Unsupervised Multimodal Magnetic Resonance Images Segmentation and Multiple Sclerosis Lesions Extraction based on Edge and Texture Features

Tannaz AKBARPOUR1*, Mousa SHAMSI1, Sabalan DANESHVAR2, Masoud POOREISA3

1Faculty of Electrical Engineering, Sahand University of Technology, New Sahand Town, PO Box 51335/1996, Tabriz, Iran
2Faculty of Electrical and Computer Engineering, University of Tabriz, Bist o Noh Bahman Bulvar, Tabriz, Iran
3Faculty of Medicine, Tabriz University of Medical Sciences, No 2 Central Building, University Street, PO Box 5165665931, Tabriz, Iran
E-mails: t_akbarpour@sut.ac.ir, shamsi@sut.ac.ir, daneshvar@tabrizu.ac.ir, pooreisam@tbzmed.ac.ir

* Author to whom correspondence should be addressed, Tel: (+98) 938-431-8058.

Received: May 11, 2017 / Accepted: June 25, 2017

Abstract
Segmentation of Multiple Sclerosis (MS) lesions is a crucial part of MS diagnosis and therapy. Segmentation of lesions is usually performed manually, exposing this process to human errors. Thus, exploiting automatic and semi-automatic methods is of interest. In this paper, a new method is proposed to segment MS lesions from multichannel MRI data (T1-W and T2-W). For this purpose, statistical features of spatial domain and wavelet coefficients of frequency domain are extracted for each pixel of skull-stripped images to form a feature vector. An unsupervised clustering algorithm is applied to group pixels and extracts lesions. Experimental results demonstrate that the proposed method is better than other state of art and contemporary methods of segmentation in terms of Dice metric, specificity, false-positive-rate, and Jaccard metric.

Keywords: Multiple Sclerosis (MS); Segmentation; Multichannel Magnetic Resonance Imaging (MRI); Wavelet; Energy; Entropy

Introduction

Multiple Sclerosis (MS) is a chronic persistent inflammatory, demyelinating and degenerative disease of the Central Nervous System (CNS), characterized by areas of inflammation, demyelination and axonal loss often causing motor, sensorial, vision, coordination and cognitive impairment [1]. Two main events related to MS are progression and relapses. Relapses are the clinical definition of inflammatory demyelination distributed over CNS. Remission of symptoms at first stages of illness is the result of compensatory mechanisms and demyelination, but they get less efficient as the illness progresses.

White matter lesions caused by inflammation are the main symbols of MS. These lesions are characterized by demyelination, axonal conduction block and axonal injury. These lesions are noticeable in magnetic resonance imaging (MRI), in form of hyperintensity in conventional T2-weighted images and hypointensity in T1-weighted images. These images are broadly utilized in detection of MS lesions and produce quantitative assessments of lesion load and inflammatory...
activity. MRI-derived information play an important role in many aspects related to MS such as diagnosis and monitoring the therapy efficacy.

The presence, distribution pattern of white matter (WM) lesions and appearance of new lesions are decisive elements in diagnosis process of MS [2]. Determining and extracting lesions are the first steps to distinguish MS disease severity and more specific measures of injury. For clinical processes, MRI images are visually evaluated for qualitative and quantitative analysis. Manual segmentation of lesions is a time consuming and difficult task, thus being subjected to large inter and intra-expert variability. Quantitative analysis has been helpful in evaluation of disease process and evaluation of therapy in recent years [3-6]. In quantitative analysis of lesions, segmentation of lesions is exploited to calculate lesion counts and volumes. Automatic segmentation methods could be an effective tool in detecting lesions with high precision.

Automatic segmentation methods could be classified according to [7] into four categories: supervised strategies based on atlas [8-11], supervised strategies based on learning from manual segmentation [12-17], unsupervised strategies to segment tissue [18-21], unsupervised strategies to segment only lesions [22-25]. The main goal of all these researches is to find reliable and fully automatic methods, which could be widely employed.

In supervised learning methods, a set of images in which the desired segmentation is known is used as training set to build and tune the algorithm. In supervised segmentation algorithms proper selection of training data is critical because improper selection of training data results in variations of computation time as well as inaccuracy of segmentation results. In the case of unlabeled data, the images could be segmented via clustering algorithms or unsupervised segmentation. Thus, no human expertise is involved in the pixel classification task.

Tissues appear in different intensity levels during different protocols of MRI. In early approaches, information of only one MRI protocol was used to detect and segment lesions, due to availability of images [26]. Merging modalities, namely, multichannel data could be beneficial in increasing feature space dimensions, leading to better discrimination of brain tissues and increasing accuracy [27-33]. Most of previous methods are based on statistical analysis, but due to [34] automatic methods outperform statistical approach.

Based on this idea, a new method is presented in this paper for segmentation of MS lesions of MRI based on T1 and T2 weighted MR images, which combines advantages of multichannel data and clustering algorithms in extracting lesions.

