Impulse Denoising via Transform Learning

Jyoti Maggu
IIIT-Delhi
jyotim@iiitd.ac.in

Ramy Hussein
UBC
ramy@ece.ubc.ca

Angshul Majumdar
IIIT-Delhi
angshul@iiitd.ac.in

Rabab Ward
UBC
rababw@ece.ubc.ca

***Abstract*— This work addresses the problem of impulse denoising. Traditionally this was removed by median filtering (and its variants). In recent times, l_1-l_1 denoising techniques that employ an l_1-norm on the data fidelity term and l_1-norm on the image's sparsifying transform has been proposed. This work proposes a transform learning based formulation for the said problem. Till date, it has been used for solving Gaussian denoising problems and MRI reconstruction problems. This is the first work, that proposes solving the impulse denoising problem using the said framework. Experimental results standard images show the superiority of our proposed method with respect to traditional (sparsity) and dictionary learning based approaches.**

***Keywords*— denoising, dictionary learning, transform learning**

I. Introduction

Impulse noise is defined as an error or aberration that affects a small portion of the signal, but the aberration is large in magnitude; thus impulse noise is sparse but large. In acoustics such noise arises from electromagnetic interference, scratches on disks or from ill synchronization during digital recording. In power electronics it arises from transients owing to the devices switching ON or OFF or from current surges. Impulse noise also affects electromagnetic signals. For example in EEG signals, sharp aberrations of small duration can be seen when the subject blinks eye. Similar artefacts corrupt ECG signals as well.

In imaging impulse noise arises when the light sensor is saturated. If the saturation is on the higher end, a white spot (salt) arises; if on the lower end it is a black spot (pepper). Usually it is more commonly known as salt-and-pepper noise; however in this work we will refer by the more generic term – impulse noise. Such noise is common in hyperspectral images arising from diffraction grating and transient dead pixels [1]. Impulse noise may arise in medical images as well [2-4]. For digital photography, it arises from malfunctioning of light sensors, faulty memory locations or timing errors in digitization

The classical approach to remove impulse noise is via median filtering [5-7]; this is because the median is not affected by the extreme values (corruption) of impulse noise. However such techniques are only useful for removing noise when a relatively small portion of the image is affected – typically less than 10%. For heavier corruption, one needs using more modern techniques. The general approach is to minimize the absolute deviations subject to sparsity penalty on the transform coefficients of the image [8]. Since impulse noise is sparse minimizing the absolute deviations is sensible; the sparsity penalty on the transform (e.g. wavelet, DCT etc.) of the image is the modelling term.

More recent papers [9-11] propose learning the sparsity promoting basis, instead of assuming the basis to be fixed (wavelet, DCT, Gabor etc.). They learn the sparsifying dictionary from the data. Such dictionary learning based approaches, yield some of the best known results.

Dictionary learning (DL) is a synthesis formulation; i.e. the dictionary is learnt so as to regenerate / synthesize the data from the coefficients. Transform learning [12-14] is its analysis equivalent. It generates the features when the learnt transform operates on the data. Dictionary learning has seen a plethora of applications in signal processing and computer visions; the number of papers on this topic runs into hundreds. Transform learning (TL) on the other hand is a new topic, with only handful of papers introducing it and proving its convergence guarantees. Its potential in signal processing is yet to be fully explored. It has only been used for Gaussian denoising and MRI reconstruction [15].

In the few areas TL has been applied, TL was found to supersede DL in performance – be it denoising or MRI reconstruction. This motivates us to apply this to the problem of impulse denoising. This is the first paper on this topic.

The rest of the paper is organized into several sections. The following section discusses some relevant prior studies. The proposed formulation is given in section 3, along with the algorithm for solution. The experimental results are described in section 4. The conclusions of this work are discussed in section 5.

II. Literature Review

Impulse noise is additive in nature. It can be expressed as,

$$y = x + n \quad (1)$$

Here x is the clean image, n is the impulse noise and y the corrupted image. The task is to recover x given y and the knowledge of the distribution n.

