Magn Reson Mater Phy (2010) 23:327–338
DOI 10.1007/s10334-010-0204-0

RESEARCH ARTICLE

Dimensionality reduction of fMRI time series data using locally linear embedding

Peter Mannfolk · Ronnie Wirestam · Markus Nilsson · Freddy Ståhlberg · Johan Olsrud

Received: 13 July 2009 / Revised: 28 January 2010 / Accepted: 29 January 2010 / Published online: 13 March 2010

Abstract
Objective Data-driven methods for fMRI analysis are useful, for example, when an a priori model of signal variations is unavailable. However, activation sources are typically assumed to be linearly mixed, although non-linear properties of fMRI data, including resting-state data, have been observed. In this work, the non-linear locally linear embedding (LLE) algorithm is introduced for dimensionality reduction of fMRI time series data.
Materials and methods LLE performance was optimised and tested using simulated and volunteer data for task-evoked responses. LLE was compared with principal component analysis (PCA) as a preprocessing step to independent component analysis (ICA). Using an example data set with known non-linear properties, LLE-ICA was compared with PCA-ICA and non-linear PCA-ICA. A resting-state data set was analysed to compare LLE-ICA and PCA-ICA with respect to identifying resting-state networks.
Results LLE consistently found task-related components as well as known resting-state networks, and the algorithm compared well to PCA. The non-linear example data set demonstrated that LLE, unlike PCA, can separate non-linearly modulated sources in a low-dimensional subspace. Given the same target dimensionality, LLE also performed better than non-linear PCA.
Conclusion LLE is promising for fMRI data analysis and has potential advantages compared with PCA in terms of its ability to find non-linear relationships.

Keywords fMRI · BOLD · Locally linear embedding · Dimensionality reduction · ICA · PCA

P. Mannfolk (✉) · R. Wirestam · M. Nilsson · F. Ståhlberg
Department of Medical Radiation Physics, Clinical Sciences, Lund University, Barngatan 2B, 22185 Lund, Sweden
e-mail: peter.mannfolk@med.lu.se

F. Ståhlberg
Department of Diagnostic Radiology, Clinical Sciences, Lund University, 22185 Lund, Sweden

J. Olsrud
MR Department, Centre for Medical Imaging and Physiology, Lund University Hospital, 22185 Lund, Sweden

Introduction

Analysis of fMRI data has classically been based upon model-driven approaches such as correlation with a model time course or statistical testing for significant differences between data corresponding to different functional conditions. Approaches such as statistical parametric mapping (SPM) [1] are based on the general linear model. In these methods, the results only reflect how well the data can be explained by a predefined model. However, the time series data for each voxel measured during an fMRI experiment is a mixture of different signals of varying origin. The different contributions to the resulting signal are caused by both task-related and non-task-related hemodynamic activations as well as other sources, such as equipment-related artefacts and subject movement [2,3]. All of these can usually not be taken into full account in a model. Furthermore, in a typical block-design fMRI experiment, changes in strategy by the subject, differences in subject hemodynamic response function (HRF) characteristics, habituation or compliance difficulties are further examples of non-predictable events possibly leading to major deviations from the expected reference function.

Methods applied to fMRI data in order to overcome the limitations of model-driven algorithms include principal

component analysis (PCA) [4], independent component analysis (ICA) [5,6] and clustering algorithms [7]. These are data-driven algorithms that operate on the functional data without previous assumptions on the expected response function and can thus be expected to recover signal variations that cannot be anticipated or modelled. PCA and ICA have become popular tools for the analysis of fMRI data and recent applications and extensions of these algorithms include, for example, analysis of resting-state data [8–11]. PCA finds the largest eigenvectors of a data covariance matrix calculated from the original data set. The eigenvectors are then used to linearly transform the original data set into a low-dimensional representation that includes most of the variance. PCA is, however, only sensitive to second-order statistics (covariance). The purpose of ICA, on the other hand, is to separate data into statistically independent sources, given only linear mixtures of the sources as well as the number of original sources. ICA can estimate higher-order statistics and is, therefore, more appropriate for data following a higher-order (non-Gaussian) distribution.

Although successfully applied to fMRI analysis, both PCA and ICA are based upon assumptions of linear data mixing which may be insufficient for capturing for example non-linear subject motion [12]. Other studies have also discussed the limitations of the assumption that sources of activation can be considered linearly mixed, suggesting the use of non-linear analysis approaches [13,14]. Furthermore, the non-linear nature of fMRI data has recently been observed in resting-state data [15,16].

A logical improvement, therefore, would be to apply non-linear dimensionality reduction methods to fMRI data. This has previously been done using the technique of Laplacian embedding [17] and by the diffusion map algorithm [18], and in the present work the method of locally linear embedding (LLE) [19] is implemented and evaluated with respect to simulated as well as experimental volunteer fMRI data. Evaluation of the ability to find task-related components for a range of settings of the neighbourhood-size parameter K is done by calculating the correlation with a reference function and by receiver operator characteristics (ROC) analysis. The capability of LLE to find other signal components of interest, such as subject movement, is also investigated. Furthermore, data analysed with LLE are used as input to ICA substituting the commonly used PCA dimensionality reduction step as previously proposed [20]. In order to investigate the non-linear characteristics of LLE, an example data set with known non-linear properties is created and used to compare LLE-ICA with PCA-ICA and also with non-linear PCA-ICA. Finally, a resting-state data set is analysed with the aim to compare LLE-ICA and PCA-ICA with respect to finding known resting-state networks.

Theory

Locally linear embedding

The purpose of the LLE algorithm is to compute a low-dimensional embedding of high dimensional data, with nearby points in the high dimensional input data remaining nearby in the resulting low-dimensional embedding. In other words, the resulting embedding reflects local geometric properties of a point and its nearest neighbours. In the following, a brief description of the LLE algorithm is given. For a more detailed description of the algorithm, the reader is referred, for example, to the paper by Roweis and Saul [19].

The input data is assumed to consist of N D-dimensional real valued vectors $\vec{X}_i$ sampled from a smooth underlying manifold with an intrinsic dimensionality $d \ll D$. For an fMRI experiment, this means that the set of N voxel-based time courses can be viewed as lying on such a low-dimensional manifold embedded in the high dimensional space, where D is the number of time points acquired. The aim of LLE is to find a compact representation $\vec{Y}_i$ of this data set, i.e. the underlying degrees of freedom, analogously to the PCA approach. However, instead of summarising the data in terms of how much of the inherent variance is explained by choosing d eigenvectors from a covariance matrix, LLE does this by finding an embedding of lower dimensionality that preserves local relationships. As for PCA, the eigenvectors obtained from the LLE algorithm are ordered so that existing dimensions do not change if more dimensions are added to the embedding space.

