

A Compressive Sensing Approach to Image Restoration

M. A. Kitchener, A. Bouzerdoum, IEEE Senior Member, and S. L. Phung, IEEE Member
School of Electrical, Computer and Telecommunications Engineering
University of Wollongong
Wollongong, NSW 2522, Australia
E-mails: mak55@uow.edu.au, a.bouzerdoum@ieee.org and phung@uow.edu.au

Abstract—In this paper the image restoration problem is solved using a Compressive Sensing approach, and the translation invariant, *a Trous*, undecimated wavelet transform. The problem is cast as an unconstrained optimization problem which is solved using the Fletcher-Reeves nonlinear conjugate gradient method. A comparison based on experimental results shows that the proposed method achieves comparable if not better performance as other state-of-the-art techniques.

I. INTRODUCTION

Image restoration techniques aim to remove the unavoidable distortions and noise that enter an image during the image capture process. These degradations are generally modeled in the following way,

$$\mathbf{g} = H\mathbf{f}_0 + \mathbf{v} \quad (1)$$

where \mathbf{f}_0 and \mathbf{g} are lexicographically ordered vectors representing, respectively, the ideal image and noisy distorted image, H is a matrix representing the blur and \mathbf{v} is additive noise. These distortions are generally dealt with by mathematically reversing the process that caused them. This equates to finding the inverse matrix of H , if in fact it exists. Even if H is invertible, the matrix is generally ill-conditioned and the noise present within the image can render the inverse solution useless. Because of this, regularization techniques are often necessary to make the problem well posed. One of the most popular regularization techniques is Tikhonov Regularization, in which \mathbf{f} is found by minimizing $\|\mathbf{g} - H\mathbf{f}\|_2^2 + \lambda\|R\mathbf{f}\|_2^2$, with λ being the regularization parameter chosen to control the trade off between the preservation of image fidelity and smoothing of the image to remove noise.

In this paper we propose an image restoration method based on Compressive Sensing (CS) theory. Research into CS has exploded in recent years with applications being developed in areas such as channel coding, data compression, inverse problems and data acquisition [4]. The benefits of CS techniques stem from the fact that many natural signals exhibit sparsity, or rather their non zero coefficients can be described as being sparse or compressed within a signal. This means that it is often possible to reconstruct images with a high degree of accuracy, despite using significantly fewer measurements than traditionally thought necessary,

provided that the signal exhibits both sparsity and incoherence.

If a signal is sparse, coefficients which are equal to, or close to, zero can in a sense be considered redundant information. A common approach to maximizing sparsity in the solution is to solve the following unconstrained minimization problem [6],

$$\text{minimize}_s \quad \|\mathbf{s}\|_1 + \lambda\|\mathbf{y} - A_0\mathbf{s}\|_2^2 \quad (2)$$

where s is the lexicographically ordered vector representing the solution to our problem, \mathbf{y} represents the observed subset, A_0 represents the “sensing matrix” or “measurement matrix” and λ represents a regularization parameter used to control the tradeoff between maintaining the fidelity of the solution and enforcing sparsity. A key part of this theory is based on the Restricted Isometry Principle (RIP)[4][2].

CS also relies on the concept of incoherence, whereby the problem must satisfy the dual requirements of being sparse in the domain it is represented and spread out in the domain in which the measurements are acquired. An excellent introduction to both these concepts are provided in [4] and [2].

In this paper, the ill-posed image restoration problem is formulated using a CS approach. This problem is then solved using the nonlinear conjugate gradient method. While the proposed method could be used for the reconstruction of incomplete noisy blurred measurements as presented in [6], the primary motivation here is to use a CS approach for the restoration of formed images.

