# Joint denoising and decompression using CNN regularization

Mario González, Javier Preciozzi, Pablo Musé, Andrés Almansa

# Joint denoising and decompression using CNN regularization

Mario González*[1], Javier Preciozzi[1], Pablo Musé[1], and Andrés Almansa[2]

[1]Universidad de la República, Uruguay
[2]CNRS & MAP5, Université Paris Descartes, France

## Abstract

*Wavelet compression schemes (such as JPEG2000) lead to very specific visual artifacts due to the quantization of noisy wavelet coefficients. They have highly spatialy-correlated structure that makes it difficult to be removed with standard denoising algorithms. In this work, we propose a joint denoising and decompression method that combines a data-fitting term which takes into account the quantization process and an implicit prior contained in a state-of-the-art denoising CNN.*

## 1. Introduction

Transform coding image compression consists of applying a linear invertible transform that sparsifies the data (like block-wise DCT for JPEG compression or a Wavelet Transform in JPEG2000) followed by a quantization of the transformed coefficients which are finally compressed by a lossless encoder. This family of compression schemes may achieve very high compression ratios but may lose some details in the quantization step. This lossy quantization is also responsible for well-known artifacts that may appear in the compressed image in the form of texture loss or Gibbs effects near edges. Many solutions have been proposed in the literature to remove some of these artifacts: many of them are variational and involve the minimization of the total variation (to minimize ringing) among all images that would lead to the observed quantized image [6, 3, 15].

However little attention has been paid in previous works to the fact that the image to be compressed may contain noise, and that noise may interact in subtle ways with the compressor, producing new kinds of artifacts that we call *outliers* (see figure 1). Those artifacts cannot be removed by the previously cited works which only aim at removing compression artifacts but not noise or its complex interactions with the compressor. However such artifacts are particularly annoying in the case of wavelet-based compressors like JPEG2000 and the CCSDS recommendation [8], which are extensively used to compress digital cinema and high-resolution remote sensing images. More recently, joint denoising and decompression procedures have been considered to remove both artifacts due to the compressor and its interaction with noise. Such methods use either TV regularization or patch-based Gaussian models in combination with relaxed versions of the quantization constraint, in order to take the effects of noise into account [7, 12, 13]. However the TV based approaches could only reliably remove isolated outliers in relatively constant areas, and patch-based approaches could only marginally improve the performance of standard denoising techniques like non-local Bayes [10].

In this work we propose a novel method for joint denoising and decompression. Our method uses a probabilistic data-fitting term similar to [13] coupled with a CNN based regularization which more closely captures natural image statistics than previously proposed patch-based methods. The proposed method is detailed in Section 3 after a short modeling of our joint denoising and decompression problem in Section 2. The rest of the paper includes the numerical implementation details (Section 4) and our experimental results (Section 5).

## 2. Modeling Noisy Compressed Coefficients

We assume that our image $u$ is corrupted by additive white Gaussian noise $n_u \sim N(0, \sigma^2 I)$.[1] The first step of the CCSDS compression applies a wavelet transform W to the noisy image. Hence the corresponding wavelet coeffi-

*This work has been partially funded by the Uruguayan Research and Innovation Agency (ANII) under grant number (FCE_1_2017_1_135458), by the French-Uruguayan cooperation program under grant number ECOS Sud U17E04, by France's Eiffel Scholarship Program of Excellence under grant number 895184E, and by the French Research Agency (ANR) under grant number ANR-14-CE27-001 (MIRIAM).

[1] Even though sensors usually produce a mixture of additive and multiplicative noise [1, ch2], a variance stabilizing transform (which also compresses the dynamic range) is usually applied before compression, and the image approximately follows our noise model after VST.

