Blind Spectra Decomposition of MRSI of the Brain with Tumor by Sparse Component Analysis

C. Chang[1], X. Mao[2], Y. S. Hung[1], J. Yang[1], P. C. Fung[1,3], D. C. Shungu[2], E. X. Wu[1]

[1]Dept. of Electrical & Electronic Engineering, The University of Hong Kong, Hong Kong, China, People's Republic of, [2]Citigroup Biomedical Imaging Center, Weill Medical College of Cornell University, New York, NY, United States, [3]Dept. of Medicine, The University of Hong Kong, Hong Kong, China, People's Republic of

Introduction

Magnetic Resonance Spectroscopy Imaging (MRSI) is extremely valuable for analyzing brain tumor metabolites *in vivo*. Several tissue types may exist in the tumor lesion area, and thus the measured MR spectra are a linear superposition of the constituent spectra of each of the different tissues. For quatitation analysis of the different tissues through MRSI, we need to first extract the individual MR spectra from the measured mixture. This is a typical linear unmixing problem or blind source separation problem. The metabolic heterogeneity in brain tumors makes this source separation problem well determined. Bayesian spectra decomposition (BSD) [1] and constrained nonnegative matrix factorization (cNMF) [2] have been successfully applied to this problem. Here in this abstract we describe a novel approach based on sparse component analysis [3], a recently developed method for blind separation of sparse sources. It is applicable since the MRS signals are generally very sparse in nature, and the results show that good spectral decomposition can be achieved by sparse component analysis.

Methods

The MR spectra for the voxels in and around the tumor lesion area are represented by a data matrix, denoted by **X**. Principal component analysis is first applied to the data matrix to determine the number of sources, *M*, and remove certain noises. The projected new data matrix is denoted as **Y**, with the same dimensions as **X**. Then, the spectra in **Y** are clustered into *M* clusters based on similarity between the spectra. Each cluster is further averaged, and the averaged spectra are represented by a new data matrix **Z** with *M* rows. Sparse component analysis is then applied to the dimension-reduced matrix **Z** to get a number of *M* source spectra, denoted as **S**. With **S** estimated, the mixing matrix **A** can be estimated by least squares based on the model **Y=AS**. Each column of **A** contains the relative magnitude of the spectra in the tumor lesion area of the corresponding tissue. Sparse component analysis tries to make the decomposed spectra as locally sparse as possible, and if the sources are both sparse and have non-overlapping peaks then the decomposed spectra will be an estimate of the true source spectra. However, it does not assume the non-negativity of the spectra, and in non-ideal cases the separated spectra may have negative part thus such decomposition should be discarded.

Experiments and Results

The sparse component analysis approach is demonstrated by a set of experimental 1.5T proton MRSI magnitude data of human brain with tumor lesion, as shown in Figure 1. The following analysis was carried out without knowing the nature or biochemical signature of the tumor. Based on the location of the tumor lesion in the MRI image, we first select an area with 11x10 voxels that contain the lesion, from the MRSI data with 32x32 pixels. Principal component analysis to the data matrix containing the spectra in the selected area shows that there are two source components that produce the mixtures. A simple clustering algorithm is then applied to the data matrix to form two clusters with centroids as shown in Figure 1(a). The sparse component analysis applied to these two transformed spectra gives two source spectra, with one shown in the bottom of Figure 1(d), denoted as source component #2, which is positive almost everywhere, and the other source spectrum has significant percentage of negative values, which is thus discarded (not shown in the Figure). We then include one of the transformed spectra to be the sources, as shown in the top of Figure 1(d), denoted as source component #1. Then the mixing matrix **A** is estimated using least squares. Promisingly, almost all the elements of **A** are non-negative, with a very small number of elements slightly less than zero, which are then forced to be zeros. With the estimated mixing matrix **A**, the relative magnitudes of the spectra in the selected area, according to each estimated source spectra, are displayed in Figure 1(e). We can see that the lesion area contains more tissues represented by component #1, but less tissues represented by component #2, than the surrounding area. Given that the tumor is a choroids plexus carcinoma (CPC), the analysis here has successfully identified the component #1 or active tumor tissue. It is dominated by the choline peak, which is the known tumor marker. CPC tumor is not of neuronal origin, thus the neuronal marker NAA is mostly absent in component #1. Component #2 represents the tissue surrounding the tumor. NAA peak is present, and choline peak is moderately elevated because there are some tumor cells in the surrounding tissue.

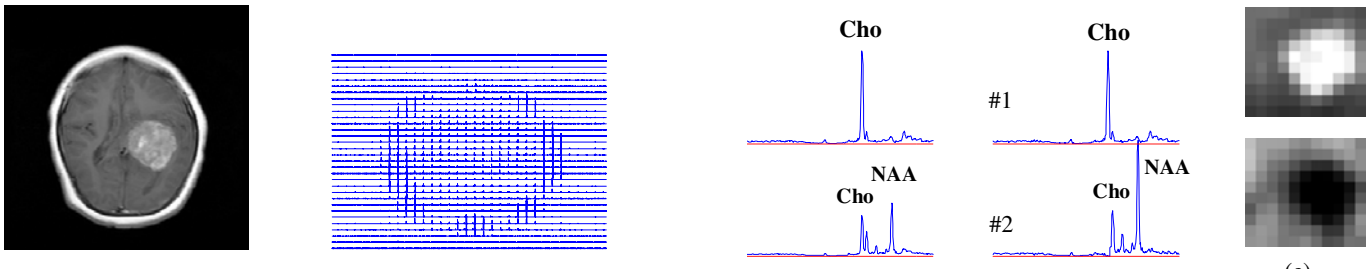

Figure 1: (a) the MRI image (256x256) of a brain with tumor; (b) a plot of the MRSI data (32x32 voxels); (c) the cluster averages of the spectra in the lesion area; (d) the estimated source spectra by sparse component analysis; (e) the distribution of the MRS around the lesion area (11x10 voxels) according to each source spectra.

Conclusion

In this abstract we present a novel sparse component analysis approach, which is a solution of blind source separation when the source signals are sparse, to the MR spectra decomposition problem, and apply it to an experimental MRSI data set of a brain with tumor lesion. The result is promising. The method is potentially valuable in the quantitative interpretation of clinical MRSI and high-resolution NMR data. Compared to the existing BSD and cNMF approaches, this method is relatively simple and expected to be robust. The result presented here is preliminary, and further study is being conducted to refine and validate this method. To further improve the result and avoid negative component, preprocessing to remove the baseline drift and advanced clustering methods to get better clustering, will be performed.

References

[1] M. F. Ochs, et al., *Journal of Magnetic Resonance*, vol. 137, pp. 161.176, 1999. [2] P. Sajda, et al., *IEEE Trans. on Medical Imaging*, vol.23, pp.1453-1465, 2004. [3] C.Q. Chang, et al, *Journal of Magnetic Resonance*, vol. 175, pp. 242-255, 2005.