A Group Sparsity Imaging Algorithm for Transient Radio Sources

Stephan Wenger and Marcus Magnor
Institut für Computergraphik, TU Braunschweig
38106 Braunschweig, Germany
http://graphics.tu-bs.de/

Urvashi Rau
National Radio Astronomy Observatory
Socorro, NM 87801, USA
http://www.aoc.nrao.edu/~rurvashi/

***Abstract*—Radio interferometers can achieve high spatial resolution for temporally constant sources by combining data observed over long periods of time. Recent imaging algorithms reconstruct smoothly varying sources by representing temporal variation in polynomial or Fourier bases. We present a novel image reconstruction algorithm that is able to reconstruct continuously and erratically varying sources as well as long as they are confined to small regions of the image. This is achieved by enforcing spatial locality and sparsity of temporally varying sources through a group sparsity prior. Numerical experiments show that the proposed approach recovers image series to high accuracy where methods without temporal consistency fail and outperforms static reconstructions of dynamic scenes even in image regions with no temporal variation.**

I. Introduction

Radio interferometers sample an image of the sky in the Fourier domain with a changing pattern due to the rotation of the Earth. When the sky region being imaged is constant over the the time of observation, the different sampling patterns can be combined to produce a single high-quality, high-resolution image. Dynamic sources, however, cannot be imaged in this way using traditional reconstruction methods: if every time frame is reconstructed separately from much smaller amounts of data, the quality of each time frame suffers. If, however, a single image is reconstructed from all available data, the transient sources can cause artifacts even in static parts of the image, and all temporal resolution is lost.

We employ ideas from compressed sensing to simultaneously reconstruct all time slices of an observation of a dynamic source by minimizing the numbers of nonzero and temporally varying pixels. Static image regions benefit from the large amount of information collected during a long observation time, while dynamic image regions are reconstructed at high temporal resolution. Since all pixels are coupled in the data through the Fourier transform, a better reconstruction of the dynamic parts can improve the reconstruction of static regions and vice versa.

II. Radio interferometry imaging

In radio interferometry, a region of the sky is captured simultaneously by an array of antennae. The measurement can be described as a sampled Fourier transform $\mathcal{F}$. We reconstruct the sky image $\mathbf{x}$ from visibilities $\mathbf{v}$ using a non-negative variant of the *fast iterative shrinkage-thresholding algorithm* by Beck and Teboulle [1] that minimizes

$$\arg\min_{\mathbf{x}} \tfrac{1}{2} \left\|\mathcal{F}\mathbf{x} - \mathbf{v}\right\|_2^2 + \tau f(\mathbf{x}) \quad \text{s.t.} \quad \mathbf{x} \geq 0\,. \tag{1}$$

The algorithm can be summarized as follows: First, initialize $\mathbf{x}_{(0)} = 0$, $\mathbf{y}_{(1)} = 0$, and $t_{(1)} = 1$, and compute L, the largest eigenvalue of $\mathcal{F}^T\mathcal{F}$. Now perform a number of iterations, starting at $k = 1$:

$$\mathbf{x}_{(k)} = \mathbf{p}_f^+(\mathbf{y}_{(k)} - \mathcal{F}^T(\mathcal{F}\mathbf{x} - \mathbf{v})/L, 1/L) \tag{2}$$

$$t_{(k+1)} = (1 + \sqrt{1 + 4t_{(k)}^2})/2 \tag{3}$$

$$\mathbf{y}_{(k+1)} = \mathbf{x}_{(k)} + (t_{(k)} - 1)/t_{(k+1)}(\mathbf{x}_{(k)} - \mathbf{x}_{(k-1)}) \tag{4}$$

The algorithm is terminated when the change between two subsequent $\mathbf{x}_{(k)}$ drops below some user-defined threshold.

