Received September 17, 2019, accepted October 18, 2019, date of publication October 23, 2019, date of current version January 7, 2020.

Temporally Consistent Segmentation of Brain Tissue From Longitudinal MR Data

PENGCHENG LI1,2, YUE ZHAO1,2, YANG LI3,4,5, QIAOYI CHEN1,2, FANGCEN LIU1,2, AND CHENQIANG GAO1,2
1School of Communication and Information Engineering, Chongqing University of Posts and Telecommunications, Chongqing 400065, China
2Chongqing Key Laboratory of Signal and Information Processing, Chongqing 400065, China
3Orthodontic department, Stomatological Hospital of Chongqing Medical University, Chongqing 401147, China
4Chongqing Key Laboratory of Oral Diseases and Biomedical Sciences, Chongqing 401147, China
5Chongqing Municipal Key Laboratory of Oral Biomedical Engineering of Higher Education, Chongqing 401147, China

Corresponding author: Yang Liu (yangliu@hospital.cqmu.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61571071 and Grant 61906025, in part by the Chongqing Research Program of Basic Research and Frontier Technology under Grant cstc2018jcyjAX0227, in part by the Science and Technology Research Program of Chongqing Municipal Education Commission under Grant KJQN201900607, in part by the Education Informatization Project of Chongqing University of Posts and Telecommunications under Grant xxyf2019-01, and in part by the Research Fundamental of Chongqing University of Posts and Telecommunications under Grant A2017-118.

ABSTRACT In this paper, we propose a new 4D segmentation formulation, aiming to improve the temporal consistency of adults’ brain tissue segmentation. In our method, tissue segmentation result is represented by the membership functions of the tissues, which are derived as a result of minimizing energy in a fuzzy C-means (FCM) framework. We first introduce a variational formulation with a temporal consistency constraint on the membership functions, then convert the constraint to a vector-valued function from which the membership functions are directly computed according to an analytically defined mapping. These vector-valued functions capture the bias field and intensity means of each tissue addressing the temporal variation in intensity inhomogeneities and the intensity means of the images. The effectiveness is demonstrated on the Baltimore Longitudinal Study of Aging (BLSA) benchmark dataset. Our method takes few parameters which are easy to tune and achieves 96.4% of TC factor for gray matter, 98.1% of TC factor for white matter, significantly superior to the compared methods on this research line.

INDEX TERMS Temporal consistency, brain tissue segmentation, longitudinal MR data.

I. INTRODUCTION Segmentation of a series of 3D data of the same subject captured at different times is a fundamental and an important task in many neuroimaging studies [1]–[5] that concentrate on normal development and aging, as well as on the evolution of pathology. The temporal consistency of the segmentation is especially important for studies analyzing subtle brain changes over time, such as in the case of normal aging and Alzheimer’s disease (AD). The task is considered difficult on MR scans as the intensity characteristics of brain tissue and pathology can greatly vary between scans of the same patient. Thus, the inconsistencies in segmentations generated from applying 3D approaches [7], [23]–[25] to a 4D data set can hide the subtle changes caused by the disease. There are various methods proposed for addressing the above issues in the past years. In [10], Xue et al. propose a method that overcomes the limitation of 3D segmentation and improves the temporal consistency of segmentation by formulating the segmentation problem in 4D. Their method is called Consistent Longitudinal Alignment and Segmentation for Serial Image Computing (CLASSIC). CLASSIC not only jointly segments a series of longitudinal 3D MR brain images of the same individual, but also estimates the longitudinal deformations in the image series, e.g. tissue atrophy. It iteratively performs two steps: (i) it jointly segments a series of 3D images using a 4D image-adaptive clustering algorithm based on the current estimate of the longitudinal deformations in the image series, and (ii) it then refines these longitudinal deformations using a 4D elastic warping algorithm. While CLASSIC generally outperforms 3D segmentation algorithms in terms of temporal consistency, it is sensitive to initialization. CLASSIC assumes that the input images are already corrected to homogeneity and have similar intensity patterns across scans. Generating those kinds of MR images is generally difficult, which negatively
impacts the robustness of the approach. Besides, CLASSIC requires the tuning of several parameters, which makes it less practical for the clinical setting. Furthermore, a longitudinal processing pipeline was proposed in FreeSurfer [11]. In this pipeline, a group-mean image is firstly generated by averaging from the rigidly-aligned longitudinal images of a subject. The cortical surfaces of the group-mean image are then used as initialization for each longitudinal image. Finally, the cortical surfaces at each time point are separately deformed to achieve longitudinal cortical surface reconstruction. The measurement of cortical thickness is of great interest in studying both normal brain development aging and a wide variety of neurodegenerative and psychiatric disorders. Neuroscience studies have suggested that various diseases such as AIDS or AD may affect cortical thickness [12].

