Separation of MR multiband images using complex independent component analysis
CS229 Final Report

Yuxin Hu, Minda Deng, Haiyu Lu

[1]*Department of Electrical Engineering, Stanford University,* [2]*Department of Applied Physics, Stanford University,* [3]*Department of Physics, Stanford University.*

Introduction

Magnetic resonance imaging (MRI) has become a powerful medical imaging modality widely used in clinical practice. However, one of the problems of MRI is that each image requires a long scan time, especially for 3D imaging. A popular solution to this problem is "controlled aliasing in parallel imaging results in higher acceleration" (CAIPIRINHA)[1] in which the images of different slices are acquired simultaneously. Therefore, the resulting image is a mixture of the real space images. A traditional way of separating images from CAIPIRINHA is parallel imaging but this method requires additional information and time. In this paper, we applied complex-valued independent component analysis (ICA) to reconstruct images, which requires no additional information. Because the linear mixture matrix is pixel dependent, images were separated into blocks. To resolve the order and amplitude ambiguities, the adjacent blocks were partially overlapped. Finally, the separated blocks were recombined using a region-growing like method.

Related Work

One kind of MRI reconstruction methods is implemented in the frequency domain (also known as the k-space)[2,3]. Since the mixture of images in the spatial domain is under-sampled in k-space, separating mixed images is equivalent to recovering the under-sampled points in k-space. Each pixel in k-space is approximately a linear combination of its adjacent points and these weightings can be applied to the whole k-space. To estimate these weightings, some fully sampled training data in the k-space is necessary, which will take extra time. Other calibration-less reconstruction methods that require no extra training data also exist[4,5,6], but these methods work well only when the center of k-space (low-frequency component) is sampled with high density. Therefore, we want to develop a real calibration-less method with no need of any training data using complex-valued ICA.

Complex-valued signals frequently appear in various application such as communications, radar, and bio imaging. The independent component analysis (ICA) has been one of the most common and successful approaches to the blind source separation problem, under the assumption of statistically independence in the source signals.

Most of the ICA algorithms exploit one of the two following properties: non-Gaussian or sample dependence[7,8], FastICA[9], Infomax[10], and RADICAL[11] work well on non-Gaussian data but neglect the sample dependence. The second-order blind identification algorithm (SOBI)[12], the efficient algorithm for blind separation using time structure (TDSEP)[13], and the weights-adjusted second-order blind separation (WASOBI)[14] exploit sample dependence in the data, however with no ability to handling the non-Gaussianity.

In our paper, we use the mutual information rate as the cost function, which takes both the non-Gaussianity and sample independence into account[15].

Data set and features

MRI takes data in the frequency domain (k-space)

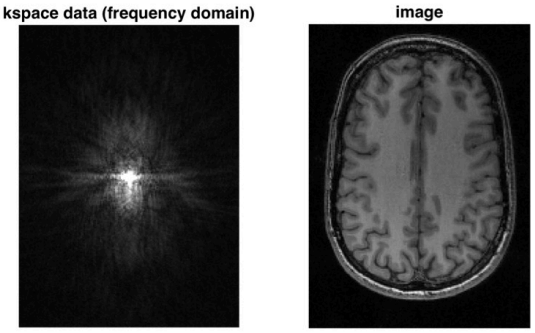

Figure 1 K-space and real images of one slice of brain combined from several recorders.

mindad@stanford.edu; yuxinh@stanfor.edu; hylu@stanfor.edu.

through phase and frequency encoding and then real space images can be constructed from the Fourier transform of the data. In the data acquisition process, multiple coils are used and they have different sensitivities due to their relative positions to the object being scanned. In MRI, these coils are like the recorders in the cocktail party problem, and the sensitivities are the same as the linear mixture matrix A. However, the sensitivities over the scanning area are pixel dependent for each coil.

As for CAIPIRINHA, two or more different slices in real space are scanned by each coil in the same time so that the scan time is remarkably decreased. However, post processing is necessary to separate different slices after scanning.
In CAPIRINHA, to make the decomposition of images easier, each slice is given a different shift as shown in Figure 2. Our goal of this project is then to separate the real space images of each slice from the mixture we get from CAIPIRINHA. Our MRI scanning data sets were collected on a 3T MR750 scanner (GE Healthcare. Waukesha, WI) with a commercially available 32-channel torso coil from the Magnetic Resonance Systems Research Laboratory at Stanford University, with acquisition matrix size = 256×256, FOV = 240mm × 240mm, slice thickness = 4mm.