In Equation 2, $\mathbf{p}_f^+(\hat{\mathbf{x}}, \beta)$ is the non-negative proximal mapping for f, which is defined as

$$\mathbf{p}_f^+(\hat{\mathbf{x}}, \beta) = \arg\min_{\mathbf{x}} \tfrac{1}{2} \left\|\mathbf{x} - \hat{\mathbf{x}}\right\|_2^2 + \beta f(\mathbf{x}) \quad \text{s.t.} \quad \mathbf{x} \geq 0\,. \tag{5}$$

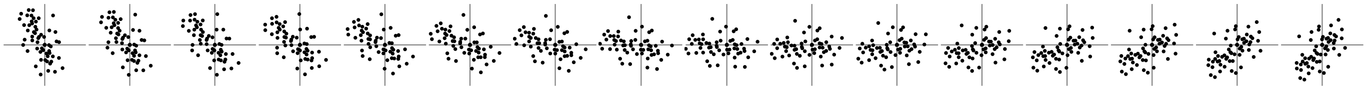
III. Sparse temporal variation

A significant part of our contribution is finding a regularizer f that appropriately describes the expected types of transient signals and is reasonably efficient to compute. We assume that most pixels exhibit no temporal variation, but those who do may change erratically. In addition, many radio images consist of small, isolated objects on a dark background. We can combine both assumptions into a *group sparsity* regularizer [3] that minimizes the number of pixels i containing nonzero intensities $\mathbf{x}_i$ at any time t. This kind of group sparsity can be achieved by minimizing the ℓ_1-norm of the ℓ_∞-norms of all groups, the so-called $\ell_{1,\infty}$-norm

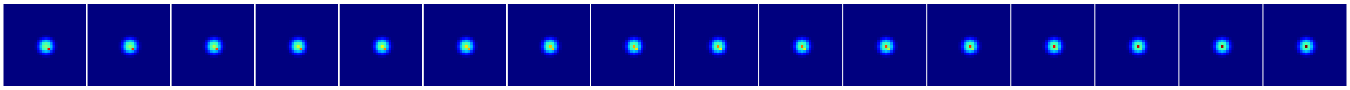
$$\left\|\mathbf{x}(t)\right\|_{1,\infty} = \sum_i \max_t \left|\mathbf{x}_i(t)\right| \,. \tag{6}$$

When only a single time slice is present, the $\ell_{1,\infty}$-norm obviously degrades to an ℓ_1-norm of $\mathbf{x}$, and the proposed approach becomes equivalent to previous approaches based on ℓ_1-norm minimization. For time-resolved data, however, we observe a number of interesting effects. First of all, minimizing

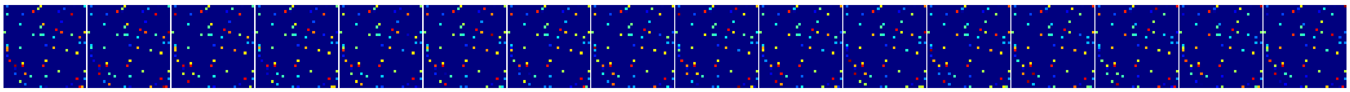
This is an abridged version of our article in *Astronomy and Computing*, volume 1, 2013, pages 40–45, http://dx.doi.org/10.1016/j.ascom.2013.02.002.



(a) UV coverage



(b) *blob*



(c) *point sources*

Fig. 1: Sampling pattern and test scenes.

the number of nonzero groups promotes pixel sparsity of each time slice. In addition, since only the maximum intensity over time is taken into account, erratic temporal behavior of intensity *amplitudes* can be reconstructed. At the same time, temporal consistency of intensity *locations* is achieved: if a pixel is bright in one time frame, the optimal choice for placing ambiguous intensity in another frame is the same pixel, and vice versa. In this way, information from multiple observations with different baseline patterns is effectively combined to resolve ambiguities and reduce sidelobes in *all* time frames. Similarly, a short flare of an otherwise faint but temporally varying source helps localize the faint source over the total duration of the observation, at the same time preventing the faint source from erroneously showing up as a side lobe in other frames. Finally, when $\ell_{1,\infty}$ minimization is applied to a time-resolved observation of a scene without any temporal variation, one can expect results very similar to those of a direct ℓ_1 reconstruction of a single image from all available data.

