

SPATIO-TEMPORAL SUPER-RESOLUTION RECONSTRUCTION OF REMOTE SENSING DATA

Igor Yanovsky^{1,2} and Jing Qin³

¹Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA

²Joint Institute for Regional Earth System Science and Engineering, University of California, Los Angeles, CA, USA

³Department of Mathematics, University of Kentucky, Lexington, KY, USA

ABSTRACT

We present a spatio-temporal super-resolution method for reconstructing a sequence of observations collected by imaging satellites. A sequence of observations is assumed to be defined on a low resolution spatio-temporal grid. It is further assumed that the sequence is generated by blurring of a captured scene with a spatio-temporal convolution kernel and is degraded by noise. Our method simultaneously exhibits deconvolution of the sequence of images from the effects of spatio-temporal blur, denoising of the data, and upsampling of the low-resolution sequence to a high resolution spatio-temporal grid. We perform the super-resolution in the space-time domain, as opposed to super-resolving the sequence separately and sequentially to a higher spatial and then temporal resolution grid. Simultaneous space-time optimization achieves a more efficient and more accurate reconstruction than reconstructing a sequence frame by frame. The proposed super-resolution methodology is based on total variation regularization and computes the solution using the alternating direction method of multipliers. Numerical results show our approach to be robust and computationally efficient.

Index Terms— Alternating direction method of multipliers, upsampling, satellite images, super-resolution

1. INTRODUCTION

The need for spatio-temporal super-resolution arises when analyzing sequences of observations captured by satellite imagery as well as when processing video sequences. In particular, many sensors have limited spatial and temporal resolution. This is especially the case for microwave sensors, which are characterized by spatial blurring and distortion. In addition, frame-rate and exposure time constrain instrument's temporal resolution.

This research was carried out in part at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration. The research of I. Yanovsky was also supported by the NSF grant DMS 2012868. The research of J. Qin was supported by the NSF grant DMS 1941197.

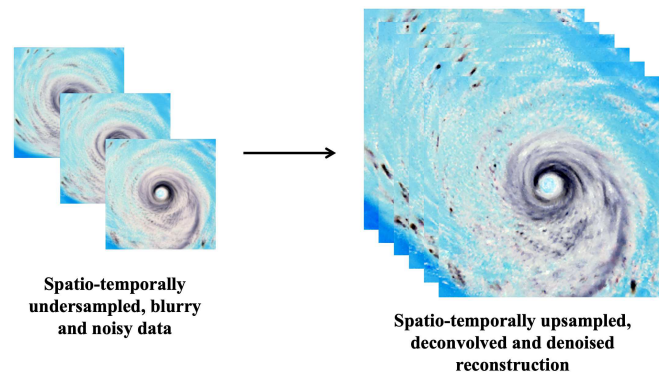


Fig. 1. Spatio-Temporal Super-Resolution: Spatio-temporally undersampled, blurry, and noisy sequence of p frames, each frame is $m \times n$ pixels, is deconvolved, denoised, and upsampled to contain $2p$ frames of size $2m \times 2n$.

There are a variety of applications of super-resolution. Spatial [1, 2, 3, 4, 5] and temporal [6, 7, 8] super-resolution methods have been proposed that perform reconstruction separately, either in space or in time. Other super-resolution approaches consider the problem of generating a high-resolution image given multiple low-resolution images [9, 10, 11, 12]. Rather than reconstructing a single high resolution image from a sequence of low-resolution images, our approach in this paper reconstructs an entire sequence of high-resolution images, by increasing spatial resolution (adding pixels) and temporal resolution (adding frames) (see Fig. 1). Moreover, in order to achieve a more accurate reconstruction than frame-by-frame processing, methods had been proposed to enhance resolution in the spatio-temporal domain [13, 14]. Our methodology improves such approaches by not only performing spatio-temporal deconvolution, but at the same time, performing upsampling.

