta$). The average accuracy rates of segmentation for simulated MR images with a noise level of 5% were 0.9532 and 0.9439 with spatial information ($G_x$, $x$, $y$, $r$, $\theta$) and ($S_x$, $x$, $y$, $r$, $\theta$), respectively, and 0.9446 and 0.9287 with spatial information ($G_x$, $x$, $y$, $S$, $r$, $\theta$) and ($W_x$, $x$, $y$, $G$, $r$, $\theta$). The accuracy rates of segmentation were highest for both simulated phantom and brain MR images, having the lowest noise levels, from a reduction of overlapping gray levels in the images. The accuracies of segmentation were higher when the spatial information included the general gray level than when it included the spatial gray level, which in turn were higher than when it included the wavelet transform. Furthermore, the performance of segmentation was also evaluated with a boundary detection methodology that is based on the Hausdorff distance to compare with the mean computer to observer difference (COD) and mean interobserver difference (IOD) for gray matter (GM), white matter (WM), and all areas (ALL) from images segmented using the decision tree. The values of mean COD are similar and around 12 mm for GM segmented using the decision tree. Our segmentation method based on a decision tree algorithm presented an easy way to perform automatic segmentation for both phantom and tissue regions in brain MR images.

1. Introduction

Magnetic resonance (MR) imaging is widely used in clinical diagnosis. Segmentation is one of the techniques used to classify the brain tissues in MR images, which is a basic problem for identifying anatomical structures in MR image processing. Several segmentation methods have been applied in the analysis of anatomical structures involving three-dimensional (3D) reconstruction, tissue-type contour definition, clinical diagnosis [1,2], and in cortical surface segmentation, volume assessment of brain tissue, tissue classification, tumor segmentation, and characterization of various brain diseases such as sclerosis, epilepsy, stroke, cancer, and Alzheimer’s disease [3,4]. The accuracy of segmenting the cortical surface for analyzing the volumes of different tissues, such as gray matter (GM) and white matter (WM), significantly affects clinical diagnoses. It is this made difficult by the presence of imaging noise.
and inhomogeneities. Several segmentation techniques have been proposed to improve the detection of brain structures in MR images and the subsequent diagnoses. Both manual and automatic segmentation methods are used to segment brain MR images. Manual segmentations, such as thresholding, is a traditional method used to distinguish among different tissues in MR brain images [5–7], but it is difficult due to a low contrast-to-noise ratio, low signal-to-noise ratio (SNR), and tissue overlapping in the gray-level distributions. It is also a very labor-intensive and time-consuming procedure [8]. Therefore, several studies have investigated automatic segmentation methods for distinguishing brain MR images structures and improving the efficiency of segmentation and tissue classification [2,9–16]. Marroquin et al. presented an automatic segmentation method based on an accurate and efficient Bayesian algorithm [17]. Automatic segmentation based on a constrained Gaussian mixture model framework employed an expectation-maximization algorithm to determine parameters and to segment both simulated and real three-dimensional, T1-weighted noisy MR images [18]. Some of these automatic segmentation methods were used to classify the tissues (GM, WM, and cerebrospinal fluid (CSF)) in brain MR images. An automatic segmentation method has also been used to segment WM lesions [19]. However, it is essential to increase accuracy in the automatic segmentation of the GM, WM, and CSF.

Several studies have improved coil sensitivities and the performance of transmitter devices [20–24], but it remains difficult and expensive to reduce imaging noise and inhomogeneity through hardware improvements. The purpose of these studies was to obtain better anatomical structures of MR images. A low cost technique to obtain MR brain structures is valuable to study. Thus, the ability through software improvements to discriminate different tissue characteristics of brain structures is increasingly important. Spatial features defined as the combination of image intensities and in-plane information in two coordinate systems (Euclidean coordinates (x, y) or polar coordinates (r, θ)) in images have generally been used to extract the spatial features of MR images [18,19]. The spatial gray information was defined in the current study by combining neighboring pixel intensities, as described in Section 2. The wavelet-transform spatial information obtained from each local area was also used. The performances of the three types of spatial information were compared using decision tree algorithms. Decision trees are easily implemented according to the attributes of a subset in the entire data set, and provide rapid analysis. Decision trees have been widely used in the analysis of symbolic data sets, classifying EEG spatial patterns [25], and different regions of digital images sensed remotely [26]. The present study compared the performance of segmentation based on an automatic decision tree with different types of spatial information—the general gray level (G), spatial gray level (S), and two-dimensional wavelet transform (W)—to improve the accuracy of segmentation in MR images.

2. Materials and methods

2.1. Preprocessing for spatial information

Spatial information on the general gray level, spatial gray level, and wavelet transform were combined in Euclidean coordinates (x, y) or polar coordinates (r, θ) with image preprocessing. Noise and RF inhomogeneities often reduce the quality of MR images; therefore, their effects on the accuracy of segmentation need to be reduced by image manipulation.