IV. Implementation

Our implementation follows the algorithm outlined in Equations 2–4, with the vectors $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{v}$ redefined as their concatenated values from each time frame. Likewise, the measurement operator $\mathcal{F}$ becomes a block-diagonal matrix consisting of the respective Fourier transforms in each time frame. The regularizer $f(\mathbf{x}) = \tau \|\mathbf{x}\|_{1,\infty}$ contains implicit information about which components of $\mathbf{x}$ and $\mathbf{y}$ belong to which time frame.

We observe that the proximal mapping $\mathbf{p}^+_{\tau\|\cdot\|_{1,\infty}}(\hat{\mathbf{x}}, \beta)$ can be computed independently for each group, that is, the $\ell_{1,\infty}$-norm is *group separable* [5]. This is because the groups, each consisting of all time frames for a single pixel, are disjoint, and the proximal mapping for the outer ℓ_1-norm is separable in its components as discussed above.

The (unconstrained) proximal mapping for the inner ℓ_∞-norm,

$$\mathbf{p}_{\tau\|\cdot\|_{1,\infty}}(\hat{\mathbf{x}}, \beta) = \arg\min_{\mathbf{x}} \tfrac{1}{2} \|\mathbf{x} - \hat{\mathbf{x}}\|_2^2 + \beta\tau \max_i \mathbf{x}_i \, , \tag{7}$$

can be computed by projecting $\hat{\mathbf{x}}$ orthogonally onto the set of all vectors with ℓ_1-norm less or equal to $\beta\tau$ and subtracting the result from $\hat{\mathbf{x}}$ [5]. Several algorithms exist for computing the orthogonal projection, including an $\mathcal{O}(n \log n)$ algorithm by Daubechies et al. [2] that our implementation uses. In order to ensure nonnegativity of the result, $\hat{\mathbf{x}}$ is thresholded against zero before applying the proximal mapping to remove any negative values.

For comparison, we implement several variants of a compressed sensing imaging algorithm [4]. The *static* method produces a single time frame from all available data under the assumption that the source is temporally constant. The ℓ_1 method reconstructs subsequent time frames individually from the data taken during the respective time frame. Finally, we implement a novel variant, the ℓ_2 method, that promotes *smooth temporal variation* in addition to ℓ_1-sparsity of each time frame. It enforces smoothness by extending the ℓ_1 method with a penalty term that is quadratic in the temporal derivative $\partial_t \mathbf{x}$, yielding the problem

$$\arg\min_{\mathbf{x}} \tfrac{1}{2} \|\mathcal{F}\mathbf{x} - \mathbf{v}\|_2^2 + \mu \|\partial_t \mathbf{x}\|_2^2 + \tau \|\mathbf{x}\|_1 \tag{8}$$

subject to $\mathbf{x}_i \geq 0$ for all i. The above statements on transitioning from the static to the time-dependent case apply. The temporal derivative is approximated by $(\partial_t \mathbf{x})_i(t) = \mathbf{x}_i(t) - \mathbf{x}_i(t-1)$, where $t > 0$ is an integer index. By substituting $(\mathcal{F}, \mu\partial_t)^T$ for $\mathcal{F}$ and $(\mathbf{v}, 0)^T$ for $\mathbf{v}$, we can reformulate the problem as