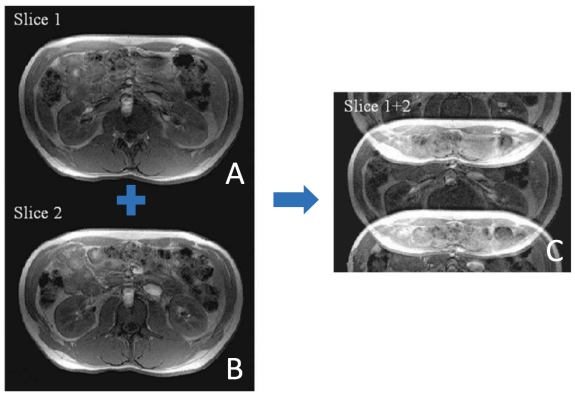

Figure 2 A typical CAIPIRINHA image C) combined from two different slices of abdomen A) and B) with shift.

We first neglected the pixel dependence of the sensitivity of each coil and tested our algorithm on two simulation data sets, which were artificially combined with two different weightings and shifts. By implementing the complex-valued ICA, it is reduced to the cocktail party problem with two speakers and two microphones.
Next, we worked on the real MRI data with pixel dependence on each coil. In other words, we can no longer use a constant linear mixture matrix A for the whole image to build the ICA model.

Methods and algorithms

Linear mixture model

Let N statistically independent complex-valued sources $s(t) = [s_1(t), \dots, s_N(t)]^T$ be mixed through an $N \times N$ nonsingular mixing matrix $A^{N\times N}$ so that we obtain the mixtures $x(t) = [x_1(t), \dots, x_N(t)]^T = A * s(t)$, $1 \le t \le T$, where the t is the discrete time index. The mixtures can be separated by constructing a demixing matrix $W^{N\times N}$ as $y(t) = W * x(t)$, where $y(t) = [y_1(t), \dots, y_N(t)]^T$ is the separation.

Cost function

A natural choice for the independent separation $y(t)$ is the mutual information rate[15], which provides a general framework to handle both non-Gaussianity and sample-dependence, among the N random variables $y_i(t)$, $i = 1, \dots, N$:

$$MIR(y_1, \dots, y_N) = \sum_{i=1}^{N} H_r(y_i) - \log\ |\det(WW^H)| - H_r(x)$$

where $H_r(y_i)$ is the entropy rate of the ith separated source defined as

$$H_r(y_i) = \lim_{T\to\infty} \frac{H(y_i)}{T} = \lim_{T\to\infty} \frac{-\sum_{n=1}^{T} P(y_i^n)\log(P(y_i^n))}{T}$$

where $H(y_i)$ is the entropy of the vector y_i, y_i^n is y_i's nth element and T is y'sdimension. Since the $H_r(x)$ is independent of W, the cost function can be written as:

$$J_r(W) = \sum_{i=1}^{N} H_r(y_i) - 2\log\ |\det(W)|$$

The role of $\log|\det(W)|$ is a regularization term. Since the entropy rate $H_r(y_i)$ is not scale invariant, i.e., $H_r(y_i) \neq H_r(ay_i)$ for $a \neq 1$, then without the regularization term the cost function can be

minimized by scaling alone.

Gradient descend VS Newton's algorithm

It's possible to use gradient descend rule to minimize the cost function $J_r(W)$ with respect to the demixing matrix W directly.

$$\frac{\partial}{\partial W} J_r(W) = \sum_{i=1}^{N} \frac{\partial}{\partial y_i} H_r(y_i) \frac{\partial y_i}{\partial W} - 2\mathrm{W}^{-\mathrm{T}}$$

To achieve faster convergence, we implement a decoupling procedure to simplify the problem into minimizing $J_r(W)$ with respect to each of its row vectors W_i. The decoupling procedure not only avoids the complicated matrix optimization problem, but also make efficient Newton algorithm become tractable after decoupling. For a Newton update, the Hessian can be computed using[15]

$$\frac{\partial^2}{\partial W_i \partial W_i} J_r(W_i) = \sum_{i=1}^{N} \frac{\partial^2}{\partial y_i^2} H_r(y_i) x x^T - \frac{H_i}{\left(W_i^T h_i\right)^2}$$

where $H_i = I - \widetilde{W}_i^T \left(\widetilde{W}_i \widetilde{W}_i^T\right)^{-1} \widetilde{W}_i$ with $\widetilde{W}_i = [W_1, W_2, \ldots, W_{i-1}, W_{i+1}, \ldots, W_N]^T$ and h_i is a unit length vector that is perpendicular to all the row vectors of W except W_i.

