BLIND SEPARATION OF TISSUES IN MULTI-MODAL MRI USING SPARSE COMPONENT ANALYSIS

Alexander M. Bronstein, Michael M.Bronstein

Department of Computer Science
Technion - Israel Institute of Technology
Haifa 32000, Israel

Michael Zibulevsky, Yehoshua Y. Zeevi

Department of Electrical Engineering
Technion - Israel Institute of Technology
Haifa 32000, Israel

ABSTRACT

We pose the problem of tissue classification in MRI as a Blind Source Separation (BSS) problem and solve it by means of Sparse Component Analysis (SCA). Assuming that most MR images can be sparsely represented, we consider their optimal sparse representation. Sparse components define a physically-meaningful feature space for classification. We demonstrate our approach on simulated and real multi-contrast MRI data. The proposed framework is general in that it is applicable to other modalities of medical imaging as well, whenever the linear mixing model is applicable.

1. INTRODUCTION

Tissue classification for diagnosis has been widely addressed from the viewpoint of machine learning and image processing [1]. Magnetic resonance imaging (MRI) is especially useful for this task owing to its ability to image tissues characterized by their magnetic properties (spin-lattice relaxation time T_1, spin-spin relaxation time T_2 and proton density PD). By appropriately choosing pulse sequence parameters (echo time T_E and repetition time T_R), tissue properties can be emphasized, producing a set of images with different contrast.

Roughly speaking, brain tissues, for example, can be thought of as consisting of water and fat in different proportions. These substances have different spin properties, and hence contribute differently to the resulting MR image when different contrasts are used. The underlying physical model in MRI suggests that such "mixing" is linear. This principle is used in the *2-point Dixon method* [2] for fat suppression. The Dixon method requires the images to be acquired exactly in-phase and out-of-phase, such that one component can be removed by simple averaging of the two images. This exact phase relation cannot always be easily achieved, e.g. due to inhomogeneities of the field.

In this study, we consider a more general *blind source separation* (BSS) framework, which can be used on generic multi-contrast MR data. We solve this problem by means of Sparse Component Analysis (SCA). This approach is based on the assumption that typical MR images, like other natural images, can be sparsely represented by an appropriate transformation. We address the problem of finding optimal sparse representation for such images. SCA produces a physically-meaningful feature space, wherein tissue classification can be carried out using simple linear methods.

This research has been supported by the HASSIP Research Network Program HPRN-CT-2002-00285, sponsored by the European Commission and by the Ollendorff Minerva Center for Vision and Image Sciences.

2. THE LINEAR MIXTURE MODEL

Let S_1 and S_2 denote two $N_x \times N_y$ *source* images, representing concentration of the basic two components (fat and water) of the brain tissue. In multi-contrast MRI, we produce a set of M *mixtures* X_i, given by linear combinations

$$X_m \quad = \quad a_{m1}S_1 + a_{m2}S_2, \quad m = 1...M, \tag{1}$$

and possibly contaminated by noise (accounting also for the presence of other substances with properties different from those of water and fat). In matrix form, (1) can be rewritten as

$$X \quad = \quad A \cdot S, \tag{2}$$

where A is a $M \times 2$ *mixing matrix*, S is $2 \times N_x Ny$ matrix consisting of source images parsed into row vectors, and X is $M \times N_x N_y$ matrix of mixtures constructed similarly. The mixing matrix represents the relative response of water and fat at each contrast. Note that unlike the Dixon method, here we assume arbitrary chosen contrasts.

We assume no *a priori* knowledge of A, except that $M \geq 2$ and $rank(A) = 2$. In addition, we assume without loss of generality that the sources have zero mean. When $M > 2$, X is a redundant representation of S (at least in the zero-noise case) with $M - 2$ linearly dependent combinations of S_1 and S_2. This redundancy can be removed by reducing the dimension of X to 2 for example by using PCA. As a result, one obtains a $2 \times N_x N_y$ matrix

$$Y \quad = \quad \Phi \cdot X = A' \cdot S, \tag{3}$$

whose rows are the first two principal components. The matrix Φ denotes the PCA projection matrix.

3. SPARSE COMPONENT ANALYSIS

Our goal is to estimate the sources S, representing water and fat concentrations, given $Y = A' \cdot S$, where A' is a 2×2 unknown invertible matrix. This problem is usually referred to as *blind source separation* and is often solved by means of Independent Component Analysis (ICA). The assumption of statistical independence of sources can be relaxed and replaced by the assumption of their sparseness. This gives rise to the Sparse Component Analysis (SCA), introduced in [3]. We use the quasi-maximum likelihood algorithm [4, 5] for SCA.