To recover the image, one assumes that the image is sparse in some transform domain (S). For orthogonal[1] (DCT,

[1] $Orthogonal: S^T S = I = SS^T$

wavelet etc.) and tight-frame[2] (redundant wavelet, complex dualtree etc.) sparsifying transforms, the following analysis and synthesis equations hold.

$$Analysis: \alpha = Sx \tag{2a}$$

$$Synthesis: x = S^T \alpha \tag{2b}$$

The synthesis equation allows (1) to be expressed in the form:

$$y = S^T \alpha + n \tag{3}$$

From (3), the sparse coefficients can be recovered by solving,

$$\min_{\alpha} \left\| y - S^T \alpha \right\|_1 + \lambda \left\| \alpha \right\|_1 \tag{4}$$

The l_1-norm on the data fidelity stems from the fact that impulse noise is sparse; the l_1-norm on the solution promotes sparsity of the transform coefficients. Once (4) is solved ($\hat{\alpha}$), the image is recovered by applying the synthesis equation $\hat{x} = S^T \hat{\alpha}$.

We have discussed the synthesis prior formulation. This is restrictive. It can only employ orthogonal or tight-frame transforms that follow the analysis-synthesis framework. It precludes powerful linear sparsifying transforms like finite differencing / gradients (leading to total variation). Any linear transform will have an analysis representation (2a), but may not have synthesis representation (2b). To accommodate such linear operators, one needs the analysis prior formulation [16].

In the co-sparse analysis prior formulation, one directly solves for the image and not the sparse transform coefficients. Evidence [16-18] suggested that the analysis prior formulation yields better recovery in inverse problems. The same holds for impulse denoising [8]. The analysis model for impulse denoising is:

$$\min_{x} \left\| y - x \right\|_1 + \lambda \left\| Sx \right\|_1 \tag{5}$$

It is easy to verify that for orthogonal transforms the analysis and synthesis forms are equivalent; but not for tight-frame transforms. Moreover the analysis formulation is generic; it has the synthesis formulation as a special case.

These studies assumed that the sparsifying basis is fixed, i.e. they used mathematically well defined transforms like DCT, wavelet etc. They can sparsely represent a wide class of signals, but may not be the best for representation a specific problem. It is general knowledge that a learnt basis will always be better at representing a particular class of signals. In dictionary learning, one adaptively learns the basis / dictionary from data. This is expressed as,

$$X = DZ \tag{6}$$

Here X is the data, D the learnt dictionary and Z the coefficients. Perhaps KSVD [19] is the most famous dictionary learning technique. However it is not the most efficient; there are more computationally efficient learning techniques for the said task. In general it is expressed as,

$$\min_{Z} \left\| X - DZ \right\|_F^2 + \lambda \left\| Z \right\|_1 \tag{7}$$

Here the l_1-norm is defined on the vectorized version of Z. In order to prevent degenerate solutions and scale ambiguity, the dictionary atoms are normalized after every iteration.

One can learn the dictionary from generic image datasets and apply it for denoising a test image. However, there is no guarantee that such a basis (trained on a different set of images) will be the best representative for the test image. Therefore, in dictionary learning, one learns the basis from the same image as it is denoising. For Gaussian denoising this is expressed as [20],

$$\min_{D,Z,x} \left\| y - x \right\|_2^2 + \sum_i \left\| P_i x - Dz_i \right\|_2^2 + \lambda \left\| z_i \right\|_1 \tag{8}$$

The first term is for data fidelity. The term within the summation sign is the same as (7) – it is for learning the dictionary and the coefficients. Here P_i represents the patch selection operator. It selects a patch (overlapping / non-overlapping) so that it can be sparse represented (z_i) by the dictionary D.

For impulse denoising, the data fidelity term for dictionary learning [9-11] is changed from the l_2-norm to the l_1-norm. The learning is expressed as,

$$\min_{D,Z,x} \left\| y - x \right\|_2^2 + \sum_i \left\| P_i x - Dz_i \right\|_1 + \lambda \left\| z_i \right\|_1 \tag{9}$$

Solving (9) is more complicated than (8). In [9, 10] a variable splitting approach is proposed. An iterative reweighted least square type technique is proposed in [11] to solve (9).

It has been found in the aforesaid studies that learning the dictionary adaptively indeed help improve denoising performance (for both Gaussian and impulse) over fixed transforms. The intuitive reason has been discussed before.

III. Transform Learning Formulation

Dictionary learning is a well studied topic, but transform learning is relatively new. Hence we discuss it briefly for ease of the reader.

A. *Transform Learning*

Transform learning analyses the data by learning a transform / basis to produce coefficients. Mathematically this is expressed as,

$$TX = Z \tag{10}$$

Here T is the transform, X is the data and Z the corresponding coefficients.