II. PROBLEM FORMULATION

The image restoration problem is underdetermined, and thus does not have a unique solution. Regularization techniques resolve this issue by using *a priori* information about the ideal image to impose constraints on the solution. Tikhonov regularization for example imposes smoothness on the solution. The objective of CS techniques is to find the solution which exhibits the greatest sparsity. Many real world signals are sparse, or have a sparse representation, so choosing the sparse solution can often provide the best approximation to the ideal solution. It follows that, provided

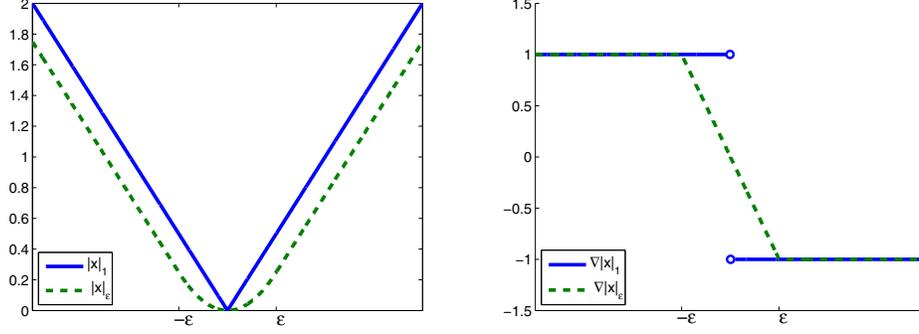


Figure 1. The Huber norm and its derivative.

the ideal image is sparse or has a sparse representation, CS techniques can be used for the image restoration problem.

While there are classes of images that are sparse, the spatial domain representation of images is generally not sparse, and as such they cannot be restored using CS directly. If however we represent an image in a domain other than the spatial domain, it is possible to find a sparse representation. For this reason, in the proposed approach we define

$$\mathbf{s} = R\mathbf{f} \quad (3)$$

where R is some transform which provides a sparse representation of the image. The problem can now be expressed as

$$\text{minimize}_{\mathbf{f}} \quad \|\mathbf{R}\mathbf{f}\|_1 + \lambda\|A_1R(\mathbf{g} - H\mathbf{f})\|_2^2 \quad (4)$$

which is in the same form as Eq. (2) where $\mathbf{y} = A_1R\mathbf{g}$ are the noisy blurred observations and \mathbf{f} is the spatial domain representation of the solution. The blurring matrix H , and matrix A_1 are in effect, factors of the measurement matrix A_0 . It can be shown that in this case $A_0 = A_1RHR^{-1}$. If R and H are commutative, see Section IV, then $A_0 = A_1H$.

III. A CONJUGATE GRADIENT APPROACH

In this paper, the unconstrained minimization problem in Eq. (2) is solved using the Fletcher-Reeves conjugate gradient method, with $0 \leq \lambda \leq \infty$. For low levels of noise the optimal regularization parameter will be very large, which can effect stability. If however we let $\lambda = \frac{\mu}{1-\mu}$ and multiply (2) throughout by $(1-\mu)$, which is a constant, the objective function becomes

$$\text{minimize} \quad J(\mathbf{s}) = (1-\mu)\|\mathbf{s}\|_1 + \mu\|\mathbf{y} - A_0\mathbf{s}\|_2^2 \quad (5)$$

where the new regularization parameter μ must satisfy $0 \leq \mu \leq 1$. With this modification the conjugate gradient iterations become,

- Step 1: Calculate the steepest direction \mathbf{p}
 $\mathbf{p}_{k+1} = \nabla_s [(1-\mu)\|\mathbf{s}\|_1 + \mu\|\mathbf{y} - A_0\mathbf{s}\|_2^2]$
 Step 2: Use Fletcher-Reeves approach to compute β
 $\beta = (\mathbf{p}_{k+1}^T \mathbf{p}_{k+1}) / (\mathbf{p}_k^T \mathbf{p}_k)$

Step 3: Update direction, \mathbf{d}

$$\mathbf{d}_{k+1} = -\mathbf{p}_{k+1} + \beta\mathbf{d}_k$$

Step 4: Find α which minimizes $J(\mathbf{s}_k + \alpha\mathbf{d}_{k+1})$ using a line search method

Step 5: Update position

$$\mathbf{s}_{k+1} = \mathbf{s}_k + \alpha\mathbf{d}_{k+1}$$

For an inexact line search, it is possible that $\beta_k \mathbf{p}_k^T \mathbf{d}_{k-1} > 0$. In this case a restart, $\mathbf{d}_k = -\mathbf{p}_k$, can be used to aid convergence; for further details on this the reader is referred to [17]. Evaluating the steepest decent direction, \mathbf{p}_k , can be problematic as the L_1 norm is not continuously differentiable. Instead we approximate the L_1 norm using the Huber norm, defined by

$$|\mathbf{x}|_\epsilon = \sum_{i=1}^n \phi(x_i) \text{ where } \phi(x) = \begin{cases} |x| - \frac{\epsilon}{2} & \text{if } |x| \geq \epsilon \\ \frac{x^2}{2\epsilon} & \text{if } |x| < \epsilon \end{cases} \quad (6)$$

