A NON-PARAMETRIC APPROACH FOR THE ESTIMATION OF INTENSITY-FREQUENCY DEPENDENT NOISE

M. Colom[⋆] *M. Lebrun*[†] *A. Buades*[⋆] *J.M. Morel* [†]

[⋆] DMI, Universitat de les Illes Balears, Crta. de Valldemossa, km 7.5, 07122 Palma de Mallorca, Spain
[†]CMLA, École Normale Supérieure de Cachan, 61 av. du Président Wilson, 94235 Cachan, France

ABSTRACT

We present a non-parametric method estimating an intensity and frequency dependent noise from a single image. The noise model is estimated on image patches and can be used consequently in all patch-based denoising methods. The method applies to cases where no access is granted to the image noise model, in particular to scanned photographs and JPEG images. The general noise model and the method to evaluate it are validated by comparing the estimations with the corresponding ground-truth curves for raw and JPEG images. Denoising experiments on scanned photographs also support the efficiency of the estimation method.

***Index Terms*—** Blind noise estimation, signal-dependent noise, frequency-dependent noise, non-parametric noise model.

1. INTRODUCTION

Noise in a digital image comes from several sources and it is transformed at each step of the processing chain of the camera. When it is acquired at the CFA, it is Poisson distributed, signal-dependent and frequency-independent. The noise at the CFA is possibly saturated and will not obey the simple linear dependency of the noise variance with the intensity [1]. It sometimes does not follow at all the linear model, even when no captor saturation occurs [2]. After demosaicing [3, 4], the noise becomes spatially correlated [5, 6] and therefore frequency-dependent. After gamma-correction, it gets even more saturated and finally, after JPEG-encoding [7] it turns into a strongly frequency-dependent noise. After JPEG compression, most of the noise at the high-frequencies is lost because of the coefficient quantization, but there remains medium and low frequency signal dependent noise. The situation is even more complex when we deal with a JPEG image (or even a raw scan) of an old photograph, in which case chemical noise is mixed with digital noise. In such cases, the assumption that the resulting noise is both signal and frequency dependent is a minimal model to cope with its complexity. Our purpose here is to find a general method for estimating such complex noise, and to validate it by comparing the estimated results to the appropriate ground truths.

Our plan is as follows. Sec. 2 reviews the literature and details the proposed noise estimation algorithm to measure intensity-frequency dependent noise on JPEG images. Sec. 3 validates the noise curves obtained with the proposed method by comparing them with the ground-truth (GT) curves for each frequency obtained from a set of 100 snapshots of the same calibration pattern. Sec. 3.1 performs a final validation by displaying denoising results [8] with the NL-Bayes algorithm [9]. Sec. 4 presents the conclusions.

2. NOISE ESTIMATION ALGORITHM

Little has been written on frequency and signal dependent noise estimation from a digital image. A method estimating a "JPEG compression history" from a single image can be found in [10]. The noise estimation method for JPEG images proposed in [11] estimates a signal dependent noise level which is not frequency dependent and therefore only gives a "noise level". Probably the most complete attempt to estimate a general noise model is contained in the blind denoising method [12], which estimates multiscale noise covariances for noise wavelet coefficients. This model is nevertheless not signal dependent. To the best of our knowledge, no method has proposed so far to estimate a general frequency and signal dependent noise patch model. The situation is nonetheless favourable, as most homoscedastic noise estimation algorithms are actually block based [13, 14, 15, 16, 17], and can therefore be adapted to measure signal and frequency dependent noise models on patches. Following the review of these methods in [18], we decided to adapt an existing frequency-dependent method by Ponamarenko et al. [5] to estimate the noise variance depending on both the intensity and the frequency. Our proposed method follows:

1. Extract from the input image $\mathbf{u}$ of size $N_x \times N_y$ all possible $M = (N_x - w + 1)(N_y - w + 1)$ overlapping $w \times w$ blocks $\mathbf{B}_k$ and compute their 2D orthonormal DCT-II, $\tilde{\mathbf{B}}_k, k \in [0, M-1]$.

2. Set $\mathbf{L} = \emptyset$ (the empty set).

3. For each DCT block $\tilde{m}_1 \in \tilde{\mathbf{B}}$,

(a) Find the block $\tilde{m}_2$ that minimizes $\text{PMSE}_{\tilde{m}_1,\tilde{m}2}$ (Eq. 3). Consider only those blocks whose horizontal and vertical distance with respect to $\tilde{m}_1$ belongs to the interval $[r_1, r_2] = [4, 14]$.

(b) Add block $\tilde{m}_1$ and its PMSE, $[\tilde{m}_1, \text{PMSE}_{\tilde{m}_1,\tilde{m}2}]$, to list $\mathbf{L}$.

4. Extract [1] from $\tilde{m}_1$ the mean of m_1.

5. Classify the elements of list **L** into disjoint bins according the mean intensity of the blocks [18, 19]. Each bin contains (with the exception of the last) 42000 DCT blocks.

Then, for each bin,

1. Consider the set $\mathbf{S}_p$ made by the DCT blocks inside the current bin whose PMSE is below the p-quantile, with $p = 0.005$.

2. Assign to the current bin the intensity I as

$$I = \underset{\tilde{m}\in\mathbf{S}_p}{\text{median}}\,(\tilde{m}[0,0]/w) \tag{1}$$

3. For each frequency $[i, j]$ with $[i, j] \in [0, w-1]^2$, $[i, j] \neq [0, 0]$,

(a) Compute the (biased[2]) variance of the noise at the current bin and frequency $[i, j]$ using the MAD estimator (Eq. 4).