In order to find similarities between points in an fMRI data set relating to BOLD signal effects, tissue contrast effects must first be removed by subtracting the mean of each time series. By doing this, the data will reside on a hyper-plane of dimensionality $D-1$. The first step of the LLE algorithm is then to select a fixed number of nearest neighbours to each point. Each point can then, in the next step, be represented by a linear patch created from its neighbouring points. For fMRI data, this can be thought of as creating patches of voxels with similar time courses. This will render weak connections between a patch built up by the time courses from active voxels and a patch built up by time courses from non-active voxels. The patches can be described geometrically by linear coefficients that reconstruct each point from its neighbours. In order to retrieve an optimal reconstruction of each data point, a cost function $\varepsilon(W)$ is minimised. The cost function is the sum of the squared distances between all the data points and their reconstructions, i.e.:

$$\varepsilon(W) = \sum_{i=1}^{N} \left\| \vec{X}_i - \sum_{j=1}^{N} W_{ij} \vec{X}_j \right\|^2 \quad (1)$$

where the weights W_{ij} contain the contribution of the j-th data point to the reconstruction of the i-th point. A linear mapping then exists consisting of a translation, rotation and rescaling that maps the high dimensional data to global coordinates on the manifold. By applying constraints to the weights W_{ij}, these are invariant to such operations and the geometric properties of the original high dimensional data can be expected to be retained in the low-dimensional mapping.

The final step is thus to obtain a mapping of the input data vectors $\vec{X}_i$ to a low-dimensional output $\vec{Y}_i$ by finding the d-dimensional coordinates of each $\vec{Y}_i$ that minimises the cost function:

$$\Phi(Y) = \sum_{i=1}^{N} \left\| \vec{Y}_i - \sum_{j=1}^{N} W_{ij} \vec{Y}_j \right\|^2 \quad (2)$$

This cost function is very similar to (Eq. 1) and is again based on the minimisation of local reconstruction errors. However, in this case, the weights W_{ij} are fixed, and the outputs $\vec{Y}_i$ are optimised. Thus, the low-dimensional outputs $\vec{Y}_i$ are reconstructed from the inherent geometric properties of the weights W_{ij} without reference to the original inputs $\vec{X}_i$.

The final step in the LLE algorithm can be thought of as unfolding the high dimensional space into the low-dimensional space, and as for PCA, this is the solution to an eigenvalue problem. This will then render corresponding patches of activated voxels in the low-dimensional space. Since these patches are weakly connected to patches corresponding to non-active voxels, they will have extreme values for one or several coordinates in the lower dimensional embedding. The d coordinates (eigenvectors) of the resulting embedding, obtained by LLE, can then be viewed as separate components of the embedding Y.

Materials and methods

Non-linear example data

In order to visualise the properties of LLE, and other methods for dimensionality reduction, a non-linear example data set was constructed. Data consisted of voxels acquired at three time points. Two groups of activated voxels were introduced. The first group simulated voxels originating from a single source, but with varying onset times. The source ($\mathbf{s}_1$) had its onset at the first time point with its time course given by $\mathbf{s}_1(t) = (1\,0\,0)$. The time course of voxel i was given by $\mathbf{y}_i(t) = \mathbf{s}_1(t - t_i) = (1 - t_i \;\; t_i \;\; 0)$ for $0 < t_i < 1$, where t_i is the onset time. This group was divided into eight clusters, with t_i spaced linearly between zero and unity and with each cluster containing eight points. The data of this group represented a non-linear modulation of $\mathbf{s}_1(t)$ and has a physiological correspondence in a relative delay of the HRF for different voxels, which is a physiologically plausible assumption [21,22]. The second group comprised a stationary cluster of 8 points with its source active during time point three, showing half the amplitude of activation relative to that of the first group, i.e. $\mathbf{s}_2(t) = (0\,0\,0.5)$. Finally, 328 points were considered as non-active voxels, stationary at the origin. The full data set thus contained 400 voxels out of which 72 were active. An offset of 0.5 was added, representing basic image contrast. Gaussian noise of variance 0.01 was added to all data points. The data set is shown in Fig. 1.

Simulated fMRI time series data

Simulated data sets were used to evaluate LLE for fMRI analysis [23]. Each set comprises 100 image volumes, consisting of 19 transversal slices (7 mm thick with 0 mm spacing). A gradient-echo echo-planar imaging (GRE-EPI) pulse sequence (matrix size 64×64, TE = 25 ms, TR = 2,000 ms, FoV = 240 mm) has been used to generate the image series. Simulated activation has subsequently been added in 10 different locations covering $3 \times 3 \times 2$ voxels each. The time course of the simulated activation is constructed using a step function corresponding to a block paradigm with five activation/rest blocks ($10 + 10$ volumes each) convolved with a canonical hemodynamic response function. In order to represent a range of typical signal change due to BOLD activation, five activation levels of 0.5, 1, 2, 4 and 6% are used. The data set also includes two models of head movement to simulate movement patterns that may be encountered during fMRI data collection. The models consist of low or high translational motions that are uncorrelated with the simulated block model. Low-grade motions correspond to subpixel shifts while high-grade motions correspond to shifts as large as several pixels.

The signal-to-noise ratio (SNR) in the simulated data sets was measured to be 121 according to:

$$\mathrm{SNR} = \frac{S}{\sigma_{\mathrm{bkg}}} \sqrt{2 - \frac{\pi}{2}} \quad (3)$$

where S is the mean signal in a region-of-interest (ROI) placed in grey matter, and σ_{bkg} is the standard deviation in a ROI placed in artefact-free background.

Volunteer data

Volunteer data were acquired using a 3T Siemens Magnetom Allegra MR unit with a standard quadrature head coil. A GRE-EPI pulse sequence (matrix size 64×64, TE = 30 ms, TR = 3,000 ms, FoV = 192 mm, 36 slices, 0.9 mm slice gap) was used to create 100 volumes during a finger tapping block paradigm consisting of five activation/rest blocks of $10 + 10$ volumes each with a total experiment length of 5 min. Five subjects participated (four males and one female, 23–37 years