Let W be an estimate of $(A')^{-1}$. Assuming that S is i.i.d., stationary and white, the likelihood of the data Y given W is

$$\begin{aligned} p(X|W) &= p_S(WX) \cdot |\det W| \\ &= \prod_{i,j} p_S\left((WX)_{ij}\right) \cdot |\det W|\,. \end{aligned} \tag{4}$$

Hence, the normalized minus log-likelihood function is

$$L(W;X) = -\log|\det W| + \frac{1}{N_x N_y}\sum_{i,j}\varphi\left((WX)_{ij}\right), \tag{5}$$

where $\varphi(s) = -\log p_S(s)$. Source images arising in MRI usually have non-log-concave, multi-modal distributions, which are difficult to model and are not suitable for optimization. However, consistent estimator of W can be obtained by minimizing $L(W;Y)$, even when $\varphi(s)$ is not exactly equal to $-\log p_S(s)$. We will choose $\varphi(s)$ to be a smooth approximation of the absolute value, which is a good choice for sparse sources. Although real MR images are usually far from being sparse, we show in the sequel how to transform general classes of images into sparse ones.

There exist various ways for minimizing $L(W;Y)$; we use the fast relative Newton algorithm, proposed in [5]. Once W is found, the sources are estimated according to

$$\hat{S} = W \cdot Y. \tag{6}$$

$\hat{S}$ are sometimes referred to as *sparse components* (SCs). It must be emphasized that the sources can be estimated up to an arbitrary scaling factor and permutation only. Particularly, this implies that SCA (like ICA) allows to find only the relative concentration of water and fat in each pixel of the MR image, and additional priors must be applied in order to decide which of the two SCs corresponds to fat and which to water.

4. SPARSE REPRESENTATION OF SOURCES

The sparsity prior used in the quasi-ML function (5) is valid for sparse sources and not valid for original MR images. On the other hand, it is especially convenient for the underlying optimization problem due to its convexity. While it is difficult to model actual distributions of MR images, it is much easier to transform an image in such a way that it fits the sparsity prior.

Let us assume that there exists a *sparsifying transformation* $\mathcal{T}_S$, which makes the sources S_1, S_2 sparse. Our algorithm is then likely to produce a good estimate of the unmixing matrix W. Not that the same transformation is applied to each of the sources, i.e.

$$S_i' = \mathcal{T}_S S_i, \tag{7}$$

using the short notation $S' = \mathcal{T}_S S$ to denote sources that underwent sparsification and arranged in rows of a $2 \times N_x N_y$ matrix.

The problem is that in the blind deconvolution setting, S is not available, and we can apply $\mathcal{T}_S$ to the PCs Y only. Hence, it is necessary that the sparsifying transformation commute with ΦA, i.e.

$$\Phi A(\mathcal{T}_S S) = \mathcal{T}_S(\Phi A S) = \mathcal{T}_S Y, \tag{8}$$

such that applying $\mathcal{T}_S$ to Y is equivalent to applying it to S. Obviously, $\mathcal{T}_S$ must be linear in order to satisfy (8).

Thus, we obtain a general BSS algorithm, which is not limited to sparse sources. We first sparsify the PCs Y using $\mathcal{T}_S$, and then apply the SCA algorithm on Y'. The obtained unmixing matrix W is then applied to Y to produce the source estimates.

4.1. Optimal sparse representation

In [6] wavelet packet transform was proposed as a candidate for $\mathcal{T}_S$ using empirical considerations. Such sparsification is usually not optimal. For best performance of the algorithm, it is required that $\mathcal{T}_S$ be the best sparsifying transformation possible. We show a way to find such a transformation. For simplicity we limit our attention to linear shift-invariant (LSI) transformations, i.e. $\mathcal{T}_S$ that can be represented by convolution with a *sparsifying kernel* $\mathcal{T}_S S_i = T * S_i$.

The sum of absolute values in the prior term in the quasi-ML function can be used as the objective to find a kernel T which will yield the sparsest image, i.e.

$$\min_T \sum_{i=1}^{2}\sum_{mn} |(T * S_i)_{mn}| \qquad s.t. \quad \|T\|_2^2 = 1, \tag{9}$$

where the constant energy constraint is posed on T to avoid the trivial (zero kernel) solution. Since in practice S_i are not known, problem (9) can be solved using some similar images instead of the actual sources [7, 8, 9].

