Fully Automatic Segmentation of the Hippocampus and the Amygdala from MRI Using Hybrid Prior Knowledge

Marie Chupin[1,2], Alexander Hammers[3], Eric Bardinet[2], Olivier Colliot[2], Rebecca S.N. Liu[4], John S. Duncan[1], Line Garnero[2], and Louis Lemieux[1]

[1] Department of Clinical and Experimental Epilepsy, IoN, UCL, London, UK
m.chupin@ion.ucl.ac.uk, l.lemieux@ion.ucl.ac.uk*
[2] Cognitive Neuroscience and Brain Imaging, CNRS UPR640, UPMC, Paris, France
[3] Faculty of Medicine, ICL, London, UK
[4] National Hospital for Neurology and Neurosurgery, UCLH, London, UK

Abstract. The segmentation of macroscopically ill-defined and highly variable structures, such as the hippocampus *Hc* and the amygdala *Am*, from MRI requires specific constraints. Here, we describe and evaluate a hybrid segmentation method that uses knowledge derived from a probabilistic atlas and from anatomical landmarks based on stable anatomical characteristics of the structures. Combined in a previously published semi-automatic segmentation method, they lead to a fast, robust and accurate fully automatic segmentation of *Hc* and *Am*. The probabilistic atlas was built from 16 young controls and registered with the "unified segmentation" of SPM5. The algorithm was quantitatively evaluated with respect to manual segmentation on two MRI datasets: the 16 young controls, with a leave-one-out strategy, and a mixed cohort of 8 controls and 15 subjects with epilepsy with variable hippocampal sclerosis. The segmentation driven by hybrid knowledge leads to greatly improved results compared to that obtained by registration of the thresholded atlas alone: mean overlap for *Hc* on the 16 young controls increased from 78% to 87% ($p < 0.001$) and on the mixed cohort from 73% to 82% ($p < 0.001$) while the error on volumes decreased from 10% to 7% ($p < 0.005$) and from 18% to 8% ($p < 0.001$), respectively. Automatic results were better than the semi-automatic results: for the 16 young controls, average overlap increased from 84% to 87% ($p < 0.001$) for *Hc* and from 81% to 84% ($p < 0.002$) for *Am*, with equivalent improvements in volume error.

1 Introduction

Volumetric analyzes of brain structures can inform on mechanisms underlying disease progression. The hippocampus *Hc* and the amygdala *Am* are of major interest, due to their implication in epilepsy and Alzheimer's disease. Manual volume measurement still remains the norm, making large cohort studies difficult. Poor boundary definition makes segmentation of these two grey matter

* MC was funded by a European post doctoral Marie Curie fellowship.

N. Ayache, S. Ourselin, A. Maeder (Eds.): MICCAI 2007, Part I, LNCS 4791, pp. 875–882, 2007.

structures from Magnetic Resonance Imaging (MRI) scans challenging. Prior knowledge from anatomical atlases is necessary to their coherent manual delineation. For automation, this knowledge can come from statistical information on shape [1][2] or deformations [3][4], which may not be suitable for diseased structures. Registering an atlas template derived from a single subject [5][6] has been shown to be influenced by the choice of the atlas [7], even if combined registration-segmentation methods [8] may prove less sensitive. Registration and segmentation using probabilistic information [9][10][11] offer more thorough global spatial knowledge. It is complementary to anatomical knowledge [12] [13], which formalizes stable global and local anatomical relationships.

Fully automatic, fast and robust segmentation of healthy and pathological hippocampi and amygdale suitable for routine use has yet to be demonstrated. On the one hand, segmentation methods based on probabilistic information require high dimensional deformations [10], in order to achieve precise extraction. On the other hand, methods based mainly on image intensity can be fast, but need a good initialisation to be accurate. We describe a new fully automatic hybrid segmentation algorithm that combines the two methods by using global spatial knowledge from a probabilistic atlases within a previously published semi-automatic algorithm driven by anatomical knowledge [13]. The new algorithm's performance is evaluated on data from healthy controls and a mixed cohort including subjects with hippocampal sclerosis associated with chronic epilepsy.

