Hierarchical clustering to measure connectivity in fMRI resting-state data

Dietmar Cordes[a,*], Vic Haughton[b], John D. Carew[c], Konstantinos Arfanakis[d], Ken Maravilla[a]

[a]Department of Radiology, University of Washington, Seattle, Washington, USA
[b]Department of Radiology, University of Wisconsin, Madison, Wisconsin, USA
[c]Department of Statistics, University of Wisconsin, Madison, Wisconsin, USA
[d]Department of Medical Physics, University of Wisconsin, Madison, Wisconsin, USA

Abstract

Low frequency oscillations, which are temporally correlated in functionally related brain regions, characterize the mammalian brain, even when no explicit cognitive tasks are performed. Functional connectivity MR imaging is used to map regions of the resting brain showing synchronous, regional and slow fluctuations in cerebral blood flow and oxygenation. In this study, we use a hierarchical clustering method to detect similarities of low-frequency fluctuations. We describe one measure of correlations in the low frequency range for classification of resting-state fMRI data. Furthermore, we investigate the contribution of motion and hardware instabilities to resting-state correlations and provide a method to reduce artifacts. For all cortical regions studied and clusters obtained, we quantify the degree of contamination of functional connectivity maps by the respiratory and cardiac cycle. Results indicate that patterns of functional connectivity can be obtained with hierarchical clustering that resemble known neuronal connections. The corresponding voxel time series do not show significant correlations in the respiratory or cardiac frequency band. © 2002 Elsevier Science Inc. All rights reserved.

Keywords: Functional imaging; Resting-state; Physiological fluctuations; Clustering

1. Introduction

Low frequency fluctuations of the cerebral hemodynamics with periods larger than 10 s have been observed in cortical and subcortical gray matter in the human brain with different imaging modalities such as fMRI [1–4], near-infrared spectroscopy [5] and transcranial Doppler sonography [6]. Regions of the brain with similar functions, even when separated by a large distance have low frequency physiological fluctuations that show a zero phase relationship, implying a high degree of temporal coherence [7–8]. This synchronicity of low frequency fluctuations in functionally related regions suggests the existence of neuronal connections characterizing a widespread cortical network. The underlying mechanism of this phase-locking behavior is still not understood. Some investigators hypothesize that synchronization between functionally connected regions might lead to a computational advantage in large distributed networks and thus keep the brain in a state of readiness in executing a future task [9–10].

In functional connectivity MR imaging (fcMRI), functionally related regions of the brain are identified by measuring the temporal correlation of spontaneous low frequency fluctuations in their MR signals while the subject is in a "resting" state. Usually, the subject lies still with eyes closed for about 10 min during which no mental exercises are performed and echo-planar images are acquired in rapid succession. Unstructured random noise is seen in the MR signal together with cardiac and respiratory noise, and other oscillations below 0.1 Hz reflecting spontaneous, physiologically driven changes in local blood flow. To analyze signal time courses for correlations, the temporal correlation coefficient of a signal from a selected voxel (so called "seed voxel") or voxels in a region of the brain is computed with signals from all other voxels in the brain. A statistical test is performed for each comparison to determine the strength of the correlation, and a suitable threshold is applied to create an image showing those regions of the brain with strong correlations to the selected seed voxels. Functional connectivity is hypothesized between the seed voxels and the

* Corresponding author. Tel.: +1-206-543-6832; fax: +1-206-543-3495.
E-mail address: dcordes@u.washington.edu (D. Cordes).

regions with high temporal correlation to the fMRI response of the seed voxels.

While the feasibility of resting-state fMRI has been demonstrated in several papers, functional connectivity is incompletely characterized with the standard method. The seed voxel method is limited since each different seed voxel may provide a different functional connectivity map. Seed voxels are usually selected on the basis of anatomic information or previously performed functional activation maps. Partial volume effects where CSF and major blood vessels both contribute to the MR signal may confound the connectivity. Therefore, if adjacent voxels to the selected seed voxels are used, the resulting cross correlation maps can have completely different features [7–8]. The seed voxel method can been criticized for bias in the selection of the seed voxels since there is no experimenter-independent selection standard.

To reduce reader bias, functional connectivity data could be analyzed by means of a model independent method. Exploratory data analysis methods have the attractive feature of being model free and thus allowing unbiased studies of brain signal responses. Examples in fMRI/PET include principal component analysis (PCA) [11], independent component analysis (ICA) [12] and cluster analysis [13–15]. PCA partitions data space into orthogonal components based on the covariance matrix. ICA relaxes the orthogonality constraint and determines components that are as statistically independent by reducing higher than second order statistical dependencies. While ICA has been applied to resting-state data [16–18], it is not clear how many of the components obtained are actually physiologic. In ICA Gaussian or near Gaussian distributions in physiological noise having a small Kurtoses cannot be separated and might contaminate other components, especially since the amplitude of low frequency oscillations is small. This is especially problematic. Furthermore, the necessary time independence of the weighting matrix as formulated in current ICA algorithms might not be true for a realistic description of resting-state data, since spatial pattern of resting-state activity might change during the long acquisition of the study. However, some successful applications have been reported recently [17].

In the present work we applied a hierarchical clustering algorithm [19–20] based on the single link (nearest neighbor) method to find clusters whose voxel members have high cross correlation coefficients that indicate low frequency synchronous fluctuations in the fMRI signal. We use the synchrony to imply functional connectivity. This method does not require prior knowledge of cluster centers or the number of clusters present in the data. Our approach represents a first attempt to define an appropriate distance measure for analyzing resting-state data to partition all possible cross correlations coefficients from multi-slice data into meaningful patterns of functional connectivity. Furthermore, this approach permits an evaluation of the effects of

hardware (i.e., gradient) instabilities and magnitudes of motion-induced correlations on connectivity data.

2. Materials and methods

2.1. Data Acquisition

Four normal male volunteers, ranging in age from 20–25 years and claiming to be in good health, participated in this study. Signed consent was obtained according to institutional guidelines (IRB approval). Head stabilization and motion control were achieved with foam pads. Each subject was instructed before the scanning session to be as motionless as possible during the EPI acquisitions, to keep his eyes closed and refrain from any cognitive exercise. MR scanning was performed in a commercial 1.5 T LX scanner (General Electric, Waukesha) equipped with high-speed gradients and a standard birdcage head coil. 2D spin echo anatomic whole brain images (coronal plane, 20 slices, 256×128 matrix, TR/TE 500 ms/8 ms, FOV 24 cm, slice thickness 7 mm, gap 2 mm) as well as 3D SPGR images (axial plane, 124 slices, 256×128, TR/TE 21 ms/7 ms, flip angle 40 deg, FOV 24 cm, slice thickness 1.1 mm, no gap) were acquired. Six resting-state gradient-recalled EPI scans (epibold) were performed in the coronal plane with parameters: 4 slices, 64×64 matrix, TR/TE 400 ms/50 ms, flip angle 50 deg, FOV 24 cm, slice thickness 7 mm, 2 mm gap, 1300 time frames, BW $\pm$ 62.5 kHz. The locations of the slices were chosen for each resting-state scan differently, so that auditory cortex, motor cortex and visual cortex were included. The high sampling rate was selected to be able to resolve cardiac oscillations. At a TR of 400 ms, the Nyquist frequency is given by $1/2TR = 1.25$ Hz. Therefore, oscillations that are below this frequency will not alias. Both respiratory and cardiac rates were recorded using a flexible respiratory belt and a pulse oximeter. Aliasing of the cardiac rate was eliminated by the choice of a TR of 400 ms. The EPI scan duration was 8 min 40 sec.

2.2. Control studies

Four different phantoms (GE QA phantom, water melon, customized brain-like phantom [21], form aldehyde fixed human brain (post-autopsy)) were scanned with identical EPI parameters to assess hardware (gradient) instabilities.

