

Unsupervised Segmentation of Head Tissues from Multi-modal MR Images for EEG Source Localization

Qaiser Mahmood · Artur Chodorowski ·
Andrew Mehnert · Johanna Gellermann ·
Mikael Persson

© Society for Imaging Informatics in Medicine 2014

Abstract In this paper, we present and evaluate an automatic unsupervised segmentation method, hierarchical segmentation approach (HSA)–Bayesian-based adaptive mean shift (BAMS), for use in the construction of a patient-specific head conductivity model for electroencephalography (EEG) source localization. It is based on a HSA and BAMS for segmenting the tissues from multi-modal magnetic resonance (MR) head images. The evaluation of the proposed method was done both directly in terms of segmentation accuracy and indirectly in terms of source localization accuracy. The direct evaluation was performed relative to a commonly used reference method brain extraction tool (BET)–FMRIB’s automated segmentation tool (FAST) and four variants of the HSA using both synthetic data and real data from ten subjects. The synthetic data includes multiple realizations of four different noise levels and several realizations of typical noise with a 20 % bias field level. The Dice index and Hausdorff distance were used to measure the segmentation accuracy. The indirect evaluation was performed relative to the reference method BET-FAST using synthetic two-dimensional (2D) multimodal magnetic resonance (MR) data with 3 % noise and synthetic EEG (generated for a prescribed source). The source localization accuracy was determined in terms of localization error and relative error of potential. The experimental results demonstrate the efficacy of HSA-BAMS, its robustness to noise and the bias field, and that it provides better segmentation

accuracy than the reference method and variants of the HSA. They also show that it leads to a more accurate localization accuracy than the commonly used reference method and suggest that it has potential as a surrogate for expert manual segmentation for the EEG source localization problem.

Keywords Magnetic resonance imaging · Multi-modal · Mean shift · Bayesian · Automatic segmentation · Head tissues · Electroencephalography (EEG) · Source localization

Introduction

Expert manual interpretation is considered the gold standard for determining accurate localization of the source of epileptic seizures from electroencephalography (EEG) scalp measurements. However, it is a time-consuming and laborious process [1]. A number of automatic non-invasive EEG source localization techniques [2] have been developed to overcome this problem. The accuracy of these techniques is not only dependent on the methods used to solve the underlying forward and inverse problems [3] but also on the quality and fidelity of the patient-specific head conductivity model used in the forward problem. The construction of a realistic head conductivity model requires accurate segmentation of magnetic resonance (MR) images of the head into tissue classes corresponding to different conductivity values. It is common practice to use five tissue classes—white matter (WM), gray matter (GM), cerebrospinal fluid (CSF), skull, and skin [4, 5] followed by manual correction for source localization, although improvements in localization accuracy can be achieved with more classes [6]. Completely manual segmentation [7], typically performed by a clinical expert, is the gold standard. Nevertheless, it is laborious and time consuming. Two refinements that can be found in the literature are semi-automatic/interactive segmentation requiring user input [6, 8–11] and fully

Q. Mahmood (✉) · A. Chodorowski · A. Mehnert · M. Persson
Department of Signals and Systems, Chalmers University of
Technology, Gothenburg 41296, Sweden
e-mail: qaiserm@chalmers.se

Q. Mahmood · A. Chodorowski · A. Mehnert · M. Persson
MedTech West, Sahlgrenska University Hospital,
Gothenburg 41285, Sweden

J. Gellermann
Universitätsklinik für Radioonkologie, Tübingen 72076, Germany

automatic segmentation followed by manual correction [5, 12, 13]. However, the need for user interaction or manual correction means that these approaches are still time consuming, subjective, tedious, and labor intensive. Moreover, they are impractical for large-scale group study. Fully automatic and accurate segmentation is thus highly desirable.

The development of a fully automated and accurate segmentation method for the construction of patient-specific head models for EEG source localization is a challenging task for several reasons including the following: (i) the complexity and variability of the underlying anatomy, (ii) noise, (iii) the bias field (spatial-intensity inhomogeneities), and (iv) the low contrast between the skull, CSF, and air in traditionally used T1-weighted (T1w) images. In this paper, we present a new fully automatic method for head tissue segmentation from multi-modal MR images suitable for EEG source localization. In common with several approaches in the literature [5, 12, 13], the method involves first partitioning the MR data into brain-tissue and non-brain-tissue sub-volumes and then independently segmenting each of these into multiple tissue classes. The idea behind this hierarchical segmentation approach (HSA) is that the detection of brain and non-brain tissue is a much simpler initial problem than the problem of segmenting the whole head into multiple tissue classes outright. For example, the brain extraction tool (BET; is a part of the FMRIB Software Library (FSL) tools developed by the FMRIB, University of Oxford) [14] can be used to robustly obtain a brain-tissue mask while simple thresholding and mathematical morphology operations [15, 16] can be used to obtain a whole-head mask (from which the non-brain-tissue mask can then be trivially obtained). What distinguishes our method is that a single segmentation approach, Bayesian adaptive mean shift (BAMS), is used to segment both the brain-tissue and non-brain-tissue sub-volumes into multiple tissue classes. Mean-shift segmentation offers two important advantages here. Firstly, it is an unsupervised segmentation technique meaning that it is not necessary to provide training examples. Secondly, it can make use of multiple magnetic resonance imaging (MRI) modalities (e.g., T1w, T2w, and proton density (PD) weighting) and indeed multiple imaging modalities (e.g., CT and MRI).

We also present direct and indirect evaluations of our proposed HSA-BAMS method. For the direct evaluation, the segmentation accuracy of HSA-BAMS is compared with that of several other instantiations of the HSA, and to a reference method BET-FMRIB's automated segmentation tool (FAST), using both synthetic multi-modal MRI data and real data (both single and multi-modal MRI data) from ten subjects. Each HSA instantiation is based on a different multi-tissue segmentation algorithm: the hidden Markov random field model and associated expectation-maximization (HMRF-EM) algorithm [17] (as implemented in the FAST in the FSL tools), the adaptive mean-shift (AMS) algorithm of Mayer and Greenspan [18], the improved Fuzzy c -means algorithm with spatial constraints

(FCM_S), proposed in [19] and the simple k -means clustering algorithm [20]. The reference method BET-FAST is based on the BET [14, 21] and FAST [17, 21] tools commonly used [5, 12, 13] in the construction of realistic head models for EEG source localization. The BET tool is used with extended functionality to extract not only the brain-tissue mask but also the skin- and skull-tissue masks (ideally both T1w and T2w images should be used although it is possible to use T1w only) and the FAST tool is employed to segment the brain tissues. For the indirect evaluation, HSA-BAMS is compared with BET-FAST and completely manual segmentation in terms of EEG source localization accuracy using synthetic two-dimensional (2D) multi-modal MRI and EEG data.