Pixel dependent linear mixture matrix

To apply the above complex-valued ICA method in separating mixed MRI images due to under-sampling, we need to make further assumptions. For a real-world MRI scanner, the linear mixture matrix A is pixel dependent. Therefore, we cannot directly implement complex-valued ICA assuming A is constant for all pixels in the image. However, A varies relatively slowly as a function of pixel position. This allows us to group close pixels together into blocks and assume the same linear mixture matrix for all the pixels in the same block. After obtaining the demixing matrix we then can combine the blocks together to reconstruct the unmixed images. However, the ICA algorithm suffers from the order/amplitude/phase ambiguities. That is, the output may differ from its "true" value by a complex constant and output order may vary. To solve these intrinsic ambiguities, the mixed images need to be divided into partially overlapped blocks. The overlapped region is used to sort the outputs. To illustrate how the sorting and correction work, we denote the overlapped regions from block 1 and block 2 as vectors a, b. Then we define how close two vectors are by the normalized inner product $\frac{a^T b}{|a||b|}$ and use this as the sorting criterion. After matching the order, we correct the amplitude and phase through a weighted average:

$b_{corrected} = \left(\sum_i \frac{a_i}{b_i} w_i\right) b$, where $w_i = \frac{|a_i|}{\sum_i |a_i|}$.

The whole process of our project can be divided into five steps.

Step 1: Compress the data using principal component analysis so we have the same number of recorders as speakers (slices)[16].

Step 2: Fourier transform raw data to get mixed real space images.

Step 3: Divide images into multiple partially overlapped blocks as described above. For each block, the images are mixed due to under-sampling.

Step 4: Apply complex-valued ICA to all the blocks.

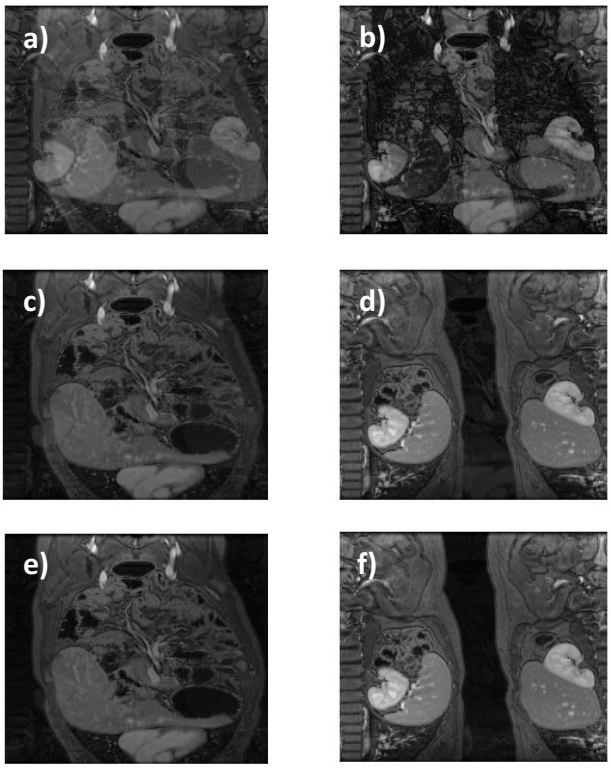

Figure 3 a) and b) are the mixture signal. c) and d) shows the independent component separated by the mutual information cost function. e) and f) shows the independent component separated by the mutual information rate cost function.

Step 5: Apply order, amplitude and phase corrections to the unmixed blocks and then recombine them block by block to reconstruct the unmixed images.

Results

In the first part of our result, we simulated ICA separation results by using mutual information and using mutual information rate as the cost function to show how mutual information rate can exploit sample dependence. We assumed that the linear mixture matrix A is constant over entire MRI image, which is usually not true in practice.