The general gray level represents the intensity of each pixel expressed in Euclidean coordinates (x, y) or polar coordinates (r, θ) for MR image segmentation. The use of more spatial information in an image improved the accuracy of image segmentation. Two types of spatial feature information were used in this study. The first was the spatial gray level:

\[
S(x, y) = \sum_{i=1}^{n} \omega_i g_i(x, y),
\]

which is the summation of combined weighting \(\omega_i\) and gray level \(g_i(x, y)\) of pixel \(i\) on the neighboring area. The neighboring area was shown in Fig. 1(a), where \(n = 5\) and \(\omega_0\) was the weighting of the gray level at the center pixel with the nearest four pixels. The second type of spatial feature information used was the coefficient of the wavelet transform transferred from each local area to represent the wavelet spatial features of the center pixel for every location. The local area consisted of every nine pixels in an MR image, as shown in Fig. 1(b).

2.2. Segmentation

The proposed automatic decision tree segmentation method used in this study was the classification and regression tree (CART) proposed by Breiman et al. [27] to model the prediction tree by statistical analysis, considering outcome variables and decision questions to assess the prediction accuracy. The method protocol was described below.
2.2.1. Decision tree classification

In a classification tree, the decision tree classification structure is constructed to distinguish different classes through statistical analysis [25, 28]. Decision trees classify multidimensional spatial data through recursive partitioning steps. Each vector consisting of \( N \) sampled data in an \( M \)-dimensional space is given by

\[
\{x_m\}, \quad m = 1, \ldots, M, \tag{2}
\]

where \( M \) represents the dimension of the data space, with the class label in the data space as

\[
f \in \{1, \ldots, J\}. \tag{3}
\]

The subspaces can be illustrated easily to maximize the overall class separation for the \( M \)-dimensional spatial data set. The class separation is maximized during the partitioning step, and is subsequently processed as the basis for further partitioning to the \( M \)-dimensional spatial data set. A two-space and two-class example is described as follows for decision tree classification. The distributions of the two spatial data sets are shown in Fig. 2.

![Diagram](imageURL)

**Fig. 2.** Decision tree configuration: (a) example of the distribution of two subspaces from an entire space, and (b) structure of the corresponding decision tree graph.

2.2.2. Decision tree construction

Descendant nodes of greater purity are desired when constructing a decision tree, achieved by maximizing an impurity function. Descendant nodes have greater purity than ancestor nodes, and an impurity function \( \phi \) is defined based on node \( N \) defined as [25]

\[
i(N) = \phi(p(o_1|N), \ldots, p(o_j|N)), \tag{4}
\]

where \( p(o_j|N) \) is the conditional probability for class \( o_j \) of node \( N \). Impurity \( i(N) \) is maximal when node \( N \) has an equal number of cases for all classes. In other words, a node is maximally pure when the node comprises of a single category. The impurity function [27, 28] can be interpreted as a general variance impurity for two or more classes, which is the Gini impurity given by

\[
i(N) = \sum_{i \neq j} p(o_i|N)p(o_j|N) = 1 - \sum_j p(o_j|N)p(o_j|N), \tag{5}
\]

where \( p(o_j|N) \) and \( p(o_j|N) \) are the proportions of patterns for classes \( o_i \) and \( o_j \) at node \( N \), respectively. The Gini impurity is 0 if all the patterns are of the same class. At the beginning of the root node, the CART calculates the node impurity with the Gini impurity function. All decision tree nodes are decided by determining the best change in the impurity from the root node down to the terminal node, as shown in Fig. 2. A node consisting of a single class has the largest purity. Thus, the terminal node is then selected when the impurity of the node is 0. The largest impurity value is 1. The best change in impurity [28] is the difference between \( i(N) \) and a sum of
the impurities of $N_L$ and $N_R$ determined by
\[ \Delta i(N) = i(N) - p_L i(N_L) - p_R i(N_R), \tag{6} \]
where $N_L$ and $N_R$ are the left and right descent nodes, $i(N)$ is their impurities, and $p_L$ and $p_R$ are their fractions of patterns at node $N$, respectively. The CART employs an iterative approach to decide the split at node $N$ based on the best numerical change in the Gini impurity, which corresponds to the maximal class separation. At the beginning of the root node, the CART estimates the node impurity using the Gini impurity function of Eq. (5). Each of these nodes is decided by maximizing $\Delta i(N)$ and minimizing $i(N)$. This repetitive approach produces the partitioning step with the highest purity at the terminal nodes. In other words, maximal class separation is equivalent to minimizing the misclassification of classes in the decision tree at node $N$ [27,28]. The Gini impurity function of Eq. (5) evaluates the probability of misclassification at node $N$. A class in node $N$ can be estimated through Eq. (5) with conditional probability $p(\omega_i|N)$ and $p(\omega_i|\bar{N})$. The conditional probability of node $N$ can also be quantified using Eq. (5). Finally, the entire decision tree structures can be decided from the data set of the entire space by algorithmically applying these rules.