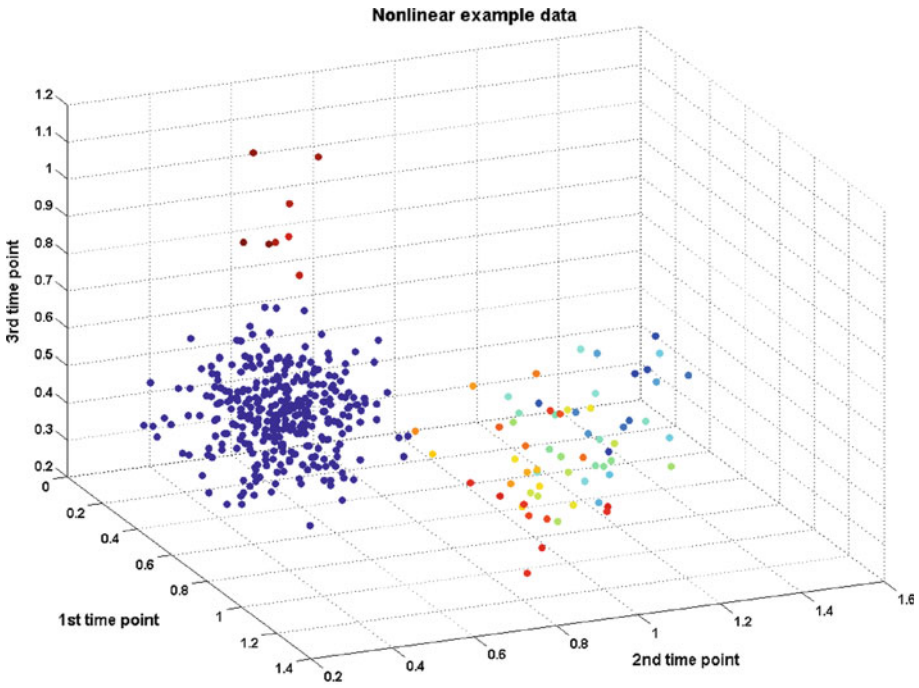

Fig. 1 Low-dimensional non-linear example data set. The group of sliding activation onset is shown in colours from *red to blue*. The stationary group at time point three is shown in *dark red*. The *dark blue* group corresponds to non-active voxels

old, average age 29). An additional volunteer was scanned in order to acquire resting-state data, and 300 volumes were acquired for this purpose using the same pulse sequence and corresponding parameters as previously described. Informed consent was obtained from each volunteer, and the project was approved by the local ethics committee. The mean SNR for the volunteer data sets was 91 (Eq. 3).

Preprocessing

The data sets (simulated and volunteer data) were realigned and volunteer data were spatially smoothed with a Gaussian kernel (FWHM = 4 mm) using SPM5 (http://www.fil.ion.ucl.ac.uk/spm) before further analysis. Spatial smoothing was not performed on the simulated data sets in order to facilitate ROC analysis. In order to evaluate the capability of the LLE algorithm to find motion-related components, data sets without realignment and spatial smoothing were also retained and used. To reduce the data sets into voxels inside the brain, a masking operation was performed by thresholding the functional volumes, retaining voxel values above the global mean of the data set. Temporal smoothing of the simulated data sets was performed using a 7th order Savitzky–Golay filter, whereas the already spatially smoothed volunteer data were analysed without application of any temporal smoothing. For all data sets, the time series for each voxel was normalised by subtracting the average signal intensity. Before applying LLE, the intrinsic dimensionality of the data sets was estimated using the Bayesian information criterion (BIC) [24].

Implementation

LLE was implemented in MATLAB (The Mathworks Inc., Natick, MA, USA) and applied to each image slice separately. The preprocessed D dynamic scans constitute the $D \times N$ matrix of data, X. The rows $X_{1k}, X_{2k}, \ldots, X_{Dk}$ corresponds to the voxels in a slice and the columns $X_{l1}, X_{l2}, \ldots, X_{lN}$ corresponds to a time series. The nearest neighbour search, i.e. the first step of LLE, was performed using Euclidean distances as described in the original implementation of LLE [19]. The low-dimensional embedding Y was then calculated with the number of output components (d) set according to the estimated intrinsic dimensionality. For a typical single slice with $d = 10$, $K = 30$ and $N \approx 1{,}500$, the computing time amounted to 110 s on a PC equipped with an Intel Xeon 2.67 GHz processor and 3.2 GB RAM. The most time consuming step of LLE is the computation of eigenvectors which scales as $O(dN^2)$. The dimensionality reduced $d \times N$ output matrix Y contains d rows, $Y_{1k}, Y_{2k} \ldots Y_{dk}$ corresponding to component images capturing underlying degrees of freedom in the original data. The number of nearest neighbours (K) was varied between 10 and 60 to investigate the sensitivity of the algorithm to the neighbourhood size. The lower limit of the interval was selected according to the criterion that LLE only can be expected to recover embeddings whose dimensionality is less than the number of neighbours [19]. This also mitigated the risk of rendering non-connected components resulting from the use of too small neighbourhoods. The depth-first-search algorithm [25] was used in order to check for the existence of non-connected components.

The upper limit was selected to avoid too large neighbourhoods which would make the algorithm lose its non-linear dimensionality reduction possibilities.

The resulting components were scaled to z-values which in this work only serve as descriptive measures without implying any statistical inferences on the results.

When creating activation maps, voxels with a z-value larger than 1 were considered activated. Characteristic time courses representing the individual component images were calculated as an average of all the voxels' time courses weighted by the contribution of each time course to the non-thresholded component, i.e. weighted by the voxel value of the output component $\vec{Y}_i$. The results from LLE processing of data were analysed separately but were also used as input to ICA (FastICA v.2.5, http://www.cis.hut.fi/projects/ica/fastica/), substituting the normal PCA preprocessing step.

Evaluation

Example data set

LLE, PCA and non-linear PCA [13] were used for dimensionality reduction of the example data set using a target dimensionality of $d = 2$ as estimated using BIC. The resulting 2D representation was plotted and inspected visually with respect to the possibility to separate activated groups of voxels from non-active voxels in the two components, i.e. in the projections on the two orthogonal axes. A group of active voxels was thus considered completely separated from the group of non-active voxels if it was possible to draw a vertical or horizontal straight line (corresponding to a threshold) between the groups.

Task-related components

Evaluation of LLE regarding the ability to find task-related components was performed by ROC analysis of the simulated data sets and by linear correlation with a reference function for the volunteer data sets (finger tapping). The correlation measure is sensitive to inter-subject variations, but since the objective here was to characterise LLE with regard to the choice of parameters, a relative measure of performance within each subject was considered adequate. For ROC analysis, one slice containing five activation areas with different activation levels was selected from the preprocessed simulated data set. The mean true-positive ratio was taken as a measure of performance and was calculated for the false-positive ratio interval [0, 0.1] [26] for each of the five known activated regions (3×3 pixels). The LLE analysis with subsequent ROC analyses was done for a varying number of neighbours ($K = 10$–60) in order to evaluate the algorithms sensitivity to the K parameter. Results from the LLE analysis was also used as an input to ICA and a corresponding ROC analyses for different number of neighbours was done. Since the LLE analysis could result in several components which visually appeared to be task related, the component with the highest true-positive ratio was always selected. In order to mitigate effects relating to random initial guesses when performing ICA, this step was repeated ten times. Hierarchical clustering was then used to verify the components.

For the volunteer finger tapping data sets ($n = 5$), LLE and LLE-ICA analysis was done on an individual basis and separately for each of the slices covering the brain. The resulting components were inspected visually in order to determine whether patterns reflected relevant activated areas for a motor paradigm as known from conventional model based analyses, i.e. primary motor cortex, supplementary motor area, premotor cortex and cerebellum. For each subject, one slice representing primary motor cortex was then selected for further evaluation with respect to task-related components. For that particular slice, the characteristic time course for each component, calculated as described earlier, was correlated with a reference function (a step function convolved with the canonical HRF used by SPM5), and the linear correlation coefficient was taken as a measure of performance. As for the simulated data, the number of neighbours was varied ($K = 10$–60), and the correlation coefficient was calculated both for LLE alone and with LLE used as an input to ICA. The component resulting in the highest correlation coefficient was always taken as the task-related component.