Materials and Methods

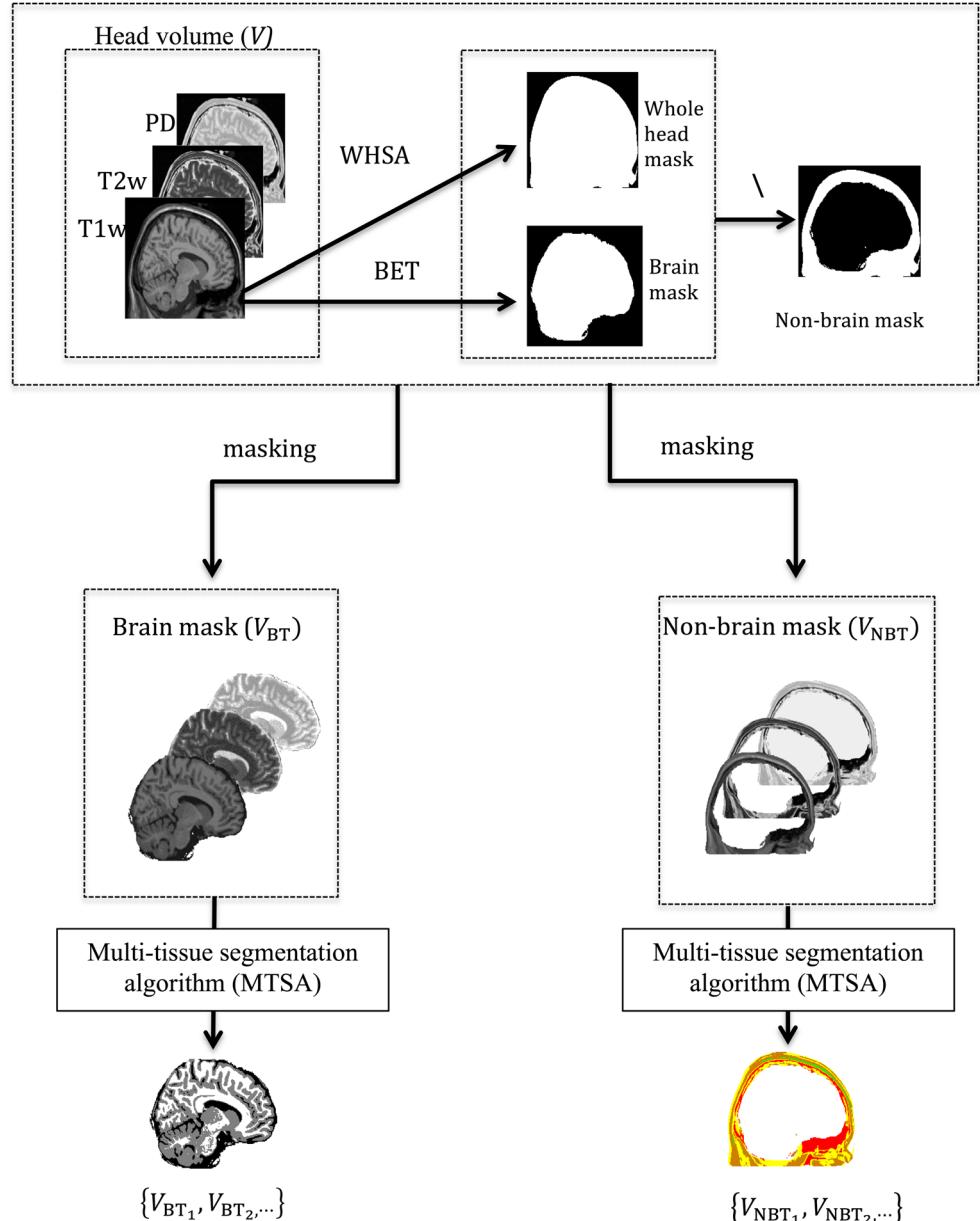
Hierarchical Segmentation Approach

A schematic of the HSA is presented in Fig. 1. The HSA takes as input a set of T1w MR images of the whole head and optionally several additional spatially co-registered sets of MR images (e.g., T2w and PD) of the whole head. This data can be modeled as a single spatial volume (V) with vector-valued voxels. In the first level of the HSA, the T1w data is used to obtain both a brain mask and a whole-head mask. The BET tool is used to obtain the former and a simple whole-head segmentation algorithm (WHSAs) is used to obtain the latter. The WHSA comprises two simple steps: (i) Otsu thresholding [15] and (ii) hole filling using morphological reconstruction and 26 connectivity [16]. The set difference between these two masks then yields a mask of the non-brain head tissue. These masks effectively partition the head volume (V) into two disjoint sub-volumes: brain tissue (V_{BT}) and non-brain tissue (V_{NBT}). In the second level of the HSA, the multi-tissue segmentation algorithm (MTSA) is applied independently to the brain-tissue (V_{BT}) and non-brain-tissue (V_{NBT}) volumes to segment them into individual tissue classes $V_{BT_1}, V_{BT_2}, \dots$ and $V_{NBT_1}, V_{NBT_2}, \dots$ respectively. In our experiments below, we used the following multi-tissue segmentation algorithms (MTSAs): our own BAMS, AMS, FAST, FCM_S, and k -means. Hereinafter, these four instantiations of the HSA are denoted HSA-BAMS, HSA-AMS, HSA-FAST, HSA-FCM_S, and HSA- k -means,

Bayesian Adaptive Mean Shift

BAMS is a variation on the AMS [22, 23] segmentation method originally proposed by Mayer and Greenspan [18] for brain tissue segmentation in MR images. Mayer and Greenspan define the adaptive kernel bandwidth for the mean-shift algorithm in terms of the distance between the current feature point and its k th nearest

Fig. 1 Schematic of the proposed hierarchical segmentation approach (HSA) for automated whole-head segmentation. In the figure, *WHSAs* stands for whole-head segmentation algorithm and *BET* stands for brain extraction tool



neighbor. However, the bandwidth value defined using this approach can be biased by outliers [24].

In [24], a global bandwidth estimation approach is proposed for the kernel that does not have this problem. In BAMS, we proposed an adaptive bandwidth estimation approach for the kernel for the mean shift algorithm, which is a variation on the bandwidth estimation approach proposed in [24]. The approach is based on a Bayesian method that involves fitting the Gamma distribution probability density function to the local variances of N sets of neighborhoods around the current feature point. See the [Appendix](#) for more details.

Mean shift is itself an iterative algorithm for finding the modes of a multivariate probability density function given discrete data sampled from it. Let $\{\mathbf{x}_i \in \mathbb{R}^d \mid i=1, \dots, n\}$ denote

this set of points. The density at point \mathbf{x} can be estimated by the Parzen window kernel density estimator

$$\hat{f}_K(\mathbf{x}) = \frac{1}{n} \sum_{i=1}^n \frac{1}{h_i^d} k\left(\frac{\|\mathbf{x}-\mathbf{x}_i\|^2}{h_i}\right) \quad (1)$$

where $h(\mathbf{x}_i) \equiv h_i > 0$ is the adaptive bandwidth of the radially symmetric kernel K for point \mathbf{x}_i and $k: [0, 1] \rightarrow \mathbb{R}$ is the kernel profile of the kernel K with bounded support defined as

$$K(\mathbf{x}) = c_{k,d} k\left(\|\mathbf{x}\|^2\right) \quad \|\mathbf{x}\| \leq 1 \quad (2)$$

and $c_{k,d}$ is a normalizing constant ensuring that the kernel K integrates to 1. The derivative of Eq. 1 leads to

$$M_{h_i, G}(\mathbf{x}) = C \frac{\widehat{f}_K(\mathbf{x})}{f_G(\mathbf{x})} = \frac{\sum_{i=1}^n \frac{1}{h_i^{d+2}} \mathbf{x}_i g\left(\frac{\|\mathbf{x}-\mathbf{x}_i\|^2}{h_i}\right)}{\sum_{i=1}^n \frac{1}{h_i^{d+2}} g\left(\frac{\|\mathbf{x}-\mathbf{x}_i\|^2}{h_i}\right)} \mathbf{x} \quad (3)$$

where $M_{h_i, G}(\mathbf{x})$ represents a mean-shift vector, which points toward the direction of maximum increase in density and also provides the basis for clustering, C is a positive constant, and $g(\mathbf{x})=k'(\mathbf{x})$ is the kernel profile of kernel G .

Given a starting point \mathbf{y}_1 of the kernel G , the following iterative rule defines successive locations of the kernel G toward a denser region or mode (local maximum):

$$\mathbf{y}_{j+1} = \frac{\sum_{i=1}^n \frac{1}{h_i^{d+2}} \mathbf{x}_i g\left(\frac{\|\mathbf{y}_j-\mathbf{x}_i\|^2}{h_i}\right)}{\sum_{i=1}^n \frac{1}{h_i^{d+2}} g\left(\frac{\|\mathbf{y}_j-\mathbf{x}_i\|^2}{h_i}\right)} \quad j = 1, 2 \quad (4)$$

The points that converge to the same mode, wherein the *magnitude* of the *mean-shift vector* converges to zero, constitute a cluster. To apply this to the problem of image segmentation, one represents each pixel as a feature point \mathbf{x}_i formed by concatenating its spatial coordinates and range values (e.g., T1w, T2w, and PD) and employs the following joint spatial-intensity domain kernel G defined

$$G(\mathbf{x}) = \frac{c}{h_s^p h_r^d} g\left(\frac{\|\mathbf{x}^s\|^2}{h_s}\right) g\left(\frac{\|\mathbf{x}^r\|^2}{h_r}\right) \quad (5)$$

where \mathbf{x}^s and \mathbf{x}^r are the spatial and range components of \mathbf{x} , and h_s and h_r are the spatial and range bandwidths, respectively, and c is the normalization constant.

In BAMS, the resulting modes are then pruned; i.e., two modes that are close to one another (with respect to the Mahalanobis distance) in the range domain are merged (decided using a window of radius R). This is done in an iterative fashion with increasing R until the variance of merging modes reaches a preset threshold value [18]. Finally, the desired number of final clusters is obtained by applying voxel-weighted k -means clustering. Cluster initialization is performed making use of prior information about tissue intensity ranges in the multi-modal MR images [18]. In this study, we used Gaussian kernels for both the spatial and range domains.

Reference Method: BET-FAST

The reference method is a combination of the BET and FAST tools commonly used for the creation of patient-specific conductivity models [5, 12, 13]. It takes the same inputs as the proposed HSA. In this study, the BET tool was applied to extract the brain-tissue mask, skull-tissue mask, and skin-

tissue mask from the T1w images in each data set. The FAST tool was then used to segment the brain-only image (masked version of the T1w images) into three tissue classes: white matter (WM), gray matter (GM), and cerebrospinal fluid (CSF).

MRI Data for Direct Evaluation

We used four different data sets to evaluate the segmentation accuracies of HSA-BAMS, the four other instantiations of the HSA (HSA-AMS, HSA-FAST, HSA-FCM_S, and HSA- k -means) and the reference method BET-FAST. These data sets are hereinafter referred to as data set 1, data set 2, data set 3, and data set 4.

Data set 1 comprises synthetic multi-modal MRI data, generated using custom MRI simulations from the BrainWeb-simulated brain database (SBD) [25]. More specifically, the data set comprises realistic [18] T1w, T2w, and PD scans of a human head for four different noise levels—3, 5, 7, and 9 % [25] with five realizations each, and also for a 20 % bias field level with five realizations of 5 % noise. Each scan is of size $181 \times 217 \times 181$ (rows by columns by axial slices) with cubic voxels of size 1 mm^3 . The parameters that were used to generate the data from the custom MRI simulations are described on the BrainWeb website [26]. The simulations also contain some MRI acquisition artifacts above the head and around the nose (described on the simulation Website). These were removed manually using the Polygon tool implemented in the ITK-SNAP software [27].

Data set 2 comprises a real T1w MRI scan of the head of a healthy subject. The volume was acquired on a 3 T Philips Achieva scanner using a gradient echo sequence with the following parameters: TR=8.17 ms, TE=3.76 ms, and flip angle of 8° . The size of the scan is $256 \times 256 \times 195$ (rows by columns by sagittal slices) with voxels of size $0.94 \times 0.94 \times 1 \text{ mm}^3$.

