An Brain Image Segmentation Method based on Non-local MRF

Zhongyuan Cui¹, Feng Wang¹ and Jin Wang²

¹School of Computer Science and Technology & School of Software, Zhoukou Normal University, Zhoukou 466001, China
²School of Computer and Software, Nanjing University of Information Science & Technology, Nanjing 210044, China

Abstract

Brain image segmentation is one of the most important parts of clinical diagnostic tools. However, accurate segmentation of brain images is a very difficult task due to the noise, inhomogeneity and sometimes deviation in brain images. Wells model incorporates the brain image segmentation and inhomogeneity correction into one framework to overcome influences from the intensity inhomogeneity and obtain good segmentation performance. However, the classical Wells model did not take spatial information into account, so its segmentation results are sensitive to the noise. In order to overcome this limitation, the MRF theory and the nonlocal information are used to construct a nonlocal Markov Random Field. With this nonlocal MRF, the improved Wells method can obtain much better segmentation results. The experimental results show that our method is robust to the noise and is able to simultaneously keep the image edge and slender topological structure very well.

Keywords: Brain image segmentation, Wells method, MRF, Nonlocal weight coefficient

1. Introduction

The application of image processing techniques has rapidly increased in recent years, especially in medical research fields. However, extracting the regions with abnormal tissue or shape from the whole data still depends on the manual segmentation by radiologists. The automatic image segmentation is a key task in many clinical applications such as surgical planning, tissue analysis, pre-surgical assessment, and abnormality detection and so on [1]. However, the automatic image segmentation is hard to accurately extract the regions due to unknown noise, poor image contrast, in homogeneity and weak boundaries that are usual in medical images. One of such is brain image segmentation which is quite complicated and challenging but its accurate result can be used to characterize neurological diseases, such as dementia, multiple sclerosis, schizophrenia and even the Alzheimer’s disease (AD) [2]. Therefore, it is meaningful for us to develop new effective brain image segmentation methods to detect the satisfactory tissue regions for clinical applications.

Magnetic resonance imaging (MRI) is a very useful medical imaging technique used in radiology to investigate the anatomy and function of the body in both health and disease. MRI scanners use strong magnetic fields and radio-waves to form images of the body. MRI acquisition parameters can be adjusted to give different grey levels for different tissues and various types of neuropathology [3-4]. In comparison to computerized tomography (CT), MRI does not use ionizing radiation, has a much greater range of available soft tissue contrast, and is more sensitive and specific for abnormalities. Due to these advantages, MRI is widely used for brain imaging in clinical diagnosis such as examining soft tissue, brain quantitative analysis and brain tumor inspection. However, it is hard to design an automated method only
based on the tissue contrast for brain MR image segmentation due to the corruption with a smoothly varying intensity in homogeneity or bias field. This bias is inherent to MR imaging and is caused by radio frequency coils or acquisition sequences and patient-induced electrodynamics interactions [5-6]. Although not always visible for a human observer, the distribution of signal intensities associated with these tissue classes is disturbed. Such a bias can cause serious misclassifications when intensity-based segmentation techniques are used [7-8].

In recent years, lots of in homogeneity or bias correction methods have been proposed [9-14]. In general, all of these methods can be categorized in two groups: prospective methods and retrospective methods [15]. The prospective methods consider bias field as an error of the imaging process that can be corrected by estimating bias field of MRI acquisition system. On the contrary, the retrospective methods don’t assume any information about acquisition methods and are more general. One classic model of such methods is segmentation based methods [15-18] which merge the tissue segmentation and bias field correction to benefit each other. The tissue classification results and bias field can be obtained simultaneously by using alternating iterative algorithm.

