MULTICHANNEL IMAGE ESTIMATION VIA SIMULTANEOUS ORTHOGONAL MATCHING PURSUIT

R. Maleh   *A. C. Gilbert*

Department of Mathematics
2074 East Hall
530 Church Street
University of Michigan
Ann Arbor, MI 48109

ABSTRACT

In modern imaging systems, it is possible to collect information about an image on multiple channels. The simplest example is that of a color image which consists of three channels (i.e. red, green, and blue). However, there are more complicated situations such as those that arise in hyperspectral imaging. Furthermore, most of these images are sparse or highly compressible. We need not measure thoroughly on all the channels in order to reconstruct information about the image. As a result, there is a great need for efficient algorithms that can simultaneously process a few measurements on all channels. In this paper, we discuss how the Simultaneous Orthogonal Matching Pursuit (SOMP) algorithm can reconstruct multichannel images from partial Fourier measurements, while providing more robustness to noise than multiple passes of ordinary Orthogonal Matching Pursuit (OMP) on every channel. In addition, we discuss the use of SOMP in extracting edges from images that are sparse in the total-variational sense and extend the ideas presented in this paper to outline how sparse-gradient multichannel images can be recovered by this powerful algorithm.

***Index Terms*—** image reconstruction, image edge analysis, algorithms, Fourier transforms

1. INTRODUCTION

As human beings, we are blessed with a visual system that allows us to see in color. Any image that is captured by the eye can be thought of as a combination of three black and white channels, namely a red channel, a green channel, and a blue channel. Yet we are unaware of this process as our visual system seamlessly combines and processes all three channels. In the digital world, a colored image is really three different images, with each image corresponding to a different primary color channel. In addition, one often encounters images with hundreds of channels in certain applications such as chemical spectroscopy. Such hyperspectral images contain many channels isolating data at different frequencies. In many cases, multi-channel images enjoy sparsity in some domain. As a result, one often obtains a huge amount of data (a large image for each channel), which contains a limited amount of actual information. We are interested in seeking efficient compressed sensing algorithms that can simultaneously handle all the given channels.

One novel technique involves the use of the Simultaneous Orthogonal Matching Pursuit (SOMP) algorithm, as presented in [5]. We assume that we receive partial Fourier measurements of a multichannel image which is sparse or compressible in some domain. Several previous works demonstrate both theoretically and empirically that the Orthogonal Matching Pursuit (OMP) algorithm recovers compressible signals and images from partial Fourier coefficients. We extend these empirical observations to multichannel signals. What constitutes a channel depends on the imaging application and, as such, we construct four examples with: (i) the standard three color channels, (ii) a variable number of channels (modeling a hyperspectral imaging system), (iii) multiple directional derivatives in a single grey-scale image, and (iv) multiple directional in a multi-channel (colored) image.

In the remainder of this paper, we shall present the SOMP algorithm as well as a discussion of how it compares to several channel-wise passes of ordinary OMP (see [4] for the algorithm). As we shall see, increasing the number of channels increases our robustness to noise. We finish by introducing a new, interesting application of SOMP, involving edge detection via multiple directional derivatives. Then we extend this idea by outlining how multi-channel images that are sparse under differentiation can be efficiently recovered from their Fourier measurements.

The authors are supported by NSF DMS 0354600 and DARPA/ONR N66001-06-1-2011. ACG is an Alfred P. Sloan Fellow.
Email:{rmaleh,annacg}@umich.edu