In [13], Wang et al. propose a 4D segmentation method for longitudinal brain MR images of adult subjects with consistent longitudinal measurement of cortical thickness. In [14], Feng et al. propose an automated method for the segmentation of brain tissues in longitudinal MR images. The method first separately segmented brain tissues into white matter, gray matter, and cerebrospinal fluid by bias correction embedded fuzzy c-means. After being normalized in interclass, the similarities are incorporated into a non-local means denoising formula to regularize the segmentation in both spatial and temporal dimensions. Non-locally regularization results are used to compute final membership functions for the segmentation. It accelerates the modified denoising algorithm using Compute Unified Device Architecture (CUDA) and obtains accelerations of more than two orders of magnitude. It demonstrates the advantages of the proposed method in terms of segmentation accuracy and the ability to consistently segment brain tissues in an arbitrary number of longitudinal brain MR image series. Recently, some other automatic machine learning and deep learning based methods are proposed for multimedia analysis [15]–[18]. These methods are adaptive and easy to fit the medical image analysis community.

To address the above challenge issues and improve the temporal consistency of tissue segmentation, we propose a new 4D segmentation formulation. In our method, tissue segmentation is represented by the membership functions of the tissues, which derived from the minimization of the energy in a fuzzy C-means (FCM) framework. Adding both the global and local spatial information into the membership function is to decrease the sensitivity problem to the noise and intensity inhomogeneity in MR data [28]–[30]. There are very few parameters in our method, and it is easy to tune the parameter to achieve desirable results. As a result, our method does not require preprocessing for intensity inhomogeneity correction and intensity normalization of the series of images. Experiments implemented on a benchmark dataset show that our method is capable of achieving 96.4% of TC factor for gray matter, 98.1% of TC factor for white matter and lower label reverse rate, which is significantly robust, efficient, and stable compared with the state-of-the-art methods.

The remainder of the paper is organized as follows. Section II provides a detailed description of a commonly 3D segmentation model and our 3D segmentation method for a single MR image. Section III provides a detailed description of our 4D segmentation method for the images in a longitudinal series. Section IV reports the experimental results. Several concluding remarks are finally given in Section V.

II. 3D SEGMENTATION

In this section, we first describe a commonly used model of MR images and our method for 3D segmentation of a single MR image, which will be then extended to 4D segmentation by introducing a regularization term for temporal consistency in Section III.

A. MEMBERSHIP FUNCTION OF IMAGE MODEL

According to a commonly used model of MR images [6]–[8], the intensity of an observed MR image at a location $x$ can be modeled as:

$$ I(x) = B(x)J(x) + N(x), $$

where $I$ is the observed image, $J$ is the true image, and $B$ is the bias field that accounts for the intensity inhomogeneity in the observed image $I$, and $N$ is white noise. Ideally, the intensities of the true image $J$ in each tissue take a specific value of the physical property being measured respectively (e.g. the proton density). Therefore, the true image $J$ is approximately a constant $c_i$ in the $i$-th tissue, which is located in a sub-region $\Omega_i$ in $\Omega$. Thus, the true image $J$ can be approximated by:

$$ J(x) = \sum_{i=1}^{N} c_i u_i(x), $$

where $c_i$ is a constant that approximates the voxel intensities from the $i$-th tissue, and $u_i$ is the membership function of the tissue/region $\Omega_i$. We call the constants $c_i$ the true intensity mean (TIM) of the voxels in the $i$-th tissue. Typically, we assume there are three types of tissues in the brain: white matter (WM), gray matter (GM), and cerebrospinal fluid (CSF).

In this paper, we represent the membership functions $u_1(x), \cdots, u_N(x)$ with a vector-valued function $U$, $(u_1, \cdots, u_N)^T$, and the TIMs $c_1, \cdots, c_N$ with a column vector $c$, $(c_1, \cdots, c_N)^T$, where $(\cdot)^T$ is the transpose operator. For convenience, the vector $U$ is referred to as a membership vector, and the vector $c$ is referred to as a TIM vector. The membership function $U$ is a mapping $U : \Omega \rightarrow U$, where $U$ is defined by:

$$ U = \left\{ U = (u_1, \ldots, u_N)^T \mid \sum_{i=1}^{N} u_i = 1, 0 \leq u_i \leq 1 \right\}, $$

which is a convex set.

B. SEGMENTATION OF A SINGLE IMAGE

Our formulation of 4D segmentation is based on a new method for joint segmentation and bias field estimation for
a single 3D MR image described as follows. In our method, the bias field \( B \) is approximated by a linear combination of a set of basis functions \( g_1, \ldots, g_M \):

\[
B(x) = \sum_{j=1}^{M} w_j g_j(x). \tag{4}
\]

In this paper, we use orthogonal polynomials as the basis functions, and those basis function satisfy:

\[
\int_{\Omega} g_i(x) g_j(x) dx = \delta_{ij}, \tag{5}
\]

where \( \delta_{ij} = 0 \) for \( i \neq j \) and \( \delta_{ij} = 1 \) for \( i = j \).