Motion-related components

In order to investigate the capability of LLE to find motion-related components, linear correlation with the known time courses of translation (normalised values between zero and one) in the x-, y- and z-directions for the simulated data sets (not motion corrected or spatially smoothed) were calculated. This was also done for the components resulting from the LLE analysis alone and with LLE as a dimensionality reduction step before ICA.

Comparison of LLE-ICA and PCA-ICA

Both LLE and PCA were used as a dimensionality reduction step before ICA, and a comparison of the results was made for simulated and volunteer data with respect to the ability to find task-related and motion-related components. The comparison was performed for $K = 30$, a value within a fairly stable interval when evaluating the sensitivity of LLE to the K parameter and the target dimensionality was chosen according to the BIC. Correlation with the known reference functions and motion parameters was used as a performance measure both for the simulated and for the volunteer data. All calculations were done for one image slice at a time.

Resting-state experiment

In order to evaluate the ability of LLE to find resting-state networks, a single slice located superior to the ventricles was selected for analysis of the resting-state data set. The location of the selected slice was selected to include areas which are part of several known resting-state networks, including the default mode network and the sensorimotor network [8].

Results

Non-linear example data

Figure 2 shows the resulting two components after dimensionality reduction plotted against each other for the three algorithms under comparison. Neither PCA nor non-linear PCA managed to separate the two groups of activation from the cluster of non-active voxels (the group corresponding to stationary activation, shown in dark red, is embedded in the group of non-active voxels). The subsequent ICA step did not improve the mappings. For LLE, a separation of both groups from the non-active voxels was obtained, although they are not properly separated in terms of the two components before the ICA step. After the ICA step, however, one group of active voxels was entirely separated in the first component, and the other group was separated in the second component. LLE managed to obtain stable decompositions for neighbourhood sizes (K) between 12 and 30.

Task-related components

For the motion corrected simulated data set, the intrinsic dimensionality was calculated to be 2 and the target dimensionality for the LLE algorithm was selected accordingly ($d = 2$). Among the output components, task-related component images were consistently present for the investigated interval of selected nearest neighbours (K), i.e.10–60. Task-related components were also found for the simulated data set with high-grade motion where the intrinsic dimensionality

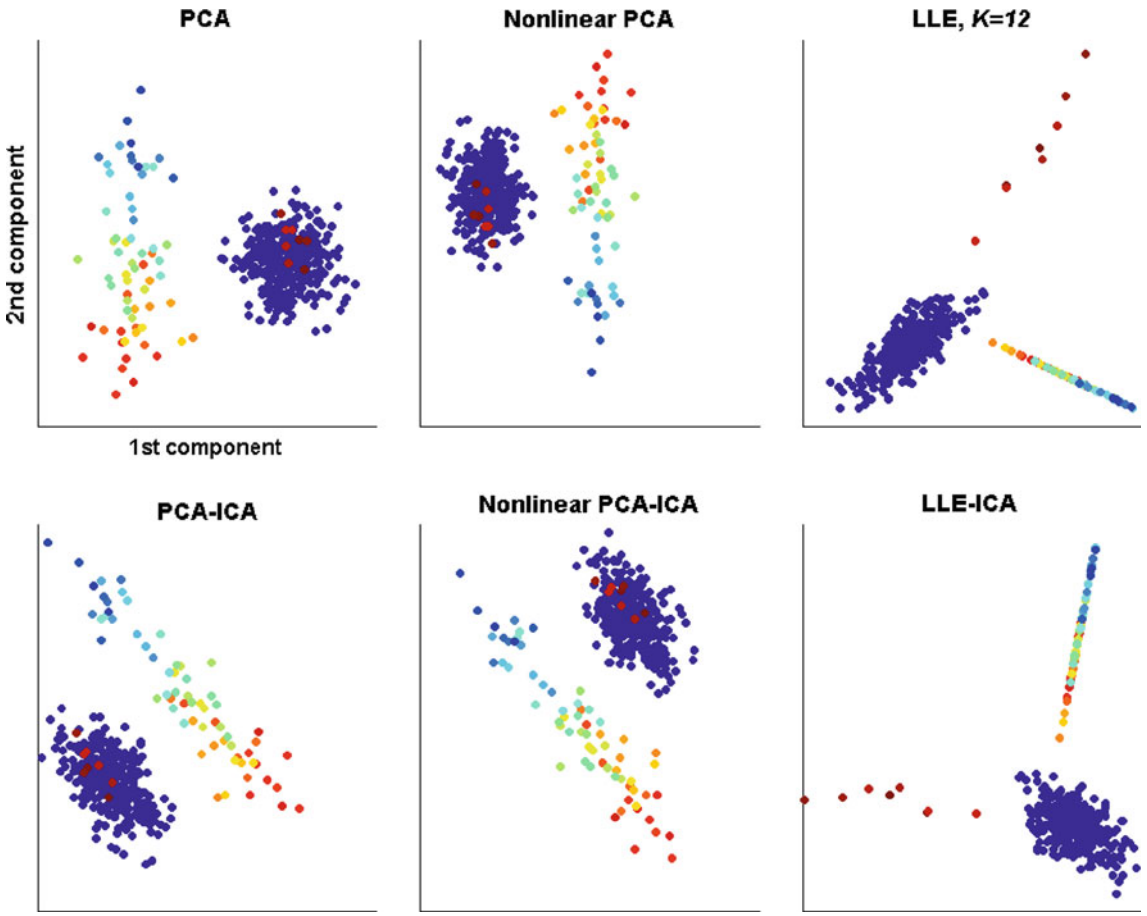

Fig. 2 Two dimensional decompositions after dimensionality reduction of the example data set (*top row*) and after an additional ICA step (*bottom row*). Dimensionality reduction was done using, from *left to right*, PCA, non-linear PCA and LLE ($K = 12$). The x- and y-axes correspond to the first and second components, respectively

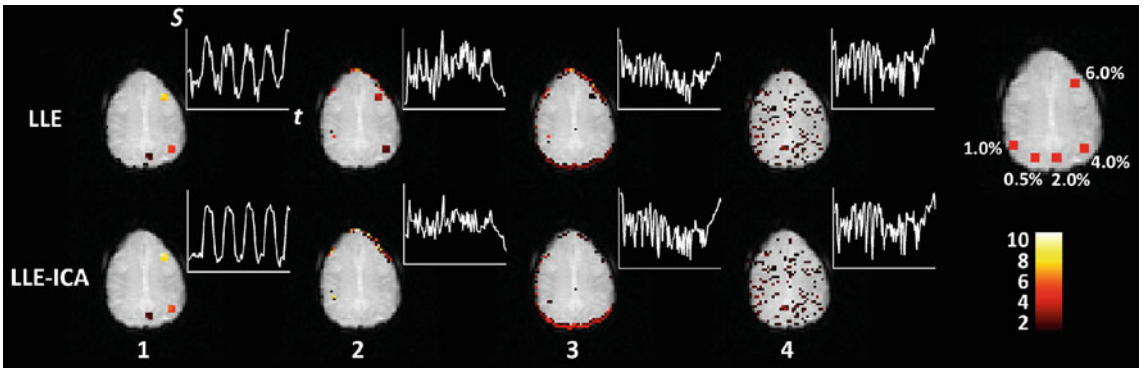

Fig. 3 Simulated data with high random motion: Thresholded example results and corresponding time courses obtained with LLE ($K/d = 30/4$, *top*) and LLE-ICA (*bottom*). The activation mask and corresponding levels of activation is also shown. Note the splitting of paradigm related activation between components 1 and 2 (*top row*), whereas only component 1 shows paradigm related activation after the ICA step (*bottom row*)