Wells, et al., described an iterative method that interleaves classification with bias field correction based on maximum likelihood parameter estimation using the expectation maximization (EM) algorithm [16]. However, the classical Wells model did not take into spatial information account, so its segmentation results are sensitive to the noise. MRF theory provides a convenient and consistent way to model context-dependent entities such as image pixels and correlated features. However, in traditional MRF models, the influence of each pixel in the neighborhood on the pixel classification is the same; this will over smooth the segmentation of brain tissue which contains lots of special structures such as corners and slender topology regions. In order to overcome this limitation, we apply the MRF theory and nonlocal method to import the spatial information of the image. Different with using the local intensity in the neighborhoods, we use the patch information to measure the pixels’ correlation in the whole neighborhoods and build an adaptive spatial constraint term. By incorporating both the spatial constraint term and Wells’ method, an accurate and robust segmentation method can be achieved.

2. Wells, et al., Method

The observed MRI signal $I$ is the product of the true signal $J$ generated by the underlying anatomy and spatially varying field factor $B$:

$$\tilde{I} = B \cdot J.$$  \hspace{1cm} (1)

Given the observed signal $\tilde{I}$, the problem is to estimate the true image $J$. A group of techniques often take the logarithmic transform of both sides:

$$\log(\tilde{I}) = \log(B) + \log(J).$$  \hspace{1cm} (2)

We set $\tilde{I} = \log(\tilde{I})$, $B = \log(B)$, and $J = \log(J)$ respectively, Eq. (2) can be written as $I = J + B$. After the preprocessing, the brain image only contains four regions: white matter (WM), gray matter (GM), cerebral spinal fluid (CSF), and background. Similar to other statistical approaches to intensity-based segmentation of MRI, Wells et al. modeled the distribution for observed values as a normal distribution with the incorporation of an explicit bias field. The probability density that class $j$ has generated the pixel value $I_i$ at position $i$ is:

$$p(I_i | \Gamma_i = j, B_i) = G_{\psi_i}(I_i - \mu(\Gamma_i = j) - B_i),$$  \hspace{1cm} (3)
where $\Gamma_i$ denotes the tissue class at position $i$, $\mu(\Gamma_i = j)$ denotes the mean value of class $j$. $G_{\Psi_{\Gamma_i}}$ is Gaussian distribution with variance $\Psi_{\Gamma_i}$:

$$G_{\Psi_{\Gamma_i}}(x) = (2\pi)^{-1/2} |\Psi_{\Gamma_i}|^{-1/2} \exp\left(-\frac{1}{2} x^T \Psi_{\Gamma_i}^{-1} x\right).$$ (4)

According to very smooth and slowly varying characteristic of the bias field $B$, Wells et al. modeled it by a zero mean Gaussian prior probability density with variance $\Psi_B$:

$$p(B) = G_{\Psi_B}(B),$$ (5)

where $G_{\Psi_B}(x) = (2\pi)^{-1/2} |\Psi_B|^{-1/2} \exp\left(-x^T \Psi_B^{-1} x / 2\right)$.

Assuming that the bias field and brain tissue classes are statistically independent and using the definition of conditional probability, the joint probability on intensity and tissue class conditioned on the bias field can be written as follows:

$$p(I_i, \Gamma_i | B_i) = p(I_i | \Gamma_i, B_i) p(\Gamma_i).$$ (6)

Then, the conditional probability of intensity alone can be obtained by calculating a marginal over tissue class:

$$p(I_i | B_i) = \sum_{\Gamma_i} p(I_i, \Gamma_i | B_i) = \sum_{\Gamma_i} p(I_i | \Gamma_i, B_i) p(\Gamma_i).$$ (7)

Assuming that the pixel intensities are statistically independent, the probability density for the whole image may then be written as:

$$p(I | B) = \prod_i p(I_i | B_i).$$ (8)

According to the Bayes’ rule, the posterior probability of the bias field given observed intensity data is able to be obtained as follows:

$$p(B | I) = \frac{p(I | B) p(B)}{p(I)}.$$ (9)