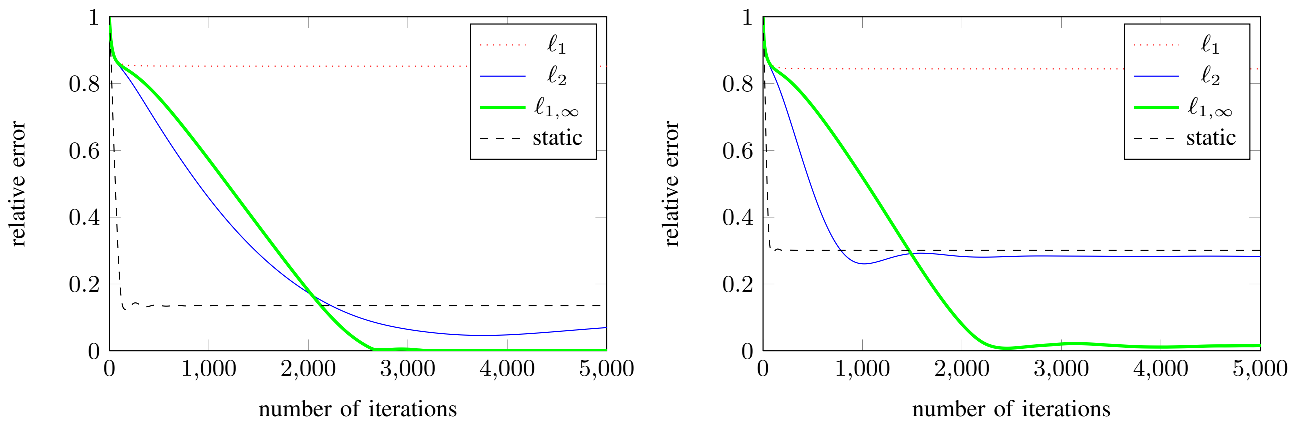
$$\arg\min_{\mathbf{x}} \tfrac{1}{2} \left\| \begin{pmatrix} \mathcal{F} \\ \mu\partial_t \end{pmatrix} \mathbf{x} - \begin{pmatrix} \mathbf{v} \\ 0 \end{pmatrix} \right\|_2^2 + \tau \|\mathbf{x}\|_1 \, , \tag{9}$$

which can be solved without major modifications to the original algorithm.

V. Results and discussion

We evaluate the reconstruction accuracy of our method on a number of simulated measurements, so that comparison to ground truth data as well as to other reconstruction methods is possible. In addition, we use a numerical experiment to investigate the circumstances under which the proposed method yields significant advantages over existing approaches.

16 subsequent 30-minute observations of different synthetic 32×32 pixel images on a hypothetical 12-antenna array were simulated, yielding 66 visibilities per time frame, Figure 1a.



(a) *blob*

(b) *point sources*

Fig. 2: Relative reconstruction errors for different test scenes.

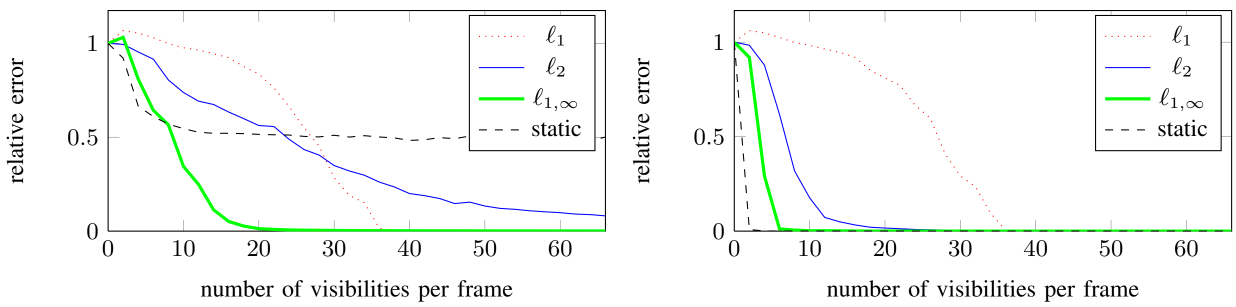
Images were reconstructed from these data using four different approaches. The *static* reconstruction is similar to a CLEAN reconstruction of a single image from all time frames, while the ℓ_1 reconstruction corresponds to the same method applied to each individual time frame. The ℓ_2 method recreates the effect of an algorithm that assumes smooth temporal variation. Finally, the proposed $\ell_{1,\infty}$ approach promotes sparsity of both the spatial intensity distribution and the set of transient pixels. All algorithms were run for 10000 iterations with $\tau = 0.1$ and (for ℓ_2) $\mu = 1$. In our experiments, results did not change significantly for different values of τ and μ within a few orders of magnitude.