### 2.3. Simulated data

Two types of simulated data were used in this study: phantom MR images and simulated brain MR images. The images were obtained from IBSR (http://www.cma.mgh.harvard.edu/ibsr). They comprised of the circle center, circle ring, and background region, as shown in row 1 of Fig. 3, with noise variations of 15 or 30 gray levels. We also added RF inhomogeneities of 20% and 40% to the two SNR phantom images. The variations in the gray levels due to noise and inhomogeneities that were added to a gold-standard phantom image are designated in Table 1. The simulated MR images obtained from BrainWeb (http://www.bic.mni.mcgill.ca/brainweb) were T1-weighted 3-mm-thick images with noise levels of 3%, 5%, 7%, 9%, and 15%. Furthermore, images of these noise levels combined with RF inhomogeneities of 20% and 40% examines the performance of segmentation with spatial information of different qualities. An expert manually derived a gold-standard brain MR image with no noise or inhomogeneity from the original image. All of the simulated data were preprocessed to extract the spatial information and then segmented using the automatic decision tree algorithm.

### 2.4. Evaluation of segmentation

A quantization index was needed to evaluate the performance of segmentation based on the accuracy of the classification. The accuracy rate was calculated based on the overlap between the gold-standard reference image and a collection of segmentation results obtained from the proposed automatic decision tree segmentation method. The accuracy rate used in this study was quantified as the overlap fraction (OF) index, defined as
\[ OF = \frac{\text{Ref}(k) \cap \text{Seg}(k)}{\text{Ref}(k)}, \tag{7} \]
which is the accuracy rate of the segmented area in class $k$ relative to the area in the gold-standard reference image [19]. Three classes of phantom MR images (circle center, circle ring, and background) and four classes of simulated brain MR images (GM, WM, CSF, and background) were used in this study. The numerator in Eq. (7) represents the number classified or intersection area of voxels in class $k$ between the proposed automatic segmentation method and the gold standard. The denominator represents the area of voxels in class $k$ in the gold standard.

Another index, a boundary detection algorithm [33,34], was used to evaluate the performance of the brain tissue segmentation. The index is based on the Hausdorff distance to calculate statistical evaluation including Williams index (WI), percent statistic (P), confidence interval of WI, and confidence interval of P. The Hausdorff distance is defined as below. Two sets of all points in two curves are $A = \{a_1, a_2, \ldots, a_n\}$ and $B = \{b_1, b_2, \ldots, b_m\}$ [33,34].
\[
\begin{align*}
&d(a_i, B) = \min_j \| b_j - a_i \|, \\
&d(b_j, A) = \min_i \| a_i - b_j \|.
\end{align*}
\tag{8}
\tag{9}
\]
The Hausdorff distance of the two curves is defined as
\[
d(A, B) = \max_i \max_j d(a_i, B), \max_j d(b_j, A).
\tag{10}
\]
It is the maximum distance of the closet points between the two curves. The Williams index (WI) is defined as
\[
I = \frac{P_0}{P_n}
\tag{11}
\]
where $P_0$ is the average level of agreements between observer 0 and reference observers. The $P_0$ is given as
\[
P_0 = \frac{1}{n} \sum_{j=1}^{n} P_{0,j},
\tag{12}
\]
where the $P_n$ is the average level between the $n$ reference observers. $P_n$ is defined as
\[
P_n = \frac{2}{n(n-1)} \sum_{j=1}^{n} \sum_{j \neq j}^{n} P_{ij}.
\tag{13}
\]
Then, the 95% CI of the WI is checked for inclusion of the expected value.

The percent statistic (P) is defined as that the number of times (or boundaries) produced by the proposed algorithm which are within the interobserver range. This is hypothesized in advance
that the computer-generated boundaries and the observer outlined boundaries are samples from the same distribution. The expected percent of times that the computer-generated boundaries lie within the interobserver range is $100\left(\frac{n}{n+1}\right)$. For example, for three human observers, this expected percentage is 75%; for four observers, it is 80%; and for five observers, it is 83%. Next, the 95% CI of the percentage statistic is checked for inclusion of the expected value.

### 3. Results

#### 3.1. Results of phantom images

All simulated phantom MR images with different SNRs and inhomogeneities (see Table 1) obtained from the IBSR website were segmented with spatial information $(G, x, y), (S, x, y), (G, x, y, r,$ $\theta), (S, x, y, r, \theta), (G, x, y, S, r, \theta), (W, x, y, G, r, \theta)$, and $(W, x, y, G$, $r, \theta, S)$. Fig. 3 shows the original phantom images with different SNRs and inhomogeneities, and the segmentation results obtained using a decision tree algorithm. The images in row 1 of Fig. 3 were the original phantom images with noise variations and RF inhomogeneities of Var15, Var30, Var15RF20, Var15RF40, Var30RF20, and Var30RF40 (in columns 1–6, respectively), as listed in Table 1. The images in row 2 of Fig. 3 corresponded to those in row 1 segmented using the automatic decision tree with spatial information $(S, x, y)$. Euclidean coordinates $(x, y)$ and polar coordinates $(r, \theta)$ were also used for spatial information in this study. The images in row 3 of Fig. 3 corresponded to those in row 1 segmented using automatic decision tree with spatial information $(S, x, y, r, \theta)$. The images in row 4 of Fig. 3 consisted of those in row 1 segmented using a decision tree with spatial information $(G, x, y)$.