2. PRELIMINARIES

Let $X \in \mathbb{C}^{d\times d\times K}$ be a K-channel image of spatial dimension d by d. We shall denote each individual channel as $X_k \in \mathbb{C}^{d\times d}$ with $1 \leq k \leq K$. Furthermore $X(n,m) \in \mathbb{C}^K$ shall denote the vector value of X at position (n,m). The sparsity T of X is defined to be the number of positions (n,m) where $X(n,m)$ is a non-zero vector. The special image $\delta_{(n,m)} \in \mathbb{C}^{d\times d}$ will denote the delta function centered at (n,m). We shall let Ω represent a set of N randomly chosen Fourier coefficients from the set $(\mathbb{Z}/d\mathbb{Z})^2$ We will not have access to X itself, but rather we will have knowledge of Ω and the Discrete Fourier Transform of each channel of X on the frequencies specified by Ω. Note that these are the *same* frequencies on each channel, as opposed to a *different* set on each channel. We denote the Fourier Transform of image X_k as $b_k = \mathcal{F}_\Omega X_k \in \mathbb{C}^N$. The collection of Fourier coefficients for all of X will similarly be denoted by $B = \mathcal{F}_\Omega X \in \mathbb{C}^{N\times K}$. We will let $\mathcal{F}_\Omega^*$ denote the Hermitian dual of $\mathcal{F}_\Omega$. Now assuming that X is sufficiently sparse and that we have knowledge of Ω and $\mathcal{F}_\Omega X$, we can use the following algorithm to recover X.

Algorithm 1. *Simultaneous Orthogonal Matching Pursuit*

Inputs:

1. An $N \times K$ matrix B consisting of Fourier measurements of X.
2. A set Ω of N frequencies.
3. Number of iterations (usually equal to the sparsity T).
4. Some norm $||\cdot||$ on $\mathbb{C}^K$ that will measure the size of X at each position.

Outputs:

1. A set Λ_T of T positions where the nonzero values of X occur.
2. A $d \times d \times K$ data structure $\tilde{X}$ that approximates X.
3. An $N \times K$ residual matrix R_T.

Procedure:

1. Initialize the residual matrix $R_0 := B$, index set $\Lambda_0 := \emptyset$, and the iteration counter $t = 1$.
2. Find the position (n,m) that maximizes $||\mathcal{F}_\Omega^* R_{t-1}(n,m)||$.
3. Set $\Lambda_t = \Lambda_t \cup \{(n,m)\}$.
4. Set A_t to be an $N \times K$ matrix whose kth column is the projection of the kth column of B onto $span_{(n,m)\in\Lambda_t}\{\mathcal{F}_\Omega \delta_{(n,m)}\}$. Set $R_t = B - A_t$.
5. If $t < T$, goto step 2.
6. Use a least-squares procedure to calculate an approximation $\tilde{X}$ with entries in locations corresponding to Λ_T that minimizes $||B - \mathcal{F}_\Omega \tilde{X}||_F$ where $||\cdot||_F$ denotes the Frobenius norm.

As a side note, if the norm that we input into the above algorithm corresponds to an ℓ_p norm with $1 \leq p \leq \infty$, we shall abbreviate the algorithm as SOMP_p. The algorithm presented in [5] is technically SOMP_1. We also emphasize that it is not a stringent requirement that $||\cdot||$ corresponds to a norm. It can be replaced by some other suitable atom selection measure. Although still an open problem, it is conjectured that if $N = O(T \,\text{polylog}(d))$, then $\tilde{X} = X$ with very high probability. With this framework, we now discuss the problem of sparse multi-channel image recovery from Fourier measurements.

3. SPARSE MULTI-CHANNEL IMAGES

Suppose first that X is some sparse 3-channel image (e.g., a color image) that is T-sparse. Unlike the setting presented in [1], we will have a fixed set of given frequencies Ω common for each channel. This setting is representative of situations that arise in imaging applications. Our objective is to compare the performance of SOMP_1 in this case against the performance of three independent channel-wise passes of ordinary OMP in the presence of noisy Fourier measurements. We repeatedly generated 40-sparse 3-channel 32 by 32 images and selected half the available frequencies ($N = 512$) for our set Ω. Then we added independent white Gaussian noise of various strengths onto each measurement channel. A plot of the average ℓ_1 error of our reconstructions as a function of the SNR of our Fourier measurements is shown in Figure 1.