2 Method

The method is based on iterative two-object competitive deformation [13]. Two regions deform following local topology-preserving transformations from two initial objects, within an extracted bounding box embedding *Hc* and *Am*. Bounding box and initial objects are automatically retrieved from probabilistic atlases. Here, the deformation process is constrained by hybrid prior knowledge derived from probabilistic atlases and neuroanatomical landmarks automatically detected during the deformation. The process is driven by minimizing a global energy functional in a Markovian framework. A competitive scheme allows the identification of the partly visible *Hc*-*Am* interface. At each iteration, the energy functional is modified according to probabilistic and anatomical likelihood.

2.1 Construction of the *Am* and *Hc* Probabilistic Atlases

MRI data from 16 young controls (S1-S16) were manually segmented by an expert following a 3D protocol [13], ensuring 3D-coherent structures. This results in a set of 32 binary labeled datasets, 16 with both *Hc* $\{Hc_i, i = 1...16\}$ and 16 with both *Am* $\{Am_i, i = 1...16\}$. The transformation from native to MNI standard space, $\{T_i, i = 1...16\}$, is computed with the unified segmentation module of SPM5 [14], which allows to iteratively optimize registration parameters (linear combination of cosine transformation bases), tissue classification, and intensity

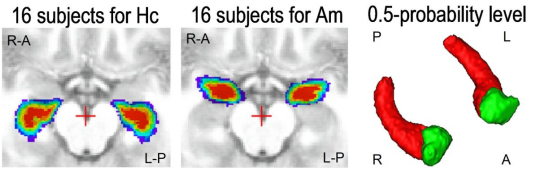

Fig. 1. Axial section showing probabilistic atlases for Hc and Am and 3D renderings of the 0.5-level of the probability maps (Hc in red, Am in green)

non-uniformity correction. The SPM5 default parameters are used. The probabilistic atlases PA_{Hc} and PA_{Am} (Fig 1) are created as follows:

$$\forall v \in \Omega, PA_{Hc}(v) = \frac{1}{16}\sum_{i=1}^{16} T_i(Hc_i)(v) \; and \; PA_{Am}(v) = \frac{1}{16}\sum_{i=1}^{16} T_i(Am_i)(v) \quad (1)$$

where v is a voxel in the MRI set Ω. $PA_{Hc}(v)$, resp. $PA_{Am}(v)$, is the probability that v belongs to Hc, resp. Am, according to the probabilistic atlas prior.

2.2 Initialization of the Deformation

For a given subject, individual atlases IPA_{Hc} and IPA_{Am} are created by back registering the atlases PA_{Hc} and PA_{Am} to the subject's space, using the inverse deformation $\{(T_i,)^{-1}, i = 1...16\}$ computed by the SPM5 unified segmentation.

Bounding box: The bounding box BB_{HcAm} must fully embed Hc and Am but is not used as a geometrical constraint. It is defined as the smallest parallelepiped subvolume in Ω around the non-null probability object $HcAm^{min} = [v \in \Omega, \; IPA_{Hc}(v) > 0 \; or \; IPA_{Am}(v) > 0]$ (Fig 2).

Initial objects: The initial object Hc^{init}, resp. Am^{init}, is created from the maximum probability object Hc^{max}, resp. Am^{max}. Hc^{max} is defined as the 1-level of the probability map IPA_{Hc}. It is built by keeping the voxels for which the probability equals one, while regularising to prevent holes ($IPA_{Hc}(v) < 1$ but v is "inside" Hc^{max}) and wires ($IPA_{Hc}(v) = 1$ but v "spikes" from Hc^{max})

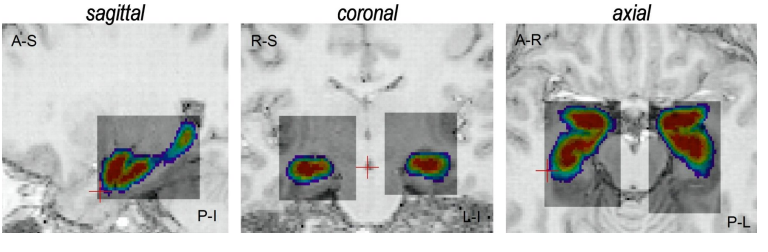

Fig. 2. Bounding box extraction, illustrated on sagittal, coronal and axial slices