By representing the basis functions \( g_1, \ldots, g_M \) as a column vector \( G = (g_1, \ldots, g_M)^T \) and the coefficients \( w_1, \ldots, w_M \) as column vector \( W = (w_1, \ldots, w_M)^T \), we can express the bias field as \( B(x) = W^T G(x) \).

Our tissue segmentation and bias field estimation is based on the adaptive fuzzy C-means (AFCM) framework proposed by Pham and Prince [8], in which the following energy was proposed in the form:

\[
\mathcal{F} = \int \sum_{i=1}^{N} (I(x) - B(x)c_i)^2 u_i^q(x) dx + \lambda_1 P_1(B) + \lambda_2 P_2(B), \tag{6}
\]

where \( q > 1 \) is the fuzzifier, \( P_1(B) \) and \( P_2(B) \) are the smoothing term on the bias field \( B \), which are introduced to penalize the \( L_2 \) norms of the first and second order derivatives of \( B \). Note that we use a continuous version of the AFCM energy here, instead of the discrete version in [8].

In our method, since bias fields are approximated by a linear combination of a set of smooth basis functions, there is no need to use the smoothing term as in the above AFCM framework, i.e. the penalty terms \( \lambda_1 P_1(B) + \lambda_2 P_2(B) \) can be removed in our method. Therefore, we propose to minimize the energy:

\[
\mathcal{F} = \int \sum_{i=1}^{N} (I(x) - W^T G(x)c_i)^2 u_i^q(x) dx, \tag{7}
\]

subject to the constraint: \( (u_1(x), \ldots, u_N(x))^T \in U \), for every \( x \in \Omega \).

The above energy can be minimized by an iterative process with respect to the variables \( U, c, \) and \( W \), respectively. The minimization with respect to \( c \) and \( W \) will be described in Section III in the context of 4D segmentation. In this section, we give the solution to the optimal membership function \( U \) that minimizes the energy \( \mathcal{F} \) for a single MR image.

The energy \( \mathcal{F} \) can be expressed as:

\[
\mathcal{F}(U, R) = \int \sum_{i=1}^{N} r_i(x) u_i^q(x) dx, \tag{8}
\]

where \( R = (r_1, \ldots, r_N)^T \) is a vector with components defined by:

\[
r_i(x) = (I(x) - W^T G(x)c_i)^2, \quad i = 1, \ldots, N. \tag{9}
\]

It can be shown that the optimal membership function \( U(x) = (u_1(x), \ldots, u_N(x))^T \) that minimizes the energy \( \mathcal{F}(U, R) \) in Eq. (8) is given by:

\[
u_i = \frac{r_i}{\sum_{j=1}^{N} r_j}, \quad i = 1, \ldots, N. \tag{10}
\]

This equation defines a mapping \( \phi : R \rightarrow U \). The membership vector \( U \) is determined by the vector \( R \) through the mapping \( \phi \).

The mapping \( \phi \) is scalar invariant, i.e., for any scalar constant \( s > 0 \), it holds that \( \phi(sR) = \phi(R) \). With this property, we can change the length of the vector \( R \) by multiplying it by a scalar constant, so that the point \( sR \) is located in a given surface \( \mathbb{R} \). In this paper, we use the surface \( \mathbb{R} \) defined by:

\[
\mathbb{R} = \left\{ R = (r_1, \ldots, r_N)^T \mid \sum_{i=1}^{N} r_i^\alpha = 1 \right\}, \tag{11}
\]

with a constant \( \alpha > 0 \). This property is meaningful in the segmentation of longitudinal data. For a location \( x \), the vectors \( R(x) \) computed from the images captured at different time points can be far away from each other, which does not allow us to perform temporal regularization on the sequence of such vectors. However, we can always multiply the vector \( R \) by a scalar \( s \) such that \( sR \in \mathbb{R} \), which corresponds to the same membership vector, i.e. \( U = \phi(sR) = \phi(R) \). Thus, the vectors \( R \) computed for different time points are “pulled” to the same surface \( \mathbb{R} \), such that they are comparable and therefore allow the application of temporal regularization to the vectors \( R \) computed for different time points. The mapping \( \phi \) and its properties will be further discussed and used in the next section for 4D segmentation.

III. 4D SEGMENTATION

In this section, we consider a series of 3D MR images captured from the same subject. The temporal consistency of segmentation of a series of 3D MR images can be improved by considering the segmentation problem in an appropriate 4D formulation. Although a series of 3D images can be stacked as a 4D data, we cannot simply extend a 3D segmentation formulation to 4D by equally treating spatial and temporal dimension in the formulation of 4D segmentation, due to the change of the image properties (e.g. histogram, contrast and intensity inhomogeneity) over time. Therefore, the extension of a 3D segmentation method to 4D is not a trivial task. The difficulty can be seen clearly from the model of serial MR images described next, which must be taken into account in the formulation of 4D segmentation.