2.3. Post-processing

In each EPI series, the images belonging to the first twenty time frames were discarded because the MR signal was not at equilibrium. During reconstruction of the raw EPI data, a low pass Hamming filter in the spatial frequency domain was applied to increase SNR at a cost of inducing a modest loss in spatial resolution [22]. All images were checked for motion with a 3D registration program included

in AFNI (Robert Cox, NIH) [23]. This registration algorithm employs Fourier interpolation in the resampling process [24]. Motion was minimized by the use of padding around the head, selection of cooperative volunteers and instruction not to move. Data sets with more than 0.6 mm maximum displacements in x, y or z, or more than 1 deg of angular motion about each axis were discarded. 3D motion correction was carried out for each EPI dataset and the file containing the 6-dimensional motion parameters at each time frame was saved. Due to interpolation of the 3D registration algorithm, the resulting data can contain motion-related components to an unknown degree. These residuals, if not removed, could lead to large correlation coefficients and might give clusters that do not represent functionally connected voxels at all. To reduce these artifacts, we used the 6-dimensional motion parameter file and projected out the two major motion components $m_i(t)$ by Gram-Schmidt orthogonalization and arrived at a motion-reduced time course $\tilde{f}(t)$ for each voxel by

$$\tilde{f}(t) = f(t) - \sum_i \frac{\langle f|m_i\rangle m_i(t)}{\langle m_i|m_i\rangle} \tag{1}$$

where $f(t)$ represents the voxel time course, $m_i(t)$ the motion components, $\langle a|b\rangle$ the scalar product of the functions $a(t)$ and $b(t)$, defined by

$$\langle a|b\rangle = \sum_t a(t)b(t). \tag{2}$$

Note, that the temporal mean of each function was removed prior to orthogonalization. For more than one motion component, Eq. (1) is only valid in the least square sense since the motion components are not necessarily orthogonal to each other. Therefore, the motion components are reduced but not completely eliminated in $\tilde{f}(t)$. For our data, we used not more than two motion components in the projection, because all other motion components were essentially zero. The two major motion components that were projected out corresponded to a rotation about the A-P axis and a displacement in the I-S direction.

Gram-Schmidt orthogonalization assumes that the motion components are orthogonal to each other. This is only approximately true for our data. In general, if the motion components have significant overlap from collinearities, equation [1] should not be used since this would lead to an over-reduction of motion components. Instead, one could use a modified projection approach, where motion components are projected out sequentially. However, we have examined our motion parameters and did not find any gross violation of the orthogonality assumption.

2.4. Residual contribution of motion components

To estimate the contribution of residual motion components (contained in $\tilde{f}(t)$), we derived an upper bound of the correlation coefficient Δcc_{fg} that is solely due to motion (see Appendix) by

$$\Delta cc_{fg} = \frac{1}{\sqrt{\langle f|f\rangle\langle g|g\rangle}} \sum_i \sum_j \frac{\langle f|m_i\rangle\langle g|m_j\rangle\langle m_i|m_j\rangle}{\langle m_i|m_i\rangle\langle m_j|m_j\rangle}$$

$$= \sum_i \sum_j cc_{fm_i} cc_{gm_i} cc_{m_i m_j} \tag{3}$$

where $f = f(t)$ and $g = g(t)$ represents two voxel time courses. The terms cc_{fm_i}, cc_{gm_i} and $cc_{m_i m_j}$ specify the correlation coefficients between the voxel time course $f(t)$ with the motion component $m_i(t)$, $g(t)$ with the motion component $m_i(t)$, and $m_i(t)$ with $m_j(t)$, respectively. Equation 3 provides only a first order estimation of motion contribution since non-linear motion effects, where changes in signal strength in a voxel caused by motion are nonlinearly related to the amount of motion, cannot be estimated using the proposed method.

2.5. Constraining the data set to brain tissue

From the total number of voxels in the four-slices with 64×64 voxels, some voxels were eliminated for consideration in the hierarchical cluster analysis. By discarding voxels with a SNR (calculated by the ratio of the mean EPI intensity and its standard deviation) in the 10th percentile or less range, the number of voxels was reduced to about 4000. Voxels outside the brain were effectively eliminated by this means. The number was reduced further by counting the number of significant pairwise correlations of low frequency contributions. We used a cross correlation cut-off threshold of 0.3 and assigned a frequency count to each voxel. This count specified how often a voxel is involved in a significant correlation with other voxels. Many voxels do not have any significant correlation with other voxels, even when a spatial Hamming filter is used, as in our approach. Fig. 1 shows a histogram of the significant correlations (i.e., correlation coefficients >0.3). We discarded all voxels that had less than four significant correlations. For a four-slice human brain data set this approach resulted typically in about 1300 to 2400 brain voxels left for the clustering analysis.

2.6. Hierarchical clustering

A hierarchical clustering algorithm based on the single link method was used [19]. As an appropriate similarity measure to group voxels into clusters for functional connectivity, we specified a new distance measure, which is based on the correlation coefficient of EPI-time series restricted to very-low frequencies (<0.1 Hz). In order to accomplish this, the spectral decomposition of all correlations between two voxels in the brain must be computed. Since four slices constitute about 4000 voxels of brain tissue at our resolution, the calculation amounts to the computation of about 10^7 different cross correlation coefficients. For each pair of correlations, the correlation coefficient has to be decomposed into frequency contributions using a method
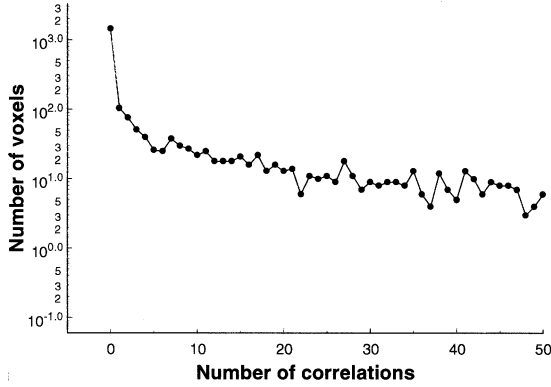
Fig. 1. The curve (logarithmically scaled) shows how many voxels above threshold have multiple correlations with other voxels. The correlation threshold is set to cc = 0.3. The number of (non-overlapping) voxels with 0,1, and 2 correlations is 1446, 104, and 76, respectively. Note, that the graph extends beyond 50 correlations on the horizontal axis. For illustrative purposes we limited the range to 50. Summation of all discrete points on the vertical axis gives 3984 voxels. Neglecting voxels with less than 3 significant correlations reduces the total number of voxels in this case to 2358.

Fig. 2. Hierarchical clustering of four objects (A, B, C, D). Note, that the cluster location is indicated by a vertical line and the distance d between objects is indicated by the horizontal link between the clusters/objects, i.e., d(A, B) = d(0, E), d(C, D) = d(0, F), d(E, F) = d(0, G) = min(d(A, C), d(A, D), d(B, C), d(B, D)), where 0 corresponds to the origin of the distance axis.

recently developed [7–8]. The spectral decomposition of the correlation coefficient "$cc_f(q, q')$" has the property that the sum over all frequencies will yield the correlation coefficient between voxels q and q':

$$cc(q, q') = \sum_f cc_f(q, q') \tag{4}$$

where

$$cc_f(q, q') = \frac{N(\mathrm{Re}(\omega_f)\mathrm{Re}(\lambda_f) + \mathrm{Im}(\omega_f)\mathrm{Im}(\lambda_f))}{D} \tag{5}$$

labels the Fourier component (with frequency f) of the correlation coefficient between voxel q and voxel q'. The terms ω_f and λ_f are the complex frequencies of the voxel time courses for q and q', respectively. The term D represents the product of the norms of the time courses for voxels q and q'. For details see Cordes [7]. As a relevant distance measure $d(q, q')$ for the clustering between voxel q and voxel q' we propose

$$d(q, q') = 1 - \sum_{f=0}^{0.1Hz} cc_f(q, q'). \tag{6}$$

The value of this distance measure ranges between 0 and 1 (for positive correlations), and it decreases as the very low frequency contributions in the correlation coefficient increase. For our clustering method, we consider only voxel pairs with low frequency correlation coefficients of at least 0.3, i.e., $0 \leq d \leq 0.7$. Strictly speaking, the distance measure $d(q, q')$ is not a metric, because it does not satisfy the triangle inequality.