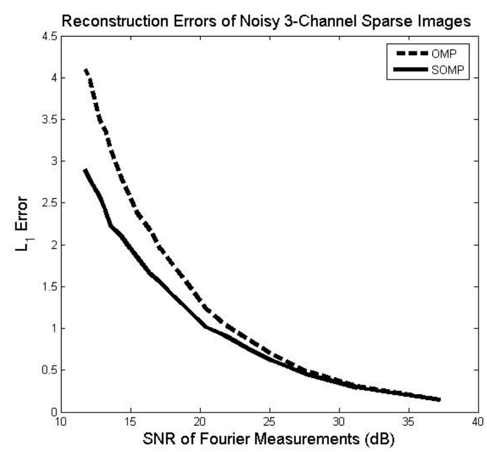


Fig. 1. The ℓ_1 error in reconstructing noisy 3-channel 40-sparse 32×32 images given 512 Fourier coefficients.

As can be seen from the plot, SOMP's gain over OMP increases considerably as noise power is increased. This is because of SOMP's simultaneous pixel selection criterion: an incorrect noisy pixel is less likely to be chosen.

For the next experiment, we compared the performance of SOMP_1 against that of ordinary OMP when the number of signal channels varied. To keep all other factors constant, we fixed $T = 40$, $N = 512$, and thus maintained a nearly fixed SNR of approximately 17 dB. A plot of the average ℓ_1 reconstruction error as a function of the number of channels is shown in Figure 2.

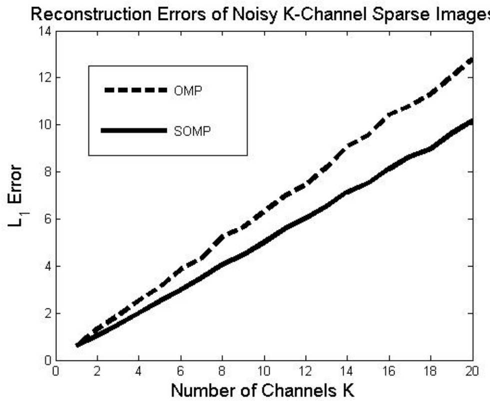


Fig. 2. The ℓ_1 error in reconstructing noisy K-channel 40-sparse 32×32 images given 512 Fourier coefficients.

In the presence of noise, using more channels reduces the reconstruction error. Since SOMP selects a common pixel for each channel at each iteration, the existence of a noisy pixel in one channel becomes insignificant when a large number of channels are present. It should be noted that we also conducted the previous two tests using SOMP$_2$ and SOMP$_\infty$, but the resulting reconstructions were essentially identical to those generated by SOMP$_1$.

4. EDGE DETECTION VIA SOMP

Another application of SOMP is that of edge detection. In this setting, let $Y \in \mathbb{C}^{d\times d}$ be a sparse gradient image, i.e. one with large constant valued regions separated by edges. Our goal is to recover the edges of Y given N frequencies Ω and $\mathcal{F}_\Omega Y$. One naive approach is to chose some direction specified by a unit vector u and then attempt to recover $\nabla Y \cdot u$ where, in the discrete world, $\nabla Y(n,m) := (Y(n,m) - Y(n-1,m), Y(n,m) - Y(n,m-1))$. This is possible using OMP; however, the edges which are parallel to the direction of u will not be recovered. If we select two or more different directions with which to examine our image, we should not miss anything. Let $K > 1$ represent the number of directions we wish to use. Set $u_k := (\cos(\pi(k-1)/2K), \sin(\pi(k-1)/2K))$ for $1 \leq k < K$ and $u_K := (0,1)$. Define the K-channel signal X by $X_k = \nabla Y \cdot u_k$. It is possible to compute $\mathcal{F}_\Omega X$ using the observation that at each frequency $(\omega_1, \omega_2) \in \Omega$, we have that

$$\mathcal{F}_\Omega X_k(\omega_1,\omega_2) = \left(1 - e^{-2\pi i\omega_1/d}, 1 - e^{-2\pi i\omega_2/d}\right) \cdot u_k \mathcal{F}_\Omega Y(\omega_1,\omega_2).$$