The method in this paper builds upon our past efforts on spatial super-resolution and spatio-temporal deconvolution. In [15], we implemented the super-resolution technique for a single image using simultaneous upsampling and deconvolu-

tion spatially, but not temporally, using the alternating direction method of multipliers (ADMM) [16]. In [17], we implemented spatio-temporal deconvolution for a sequence of images, without performing upsampling, using the Split Bregman method [18]. In this paper, we implement the spatio-temporal super-resolution method, via spatio-temporal deconvolution and upsampling, using a total variation (TV) based ADMM technique to increase resolution of spatio-temporally distorted, noisy and downsampled image sequence.

2. SIMULTANEOUS DECONVOLUTION AND UPSAMPLING VIA ADMM

We assume that $u_0 \in \mathbb{R}^{m \times n \times p}$ is a ground truth image sequence, consisting of p frames, each frame is $m \times n$, corresponding to a physical scene. While being captured, the scene is convolved (or blurred) by a point spread function K and downsampled by an operator D . It is also degraded by additive noise κ . The forward imaging model can be written as

$$f = DKu_0 + \kappa, \quad (1)$$

where f is a blurry, noisy and subsampled sequence of observations.

Since the deconvolution and upsampling problem is very ill-posed, we apply regularization to guarantee the existence and uniqueness of solution in order to preserve the geometric qualities of the solution. We use the TV norm $\|u\|_{TV} = \|\nabla u\|_1 = \int |\nabla u|$ as a regularization term, which allows for the retrieval of images with sharp edges. Given a sequence of blurry, noisy, and subsampled observations f , we solve the following TV-based minimization problem to reconstruct image sequence u_0 :

$$\min_u \|\nabla u\|_1 + \frac{\mu}{2} \|DKu - f\|_2^2, \quad (2)$$

where $\mu > 0$ is a weighting parameter, and u is a video to be recovered. Since D is not circulant, we can not solve the least squares subproblem for u by directly applying ADMM. In order to minimize (2), we introduce additional variables. We let $\mathbf{v} = \nabla u$ in order to transfer ∇u out of nondifferentiable terms and let $w = Ku$ to split D and K . (Here, and in what follows, we denote vector quantities in bold font.) We write the minimization formulation as

$$\min_{u, \mathbf{v}, w} \|\mathbf{v}\|_1 + \frac{\mu}{2} \|Dw - f\|_2^2, \quad \text{such that } \mathbf{v} = \nabla u, w = Ku.$$

The corresponding augmented Lagrangian is

$$L(u, \mathbf{v}, w, \mathbf{x}, y) = \|\mathbf{v}\|_1 + \frac{\mu}{2} \|Dw - f\|_2^2 + \frac{\rho_1}{2} \|\nabla u - \mathbf{v} + \mathbf{x}\|_2^2 + \frac{\rho_2}{2} \|w - Ku + y\|_2^2,$$

where $\mathbf{x}$ and y are dual variables, scaled by parameters ρ_1 and ρ_2 , respectively. The dual variables are updated as follows:

$$\mathbf{x} \leftarrow \mathbf{x} + \gamma(\nabla u - \mathbf{v}) \text{ and } y \leftarrow y + \gamma(w - Ku).$$

For fixed u and w , the minimization problem for $\mathbf{v}$ is

$$\mathbf{v}^* = \arg \min_{\mathbf{v}} \left\{ \|\mathbf{v}\|_1 + \frac{\rho_1}{2} \|\nabla u - \mathbf{v} + \mathbf{x}\|_2^2 \right\},$$

which can be solved in the closed form using the shrinkage operator.

For fixed u and $\mathbf{v}$, the minimization problem for w is

$$w^* = \arg \min_w \frac{\mu}{2} \|Dw - f\|_2^2 + \frac{\rho_2}{2} \|w - Ku + y\|_2^2,$$

leading to normal equations which can be solved in the closed form.

For fixed $\mathbf{v}$ and w , the minimization problem for u is

$$u^* = \arg \min_u \frac{\rho_1}{2} \|\nabla u - \mathbf{v} + \mathbf{x}\|_2^2 + \frac{\rho_2}{2} \|w - Ku + y\|_2^2,$$

which can be solved by the 3-D discrete Fourier transform.