Data set 3 comprises real multi-modal MRI data (T1w and T2w scans) of the head of a second healthy subject (acquired using the same scanner as used for data set 2). The T1w scan was performed using a gradient echo sequence with the same parameters as for data set 2. The size of the T1w scan is $256 \times 256 \times 195$ (rows by columns by sagittal slices) with voxels of size $0.94 \times 0.94 \times 1 \text{ mm}^3$. The T2w scan was performed using a spin echo sequence with the following parameters: TR=2500 ms, TE=233 ms, and flip angle of 90° . The size of the T2w scan is $256 \times 256 \times 327$ (rows by columns by sagittal slices) with voxels size of $0.94 \times 0.94 \times 0.55 \text{ mm}^3$.

Finally, data set 4 comprises real multi-modal MRI data (T1w, T2w and PD scans) of the heads of eight healthy subjects, of differing sex and age, obtained from the IXI datasets [28]. The details of these subjects are presented in Table 1. The data

Table 1 Details of the eight healthy subjects for data set 4

| Data set 4 (IXI datasets) | | | |
|---------------------------|------------------|--------|-----|
| No. | Name | Sex | Age |
| 1 | IXI040_Guys_0724 | Female | 44 |
| 2 | IXI109_Guys_0732 | Female | 36 |
| 3 | IXI121_Guys_0772 | Male | 35 |
| 4 | IXI144_Guys_0788 | Female | 29 |
| 5 | IXI191-Guys-0801 | Male | 44 |
| 6 | IXI249-Guys-1072 | Male | 68 |
| 7 | IXI491-Guys-1032 | Female | 75 |
| 8 | IXI587-Guys-1128 | Male | 33 |

of each subject were acquired using a Philips Medical Systems Intera 1.5 T scanner. The T1w scan for each subject was performed with the following parameters: TR=9.80 ms, TE= 4.60 ms, and flip angle of 8°. The size of the T1w scan is 256×256×150 (rows by columns by sagittal slices) with voxels of size 0.95×0.94×1.20 mm³. The T2w scan for each subject was performed with the following parameters: TR=8178 s, TE= 100 ms, flip angle of 90°. The size of the T2w scan is 256×256×130 (rows by columns by sagittal slices) with voxels of size 0.94×0.94×1.25 mm³. The PD scan for each subject was performed with the following parameters: TR=8178 ms, TE= 8.0 ms, and flip angle of 90°. The size of the PD scan is 256×256×130 (rows by columns by sagittal slices) with voxels of size 0.94×0.94×1.25 mm³.

A reference or gold standard segmentation for each data set was obtained as follows. For data set 1, the gold standard was obtained from the nine tissue classes defined in the synthetic data [25]. This was reduced to seven classes by merging tissues with similar conductivity values. In particular the connective and muscle tissue classes were merged, and the glial matter and GM classes were merged. For data sets 2, 3, and 4, a radio-oncologist, with more than 17 years experience, manually segmented each subject data into five tissue classes: WM, GM, CSF, skull and skin. This represented a tradeoff between the time needed to manually segment the images and delineating the essential classes for EEG source localization (these five classes are commonly used for EEG source localization [4, 5, 7]). The radio-oncologist included fat, muscle,

and skin in the “skin” class. The manual segmentation took about 170 h for 200 slices. Another clinical expert independently reviewed these segmentations and queried/resolved any perceived anomalies.

In EEG source localization it is generally accepted that weak volume currents in the face region (including the nose and pharynx) of subjects have a negligible influence on the measurement fields [4]. The reason is that these regions are located far away from the EEG electrodes. Given that they are also difficult to segment because the air cavities and bone cannot be distinguished in the T1w and T2w images, and because of susceptibility artifacts, we excluded these regions (using the ITK-SNAP polygon tool) from each real data set according to the cutting procedure described in [4, 29]. This was performed in consultation with a clinical expert.

Pre-processing: Co-registration and Intensity Scaling

The use of multiple modalities with differing spatial resolutions necessitates the use of spatial co-registration and resampling. This ensures that each voxel can be assigned a vector of multi-modal values. For this purpose, the T2w image was spatially co-registered with the T1w image for the subject in data set 3 and both the T2w and PD images were spatially co-registered with the T1w images for each subject in data set 4.

In this study, we employed the well-known and publicly available FLIRT tool for co-registration. It is “a fully automated robust and accurate tool for linear (affine) intra- and inter-modal brain image registration.” It is based on a hybrid global-local optimization technique [30] that provides a fast and accurate multi-modal affine registration.

Finally for each image (T1w, T2w, and PD) of each subject in each data set the intensity values were scaled to the interval [0, 1] using linear contrast stretching [18].

MRI and EEG Data for Indirect Evaluation

We used a 2D slice (slice number 100) from each scan (T1w, T2w and PD) in the 3 % noise data from data set 1 to indirectly evaluate the performances of the proposed method and the reference BET-FAST method in terms of EEG source localization accuracy.

Table 2 McNemar tests for the GM tissue class for data set 3

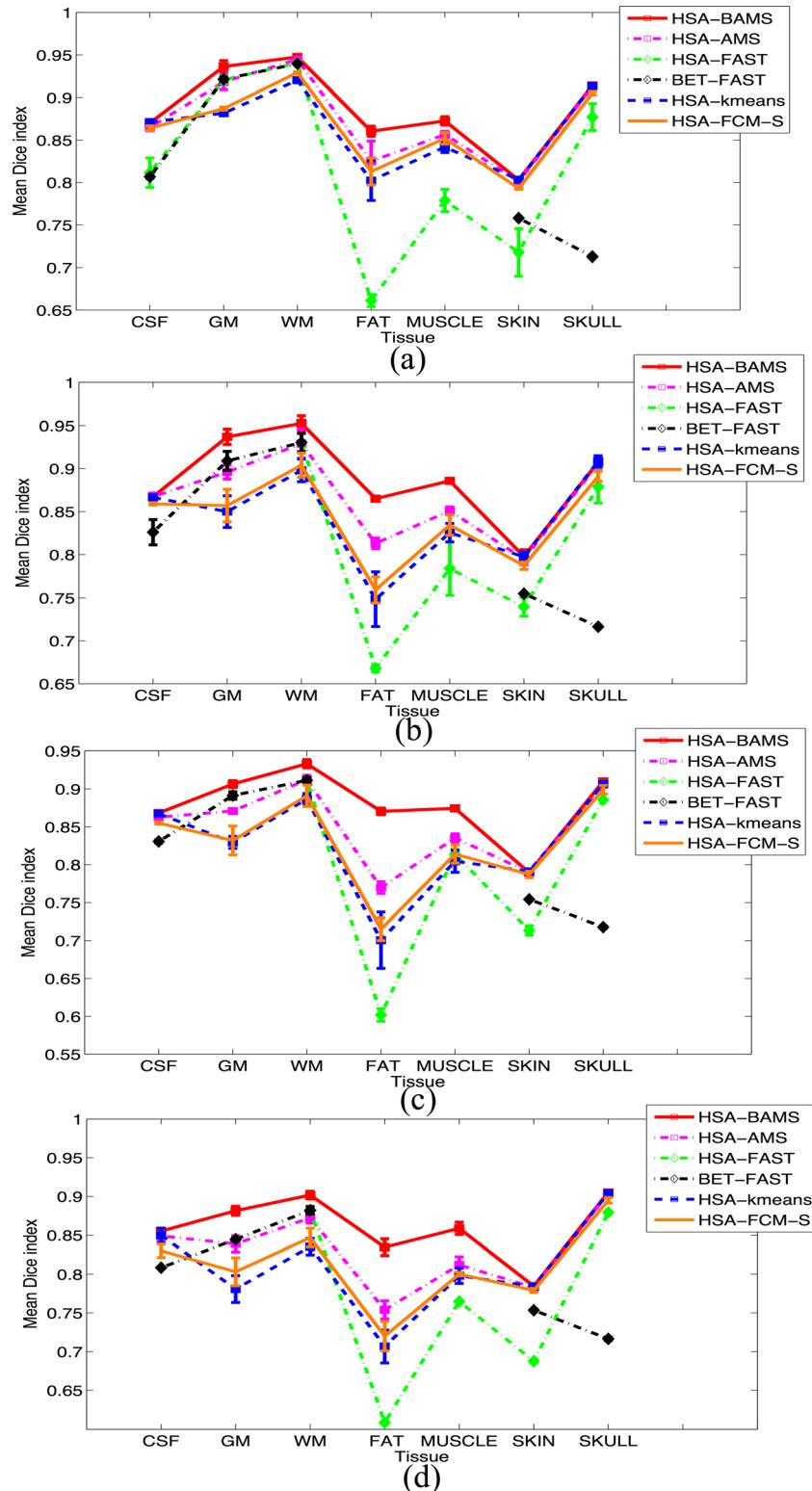
| Comparison test | n_{11} | n_{12} | n_{21} | n_{22} | p value |
|----------------------------------|----------|----------|----------|----------|-----------|
| Proposed vs HSA-AMS | 484,118 | 82,252 | 40,810 | 88,662 | <0.01 |
| Proposed vs HSA-FAST | 516,421 | 49,949 | 36,967 | 92,505 | <0.01 |
| Proposed vs HSA- <i>k</i> -means | 455,050 | 111,320 | 45,687 | 83,785 | <0.01 |
| Proposed vs HSA-FCM_S | 471,030 | 99,340 | 35,687 | 53,785 | <0.01 |
| Proposed vs BET-FAST | 516,421 | 49,949 | 36,967 | 92,505 | <0.01 |

Synthetic EEG was generated using the 2D ground truth by placing a source in the GM tissue and calculating EEG signals from 30 electrodes placed equidistantly around the head, based on the 10/10 system [1, 5, 7].