The images segmented with spatial information $(S, x, y)$ and $(S, x, y, r, \theta)$ (rows 2 and 3 of Fig. 3) showed better performance than those segmented with spatial information $(G, x, y)$. Images with a noise variation of 30 gray levels and 40% RF inhomogeneities constituted a very large fraction of the source phantom images. The performance for images with Var30, Var30RF20, and Var30RF40 (row 4 of Fig. 3) segmented with spatial information $(G, x, y)$ was

---

**Table 2**

Designations of the original simulated MR images obtained by combining the noise levels and inhomogeneities parameters.

<table>
<thead>
<tr>
<th>Designation Combined noise level and inhomogeneities parameter</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1n3</td>
</tr>
<tr>
<td>T1n5</td>
</tr>
<tr>
<td>T1n7</td>
</tr>
<tr>
<td>T1n9</td>
</tr>
<tr>
<td>T1n15</td>
</tr>
<tr>
<td>T1n3RF20</td>
</tr>
<tr>
<td>T1n5RF20</td>
</tr>
<tr>
<td>T1n7RF20</td>
</tr>
<tr>
<td>T1n9RF20</td>
</tr>
<tr>
<td>T1n15RF20</td>
</tr>
<tr>
<td>T1n3RF40</td>
</tr>
<tr>
<td>T1n5RF40</td>
</tr>
<tr>
<td>T1n7RF40</td>
</tr>
<tr>
<td>T1n9RF40</td>
</tr>
<tr>
<td>T1n15RF40</td>
</tr>
</tbody>
</table>

---

Fig. 4. Average accuracy rates of segmentation obtained using a decision tree with different spatial information from original phantom images with different noise variations and inhomogeneities.

Fig. 5. Results of segmentation using a decision tree for simulated MR images obtained from BrainWeb. Upper row contains the original MR images with noise levels of T1n3, T1n5, T1n7, T1n9, and T1n15. Lower row contains the corresponding images resulting from segmentation with spatial information $(G, x, y, r, \theta)$. 

---
unclear. The segmentation of phantom MR images with Var15, Var30, Var15RF20, Var15RF40, Var30RF20, and Var30RF40 using a decision tree with spatial information \((G, x, y, r, \theta), (G, x, y), (G, x, y, r, \theta), (G, x, y, r, \theta), (S, x, y), (W, x, y, G, r, \theta), (W, x, y, G, r, \theta, S)\) produced better performance. Fig. 4 shows the average accuracy rate of phantom images with different SNRs and inhomogeneities segmented by a decision tree algorithm with different spatial information. The average accuracy rates were averaged across all phantom image regions (circle ring, circle center, and background). They were evaluated by the OF index as described in Section 2. The average accuracy rates of segmentation for phantom images with Var15, Var30, Var15RF20, Var15RF40, Var30RF20, and Var30RF40 and segmentation spatial information \((G, x, y, r, \theta), (G, x, y), (G, x, y, r, \theta), (G, x, y, S, r, \theta), (S, x, y), (W, x, y, G, r, \theta), (W, x, y, G, r, \theta, S)\) were shown in Fig. 4. The highest average accuracy rates of segmentation are in the range of 0.9819–0.9999 for phantom images with Var15 and Var15RF20 segmented using a decision tree for all of the used spatial information. The higher average accuracy rates of segmentation are shown in Fig. 4 for phantom images with Var15RF40 for all of the used spatial information. The accuracy rates of phantom images with Var30 and Var30RF20 segmented by a decision tree for spatial information \((G, x, y, r, \theta), (G, x, y), (G, x, y, r, \theta), (G, x, y, S, r, \theta), (S, x, y), (W, x, y, G, r, \theta), (W, x, y, G, r, \theta, S)\) were moderate, ranging from 0.8964 to 0.9872. The average accuracy rates of phantom images with Var30 and Var30RF20 segmented by a decision tree for spatial information \((S, x, y, r, \theta)\) and \((S, x, y)\) were also close to the highest values. Segmenting images with Var30RF40 with spatial information \((G, x, y, r, \theta), (G, x, y), (G, x, y, S, r, \theta), (W, x, y, G, r, \theta), (W, x, y, G, r, \theta, S)\) produced the lowest average accuracy rates, although segmentation with spatial information \((S, x, y, r, \theta)\) and \((S, x, y)\) produced a higher average accuracy rate of 0.9461. The segmentation results shown in Fig. 3 indicated that the automatic decision tree successfully segmented phantom images with different noise variations and RF inhomogeneities.