A. MODEL OF SERIAL MR IMAGE

According to a commonly accepted model in Eq. (1), a series of observed MR images \( I_t(x) \) can be modeled as:

\[
I_t(x) = B_t(x)J_t(x) + N_t(x), \tag{12}
\]
where \( t \) in the subscript represents the scanning time, \( J_t(x) \) is the unknown true image, and \( B_t(x) \) is the unknown bias field, and \( N_t(x) \) is noise. The true image \( J_t(x) \) can be approximated by:

\[
J_t(x) = \sum_{i=1}^{N} c_i^{(t)} u_i^{(t)}(x),
\]

(13)

where \( c_i^{(t)} \) is the true intensity mean (TIMs) of the voxels in the \( i \)-th tissue, and \( u_i^{(t)}(x) \) is the membership function of the \( i \)-th tissue at time point \( t \). Since this series of images are captured from the same subject, there is a certain temporal consistency of the membership functions \( u_i^{(t)} \) at different time points, which will be taken into account in our formulation of 4D segmentation.

Note that the bias fields and the TIMs of the images at different time points are not related to each other. They may vary over time, and thereby lead to the variability of the images in a longitudinal series, which renders it difficult for a 3D method to achieve temporally consistent segmentation.

### B. FORMULATION OF 4D SEGMENTATION

Our 4D segmentation method is formulated on a series of MR images \( I_t(x) \) with \( t \) denoting the scanning time. We assume that this series of images is registered to the first image in the series. From \( I_t(x) \), we need to find the membership functions \( u_i^{(t)}(x), \ldots, u_N^{(t)}(x) \) as the segmentation result for each time point \( t \), which should satisfy a certain temporal consistency property. \( I_t(x) \) can be represented as a 4D data \( I(x,t) \). The series of estimated bias fields \( B_t(x) \) and the membership functions can be written as \( B_t(x) = B_t(x) \) and the membership function \( U(x,t) = (u_1^{(t)}(x), \ldots, u_N^{(t)}(x))^T \). The bias field \( B_t(x,t) \) is given by \( B_t(x,t) = W_t^T(t)G(x) \) with \( W(t) = (w_1(t), \ldots, w_M(t))^T \).

The TIMs, \( c_1^{(t)}, \ldots, c_N^{(t)} \) can be represented as \( c_t = (c_1^{(t)}, \ldots, c_N^{(t)})^T \).

Our 4D segmentation goal is to obtain an optimal membership function \( U(x,t) \) with temporal consistency in some sense. We use an energy minimization approach to find the optimal membership function \( U(x,t) \). The vectors \( W(t) \) and \( c(t) \) at all time point \( t \) are jointly computed by minimizing energy in our energy minimization framework.

Given a series of images \( I(\cdot,t) \), we define an energy:

\[
\mathcal{G}(U, c, W) = \int_{L_0}^{L} \int_{\Omega} \sum_{i=1}^{N} (I(x,t) - W_t^T(t)G(x)c_i(t))^2 u_i^q(x,t) dx dt,
\]

(14)
in which \( \mathcal{G} \) is an extension of the energy \( \mathcal{F} \) is defined in Eq. (7).

Let \( R(x,t) \) be the vector \( R(x,t) = (r_1(x,t), \ldots, r_N(x,t))^T \) with \( r_i(x,t) = (I(x,t) - W_t^T(t)G(x)c_i(t))^2 \). Then, by writing \( y = (x,t) \), the functions \( r_i(x,t) \) can be expressed as \( r_i(y) = r_i(y) \). Thus, given the vector \( R(x,t) \), the above energy \( \mathcal{G} \) can be simplified as:

\[
\mathcal{G}(U, R) = \int_{\Omega \times [0,L]} \sum_{i=1}^{N} r_i(y)u_i^q(y)dy,
\]

(15)

which is in a 4D formulation. For the computation of the membership function \( U \) for given vectors \( c(t) \) and \( W(t) \), we can directly work with the energy \( \mathcal{F} \) in the form of Eq. (15). To compute the vectors \( c(t) \) and \( W(t) \), we use the energy \( \mathcal{G} \) in the form of Eq. (14).

The 4D formulation energy in Eq. (15) still does not take into account the temporal consistency. We will add a temporal regularization term \( \mathcal{P} \) to define an energy \( \mathcal{E} \) by:

\[
\mathcal{E} = \mathcal{G} + \lambda \mathcal{P},
\]

(16)

with \( \lambda > 0 \) as the weight of the temporal regularization term. According to the model of serial MR images in Section III-A, the vectors \( c(t) \) and \( W(t) \) do not have temporal consistency. Therefore, the computation of the vectors \( c(t) \) and \( W(t) \) is not subject to constraint by the temporal regularization term \( \mathcal{P} \). Thus, we can first describe the minimization of the energy \( \mathcal{E} \) with respect to the vectors \( \mathbf{W} \) and \( c \), which is equivalent to the minimization of the energy \( \mathcal{G} \) with respect to \( \mathbf{W} \) and \( c \).