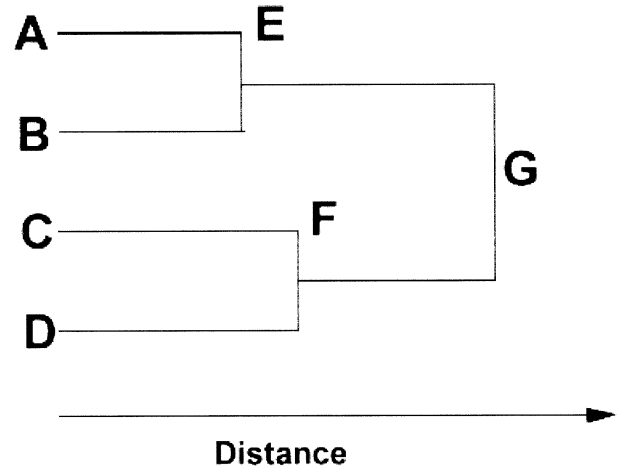Clustering methods such as the well-known K-means

algorithm [19] are commonly used in data analysis. A drawback of the (pure) K-means algorithm is that the number of clusters has to be identified in advance. As an alternative, hierarchical clustering methods proceed by stages producing a binary cluster tree, which is also called a dendrogram. At each stage in the cluster tree, objects with similar features are linked. The number of clusters at the final stage can be determined by selecting a linkage inconsistency threshold, which is one method to find the natural cluster divisions in the data set. If a link between clusters in the dendrogram has approximately the same length as neighboring links, then the objects that are joined at this level have similar features, which constitute the definition of a cluster. The corresponding links are said to exhibit a high level of consistency. Vice versa, if the length of a link differs from neighboring links, the level of inconsistency of the members at this level grouped in a cluster is high. The linkage inconsistency coefficient at a depth of two in the dendrogram is defined as the length of a link in a cluster hierarchy minus the average length of all links divided by the standard deviation of all links. As an example, consider four objects labeled A, B, C, D, which have a distance d(P, Q) to each other, and P, Q ϵ {A, B, C, D} (see Fig. 2). A and B are linked forming cluster E and similarly, C and D are linked to cluster F. Next, E and F is linked to cluster G. Distances between objects (voxel time series) are measure by Eq. 6. Distances between clusters consisting of objects are based on the minimum path length among all pairs of objects in the clusters (single link method). For example, the distance between clusters E and F in Fig. 2 is given by the minimum distance between A, B and C, F, which is min(d(A, C), d(A, D), d(B, C), d(B, D)). The linkage inconsistency coefficient I of the last link (cluster G = E + F) is given by
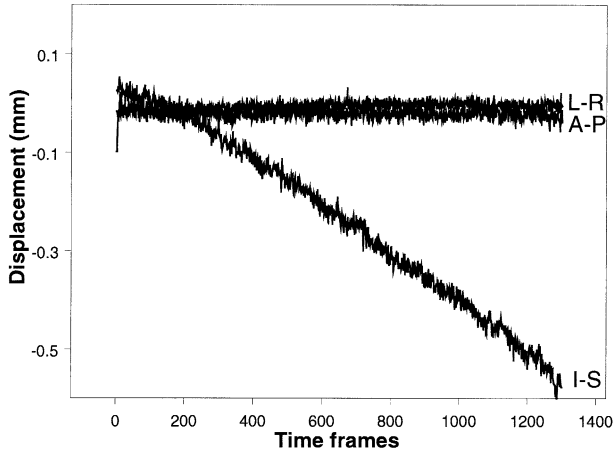
Fig. 3. Linear displacement parameters as calculated by the 3D registration algorithm of customized brain phantom (L-R direction, A-P direction, I-S direction). The interval between two time frames is 400 ms.
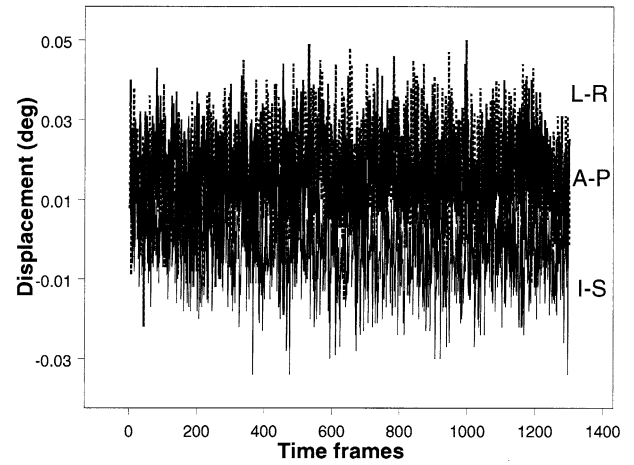

Fig. 4. Angular displacement parameters as calculated by the 3D registration algorithm of customized brain phantom (rotation about L-R axis, A-P axis, I-S axis).

$$I = \frac{d(E, F) - \dfrac{d(A, B) + d(C, D) + d(E, F)}{3}}{std} \qquad (7)$$

where *std* is the standard deviation of the distance for the three links (A + B, C + D, E + F). As a suitable threshold for cluster size identification, we used an inconsistency coefficient of $I = 0.9$ for all resting-state fMRI data. This value was found empirically by measuring the degree of spatial extent between task-defined primary motor areas and obtained cluster size in resting-state data. An inconsistency coefficient of 0.8 or smaller could not produce acceptable cluster sizes.

2.7. Characterizing clusters by frequency-specific contributions to the correlation coefficient

The quantity "$cc_f(q, q')$" in Eq. (5) defines the contribution of each frequency f to the correlation coefficient of the voxel time courses q and q'. For each cluster, the average value of $cc_f(q, q')$ can be computed by

$$\langle cc_f \rangle = \frac{1}{N} \sum_{q \neq q'} cc_f(q, q'), \qquad (8)$$

where N counts the number of pairwise correlations in the cluster. The term $\langle cc_f \rangle$ can also be called the average correlation coefficient at frequency f. In order to visualize the contributions of the low frequency band versus respiratory and cardiac bands, we computed the integral of Eq. (8) by

$$CC(f) = \sum_{f'=0Hz}^{f} \langle cc_{f'} \rangle \qquad (9)$$

as a function of *f*. This function quantifies the influence of several frequency bands on the total correlation coefficient. A steep slope in *CC(f)* describes frequency bands that give

major contributions, whereas a slope of zero specifies non-contributing frequencies.

3. Results

3.1. Control studies

Artifacts resembling motion were present in all phantom data due to hardware instabilities. For the customized brain phantom [21] linear and angular displacements were required to register the data (Figs. 3 and 4). The linear displacement in the I-S direction is the dominant artifact. The scanner drift is present during the entire scan and has a constant slope of 0.067 mm/min. A slope of similar magnitude has also been found in human data. Fourier analysis (also evident in Fig. 4) shows that the data are contaminated by very low frequencies (period > 70 sec) with small
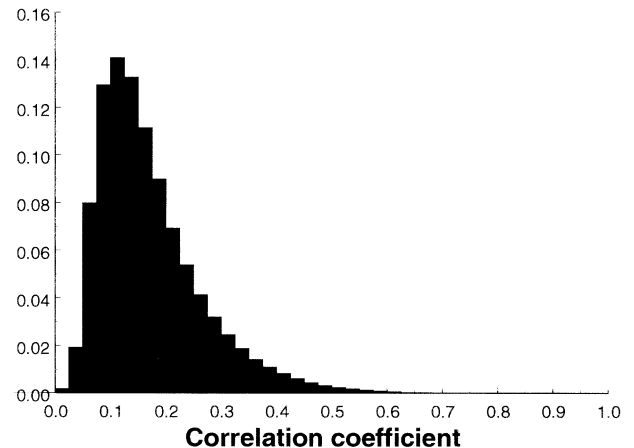