Fig. 2 Mean Dice index with 0 % bias field level for **a** 3, **b** 5, **c** 7, and **d** 9 % noise level. The whiskers show ± 1 standard deviation

Implementation Details and Optimal Parameters Settings for the Segmentation Algorithms

MATLAB, together with FSL, was used to implement all of the algorithms. The synthetic multi-modal MRI data with 0 %



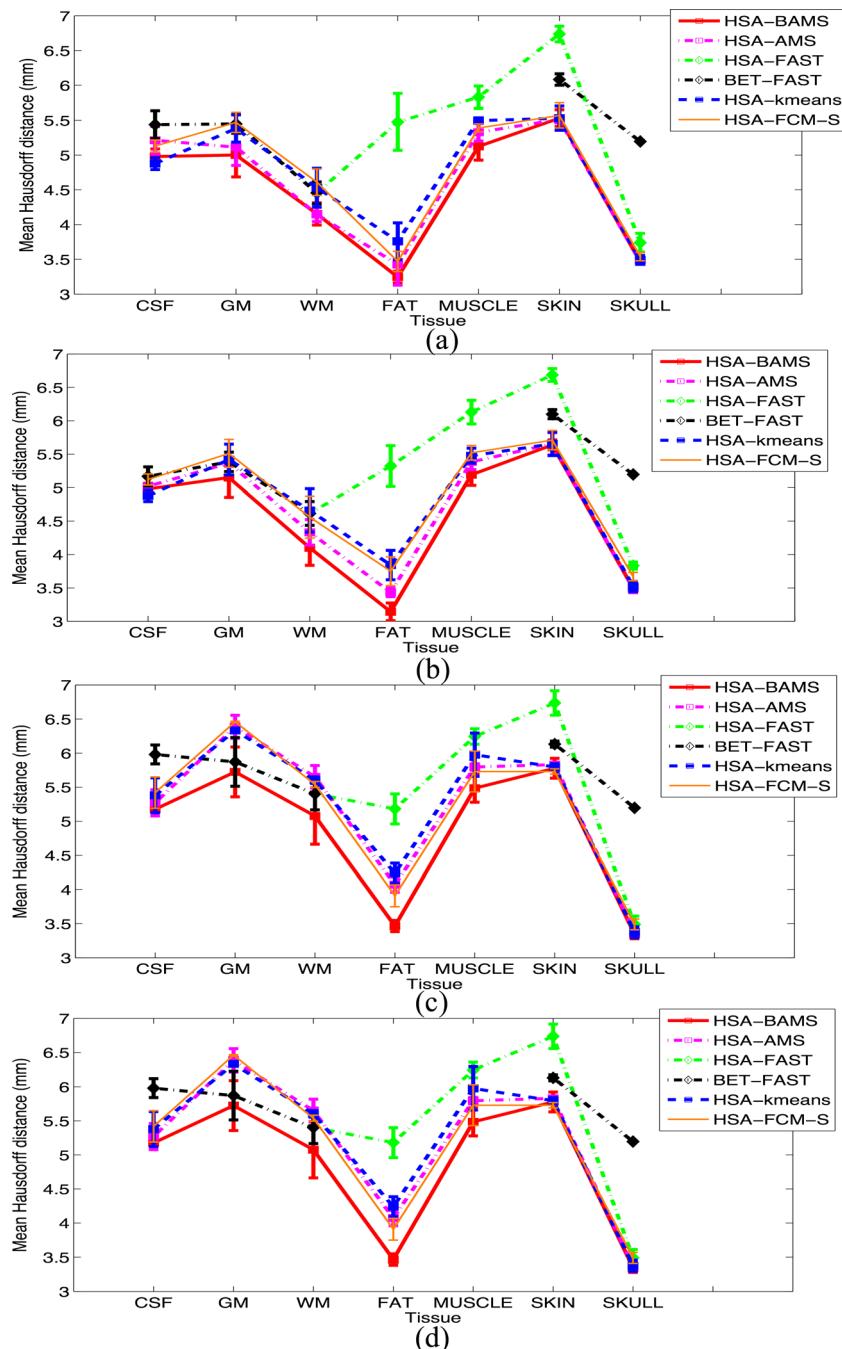
noise level and 0 % bias field level from the BrainWeb database [25] was used to empirically determine the optimal parameter settings for all algorithms.

The HSA-*k*-means method was implemented using MATLAB's “*k*-means” function with user-defined initialization. The initialization was based on the prior knowledge of tissue intensity ordering in MR images. For example, in the T1w brain images the darkest region corresponds to the CSF tissue, the brightest region corresponds to the WM tissue and the second brightest region corresponds to the GM tissue.

For the brain, skull, and skin mask extraction, the BET tool was applied with the default threshold parameter setting $f=0.5$.

In the HSA-FAST and BET-FAST methods, the FAST tool was employed with the default MRF beta value setting $H=0.1$ and user-defined initial tissue-type means (to initialize the clustering centers for tissue segmentation). These means were determined using the prior knowledge of tissue intensity ordering in MR images. For example, in the T1w brain image the highest intensity value was assigned as the mean for WM tissue, the lowest intensity value was assigned as the mean for

Fig. 3 Mean Hausdorff distance (mm) for 0 % bias field level with **a** 3, **b** 5, **c** 7, and **d** 9 % noise levels. The whiskers show plus/minus one standard deviation



CSF tissue, and the mean of the highest and lowest intensity values was chosen as the mean for GM tissue.

The optimal values of the parameters h_s , and M_1 , M_2 , and N (described in the Appendix) for the proposed method, HSA-BAMS, were obtained using a grid search over a predefined range of values. For each 4-tuple (h_s, M_1, M_2, N) , the Dice index (DI) was computed for all seven tissue types. The optimal 4-tuple was chosen to be that which yielded the maximum mean DI over the all seven tissue types. The optimal 4-tuple for brain-tissue (V_{BT}) segmentation was (3, 100, 330, 10) and (3.7, 100, 1300, 10) for non-brain-tissue (V_{NBT}) segmentation.

In the HSA-AMS method, the same values of h_s were chosen for brain-tissue (V_{BT}) and for non-brain-tissue (V_{NBT}) segmentation as determined for HSA-BAMS. The parameter ‘ k ’ (nearest neighborhoods) was set to 120 for brain-tissue (V_{BT}) segmentation as suggested in [18]. A gridded search was used to determine the optimal k for non-brain-tissue (V_{NBT}) segmentation (using the same optimality criterion used for HSA-BAMS). The optimal value of k for non-brain-tissue (V_{NBT}) segmentation was determined to be 120.

In the HSA-FCM_S method, the FCM_S [19] was employed with following parameters settings: $m=2$ and $\alpha=0.8$, wherein the parameter m was used to control the fuzziness of the resulting clusters and α was used to control the tradeoff between the original image and the

corresponding median-filtered image. Moreover, in FCM_S, the clusters was initialized using the prior knowledge of tissue intensity ordering in MR images

EEG Source Localization

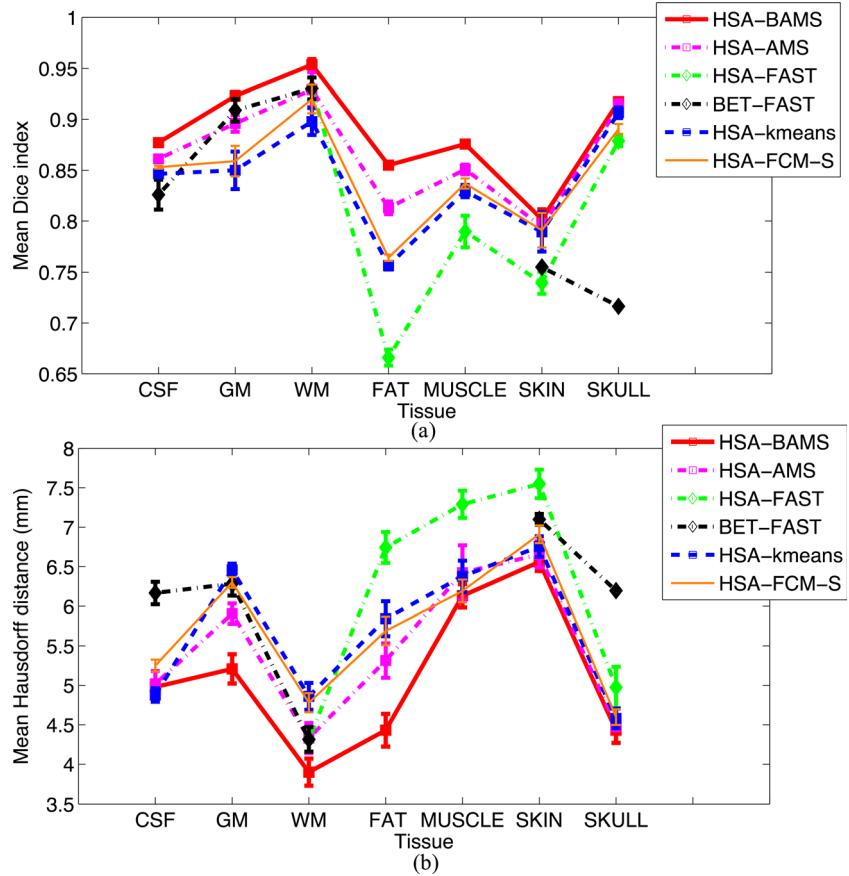
Three head conductivity models were constructed from the ground truth (GT), HSA-BAMS, and HSA-FAST segmentations respectively. The localization procedure is that described in [1, 5, 7]. In summary, given a head conductivity model and the EEG data, the procedure involves the solution of the following two problems: (i) the forward problem that deals with finding the scalp potentials for the given current sources and (ii) the inverse problem that deals with estimating the sources to fit the given potential distributions at the scalp electrodes.