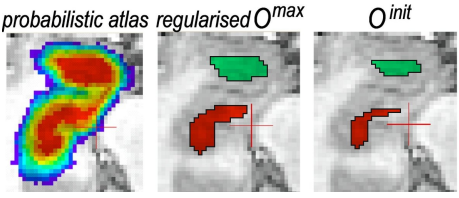

Fig. 3. Atlas, regularised object and initial object, for Hc and Am (axial slice)

to appear. Let $N^{6N}_{Hc^{max}}$ be the number of 6-neighbours of v labelled in Hc^{max}. If $N^{6N}_{Hc^{max}}$ is larger than 3, rejecting v from Hc^{max} will result in a hole in Hc^{max}; if $N^{6N}_{Hc^{max}}$ is smaller than 1, including v in Hc^{max} will result in a wire. Combining regularity rule and probability threshold iteratively, we get:

$$\begin{cases} [Hc^{max}]^0 = [v \in BB_{HcAm}, IPA_{Hc}(v) > 0] \\ [Hc^{max}]^i = \begin{bmatrix} IPA_{Hc}(v) = 1 \; and \; N^{6N}_{[Hc^{max}]^{i-1}}(v) \geq 1 \\ or \; IPA_{Hc}(v) \neq 1 \; and \; N^{6N}_{[Hc^{max}]^{i-1}}(v) \geq 3 \end{bmatrix} \end{cases}. \tag{2}$$

The iterations proceed until Hc^{max} (resp. Am^{max}) remains unchanged. Hc^{max} (resp. Am^{max}) is then eroded with a 1mm-structuring element, and the largest connected component is kept to create the initial object Hc^{init} (resp. Am^{init}) (Fig 3). The erosion step is introduced to increase robustness in case of atrophy.

2.3 Introduction of Hybrid Prior Knowledge

The regularisation term of the energy functional in [13] is modified to take into account the probability of the voxel v to belong to the deforming object O (Hc or Am), derived from IPA_{Hc} and IPA_{Am}. This term is locally expressed as the comparison of the number of O-labelled neighbours of v, $N_O(v)$, and a standard number of neighbours $\widetilde{N}$, with σ^I a standard deviation around $\widetilde{N}$:

$$E^I_O = \left(\frac{\widetilde{N} - \gamma^{PZ}_O(v)\gamma^{AZ}_O(v)N_O(v)}{\sigma^I} \right)^5. \tag{3}$$

The γ parameters influence the classification according to prior probabilities for v to belong to O; for γ superior to 1, $N_O(v)$ is artificially increased, decreasing the energy, and vice versa for γ inferior to 1. γ^{AZ}_O models anatomical zones AZ defined by the anatomical landmarks ($\gamma^{AZ}_O(v) = 2$ if v is likely for O and 0.5 if v is unlikely for O). γ^{PZ}_O models probability zones PZ given by IPA_O: $\gamma^{PZ}_O(v) = 0.75$ if $IPA_O(v) = 0$, $\gamma^{PZ}_O(v) = 1.5$ if $0.75 \leq IPA_O(v) < 1$ and $\gamma^{PZ}_O(v) = 2$ if $IPA_O(v) = 1$; otherwise, $\gamma^{PZ}_O = 1$. Values are chosen empirically.

2.4 Atlas Mismatch: Automatic Detection and Correction Strategies

Atlas-based segmentation methods face a common problem: a possible initial mismatch of the atlas. An automatic strategy is used to detect such occurrences,

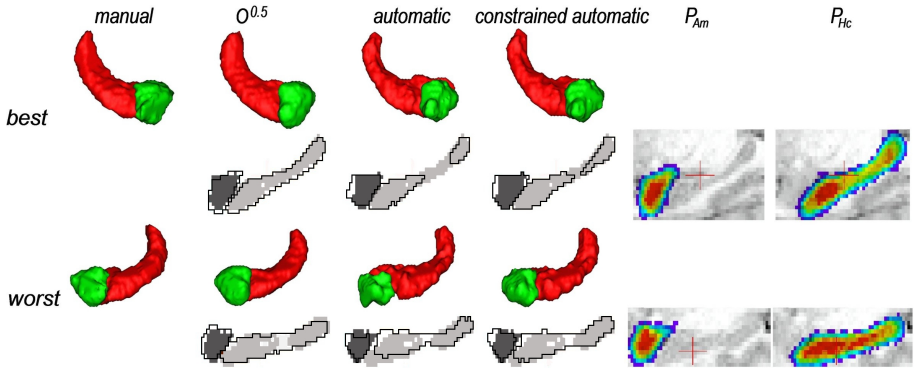

Fig. 4. 3D-renderings of automatic and manual segmentations, and overlap between segmentations (manual segmentations in shades of grey) and probabilistic atlases