1) OPTIMIZATION OF \( \mathbf{W} \) AND \( c \)

Minimization of the energy \( \mathcal{G} \) with respect to the variable \( \mathbf{W} \) can be achieved by solving the equation \( \partial \mathcal{G}/(\partial \mathbf{W}(t)) = 0 \), which yields the following solution:

\[
W(t) = \left( \int_{\Omega} G(x)G^T(x) \sum_{i=1}^{N} c_i^2 u_i^q(x,t) dx \right)^{-1} \times \left( \int_{\Omega} G(x)I(x,t) \sum_{i=1}^{N} c_i^2 u_i^q(x,t) dx \right)
\]

(17)

thus, the bias field is estimated by:

\[
B(x,t) = B_t(x) = W_t^T(t)G(x).
\]

(18)

Minimization of the energy \( \mathcal{G}(U, c, \mathbf{W}) \) with respect to \( c \) yields an optimal TIMs \( c(t) = (c_1^{(t)}, \ldots, c_N^{(t)})^T \) given by:

\[
c_i(t) = \frac{\int_{\Omega} I(x,t)B_t(x,t)u_i^q(x,t) dx}{\int_{\Omega} B_t^2(x,t)u_i^q(x,t) dx}, \quad i = 1, \ldots, N.
\]

(19)

2) OPTIMIZATION OF 4D MEMBERSHIP FUNCTION

As same in the 3D formulation, we define:

\[
r_i(x,t) = (I(x,t) - W_t^T(t)G(x)c_i(t))^2, \quad i = 1, \ldots, N.
\]

(20)

which are represented by the function \( R(x,t) = (r_1(x,t), \ldots, r_N(x,t))^T \). By writing \( y = (x,t) \), the function \( r_i(x,t) \) can be further expressed as \( R(y) = R(x,t) \). From this 4D function \( R(y) \), we will compute the 4D membership function \( U(y) = U(x,t) \).

As described in Section II-B, the function \( R(x,t) = R(x,t) \) at each time point \( t \) determines the membership function \( U \) given by Eq. (10) as a segmentation result for image
It is clear that the energy is a convex function in each of the variables. The minimization of $G$ is solved in an iterative process of interleaved minimization with respect to each variable. The whole process can be described by the following steps:

Step 1. Initialize the membership functions $U$ by random matrices where each entry follows a uniform distribution in $[0, 1]$, and initialize $c$.

Step 2. Update bias field $B$ by Eq.(17) to Eq.(18).

Step 3. Update $c$ by Eq.(19).

Step 4. Update membership function $U$ by Eq.(20).

Step 5. Check convergence criterion $|c^n - c^{n-1}| < \epsilon$. If convergence has been reached or the iteration number exceeds a prescribed maximum number, stop the iteration, otherwise, go to Step 2.

In the above iteration process, each of the variables is updated with the other variables computed in the previous iteration, and we set the $\epsilon$ in Step 5 to 0.001, and $n$ denotes the $n$-th iteration.

C. TEMPORAL REGULARIZATION OF U AND R

The Eq. (10) defines a mapping from the space of $R$ to the space of membership vector $U$. Denoting the Jacobian matrix of $\phi$ as $J_\phi$, it can be inferred that the membership function $U(x, t)$ that minimizes the energy $G(U, R)$ is given by $U(x, t) = \phi(R(x, t))$. Therefore, we have the following relationship:

$$\frac{\partial U(x, t)}{\partial t} = J_\phi(R(x, t)) \frac{\partial R(x, t)}{\partial t}.$$ (25)

It then follows that:

$$P(U) \leq \beta P(R),$$ (26)

where $\beta = (\sup_{R} |J(R)|)^k$ and $\sup$ means supremum. This inequality shows that the regularization of $U$ can be achieved by regularization on $R$.

IV. EXPERIMENTAL RESULTS

A. MR DATASET DESCRIPTION

Our method has been tested on the Baltimore Longitudinal Study of Aging (BLSA) data [5]. The BLSA study is a longitudinal study of aging and early markers of Alzheimer’s Disease. Researchers measure physical and cognitive changes associated with aging in real-time among a dedicated group of BLSA participants who come in for testing at regular intervals throughout their lives.

In this experiment, 5 sets of longitudinal 3D T1-SPGR MR brain images of healthy older adults from BLSA dataset are selected in the experiments, each set of MR images contain 11 serial scans where they were obtained during 11 consecutive years.

B. RESULTS

To demonstrate the temporal consistency of the segmentation results of our method, we compare it with a 3D segmentation method, called FAST, implemented by a popular software FSL based on the work of Zhang et al. [7]. The FSL
software is publicly available in the Analysis Group, FMRIB, Oxford, UK.\footnote{http://www.fmrib.ox.ac.uk/fsl/}

We first show the results of an experiment with the BLSA data. Fig. 1 shows the segmentation results for the images acquired in years 5, 6, and 7 (shown from left to right on the top row) obtained by applying FAST to each image individually and our 4D segmentation method to the 11 images jointly.

The middle row and the bottom row show the results of FAST and our method, respectively. It can be clearly seen that the results of our method for these three images are temporally more consistent than the results of FAST. To see this more clearly, we enlarge a portion of the images in Fig. 2 (marked by the red circle) and show the enlarged images in Fig. 2.