5. GEOMETRIC BLIND SOURCE SEPARATION

Another remarkable property of sparse components approach is that it allows performing very simple *geometric* BSS. Since most pixels in the source images have a near-zero magnitude and the locations of the non-zero values in the sources are usually independent[1], there is a high probability that only a single source will contribute to a given pixel in each mixture [3, 10]. Consequently, the majority of the non-zero-valued pixels in each mixture will be influenced by one source only and have a magnitude equal to that of the source multiplied by the corresponding coefficient of the mixing matrix.

In the scatter plot of one mixture versus the other these pixels will therefore be clustered along lines (each corresponding to a source) at a distance from the center depending on source magnitude. Hence, it is possible to reveal the ratios of each source's contribution to the mixtures by measuring the angles of each of the lines. Figure 2 depicts scatter plots of the mixtures. The orientations corresponding to the columns of the mixing matrix are clearly visible, so that one can measure the obtained angles and thereby restore the matrix entries. [11].

6. SIMULATION RESULTS

The performance of the proposed methods was assessed using simulated brain MRI data, obtained from the BrainWeb database [12]. Four 1 mm thick slices were acquired using the spin echo protocol with $T_R = 2500$ msec and different values of T_E, thus giving different weights to the underlying substances (Figure 1). Skull bones were manually removed from the images.

According to our underlying assumption, the four observed images $X_1, ..., X_4$ are linear mixtures of independent sources which,

[1] Note that strict statistical independence is not required here.

in our case, can be roughly divided into "water" and "fat". We have therefore attempted to extract these sources from the four mixtures using SCA. After subtracting the mean value, each of the mixtures was sparsified by convolution with a 2×2 corner-detection kernel,

$$T = \begin{pmatrix} +1 & -1 \\ -1 & +1 \end{pmatrix}, \quad (10)$$

which is obtained by solving (9) and yields the best performance for sources containing sharp edges and corners. Scatter plots of the sparsified mixtures (Figure 2) reveal two clearly visible independent sources in the data. PCA was used to project the four sparsified mixtures onto a two-dimensional space; SCA algorithm was then applied to the two principal components in order to estimate the two sources (see Figure 3, first row).

The obtained sources have obvious physical meaning: $\hat{S}_1$ is proportional to the amount of water found in each voxel, whereas $\hat{S}_2$ is proportional to the amount of fat. We normalize $\hat{S}_1, \hat{S}_2$ to the interval $[0, 1]$ and perform classification by hard thresholding: pixels where cerebral spinal fluid (CSF) is present usually have large concentration of water, therefore, we classify all pixels for which $\hat{S}_1 > 0.5$ as CSF. Similarly, pixels with $\hat{S}_1 \in [0.05, 0.5]$, $\hat{S}_2 \in [0.05, 0.5]$ are classified as gray matter, and $\hat{S}_1 \in [0.05, 0.5]$, $\hat{S}_2 > 0.5$ as white matter (Figure 4). The low threshold 0.05 is used to distinguish the tissues from the dark background. Figure 3 (second row) depicts the resulting tissue segmentation compared to the ground truth.

A similar experiment was repeated with three registered T_1, T_2 and PD weighted scans of a human normal brain (Figure 5), acquired using the spin echo sequence with $T_R/T_E = 700/20msec$, $T_R/T_E = 2200/30msec$, and $T_R/T_E = 2200/80msec$, respectively [2]. The estimated sources corresponding to "water" and "fat" are depicted in Figure 6 (left and middle, respectively). Instead of hard-threshold segmentation, fuzzy segmentation was applied in this experiment. Each pixel was assigned three values ρ_W, ρ_G, ρ_C, ranging from 0 to 1 and corresponding to the possibility to find white matter, gray matter and CSF, respectively, in that pixel. These values were computed by applying sigmoid-like functions to the relative concentration of "fat" in the pixel,

$$\rho_F = \frac{\hat{S}_2}{\hat{S}_1 + \hat{S}_2}. \quad (11)$$

Figure 6 (right) depicts the results obtained by fuzzy segmentation of brain tissues using pseudo-colors, where R,G,B represent ρ_W, ρ_G, ρ_C, respectively.

7. CONCLUSIONS

The fact that multi-contrast MR images can be considered as weighted linear mixtures of physically-meaningful source components lends itself to the BBS framework and, in turn, to the application of SCA as a tool for extraction of these components. Since the latter requires knowledge of source distributions, which are generally hard to model, we used a simple sparsity-based prior in combination with a sparsifying transformation.