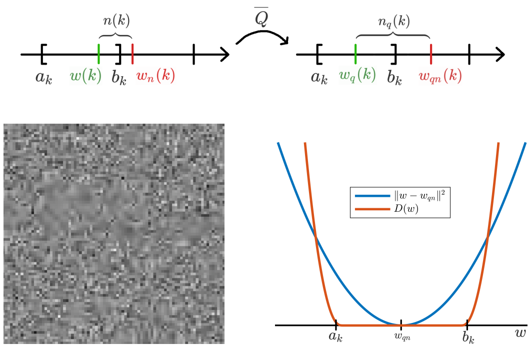


Figure 1: Above: When the noise $n(k)$ added to a coefficient $w(k)$ changes its quantization interval, we get an *outlier*, and the noise may be amplified. Below: as $q(k)/\sigma$ varies, the noise is heterogeneous. So we use a better datafit term than $L^2$ norm.

cients are corrupted by Gaussian noise

$$n := \underbrace{W(u+n_u)}_{w_n} - \underbrace{W(u)}_{w} = Wn_u \sim N(0, \sigma^2 WW^T).$$

If the wavelet transform was orthogonal then $WW^T = I$ and the $n_w$ would be white. Most compression algorithms use, however, the CDF 9/7 *biorthogonal* wavelet transform [5], but even in that case $WW^T \simeq I$ is a good approximation [2] since $W$ is *nearly orthogonal*.

Each wavelet coefficient is then quantized by setting to 0 its $m(k)$ least significant bits:

$$Q(w_n(k)) := \operatorname{sign}(w_n(k)) \left\lfloor \frac{|w_n(k)|}{2^{m(k)}} \right\rfloor 2^{m(k)}$$

The values of $m(k)$ are chosen by the compression algorithm to optimize the rate/distortion trade-off, and can be recovered from the compressed image. From these values we can recover the quantization intervals $Q^{-1}(w_n(k)) = [a(k), b(k)]$[2], as well as their centers $\overline{Q^{-1}}(w_n(k))$. The standard decoder yields

$$u_{qn} := W^{-1} w_{qn} = W^{-1} \underbrace{\overline{Q^{-1}} Q}_{\overline{Q}}(w+n).$$

As illustrated in Figure 2, $w_{qn}$ may be corrupted by *outliers* due to the interaction between $n$ and the codec $\overline{Q}$. To understand why this occurs consider the situation depicted in Figure 1. If there was no noise we would obtain the quantized coefficients $w_q := \overline{Q}(w)$. When the noise level $\sigma \ll q(k)$ is relatively small, the noisy quantized coefficient $w_{qn} := \overline{Q}(w+n)$ is most often equal to $w_q$, and the quantizer has a denoising effect. However, occasionally the noise may be large enough to change the quantization interval. In that case quantization may amplify the noise $|n_q| = |w_{qn} - w_q| > |w_n - w| = |n|$, and we get a visible (wavelet shaped) artifact that we call an *outlier*. Outliers are particularly annoying when they are isolated. When the noise level $\sigma \gtrsim q(k)$ is similar to or larger than the quantization level then outliers occur everywhere and they appear indistinguishable from white noise.

In the next section we propose a Bayesian approach to estimate the original image $u$ from its noisy, quantized observation. The datafit term will be formulated in the wavelets domain; this is the natural choice since quantization is performed on this domain.

[2] of length $q(k) = 2^{m(k)}$ except for the case $Q(w_n(k)) = 0$ where the quantization interval is of length $q(k) = 2^{m(k)+1}$.

# 3. Proposed restoration method

## 3.1. Motivation via MAP estimation

The *Maximum A Posteriori (MAP)* estimation of the non-degraded image $u$ knowing its degraded version $u_{qn}$ is stated as

$$\hat{u} = \arg\max_u p(u|u_{qn}) = \arg\max_u p(u_{qn}|u)p(u) \tag{1}$$

$$= \arg\min_u -\log(p(u_{qn}|u)) - \log(p(u)), \tag{2}$$

where $\hat{u}$ is the MAP estimator of $u$. Finding $\hat{u}$ amounts to solve the optimization problem

$$\hat{u} = \arg\min_u D(u) + \lambda R(u), \tag{3}$$

where $D(u)$ is a data-fitting term that depends on the *forward* operator and the noise model, $R$ is the regularization ($-\log(\text{prior})$) to be used in the restoration, and $\lambda > 0$ is the strength of the regularization.