based on two successive tests comparing intensity distributions for the grey matter (GM) and for the 0.5-level object for Hc: $Hc^{0.5} = [v \in \Omega,\ IPA_{Hc}(v) \geq 0.5]$. The first test detects cases when the deformed atlas fails to match the Hc sclerosis. It compares the average intensity on $Hc^{0.5}$ to that of GM; the assumption is that an overestimated $Hc^{0.5}$ when Hc is sclerotic will include cerebrospinal fluid dark voxels. The second test detects atlas misalignment in the region of interest. It compares the standard deviation on $Hc^{0.5}$ to that on an eroded version of $Hc^{0.5}$. The assumption is that a misaligned $Hc^{0.5}$ will include voxels of several tissues; for large misplacements, erosion will not reduce the standard deviation. An automatic correction strategy was introduced: if the atlas IPA_{Hc} is misaligned, the correction relies on decreasing the probabilistic constraint; if IPA_{Hc} is overestimated, the correction relies on shrinking IPA_{Hc} by erosion.

3 Performance Evaluation

The impact of the new automatic initialisation process and probabilistic atlas constraint on segmentation performance was evaluated with qualitative and quantitative comparisons between automatic (without and with atlas constraint), semi-automatic (with manual initialisation [13]) and manual segmentations [13] together with comparisons to the 0.5-level object derived from the registered atlas in subject's space ($Hc^{0.5}$ and $Am^{0.5}$), as a simple atlas-based segmentation.

3.1 Validation Data

The atlas is created using the 16 controls data (S1-S16 included in [13]), acquired in the axial plane; a leave-one-out procedure is followed, the atlas being derived from the remaining 15 subjects for each subject. Data from 23 subjects (mixed cohort), acquired in the coronal plane [15] were split into 3 groups: 8 healthy controls (C1-C8, Hc volume: $2.9 \pm 0.5cm^3(1.8 - 3.6)$), 8 subjects with epilepsy

Table 1. Segmentation performance index values in data from 16 healthy controls (top) and from the 23 subjects in the mixed cohort (bottom)

		semi-auto	$0^{0.5}$	automatic	constrained	corrected
16 young controls						
Hc	*RV*	7 ± 4(0-14)	10 ± 6(1-26)	9 ± 6(0-25)	7 ± 4(0-15)	7 ± 4(0-15)
	K	84 ± 3(78-89)	78 ± 4(64-84)	82 ± 4(74-89)	87 ± 2(80-90)	87 ± 2(80-90)
	MIV	1.1 ± 1(0-4)	0.8 ± 1(0-5)	1.6 ± 2.3(0-9)	0.8 ± 0.8(0-4)	0.8 ± 0.8(0-4)
	DM	4.5 ± 1.5(3-9)	4 ± 1.3(3-9)	5.1 ± 2.4(2-15)	3.5 ± 1.2(2-8)	3.5 ± 1.2(2-8)
Am	*RV*	12 ± 7(1-27)	10 ± 8(0-33)	14 ± 9(1-35)	10 ± 6(0-26)	10 ± 6(0-26)
	K	81 ± 4(69-88)	82 ± 4(70-89)	77 ± 6(62-86)	84 ± 4(76-91)	84 ± 4(76-91)
	MIV	1.5 ± 1(0-4)	1.9 ± 2.2(0-9)	1.1 ± 0.7(0-2)	1.1 ± 1(0-5)	1.1 ± 1(0-5)
	DM	3.9 ± 0.9(3-6)	2.8 ± 0.5(2-4)	4.5 ± 1.1(3-7)	3.3 ± 0.7(2-7)	3.3 ± 0.7(2-7)
mixed cohort						
Hc	*RV*		18 ± 15(0-74)	25 ± 39(0-200)	10 ± 11(0-64)	9 ± 7(0-33)
	K		73 ± 11(41-86)	70 ± 18(0-87)	81 ± 8(59-89)	82 ± 6(64-89)
	MIV		2.3 ± 4.7(0-28)	11 ± 14(0-68)	3.5 ± 8(0-48)	3.0 ± 5.6(0-33)
	DM		4.5 ± 2.9(2-20)	10 ± 5.5(3-27)	5.3 ± 2.9(3-14)	5.1 ± 2.6(3-14)
Am	*RV*		15 ± 10(0-44)	32 ± 42(0-200)	20 ± 14(1-58)	20 ± 13(1-54)
	K		75 ± 8(35-86)	63 ± 18(0-85)	77 ± 9(34-88)	77 ± 7(50-88)
	MIV		2.9 ± 3.4(0-13)	1.2 ± 0.9(0-10)	1.8 ± 2.3(0-10)	1.8 ± 2.3(0-10)
	DM		2.8 ± 3.6(2-7)	6.2 ± 2.7(3-16)	3.7 ± 0.9(2-6)	3.7 ± 0.9(2-6)