Material and Method

As mentioned above; our purpose is to introduce a method that could improve accuracy of segmentation results especially in presence of noise. Proposed method is a combination of a well-known clustering algorithm and time and frequency domain features, as shown in Figure 1. The method takes T1 and T2 weighted images of brain with MS lesions as input and provides a segmented image as output. Images are co-registered before performing the main process of the algorithm.

To prepare images for processing some preprocessing should be performed on input images. Skull stripping as the preprocessing step is done by means of region properties and thresholds. Steps to remove skull and preparing image for main algorithm are:

1. binarizing image by means of thresholding
2. finding percentage of filled area in image
3. labeling this area and achieving a mask
4. multiplying this mask into original image and stripping skull

Main loop of algorithm consists of following steps (Figure 1): 

**Step 1:** Swipe each image with a $3 \times 3$ window and extract statistical features described in next section.

**Step 2:** Transform each image to wavelet space based on stationary wavelet transform (SWT) of db5 filter and extract approximate, horizontal, vertical and diagonal coefficients of each pixel.
Step 3: Perform feature fusion and form feature vector, including four statistical and eight wavelet features for each pixel.

Step 4: Exploit FCM to cluster pixels into four groups.

Step 5: Label pixels and form segmented image.

Figure 1. Block diagram of the proposed method

Fuzzy C-Means Clustering

FCM is the most famous fuzzy clustering algorithm [35]. Consider a set of $n$ samples, $X = (x_1, \ldots, x_n)$. These samples are to be divided into $c$ clusters. The algorithm optimizes the following cost function, iteratively [36]:

$$J_m = \sum_{i=1}^{n} \sum_{j=1}^{c} u_{ij}^m \|x_i - v_j\|^2$$

(1)

where $u_{ij}$ is the membership of pixel $x_i$ in the $j$th cluster, $v_j$ is the center of $j$th cluster; $\| \|$ is the metric norm and $m$ is the parameter which controls degree of fuzziness, which is usually set to 2.

The membership function shows the probability that a pixel is associated to a cluster. At each iteration, values of membership function and cluster centers are updated as follows:

$$u_{ij} = \frac{1}{\sum_{k=1}^{c} \left( \frac{\|x_i - v_k\|^2}{\|x_i - v_j\|^2} \right)^{\frac{2}{m-1}}$$

(2)

$$v_j = \frac{\sum_{i=1}^{n} u_{ij}^m x_i}{\sum_{i=1}^{n} u_{ij}^m}$$

(3)

Feature Space

Members of feature vector are introduced in this section. Features are extracted in both frequency and spatial domains. Stationary wavelet transform (SWT) is exploited to extract features of frequency domain due to its better performance in removing noise effects and extracting edges.
compared to other wavelet transforms [37]. Statistical features are extracted in spatial domain to add information of surrounding pixels.

**Stationary Wavelet Transform (SWT)**

In traditional wavelet transform, each level is downsampled and filtered to get next level. Afterwards, there is a reduction of $1/2^m$ in number of samples, $m$ is the number of decomposition levels. Thus, pixel based segmentation cannot be performed based on features of decomposed signal [38]. A new method called SWT to extract features of texture at different scales by means of wavelet was proposed in [39]. To use the decomposition, filter is downsampled at each level based on equation (4). Thereafter it is convolved with the signal to achieve signal at next level using equation (5):

$$h_{r+1}(k) = [h]_{2^r} h_k(k)$$

$$g_{r+1}(k) = [g]_{2^r} g_k(k)$$

The equation above means the dilution of filters $h$ and $g$ by a factor 2 at each iteration.

$$s_{r+1}(k) = h_{r+1}(k)s_k(k)$$

$$d_{r+1}(k) = g_{r+1}(k)d_k(k)$$

Each step includes convolving with basic filters $h$ and $g$, which are expanded by inserting enough of zeros between filter taps. The complexity of this process is proportional to sample count. Achieved subsignals have the same lengths the original signal. They have information of different frequency regions effective in segmentation.

**Texture Features**

Statistical features are employed here to add information of surrounding pixels to feature space. Two statistical features, namely, energy and entropy are used in this paper[40]:

$$\text{energy} = \frac{1}{N} \sum_{n \in N} x_i^2$$

$$\text{entropy} = -\sum_{n \in N} x_i^2 \log x_i^2$$

where $x_i$ is the pixel intensity and $N$ is a neighborhood of $n$ pixels centered at $x_i$. Equal values of energy and entropy for two adjacent pixels means the same values in neighborhood, thus, belonging to same type of texture.

Feature fusion is the basic idea of proposed method. For this aim, four coefficient of wavelet transform for each pixel are augmented with statistical feature derived based on a $3 \times 3$ window for both input images to form a twelve member feature vector for each pixel. This feature vector is fed into an unsupervised clustering algorithm to group pixels. Labeling members of group yields final image.