Quantification of Segmentation Performances

The segmentation performances of HSA-BAMS, the four instantiations of the HSA, and the reference method BET-FAST were evaluated quantitatively using the DI [31] and Hausdorff distance (H) [32].

The DI was computed for each tissue type, data set, and segmentation method. The DI measures the degree of overlap

Fig. 4 **a** Mean Dice index. **b** Mean Hausdorff distance (mm) for five realizations of 5 % noise level with 20 % bias field level. The whiskers show ± 1 standard deviation



between the ground truth and the segmentation result. It is defined as

$$DI = \frac{2V_{ae}}{(V_a + V_e)} \quad (6)$$

where V_{ae} is the number of voxels the segmentation result and the ground truth have in common, and V_a and V_e denote the number of voxels in the segmentation result and the ground truth respectively. The DI has value one for perfect

segmentation and zero when there is no overlap between the segmentation result and ground truth. Likewise, H was measured for each tissue type, data set, and segmentation method. It is defined as

$$H = \max\{H_{SG}, H_{GS}\} \quad (7)$$

where S and G are two sets of points that belong to the segmentation result and the ground truth, respectively.

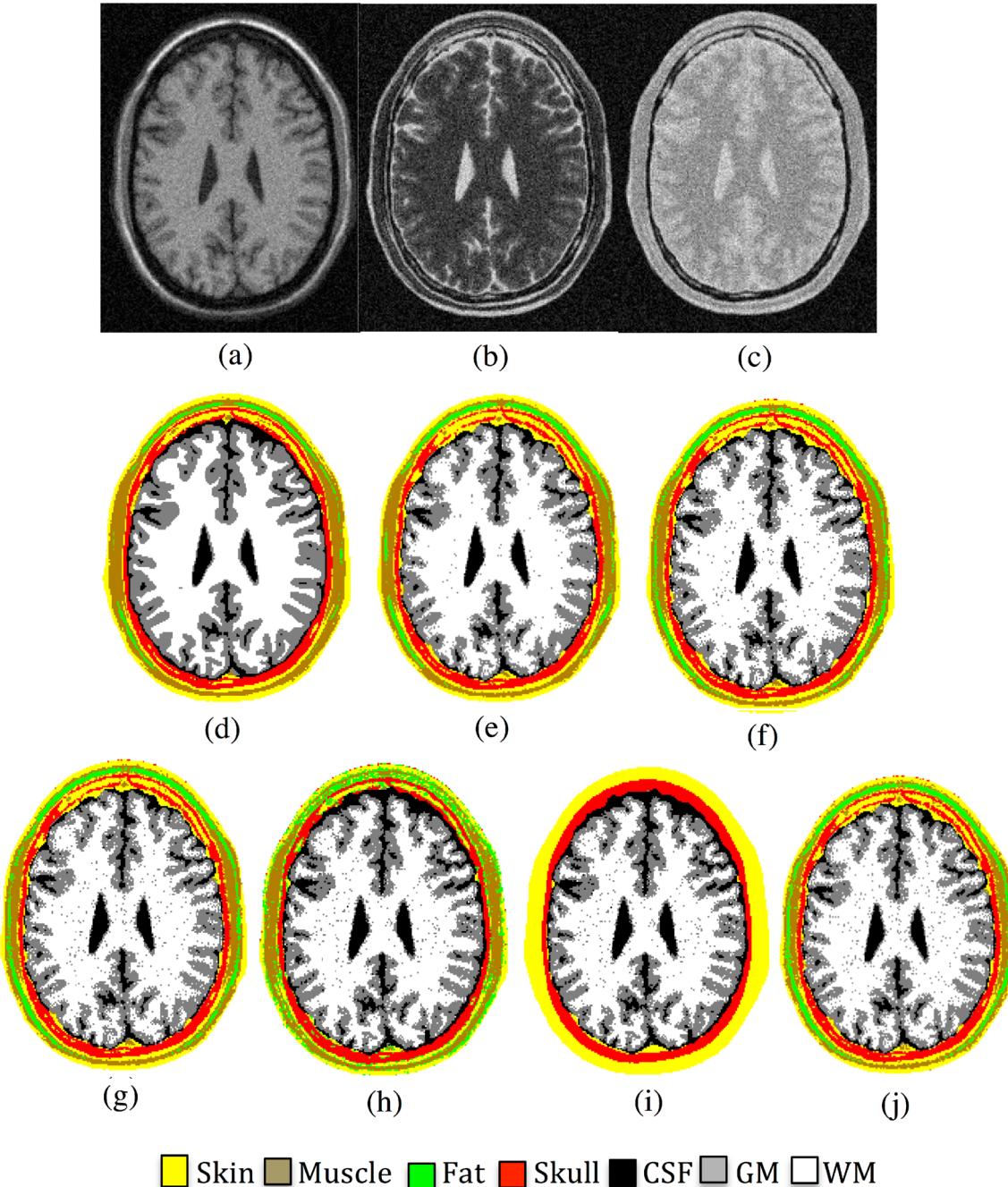


Fig. 5 Segmentation example from data set 1 for the 9 % noise level: **a** T1w, **b** T2w, **c** PD, **d** ground truth, **e** HSA-BAMS, **f** HSA-AMS, **g** HSA-*k*-means, **h** HSA-FAST, **i** BET-FAST, **j** HSA-FCM_S (WM in white, GM in gray, CSF in black, skull in red, fat in green, muscle in brown, and skin in yellow)

$H_{SG} = \max\{d_i^{SG}\}$ is the maximum value of the surface distance (Euclidian distance) of all surface voxels in S , and d_i^{SG} represents the minimum distance for the i th surface voxel in S to the set of surface voxels in G . Similarly, $H_{GS} = \max\{d_i^{GS}\}$ is the maximum value of the surface distance of all surface voxels in G and d_i^{GS} represents the minimum distance for the i th surface voxel in G to the set of surface voxels in S .

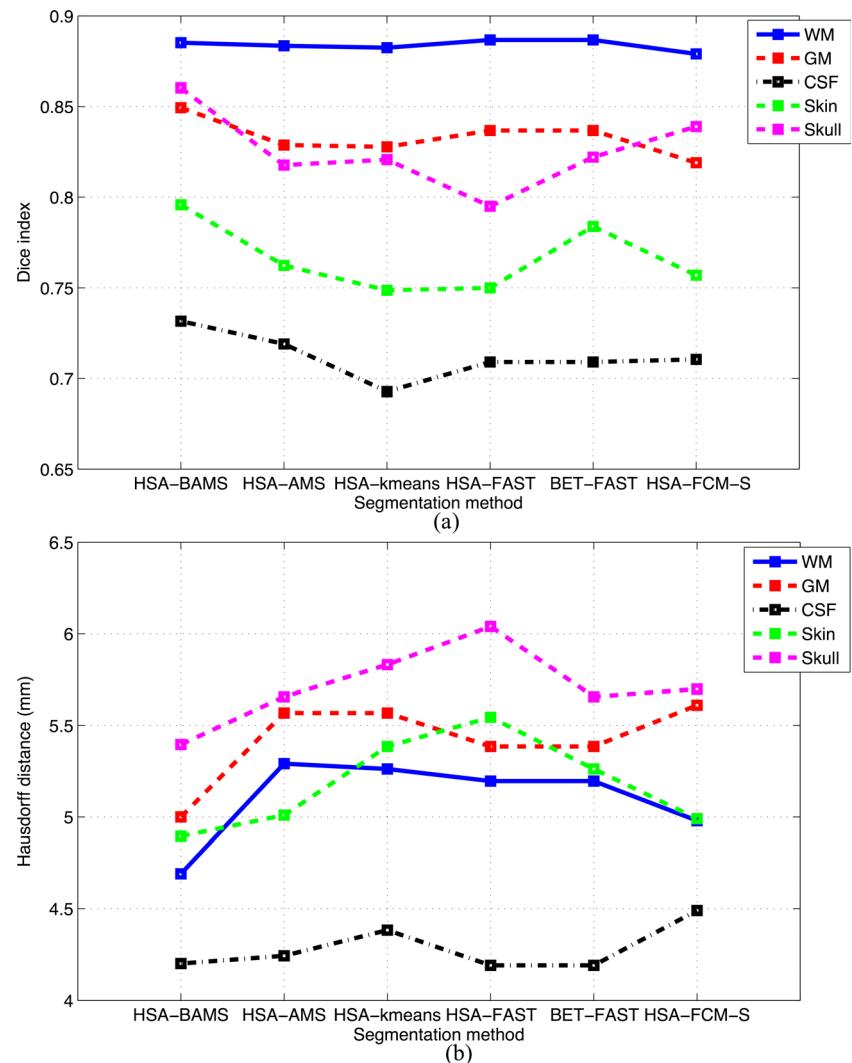
Tests for Statistical Significance in Voxel-Wise Classification Performances

To determine whether there exists a statistically significant difference in the voxel-wise classification performance between the proposed HSA-BAMS and each of the other methods for each tissue type and each data set, several multiple comparison tests were performed. Each multiple comparison test involved performing five McNemar tests [33], each comparing HSA-BAMS with one of the other methods. Each

McNemar test involved computing a 2×2 contingency table $\begin{bmatrix} n_{11} & n_{12} \\ n_{21} & n_{22} \end{bmatrix}$ where n_{11} is the number of voxels correctly classified by both methods, n_{12} is the number of voxels correctly classified by HSA-BAMS but not the other method, n_{21} is the number of voxels incorrectly classified by HSA-BAMS but correctly classified by the other method, and n_{22} is the number of voxels incorrectly classified by both methods.