and known *Hc* sclerosis (HS1-HS8, $2.0 \pm 0.8 cm^3 (0.7 - 3.5)$), 7 subjects with temporal lobe epilepsy and normal *Hc* volumes (TL1-TL7, $2.6 \pm 0.5 cm^3 (1.6 - 3.4)$)). All datasets were manually segmented according to the same protocol as that used to create the atlas, by the same investigator. All processing apart from registration is run within the Anatomist software environment [16]. The whole segmentation procedure for both *Hc* and *Am* requires around 15min. Four indices are used to compare the segmentation S with the reference R [13]: $RV(S,R) = 2.|V_S - V_R|/(V_S + V_R)$, the error on volumes; $K(S,R) = 2.V_{S \cap R}/(V_S + V_R)$, the Dice overlap; $MIV(S_1, R_1, R_2) = 2.V_{S_1 \cap R_2}/(V_{S_1} + V_{R_1})$, the missclassified interface voxels and $DM(S,R) = \max[\max_{v \in \hat{S}}(d(v, \hat{R})), \max_{v \in \hat{R}}(d(v, \hat{S}))]$, the symmetric maximal distance, with $\hat{S}$ the 6-connectivity border of S. Significance for the variation of these values was tested using a two-tailed Student's t-test.

3.2 Validation Results in Young Healthy Controls

The segmentation results were found to be qualitatively correct. The two cases chosen for illustration (Fig 4) are those with the best and worst results in [13]. Table 1, top rows, summarizes the quantitative results. The automatic initialisation gives better results than the registered 0.5-level object, but inferior to the semi-automatic results [13]; the automatic results with atlas constraint are better than those from all other methods. No need for correction was detected.

3.3 Evaluation on Data from Mixed Cohort

Quantitative comparisons between automatic and manual segmentations (Table 1, bottom rows), show the importance of the atlas constraint on the results, even if the automatic results for Am were not significantly better than $Am^{0.5}$. Occasional initial atlas misalignments were observed; most were not detected as atlas mismatch but were nonetheless successfully corrected in the course of the automatic segmentation. Atlas mismatch was detected in 3 Hc out of 78, as illustrated in Fig. 5: atlas overestimation for two highly sclerotic Hc and atlas misalignment in another one. Segmentation failures were prevented with the correction strategy, as shown in the right column of Table 1. Note that the worst value for RV (33%) in our method was obtained for the smallest Hc in the group studied according to the manually estimated volume. $0.7cm^3$.

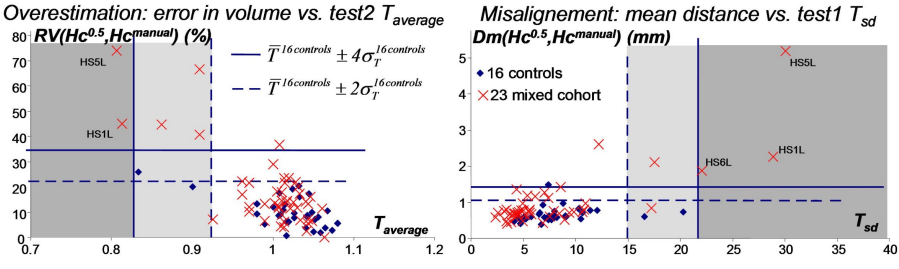

Fig. 5. Qualitative indices vs. detection test characterizing atlas mismatch (defined vs. the test value average on the 16 young controls with a range of 4 standard deviations).