We can see multiple locations where the segmentation results of our method are more consistent than those of FAST. One of such locations is marked with a red arrow in Fig. 2 as an example.

To quantitatively evaluate the segmentation results, we use two metrics to evaluate the temporal consistency of the segmentation results. The first metric is the temporal consistency ($TC$) factor proposed in [10], which is defined as:

$$
TC = \frac{1}{|\Omega|} \sum_{x \in N} \left(1 - \frac{k_x}{(Y - 1)}\right),
$$

where $Y$ is the number of scans. The $TC$ factor is always between 0 and 1. A larger $TC$ factor indicates a better temporal consistency, and vice versa.

Each voxel $x$ in the region of interest $\Omega$ is labeled as WM, GM, and CSF as the result of tissue segmentation. We denote by $k_i$ the number of label changes of corresponding voxels across time, then the segmentation of the corresponding voxels considered to be consistent if $k_i$ is small, and vice versa.

![FIGURE 1. Segmentation results of our method (in bottom row) with FAST (in middle row) for images of years 5, 6, and 7 in a series of MR images, the original images are shown in the top row.](image)

![FIGURE 2. An enlarged view of the original images and segmentation results shown in Fig. 1. The results of our method (in bottom row) with FAST (in middle row) for images of years 5, 6, and 7 in a series of MR images as shown in the top row.](image)

![FIGURE 3. Evaluation of segmentation temporal consistency in terms of $TC$ factor: (a) $TC$ for GM; (b) $TC$ for WM.](image)

![TABLE 1. The average $TC$ factor results in five subjects of different methods.](image)
the proposed method can enforce better temporal consistency than those compared methods.

By definition, the TC factor is in favor of the segmentation results that are the same for all the images in the longitudinal series, even though the brains usually change over time.

In this paper, we propose a different metric to evaluate the temporal consistency, with consistent changes of the label assigned to corresponding voxels. This metric is evaluated by counting the voxels that are labeled incorrectly at three consecutive time points \( t_1, t_2, \) and \( t_3 \). We know that, due to the aging of the subjects or some pathological changes, the brains would shrink and deform gradually. Therefore, when a voxel \( x \) is labeled as a tissue A (e.g. GM) at time point \( t_1 \), it could be labeled as a different tissue B (e.g. WM) at the time point \( t_2 \), due to the changes of the brain. At the third time point \( t_3 \), this voxel would be labeled as either tissue B or back to A, while the latter case is not supposed to occur when the segmentation algorithm performs well. Otherwise, this voxel is labeled back as tissue A again at the time point \( t_3 \) as at the time point \( t_1 \). We call such label change as a label reverse. We assume that the label reverse is largely due to the result of inconsistent segmentation. We count the number of voxels that undergo label reverse in three consecutive time points, and compute the percentage of such voxels, which is referred to as a label reverse (LR) rate. A temporally consistent segmentation result should have a low LR rate. The LR rates our method and the FAST algorithm for the five subjects for time points \( t_1 = 4, t_2 = 5, \) and \( t_3 = 6 \) are plotted in Fig. 4(a), while the LR rates for \( t_1 = 5, t_2 = 6, \) and \( t_3 = 7 \) are plotted in Fig. 4(b). Clearly, the LR rates of our method are all below 0.6%, which are significantly lower than those of the FAST algorithm, which are all above 2%.

With more consistent segmentation results of our method, the brain tissue volume can be measured more precisely and stably over time. We use an open-source software platform called 3D-Slicer [31] to calculate the tissue volume. This is demonstrated by plotting the volumes of the WM and GM measured from a series of 11 images of a subject by using our method and a 3D segmentation algorithm, as shown in Fig. 5. It can be obviously observed from this figure that the results of our method show a more stable decreasing of the WM and GM volume over time, compared with the result of the FAST algorithm.
3 levels of noise, and 3 levels of intensity non-uniformity, Table 2 shows the detail of the used data.

Fig. 6 shows the robustness of our method with data with varying inhomogeneities. Fig. 7 shows the bias correction and segmentation result. And Fig. 8 shows our energy function minimization, after about 5 iterations, the energy changed to be steady, which shows that our energy function can converge quickly.

V. CONCLUSION

We propose a robust and efficient temporally consistent segmentation method of longitudinal MR data. In our method, tissue segmentation result is represented by the membership functions of the tissues, which are derived as a result of minimizing energy in a fuzzy C-means (FCM) framework. Our method does not require pre-processing for intensity inhomogeneity correction and intensity normalization of the series of input images. There are very few parameters in our method, and it is easy to tune the parameter to achieve desirable results. Our method is robust, efficient, and fully automatic, with stable performance and sufficient accuracy, which have been verified by experiments on a set of real MR data.

ACKNOWLEDGMENT

(Pengcheng Li and Yue Zhao contributed equally to this work.)