Then, the maximum-a-posteriori (MAP) principle was used to formulate the bias field estimation as the value of $B$ having the largest posterior probability:

$$B = \arg \max_B p(B | I).$$ (10)

The EM algorithm then interleaved bias field calculation, class-conditional distribution parameter estimation and a statistical classification of the image pixels into classes. Using EM algorithm based on the statistical probability framework, Wells et al. method is able to obtain the tissue class and correct the bias field simultaneously and is especially useful for images derived from surface coils, where the large intensity variations make it difficult to accommodate the image data on films for viewing. However, this method does not exploit the information about spatial connectedness of neighboring pixels belonging to the same class, which leads the methods sensitive to noise. When the image is seriously corrupted by the noise, the segmentation results and bias field estimated by this method are not accurate enough. In order to reduce the effect of the noise, the spatial information needs to be taken into account during the brain tissue segmentation.

In order to address this problem, we propose an improved method based on Markov random field and non-local mean theory. Firstly, the Markov random field theory is analyzed in next section. Then, under the MRF theory, we use the non-local mean theory to measure the similarity between the pixels in the neighborhood and build an adaptive spatial constraint term (we call it NLMRF). By incorporating both the spatial constraint term and Wells’ method, an accurate and robust segmentation method can be achieved, which is demonstrated in following section.
3. Our Method

3.1. MRF Theory

MRF theory provides a convenient and consistent way to model context-dependent entities such as image pixels and correlated features [17]. This is achieved by characterizing mutual influences among such entities using conditional MRF distributions. In an MRF, the label of one pixel is decided not only by itself but also the characteristics of other pixels in its neighborhood. According to MRF theory, setting \( N(i) \) as the set of sites neighboring \( i \), we have \( i \notin N(i) \) and \( i \in N(j) \iff j \in N(i) \). According to Hammersley-Clifford theorem, an MRF can equivalently be characterized by a Gibbs distribution. Thus

\[
P(x) = \frac{1}{Z} \exp(-\sum_{C} V_c(x)),
\]

where \( C \) is a clique, and \( V_c(x) \) is the clique potential over \( C \). In this paper, we only consider the clique potential of two pixels:

\[
V_c(i, j) = \beta |1 - \delta(x, y)|.
\]

\( \delta \) is Dirac function, \( Z = \sum_{x \in F} \sum_{C} \exp(-\sum_{c} V_c(x)) \) is a normalizing constant called the partition function, and \( \beta \) is a constant.

3.2. Nonlocal-MRF Wells’ Method

With MRF theory, a prior of pixel \( i \) belongs to class \( j \) is decided by the classes of all pixels in its neighborhood. Thus,

\[
p(\Gamma_i = j) = \frac{1}{Z} \exp(-\sum_{\forall k \in N(i)} \beta |1 - \delta(\Gamma_i, \Gamma_k)|).
\]

From above equation, we can see that the traditional MRF theory uses all of pixels in the neighborhood to calculate the prior, that means the MRF model may be spatially dependent and is more flexible for image modeling in the sense that integrating the statistical and spatial properties. However, in traditional MRF models, the influence of each pixel in the neighborhood on the pixel classification is the same. This is not suit for segmenting the brain tissue which contains lots of special structures such as corners and slender topology regions. Besides this, it is also very hard to decide the size of the neighborhood to make it is adaptive to the noise and the brain structures. The effects on suppressing the noise is not good enough if we chose a small size, on the contrary, the edge or special structure may be over smoothed if the size is too large. Aiming to solve these problems, we add a similarity weight in Eq. (12) to improve the prior probability:

\[
p(\Gamma_i = j) = \frac{1}{Z} \exp(-\sum_{\forall k \in N(i)} NL_{ik} \beta |1 - \delta(\Gamma_i, \Gamma_k)|).
\]