The simulation results are shown in figure 3. e) and f) indicates that the mutual information rate cost function successfully separates two independent components without artificial features which appears in c) and d).

	MI as cost function	MIR as cost function
1	0.122474487	0.023043437
2	0.119582607	0.020124612
3	0.187082869	0.020832667
4	0.154919334	0.021610183
5	0.170293864	0.022022716
6	0.126491106	0.022847319
7	0.154919334	0.022203603
8	0.187082869	0.020469489
9	0.192353841	0.021142375
10	0.161245155	0.022405357

Table1. Ten different data sets separated by mutual information cost function and mutual information rate function. Root mean square differences $(RMS(real\ IC - separated\ IC))$ *are listed.*

In this simulation, we know the real independent component. Therefore, we calculated the root mean square of the differences between separated IC with the real images to compare the mutual information and mutual information rate. We observed one order of magnitude enhancement in mutual information rate cost function.

Till this point, we can draw the conclusion that because of the sample dependence in MRI data it is necessary to implement the mutual information rate as the cost function.

In the second part of our result, we applied this complex-valued ICA method to real coil data, first without dividing them into blocks. As shown in figure 4 a)-d), clearly, the edges are not well separated, especially for the right half of the images.

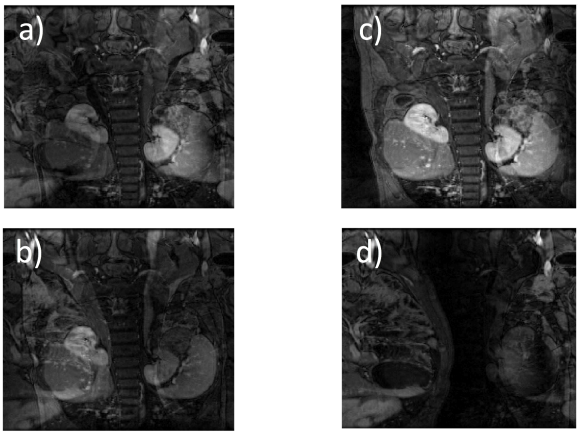

Figure 4 a) and b) show the mixed images of two coils; c) and d) show the separated images.

We then divided images into blocks before we apply the same algorithm.
As shown in figure 4 e)-h), detailed features are better separated and resolved. Also, edges are more clear. Comparing the resulting separated images, we concluded that for most part of the unmixed images, grouping pixels into blocks before applying complex-valued ICA did improve the overall performance.

Also, as shown in figure 5, dividing into blocks greatly reduces the root mean square error of the separated images compared to rescanned "true" images by 25%. The variance in the root mean square error also decreases, indicating that

complex-valued ICA with divided blocks is more robust.

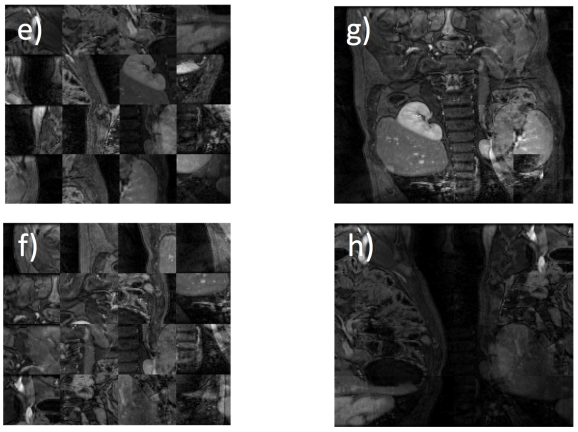

Figure 4 e) and f) show the separated blocks without order, amplitude and phase corrections; g) and h) show the recombined images by the region-growing method.

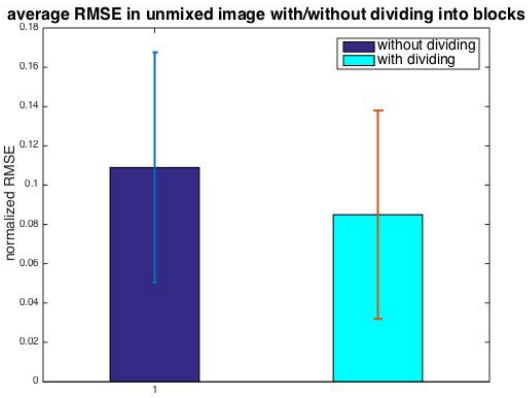

Figure 5 Average RMS error without/with dividing into blocks

Conclusion and future work

In this project, we applied cICA algorithm to separate mixed MR images due to under-sampling. Results from simulated mixed images showed that mutual information rate works better as the cost function, since it does not require independent data sets. Also, we showed that dividing original mixed images into blocks and then applying complex-valued ICA improved the performance because it accounts for the fact that the mixing matrices are pixel dependent. Unmixed images were successfully reconstructed from blocks by region growing method and detailed features were well separated and resolved.