As ϵ approaches zero, the Huber norm approaches the L_1 norm, however unlike the L_1 norm the Huber norm is differentiable, with its derivative being

$$\nabla|\mathbf{x}|_\epsilon = \sum_{i=1}^n \varphi(x_i) \text{ where } \varphi(x) = \begin{cases} \text{sign}(x) & \text{if } |x| \geq \epsilon \\ \frac{x}{\epsilon} & \text{if } |x| < \epsilon \end{cases} \quad (7)$$

As an additional benefit, unlike the L_1 norm which is contrast invariant, the Huber norm smooths small gradients while still allowing sharp discontinuities at large gradients. As noted in [18], for this reason the Huber norm has been used in denoising applications to overcome staircasing effects.

IV. DOMAIN SELECTION

The success of the proposed approach relies on selecting R , so that $R\mathbf{f} = \mathbf{s}$ is sparse. While there are many transforms which can achieve this, for this paper we have chosen the *a Trous*, Undecimated Discrete Wavelet Transform (UDWT). Wavelet transforms are very popular in image processing and compression, because of their ability to represent both spatial and temporal information in a sparse manner [7][13]. In fact, it is this ability to promote sparsity which has made wavelet-based image restoration methods

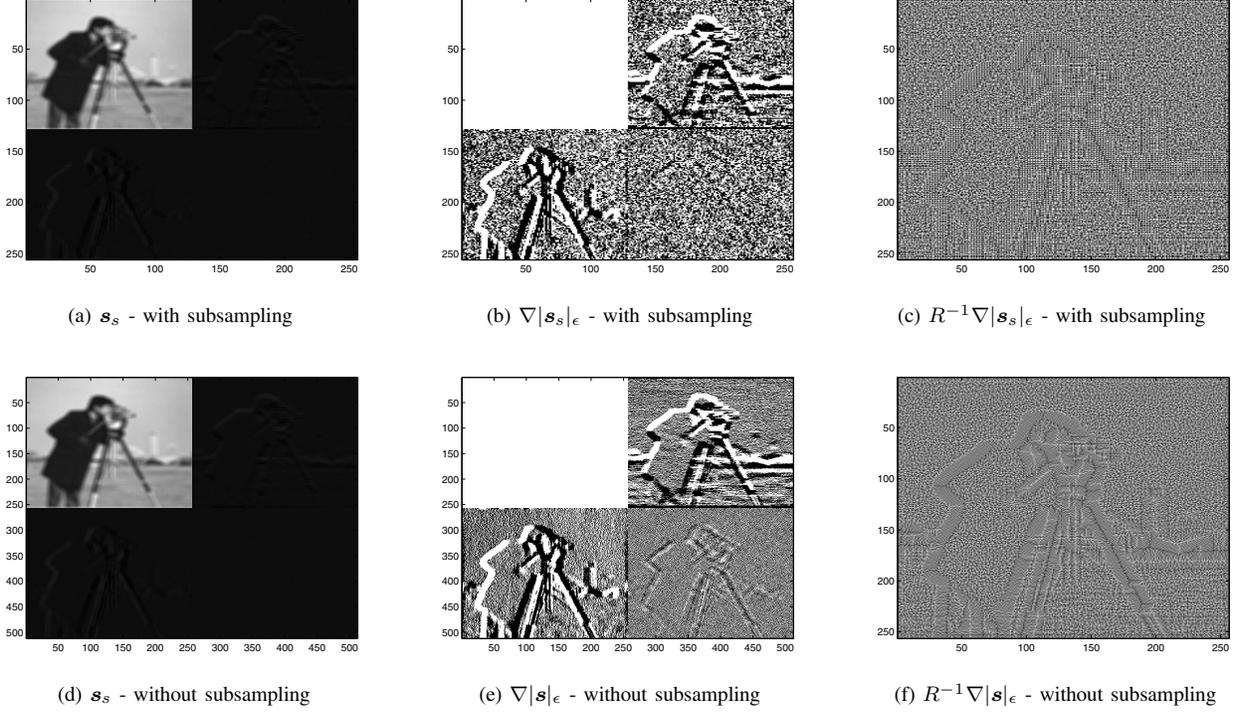


Figure 2. The effect of subsampling

successful. Undecimated Discrete Wavelet Transforms are so named because, unlike the traditional Discrete Wavelet Transform (DWT), they do not perform down sampling of the output signal. While the vast majority of literature is dedicated to the traditional DWT, the redundant nature of the UDWT has been shown to give improved performance in denoising applications [15][16][12][11][5], in addition to providing a clear relationship between the spatial domain and wavelet domain layers. Furthermore the UDWT, unlike the DWT, has convolutional commutativity, so that

$$RHf = HRf = Hs \quad (8)$$

where H and R are the blurring matrix and transform matrix respectively. In other wavelet based image restoration methods which use the DWT [10][14][9][8], this loss of commutativity is dealt with by performing the inverse DWT before reblurring, so that

$$RHf = RHR^{-1}s \quad (9)$$

While this provides a solution, which can be concisely represented in matrix form, using the commutative UDWT is more efficient.

Another significant advantage with using the UDWT is found when minimizing $|s|_\epsilon$. Consider for example the results in Figure 2. Figures 2(a) and 2(d) show the wavelet domain representations of g . While the UDWT and DWT