Thus, we can now obtain the edges of Y by recovering X via SOMP$_\infty$ and then setting the edge image Y_e equal to 1 at all positions where $\tilde{X}(n,m) \neq 0$ and 0 otherwise. The reason for selecting the ℓ_∞ norm is because at every position (n,m), we have the following approximation for large K:

$$\underset{k\leq K}{max} \, ||\nabla Y \cdot u_k||_\infty \approx |\nabla Y| \tag{4.1}$$

where the later expression is precisely the total variation operator (TV), which at each edge position, returns the magnitude of the largest jump over that edge. Thus, by using multiple derivative channels, we are effectively working with Fourier observations of the image under the total variation operator, which would be impossible to directly obtain otherwise from the original Fourier measurements due to nonlinearity. Also, by approximating the edge jump of largest magnitude, we are strengthening our pixel selection criterion against the effect of noisy observations. We note that it is also possible to reconstruct the original image from its gradients via a simple system of differential equations.

We tested this procedure on a noiseless 64×64 version of the famous Shepp-Logan Phantom image (the original is reproduced below in figure 3) with only 15% of all Fourier coefficients given. We repeatedly ran SOMP$_\infty$ on this image with different numbers of directional derivatives chosen for the edge detection (with a fixed sampling percentage of 15%). Figure 4 shows a plot of the average ℓ_1 reconstruction error of the edges as a function of the number of directions K.

Fig. 3. The Original Shepp-Logan Phantom

When $K = 1$, the ℓ_1 error was on the order of 400. This is no surprise since a single directional derivative will miss all parallel edges. We obtain a rapid drop in error after the second direction is introduced. Introducing additional derivatives does not significantly reduce the error (as 15% Fourier coefficients is probably an insufficient sampling rate for this image, regardless of the number of derivatives).

Next, we took a noiseless 128×128 Shepp-Logan phantom and reconstructed its edges from 15% of its Fourier Coefficients using 1, 2, and 3 directional derivative channels. The reconstructions are shown in figure 5. Again, we see that the biggest drop in error occurs when transitioning from one to two derivatives. With one derivative, which happens to be acting in the vertical direction, all purely vertical edges are missed, contributing to the large error. The second direction alleviates this problem. Adding a third derivative actually increases the ℓ_1 error by a trivial amount. We seem to have

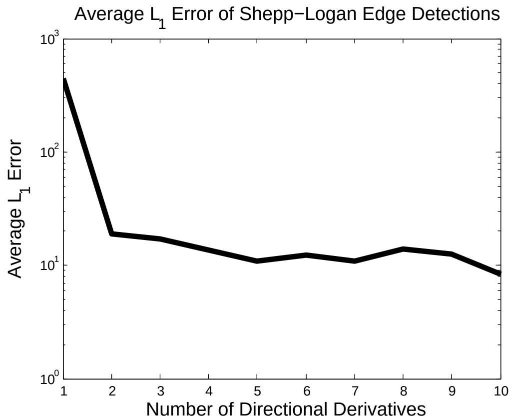
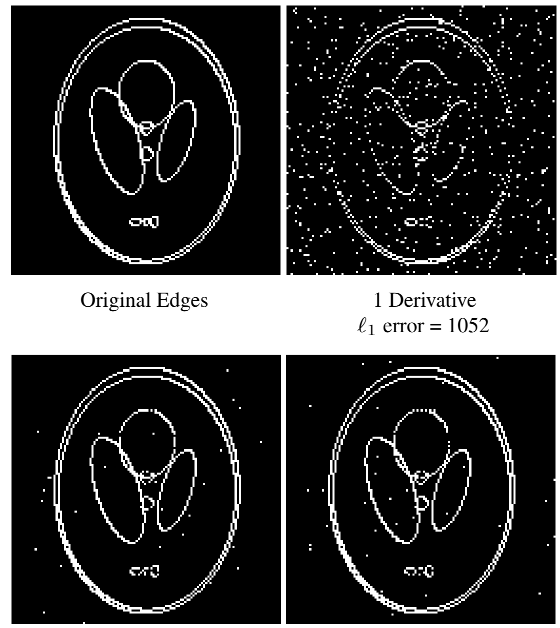