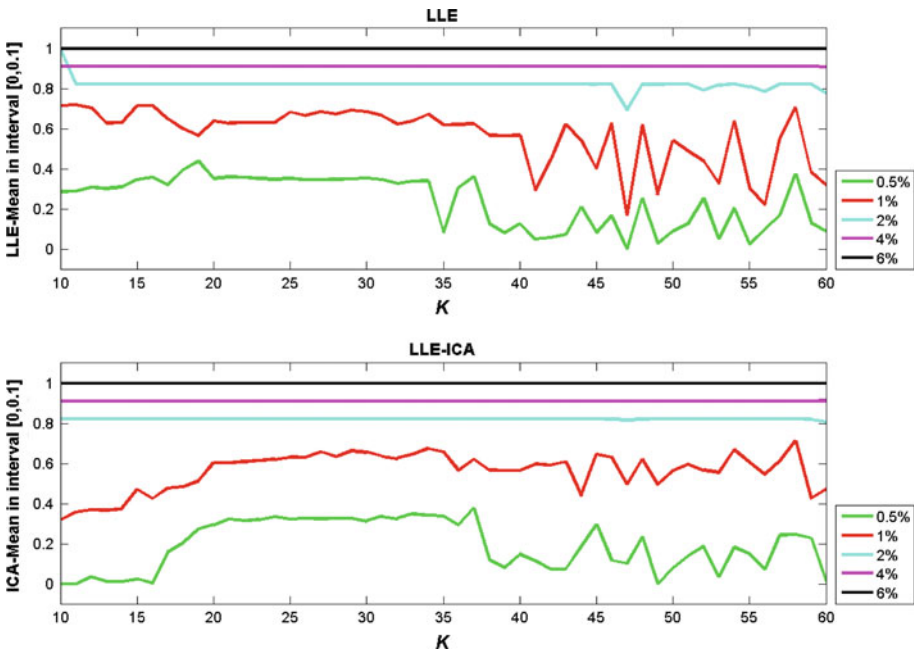

Fig. 4 Simulated data: ROC analysis of the five different levels of activation for neighbourhood sizes between 10 and 60. The graphs show mean value in false-positive range 0–0.1 for LLE processed data (*top diagram*) and for ICA decomposition of the components obtained from the LLE dimensionality reduction step (*bottom diagram*)

was calculated to be 4. Figure 3 shows, as an example, the resulting thresholded components ($z > 1$ considered active) and their corresponding time courses obtained using LLE alone and when performing ICA on the resulting components from LLE for this data set. As a reference the activation mask (five regions with different known percent signal change) is also shown. Task-related components were consistently found also when performing ICA on the components obtained with LLE. A difference was noted in that LLE could result in more than one task-related component (Fig. 3, components 1 and 2) while LLE-ICA resulted in only one single task-related component image (Fig. 3, component 1). It can be noted that the time course for component 1 has improved with regard to correspondence with the block paradigm after the ICA step.

For simulated data, the performance was found to be somewhat dependent on the choice of neighbourhood size, which can be seen in Fig. 4. For both LLE alone and for LLE as a preprocessing step to ICA, there was an interval with quite high and stable ROC mean values for values of K between 20 and 35. A noticeable drop off was then seen for the lowest activation levels, with lower and less stable ROC scores at K values above approximately 35. It can also be noted that for K lower than approximately 20, the ROC scores for LLE-ICA were dropping off for the lowest levels of activation.

For the volunteer data sets, the intrinsic dimensionality was estimated to be between 7 and 11 and LLE results again showed consistent task-related activation patterns. Figure 5 shows components with patterns corresponding to expected activated areas in primary motor cortex and cerebellum.

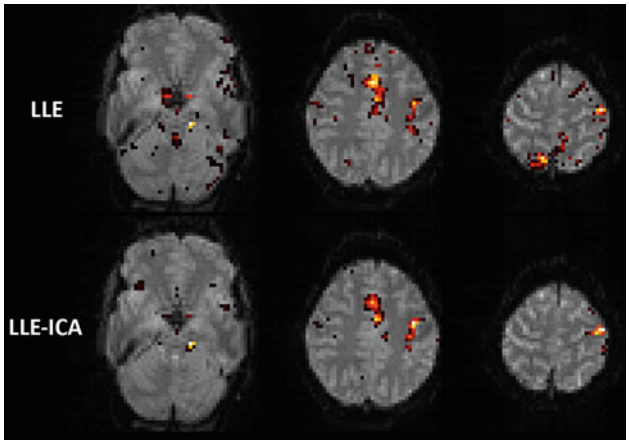

Fig. 5 Example slices from volunteer data. Results showing task-related component images obtained with LLE ($K/d = 30/10$, *top row*) and with ICA performed on the LLE components (*bottom row*). The components shown are the ones with the highest correlation with the reference function

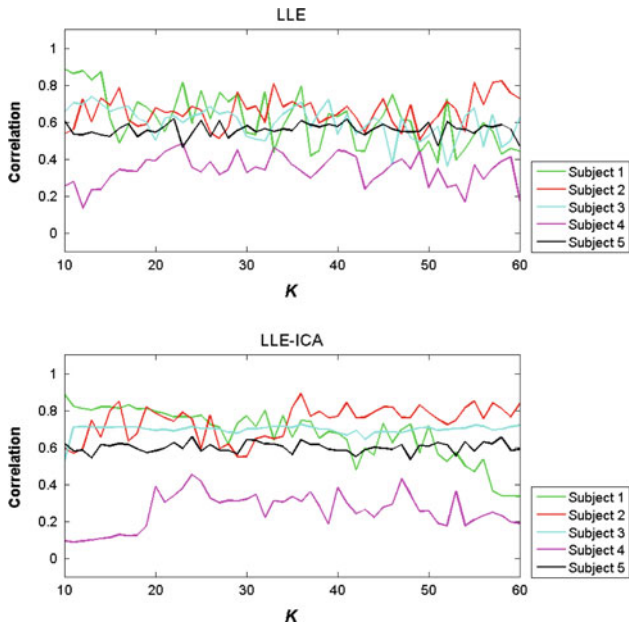

Fig. 6 Volunteer data. Performance (i.e. correlation with reference function) for different settings of the K parameter ($10 \leq K \leq 60$) for LLE processed data (*top*) and LLE-ICA processed data (*bottom*)

As was noted for simulated data, the task-related activation was sometimes present in more than one component. Correlation with the reference function showed no obvious deterioration for higher values of K (Fig. 6).