Fig. 5. Distribution of the correlation coefficient between each pair of voxels in the major cluster after registration but without orthogonalization for customized brain phantom.
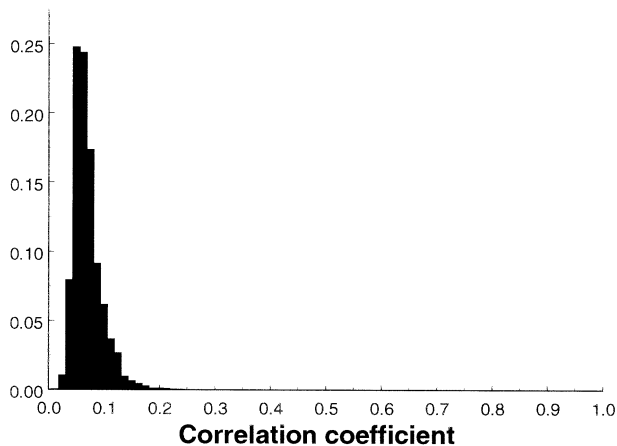
Fig. 6. Distribution of the correlation coefficient between each pair of voxels in the major cluster after registration and orthogonalization with the major motion correction functions for customized brain phantom.

amplitudes. Clustering of the registered phantom data produced two clusters. Fig. 5 shows a histogram of the low-frequency (<0.1 Hz) correlation coefficient for all pairs of voxels in the dominant cluster. The estimated contamination of the correlation coefficient due to motion components in the data according to Eq. (3) is 0.39 for both clusters. After Schmidt-orthogonalization of the motion components from the registered data, this value is reduced to 0.02. Fig. 6 shows the histogram of the corresponding correlation coefficients in the major cluster. Fig. 7 shows the frequency-specific information from two clusters of the correlation coefficient indicating that very low frequencies with periods larger than 70 sec are still present in the data. Fig. 8 shows the integrated correlation coefficient according to Eq. (9). The effect of Gram-Schmidt orthogonalization is clearly visible for the major two clusters. The average correlation coefficient is reduced to less than 0.1 after Gram-Schmidt orthogonalization.

3.2. Human studies

For human studies, the number of voxels in the four slices in the cluster analysis was 1300 to 2400. In these four slices typically 15 to 25 clusters containing at least five voxels were found by hierarchical cluster analysis (Fig. 9). Few clusters had more than 10 voxels. Occasionally, a single cluster of size 30 to 50 was obtained, which could readily be identified as CSF in the ventricles or in the sylvian fissure. The corresponding frequency spectrum of the total correlation coefficient contained large contributions in the frequency range larger than 0.1 Hz due to respiratory and cardiac effects. Large portions of clusters (1200 in this case) were single voxel clusters, i.e., the corresponding voxels could not be clustered at all with our algorithm.

Most of the major clusters (that contained more than 4 voxels) gave reproducible pattern in the sensorimotor cortex, thalamus, primary visual cortex, fusiform gyrus, primary auditory cortex and Broca's area. As typical examples, Fig. 10 shows readily identifiable patterns in one subject. All clusters were obtained by considering only frequency contributions less than 0.1 Hz (after Gram-Schmidt orthogonalization with the major motion functions from the registration). These patterns could be obtained in several different resting-state scans and the cluster locations were stable and reliable. Besides these patterns, other non-stationary patterns could be found. These patterns were obtained in a particular resting-state data set and could not be reproduced in other data sets of the same subject. Fig. 11 shows examples of these clusters. Motion artifact may explain clusters like those in Fig. 11, which occur at the edge of the inferior temporal lobes, where motion artifacts are common. However, analysis of possible motion contribution for these clusters gave insignificant numbers (see Table 1) and does not provide a sufficient explanation.

The role of cardiac and respiratory frequencies in the clusters was analyzed by computing the integral of the average frequency specific correlation component from 0 to frequency f for all clusters (Fig. 12). If respiratory and cardiac effects contaminate voxels in specific clusters, a step-like increase in the slope at the respiratory range (0.2–0.3 Hz) and cardiac range (0.9 to 1.0 Hz) would occur. This is in general not the case. An exception is cluster D, which shows that its voxel time courses are significantly contam-
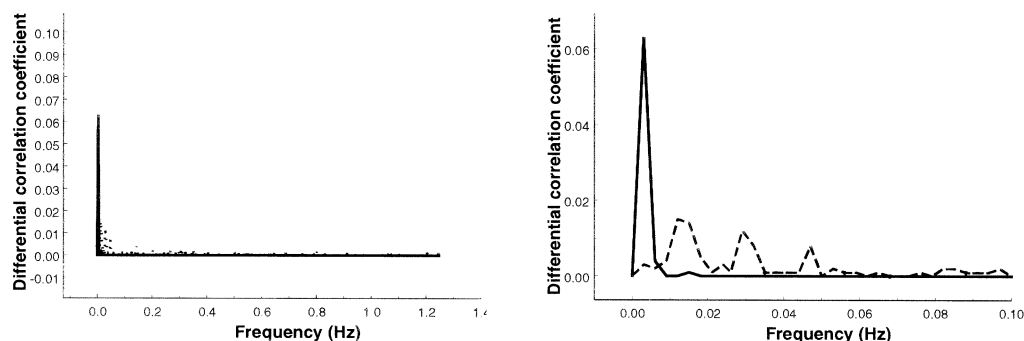

Fig. 7. Average frequency specific correlation component (Eq. (8)) for two clusters in customized brain phantom as a function of frequency components after registration and orthogonalization to the dominant motion components. Note, that a peak close to zero frequency still remains, even with orthogonalization. The enlarged figure on the right shows the graph in the low frequency range < = 0.1 Hz.
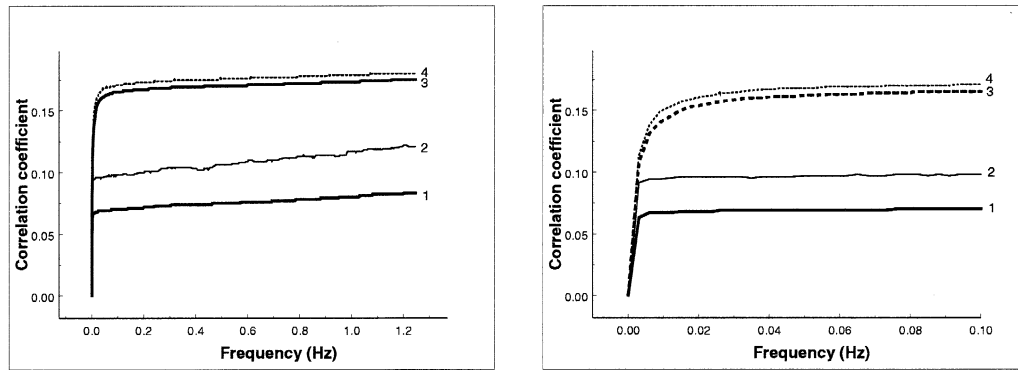
Fig. 8. Integrated correlation coefficient from 0 to frequency f (Eq. (9)) for customized brain phantom. An enlarged graph for the small frequency range is displayed to the right. The curves (labeled 3 and 4) represent two clusters of registered data without motion orthogonalization. The clusters labeled 1 and 2 correspond to registered and motion orthogonalized data as in Fig. 7. It is apparent that only low frequencies less than 0.1 Hz contribute significantly and that non-orthogonalized voxel time courses still have large contaminations from gradient instabilities (low frequency contamination, especially linear drift). The enlarged figure on the right suggests that artifacts can be reliably removed by deleting frequency contributions less than 0.01 Hz. However, these contributions are only minor after Gram-Schmidt orthogonalization.

inated by respiratory and cardiac effects. Therefore, even if the clustering was reduced to the low frequency range ($\leqq$ 0.1 Hz), one has to apply caution in interpreting this cluster D as an example for functional connectivity since low frequency contribution in CSF and major blood vessels do occur. All other clusters show that the dominant correlations are indeed frequencies less than 0.1 Hz. The slope in Fig. 12 shows a steep increase until 0.1 Hz and then is almost zero for higher frequencies. Since the average contribution to the correlation coefficient of the low frequency range is larger than 0.3 and the estimated artifacts are small (Table 1), gradient instabilities or subject motion are not possible explanations.