## 3.2. Data fitting

Let $w = Wu$ the coefficients of the original (unknown) image, and $w_{qn} = Wu_{qn}$ the wavelet coefficients of the corrupted image. As stated before, the quantization intervals of each of these coefficients can be retrieved as $[a(k), b(k)] = Q^{-1}(w_{qn}(k)) = \overline{Q}^{-1}(w_{qn}(k))$. Using this notation, and given that the noise in the wavelet domain is $N(0, \sigma^2 I)$ (Section 2), the conditional probability of the corrupted coefficients given the original ones is

$$p(w_{qn}|w = \omega) = \prod_k p(w_{qn}(k)|w(k) = \omega(k)) \tag{4}$$

$$= \prod_k p(\overline{Q}(\omega(k) + n(k)) = w_{qn}(k)) \tag{5}$$

$$= \prod_k p(\omega(k) + n(k) \in [a(k), b(k)]) \tag{6}$$

$$= \prod_k p\left(\frac{n(k)}{\sigma} \in \left[\frac{a(k) - \omega(k)}{\sigma}, \frac{b(k) - \omega(k)}{\sigma}\right]\right). \tag{7}$$

In the following we consider the log-likelihood function

$$D(\omega) = -\log p(w_{qn}|w=\omega) \tag{8}$$

$$= -\sum_k \log\left(\phi\left(\frac{b(k)-\omega(k)}{\sigma}\right) - \phi\left(\frac{a(k)-\omega(k)}{\sigma}\right)\right), \tag{9}$$

where $\phi$ is the normal cumulative distribution function. This data term in the wavelet domain carefully takes into account the quantization process of the coefficients. Although this term is not quadratic as in most inverse problems, at least it is convex and we have an analytic expression for its gradient and its Hessian matrix [17].

### 3.3. Minimization with ADMM

Finally, problem (3) can be written as

$$\min_{w,u} D(w) + \lambda R(u) \qquad \text{s.a.}\;\; W^{-1}w = u$$

where $W^{-1}$ is the inverse wavelet transform (synthesis). The ADMM algorithm [4] becomes (subscripts indicate the iteration number):

$$\begin{cases} w_{k+1} = \underset{w}{\operatorname{argmin}}\;\; D(w) + \frac{\rho}{2}\|W^{-1}w - u_k + \frac{1}{\rho}y_k\|^2 \\ u_{k+1} = \underset{u}{\operatorname{argmin}}\;\; \lambda R(u) + \frac{\rho}{2}\|W^{-1}w_{k+1} - u + \frac{1}{\rho}y_k\|^2 \\ y_{k+1} = y_k + \rho(W^{-1}w_{k+1} - u_{k+1}). \end{cases} \tag{10}$$

### 3.4. Regularizing by denoising

The second subproblem can be rewritten as

$$u_{k+1} = \underset{x}{\operatorname{argmin}}\;\; \frac{1}{2(\lambda/\rho)}\left\|(W^{-1}w_{k+1} + \frac{1}{\rho}y_k) - u\right\|^2 + R(u).$$

In terms of MAP estimation, this step can be seen as a *Gaussian denoising* of $W^{-1}w_{k+1} + \frac{1}{\rho}y_k$ with noise variance $\sigma^2_{\mathcal{G}} = \lambda/\rho$. The solution can be computed using a good denoiser $\mathcal{G}$ as the proximal operator of an *implicit prior* $R(u)$ [11]:

$$u_{k+1} = \mathcal{G}(W^{-1}w_{k+1} + \frac{1}{\rho}y_k,\;\; \sigma^2_{\mathcal{G}} = \lambda/\rho).$$

## 4. Numerical implementation

For the first subproblem in (10), let $v = -u_k + \frac{1}{\rho}y_k$, then define $F(w) := D(w) + \frac{\rho}{2}\|W^{-1}w + v\|^2$. The first and second derivatives of $F(w)$ are given by

$$\nabla F(w) = \nabla D(w) + \rho W^{-T}(W^{-1}w + v)$$
$$\nabla^2 F(w) = \nabla^2 D(w) + \rho W^{-T}W^{-1}.$$

As pointed out before, for the CDF9/7 the term $W^TW$ can be fairly approximated by the identity matrix $I$, yielding

$$\nabla^2 F(w) \simeq \nabla^2 D(w) + \rho\, I.$$

Now, since $D(w)$ is separable in terms of the elements $w(k)$ of $w$, it follows that $\nabla^2 D(w)$ is a diagonal matrix. It is also positive definite, since function $D(w)$ is strictly convex[3]. It follows that $\nabla^2 F(w)$ is a diagonal, positive definite matrix, and therefore the minimization of $F(w)$ can be computed very efficiently using a Newton method:

$$\begin{aligned} w^0 &= W^{-1}u_k \\ w^{j+1} &= w^j - \alpha_j(\nabla^2 F(w^j))^{-1}\nabla F(w^j) \\ &= w^j - \alpha_j(\nabla^2 D(w^j) + \rho\, I)^{-1} \\ &\qquad\qquad (\nabla D(w^j) + \rho W^T(Ww^j + v)). \end{aligned}$$