**Experimental Set-Up**

In order to assess the performance of proposed algorithm, it is tested on two sets of multichannel data, available at brainweb dataset [41] and the segmented outputs are compared with three methods: Ortiz et al. [42], Demirhan and Guler [40], Sarti et al. [43], Anbeek et al. [14] and single-channel T2-weighted segmentation of proposed method [44]. Numerical evaluation plays an important role in deciding about performance of method. Performance is investigated based on four metrics, described briefly as follows:

- **Dice metric**

  Dice metric quantifies overlap of segmentation results and manually segmented images. For two regions of $S_1$ and $S_2$, dice metric is defined as [45]:

$$DC = \frac{2|S_1 \cap S_2|}{|S_1| + |S_2|}$$

  High amount of this metric means more overlap with ground truth, thus better performance.
Specificity
Specificity measures the proportion of negative pixels that are correctly identified as equation (9). Higher specificity means better performance in omitting negative pixels from regions.

\[
specificity = \frac{TN}{TN + FP}
\]  (9)

where

\[
TN = |\text{ref}\cap\text{seg}|
\]  (10)

\[
FP = |\text{seg} - \text{ref}|
\]  (11)

False Positive Rate (FPR)
FPR measures the proportion of negative pixels that are classified as positive, as follows:

\[
FPR = \frac{FP}{FP + TN}
\]  (12)

Jaccard metric
Jaccard metric is another metric to measure the overlap of segmented region and ground truth. Its difference with Dice metric is in taking difference of two regions into account. It is calculated as follows:

\[
J = \frac{|s_1 \cap s_2|}{|s_1 \cup s_2| - |s_1 \cap s_2|}
\]  (13)

The range of values that these metrics can take are summarized in Table 1.

<table>
<thead>
<tr>
<th>minimum</th>
<th>maximum</th>
<th>metric</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>Dice metric</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>Specificity</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>False positive rate (FPR)</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>Jaccard metric</td>
</tr>
</tbody>
</table>

All executions are done using Matlab 2014a. First set includes 8 pairs of simulated T1-weighted and T2-weighted images of brain with MS lesions with slice thickness of 1mm, 3mm, 5mm, and 7mm, 0 percent of noise and intensity inhomogeneity. Second dataset consists of 30 pairs of simulated T1-weighted and T2-weighted images of brain with MS lesions with slice thickness of different combinations of slice thickness, intensity inhomogeneity and noise levels. Two radiologists have segmented images manually and ground truth is formed based on results of manual segmentation. Registration is done on T2-weighted image. Skull stripping is done before execution of main algorithm. An example is presented in Figure 2.

**Figure 2.** Skull stripping (a) original image. (b) skull-stripped image.
Results and Discussion

The results obtained in the segmentation by applying the developed algorithm are exemplified in Figure 3 and 4. Each image is segmented into four classes: background, white matter (WM), gray matter (GM), and cerebrospinal fluid and MS plaques. Any member of forth class is considered as a lesion only if it is located in WM.

Figure 3. Segmentation results of first dataset (a) T1-weighted, (b) T2-weighted, (c) Ground truth, (d) Anbeek et al. [14], (e) Sarti et al. [43], (f) Bezdek et al. [44], (g) Demirhan and Guler [40], (h) Ortiz et al. [42], (i) Proposed method
Figure 4. Segmentation results of second dataset (a) T1-weighted, (b) T2-weighted, (c) Ground truth, (d) Anbeek et al. [14], (e) Sarti et al. [43], (f) Bezdek et al. [44], (g) Demirhan and Guler [40], (h) Ortiz et al. [42], (i) Proposed method.

Average values of all metrics and their standard deviations are depicted in Figure 5 and 6. As depicted in Figure 5(a), the proposed method is capable of finding the most overlap between ground truth and segmented image. This proves better performance of proposed method in comparison to other methods. As mentioned earlier, Specificity shows performance of method certainty in detecting points which do not belong to region of interest. Figure 5(b) shows the ability of proposed method in detecting these pixels better than other methods. The least amount of wrong pixels (FPR) is detected by proposed method. According to Jaccard metric, the proposed method performs the best in finding intersection with ground truth.
Figure 5. Average values of metrics over dataset in Figure 3

Figure 6. Average values of metrics for dataset in Figure 4

Another important issue here, is the disability of methods proposed by Ortiz and Demirhan, in finding the GM or our forth class. In spite of their acceptable performance in segmenting lesion, their general performance is not acceptable.

A disadvantage of FCM is its sensitivity to noise. A predominance of proposed method is its ability in compensating for noise effects. As shown in Figure 5 and 6, the proposed method has the best performance compared to other methods. Utilizing information of difference channels would increase accuracy and reliability of method even in presence of noise and intensity inhomogeneity. Average values yield the same results about performance of method.
Standard deviation is the measure of diversity in attained results. Less standard deviation means less change with data variations, thus higher reliability. Values of standard deviation for Dice metric and specificity are depicted in Figure 7. Superiority of method against data variations is also confirmed.

![Figure 7. Standard deviation of metrics (a) Dice metric and (b) Specificity](image)

**Conclusion**

Quantitative analysis indicates that our method has the best performance compared to contemporary state of art methods, with average improvement of 16.38% in Dice metric, 0.03% in specificity, 0.04% in FPR and 35.6% in Jaccard index.

**References**