where \( NL_{ik} \) is the structure similarity weight of pixel \( i \) and \( k \). We can set \( NL_{ik} \) a big value when the structures of these two pixels are very similar and a small value in the opposite case. This weight is very well to adjust the influences of neighbor pixels on the classification. Inspired by non-local mean filter method [19-20], we use the patch to calculate the similarity weight in this paper. The non-local (NL) regularization is a strategy that has been proposed first as a denoising tool and named as NL Mean denoising. Essentially, it aims to take advantage of the redundancy present in natural structures: broadly speaking a small patch around a pixel may match patches around other pixels within the same scene. The weight \( NL_{ik} \) for the pixel \( i \) and \( j \) is defined as:
\[ NL_{ik} = \frac{1}{Z_i} \exp\left(-\|I(N(i)) - I(N(k))\|^2/h^2\right), \]  

(14)

where \( N_i \) is the neighborhood of pixel \( i \), \( Z_i \) is a normalization constant, and \( h \) is a smoothing parameter. The vector \( I(N(i)) \) contains the grey-level profile in the neighborhood of the pixel \( i \) and \( I(N(k)) \) is the \( k \)-th component of this vector.

Compared to the Wells method, we use MRF to calculate the prior probability in this paper, so that the label of one pixel is decided not only by itself but also the characteristics of other pixels in its neighborhood. Beside this advantage, each point of the extended neighborhood in our method does not necessarily have the same influence, based on a specific non-local weight modeling the similarity between image patches. In particular, this weight penalizes the points \( k \) which are surrounded by a patch less similar to the patch around the current point \( i \) (\( NL_{ik} \) is low), allowing the points surrounded by the same kind of patches as \( k \) to have greater influence. Also, with this adaptive similarity weight, we do not need to adjust the size of the neighborhood to make it suit for different brain structures.

**3.3. NLMRF-Wells Algorithm**

In this paper, we firstly use the non-local weight to calculate the structure similarity and improve the traditional MRF approaches. Then, the non-local prior probability is applied to improve the Wells method for brain MR image segmentation (we can call it NLMRF-Wells method), the final method using non-local MRF and Wells method can be summarized as follows:

- **Step 1:** Taking the logarithmic transform of the bias corrupted image and using MAP method to segment the brain image as preprocessing step for providing the initial parameters;
- **Step 2:** Estimating the bias field:
  \[ R = H\bar{R}, \]  
  (15)

where, \( H = [\psi_{ij}^{-1} - \psi_{ij}^{-1}] \), \( \psi \) is the mean variance of each label and \( \bar{R} = \sum_j W_j[\psi_{ij}^{-1}(I_i - \mu_j)]. \)

- **Step 3:** Calculating the structure similarity weight of each pixel using Eq. (14);
- **Step 4:** Using Eq. (13) to calculate the prior probability of each pixel belonging to each label;
- **Step 5:** Calculating the probability of each pixel belonging to each label:
  \[ W_{is} p(I_i = j)G_{is}(I_i - \mu_j - B_j)/\sum_j p(I_i = s)G_{is}(I_i - \mu_s - B_s); \]  
  (16)

- **Step 6:** Calculating the parameters in Gaussian distribution:
  \[ \mu(j) = \sum_i W_{ij}(I_i - B_j)/\sum_i W_{ij} \]  
  (17)
  \[ \psi_j = \sum_i W_{ij}(I_i - \mu_j)^2/\sum_i W_{ij} \]  
  (18)

- **Step 7:** Go to Step 2 if the result is not convergence, otherwise stop the iteration.

**4. Implementation and Results**

Our method was applied in MATLAB on a PC with Inter Pentium IV 2.4G processor and 2048 RAM. In order to test the reliability and validity of the algorithm, we describe the segmentation on both synthetic images and brain MRIs.
4.1. Evaluation with Synthetic Data

In order to show the advantage of non-local similarity weight, we use the MAP, MRF-MAP and NLMRF-MAP methods to test a synthetic image which size is $128 \times 128$. This synthetic image is seriously corrupted by the noise and also contains some special structures. The segmentation result of the traditional MAP method is shown in Figure 1(b). Without any spatial information, this method is very sensitive to the noise. MRF-MAP method extended the MAP method by using local neighbor information to reduce the effect of the noise. However, this method is an isotropic method, which makes it unable to overcome all the noises. Furthermore, the special structures such as corner or slender region in the image are not segmented out very well. Compared with these two methods, the NLMRF-MAP method using non-local weight to improve the MRF can overcome the noise influences and segment the special structures very well. Therefore, our method can be considered as a superior approach.