Finally, the second subproblem in (10) can be computed by means of a Gaussian denoiser, as described in Section 3.4. We choose $\mathcal{G}$ to be the residual network of Zhang et al. [16], which is state of the art in Gaussian denoising.

## 5. Results and Conclusions

Figure 2 presents the results on the Lena image. This example clearly shows one of the major problems derived from the compression of noisy images: the presence of wavelet outlier coefficients introduces visual artifacts (compare the second and third images on top). It is clear that a Gaussian denoiser like [16] cannot completely remove all the artifacts, which are still present in the scene.

We can distinguish two different regimes: in regions where $q \ll \sigma$ almost all coefficients contain wavelet outliers induced artifacts and the degradation is very close to white noise. In this situation, which corresponds to the experiment presented here, the performances of [16] and our method are similar. However, when $q \gg \sigma$ artifacts become more isolated events, and the degradation deviates significantly from white Gaussian noise. In this case, not shown here, [16] cannot get its full potential, but our method performs particularly well.

In Table 1 we present the quantitative analysis of the method by comparing the PSNR, SSIM [14] and NLP [9] obtained for the proposed method against WNLB and [16].

Even though [16] exhibits slightly better objective quality measures, a visual inspection of the results in Figure 2 reveals that the proposed method removes more outliers while preserving more image detail. Previous state of the art method WNLB is clearly far behind both in terms of visual quality and objective measures.

[3] It can be shown that when $\sigma \ll q$, or when the $w(k)$ are far from its interval bounds, the inversion of $\nabla^2 D(w)$ is ill-conditioned. However, the matrix to be inverted is $\nabla^2 F(w)$, a regularized version of $\nabla^2 D(w)$.

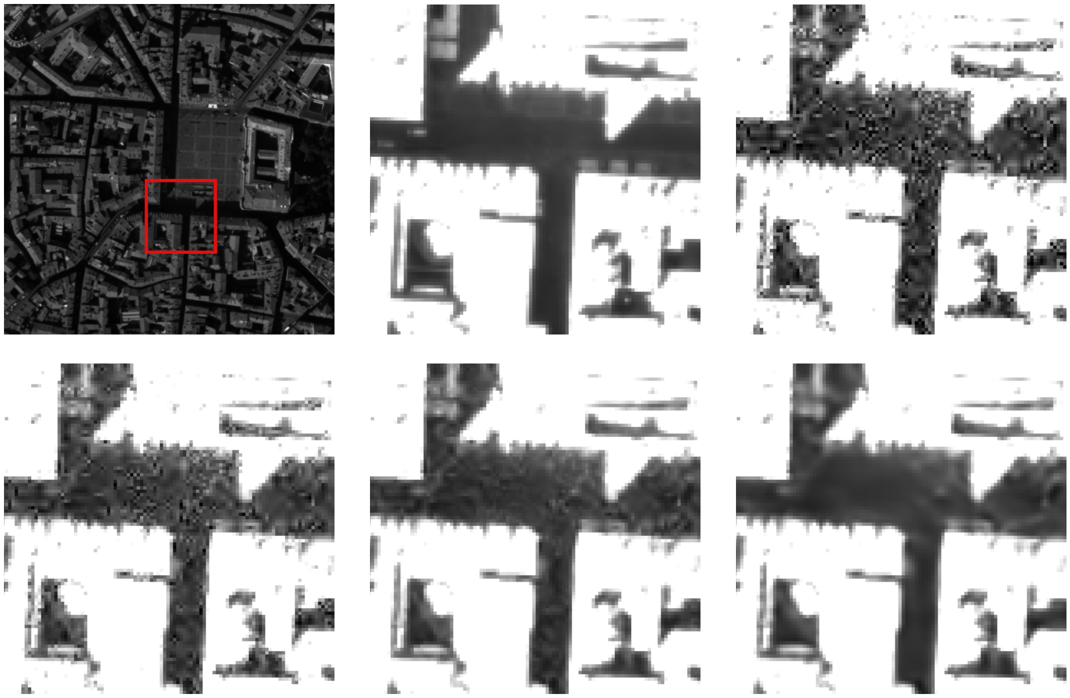

Figure 2: Top: original (St.Michel), zoomed region, compressed with noise ($\sigma = 4$, BPP = 2). Below: results from WNLB, denoiser [16] only, and our method. Dynamic range on zoomed image was saturated for better visualization.