Applying ICA to the results obtained by LLE did not generally affect the dependence on K for the volunteer data, but the correlation with a reference function became more stable with less fluctuations as K varies (Fig. 6, bottom). The fluctuation (standard deviation) of the correlation coefficient for the K interval 10–60 (50 data points) was calculated for each volunteer and the average fluctuation for all volunteers (mean $\pm$ 1SD, $n = 5$) was 0.085 ± 0.038 for LLE alone and 0.077 ± 0.049 for LLE-ICA. Furthermore, in cases where LLE found several task-related components, ICA decomposition of these typically resulted in only one task-related component as was also observed for the simulated data.

Motion-related components

Some of the non-task-related components resulting from the LLE analysis showed peripheral activation with characteristic time courses that correlate well with the motion parameters. Component images with false activation due to translation can be seen in Fig. 3 for the simulated data set with high-grade random motion (components 2–4). Component 2 shows a typical motion-related component (x-direction, right-left) characterised by false activation in the right periphery, whereas components 3 and 4 shows motion-related activations in the y- (anterior–posterior) and z-direction (through-plane), respectively. As was observed earlier (Sect. Task-related components), component 2 shows a mix of task and motion-related activation, which is remedied after the subsequent ICA step.

Comparison of LLE-ICA and PCA-ICA

The comparison between LLE-ICA and PCA-ICA processing revealed that the two approaches compares well regarding correlation with the reference function for a blocked finger tapping paradigm as well as for the simulated data set (Table 1). The two methods were also comparable regarding the detection of motion (Table 1).

Resting-state experiment

The resulting 10 components obtained by ICA after dimensionality reduction by means of either LLE or PCA are shown in Fig. 7, in which the components are ordered based on visually observed spatial similarities between the results of the two methods. Patterns taken to correspond to the sensorimotor network were observed for both methods in components 1 and 2. Plausible spatial patterns in areas associated with the default mode network were found in components 3 and 4 for both methods. Activation corresponding to the medial frontal network as described in [8] was found for both methods (component 6). Both methods also resulted in obvious artefactual components with, for example, activation located in the ventricles (components 9 and 10). It is noticeable that LLE-ICA yielded sparser spatial patterns than PCA-ICA for the sensorimotor network and also in the component showing the medial frontal network. For the sensorimotor network, the right and left sensorimotor areas were split between components 1 and 2 for LLE-ICA whereas bilateral activation was captured in component 1 for PCA-ICA.

Table 1 Performance in terms of correlation with a known reference function (task response and motion parameters)

	Subject 1	Subject 2	Subject 3	Subject 4	Subject 5	Simulated data
LLE-ICA						
Task response	**0.77**	0.59	**0.70**	0.41	**0.58**	**0.93**
x-translation	0.36	**0.57**	0.59	0.63	0.33	0.59
y-translation	**0.83**	0.68	**0.84**	0.91	0.71	**0.93**
z-translation	**0.64**	**0.51**	0.46	**0.98**	0.46	**0.97**
PCA-ICA						
Task response	0.74	**0.72**	**0.70**	**0.49**	**0.58**	**0.93**
x-translation	**0.44**	0.49	**0.61**	**0.66**	**0.60**	**0.83**
y-translation	0.77	**0.70**	0.66	**0.92**	**0.77**	0.86
z-translation	0.61	0.38	**0.50**	0.97	**0.57**	**0.97**

The method (LLE-ICA or PCA-ICA) resulting in the best performance is indicated for each subject and reference function in boldface

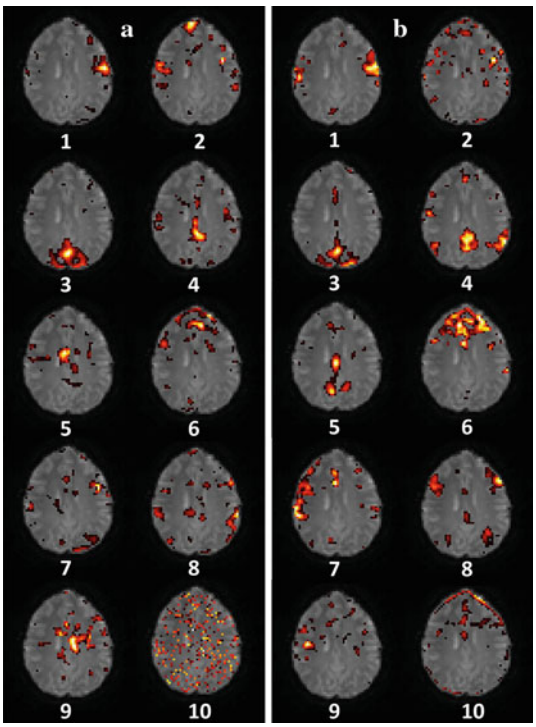

Fig. 7 Resting-state data. Panel **a** shows the 10 components obtained after ICA decomposition using LLE as dimensionality reduction method ($K = 20, d = 10$). Panel **b** shows corresponding independent components after dimensionality reduction using PCA. The component maps are thresholded at $z > 1$

Discussion

Locally Linear Embedding is a dimensionality reduction algorithm, which in this work is proposed as a novel tool for fMRI data analysis. The algorithm is easy to implement, and there are only two free parameters to set, i.e. d and K. It is data-driven and hence does not depend on any predetermined model of the hemodynamic response. An example data set with known non-linearities was used to investigate some characteristics of LLE compared with PCA and non-linear PCA, both with and without a subsequent ICA step. Simulated data sets with known added signal time courses to various voxel clusters were used to evaluate the performance of LLE alone and also LLE-ICA with respect to finding task-related components and motion. The performance with respect to finding task-related components and motion was also evaluated for volunteer data using a finger tapping block paradigm. We have shown that in a case with known non-linearities within the example data set, LLE still managed to find meaningful decompositions of data, in contrast to linear or non-linear PCA given the same target dimensionality. Our results also show that LLE was capable of finding task-related as well as motion-related components for both simulated and volunteer data, without any obvious drawbacks compared with the more commonly used linear PCA. Additionally, the use of LLE as a dimensionality reduction step before ICA on a resting-state data set results in plausible spatial activation patterns corresponding to known intrinsic functional networks [8].