In traditional reconstruction methods, the linear relationship between sampling points in the frequency domain is assumed and learned from the training data. We can apply more complex models to reconstruct more details. In addition, in this complex ICA algorithm, we can add some other constraints from some properties of MR images, e.g. sparsity and locally low rank[5,17], to achieve faster convergence and better reconstruction results.

Reference

[1] Breuer, Felix, et al. "Controlled aliasing in parallel imaging results in higher acceleration (CAIPIRINHA)." *Proceedings of the 20th Annual Meeting of ESMRMB, Rotterdam, Netherlands*. 2003.

[2] Griswold MA, Jakob PM, Heidemann RM, Nittka M, Jellus V, Wang J, Kiefer B, Haase A. Generalized autocalibrating partially parallel acquisitions (GRAPPA). Magn Reson Med 2002;47:1202–1210.

[3] Lustig, Michael, and John M. Pauly. "SPIRiT: Iterative self-consistent parallel imaging reconstruction from arbitrary k-space." *Magnetic resonance in medicine* 64.2 (2010): 457-471.

[4] Chen, Chen, Yeqing Li, and Junzhou Huang. "Calibrationless parallel MRI with joint total variation regularization." *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer Berlin Heidelberg, 2013.

[5] Trzasko, Joshua D., and Armando Manduca. "Calibrationless parallel MRI using CLEAR." *2011 Conference Record of the Forty Fifth Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*. IEEE, 2011.

[6] Majumdar, Angshul, Kunal Narayan Chaudhury, and Rabab Ward. "Calibrationless parallel magnetic resonance imaging: a joint sparsity model." *Sensors* 13.12 (2013): 16714-16735.

[7] Comon, Pierre, and Christian Jutten, eds. *Handbook of Blind Source Separation: Independent component analysis and applications*. Academic press, 2010.

[8] Adali, Tulay, Matthew Anderson, and Geng-Shen Fu. "Diversity in independent component and vector analyses: Identifiability, algorithms, and applications in medical imaging." *IEEE Signal Processing Magazine* 31.3 (2014): 18-33.

[9] Hyvarinen, Aapo. "Fast and robust fixed-point algorithms for independent component analysis." *IEEE transactions on Neural Networks* 10.3 (1999): 626-634.
[10] Bell, Anthony J., and Terrence J. Sejnowski. "An information-maximization approach to blind separation and blind deconvolution." *Neural computation* 7.6 (1995): 1129-1159.
[11] Learned-Miller, Erik G., and W. Fisher John III. "ICA using spacings estimates of entropy." *Journal of Machine Learning Research* 4.Dec (2003): 1271-1295.
[12] Belouchrani, Adel, et al. "A blind source separation technique using second-order statistics." *IEEE Transactions on signal processing* 45.2 (1997): 434-444.
[13] Ziehe, Andreas, and Klaus-Robert Müller. "TDSEP—an efficient algorithm for blind separation using time structure." *ICANN 98*. Springer London, 1998. 675-680.
[14] Yeredor, Arie. "Blind separation of Gaussian sources via second-order statistics with asymptotically optimal weighting." *IEEE Signal Processing Letters* 7.7 (2000): 197-200.
[15] Fu, Geng-Shen, et al. "Complex Independent Component Analysis Using Three Types of Diversity: Non-Gaussianity, Nonwhiteness, and Noncircularity." *IEEE Transactions on Signal Processing* 63.3 (2015): 794-805.
[16] Zhang, Tao, et al. "Coil compression for accelerated imaging with Cartesian sampling." *Magnetic resonance in medicine* 69.2 (2013): 571-582.
[17] Lustig, Michael, David Donoho, and John M. Pauly. "Sparse MRI: The application of compressed sensing for rapid MR imaging." *Magnetic resonance in medicine* 58.6 (2007): 1182-1195.