For each McNemar test, the null hypothesis was that the two methods have the same performance or error rate, i.e., $n_{12} = n_{21}$, versus the alternative hypothesis that they do not. The level of significance for each multiple comparison test was taken to be $\alpha = 0.05$ and so, using Bonferroni correction, the level of significance for each McNemar test was $\alpha = 0.05 / 5 = 0.01$ (The results for the GM tissue class for data set 3 are shown in Table 2 for illustrative purposes). Each McNemar test in essence tests whether the two methods classify in the same way; i.e., make the same misclassification errors.

Fig. 6 Data set 2: **a** Dice index and **b** Hausdorff distance (mm) for each tissue and method



Quantification of EEG Source Localization Performances

The EEG source localization performances of the proposed method HSA-BAMS and the reference method BET-FAST were evaluated quantitatively in terms of relative error (*RE*) of potential and localization error (*LE*) [1, 5, 7]. The relative error is defined:

$$RE = \frac{\|\mathbf{U}_{\text{measured}} - \mathbf{U}_{\text{estimated}}\|}{\|\mathbf{U}_{\text{measured}}\|} \quad (8)$$

where $\mathbf{U}_{\text{measured}}$ is a vector of the measured potential on the head and $\mathbf{U}_{\text{estimated}}$ is a vector of the potential generated from the simulated source. The localization error is defined:

$$LE = \|\mathbf{X}_{\text{true}} - \mathbf{X}_{\text{estimated}}\| \quad (9)$$

where \mathbf{X}_{true} is the real source position and $\mathbf{X}_{\text{estimated}}$ is the estimated source position.

Results and Discussion

Segmentation Evaluation for Data Set 1

The mean DI and mean *H* values over the five realizations of four different noise levels (with 0 % bias field) for each

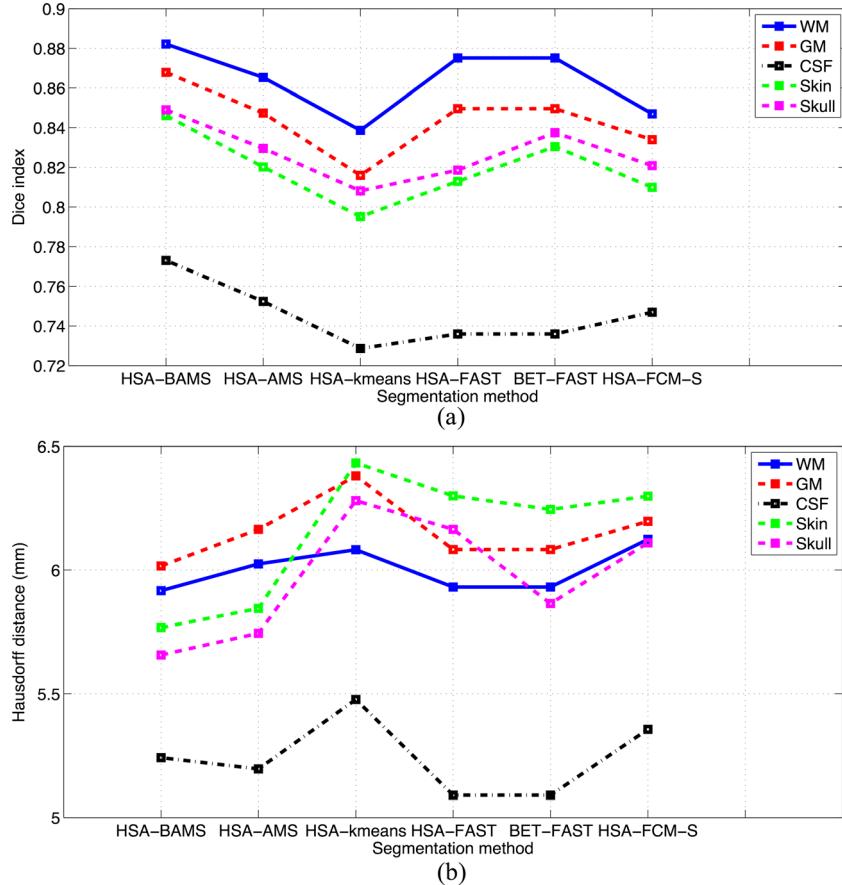
method and each tissue are shown in Figs. 2 and 3, respectively.

The mean DI values show that the segmentation performance of HSA-BAMS is consistently better than that of the HSA-FCM_S and HSA-FAST methods, and the reference method BET-FAST for all tissue types. They also show that HSA-BAMS is consistently better than the HSA-*k*-means and HSA-AMS methods for the WM, GM, fat, and muscle tissue and achieves essentially the same performance as these two methods for the CSF, skin, and skull.

The mean *H* values show that the segmentation performance of HSA-BAMS is consistently better than that of the HSA-FAST method and the reference method BET-FAST for all tissue types. They also show that HSA-BAMS performs consistently better than the HSA-*k*-means and HSA-FCM_S methods for the segmentation of WM, GM, fat, and muscle tissue at all noise levels and the HSA-AMS method for the segmentation of WM, GM, fat, and muscle tissue especially at higher noise levels (5, 7, and 9 %).

For each realization of each noise level, the McNemar tests provide evidence that the classification behavior of HSA-BAMS is different to that of HSA-*k*-means and HSA-AMS methods (*p* values <0.01) for WM, GM, fat, and muscle and also different to that of HSA-FCM_S, HSA-FAST, and BET-FAST methods for all tissue types. The tests do not provide

Fig. 7 Data set 3: **a** Dice index and **b** Hausdorff distance (mm) for each tissue and method



evidence that the classification behavior of HSA-BAMS is different to that of the HSA-*k*-means and HSA-AMS methods for the segmentation of CSF, skull, and skin tissue.

The mean DI and mean *H* values over five realizations of 5 % noise with 20 % bias field for each method and each tissue are shown in Fig. 4. The mean DI values for HSA-BAMS are higher (better) than that for the HSA-AMS, HSA-*k*-means, HSA-FCM_S, HSA-FAST and BET-FAST for all tissues except skin and skull where HSA-BAMS is on a par with HSA-AMS.

The proposed method HSA-BAMS has consistently lower (better) mean *H* values compared with the reference method BET-FAST and HSA-FAST and HSA-FCM_S methods for all tissue types. Compared with the HSA-*k*-means and HSA-AMS methods, HSA-BAMS has consistently lower mean *H* values for the segmentation of WM, GM, fat, and muscle tissue.

The McNemar tests provide evidence that HSA-BAMS performs differently (*p* values <0.01) to all the other methods for all tissue types. The tests do not provide evidence that the classification behavior of HSA-BAMS is different to that of the HSA-AMS method for the segmentation of skin tissue.

Figure 5 shows a segmentation example for each method for the 9 % noise level. It shows that HSA-BAMS is less sensitive to noise; in particular for the segmentation

of GM and WM tissue. It also shows that it is less likely to misclassify fat as muscle or skin tissue compared with all other methods.

Segmentation Evaluation for Data Set 2

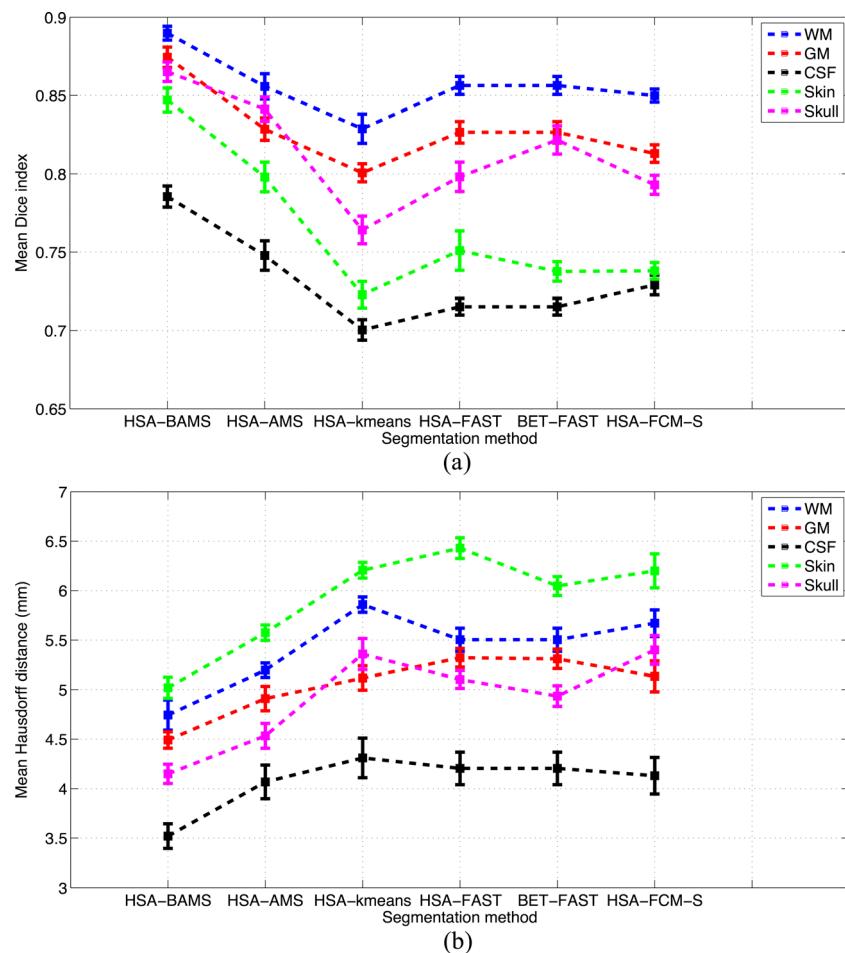
The DI and *H* values for each method and each tissue are shown in Fig. 6. The DI values show that the performance of HSA-BAMS is consistently better than that of the HSA-AMS, HSA-*k*-means, HSA-FCM_S, HSA-FAST, and BET-FAST methods for the GM, CSF, skin, and skull tissue. The *H* values show that the proposed method is consistently better (lower values) than all of the competing methods for the segmentation of WM, GM, skin, and skull tissue.