can both be used for perfect reconstruction of an image using $R^{-1}(Rf)$, when calculating the search direction $R^{-1}\nabla|Rf|_\epsilon$, as in (7), the results are very different. The redundant nature of the UDWT ensures that more information is retained (See Figures 2(c) and 2(f)), which aids in the restoration.

The only problem with using the UDWT is that the size of s increases with every application of the wavelet transform. This problem is elegantly solved by the proposed technique however by using CS, and selecting A_1 such that there are fewer rows than columns ($m \ll n$). In this way the size of $y - A_0s$ can be controlled without using subsampling. Furthermore, if A_1 is designed to take advantage of the large number of coefficients close to, or equal to, zero, this will correspond to an increase in performance and efficiency.

V. RESULTS

The performance of the proposed approach was evaluated by performing a range of experiments using different images and levels of noise. In each of these experiments the Improvement in Signal to Noise Ratio (ISNR) was used as a measure of performance, defined by:

$$\text{ISNR} = 10 \log_{10} \frac{\|g - f_o\|_2^2}{\|f - f_o\|_2^2} \text{ dB} \quad (10)$$

In line with the experiments performed in [7],[10],[9], a simple Daubechies wavelet was used for the following experiments, specifically the Daubechies 1 wavelet.

The performance of the proposed approach was evaluated by repeating the experiments performed in [1]. This involved restoring the Cameraman, Lena and Shepp-Logan Phantom images, all of which had been degraded by a 9×9 uniform blur with three different levels of additive zero mean gaussian noise corresponding to a BSNR of $40dB$, $30dB$ and $20dB$. Here $BSNR = 10 \log_{10}(\frac{\sigma_{Hf}^2}{\sigma_n^2}) dB$, where σ_{Hf}^2 is the variance of Hf and σ_n^2 is the variance of the noise, v . The regularization parameter λ was chosen in an ad-hoc manner.

A. Experiment 1

In the first set of experiments, the proposed approach was tested with the Measurement Matrix A_1 set to the identity matrix. A summary of these results can be found in Table I, labeled “CS-WT(1)”. The results show that the proposed method performs as well as state-of-the-art methods with particularly good results obtained for the smoother Phantom and Cameraman images. This success can be attributed to the fact that smooth images naturally have sparse wavelet domain representations.

B. Experiment 2

In the second set of experiments, the wavelet decomposition was applied an additional 2 times to the “Low Low” region, with A_1 set to the identity matrix. These results are included in Table I under the label “CS-WT(3)”. These results show an increase in performance, at the expense of computation time, with more texture information being retained as the number of UDWT layers is increased.