![Figure 1. Segmentation Results on a Simulated Image: (a) Initial Image; (b) Segmentation Results of MAP Method; (c) Segmentation Results of MRF-MAP Method; (d) Segmentation Results of NLMRF-MAP Method](image)

4.2. Evaluation with Brain Image

In order to compare the presentations of different segmentation methods, we use the brain images from open database generated by the McConnell Brain Imaging Center at the Montreal Neurological Institute, McGill University. This database contains lots of normal human brain images with different noise and bias field. The corresponding standard segmentation results are also given in this database for validation. The experiment can download the data as they want and apply their method to obtain the segmentation results and verify their results by using the standard segmentation. For comparative purpose, we use 0, 90, 160, 240 to denote the intensities of background, CSF, GM and WM respectively in our segmentation results.

The first line of Figure 2 demonstrates the brain MR image with 100% bias field and 0%, 3% and 5% random noise. The images from the second line to the last line demonstrate the corresponding results of MAP, Wells, MRF-Wells and NLMRF-Wells methods respectively. From the results of the MAP method, we can see that without considering the effect of the bias field, this method misclassified some WM tissues into GM. Furthermore, it is sensitive to the noise without any spatial information. Compared to the MAP method, the Wells’ method reduced the effect of bias field, however, only using the intensity distribution information during tissue segmentation; this method is sensitive to the noise too. The MRF-Wells method extended the Wells method by using local neighbor information to reduce the effect of the noise and bias field. However, this method is an isotropic method, which make it unable to...
reduce the effect of strong noise and usually loses some special structure information of brain tissues. Different from traditional MRF, NLMRF-Wells method uses non-local patch to calculate the weight, the neighborhoods of pixels belong to the same tissue will have higher weight and have higher probability that have a strong influence on the classification of current pixel, which make our method much less sensitive to the noise.

Figure 2. Comparison of Segmentation Results using MAP, Wells, MRF-Wells, NLMRF-Wells Methods

In order to demonstrate the adaptive ability of our method, we use different sizes of neighborhood to test top-left image with 100% intensity in homogeneity and 5% random noise in Figure 3. The images on the first line and second line demonstrate the segmentation results of MRF-Wells and NLMRF-Wells method by using neighborhood with the size 3×3, 5×5, 7×7, 9×9. From the results, we can see that the size of neighborhood has a serious impact to the segmentation when using MRF-Wells method. The gray matter and CSF regions contain lots of slender structures, so that the edges of these tissues are over smoothed when neighbor size increase, which causes some tissue regions misclassification. Compared to MRF-Wells method, applying the patch information to calculate the weight of each pixel in the neighborhood, the neighborhoods of pixels belong to the same tissue will have higher weight and have higher probability that have a strong influence on the classification of current pixel, which make our NLMRF-Wells method contains not only spatial information but also the structure similarity. This advantage demonstrates that our method is self-adaptive, and a superior approach.
4.3. Quantitative Analysis

For comparison, we use Jaccard similarity (JS) coefficient as a metric to evaluate the performance of these methods. The Jaccard coefficient measures similarity between finite sample sets, and is defined as the size of the intersection divided by the size of the union of the sample sets: \( J(S_1, S_2) = \frac{|S_1 \cap S_2|}{|S_1 \cup S_2|} \), where \( S_1 \) and \( S_2 \) demonstrate the segmentation result using each segmentation method and groundtruth respectively. Higher JS value denotes more accurate segmentation results and lower JS value denotes worse segmentation.