The McNemar tests provide evidence that HSA-BAMS performs differently (*p* values <0.01) to all other methods for all tissue types.

Segmentation Evaluation for Data Set 3

The DI and *H* values for each method and each tissue are shown in Fig. 7. The DI values show that the performance of HSA-BAMS is consistently better than that of all competing

Fig. 8 Data set 4: **a** Mean Dice index and **b** mean Hausdorff distance (mm) over the eight subjects for each tissue and method. The whiskers show ± 1 standard deviation



methods for all tissue types. The H values show that HSA-BAMS is better (lower H values) than all competing methods for the WM, GM, skin, and skull tissues.

The McNemar tests provide evidence that HSA-BAMS performs differently (p values <0.01) to all other methods for all tissue types.

Segmentation Evaluation for Data Set 4

The mean DI and mean H values over the eight subjects for each method and each tissue are shown in Fig. 8. The mean DI and mean H values show that the performance of HSA-BAMS is consistently better than that of the reference method BET-

FAST as well as that of the HSA-FAST, HSA- k -means, HSA-FCM_S, and HSA-AMS methods for all tissue types.

The McNemar tests for each subject provide evidence that HSA-BAMS performs differently (p values <0.01) to all other methods for all tissue types.

Figure 9 shows a segmentation example for each method for a single sagittal slice (face region has been excluded) from the subject (IXI040_Guys_0724) in data set 4. It shows that HSA-BAMS more accurately segments all of the tissue types than the competing methods compared with the ground truth. The HSA- k -means, HSA-FCM_S, HSA-FAST, and HSA-AMS methods over-segment the skull tissue while the BET-FAST method over-segments the skin tissue. Figure 10 also

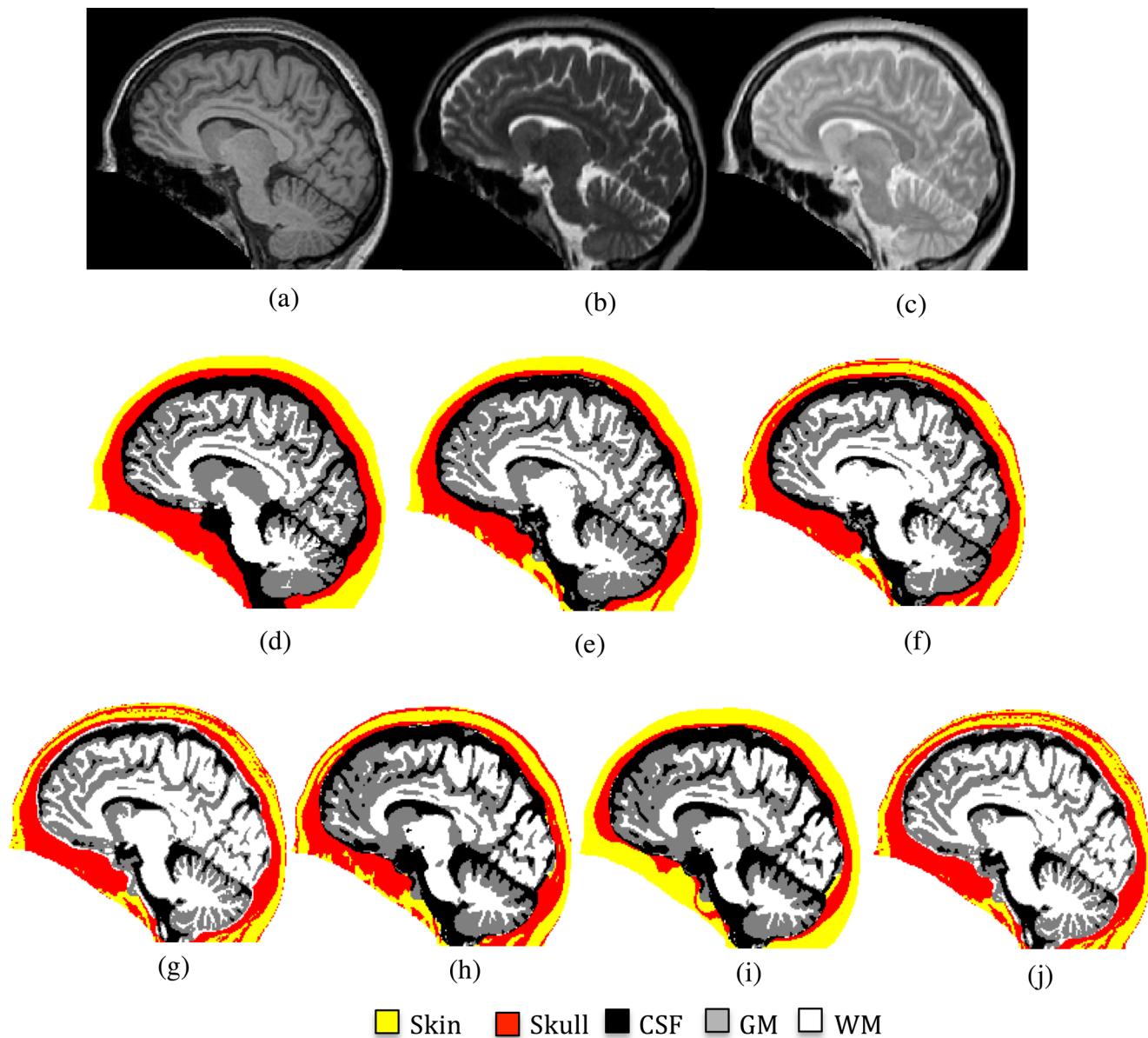


Fig. 9 Segmentation example from the subject (IXI040_Guys_0724) in data set 4 for a single sagittal slice **a** T1w image, **b** T2w image, **c** PD image, **d** ground truth, **e** HSA-BAMS, **f** HSA-AMS, **g** HSA- k -means, **h** HSA-FAST, **i** BET-FAST, **j** HSA-FCM_S (WM in white, GM in gray, CSF in black, skull in red, and skin in yellow)

suggests that all of the competing methods (HSA-*k*-means, HSA-FCM_S, HSA-FAST, HSA-AMS, and BET-FAST) have a higher tendency to over-segment the WM tissue compared with the proposed method HSA-BAMS.

Source Localization Evaluation

Figure 10 shows the location of (a) real (simulated) source (in green), (b) estimated source using the proposed method HSA-BAMS (in red), and (c) estimated source using the reference method BET-FAST (in blue), superimposed on the ground truth image. It can be seen that the estimated location of the source obtained using HSA-BAMS is in good agreement with the actual location of the source while that obtained using BET-FAST is not.

The RE of the potential and the LE for the proposed method HSA-BAMS are 0.01 and 0.0 mm, respectively, and for the reference method BET-FAST are 0.04 and 4.2 mm, respectively. They also show that HSA-BAMS yields better EEG source localization than BET-FAST.

Summary and Conclusion

In this paper, we presented a new fully automatic unsupervised method for head tissue segmentation from multi-modal MR images suitable for EEG source localization. In common with several existing approaches, it is a hierarchical segmentation approach wherein the MR data is first partitioned into

brain-tissue and non-brain-tissue sub-volumes and then each sub-volume is independently segmented into multiple tissue classes. What distinguishes our method is that a single segmentation approach, BAMS, is used to segment both the brain-tissue and non-brain-tissue sub-volumes into multiple tissue classes. The two main advantages of mean-shift segmentation are that it is an unsupervised technique (it does not require training examples) and that it can make use of multiple MRI modalities and indeed multiple other imaging modalities.

Several evaluations of the performance of the proposed method and of reference and variant methods were also presented. This included both direct evaluation in terms of segmentation accuracy and indirect evaluation in terms of EEG source localization accuracy. The direct evaluation results, based on both synthetic data and real data from ten subjects, show that the proposed method generally performs better than the competing methods and is more tolerant to the noise and the bias field. Importantly, for the real data, the proposed method, HSA-BAMS, outperforms the reference method BET-FAST and variations on the proposed method for all tissue types considered important for the EEG source localization problem. The EEG source localization results show that the proposed method outperforms the reference method BET-FAST commonly used for the construction of a realistic head model for EEG source localization. Overall, the experimental results suggest that the proposed method HSA-BAMS can be used as a surrogate for the reference method BET-FAST, as well as manual segmentation for the construction of patient-specific head models for EEG source localization.

Acknowledgments The author Mr. Mahmood acknowledges scholarship funding from the Higher Education Commission of Pakistan (HEC) and Chalmers University of Technology in support of this work.