The two free parameters, K and d, need to be considered with respect to 1) redundancy of the results which would be the case when d is set too high and 2) possible instability of the algorithm. The target dimensionality (d) is usually estimated when using data-driven fMRI analysis and a number of methods to accomplish this have been investigated for simulated and real fMRI data [27]. Proper estimation of the target dimensionality is a quite difficult problem, and most approaches tend to overestimate the dimensionality due to the presence of temporal autocorrelations. This is also true for the BIC algorithm used in this work. However, the estimated dimensionality for the artificial data sets used in this work agreed quite well with the prior knowledge (e.g. $d = 2$ for the example data set as well as for simulated data after realignment). For the volunteer data

(finger tapping block paradigm), it was noticed that BIC resulted in similar target dimensionality as reported in other fMRI studies [27]. Also, since the output components from LLE (as for PCA) are ordered and remain the same when new components are added and since we consistently found task and motion-related components in the present work with d estimated according to the BIC, we did not further evaluate the influence of d. There is of course a possibility that other methods to estimate the intrinsic dimensionality would imply other choices of d. It may, for example, be a limitation that BIC, and other commonly used methods for dimensionality estimation, model data linearly, but an optimal method for estimating the intrinsic dimensionality in the case of non-linear mixing remains to be developed.

With respect to the number of nearest neighbours (K) the LLE algorithm could result in smoothing or elimination of small-scale structures if a too large number is set for this parameter. In the present study, as in other studies concerning LLE [20,28], the results were typically stable over a wide range of neighbourhood sizes. However, for the simulated data sets, with added signal time courses to a known set of voxels, we found that K values above approximately 35 rendered lower ROC scores for the lower activation levels. This could be explained by the way in which the algorithm creates locally linear patches. Given the ideal case that the time courses for activated voxels are very different from those of non-activated voxels, the reconstruction of an activated voxel would be increasingly better as long as other activated voxels are selected as neighbours. If the neighbourhood size is then selected to be larger than the number of activated voxels, the reconstruction of an activated voxel would become worse since also time courses from non-activated voxels would contribute. The fact that the curves for the different activation levels split up, with a more pronounced drop off for lower activation levels as K increases (Fig. 4), could be explained by the increased possibility of the selection of a non-activated voxel when more noise is present, i.e. for voxels with the lowest levels of activation.

In contrast, too small neighbourhoods can divide the data into disjoint clusters with the implication that LLE must be applied separately to each connected component. We tested for this as part of applying the LLE algorithm and did not find disjoint clusters for any of the data sets for K between 10 and 60. Too low values of K with respect to the true number of activated voxels could also potentially render several neighbourhoods containing activated voxels which could result in true activation split over several LLE components. This was noted in some cases, in which more than one resultant component contained task-related activation patterns. In Fig. 4 (top), this is probably what can be seen as a fluctuation of ROC scores for values of K between approximately 10 and 20. This effect was, however, only observed for the lowest levels of activation. The dependence on K was also investigated after a subsequent ICA step, and in this case, the ICA step induced a few false active voxels which explains the lower ROC scores. The algorithm is hence not critically dependent on the choice of neighbourhood sizes but the selected value of K does seem to have some dependency on the expected number of true active voxels.

For the volunteer data, the behaviour described earlier concerning the choice of K was not as obvious, with larger fluctuations observed for all values of K (Fig. 6). One reason could be the lower SNR in the volunteer data, making the selection of a non-activated voxel as a neighbour to an activated voxel more likely. In vivo, a more continuous scale of different time courses, with varying relative percent signal change and varying correlation with the reference function, can also be expected. According to these observations and the fact that LLE may lose some of its non-linear capabilities for large neighbourhood sizes, we considered a choice of K within the range of 20–30 to be optimal for the simulated and volunteer data used in the present study.

LLE was also used in this work as a substitute for PCA in the preprocessing step before applying ICA. The comparison between LLE-ICA and PCA-ICA, both for simulated data with added signal time courses and for volunteer data, showed a similar performance regarding the respective algorithm's ability to find both task-related and motion-related components (Table 1). Thus, the results suggest that LLE may well be used as an alternative to the PCA method of dimensionality reduction without any obvious drawbacks. However, in this study, we have used paradigms that are known to be robust and the reference functions used for evaluation are not expected to contain non-linearities and do, therefore, not really challenge the non-linear capabilities of LLE. Other paradigms exist that could pose a problem for a linear dimensionality reduction method and which would benefit from a non-linear method. For example, Friston et al. [13] conducted an fMRI study of visual processing that addressed the interaction between colour and motion processing in order to demonstrate non-linear behaviour of fMRI data. To address the possible non-linear nature of fMRI data, we constructed an example data set with non-linear properties and we applied LLE, PCA and non-linear PCA with and without a subsequent ICA step. It was found that LLE was capable of separating the two groups of activated voxels using only two components as indicated by BIC. In contrast, neither PCA nor non-linear PCA could separate the two groups since the first group contained a large variance in two time points (dimensions). A target dimensionality of 3 would have been needed for PCA and non-linear PCA to separate the second group as well. Hence, LLE is capable of separating sources using a lower target dimensionality. This illustrates that since PCA and non-linear PCA are variance based, a large number of components might be required to find task-related components in real fMRI data that account for relatively low

variance. This would also hamper PCA as a dimensionality-reduction preprocessing step to ICA since a redundancy of input components would be generated. From the example data set, it was also seen that after the LLE step alone the two groups of active voxels were not separated into two distinct components. This was resolved by the ICA step (Fig. 2). The ICA step could, therefore, be important after the LLE step since it rotates the coordinate system so that the statistical independence of the components is maximised.

In order to further investigate the properties of LLE we also analysed a resting-state data set, and similar components were obtained for both LLE and PCA, when used as dimensionality reduction method prior to ICA. Several components showed activation patterns corresponding to known resting-state networks including the sensorimotor, medial frontal and the default mode networks. Noticeable was that LLE in some components produced sparser activation patterns and that more than one component corresponded to a single network (Fig. 7). This was interpreted as a higher sensitivity to differences in the individual time courses, although this characteristic might also indicate a lower overall sensitivity. In analogy with the results from the example data set, it is possible that LLE captures the activation in fewer components as compared to PCA, which might explain why the tenth LLE component fails to show any spatial activation. This remains, however, an open question that could be addressed in future work.

A case when LLE may potentially fail to provide sufficient characterisation of the activation sources is when some of the neighbourhoods that are integrated in the final step of LLE consist of neighbours belonging to more than one source, especially in cases where sources occur simultaneously. If the nearest neighbours do not consistently come from the same source as a result of similar time course characteristics, then the LLE algorithm will fail to characterise the sources adequately. This is the case of the algorithm mapping distant inputs to nearby outputs, a type of failure that can also occur if the original data are too noisy or if data are insufficient to ensure that the underlying manifold is well-sampled. In this work, a repetition time of 3 s was used for the volunteer data and a shortening of the repetition time could be beneficial from the perspective of sampling density.

Another potential problem can occur if the hemodynamic response in terms of signal increase and delay is more complex than, for example, in the simulated data. For volunteer or patient data, this could result in poorer correlation, although slight differences in time courses could still be captured as one component. This is illustrated in the analysis of the example data set.