3. RESULTS

In our investigations, we used simulated microwave 157 GHz, 166 GHz, and 176 GHz channel image sequences (see Fig. 2(a)) to test the spatio-temporal super-resolution method with deconvolution and upsampling. The sequences capture hurricane Rita in the Gulf of Mexico in 2005. The Weather Research and Forecast model [19] was used to generate simulations. The spatio-temporal image sequence consists of 36 frames at 10-minute intervals. Each frame is of size 402×402 pixels, with the resolution of a pixel at 1.3 km. The image intensities in figures represent brightness temperatures in units of Kelvin (K). Note that colors in Figs. 1 and 2 correspond to the three microwave frequencies, i.e. 157 GHz (R), 166 GHz (G), and 176 GHz (B), the “false color” rendering.

In order to blur, add noise to, and downsample the original sequence in Fig. 2(a), we performed the following operations: (1) We used the spatio-temporal convolution kernel, which amounted to blurring the sequence spatially with the Gaussian point spread function of standard deviation 1, and temporally with the temporal rectangular kernel of 5 frames wide. (2) The spatio-temporal image sequence was also degraded with additive Gaussian noise of standard deviation $\sigma = 1$ K. (3) The image sequence was downsampled to 18 frames of 201×201 pixels.

The operations described above produce degraded image sequences in Fig. 2(b) that significantly differ from the ground truth image sequences in Fig. 2(a). We note that our algorithm can reconstruct a spatio-temporal image sequence with an arbitrary integer upsampling factor, but these results only explore an upsampling factor of 2.

We use the efficient ADMM for spatio-temporal super-resolution to obtain the result in Fig. 2(c). The results were processed channel-by-channel for the three frequencies. Our method has recovered an image sequence which visually