Appendix

A.1 Bayesian-Based Adaptive Bandwidth Estimator

The bandwidth is modeled [24] by the a posteriori probability density function $p(s|\mathbf{x})$ of local data spread or variance s given the data (feature) point \mathbf{x} . Let $M < n$ (total number of data points) be the number of nearest neighbors to a data sample \mathbf{x}_i . We can then define the pseudolikelihood

$$P\left(s \mid \mathbf{x}\right) = \prod_{j=1}^N P\left(s \mid \mathbf{x}_{M_j}\right) \quad (10)$$

where $P\left(s \mid \mathbf{x}_{M_j}\right)$ is the probability of local data spread s depending on the M_j nearest neighborhood samples to \mathbf{x}_{M_j} and $\{M_j\}_{j=1,\dots,N}$ is the set of N such neighborhoods of various

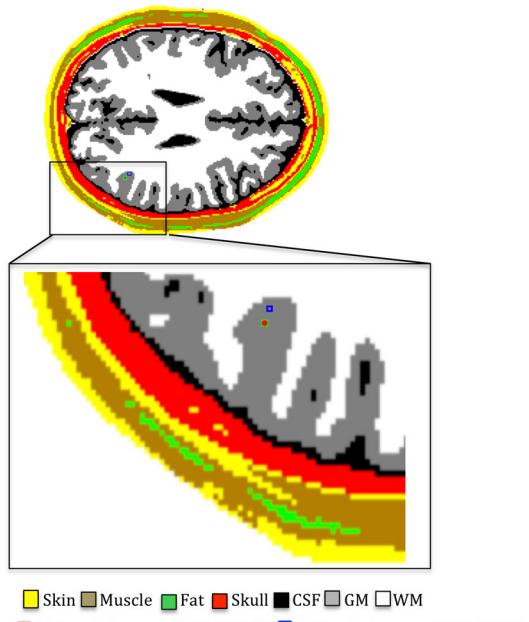


Fig. 10 EEG source localization results: ground truth superimposed with the location of the real (simulated) source (green), HSA-BAMS-estimated source (red), and BET-FAST-estimated source (blue)

sizes. The evaluation of these probabilities over the entire set of M_j is then given by

$$P\left(s \mid \mathbf{x}_{M_j}\right) = \int P\left(s \mid M_j, \mathbf{x}_{M_j}\right) P\left(M_j \mid \mathbf{x}_{M_j}\right) dM_j \quad (11)$$

Applying Bayes rule we get

$$P\left(M_j \mid \mathbf{x}_{M_j}\right) = \frac{P\left(\mathbf{x}_{M_j} \mid M_j\right) P(M_j)}{P(\mathbf{x}_{M_j})} \quad (12)$$

where $P(\mathbf{x}_{M_j} \mid M_j)$ is the probability of the data sample \mathbf{x}_{M_j} given the M_j nearest neighborhood. Hereinafter, $P(M_j)$ is considered to have uniform distribution on the interval $[M_1, M_2]$. Several values are selected for M_j in this interval according to

$$M_j = M_1 + j \frac{M_2 - M_1}{N} \quad (13)$$

For a given M_j , the local variance s_j is computed as

$$s_j = \frac{\sum_{l=1}^{M_j} \|\mathbf{x}_{i,l} - \mathbf{x}_i\|^2}{M_j - 1} \quad i = 1, 2, \dots, n \quad j = 1, 2, \dots, N \quad (14)$$

where $\mathbf{x}_{i,l}$ is the l th nearest neighbor to the data point \mathbf{x}_i . The distribution of variances is modeled as the Gamma distribution defined as

$$p(s \mid \alpha, \beta) = \frac{\beta^\alpha s^{\alpha-1}}{\Gamma(\alpha)} e^{-\beta s} \quad s \geq 0 \quad \alpha, \beta > 0 \quad (15)$$

where

$$\Gamma(t) = \int_0^\infty r^{t-1} e^{-r} dr \quad (16)$$

is the Gamma function, and α and β define the shape and the scale of the Gamma distribution, respectively.

These parameters are estimated using the maximum likelihood approach [24]. The estimate of the adaptive bandwidth is identically the mean of this distribution, i.e.,

$$\hat{h}(\mathbf{x}_i) = \hat{\alpha}\hat{\beta} \quad i = 1, 2, \dots, n \quad (17)$$

References

- Shirvany Y: Non-invasive EEG functional neuroimaging for localizing epileptic brain activity, PhD dissertation, ISBN: 978-91-7385-810-6, Sweden: Chalmers University of Technology, 2013
- Grech R, Cassar T, Muscat J, Camilleri K, Fabri SG, Zervakis M, Xanthopoulos P, Sakkalis V, Vanrumste B: Review on solving the inverse problem in EEG source analysis. *J Neuro Eng Rehab* 5:25, 2008
- Shirvany Y, Porras AR, Kowkabzadeh K, Mahmood Q, Lui H-S, Persson M: Investigation of brain tissue segmentation error and its effect on EEG source localization, *Conf. Proc. IEEE Eng Med Biol Soc (EMBS)*, 2012, 1522–25
- Wolters CH, Anwander A, Tricoche X, Weinstein D, Koch MA, MacLeod RS: Influence of tissue conductivity anisotropy on EEG/MEG field and return current computation in a realistic head model: a simulation and visualization study using high-resolution finite element modeling. *Neuroimage* 54:813–26, 2006
- Shirvany Y, Edelvik F, Jakobsson S, Hedström A, Mahmood Q, Chodorowski A, Persson M: Non-invasive EEG source localization using particle swarm optimization: a clinical experiment, *Conf. Proc. IEEE Eng Med Biol. Soc (EMBS)*, 2012, 6232–5
- Ramon C, Schimpf PH, Haueisen J: Influence of head models on EEG source localizations and inverse source localizations. *Biomed Eng Online* 5:10, 2006. doi:10.1186/1475-925X-5-10
- Shirvany Y, Mahmood Q, Edelvik F, Persson M, Hedstrom A, Jakobsson S: Particle swarm optimization applied to EEG source localization of somatosensory evoked potentials. *IEEE Trans Neural Syst Rehab Eng* 11:20, 2013. doi:10.1109/TNSRE.2013.2281435
- Yvert B, Bertrand O, Echallier J: Improved forward EEG calculations using local mesh refinement of realistic head geometries. *Electroencephalogr Clin Neurophysiol* 5:381–392, 1995
- Heinonen T, Eskola H: Segmentation of T1 MR scans for reconstruction of resistive head models. *Comput Methods Prog Biomed* 54: 173–81, 1997
- Heinonen T, Dastidar P, Frey F, Eskola H: Applications of MR image segmentation. *Int J Bioelectromagnet* 1:5–39, 1999
- Acar ZA, Makeig S: Neuroelectromagnetic forward modeling toolbox. *J Neurosci Methods* 190:258–270, 2010
- Rullmann M, Anwander A, Dannhauer M, Warfield S, Duffy F, Wolter CH: EEG source analysis of epileptiform activity using a 1 mm anisotropic hexahedra finite element head model. *Neuroimage* 44:399–410, 2009
- Lanfer B, Scherg M, Dannhauer M, Knösche TR, Burger M, Wolters CH: Influences of skull segmentation inaccuracies on EEG source analysis. *Neuroimage* 62:418–31, 2006
- Smith SM: Fast robust automated brain extraction. *Hum Brain Mapp* 17:143–155, 2002
- Otsu N: A threshold selection method from gray level histogram. *IEEE Trans Systems Man Cybernet* 9:62–66, 1979
- Soille P: Morphological Image Analysis: Principles and Applications. Springer-Verlag 173–174, 1999
- Zhang Y, Brady M, Smith S: Segmentation of brain MR images through a hidden Markov random field model and the expectation maximization algorithm. *IEEE Trans Med Imag* 20:45–57, 2001
- Mayer A, Greenspan H: An adaptive mean-shift framework for MRI brain segmentation. *IEEE Trans Med Imag* 28:1238–1249, 2009
- Wen Y, He L, von Deneen KM, Lu Y: Brain tissue classification based on DTI using an improved fuzzy C-means algorithm with spatial constraints. *Magn Reson Imaging* 31(9):1623–30, 2013
- Seber GAF: Multivariate Observations: Hoboken. Wiley, NJ, 1984
- Jenkinson M, Beckmann CF, Behrens TE, Woolrich MW: Smith SM: FSL. *Neuroimage* 62:782–790, 2012
- Comaniciu D, Meer P: Mean shift: a robust approach toward feature space analysis. *IEEE Trans Pattern Anal Mach Intell* 24:603–619, 2002
- Comaniciu D, Ramesh V, Meer P: The variable bandwidth mean-shift and data-driven scale selection, *ICCV* 438–445, 2001
- Bors AG, Nasios N: Kernel bandwidth estimation for nonparametric modeling. *IEEE Trans Systems Man Cybernet* 39:1543–1555, 2009
- Cocosco CA, Kollokian V, Kwan RK-S, Evans AC: BrainWeb: online interface to a 3-D MRI simulated brain database. *Neuroimage* 5:S425, 1997

26. BrainWeb. Available at: http://brainweb.bic.mni.mcgill.ca/brainweb/about_sbd.html. Accessed August 2013.
27. ITK-SNAP software. Available at: <http://www.itksnap.org>. Accessed August 2013.
28. IXI datasets. Available at: <http://www.brain-development.org>. Accessed Sept. 2013.
29. Wagner M: Rekonstruktion neuronaler Ströme aus bioelektrischen und biomagnetischen Messungen auf der aus MR-Bildern segmentierten Hirnrinde, PhD thesis, ISBN:3-8265-4293-2, Shaker-Verlag Aachen 1998
30. Jenkinson M, Bannister PR, Brady JM, Smith SM: Improved optimisation for the robust and accurate linear registration and motion correction of brain images. *Neuroimage* 17:825–841, 2002
31. Dice LR: Measures of the amount of ecologic association between species. *Ecology* 26:297–302, 1945
32. Babalola KO, Patenaude B, Aljabar P, Schnabel J, Kennedy D, Crum W, Smith S, Cootes TF, Jenkinson M, Rueckert D: Comparison and evaluation of segmentation techniques for subcortical structures in brain MRI, *Med Image Comput Assist Interv (MICCAI)*, 2008, 409–416
33. Dietterich TG: Approximate statistical tests for comparing supervised classification learning algorithms. *Neural Comput* 10:1895–19, 1998