C. Experiment 3

Finally, the experiments were repeated with the introduction of a $m \times n$ measurement matrix, A_1 , in which there are fewer rows than columns ($m \ll n$). Since the “Low Low” region of s is not sparse, here we only apply the measurement matrix to the other regions, \bar{s} . As done in [3], A_1 was chosen so that the measurements were taken from the fourier domain of s , with the samples corresponding to the $m = 0.25n$ largest fourier coefficients of y being chosen for the reconstruction. By using the Fast Fourier Transform (FFT), it was not necessary to implement the measurement matrix explicitly, which would have been impractical for such a large scale problem. The results obtained are in Table I, labeled “CS-WT- A_1 ”. The introduction of A_1 reduced the size of $y - A_0s$, with minimal effect on performance. In some cases the introduction of a compressive A_1 actually produced better results, however no conclusions can be drawn about this without an accurate means of calculating the optimal λ in each situation. Figure 3 shows the evolution, in terms of ISNR, of the restored Phantom image, for the three levels of noise. For this experiment the iterations were not terminated until the 500th iteration to demonstrate both, convergence and stability. The results for BSNR of 20 and 30 clearly show convergence at the optimal value of ISNR.

Table I
SUMMARY OF RESULTS WITH COMPARISONS FROM [1]

BSNR	Method	Lena		Cameraman		Shepp-Logan	
		ISNR	σ^2	ISNR	σ^2	ISNR	σ^2
40dB	CS-WT (1)	7.35	0.17	8.62	0.308	18.75	0.48
	CS-WT (3)	7.5	0.17	8.86	0.308	18.5	0.48
	CS-WT- A_1 (1)	7.33	0.17	8.49	0.308	18.90	0.48
	MOL	6.79	0.17	6.16	0.31	5.82	0.48
	BF01	8.34	-	8.55	-	14.22	-
	BF02	8.35	-	8.25	-	12.01	-
	ALG1	8.42	0.12	8.57	0.25	13.69	0.40
	ALG2	8.37	0.12	8.46	0.27	13.26	0.44
30dB	CS-WT (1)	5.11	1.66	5.80	3.08	11.87	4.80
	CS-WT (3)	5.22	1.66	6.0	3.08	11.40	4.80
	CS-WT- A_1 (1)	5.13	1.66	5.69	3.08	11.12	4.80
	MOL	4.62	1.71	3.98	3.22	4.10	4.69
	BF01	6.08	-	5.68	-	8.88	-
	BF02	5.64	-	4.65	-	6.91	-
	ALG1	5.89	1.49	5.41	3.04	7.77	4.60
	ALG2	5.58	1.64	4.38	3.60	6.50	4.84
20dB	CS-WT (1)	3.43	16.6	3.64	30.8	6.7	47.96
	CS-WT (3)	3.53	16.6	3.76	30.8	6.7	47.96
	CS-WT- A_1 (1)	3.55	16.6	3.65	30.8	6.6	47.96
	MOL	2.94	17.07	2.26	32.42	2.66	45.68
	BF01	4.09	-	3.31	-	5.57	-
	BF02	4.14	-	2.12	-	2.95	-
	ALG1	3.72	17.00	2.42	34.05	3.01	49.97
	ALG2	3.15	17.93	1.94	36.52	2.64	53.67

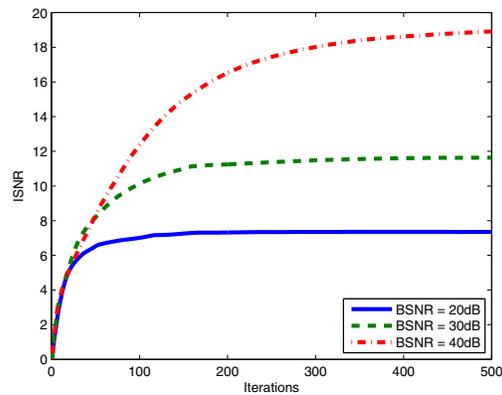


Figure 3. The evolution of ISNR over 500 iterations for the CS-WT- A_1 method, applied to Phantom image with 3 different levels of noise

VI. CONCLUSION

In this paper an image restoration method was formulated using the translation invariant, a Trous, undecimated wavelet transform. It was shown that a Compressive Sensing approach to the image restoration problem can be used as an alternative to traditional subsampled DWT image restoration approaches. A comparison based on experimental results has shown that the proposed method performs competitively with other state-of-the-art image restoration methods, particularly for smoother images with naturally sparse wavelet domain representations.