### 3.2. Results of simulated brain MR images

All simulated MR brain images with different noise levels and inhomogeneities (see Table 2), as described in Section 2, were also segmented using the automatic decision tree with different spatial information \((G, x, y), (G, x, y, r, \theta), (G, x, y, r, \theta), (G, x, y, S, r, \theta), (W, x, y, G, r, \theta), (W, x, y, G, r, \theta, S)\) produced the highest average accuracy rates (0.9374–0.9598), with the resulting images shown in Fig. 5. Segmentation these MR images with spatial information \((G, x, y, S, r, \theta)\) and \((W, x, y, G, r, \theta)\) produced moderate and low average accuracy rates of 0.9132–0.9626 and 0.8920–0.9297, respectively. Fig. 5

![Fig. 6. Average accuracy rates of segmentation with different spatial information from the original MR images with noise levels of T1n3, T1n5, T1n7, T1n9, and T1n15.](image-url)
Fig. 8. Average accuracy rates of segmentation with different spatial information from MR images with noise parameters of T1n3RF20, T1n5RF20, T1n7RF20, T1n9RF20, and T1n15RF20.

shows that all of the simulated brain MR images with these noise levels were successfully segmented with the automatic decision tree with spatial information \((G, x, y, r, \theta)\), \((G, x, y)\), \((S, x, y, r, \theta)\), \((S, x, y)\), \((W, x, y, G, r, \theta)\), and \((W, x, y, G, r, \theta)\).

Fig. 7 shows original simulated brain MR images with different noise levels and an RF inhomogeneity of 20% obtained from BrainWeb (upper row) and the images resulting from segmentation with spatial information \((G, x, y, r, \theta)\) (lower row). The segmented images showed better visual rendition. Fig. 8 shows the average accuracy rates of segmentation with different spatial information from the simulated brain MR images shown in Fig. 7. The average accuracy rates decreased as the noise level increased from T1n3RF20 to T1n15RF20 for most of the segmentations with this spatial information. They do not differ greatly between Figs. 6 and 8, ranging from 0.96 to 0.89. The presence of 20% RF inhomogeneities had little effect on segmentation of these simulated brain MR images. The average accuracy rates of segmentation in these brain MR images (Fig. 8) with spatial information \((G, x, y, r, \theta)\), \((G, x, y, S, r, \theta)\), and \((W, x, y, G, r, \theta)\) were 0.9376–0.9587, 0.9144–0.9538, and 0.8865–0.9285, respectively. All of the simulated brain MR images with these noise levels and inhomogeneities were successfully segmented with this spatial information for all of the accuracy rates of automatic decision tree segmentation.

Fig. 9 shows original simulated brain MR images with different noise levels and an RF inhomogeneity of 40% from BrainWeb (upper row) and the images resulting from segmentation with spatial information \((G, x, y, r, \theta)\) (lower row). The segmented images showed better visual rendition. Fig. 10 shows the average accuracy rates of segmentation with different spatial information from the simulated brain MR images shown in Fig. 9. The average accuracy rates decreased as the noise level increased from T1n3RF40 to T1n15RF40 for most of the segmentations with this spatial information. They do not differ greatly between Figs. 8 and 10 except in lower values of the range. The 40% RF inhomogeneities have a greater effect on the segmentation than that shown for the 20% RF inhomogeneities in Figs. 7 and 8. The average accuracy rates of segmentation of these MR images with spatial information \((W, x, y, G, r, \theta)\) as shown in Fig. 10 changed from 0.8810 to 0.9261. They...
were also lower than that in brain MR images with T1n3RF20 to T1n15RF20 because of the larger fraction of the combined inhomogeneities. A higher average accuracy rate of segmentation with spatial information \((G, x, y, r, \theta)\) for simulated brain MR images made it easier to classify the GM, WM, and CSF in simulated MR brain images with different noise levels and inhomogeneities, as shown in Fig. 11. Fig. 11(a) shows the accuracy rates of segmentation with spatial information \((G, x, y, r, \theta)\) for simulated brain MR images with T1n3, T1n5, T1n7, T1n9, and T1n15; T1n3RF20, T1n5RF20, T1n7RF20, T1n9RF20, and T1n15RF20; and T1n3RF40, T1n5RF40, T1n7RF40, T1n9RF40, and T1n15RF40 (c).

Fig. 11. Accuracy rates of GM, WM, and CSF in simulated MR images segmented using a decision tree with spatial information \((G, x, y, r, \theta)\) for T1n3, T1n5, T1n7, T1n9, and T1n15 (a); T1n3RF20, T1n5RF20, T1n7RF20, T1n9RF20, and T1n15RF20 (b); and T1n3RF40, T1n5RF40, T1n7RF40, T1n9RF40, and T1n15RF40 (c).