4 Discussion and Conclusion

The combination of probabilistic knowledge and anatomical prior knowledge within the competitive deformation algorithm [13] allows accurate fully automatic segmentation of Hc and Am on healthy controls. The automatic detection and correction of initial atlas mismatch resulted in highly encouraging Hc segmentation on data from a representative group of subjects with epilepsy, including cases with high degrees of hippocampal sclerosis.

The fully automatic method performed better than both semi-automatic segmentation and 0.5-level probability object. For controls, the segmentation results (a K value of 87% for Hc and 84% for Am) compare favourably with the literature. In fact, K values are 83% for Hc [5] after manual placement of 28 landmarks, or 80% for Hc and 65% for Am [9], with similar errors on volumes. Using a method that requires the placement of 50 manual landmarks, a K of 88% for Hc was obtained [2]. The results for an entirely probabilistic method were 82% for Hc and 81% for Am [10]. Moreover, our results on patient data compare favourably with published results in subjects with epilepsy. Values for the 9 sclerotic Hc in HS1-8 (average volume: $1.4cm^3$ (0.7-2)) are K=76% (64-83) and RV=16% (4-33) for Hc, while in [5], they were, for 5 sclerotic Hc (average

volume: $1.3cm^3$ (1.2-1.4), K=67% (57-75) and RV=16% (6-19). A recent study in sclerotic Hc ($1.2cm^3$ (1.1-1.6)) achieved similar overlap (K=76% (71-83)) but at the expense of about 2 weeks of CPU time [17].

References

1. Kelemen, A., et al.: Elastic model-based segmentation of 3-D neuroradiological data sets. IEEE TMI 18, 828–839 (1999)
2. Shen, D., et al.: Measuring size and shape of the hippocampus in MR images using a deformable shape model. Neuroimage 15, 422–434 (2002)
3. Duchesne, S., et al.: Appearance-based segmentation of medial temporal lobe structures. Neuroimage 17, 515–531 (2002)
4. Klemenčič, J., et al.: Non-rigid registration based active appearance models for 3D medical image segmentation. J. Imag. Sci. and Tech. 48(2), 166–171 (2004)
5. Hogan, R., et al.: Mesial temporal sclerosis and temporal lobe epilepsy: MR imaging deformation-based segmentation of the hippocampus in five patients. Radiology 216, 291–297 (2000)
6. Dawant, B., et al.: Automatic 3-D segmentation of internal structures of the head in MR images using a combination of similarity and free-form transformation. IEEE TMI 18, 909–916 (1999)
7. Carmichael, O., et al.: Atlas-based hippocampus segmentation in Alzheimer's disease and mild cognitive impairment. Neuroimage 27(4), 979–990 (2005)
8. Wang, F., et al.: Joint registration and segmentation of neuroanatomic structures from brain MRI. Academic Radiology 13, 1104–1111 (2006)
9. Fischl, B., et al.: Whole brain segmentation: Automated labelling of neuroanatomical structures in the human brain. Neuron. 33, 341–355 (2002)
10. Heckemann, R., et al.: Automatic anatomical brain MRI segmentation combining label propagation and decision fusion. Neuroimage 33, 115–126 (2006)
11. Pitiot, A., et al.: Expert knowledge-guided segmentation system for brain MRI. Neuroimage 23, S85–S96 (2004)
12. Bloch, I., et al.: Fusion of spatial relationships for guiding recognition, example of brain structure recognition in 3D MRI. Pattern Recogn. Let. 26, 449–457 (2005)
13. Chupin, M., et al.: Automated segmentation of the hippocampus and the amygdala driven by competition and anatomical priors: Method and validation on healthy subjects and patients with Alzheimer's disease. Neuroimage 34, 996–1019 (2007)
14. Ashburner, J., Friston, K.: Unified segmentation. Neuroimage 26, 839–851 (2005)
15. Liu, R., et al.: A longitudinal quantitative MRI study of community-based patients with chronic epilepsy and newly diagnosed seizures: Methodology and preliminary findings. Neuroimage 14, 231–243 (2001)
16. Rivière, D., et al.: A structural browser for human brain mapping. In: Human Brain Mapping p. 912 (2000)
17. Hammers, A., et al.: Automatic detection and quantification of hippocampal atrophy on MRI in temporal lobe epilepsy: A proof-of-principle study. Neuroimage 36, 38–47 (2007)