Surprisingly, as we found, the optimal LSI sparsfying transformation for brain MR images is a simple corner detector. The latter performs better than non-optimal, though more general, sparse representations (e.g. wavelet packets [10]), derived from empirical considerations and used in previous studies. Further generalization of this result to a wider class of non-LSI transformations will be presented elsewhere. In addition, since our approach is not limited to MRI, we search for optimal sparsifying transformations suitable for other classes of medical images.

SCA produces a feature space which is useful in classification. In this study we used linear hard-threshold and fuzzy soft-threshold classification as an example; other classifiers can be used instead if necessary. In the case of MR images, two sparse components sufficed for accurate classification. In the general case of classification, feature spaces of higher dimensions may be required.

[2] MR brain data set 657 was provided by the Center for Morphometric Analysis at Massachusetts General Hospital and is available at `http://www.cma.mgh.harvard.edu/ibsr`.

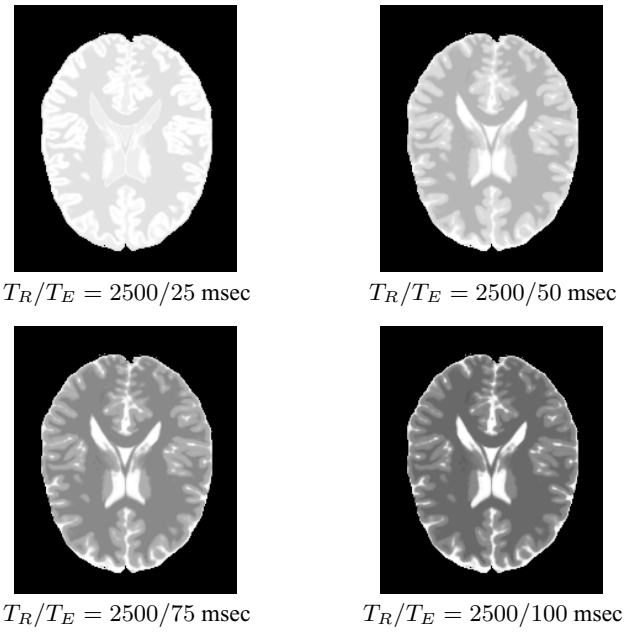

Fig. 1. Simulated MRI brain data.

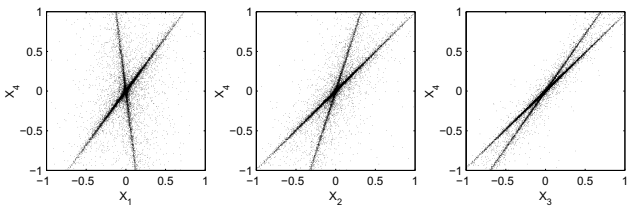

Fig. 2. Scatter plots of the normalized sparsified mixtures.

8. REFERENCES

[1] C.-M. Wang, C. Chen, Y.-N. Chung, S.-C. Yang, P.-C. Chung, C.-W. Yang, and C.-I. Chang, "Detection of spectral signatures in multispectral MR images for classification," *IEEE Trans. Medical Imaging*, vol. 22, no. 1, pp. 50–61, 2003.

[2] W. T. Dixon, "Simple proton spectroscopic imaging," *Radiology*, pp. 189–194, 1984.

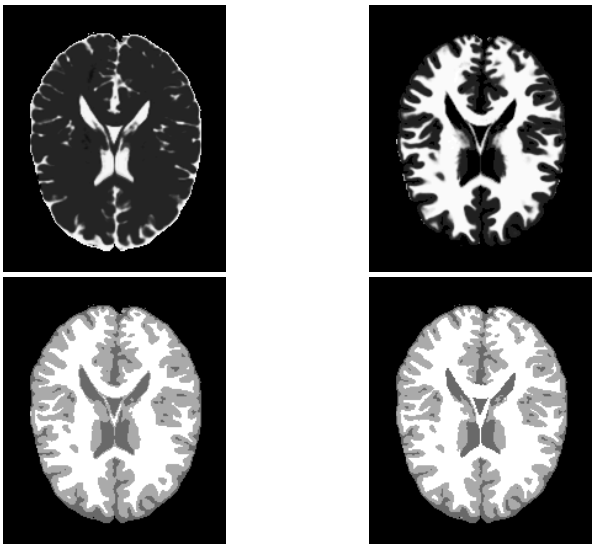

Fig. 3. Sources separated by the application of the SCA-based approach. First row: sources corresponding to "water" (left) and "fat" (right). Second row: segmentation according to the estimated sources (left) compared to the ground truth (right). The colors stand for: background (black), fluid (dark gray), gray matter (light gray) and white matter (white).