It is also interesting to see that some major contributions to the correlation coefficients arise from frequencies with periods larger than 70 sec. In phantoms, these contributions were artifacts from gradient instabilities. However, for human subjects, associating this very low frequency range to gradient/motion artifacts can give only a partial explanation,

since registration and motion orthogonalization will reduce motion contamination (see Table 1). Therefore, a significant portion of the very low frequencies should be attributed to slow physiological changes rather than motion artifacts. For example, the pattern of the fusiform gyrus (cluster G) arises from very low frequencies (T $>$ 70 sec) that contribute significantly to the correlation coefficient (0.46). Thus, these numbers cannot be explained entirely by correlations from gradient drifting or remaining motion artifacts, however a partial explanation due to motion might be valid.

4. Discussion

We have demonstrated the application of a clustering methodology based on temporal correlations in the low frequency range that reveals spontaneous hemodynamic processes in the resting brain. This approach does not rely on an Euclidean distance measure, which is often used in clustering Gaussian-distributed data [15]. Rather, our method is designed to explore correlations in structured non-Gaussian physiological noise without any contaminations from respiratory or cardiac effects. In order to separate physiological artifacts such as respiratory and cardiac components from true functionally connected regions, we used a high sampling rate and computed the frequency-specific contribution of the correlation coefficient. Clustering based on low frequency correlations does not produce only groups of functionally connected voxels that belong to the same network since CSF as well as larger blood vessels may have some low frequency correlations [8]. However, the amplitude of low frequency oscillations in CSF and major blood vessels is in general much smaller than in functional connected tissue. By computing the integral of the frequency specific correlation component according to Eq. (9), contributions from physiological noise sources in the respiratory
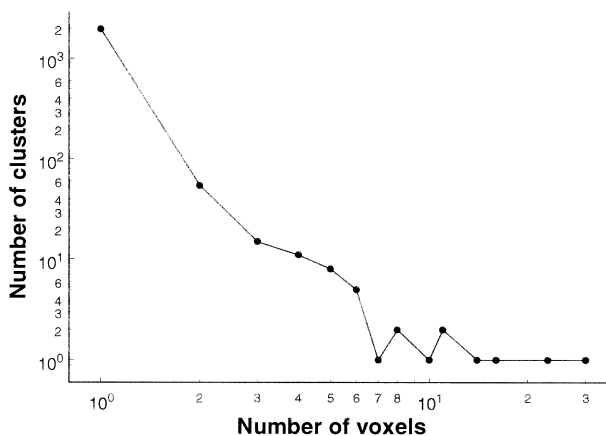

Fig. 9. Histogram of the cluster size (number of voxels) in a typical human study. Both axes are scaled logarithmically.
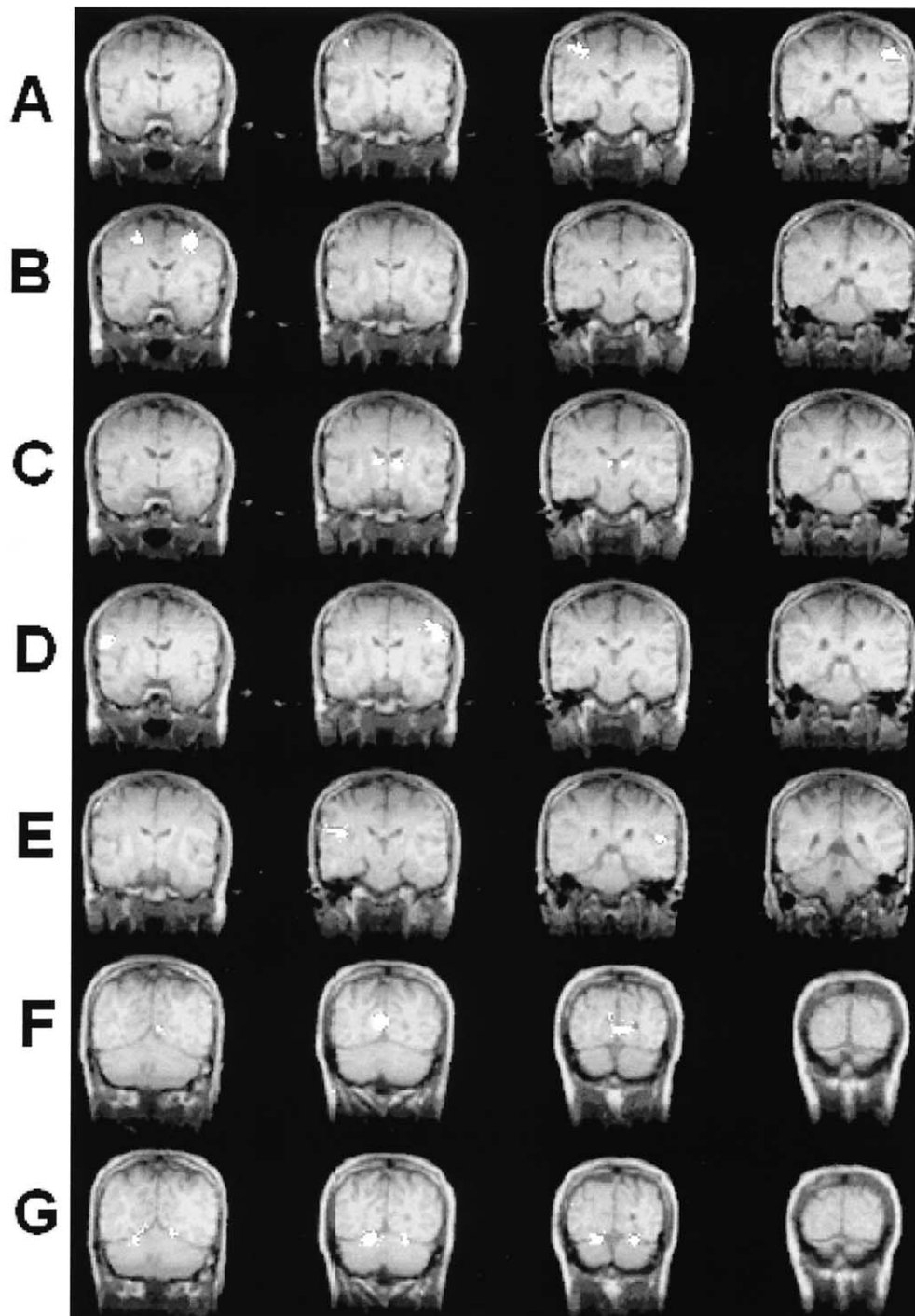
Fig. 10. Voxels identified by hierarchical cluster analysis in a resting data set in the sensorimotor cortex (cluster A, cluster B), thalamus (cluster C), frontal lobes (cluster D), primary auditory cortex (cluster E), primary visual cortex (cluster F), and fusiform gyrus (cluster G).

and cardiac range can be readily identified. It is expected that CSF contains many different frequencies in the correlation spectrum, mostly from respiratory and cardiac pulsations [4,8]. In arteries, major frequency contributions arise from the cardiac beat and its harmonics. In veins most frequencies are in the respiratory range with smaller contributions in the cardiac range [8,25]. Examples from our results are cluster D and E in Fig. 12. Cluster D is contam-