The performances of simulated brain MR image segmentation were also evaluated with a well-known methodology of boundary detection algorithm [33,34]. The evaluations based on the Hausdorff distance in comparison with the mean computer to observer difference (COD), mean interobserver difference (IOD), Williams index (WI), WI confidence interval (CI), percent statistic (P) for gray matter (GM), white matter (WM), and all areas (All) from images segmented with the decision tree with spatial information \((G, x, y, r, \theta)\), \((S, x, y, r, \theta)\), \((G, x, y, S, r, \theta)\), \((W, x, y, G, r, \theta)\), and \((W, x, y, G, r, \theta, S)\) from simulated brain MR images with all noise and inhomogeneity levels were shown in Table 3. The WI was close to one, indicating similar differences between COD boundaries and expert gold-standard boundaries, and that expert generated boundaries. All areas (All) in Table 3 represents that the values of GM, WM, CSF, and background of the brain MR images were calculated. The values of mean COD were approximately 12 mm. The values of mean COD were all close to one for GM, WM and All segmented using decision tree (Table 3). Most of the upper limit of 95% CI were larger than one, except 0.95 for the All (all areas) segmented with \((S, x, y, r, \theta)\). 0.96 for the WM segmented with \((G, x, y, S, r, \theta)\), and 0.95 for the WM segmented with \((W, x, y, G, r, \theta)\). The expected value of P in Table 3 was 66.7%. All of the upper limits of 95% of P in Table 3 for Hausdorff distance were lower than its expected value. Discrepancies among areas were shown in the percent statistics (P). The mean COD of Hausdorff distance was not smaller due to the variations of GM, WM, and CSF brain MR image boundaries. The edge variations of GM, WM, and CSF are more complex than the edge of other organs in human body. Evaluation results with Hausdorff distance from the segmented tissues might not show better performance than those of other organs. The performance of segmentation in the present study might have been affected by the segmentation gold-standard reference established by an expert.

4. Discussion

A proposed automatic segmentation method using a decision tree was used in the present study to classify different tissue types in brain MR images. The phantom and simulated MR images obtained from IBSR and BrainWeb, respectively, were both successfully segmented by the proposed decision tree algorithm. The performance of the proposed segmentation technique was evaluated using a previously described index [19,29]. The gray-level distributions in the phantom MR images differed more between the various regions. The spatial gray-level information had a greater effect on the performance of phantom image segmentation. The gray-level distributions of each tissue overlapped more in different regions in simulated brain MR images than in the phantom MR images. Therefore, the spatial gray-level information is useful for assessing the performance of segmentation, with local features of the spatial information being more suitable for assessing the accuracy of segmentation by a decision tree of a simulated MR image. The average accuracy rates were higher with spatial information \((S, x, y, r, \theta)\) and \((S, x, y)\) for the simulated phantom MR images with all
the used noise levels and inhomogeneities, which was due to the gray levels of the spatial information being the main factor affecting segmentation of the phantom images. The average accuracy rates were lower for all the used spatial information when the simulated phantom MR images were combined with a noise variation of 30 gray levels. Furthermore, the average accuracy rates were highest with spatial information \((G, x, y, r, \theta)\) for simulated brain MR images with all the used noise levels and inhomogeneities due to the location attribute of the spatial information being more important than the gray-level information. The average accuracy rates were lowest for all the used spatial information when the simulated brain MR images contained 15% noise, which represented the largest fraction of images. The best results of segmentation were obtained in this study for simulated and brain MR images with the lowest noise levels.

The noise level is the main factor responsible for overlapping of the gray-level distribution in MR images. Also, the gray level is the main spatial feature that affects the performance of segmentation in phantom MR images, and hence it is the main decision attribute of tree structures. These characteristics were confirmed in both phantom and simulated MR images. The average accuracy rates of segmentation with spatial information \((G, x, y, r, \theta)\) were highest for phantom images with Var15, Var15RF20, and Var15RF40 (0.9999, 0.9999, and 0.9998, respectively), and were lowest for phantom images with Var30, Var30RF20, and Var30RF40 (0.9938, 0.9164, and 0.8778). The decrease (0.06) in the average accuracy rate for phantom images with Var15 and Var30 was more than that (0.0099) for phantom images with Var15 and Var15RF40 (see Fig. 4). The noise variation was the main factor affecting the accuracy rate of phantom image segmentation. In simulated MR images, the average accuracy rates of segmentation with spatial information \((G, x, y, r, \theta)\) were highest for simulated MR images with T1n3 and T1n15 (0.9598 and 0.9374, respectively) (Fig. 6), and lowest for simulated MR images with T1n3RF40 and T1n15RF40 (0.9582 and 0.9371, respectively). The decrease (0.0224) in the average accuracy rate for simulated images with T1n3 and T1n15 was more than that (0.0016) for phantom images with T1n3 and T1n3RF40 (see Fig. 10). The noise level was also the main factor affecting the accuracy rate of simulated MR image segmentation. Noise had similar effects on the trends in the performance of tissue (GM, WM, and CSF) segmentation of simulated MR images (see Fig. 11), and those on GM and WM segmentation were similar to those of previously reported approaches [17,18,30]. The accuracy rates of our segmentation method increased with decreasing noise level, which is consistent with previous results for tissues or the overall cortical surface [17,18,30,31]. Our method was also suitable for segmenting MR images, although its performance decreased as the noise level in the images increased.