Fig. 4. The ℓ_1 error in detecting the edges of the Shepp-Logan phantom as a function of the number of directional derivatives chosen.

hit a limit as to the accuracy we can obtain using our current Fourier sampling scheme. All our experiments seem to suggest that given an edge image of some sparsity, we will require roughly the same number of Fourier measurements of that image to satisfactorily reconstruct it, regardless of the number of derivative channels that we choose. However, this should not discredit the use of multiple derivatives. They may become invaluable in situations where Fourier measurements of some set of directional derivatives are directly given a priori and one or more of these channels are corrupted by noise or other artifacts. In addition, Figure 5 also suggests that the incorrectly chosen pixels using two and three derivative channels are different. This suggests that we could combine the two respective reconstructions to obtain an even more accurate edge representation.

5. DECOMPOSITION OF CHANNELS

We can combine the two primary uses of SOMP outlined earlier in order to derive a third interesting application, which involves breaking down image channels into different channels that are easier to process. For example, consider a colored sparse gradient image X which consists of the three channels red, green, and blue. Unfortunately, none of these channels are themselves sparse. They are only sparse in gradient. In other words, $\nabla X_k \cdot u$ is a sparse image for each $k = 1, 2, 3$ and any unit vector u. For the sake of simplicity, let us work with the vertical and horizontal directions only and denote the partial derivatives of any channel X_k as $D_v X_k$ and $D_h X_k$ respectively. We can decompose each color channel into these two derivative channels, giving us 6 channels all together onto which we can apply the SOMP algorithm. For each color channel k, this will return reconstructions of $D_v X_k$ and $D_h X_k$. Then we can use the following expression, which is derived in [3], in order to recover the original colored image X_k.



Original Edges

1 Derivative
ℓ_1 error = 1052

2 Derivatives
ℓ_1 error = 56

3 Derivatives
ℓ_1 error = 68

Fig. 5. The edges of the original Shepp-Logan phantom plus reconstructions using one, two, and three directional derivatives.

$$X_k \approx D_v^{-1}[D_v X_k] + D_h^{-1}[D_h X_k] - D_h^{-1} D_v^{-1} D_h [D_v X_k].$$

In this expression, D_v^{-1} is the vertical summation operator given by

$$D_v^{-1} X_k(n,m) = \sum_{j=1}^{n} X_k(j,m). \tag{5.1}$$

D_h^{-1} is similarly defined. All items shown in brackets were estimated directly using SOMP.

Figure 6 shows various reconstructions of a 128 x 128 colored Shepp-Logan phantom. After taking the Fourier Transform of the original image and randomly discarding 80% of the coefficients, we were able to use the technique outlined above to generate an exact reconstruction, which is the lower-left image. Compare this to the naive estimate (upper-right) generated by applying the Fourier back-projection operator $\mathcal{F}_\Omega^*$ onto our given measurements. In addition, we corrupted the original colored phantom with white Gaussian noise (SNR $\approx$ 40dB) and then used the same procedure to obtain the lower-right image.