The following transform learning formulation was proposed [12, 13] –

[2] $Tight-frame: S^T S = I \neq SS^T$

$$\min_{T,Z} \|TX - Z\|_F^2 + \lambda\left(\|T\|_F^2 - \log\det T\right) + \mu\|Z\|_1 \quad (12)$$

The factor $-\log\det T$ imposes a full rank on the learned transform; this prevents the degenerate solution (*T=0, Z=0*). The additional penalty $\|T\|_F^2$ is to balance scale; without this $-\log\det T$ can keep on increasing producing degenerate results in the other extreme.

In [12, 13], an alternating minimization approach was proposed to solve the transform learning problem. This is given by –

$$Z \leftarrow \min_{Z} \|TX - Z\|_F^2 + \mu\|Z\|_1 \quad (13a)$$

$$T \leftarrow \min_{T} \|TX - Z\|_F^2 + \lambda\left(\varepsilon\|T\|_F^2 - \log\det T\right) \quad (13b)$$

Updating the coefficients (13a) is straightforward. It can be updated via one step of soft thresholding. This is expressed as,

$$Z \leftarrow signum(TX) \cdot \max\left(0, abs(TX) - \mu\right) \quad (14)$$

Here $\odot$ indicates element-wise product.

In the initial paper on transform learning [12], a non-linear conjugate gradient based technique was proposed to solve the transform update. In the more refined version [13], with some linear algebraic tricks they were able to show that a closed form update exists for the transform.

$$XX^T + \lambda\varepsilon I = LL^T \quad (15a)$$

$$L^{-1}XZ^T = USV^T \quad (15b)$$

$$T = 0.5R\left(S + (S^2 + 2\lambda I)^{1/2}\right)Q^T L^{-1} \quad (15c)$$

The first step is to compute the Cholesky decomposition; the decomposition exists since $XX^T + \lambda\varepsilon I$ is symmetric positive definite. The next step is to compute the full SVD. The final step is the update step. One must notice that L^{-1} is easy to compute since it is a lower triangular matrix. The proof for convergence of such an update algorithm can be found in [14].

B. Impulse Denoising

Let us reiterate the problem statement. We assume an additive noise model, where *x* is the original image which is corrupted by impulse noise n generating a noisy image *y*.

$$y = x + n$$

For additive Gaussian noise, a transform learning based denoising solution has been proposed in prior studies [12, 13]. It is expressed as follows,

$$\min_{D,Z,x} \|y - x\|_2^2 + \sum_i \|TP_i x - z_i\|_2^2 + \mu\|z_i\|_1 \quad (16)$$

As in the case of dictionary learning the data fidelity term $\|y - x\|_2^2$ is an Euclidean norm owing to the Normal distribution of noise. The term $\sum_i \|TP_i x - z_i\|_2^2 + \mu\|z_i\|_1$ is the transform learning term.

Since we want to remove impulse noise, the only change in the formulation (16) be in the data fidelity term; we have to change it to the more robust l_1-norm. The final expression for impulse denoising will be,

$$\min_{T,Z,x} \|y - x\|_1 + \sum_i \|TP_i x - z_i\|_2^2 + \mu\|z_i\|_1 \quad (17)$$

As we all know, changing the smooth l_2-norm with the non-differentiable l_1-norm increases the complexity of the solution. Solving (17) is not as straightforward as (16). We propose to derive a solution to (17) using the Split Bregman approach.

We introduce a proxy $p = y - x$. We add terms relaxing the equality constraints between the variable (x) and its proxy (p); in order to enforce equality at convergence, we introduce Bregman variables b. The new objective function is:

$$\min_{T,Z,x,p} \|p\|_1 + \gamma\|p - (y - x) - b\|_2^2 + \sum_i \|TP_i x - z_i\|_2^2 + \mu\|z_i\|_1 \quad (18)$$

Alternating minimization allows (18) to be expressed in terms of the following sub-problems:

$$\text{P1}: \min_{T,Z} \sum_i \|TP_i x - z_i\|_2^2 + \mu\|z_i\|_1$$

$$\text{P2}: \min_{x} \gamma\|p - (y - x) - b\|_2^2 + \sum_i \|TP_i x - z_i\|_2^2$$

$$\text{P3}: \min_{p} \|p\|_1 + \gamma\|p - (y - x) - b\|_2^2$$

Sub-problem P1 is the standard transform learning problem. We have already studied the solution in the previous section. Sub-problem P2 is a simple least squares problem. It has a closed form solution in the form of pseudoinverse; however in this work we use conjugate gradient to solve for *x*. Sub-problem P3 is an l_1-minimization problem. It has a closed form update in the form of a single step of soft thresholding.