inated by frequency components in the respiratory and cardiac range. For example, the correlation coefficient shows a jump of 0.2 in the range 0.1 to 0.3 Hz and a jump of 0.1 at 0.95 Hz. Since the cardiac contribution is small compared to the respiratory contribution, it is reasonable to assume that cluster D contains artifacts from larger veins or CSF. Cluster E does not show any respiratory frequency components, but a small contribution at the cardiac rate (0.95 Hz) is
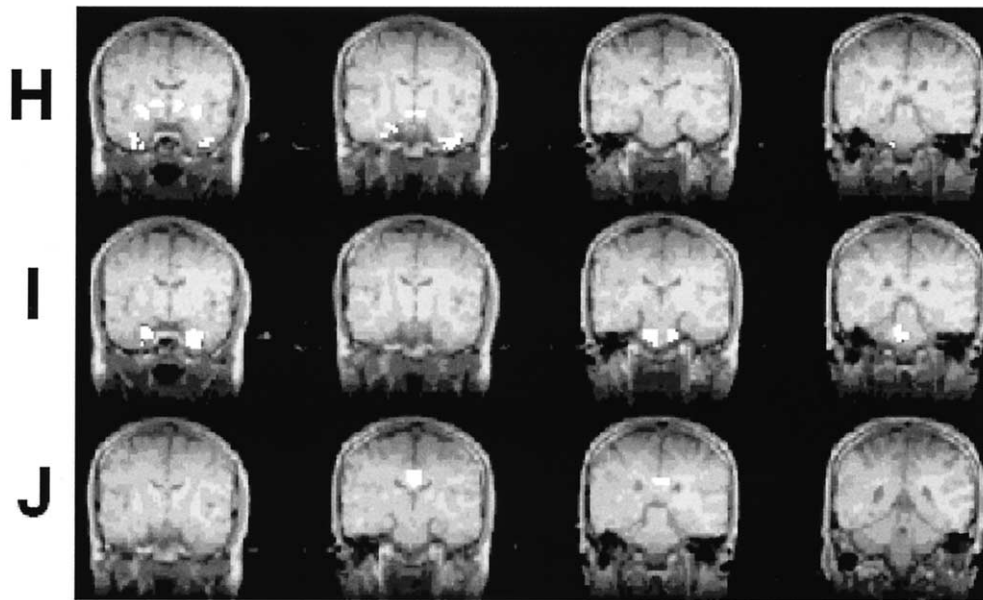
Fig. 11. Voxels identified by hierarchical cluster analysis in a resting data set. These patterns are not previously reported: Cluster H illustrating symmetrical regions in the temporal lobes and basal ganglia, cluster I symmetrical regions in the mesial temporal lobes and brain stem, cluster J symmetrical regions in the mesial frontal lobes.

apparent. Therefore, cluster E is probably free of CSF but contains a small contribution of an arterial vessel.

Low frequency motion artifacts arising either from subject motion or from instabilities of the gradient system can in principle be reduced by a highpass filter or by linear (or piecewise) linear detrending of the time series affecting only frequencies with period larger than T = 70 sec. Without elimination of these frequencies, large artificial correlations can be observed in phantom data from gradient drifting, where frequencies below 0.015Hz (T = 65 sec) are found [26]. Gradient drifting will affect all voxels. However, correlations due to gradient drifting will be more pronounced near regions with large spatial intensity gradients. For our registered phantom data (but without motion orthogonalization), correlations with magnitudes between 0.3 and 0.4 could be observed. We have chosen a more powerful method to reduce these artifacts. This approach reduces the maximum correlations in phantom data to values less than 0.1 and has the advantage that only true artifact components are removed from the data set. Frequency filtering or detrending algorithms remove all information regarding very low frequency components and might therefore eliminate true physiological fluctuations of very low frequencies.

In this study we discarded two volunteer data sets be-

cause of a significant contribution of residual motion artifact was detected. All data reported here has very small contributions from possible motion artifacts. However, some of the clusters were located near edges of the brain (lower temporal lobes and brain stem in clusters H and I) where motion artifact is often detected. However none of these clusters showed contributions from either respiratory (0.2–0.3 Hz) or cardiac motion (0.95 Hz). Also subject motion in these data sets was too small to explain the magnitude of the correlation coefficient (see Table 1). In order to select voxels for clustering we have selected a threshold of 0.3, which is large enough to suppress motion and gradient-drifting artifacts (after registration and motion reduction by orthogonalization).

We find that several clusters are fairly stable in different resting state data sets of the same subject. These clusters are located where previous fMRI studies have found functional connectivity. For example, primary cortices such as auditory cortex, Broca's area, visual cortex, somatosensory cortex but also subcortical regions such as the thalamus and possibly the hippocampus are fairly robust. Other patterns are variable and have been seen in some data sets but not in others. Since the "resting-state" does not preclude cognitive activity or sensory stimulation, many spontaneous activities may occur during a 10 min MR scan. Even in the same

Table 1

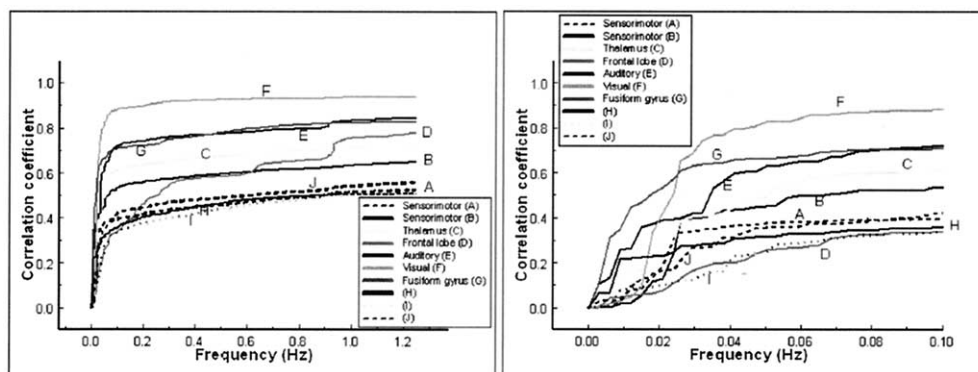| Cluster	A	B	C	D	E	F	G	H	I	J
$\langle cc \rangle$	0.40	0.53	0.61	0.34	0.72	0.88	0.71	0.36	0.34	0.42
Δcc	0.10	0.05	0.16	0.02	0.04	0.18	0.31	0.17	0.05	0.11

Fig. 12. Contribution of frequency ranges to the correlation coefficient for all clusters (A-J) in Figs. 10–11. An enlarged graph for the small frequency range is shown on the right.

volunteer, substantial variations in activity patterns are possible. Since an MR session for each volunteer lasted about two hours, it is understandable that the volunteers' thought patterns at the beginning of the session might have been different than two hours later. After the sessions, volunteers were asked what they thought during scanning. The responses varied. Most were able to relax and refrain from any particular cognitive activities; others reported impatience toward the end of the session. We noticed that the motion increased toward the end of the session. Despite the uncertainties of the "resting-state," functional connectivity analysis might in principle allow the detection of distributed networks in a way not achieved by task activation [27]. Also, an understanding of low frequency processes in the resting-state can lead to an improvement of task paradigms in fMRI where the resting-state is considered to be the null state.

In optical imaging methods [5], frequencies of 0.04 Hz and 0.1 Hz in physiological noise of the visual cortex were detected with near-infrared spectroscopy. These frequencies were proven to be unrelated to heart rate and respiration. The frequency range of low frequency oscillations in [oxy-Hb], [deoxy-Hb], and [Cyt-ox] does agree with fMRI connectivity studies where frequency peaks less than 0.1 Hz were found. However, the particular locations of the peaks detected by optical imaging do not match with the fMRI peaks. It has been hypothesized that the BOLD contrast is generated by a complex interplay of oxygen uptake, blood volume and flow velocity, which may alter its spectral profile.

A hierarchical clustering algorithm based on the single link method has the advantage of easy implementation. Furthermore, a cluster validity coefficient, called the cophenetic correlation number, is calculated from the linkage function leading to the dendrogram. The cophenetic correlation number measures the performance of the clustering solution and ranges between 0 and 1 (1 for perfect cluster solutions). This number also depends on the value for the cluster inconsistency threshold, which is set to 0.9 to achieve acceptable cluster sizes for the largest ten clusters.

For our data and thresholds set, the cophenetic correlation number is 0.48. A large portion of voxels with sufficient correlations could not be clustered at all, even when the correlation inconsistency threshold was raised. However, interesting patterns in a small group of voxels could be found. For example, functional connectivity in subcortical voxels belonging to the thalamus was found consistently. In the past using the seed voxel method, contra and ipsilateral functional connectivity in the thalamus was detected [28]. With our method, we are able to provide more detail about the physiological characteristics of the corresponding clusters, such as frequency specificity of correlation coefficients and degree of motion contamination.

Not all images generated by the "seed voxels" method are in perfect concordance with the clusters obtained by our new method. In particular, auditory as well as somatosensory connectivity show different patterns. Primary visual cortex and thalamus connectivity have better concordance. However, to find differences is not surprising since both methods are difficult to compare. Maps from the seed voxels method can be different even in the same data set, because the seed voxels have to be chosen first by the experimenter. Therefore, a quantitative comparison of both methods can be very subjective.