<table>
<thead>
<tr>
<th>Brain tissue</th>
<th>method</th>
<th>0% noise</th>
<th>3% noise</th>
<th>5% noise</th>
</tr>
</thead>
<tbody>
<tr>
<td>White Matter</td>
<td>MAP method</td>
<td>0.55</td>
<td>0.51</td>
<td>0.42</td>
</tr>
<tr>
<td></td>
<td>Wells method</td>
<td>0.90</td>
<td>0.81</td>
<td>0.78</td>
</tr>
<tr>
<td></td>
<td>MRF-Wells method</td>
<td>0.89</td>
<td>0.85</td>
<td>0.82</td>
</tr>
<tr>
<td></td>
<td>NLMRF-Wells method</td>
<td>0.90</td>
<td>0.89</td>
<td>0.89</td>
</tr>
<tr>
<td>Gray Matter</td>
<td>MAP method</td>
<td>0.52</td>
<td>0.47</td>
<td>0.41</td>
</tr>
<tr>
<td></td>
<td>Wells method</td>
<td>0.87</td>
<td>0.76</td>
<td>0.74</td>
</tr>
<tr>
<td></td>
<td>MRF-Wells method</td>
<td>0.86</td>
<td>0.82</td>
<td>0.81</td>
</tr>
<tr>
<td></td>
<td>NLMRF-Wells method</td>
<td>0.88</td>
<td>0.87</td>
<td>0.87</td>
</tr>
</tbody>
</table>

In this paper, we use MAP, Wells, MRF-Wells, NLMRF-Wells methods to test 20 images with 100% intensity in homogeneity and 0%, 3%, 5% random noise in McGill database respectively. The performances of these methods are shown in Table 1. We can see that without considering intensity in homogeneity affection the MAP method cannot obtain correct results. The results of Wells, MRF-Wells, and NLMRF-Wells methods are much the same when the images are not corrupted by the noise. However, with the increase of the noise strength, the performance of Wells method dropped rapidly, this because the Wells method...
did not consider any spatial information during the tissue segmentation. Using spatial constraint, the results of MRF-Wells and NLMRF-Wells methods are both robust to the noise. Furthermore, considering the structure similarity between the pixels in the neighborhood, the performance of our NLMRF-Wells method is much better than that of MRF-Wells method. Also, our method is more stable and reliable to different noise intensity. This comparison demonstrates that our method has good performance and strong stability for brain tissue segmentation.

5. Conclusions

In this paper, we propose an improved method based on Wells method for simultaneous estimation of the bias field and segmentation of tissues in MRIs. The MRF theory and the nonlocal information are used to construct a nonlocal Markov Random Field. With this nonlocal MRF, the improved Wells method is robust and stable enough to the noise and is able to obtain much better segmentation results. The experimental results show that our method is robust to the noise and can solve the image edge and slender topological structure very well when segmenting brain images with intensity in homogeneities and noise.

Acknowledgments

This paper is a revised and expanded version of a paper entitled “Improved Wells Method using MRF Radom Field and Non-local Weight” presented at Soft Tech 2014, Yeosu, Korea, May 8-10, 2014. This work was supported by the Industrial Strategic Technology Development Program (10041740) funded by the Ministry of Trade, Industry and Energy (MOTIE) Korea, the National Natural Science Foundation of China (61402234) and the Natural Science Foundation of Jiangsu Province (No. BK2012461). It was also supported by China Meteorological Administration under CMA grants (2014 MC 16, M43).

References


Authors

Zhongyuan Cui, lecturer of the school of computer science and technology in Zhoukou Normal University, his main research interests include image processing, pattern recognition and numerical analysis.

Feng Wang, lecturer of the school of computer science and technology in Zhoukou Normal University, his main research interests include pattern analysis, computer vision and artificial intelligence.