The final step of the Split Bregman technique is to update the relaxation variable. This is done by simple gradient descent.

$$b \leftarrow p - (y - x) - b$$

The iterations continue till the objective function converges to a local minima, or till a maximum number of specified iterations.

IV. Experimental Results

In this work experiments are carried out on some standard test images – Barbara, Cameraman, Peppers and Rice. Tests were carried out with 2%, 5% and 10% salt and pepper noise.

We have compared with the dictionary learning (DL) based method for impulse denoising [9]. We have also compared with the synthesis (using wavelet transform) and analysis (using total variation – TV) prior formulations proposed in [8]. For evaluation, the metric we have used is structural similarity index (SSIM); this is because SSIM is known to correlate better with human evaluation compared to PSNR.

The quantitative results are shown in the following tables. The optimal settings (number of dictionary atoms, parameters, sparsity etc.) for the methods compared against have been taken from the corresponding papers; the parameters vary for each noise level and we have used prescribed in the papers that are supposed to yield the best possible results.

For our proposed transform learning (TL) based method, a redundant transform (2x) is used. The patch sizes are 8x8; non-overlapping patches are considered. It is initialized with concatenated wavelet and DCT matrices. Our method requires specification of one parameter (μ) and one hyper-parameter (γ); μ has been fixed on a validation image (Lena) for each noise level; the values used are 0.25 for 2% and 5% noise and 0.5 for 10% noise. We found that the value of the hyperparameter did not have much effect on the denoising performance; we kept it fixed at 0.1.

TABLE I. SSIM FOR 2% IMPULSE NOISE

Image	Proposed	DL [9]	TV [8]	Wavelet [8]
Barbara	**0.93**	**0.93**	0.90	0.88
Cameraman	**0.92**	**0.92**	**0.92**	0.89
Peppers	0.93	**0.95**	**0.95**	0.92
Rice	0.92	**0.94**	**0.94**	0.92

TABLE II. SSIM FOR 5% IMPULSE NOISE

Image	Proposed	DL [9]	TV [8]	Wavelet [8]
Barbara	**0.90**	0.86	0.89	0.86
Cameraman	**0.91**	0.88	0.90	0.87
Peppers	**0.91**	0.89	**0.91**	**0.91**
Rice	**0.90**	0.88	0.89	**0.91**

TABLE III. SSIM FOR 10% IMPULSE NOISE

Image	Proposed	DL [9]	TV [8]	Wavelet [8]
Barbara	**0.84**	0.80	0.78	0.80
Cameraman	**0.83**	0.81	0.79	0.79
Peppers	**0.87**	0.82	0.83	0.83
Rice	**0.85**	0.81	0.81	0.80

The results are interesting. We find that for low noise, the existing techniques perform better overall. Dictionary Learning based removal yields the best results. TV denoising also yields results almost at par with DL.

When the noise increases, i.e. with moderate (5%) corruption of pixels, our proposed transform learning based formulation takes over. It yields the best results. DL and TV suffer significant drop in SSIM. Comparatively, the drop in performance is less in the wavelet based synthesis formulation.

With even more noise – 10% corruption of pixels, our method performs the best. The results of transform learning do not deteriorate much from the 5% noise. But there is a significant fall in denoising performance for all the other techniques.

For qualitative evaluation we have shown the original, noisy and denoised images for 10% noise in the following figure. Visual evaluation corroborates the numerical results. One notices that the proposed method produces the cleanest results with almost negligible smoothing. The dictionary learning based method preserves contours but cannot get rid of all the noise. The TV based denoised method removes noise but is overtly smooth. The wavelet based denoising produces the worse results; it cannot remove significant portion of noise.

Fig. 1. Left to Right – Original , Noisy, Proposed, Dictionary Learning, TV, Wavelet

V. CONCLUSION

This work proposes a technique for removing impulse noise based on a transform learning based formulation. Results show that our method yields the best results when compared with state-of-the-art dictionary learning based and sparsity based (wavelet denoising and total variation regularization) methods. This work experiments on simulated noise. In future, we would like to see how this technique fairs in real life scenarios like hyperspectral denoising and EEG / ECG artifact removal.

ACKNOWLEDGEMENT

This work was made possible by NPRP grant 7-684-1-127 from the Qatar National Research Fund (a member of Qatar Foundation). The statements made herein are solely the responsibility of the authors.