| Image | PSNR | SSIM | NLP |
|---|---|---|---|
| Corrupted | 35.92 | 0.8320 | 5.28 |
| WNLB | 36.67 | 0.8537 | 30.70 |
| [16] | **39.59** | 0.9169 | 6.39 |
| Ours | 39.52 | **0.9241** | **2.95** |

Table 1: Results. For PSNR and SSIM, higher is better. For NLP, lower is better.

## References


[1] C. Aguerrebere. *On the Generation of High Dynamic Range Images Theory and Practice from a Statistical Perspective*. Phd, Telecom ParisTech, 2014. 1

[2] J. Akhtar. Optimization of biorthogonal wavelet filters for signal and image compression. Master's thesis, Hovedoppgave, University of Oslo, 2001. 2

[3] F. Alter, S. Durand, and J. Froment. Adapted total variation for artifact free decompression of jpeg images. *JMIV*, 23:199–211, 2005. 1

[4] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein. Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers. 3(1):1–122, 2011. 3

[5] A. Cohen, I. Daubechies, and J. Feauveau. Biorthogonal bases of compactly supported wavelets. *Communications on Pure and Applied Mathematics*, 45(5):485–560. 2

[6] S. Durand and J. Froment. Reconstruction of wavelet coefficients using total variation minimization. *SIAM, Journal on Scientific Computing*, 24(5):1754–1767, 2003. 1

[7] S. Durand and M. Nikolova. Denoising of frame coefficients using $\ell^1$ Data-Fidelity term and Edge-Preserving regularization. *SIAM Multiscale Modeling & Simulation*, 6(2):547–576, 2007. 1

[8] T. C. C. for Space Data Systems (CCSDS). Technical report, image data compression. 122.0-b-1, 2005. [URL]. 1

[9] V. Laparra, J. Ballé, A. Berardino, and E. Simoncelli. Perceptual image quality assessment using a normalized laplacian pyramid. In B. Rogowitz, T. N. Pappas, and H. de Ridder, editors, *Proc. IS&T Int'l Symposium on Electronic Imaging, Conf. on Human Vision and Electronic Imaging*, volume 2016, pages 1–6, San Francisco, CA, 14-18 Feb 2016. Society for Imaging Science and Technology. 3

[10] M. Lebrun, A. Buades, and J.-M. Morel. A nonlocal bayesian image denoising algorithm. *SIAM Journal on Imaging Sciences*, 6(3):1665–1688, 2013. 1

[11] T. Meinhardt, M. Moeller, C. Hazirbas, and D. Cremers. Learning Proximal Operators: Using Denoising Networks for Regularizing Inverse Imaging Problems. apr 2017. 3

[12] J. Preciozzi. *Two Restoration Problems In Satellite Imaging*. PhD thesis, Universidad de la República, 2016. [URL]. 1

[13] J. Preciozzi, M. Gonzalez, A. Almansa, and P. Muse. Joint denoising and decompression: A patch-based Bayesian approach. In *ICIP*, pages 1252–1256. IEEE, sep 2017. 1

[14] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE Trans. Image Processing*, 13(4):600–612, 2004. 3

[15] P. Weiss, L. Blanc-Feraud, T. Andre, and M. Antonini. Compression artifacts reduction using variational methods : Algorithms and experimental study. In *ICASSP*, pages 1173–1176, March 2008. 1

[16] K. Zhang, W. Zuo, S. Gu, and L. Zhang. Learning deep cnn denoiser prior for image restoration. In *CVPR*, pages 3929–3938, 2017. 3, 4

[17] A. Zymnis, S. Boyd, and E. Candes. Compressed sensing with quantized measurements. *IEEE Signal Processing Letters*, 17(2):149–152, 2010. 3