Unfortunately, errors in the edge detection step lead to the vertical stripe artifacts seen in the last reconstruction. Fortunately, these undesired lines do not introduce blurring of the edges, and therefore, all major features of the original image, including the small ellipses, are still present. Furthermore,

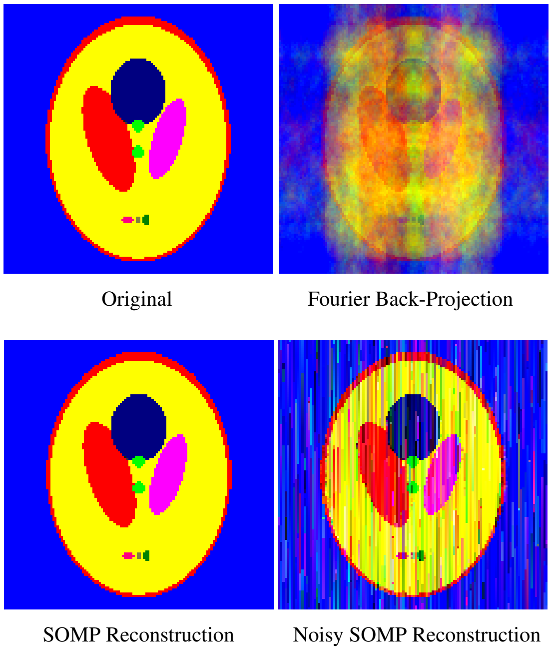


Fig. 6. The original colored phantom, its trivial back-projection reconstruction, as well as noiseless and noisy SOMP-based estimates.

post-processing of the image could easily remove the vertical artifacts.

6. CONCLUSION

In the noisy world that we live in, any image or signal processing algorithm must be robust to noise. Often times, we are interested in collecting and analyzing data that arrives in multiple streams or channels, each of which is corrupted with noise. Fortunately, there are algorithms such as SOMP that efficiently and accurately recover such signals from a small number of Fourier measurements when they are sparse in some domain. As either noise levels or the number of signal channels increase, SOMP's edge over ordinary OMP begins to grow, making it quite robust to noise if enough channels are present. SOMP also has other applications in image processing: it is a powerful edge detector. By utilizing two or more derivative channels, the likelihood of missing an edge pixel is greatly reduced. Furthermore, by decomposing ordinary image channels into derivative channels, it is possible for SOMP to detect the edges of multi-channel sparse gradient images. Then by using the simple integration and correction procedure introduced earlier, these images can subsequently be recovered from their edges. The results of this paper present several challenges that should be addressed in the future. First, how can we efficiently combine edge reconstructions generated by different numbers or sets of directional derivatives in order to create an even more accurate representation of an image's edges. Solving this problem will be beneficial as the vertical line artifacts seen in the reconstructions of sparse gradient images will be greatly reduced. Secondly, how can we speed up this algorithm? There are sublinear time Fourier algorithms (see [2]) that essentially can perform the same function as OMP with Fourier coefficients. It may be worthwhile to investigate whether such algorithms can perform well in a simultaneous sparse approximation setting. It is our ultimate goal to greatly improve the speed at which we can process and reconstruct large multi-channel data sets.

7. REFERENCES

[1] M. F. Duarte, S. Sarvotham, D. Baron, M. Wakin, R. Baraniuk, *Distributed Compressed Sensing of Jointly Sparse Signals*, Conf. Record of the 39th Asilomar Conference on Signals, Systems and Computers, 2005.

[2] A. C. Gilbert, S. Guha, P. Indyk, S. Muthukrishnan, M. J. Strauss, *Near-Optimal Sparse Fourier Representations via Sampling*, Proc. of the 2002 ACM Symposium on Theory of Computing STOC, pp. 152–161, 2002.

[3] R. Maleh, A. C. Gilbert, *Sparse Gradient Image Reconstruction Done Faster*, To Appear in Proc. ICIP, 2007.

[4] J. Tropp, A. C. Gilbert, *Signal Recovery from Partial Information via Orthogonal Matching Pursuit*, Submitted for publication, April 2005.

[5] J. Tropp, A. C. Gilbert, M. J. Strauss, *Simultaneous Sparse Approximation via Greedy Pursuit*, invited paper, special session on "Sparse representations in signal processing", in Proc. 2005 IEEE International Conference on Acoustics, Speech, and Signal Processing ICASSP, Philadelphia, PA, March 2005.