REFERENCES

[1] S. Tariyal, H. Agrawal and A. Majumdar, “"Removing Sparse Noise from Hyperspectral Images with Sparse and Low-rank Penalties”, SPIE Journal of Electronic Imaging

[2] Sadri, A. R., Zekri, M., Sadri, S., & Gheissari, N. (2012). Impulse noise cancellation of medical images using wavelet networks and median filters. Journal of medical signals and sensors, 2(1), 25.

[3] Toprak, A., & Güler, İ. (2007). Impulse noise reduction in medical images with the use of switch mode fuzzy adaptive median filter. Digital signal processing, 17(4), 711-723.

[4] Toprak, A., & Güler, İ. (2006). Suppression of impulse noise in medical images with the use of fuzzy adaptive median filter. Journal of Medical systems, 30(6), 465-471.

[5] Loupas, T., McDicken, W. N., & Allan, P. L. (1989). An adaptive weighted median filter for speckle suppression in medical ultrasonic images. IEEE transactions on Circuits and Systems, 36(1), 129-135.

[6] Sun, T., & Neuvo, Y. (1994). Detail-preserving median based filters in image processing. Pattern Recognition Letters, 15(4), 341-347.

[7] Abreu, E., Lightstone, M., Mitra, S. K., & Arakawa, K. (1996). A new efficient approach for the removal of impulse noise from highly corrupted images. IEEE transactions on image processing, 5(6), 1012-1025.

[8] H. Agarwal and A. Majumdar, "Generalized Synthesis and Analysis Prior Algorithms with Application to Impulse Denoising", ICVGIP 2014.

[9] Wang, S., Liu, Q., Xia, Y., Dong, P., Luo, J., Huang, Q., & Feng, D. D. (2013). Dictionary learning based impulse noise removal via L1–L1 minimization. Signal Processing, 93(9), 2696-2708.

[10] A. Majumdar, N. Ansari, H. Agarwal and P. Biyani, “Impulse Denoising for Hyper-Spectral Images: A Blind Compressed Sensing Approach”, Signal Processing, Vol. 119, pp. 136-141, 2016.

[11] S. Mukherjee, R. Basu, and C. S. Seelamantula, "L1-K-SVD: A robust dictionary learning algorithm with simultaneous update," vol. 123, pp. 42-52, June 2016, Elsevier (Signal Processing).

[12] Ravishankar, Saiprasad, and Yoram Bresler. "Learning sparsifying transforms." IEEE Transactions on Signal Processing 61.5 (2013): 1072-1086.

[13] Ravishankar, Saiprasad, Bihan Wen, and Yoram Bresler. "Online sparsifying transform learning-Part I: Algorithms." IEEE Journal of Selected Topics in Signal Processing 9.4 (2015): 625-636.

[14] Ravishankar, Saiprasad, and Yoram Bresler. "Online Sparsifying Transform Learning-Part II: Convergence Analysis." IEEE Journal of Selected Topics in Signal Processing 9.4 (2015): 637-646.

[15] S. Ravishankar and Y. Bresler, “Efficient Blind Compressed Sensing using Sparsifying Transforms with Convergence Guarantees and Application to MRI,” SIAM Journal on Imaging Sciences, 2015.

[16] I. W. Selesnick and M. A. T. Figueiredo, “Signal restoration with overcomplete wavelet transforms: comparison of analysis and synthesis priors”, Proceedings of SPIE, Vol. 7446 (Wavelets XIII), 2009.

[17] A. Majumdar and R. K. Ward, “On the Choice of Compressed Sensing Priors: An Experimental Study”, Signal Processing: Image Communication, Vol. 27 (9), pp. 1035–1048, 2012.

[18] A. Majumdar and R. K. Ward, “Accelerating Multi-Echo T2 Weighted MR Imaging: Analysis Prior Group Sparse Optimization”, Journal of Magnetic Resonance, Vol. 210 (1), pp. 90-97, 2011.

[19] M. Aharon, M. Elad and A. Bruckstein, "K-SVD: An Algorithm for Designing Overcomplete Dictionaries for Sparse Representation", IEEE Transactions on Signal Processing, Vol. 54 (11), pp. 4311-4322, 2006.

[20] M. Elad and M. Aharon, "Image Denoising Via Sparse and Redundant Representations Over Learned Dictionaries," IEEE Transactions on Image Processing, Vol.15 (12), pp. 3736-3745, 2006.