VII. ACKNOWLEDGEMENT

This work has been supported, in part, by a grant from the Australian Research Council.



(a) Degraded Cameraman



(b) Restored Cameraman



(c) Degraded Phantom



(d) Restored Phantom

Figure 4. Restoration Results - CS-WT- A_1 - BSNR 30dB

REFERENCES

- [1] S. D. Babacan, R. Molina, and A. K. Katsaggelos. "parameter estimation in tv image restoration using variational distribution approximation". *IEEE Trans. on Image Processing*, 17(3):326–339, 2008.
- [2] R. G. Baraniuk. "compressive sensing". *IEEE Signal Processing Magazine*, 24(4):118–121, 2007.
- [3] E. J. Candes, J. Romberg, and T. Tao. "robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information". *IEEE Trans. on Information Theory*, 52(2):489–509, 2006.
- [4] E. J. Candes and M. B. Wakin. "an introduction to compressive sampling". *IEEE Signal Processing Magazine*, 25(2):21–30, 2008.
- [5] F. Chaillan and P. Courmontagne. "coupling the stochastic matched filter and the trous algorithm for sas image denoising". In *OCEANS 2006 - Asia Pacific*, pages 1–8, 2006.
- [6] M. J. Dimet and F. Le. "deblurring from highly incomplete measurements for remote sensing". *IEEE Trans. on Geoscience and Remote Sensing*, 47(3):792–802, 2009.
- [7] M. A. T. Figueiredo, J. M. Bioucas-Dias, and R. D. Nowak. "majorization-minimization algorithms for wavelet-based image restoration". *IEEE Trans. on Image Processing*, 16(12):2980–2991, 2007.
- [8] M. A. T. Figueiredo and R. D. Nowak. "image restoration under wavelet-domain priors: an expectation-maximization approach". In *Proc. 2002 Int. Conf. on Image Processing*, volume 1, pages I-337–I-340 vol.1, 2002.
- [9] M. A. T. Figueiredo and R. D. Nowak. "an em algorithm for wavelet-based image restoration". *IEEE Trans. on Image Processing*, 12(8):906–916, 2003.
- [10] M. A. T. Figueiredo and R. D. Nowak. "a bound optimization approach to wavelet-based image deconvolution". In *IEEE Int. Conf. on Image Processing*, volume 2, pages II-782–5, 2005.
- [11] J. Fowler. "the redundant discrete wavelet transform and additive noise". *Signal Processing Letters, IEEE*, 12(9):629–632, 2005.
- [12] M. Lang, H. Guo, J. Odegard, C. Burrus, and R. Wells, Jr. "noise reduction using an undecimated discrete wavelet transform". *IEEE Signal Processing Letters*, 3(1):10–12, 1996.
- [13] Q. Leming. "wavelet image restoration and regularization parameters selection". In *Fourth Int. Conf. on Frontier of Computer Science and Technology*, pages 241–247, 2009.
- [14] R. D. Nowak and M. A. T. Figueiredo. "fast wavelet-based image deconvolution using the em algorithm". In *Conf. Record of the Thirty-Fifth Asilomar Conf. on Signals, Systems and Computers.*, volume 1, pages 371–375, 2001.
- [15] E. Pardo, J. San Emeterio, M. A. Rodriguez, and A. Ramos. "noise reduction in ultrasonic ndt using undecimated wavelet transforms". *Ultrasonics*, 44(Supplement 1):e1063–e1067, 2006.
- [16] M. Shensa. "the discrete wavelet transform: wedding the a trous and mallat algorithms". *IEEE Trans. on Signal Processing*, 40(10):2464–2482, 1992.
- [17] W. Sun and Y.-x. Yuan. *Optimization theory and methods nonlinear programming*. Springer Science+Business Media, New York, 2006.
- [18] M. Unger, T. Pock, M. Grabner, A. Klaus, and H. Bischof. "a variational approach to semiautomatic generation of digital terrain models". In *Proc. of the 5th Int. Symp. on Advances in Visual Computing: Part II*, pages 1119–1130, Las Vegas, Nevada, 2009. Springer-Verlag.