REFERENCES

[1] Y. Guo, Z. Wu, and D. Shen, “Learning longitudinal classification-regression model for infant hippocampus segmentation,” Neurocomputing, to be published.

[2] G. Erus, J. Doshi, Y. An, D. Vargelakas, S. M. Resnick, and C. Davatzikos, “Longitudinally and inter-site consistent multi-atlas based parcellation of brain anatomy using harmonized atlases,” Neuroimage, vol. 166, pp. 71–78, Feb. 2018.

[3] A. Carass, S. Roy, A. Jog, J. L. Cuzzoceo, E. Magrath, A. Gherman, J. Button, J. Nguyen, F. Prados, and C. H. Sudre, “Longitudinal multiple sclerosis lesion segmentation: Resource and challenge,” Neuroimage, vol. 148, no. 1, pp. 71–102, 2017.

[4] R. Nylander, M. Fahlström, E. Rostrup, J. Kullberg, S. Damangir, H. Ahlström, L. Lind, and E.-M. Larsson, “Quantitative and qualitative mri evaluation of cerebral small vessel disease in an elderly population: A longitudinal study,” Acta Radiolog., vol. 59, no. 5, pp. 612–618, 2018.

[5] S. M. Resnick, A. F. Goldsztajn, C. Davatzikos, S. Golski, M. A. Kraut, E. J. Metter, R. N. Bryan, and A. B. Zonderman, “One-year age changes in MRI brain volumes in older adults,” Cerebral Cortex, vol. 10, no. 5, pp. 464–472, 2000.

[6] W. M. Wells, III, W. E. L. Grimson, R. Kikinis, and F. A. Jolesz, “Adaptive segmentation of MRI data,” IEEE Trans. Med. Imag., vol. 15, no. 8, pp. 429–442, Aug. 1996.

[7] Y. Zhang, M. Brady, and S. Smith, “Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm,” IEEE Trans. Med. Imag., vol. 20, no. 1, pp. 45–57, Jan. 2001.

[8] D. L. Pham and J. L. Prince, “Adaptive fuzzy segmentation of magnetic resonance images,” IEEE Trans. Med. Imag., vol. 18, no. 9, pp. 737–752, Sep. 1999.

[9] K. Van Leemput, F. Maes, D. Vandermeulen, and P. Suetsens, “Automated model-based bias field correction of MR images of the brain,” IEEE Trans. Med. Imag., vol. 18, no. 10, pp. 885–896, Oct. 1999.

[10] Z. Xue, D. Shen, and C. Davatzikos, “CLASSIC: Consistent longitudinal alignment and segmentation for serial image computing,” Neuroimage, vol. 30, no. 2, pp. 388–399, 2006.

[11] M. Reuter and B. Fischl, “Avoiding asymmetry-induced bias in longitudinal image processing,” Neuroimage, vol. 57, no. 1, pp. 19–21, 2011.

[12] P. M. Thompson, A. D. Lee, R. A. Dutton, J. A. Geaga, K. M. Hayashi, M. A. Eckert, U. Bellugi, A. M. Galaburda, J. R. Korenbrot, D. L. Mills, A. W. Toga, and A. L. Reiss, “Abnormality of cortical complexity and thickness profiles mapped in Williams syndrome,” J. Neurosci., vol. 25, no. 16, pp. 4146–4158, 2005.

[13] L. Wang et al., “4D segmentation of longitudinal brain MR images with consistent cortical thickness measurement,” in Proc. Int. Workshop Spatio-Temporal Image Anal. Longitudinal Time-Ser. Image Data. Berlin, Germany: Springer, 2012, pp. 63–75.

[14] C. Feng, D. Zhao, and M. Huang, “Segmentation of longitudinal brain MR images using bias correction embedded fuzzy C-means with non-locally spatio-temporal regularization,” J. Vis. Commun. Image Represent., vol. 38, pp. 517–529, Jul. 2016.

[15] K. Zhan, X. Chang, J. Guan, L. Chen, Z. Ma, and Y. Yang, “Adaptive structure discovery for multimedia analysis using multiple features,” IEEE Trans. Cybern., vol. 49, no. 5, pp. 1826–1834, Mar. 2019.

[16] W. Dai et al., “SCAN: Structure correcting adversarial network for organ segmentation in challenging registration,” in Proc. Int. Conf. Med. Image Analysis and Multimodal Learning for Clinical Decision Support. Cham, Switzerland: Springer, 2018, pp. 263–273.

[17] Z. Zhang, Y. Zou, and C. Gan, “Textual sentiment analysis via three different attention convolutional neural networks and cross-modality consistent regression,” Neurocomputing, vol. 275, pp. 1407–1415, Jan. 2018.

[18] C. Q. Gao, D. Meng, Y. Yang, Y. Wang, X. Zhou, and A. G. Hauptmann, “Infrared patch-image model for small target detection in a single image,” IEEE Trans. Image Process., vol. 22, no. 12, pp. 5096–5095, Dec. 2013.