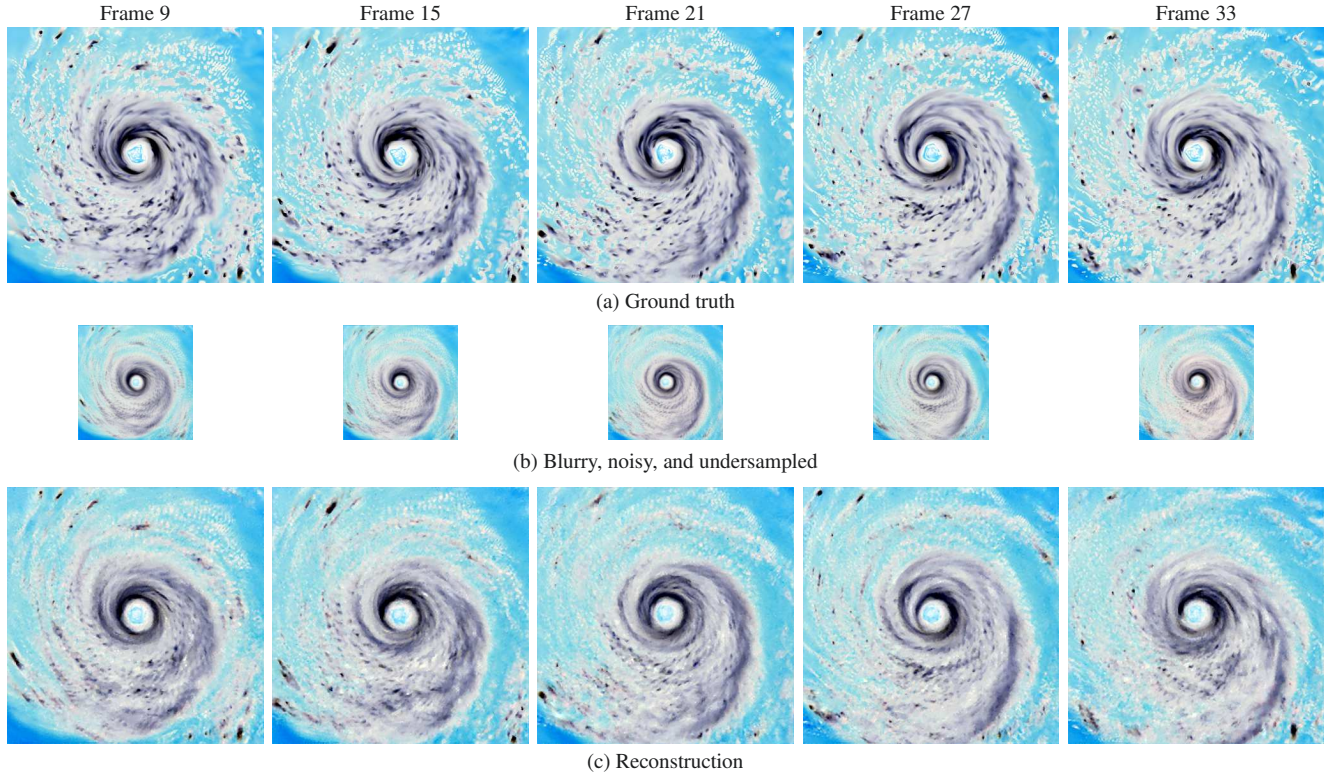


Fig. 2. Spatio-temporal super-resolution of a three-channel microwave image sequence of the simulated hurricane Rita. (a) Five of 36 original simulated ground truth 402×402 pixel frames are shown. (b) Ground truth sequence from (a) is spatially convolved with the Gaussian kernel of standard deviation 1 and temporally averaged with a rectangular function of five frames wide. The sequence is further subsampled to 18 frames of 201×201 pixels and degraded with additive Gaussian noise of standard deviation $\sigma = 1$ K. (c) Reconstruction result.

seems to be similar to the ground truth image sequence. To quantify the performance, we use root-mean-square error (RMSE), signal-to-noise ratio (SNR), and peak signal-to-noise ratio (PSNR) defined as

$$\text{RMSE} = \sqrt{\|u_0 - u\|^2 / N},$$

$$\text{SNR} = 10 \log_{10} (\sigma_{u_0}^2 \cdot N / \|u_0 - u\|^2),$$

$$\text{PSNR} = 10 \log_{10} (\max(u)^2 \cdot N / \|u_0 - u\|^2),$$

where $N = m \cdot n \cdot p$ is the number of pixels in the image sequence, u_0 is the ground truth image sequence, $\sigma_{u_0}^2$ is the variance of u_0 , and u is the image sequence we quantify the signal of.

Table 1 gives RMSE, SNR, and PSNR values for (i) the subsampled, blurry, and noisy image sequence in Fig. 2(b) and (ii) reconstructed sequence in Fig. 2(c) both relative to the ground truth image sequence in Fig. 2(a). The comparison between the subsampled, blurry, and noisy sequence and the ground truth sequence is performed by upsampling the former sequence to a higher resolution 402×402 grid using cubic interpolation spatio-temporally. The values demonstrate that the result in Fig. 2(c) reduces errors and increases the signal in image sequences compared to Fig. 2(b) quantitatively.

4. CONCLUSIONS

We proposed a spatio-temporal super-resolution method for reconstructing a sequence of observations that is assumed to be defined on a low resolution spatio-temporal grid, generated by blurring of a captured scene with a spatio-temporal convolution kernel, and degraded by noise. The approach is based on ADMM and is shown to be robust to noise. Our experiments show that we can reconstruct more accurate images both visually and based on common quality metrics.

5. ACKNOWLEDGMENT

The authors would like to thank Bjorn Lambrigtsen and Luminita Vese for their helpful discussions.

6. REFERENCES

- [1] M. A. T. Figueiredo and R. D. Nowak, "A bound optimization approach to wavelet-based image deconvolu-

Table 1. RMSE, SNR, and PSNR values in subsampled blurry and noisy as well as reconstructed image sequences shown in Fig. 2 (b) and (c). The values for the three channels are shown.

	RMSE	SNR	PSNR
Low res. 157 GHz	17.95	4.41	24.30
High res. reconstruction	14.53	6.25	26.14
Low res. 166 GHz	12.44	4.97	27.49
High res. reconstruction	9.44	7.37	29.88
Low res. 176 GHz	9.76	5.52	29.46
High res. reconstruction	7.10	8.28	32.22

tion,” in *IEEE International Conference on Image Processing 2005*, vol. 2, 2005, pp. II–782.