Hierarchical clustering is computationally demanding compared to traditional approaches (K-means and related algorithms) [19]. The memory requirements are proportional to the square number of groups and the execution time is expected to increase substantially for a whole brain data set. We were able with a 900 MHz Pentium III PC with 1.5 GB RAM to handle a 4-slice data set in one hour.

Analysis of resting-state data has both biologic and mathematical complexities. From the biologic side, the origin of the vasomotor oscillations is still not understood. However, we have shown that respiratory and cardiac beats are not the cause of correlated signal fluctuations in clusters across hemispheres. Since these low frequency correlations do occur also between groups of voxels that are far apart, the oscillations must have been modulated by neuronal activity. For future studies, patients with well-characterized

anatomic disconnections should be studied to shed more light on the nature of slow vasomotor oscillations in relation to neuronal activities [29–31]. Also, it may be interesting to study functional connectivity in patients with psychiatric disorders (dementia or schizophrenia). From the mathematical side, new methods of analyzing resting-state data are rarely found. The methodological issues in functional connectivity analysis are complex. Classification schemes have not been explored. For each of the numerous available clustering schemes there exist methods to define a set of thresholds based on information maximization criteria [32]. Different clustering schemes have different power for the same application. It would be worthwhile to explore more sophisticated clustering algorithms to partition resting-state data into meaningful patterns since with our method many voxels could not be clustered (Fig. 9) and we were able to obtain mostly patterns involving largely eloquent areas in the brain. Ultimately, if the theory of functional connectivity is valid, there should be more concordance between anatomic connections, as found by diffusion tensor imaging, and functional connections. However, functional connectivity may not concord entirely with the anatomic connectivity. Psychiatric illness may represent in some cases failure of the normal connectivity to develop. In others it may represent the abnormal over or under-utilization of some pathways. An example of the former may be schizophrenia [33]. During the second decade of life, connections between brain regions undergo some pruning [34,35]. In schizophrenia, some researchers hypothesize, the pruning is excessive, or the connections prior to pruning are less well developed [36]. On the other hand, in depression, connections between the frontal lobe and amygdala are hypothetically decreased or other connections increased, due in some cases to multiple sad experiences [37].

Nonetheless, before the potential of resting-state data for whole brain functional connectivity analysis can be fully realized, improvements in the theoretical methods and more careful studies are needed.

5. Conclusion

The present study demonstrates that physiologically predictable patterns of correlations in low-frequency oscillations can be found in resting-state fMRI data by a hierarchical clustering methodology. A requirement for measuring connectivity with the method has been described. The method is "data-driven," meaning that the data themselves determine the natural divisions in the data set for functional connectivity. The approach is more powerful than the "seed-voxel" method in resting-state data analysis where the user selects a group of voxels and probes for all possible neuronal connections via cross correlation to the seed voxels. We have shown that most clusters are not

contaminated by respiratory or cardiac noise sources and are characterized only by large correlations of low frequency components. Furthermore, we investigated the contribution of motion artifacts from scanner instabilities as well as subject motion on functional connectivity maps and illustrated a method to eliminate or reduce artifacts from resting-state data sets.

Appendix

Derivation of the Contribution to the Correlation Coefficient from Motion (Eq. (3))

The time course for a voxel can be expressed approximately as

$$f(t) = \tilde{f}(t) + \sum_i \frac{\langle f|m_i\rangle m_i(t)}{\langle m_i|m_i\rangle} \tag{10}$$

where $\tilde{f}(t)$ is an ideal time course without any motion components $m_i(t)$ present and the temporal mean has been removed from each function. The terms in brackets are scalar products according to Eq. (2). For orthogonal functions $m_i(t)$, Eq. (10) is exactly true. However, in our case, the motion components are usually not orthogonal to each other and Eq. (10) will be the best approximation in the least square sense. Similarly, let $g(t)$ specify a different voxel with

$$g(t) = \tilde{g}(t) + \sum_i \frac{\langle g|m_i\rangle m_i(t)}{\langle m_i|m_i\rangle} \tag{11}$$

where $\tilde{g}(t)$ is an ideal time course without any motion components $m_i(t)$ present. Then the scalar product between $f(t)$ and $g(t)$ becomes

$$\langle f|g\rangle = \langle \tilde{f}|\tilde{g}\rangle + \sum_i \frac{\langle f|m_i\rangle\langle m_i|\tilde{g}\rangle}{\langle m_i|m_i\rangle} + \sum_i \frac{\langle g|m_i\rangle\langle m_i|\tilde{f}\rangle}{\langle m_i|m_i\rangle}$$
$$+ \sum_i \sum_j \frac{\langle f|m_i\rangle\langle g|m_j\rangle\langle m_i|m_j\rangle}{\langle m_i|m_i\rangle\langle m_j|m_j\rangle}. \tag{12}$$

Since $\tilde{f}(t)$ and $\tilde{g}(t)$ are motion free, the scalar products $\langle m_i|\tilde{g}\rangle$ and $\langle m_i|\tilde{f}\rangle$ vanish. Hence

$$\langle f|g\rangle = \langle \tilde{f}|\tilde{g}\rangle + \sum_i \sum_j \frac{\langle f|m_i\rangle\langle g/m_j\rangle\langle m_i|m_j\rangle}{\langle m_i|m_i\rangle\langle m_j|m_j\rangle}. \tag{13}$$

Introducing the correlation coefficient between the functions a(t) and b(t) as the normalized scalar product, i.e.,

$$cc_{ab} = \frac{\langle a|b\rangle}{\sqrt{\langle a|a\rangle\langle b|b\rangle}}, \tag{14}$$

Eq. (13) can be expressed as

$$cc_{fg} = \frac{1}{\sqrt{\langle f|f\rangle\langle g|g\rangle}}\left(\langle\tilde{f}|\tilde{g}\rangle + \sum_i\sum_j\frac{\langle f|m_i\rangle\langle g|m_j\rangle\langle m_i|m_j\rangle}{\langle m_i|m_i\rangle\langle m_j|m_j\rangle}\right) = \frac{\langle\tilde{f}|\tilde{g}\rangle}{\sqrt{\langle f|f\rangle\langle g|g\rangle}} + \sum_i\sum_j cc_{fm_i}cc_{gm_j}cc_{m_im_j}. \tag{15}$$

The term $\langle f|f\rangle$ in the denominator becomes using Eq. (13)

$$\langle f|f\rangle = \langle\tilde{f}|\tilde{f}\rangle + \sum_i\sum_j\frac{\langle f|m_i\rangle\langle f|m_j\rangle\langle m_i|m_j\rangle}{\langle m_i|m_i\rangle\langle m_j|m_j\rangle} \tag{16}$$

and can be written in terms of correlation coefficients by

$$\langle f|f\rangle = \langle\tilde{f}|\tilde{f}\rangle + \langle f|f\rangle\sum_i\sum_j cc_{fm_i}cc_{fm_j}cc_{m_im_j}. \tag{17}$$

Solving for $\langle f|f\rangle$ gives

$$\langle f|f\rangle = \frac{\langle\tilde{f}|\tilde{f}\rangle}{1 - \sum_i\sum_j cc_{fm_i}cc_{fm_j}cc_{m_im_j}}. \tag{18}$$

A similar expression holds for $\langle g|g\rangle$:

$$\langle g|g\rangle = \frac{\langle\tilde{g}|\tilde{g}\rangle}{1 - \sum_i\sum_j cc_{gm_i}cc_{gm_j}cc_{m_im_j}}. \tag{19}$$

Using Eqs. (18–19), the first term on the right side in Eq. (15) becomes

$$\frac{\langle\tilde{f}|\tilde{g}\rangle}{\sqrt{\langle f|f\rangle\langle g|g\rangle}} = cc_{\tilde{f}\tilde{g}}\sqrt{\left(1 - \sum_i\sum_j cc_{fm_i}cc_{fm_j}cc_{m_im_j}\right)\left(1 - \sum_i\sum_j cc_{gm_i}cc_{gm_j}cc_{m_im_j}\right)}. \tag{20}$$

Since the products of terms under the double sum are all positive, we obtain the relation

$$\frac{\langle\tilde{f}|\tilde{g}\rangle}{\sqrt{\langle f|f\rangle\langle g|g\rangle}} \le cc_{\tilde{f}\tilde{g}}. \tag{21}$$

From Eq. (15), we then find

$$cc_{fg} \le cc_{\tilde{f}\tilde{g}} + \sum_i\sum_j cc_{fm_i}cc_{gm_j}cc_{m_im_j} \tag{22}$$

or

$$cc_{fg} - cc_{\tilde{f}\tilde{g}} \le \sum_i\sum_j cc_{fm_i}cc_{gm_j}cc_{m_im_j} \tag{23}$$

The term

$$\Delta cc_{fg} = \sum_i\sum_j cc_{fm_i}cc_{gm_j}cc_{m_im_j} \tag{24}$$

then serves as an upper bound of the motion contribution to the correlation coefficient.