[19] Y. Zhao, S. Guo, M. Luo, X. Shi, M. Bilello, S. Zhang, and C. Li, “A level set method for multiple sclerosis lesion segmentation,” Magn. Reson. Imag., vol. 49, pp. 94–100, Jun. 2018.

[20] C. Gao, L. Wang, Y. Xiao, Q. Zhao, and D. Meng, “Infrared small-dim target detection based on Markov random field and guided noise modeling,” Pattern Recognit., vol. 76, pp. 453–465, Apr. 2018.

[21] Y. Yan, C. Xu, D. Cai, and J. J. Corso, “Weakly supervised actor-action segmentation via robust multi-task ranking,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2017, pp. 1298–1307.

[22] L. Nie, L. Zhang, Y. Yan, X. Chang, M. Liu, and L. Shao ling, “Multi-view physician-specific attributes fusion for health seeking,” IEEE Trans. Cybern., vol. 47, no. 11, pp. 3680–3691, Nov. 2017.

[23] C. Lodig, R. A. Heckemann, A. Hamnners, J. C. Lopez, V. F. J. Newcombe, A. Makropoulos, J. Lötjönen, D. K. Menon, and D. Rueckert, “Robust whole-brain segmentation: Application to traumatic brain injury,” Med. Image Anal., vol. 21, no. 1, pp. 40–58, 2015.

[24] V. P. Salve, A. K. Salve, and K. C. Jondhale, “Brain tumor segmentation using MS algorithm,” Int. J. Eng. Technol. Sci. Res., vol. 4, no. 4, pp. 118–123, 2017.

[25] C. A. Pluass, A. J. Pluassard, A. Carass, S. M. Resnick, D. L. Pham, J. L. Prince, and B. A. Landman, “Consistent cortical reconstruction and multi-atlas brain segmentation,” Neuroimage, vol. 138, pp. 197–210, Sep. 2016.

[26] S. Liao, H. Jia, G. Wu, and D. Shen, “A novel framework for longitudinal atlas construction with groupwise registration of subject image sequences,” Neuroimage, vol. 59, no. 2, pp. 1275–1289, 2012.

[27] B. C. Davis, P. T. Fletcher, E. Bullitt, and S. Joshi, “Population shape regression from random design data,” Int. J. Comput. Vis., vol. 90, no. 2, pp. 255–266, 2010.
[28] S. K. Adhikari, J. K. Sing, D. K. Basu, M. Nasipuri, and P. K. Saha, “A non-parametric method for intensity inhomogeneity correction in MRI brain images by fusion of Gaussian surfaces,” *Signal Image Video Process.*, vol. 9, no. 8, pp. 1945–1954, 2015.

[29] Y. K. Dubey, M. M. Mushrif, and K. Mitra, “Segmentation of brain MR images using rough set based intuitionistic fuzzy clustering,” *Biocybern. Biomed. Eng.*, vol. 36, no. 2, pp. 413–426, 2016.

[30] M. M. George and S. Kalaivani, “Retrospective correction of intensity inhomogeneity with sparsity constraints in transform-domain: Application to brain MRI,” *Magn. Reson. Imag.*, vol. 61, pp. 207–223, Sep. 2019.

[31] R. Kikinis, S. D. Pieper, and K. G. Vosburgh, “3D slicer: A platform for subject-specific image analysis, visualization, and clinical support,” in *Intraoperative Imaging and Image-Guided Therapy*. New York, NY, USA: Springer, 2014, pp. 277–289.

PENGCHENG LI received the bachelor’s degree in telecommunication engineering from the Chongqing University of Posts and Telecommunications, China, in 2017, where he is currently pursuing the master’s degree with the School of Communication and Information Engineering. His research interests include image processing, deep learning, and medical image analyzing.

YUE ZHAO received the Ph.D. degree from Jilin University, China, in 2017. She is currently a Lecturer with the School of Communication and Information Engineering, Chongqing University of Posts and Telecommunications. Her research interests include image processing and analysis, pattern recognition, and medical image processing.

YANG LIU received the M.D. and Ph.D. degrees from Peking University, China, in 2014 and 2017, respectively. He was a Visiting Scholar with the University of Pennsylvania from 2015 to 2017. He has been a member of Orthodontists of the Royal College of Surgeons of Edinburgh. He is currently an Orthodontist with the Stomatology Hospital of Chongqing Medical University. His research interests include CBCT image reconstruction, CBCT tooth segmentation analysis, and orthodontic tooth movement mechanism.

FANGCEN LIU received the bachelor’s degree in information engineering from the Chongqing University of Posts and Telecommunications, China, in 2018, where she is currently pursuing the master’s degree with the School of Communication and Information Engineering. Her research interests include image processing, deep learning, and infrared image processing.

CHENQIANG GAO received the Ph.D. degree from the Huazhong University of Science and Technology, China, in 2009. He is currently a Professor with the School of Communication and Information Engineering, Chongqing University of Posts and Telecommunications. His research interests include infrared image processing, action recognition, event detection, computer vision, and machine learning.