fMRI data analysis by means of Locally Linear Embedding

P. Mannfolk[1], A. Waites[1,2], A. Nilsson[1], J. Olsrud[1]

[1]MR Department, Lund University Hospital, Lund, Sweden, [2]Brain Research Institute, Melbourne, Australia

Purpose/Introduction: Data driven analysis methods have become popular tools for the analysis of fMRI data, such as Principal Component Analysis (PCA) and Independent Component Analysis (ICA). We describe a method for analyzing fMRI data based on the Locally Linear Embedding (LLE) algorithm of Roweis and Saul [1]. The proposed method tries to preserve the local geometry of the data by approximating each data point with a linear combination of its neighbours. These weights are then used to compute a low-dimensional embedding. LLE is appealing in that there are only three parameters that need to be specified: K – the number of nearest neighbours used to reconstruct each data point, d – the dimensionality to map to (corresponding to the number of components used in e.g. ICA) and finally, r – the regularisation parameter, only needed when the original dimensionality of the data set is lower than K which is not the case in fMRI data analysis. The algorithm is easily implemented and extensions have been proposed in order to automatically determine the free parameters [2]. In this work we investigate the use of LLE as an fMRI time series analysis tool and we show this using simulated and motor activation fMRI data sets.

Subjects and Methods: For the analysis, a simulated fMRI data set comprised of 100 dynamic scans (Siemens Magnetom Allegra 3T, TR/TE=3000/30ms, 64x64 matrix) was used with 2.5% added signal in different parts of the brain depicting a block paradigm with 5 blocks of activation/resting. We also used data from a motor activation experiment acquired at the same scanner and with the same parameters as for the simulated set. The data sets were realigned, slice time corrected, normalized and smoothed (FWHM = 8 mm) using SPM2 before data analysis. Locally Linear Embedding was then performed with different settings of the parameters K and d. The resulting components were scaled to z-scores, which in this case only were used for descriptive purposes. Active voxels were then assigned as those exceeding an arbitrary z-value threshold.

Results: Task related components were consistently found when using between 10 and 25 components ($25>d>10$). Figure 1 shows two slices with the location of the simulated activation (top row) and the activation pattern determined by LLE (K=15, d=15, bottom row). Noise- and motion-related components were also found as is often seen in ICA results (not shown here). Figure 2 shows activation maps from the motor paradigm (K=25, d=25).

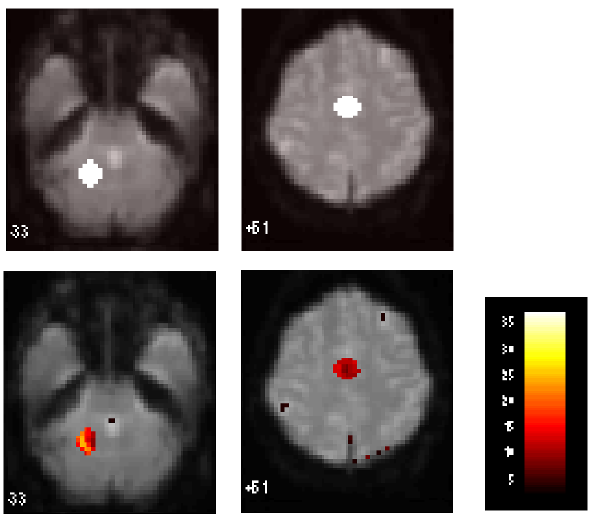

Figure 1: Simulated data

Discussion/Conclusion: We have presented a novel method for fMRI analysis. Locally Linear Embedding is straightforward to implement and when used as described above only involves two parameters: K, controlling the nonlinearity of the mapping and d, controlling the number of components used. It has been shown [3] that a reliable dimensionality reduction can be obtained over a wide range of values of K which is also observed in our experiments. However, if K is too low the components will only reflect local properties. If K is too high, the components will model the data linearly and behave like traditional PCA. Also the choice of d affects the results; if it is set too low, data from otherwise separated components might be mapped on top of each other. If d is set too high, the algorithm will enhance noise. Several methods to estimate the optimal values of K and d have been proposed, as well as extensions to make LLE even more robust. It would also be feasible to use other dimensionality reduction methods such as PCA together with LLE.

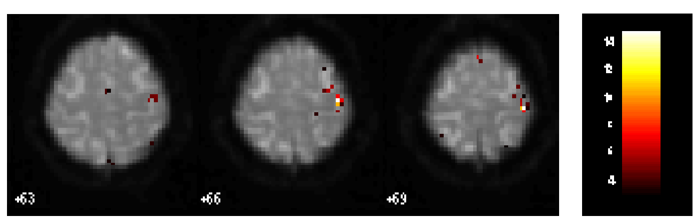

Figure 2: Motor activation results

In summary, we have shown the feasibility of Locally Linear Embedding as a means of analysing fMRI data. The algorithm manages to find activation patterns using a simple implementation and involves only two free parameters. Further work will investigate the algorithms performance in comparison with other methods and its optimization possibilities.

References:

[1] Roweis S T, Saul L K [2000] Science. 290:2323-2326
[2] Ridder D d, Duin R P W [2002] Technical Report Number PH-2002-01, Delft University of Technology, Netherlands
[3] Roweis S T, Saul L K [2003] Journal of Machine Learning Research. 4:119-155