For comparison, consider the spatial information approach proposed by Anbeek et al. [19]. In phantom images (see Fig. 4), the average accuracy rates of segmentation for phantom images with Var15 were 0.9999 and 0.9973 with spatial information \((G, x, y, r, \theta)\) and \((S, x, y, r, \theta)\), respectively, 0.9999 and 0.9973 with spatial information \((G, x, y)\) and \((S, x, y)\), and 0.9999 and 0.9819 with spatial information \((G, x, y, S, r, \theta)\) and \((W, x, y, G, r, \theta)\). In simulated MR images (see Fig. 6), the average accuracy rates of segmentation for simulated MR images with T1n5 were 0.9532 and 0.9439 with spatial information \((G, x, y, r, \theta)\) and \((S, x, y, r, \theta)\), respectively, 0.9480 and 0.9369 with spatial information \((G, x, y)\) and \((S, x, y)\), and 0.9446 and 0.9287 with spatial information \((G, x, y, S, r, \theta)\) and \((W, x, y, G, r, \theta)\). The overlapping of gray levels of noise was greater for spatial information \((S)\) obtained from five neighboring pixels (see Fig. 1(a)) in a local region than for spatial information \((G)\) obtained from a single gray-level intensity. Therefore, the accuracy rates of segmentation with spatial information \((S, x, y)\) and \((S, x, y, r, \theta)\) were lower than those of segmentation with spatial information \((G, x, y)\) and \((G, x, y, r, \theta)\). The overlapping of gray levels of noise was greater for spatial information \((W)\) obtained from nine neighboring pixels (see Fig. 1(b)) in a local region than for spatial information \((G)\) obtained from a single gray-level intensity. Thus, the accuracy rates of segmentation with spatial information \((W, x, y, G, r, \theta)\) were lower than those of segmentation with spatial information \((G, x, y, S, r, \theta)\).

<table>
<thead>
<tr>
<th>Spatial information</th>
<th>Tissue</th>
<th>COD (mm)</th>
<th>IOD (mm)</th>
<th>WI</th>
<th>95% CI</th>
<th>P (%)</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>((G, x, y, r, \theta))</td>
<td>GM</td>
<td>12.66</td>
<td>13.60</td>
<td>1.10</td>
<td>(1.10, 1.09)</td>
<td>64.4</td>
<td>(61.5, 67.3)</td>
</tr>
<tr>
<td></td>
<td>WM</td>
<td>12.37</td>
<td>12.25</td>
<td>0.98</td>
<td>(0.99, 0.97)</td>
<td>24.4</td>
<td>(3.5, 45.4)</td>
</tr>
<tr>
<td></td>
<td>All</td>
<td>11.62</td>
<td>14.99</td>
<td>1.12</td>
<td>(1.18, 1.05)</td>
<td>40.0</td>
<td>(18.4, 61.6)</td>
</tr>
<tr>
<td>((S, x, y, r, \theta))</td>
<td>GM</td>
<td>12.08</td>
<td>13.60</td>
<td>1.14</td>
<td>(1.14, 1.13)</td>
<td>62.6</td>
<td>(56.6, 67.8)</td>
</tr>
<tr>
<td></td>
<td>WM</td>
<td>10.98</td>
<td>12.25</td>
<td>1.04</td>
<td>(1.05, 1.03)</td>
<td>40.0</td>
<td>(18.4, 61.6)</td>
</tr>
<tr>
<td></td>
<td>All</td>
<td>15.05</td>
<td>14.99</td>
<td>0.96</td>
<td>(0.97, 0.95)</td>
<td>20.0</td>
<td>(11.3, 38.9)</td>
</tr>
<tr>
<td>((G, x, y, S, r, \theta))</td>
<td>GM</td>
<td>12.85</td>
<td>13.60</td>
<td>1.09</td>
<td>(1.10, 1.08)</td>
<td>57.7</td>
<td>(47.4, 68.2)</td>
</tr>
<tr>
<td></td>
<td>WM</td>
<td>12.71</td>
<td>12.25</td>
<td>0.97</td>
<td>(0.98, 0.96)</td>
<td>26.6</td>
<td>(5.0, 48.3)</td>
</tr>
<tr>
<td></td>
<td>All</td>
<td>15.15</td>
<td>14.99</td>
<td>1.07</td>
<td>(1.14, 1.06)</td>
<td>35.5</td>
<td>(11.3, 58.0)</td>
</tr>
<tr>
<td>((W, x, y, G, r, \theta))</td>
<td>GM</td>
<td>12.87</td>
<td>13.60</td>
<td>1.08</td>
<td>(1.08, 1.07)</td>
<td>62.2</td>
<td>(56.6, 67.8)</td>
</tr>
<tr>
<td></td>
<td>WM</td>
<td>12.10</td>
<td>12.25</td>
<td>0.97</td>
<td>(0.90, 0.95)</td>
<td>22.2</td>
<td>(2.2, 42.2)</td>
</tr>
<tr>
<td></td>
<td>All</td>
<td>11.78</td>
<td>14.99</td>
<td>1.13</td>
<td>(1.19, 1.07)</td>
<td>53.3</td>
<td>(38.9, 67.8)</td>
</tr>
<tr>
<td>((W, x, y, G, r, \theta))</td>
<td>GM</td>
<td>12.29</td>
<td>13.60</td>
<td>1.12</td>
<td>(1.11, 1.13)</td>
<td>57.8</td>
<td>(47.4, 68.2)</td>
</tr>
<tr>
<td></td>
<td>WM</td>
<td>11.93</td>
<td>12.25</td>
<td>0.99</td>
<td>(0.97, 1.00)</td>
<td>26.7</td>
<td>(5.1, 48.3)</td>
</tr>
</tbody>
</table>
images with T1n3, T1n3RF20, and T1n3RF40 were 0.9077, 0.9054, and 0.9038, respectively (see Fig. 11). These data indicated there were no significant differences among accuracy rates of segmentation of tissues in inhomogenous MR images. Thus, the presence of inhomogeneity in MR images might not decrease the accuracy rates of segmentation for both phantom MR images and simulated MR images. Furthermore, segmentation performance was also evaluated by a boundary detection methodology. Evaluation results with Hausdorff distance from the segmented tissues demonstrated that mean computer-generated and mean expert gold standard showed similar differences from the expert gold-stand. The percent statistics (P) was the largest when tissues were segmented using decision tree with spatial information (G, x, y, r, θ). Other approaches could also be used to determine the gold-standard reference image. A more sophisticated method [32] is to have a group of experts construct the reference gold-standard image, which might improve the accuracy of segmentation.