- [2] J. H. Money and S. H. Kang, “Total variation minimizing blind deconvolution with shock filter reference,” *Image and Vision Computing*, vol. 26, no. 2, pp. 302–314, 2008.
- [3] S. D. Babacan, R. Molina, and A. K. Katsaggelos, “Variational bayesian blind deconvolution using a total variation prior,” *IEEE Transactions on Image Processing*, vol. 18, no. 1, pp. 12–26, 2009.
- [4] I. Yanovsky, B. Lambrigtsen, A. Tanner, and L. Vese, “Efficient deconvolution and super-resolution methods in microwave imagery,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 8, no. 9, pp. 4273–4283, 2015.
- [5] I. Yanovsky and B. Lambrigtsen, “Multispectral super-resolution of tropical cyclone imagery using sparsity-based approaches,” *International Journal of Remote Sensing*, vol. 37, no. 11, pp. 2494–2509, 2016.
- [6] S. K. Nayar and M. Ben-Ezra, “Motion-based motion deblurring,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 6, pp. 689–698, 2004.
- [7] R. Raskar, A. Agrawal, and J. Tumblin, “Coded exposure photography: motion deblurring using fluttered shutter,” *ACM Transactions on Graphics*, vol. 25, no. 3, pp. 795–804, 2006.
- [8] I. Yanovsky and B. Lambrigtsen, “Temporal super-resolution of microwave remote sensing images,” in *2018 IEEE 15th Specialist Meeting on Microwave Radiometry and Remote Sensing of the Environment (MicroRad)*, 2018, pp. 110–115.
- [9] M. Batz, J. Koloda, A. Eichenseer, and A. Kaup, “Multi-image super-resolution using a locally adaptive denoising-based refinement,” in *2016 IEEE 18th International Workshop on Multimedia Signal Processing (MMSP)*, 2016, pp. 1–6.
- [10] R. Nayak, S. Harshavardhan, and D. Patra, “Morphology based iterative back-projection for super-resolution reconstruction of image,” in *2014 2nd International Conference on Emerging Technology Trends in Electronics, Communication and Networking*, 2014, pp. 1–6.
- [11] H. Zhu, X. Tang, J. Xie, W. Song, F. Mo, and X. Gao, “Spatio-temporal super-resolution reconstruction of remote-sensing images based on adaptive multi-scale detail enhancement,” *Sensors*, vol. 18, no. 2, p. 498, 2018.
- [12] J. Qin and I. Yanovsky, “Robust super-resolution image reconstruction method for geometrically deformed remote sensing images,” in *IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium*, 2018, pp. 8050–8053.
- [13] E. Shechtman, Y. Caspi, and M. Irani, “Space-time super-resolution,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 4, pp. 531–545, 2005.
- [14] S. H. Chan, R. Khoshabeh, K. B. Gibson, P. E. Gill, and T. Q. Nguyen, “An augmented lagrangian method for total variation video restoration,” *IEEE Transactions on Image Processing*, vol. 20, no. 11, pp. 3097–3111, 2011.
- [15] J. Qin, I. Yanovsky, and W. Yin, “Efficient simultaneous image deconvolution and upsampling algorithm for low resolution microwave sounder data,” *Journal of Applied Remote Sensing*, vol. 9, no. 1, pp. 095 035/1–15, 2015.
- [16] R. Glowinski and A. Marrocco, “Sur l’approximation, par elements finis d’ordre un, et la resolution, par, penalisation-dualité, d’une classe de problemes de Dirichlet non lineares,” *Revue Française d’Automatique, Informatique, et Recherche Operationelle*, vol. 9, pp. 41–76, 1975.
- [17] I. Yanovsky, J. Qin, and B. Lambrigtsen, “Spatio-temporal resolution enhancement for geostationary microwave data,” in *2020 IEEE 16th Specialist Meeting on Microwave Radiometry and Remote Sensing of the Environment (MicroRad)*, 2020.
- [18] T. Goldstein and S. Osher, “The split Bregman method for L1-regularized problems,” *SIAM Journal on Imaging Sciences*, vol. 2, no. 2, pp. 323–343, 2009.
- [19] J. Michalakes, J. Dudhia, D. Gill, J. Klemp, and W. Skamarock, *Design of a next-generation regional weather research and forecast model: Towards Teracomputing*. River Edge, New Jersey: World Scientific, 1998.