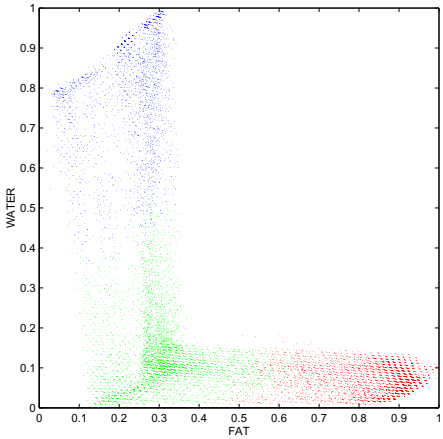

Fig. 4. Feature space spanned by $\hat{S}_1, \hat{S}_2$. Points represent pixels in the images, projected onto the estimated sources. Ground truth is shown by red (white matter), green (gray matter) and blue (gray matter).

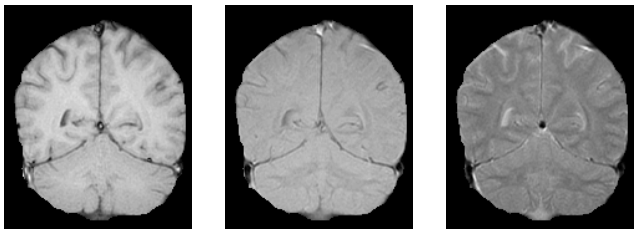

Fig. 5. Real MRI brain data. Left to right: T_1, T_2 and PD-weighted images.

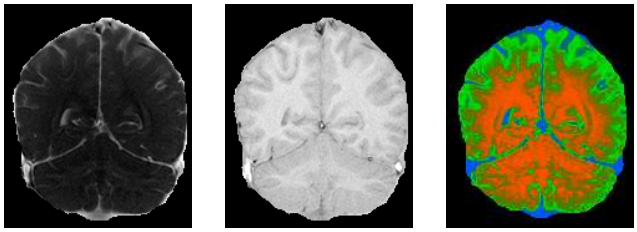

Fig. 6. Estimated sources in real brain data: "water" (left) and "fat" (middle). Right: fuzzy segmentation of brain tissues to white matter (red), gray matter (green) and cerebral spinal fluid (blue), presented in pseudo-color.

[3] M. Zibulevsky and B. A. Pearlmutter, "Blind source separation by sparse decomposition," *Neural Computation*, vol. 13, no. 4, 2001.

[4] D. Pham and P. Garrat, "Blind separation of a mixture of independent sources through a quasi-maximum likelihood approach," *IEEE Trans. Sig. Proc.*, vol. 45, pp. 1712–1725, 1997.

[5] M. Zibulevsky, "Sparse source separation with relative Newton method," in *Proc. ICA2003*, April 2003, pp. 897–902.

[6] M. Zibulevsky, P. Kisilev, Y. Y. Zeevi, and B. A. Pearlmutter, "Blind source separation via multinode sparse representation," in *Proc. NIPS*, 2001.

[7] A. M. Bronstein, M.M. Bronstein, Y. Y. Zeevi, and M. Zibulevsky, "Quasi maximum likelihood deconvolution of images using optimal sparse representations," Tech. Rep., Technion, Israel, November 2003.

[8] M. M. Bronstein, A. M. Bronstein, M. Zibulevsky, and Y. Y. Zeevi, "Optimal sparse representations for blind source separation and blind deconvolution: A learning approach," in *Proc. IEEE ICIP04*, 2004.

[9] M. M. Bronstein, A. M. Bronstein, M. Zibulevsky, and Y. Y. Zeevi, "Blind deconvolution of images using optimal sparse representations," *IEEE Image Proc.*, 2004, in press.

[10] P. Kisilev, M. Zibulevsky, and Y.Y. Zeevi, "Multiscale framework for blind source separation," *JMLR*, 2003.

[11] A. M. Bronstein, M.M. Bronstein, M. Zibulevsky, and Y. Y. Zeevi, "Separation of reflections via sparse ICA," in *Proc. ICIP*, 2003.

[12] "BrainWeb 3D MRI simulated brain database," `http://www.bic.mni.mcgill.ca/brainweb`.