In conclusion, our segmentation method based on a decision tree algorithm presented a useful way to perform automatic segmentation for both phantom and tissue (GM, WM, and CSF) regions in brain MR images. It provided an easy method, through the partition of each spatial information (spatial feature), to form a decision tree for the brain tissues segmentation in MR images. The accuracy rates of segmentation were highest for both simulated phantom and brain MR images, having the lowest noise levels, from a reduction of overlapping gray levels in the images. The accuracies of segmentation were higher when the spatial information included the general gray level (G) than when it included the spatial gray level (S), which in turn were higher than when it included the wavelet transform (W). Finally, the accuracy rate of our segmentation method was not affected by inhomogeneity in MR images.

Acknowledgments

The authors thank Mr. Kuan-Hung Lin for his suggestions and comments concerning this manuscript. This study was partially supported by grants NSC95-2221-E-009-171-MY3 from the National Science Council of the Republic of China and grant VGHUST96-P5-19 from VGHUST Joint Research Program, Tsou's Foundation, Taiwan.

References


Wen-Hung Chao  
Wen-Hung Chao received the M.S. degree in biomedical engineering from National Cheng Kung University, Tainan, Taiwan, in 1986. He also received Ph.D. degree in the Department of Electrical and Control Engineering from National Chiao-Tung University, Hsinchu, Taiwan in 2008. Since August 1996, he has been a Lecturer with the Department of Biomedical Engineering, Yuanpei University, Hsinchu, Taiwan. His current research interests are biomedical signal processing, image processing, fuzzy neural networks, and designing biomedical instruments.

You-Yin Chen is an Assistant Professor of the Department of Electrical and Control Engineering at National Chiao-Tung University in Hsinchu, Taiwan. He received his Ph.D. degree in 1994 and M.S. (1992) degrees, respectively, in Electrical Engineering at National Taiwan University and B.S. degree (1996) in the Department of Electrical Engineering from the National Chiao-Tung University. Dr. Chen’s research interest is in the field of Neuroengineering. Topics of Neuroengineering include the analysis of neuroimages, neural ensemble recordings and brain-machine interfaces (BMI).

Sheng-Huang Lin received his M.D. degree from Kaohsiung Medical School, Taiwan, in 1992, and he obtained his M.Sc. degree in the Institute of Medical Science at Tzu Chi
University, Taiwan in 2000. Currently, he is an attending neurologist in the Buddhist Tzu Chi General Hospital and focusing on the research of Parkinson disease and movement disorders.

Yen-Yu I. Shih received his B.S. degree in Biomedical Imaging and Radiological Sciences from National Yang-Ming University, Taiwan, in 2004, and he obtained his Ph.D. degree in the Institute of Biomedical Engineering at National Taiwan University, Taiwan in 2007. Currently, he is a post-doctoral researcher in the Institute of Biomedical Sciences at Academia Sinica and focusing on the research of imaging nociceptive responses in rodent brain using MRI and PET.

Siny Tsang completed her B.S. in Agriculture from National Taiwan University in 2007. She is now a master student in the College of Criminal Justice at Sam Houston State University, U.S.A. Her research interests include psychopathy, serial offenders, gender differences between offenders, and the neurological mechanism of criminal minds.