For the first experiment, we simulate a small moving object on an extended background, Figure 1b. The background is modeled by a cosine-shaped blob, overlaid by a single bright pixel moving diagonally into the center. The static reconstruction recovers the main features of the source, but is unable to temporally resolve the movement. The ℓ_1 reconstruction, on the other hand, suffers from severe artifacts because each individual time frame does not contain enough information to reconstruct the whole image. The ℓ_2 method is well adapted to the continuous type of temporal variation and therefore reconstructs the scene well. However, it does not strongly penalize smooth temporal variation even in static parts of the image, leading to faint sidelobes in the background. Finally, the $\ell_{1,\infty}$ reconstruction is visually almost indistinguishable from ground truth.

Relative reconstruction errors for the different approaches are shown in Figure 2a. Because the static method only reconstructs a single time slice, it converges quickly. However, a significant error remains due to the temporally varying parts of the image. The independent reconstructions of the ℓ_1 method converge at similar speed, but leave an even higher error due to the data-starved setting. For the ℓ_2 approach, the sidelobes in the background and slight temporal fluctuations in static regions, even when not visually conspicuous, lead to noticeable residual error. Finally, the $\ell_{1,\infty}$ method converges more slowly at first, but finally reaches a relative error of the order of 10^{-4}.

In a second experiment, we investigate how the different approaches perform in the presence of static, smoothly varying and erratically varying sources. The synthetic source consists of a dark background with randomly placed point sources, 30 of which are static, 15 change linearly over time and 15 vary erratically, Figure 1c. While the static method recovers the locations and average intensities of both static and varying sources reasonably well in most cases, visibility data that is not explained by the static model leads to faint bogus sources in background regions. The ℓ_1 method correctly localizes many of the sources, but without exploiting temporal coherence, each individual frame does not contain enough information to be completely reconstructed. The ℓ_2 method reconstructs many static or smoothly varying sources rather well, but attenuates the temporal variation of erratically varying sources. Similar to the static method, this leads to bogus sources in the background. In addition, even many static sources fluctuate slightly over time because no penalty is used to enforce their being completely static. Finally, the $\ell_{1,\infty}$ method correctly recovers the location and behavior of static, smoothly varying and erratic sources with only minor errors in the absolute intensities.

Relative reconstruction errors for the different approaches are shown in Figure 2b. The static and ℓ_1 performance resembles the previous experiment, albeit the residual error for the static method is higher due to a larger amount of temporally varying sources. Like the static method, the ℓ_2 approach suffers from artifacts due to visibility data from erratic sources that cannot be explained by the model. It reaches a slightly lower reconstruction error, presumably because the smoothly varying sources are better reproduced. Finally, the $\ell_{1,\infty}$ method correctly localizes all sources, but the bias introduced by the regularization term leads to inaccuracies in absolute intensity of about one percent. This over-regularization can be counterbalanced with a subsequent *debiasing* step that keeps all zero pixels fixed and solves for the remaining intensities in a least-squares sense.



(a) dynamic scenes

(b) static scenes

Fig. 3: Median relative reconstruction errors for different scene types.

The performance of different reconstruction approaches can vary widely for different observation situations. While a comprehensive quantitative study of the influence of image content on reconstruction quality is out of the scope of this paper, we investigate how quality varies with the amount of available information per time frame. To do so, we use the same setup as before, but select only a random subset of all available visibilities in each frame. We then reconstruct several different, randomly generated source images from these shortened data using the four abovementioned methods. By varying the number of visibilities retained in the data, we can graph the relationship between the amount of available information and the reconstruction quality of different algorithms.