Several other non-linear dimensionality reduction methods such as neural networks and self-organising maps do not share some of the favourable properties of LLE, such as guarantees of global optimality or convergence. Alternative non-linear methods also tend to involve many more free parameters, such as learning rates or convergence criteria. It is also worth noticing that, in the present study, LLE was employed in its original form and the use of different distance metrics [29] or other neighbour selection schemes might improve performance in cases were LLE may otherwise fail to capture the sources adequately. Instead of selecting a fixed number of nearest neighbours, LLE could for example select all neighbours of a given point that falls within a certain radius, which potentially could further reduce the algorithms sensitivity to the neighbourhood size. Accordingly, several extensions have been proposed in order to optimise the algorithm including a kernelized version of the algorithm. Finally, in this work, LLE was performed slice-by-slice. The last step of the algorithm requires the calculation of eigenvalues from a sparse matrix of size $N \times N$, where N is the number of voxels. This is intractable for large whole-brain data sets when implemented in MATLAB as in this work. However, more memory-efficient algorithms can mitigate this problem, making an extension to whole-brain analysis possible [30].

Conclusion

LLE is a promising tool for dimensionality reduction of fMRI time series, capable of finding meaningful activation patterns that can also be used as input to ICA as a substitute for a dimensionality reducing PCA step.

Acknowledgments This study was performed with generous support from the Swedish Research Council (Grant no. 13514), the Swedish Cancer Society, grant number 070198, the Swedish Society of Medicine, the Lund Health Care District's Research Foundations and the Knut and Alice Wallenberg Foundation. We would also like to gratefully acknowledge Dr. Anthony Waites for valuable support.

References

1. Friston K (1996) Statistical parametric mapping and other analyses of functional imaging data. In: Toga AWMJ (ed) Brain mapping: the methods. Academic Press, San Diego, pp 363–396
2. Aguirre GK, Zarahn E, D'Esposito M (1998) The variability of human, BOLD hemodynamic responses. Neuroimage 8:360–369
3. Glover GH (1999) Deconvolution of impulse response in event-related BOLD fMRI. Neuroimage 9:416–429
4. Moeller JR, Strother SC (1991) A regional covariance approach to the analysis of functional patterns in positron emission tomographic data. J Cereb Blood Flow Metab 11:A121–A135
5. Makeig S, Jung TP, Bell AJ, Ghahremani D, Sejnowski TJ (1997) Blind separation of auditory event-related brain responses into independent components. Proc Natl Acad Sci USA 94:10979–10984
6. McKeown MJ, Makeig S, Brown GG, Jung TP, Kindermann SS, Bell AJ, Sejnowski TJ (1998) Analysis of fMRI data by blind separation into independent spatial components. Hum Brain Mapp 6:160–188

7. Moser E, Baumgartner R, Barth M, Windischberger C (1999) Explorative signal processing in functional MR imaging. Int J Imaging Syst Technol 10:166–176
8. Calhoun VD, Kiehl KA, Pearlson GD (2008) Modulation of temporally coherent brain networks estimated using ICA at rest and during cognitive tasks. Hum Brain Mapp 29:828–838
9. Kim DI, Manoach DS, Mathalon DH, Turner JA, Mannell M, Brown GG, Ford JM, Gollub RL, White T, Wible C, Belger A, Bockholt HJ, Clark VP, Lauriello J, O'Leary D, Mueller BA, Lim KO, Andreasen N, Potkin SG, Calhoun VD (2009) Dysregulation of working memory and default-mode networks in schizophrenia using independent component analysis, an fBIRN and MCIC study. Hum Brain Mapp 30:3795–3811
10. Lin QH, Liu J, Zheng YR, Liang H, Calhoun VD (2009) Semiblind spatial ICA of fMRI using spatial constraints. Hum Brain Mapp
11. Sui J, Adali T, Pearlson GD, Calhoun VD (2009) An ICA-based method for the identification of optimal fMRI features and components using combined group-discriminative techniques. Neuroimage 46:73–86
12. Lund TE, Norgaard MD, Rostrup E, Rowe JB, Paulson OB (2005) Motion or activity: their role in intra- and inter-subject variation in fMRI. Neuroimage 26:960–964
13. Friston K, Phillips J, Chawla D, Buchel C (2000) Nonlinear PCA: characterizing interactions between modes of brain activity. Philos Trans R Soc Lond B Biol Sci 355:135–146
14. McKeown MJ, Sejnowski TJ (1998) Independent component analysis of fMRI data: examining the assumptions. Hum Brain Mapp 6:368–372
15. Lahaye PJ, Poline JB, Flandin G, Dodel S, Garnero L (2003) Functional connectivity: studying nonlinear, delayed interactions between BOLD signals. Neuroimage 20:962–974
16. Xie X, Cao Z, Weng X (2008) Spatiotemporal nonlinearity in resting-state fMRI of the human brain. Neuroimage 40:1672–1685
17. Thirion B, Faugeras O (2004) Nonlinear dimension reduction of fMRI data: the Laplacian embedding approach. In: Proc 2st Proc IEEE ISBI, Arlington, VA, pp 372–375
18. Shen X, Meyer FG (2006) Nonlinear dimension reduction and activation detection for fMRI dataset. In: Proceedings of the 2006 conference on computer vision and pattern recognition workshop
19. Roweis ST, Saul LK (2000) Nonlinear dimensionality reduction by locally linear embedding. Science 290:2323–2326
20. FitzGerald D, Coyle E, Lawlor B (2003) Independent subspace analysis using locally linear embedding. In: Proceedings of the 6th international conference on digital audio effects
21. Bonakdarpour B, Parrish TB, Thompson CK (2007) Hemodynamic response function in patients with stroke-induced aphasia: implications for fMRI data analysis. Neuroimage 36:322–331
22. Peck KK, Bradbury MS, Hou BL, Brennan NP, Holodny AI (2009) The role of the supplementary motor area (SMA) in the execution of primary motor activities in brain tumor patients: functional MRI detection of time-resolved differences in the hemodynamic response. Med Sci Monit 15:MT55–62
23. Pickens DR, Li Y, Morgan VL, Dawant BM (2005) Development of computer-generated phantoms for fMRI software evaluation. Magn Reson Imaging 23:653–663
24. Hansen LK, Larsen J, Kolenda T (2001) Blind detection of independent dynamic components. In: 2001 IEEE International conference on acoustics, speech, and signal processing, vols I–VI, Proceedings. pp 3197–3200, 4053
25. Lewis HR, Denenberg L (1991) Data structures & their algorithms. HarperCollins, New York, NY
26. Skudlarski P, Constable RT, Gore JC (1999) ROC analysis of statistical methods used in functional MRI: individual subjects. Neuroimage 9:311–329
27. Cordes D, Nandy RR (2006) Estimation of the intrinsic dimensionality of fMRI data. Neuroimage 29:145–154
28. Saul LK, Roweis ST (2004) Think globally, fit locally: unsupervised learning of low dimensional manifolds. J Mach Learn Res 4:119–155
29. DeCoste D (2001) Visualizing mercer kernel feature spaces via kernelized locally-linear embeddings. In: The 8th international conference on neural information processing
30. Heller EJ, Kaplan L, Pollmann F (2008) Inflationary dynamics for matrix eigenvalue problems. Proc Natl Acad Sci 105:7631–7635