References

[1] Haughton VM, Biswal B. Clinical application of basal regional cerebral blood flow fluctuation measurements by FMRI. In: Hudetz A, Bruley D, editors. Oxygen Transport to Tissue. New York: Plenum Publishing Corp., 1998. p. 583–90.

[2] Biswal B, Yetkin FZ, Haughton VM, Hyde JS. Functional connectivity in the motor cortex of resting human brain using echo-planar MR imaging. Magn Res Med 1995;34:537–41.

[3] Lowe MJ, Mock BJ, Sorenson JA. Functional connectivity in single and multislice echoplanar imaging using resting state fluctuations. Neuroimage 1998;7:119–32.

[4] Hyde JS, Biswal BB. Functional related correlations in the noise. In: Moonen C, Bandettini PA, editors. Medical radiology: diagnostic imaging and radiation oncology. Berlin, Heidelberg, New York: Springer-Verlag, 1999. p. 263–75.

[5] Obrig H, Neufang M, Wenzel R, Kohl M, Steinbrink J, Einhäupl K, Villringer A. Spontaneous low frequency oscillations of cerebral hemodynamics and metabolism in human adults. Neuroimage 2000; 12:623–39.

[6] Giller CA, Hatab MR, Giller AM. Oscillations in cerebral blood flow detected with transcranial Doppler index. J Creb Blood Flow Metab 1999;19:452–9.

[7] Cordes D, Haughton VM, Arfanakis K, Wendt GJ, Turski PA, Moritz CH, Quigley MA, Meyerand ME. Mapping functionally related regions of brain with functional connectivity MRI (fcMRI). AJNR Am J Neuroradiol 2000;21:1636–44.

[8] Cordes D, Haughton V, Arfanakis K, Carew JD, Turski PA, Moritz CH, Quigley MA, Meyerand E. Frequencies contributing to functional connectivity in the cerebral cortex in "resting-state" data. Am J Neuroradiol AJNR 2001;22:1326–33.

[9] Bressler SL. The dynamic manifestation of cognitive structures in the cerebral cortex. In ASoCS Technical Report 97–01. New Trends in Cognitive Science. Does representation need reality. Riegler A, Peschl M, editors. 1997. pp. 37–41.

[10] Bressler SL, Kelso JAS. Cortical coordination dynamics and cognition. Trends Cogn Science 2001;5:26–36.

[11] Friston KJ, Frith CD, Liddle PF, Frackowiak RS. Functional connectivity: the principal-component analysis of large (PET) data sets. J Cereb Blood Flow Metab 1993;13:5–14.

[12] McKeown MJ, Makeig S, Brown GG, Jung TP, Kindermann SS, Bell AJ, Sejnowski TJ. Analysis of fMRI data by blind separation into independent spatial components. Hum Brain Mapp 1998;6:160–88.

[13] Filzmoser P, Baumgartner R, Moser EA. Hierarchical clustering method for analyzing functional MR images. Magn Reson Imaging 1999;17:817–26.

[14] Goutte C, Toft P, Rostrup E, Nielsen FA, Hanson LK. On clustering fMRI time series. Neuroimage 1999;9:298–310.

[15] Baune A, Sommer FT, Erb M, Wildgruber D, Kardatzki B, Palm G, Grodd W. Dynamical cluster analysis of cortical fMRI activation. Neuroimage 1999;9:477–89.

[16] Cordes D, Carew JD, Eghbalnia H, Meyerand E, Quigley M, Arfanakis K, Assadi A, Turski PA, Haugthon V. Resting-State Functional Connectivity Study using ICA. Proc ISMRM 7th annual meeting, Philadelphia, 1999. p. 1706.

[17] Kiviniewi V, Kantola JH, Biswal BB, JauHiainen J, Hyvärinen A, Tervonen O. Localization of the resting-state vasomotor fluctuation with FFT, cross correlation, Principal Component and Independent Component Analysis of fMRI data. Proc ISMRM 9th annual meeting, Glasgow, 2001. p. 1708.

[18] Arfanakis K, Cordes D, Haughton V, Moritz C, Quigley M, Meyerand M. Combining Independent Component Analysis and correlation

analysis to probe interregional connectivity in fMRI task activation datasets. Magn Reson Imaging 2000;18:921–30.

[19] Hartigan JA. Clustering algorithms. John Wiley & Sons, 1975.

[20] Matlab, Statistical toolbox. The MathWorks, Inc. HTTP://www-.mathworks.com.

[21] Madsen EL, Blechinger JC, Frank GR. Low-contrast focal lesion detectability phantom for 1H MR imaging. Med Phys 1991;18(3):549–54.

[22] Lowe MJ, Sorenson JA. Spatially filtering functional magnetic resonance imaging data. Magn Reson Med 1997;37(5):723–9.

[23] Cox RW. AFNI: software for analysis, and visualization of functional magnetic resonance neuroimages. Comput Biomed Res 1996;29:162–73.

[24] Cox RW, Jesmanowicz A. Real-time 3D image registration for functional MRI. Magn Reson Med 1999;42:1014–18.

[25] Lowe MJ. Gram-Schmidt orthogonalization to reduce aliased physiologic noise in low sampling rate fMRI data. Proc ISMRM 7th annual meeting, Philadelphia, 1999. p. 1711.

[26] Smith AM, Lewis KL, Ruttimann UE, Ye FQ, Sinnwell TM, Yang Y, Duyn JH, Frank JA. Investigation of low frequency drift in fMRI signal. Neuroimage 1999;9:526–33.

[27] Xiong J, Parsons LM, Gao JH, Fox PT. Interregional connectivity to primary motor cortex revealed using MRI resting state images. Hum Brain Mapp 1999;8:151–6.

[28] Stein T, Moritz C, Quigley M, Cordes D, Haughton V, Meyerand E. Functional connectivity in the thalamus and hippocampus studied with functional MR imaging. AJNR Am J Neuroradiol 2000;21:1397–401.

[29] Quigley M, Cordes D, Wendt G, Turski P, Moritz C, Haughton V, Meyerand E. Effect of focal and nonfocal cerebral lesions on functional connectivity studied with MR imaging. AJNR Am J Neuroradiol 2001;22:294–300.

[30] Maldjian JA. Functional connectivity MR imaging: fact or artifact? AJNR Am J Neuroradiol 2001;22:239–40.

[31] Geschwind N. Disconnection syndromes in animals and man. Brain 1965;88:231–94.

[32] Goutte C, Hansen LK, Liptrot MG, Rostrup E. Feature-space clustering for fMRI meta-analysis. Hum Brain Mapp 2001;13:165–83.

[33] Friston KJ, Frith CD. Schizophrenia: a disconnection syndrome? Clin Neurosci 1995;3:89–97.

[34] Hoffman RE, McGlashan TH. Parallel distributed processing and the emergence of schizophrenic symptoms. Schizophr Bull 1993;19:119–40.

[35] Randall PL. Schizophrenia, abnormal connection, and brain evolution. Med Hypotheses 1983;10:247–80.

[36] Weinberger DR, Berman KF, Suddath R, Torrey EF. Evidence of dysfunction of a prefrontal-limbic network in schizophrenia: a magnetic resonance imaging and regional cerebral blood flow study of discordant monozygotic twins. Am J Psychiatry 1992;149:890–7.

[37] Drevets WC, Videen TO, Price JL, Preskorn SH, Carmichael ST, Raichle ME. A functional anatomical study of unipolar depression. J Neurosci 1991;12:3628–41.