Figure 3a shows the median relative error computed from reconstructions of 20 source images, each containing 10 randomly placed, erratically varying point sources on a dark background. The ℓ_1 method is able to reconstruct each frame individually (without exploiting temporal coherence) from about 35 visibilities per frame, while the proposed $\ell_{1,\infty}$ approach achieves the same accuracy with as little as 20 visibilities. The ℓ_2 method fails to reach high accuracy because it inevitably smoothes out the erratic variation. Reconstructing a single image using the *static* method fails completely since the scene is dynamic and cannot be represented by a single image. In conclusion, the $\ell_{1,\infty}$ method always performed to the best competitor, independent of the number of visibilities used. However, since it involves a comparatively high computational load, the static or ℓ_1 approaches may be more convenient to use for very low and very high numbers of visibilities, respectively.

After demonstrating that the proposed $\ell_{1,\infty}$ approach is at least on par with the reference methods for erratic sources, we investigate how it performs when no temporal variation is present in the data. The results are shown in Figure 3b, where 20 randomly generated images were reconstructed, each containing 10 randomly placed static point sources. First, we observe that the ℓ_1 performance on a static scene is indistinguishable from that on a dynamic scene because the temporal coherence between frames is not exploited. The $\ell_{1,\infty}$ approach, on the other hand, benefits from the additional coherence; satisfactory reconstruction quality is reached with as little as 10 visibilities. The ℓ_2 method requires about 20 visibilities to achieve similar accuracy. This might be caused by the fact that pixel intensities are allowed to fluctuate over time because neither sparsity of the set of transient pixels nor of the temporal variation itself are enforced. Unsurprisingly, the *static* method excels at reconstructing a static scene. Unless the number of visibilities is very low, however, the performance of the proposed approach seems comparable to the static reconstruction method even on completely static images.

VI. Conclusion

We have presented a novel image reconstruction algorithm for transient radio sources based on group sparsity. Numerical experiments show that the proposed approach outperforms existing methods on data-starved observations of sources with a sparse pattern of smooth or erratic temporal variation. Outside this realm, it degrades gracefully: for data-starved observations of static scenes, its performance is comparable to sparse reconstruction of a single static image, while for data-rich observations of dynamic scenes, it performs comparably to sparse reconstruction of individual frames.

Acknowledgments

This work was supported in part by a Feodor Lynen alumni grant from the Alexander von Humboldt foundation, Germany.

References

[1] A. Beck and M. Teboulle, “A fast iterative shrinkage-thresholding algorithm for linear inverse problems,” *SIAM J. Imaging Sciences*, vol. 2, no. 1, pp. 183–202, 2009.

[2] I. Daubechies, M. Fornasier, and I. Loris, “Accelerated projected gradient method for linear inverse problems with sparsity constraints,” *J. Fourier Analysis and Applications*, vol. 14, no. 5, pp. 764–792, 2008.

[3] M. Fornasier and H. Rauhut, “Recovery algorithms for vector-valued data with joint sparsity constraints,” *SIAM J. Numerical Analysis*, vol. 46, no. 2, pp. 577–613, 2008.

[4] S. Wenger, M. Magnor, Y. Pihlström, S. Bhatnagar, and U. Rau, “SparseRI: A compressed sensing framework for aperture synthesis imaging in radio astronomy,” *Publications of the Astronomical Society of the Pacific*, vol. 122, no. 897, pp. 1367–1374, 2010.

[5] S. J. Wright, R. D. Nowak, and M. A. T. Figueiredo, “Sparse reconstruction by separable approximation,” *IEEE Transactions on Signal Processing*, vol. 57, no. 7, pp. 2479–2493, 2009.