(b) Correct the biased variance and obtain the final estimate

$$\tilde{\sigma}[I][i, j] = 1.967\hat{\sigma}[I][i, j] - 0.2777. \tag{2}$$

The correction function in Eq. (2) is obtained by adding simulated homoscedastic noise to a set of noise-free images and afterwards adjusting a linear function that returns the theoretical STD given the biased estimate $\hat{\sigma}[I][i, j]$. The above self-explanatory algorithm involves the following straighforward formulas:

$$\text{PMSE}_{\tilde{m}_1,\tilde{m}_2} := \frac{1}{w^2}\sum_{i=0}^{w}\sum_{j=0}^{w}(\tilde{m}_1[i,j]-\tilde{m}_2[i,j])^2(w^2+1-i-j)^2; \tag{3}$$

$$\hat{\sigma}[I][i, j] = \text{MAD}(\mathbf{S}_p) = \underset{\tilde{n}\in\mathbf{S}_p}{\text{median}}\left(\left|\tilde{n}[i, j] - \underset{\tilde{m}\in\mathbf{S}_p}{\text{median}}\,(\tilde{m}[i, j])\right|\right). \tag{4}$$

[1] This operation is fast since the mean of m_1 can be obtained as $\tilde{m}_1[0, 0]/w$.

[2] The estimate is biased because of the MAD estimator and because the variance is measured using a finite number of samples from **L**.

Unlike the original method that directly computes the MSE between the DCT blocks, we propose to use a ponderated-MSE (PMSE, Eq. 3) that gives more importance to the low-frequencies of the blocks in the comparison, since most of the geometric information is located there, whereas the random variation at high-frequencies is mostly explained by the noise. It should therefore be given less weight in the comparison.

Also, we found that storing block $\tilde{m}_1$ instead of $\tilde{m}_1 - \tilde{m}_2$ in step 3b (as the original method does), increases the accuracy. A deeper discussion about the effect of the subtraction is beyond the scope of this short paper.

3. VALIDATION OF THE METHOD

The above proposed method gives an estimation of the standard deviation (STD) of the noise that depends both on the intensity and frequency in a single image. It uses the observation of blocks at many spatial locations and is therefore called the *spatial* estimation.

We can validate the spatial estimation method by taking raw and JPEG photographs with a given camera. The value of the spatially estimated STD *on a single image* should match the ground-truth STD for that camera for the configured ISO speed [1], obtained from multiple frames. For that purpose, consider a sequence of images of the same scene taken with fixed camera position and constant lighting. Under these conditions, any variation of the intensity in any pixel through the sequence is only attributable to the effect of the noise. It is therefore possible to build a GT noise curve for both raw and JPEG-encoded images, associating with each observed mean signal value the corresponding standard deviation of its observed samples. Similarly, by **frequency noise curve** we mean a numerical function associating with each value of the block mean a standard deviation (STD) of the DCT coefficient of the noise at that frequency. Thus, there are as many noise curves as DCT coefficients. To obtain such curves, instead of measuring the variation of the intensity of the pixels in a fixed position along the sequence, we consider all M overlapping $w \times w$ blocks in the image, compute their orthonormal DCT-II, and measure the variance at the intensity of the bin and frequency $[i, j] \in [0, w - 1]^2$, $[i, j] \neq [0, 0]$ along the coefficients of the blocks at the same spatial position and with varying image index.

The noise curve obtained this way for *each DCT frequency* is called the *temporal* estimation and can be used as a ground truth (GT) to compare with the spatial estimation. Even if a noise model for JPEG images has never been proposed in the literature, it is therefore possible to obtain reliable empirical GT curves for JPEG images. To obtain them, it suffices to JPEG-encode each image of the snapshot with the same quality parameter, and to apply the above described procedure.

The objective of this section is to verify that the spatial standard deviation (STD) measured at any frequency $[i, j] \in$

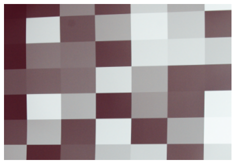


Fig. 1. One snapshot of the calibration pattern used to measure the temporal oscillation of the pixel intensities. The temporal STD is the GT of the spatial estimation.

$[0, w-1]^2, [i, j] \neq [0, 0]$ using the algorithm in Sec. 2 coincides with the STD of the temporal series measured only at that frequency for all intensities. To build the temporal STD noise curve we used 100 snapshots of the same calibration pattern (see Fig. 1), for both raw and JPEG-encoded images. In principle, any image might be used to get the temporal STD of the noise, but it is preferable to use an object with large flat regions of different gray levels, in order to avoid the effect of textures in the temporal estimation.

In the sequel, we compare the results of the spatial estimation to the GT, for both raw and JPEG-encoded images taken with a Canon EOS 30D camera with exposure time $t = 1/30s$, ISO speed 1600, and blocks of $w \times w$ DCT coefficients with $w = 4$. (This block size was chosen as particularly adapted for JPEG denoising algorithms). Fig. 4 compares the temporal and the spatial STDs for raw images and Fig. 5 shows the same for JPEG-encoded images with compression factor $Q = 92$. Only coefficients $[1, 1]$, $[2, 2]$, and $[3, 3]$ are shown, but equivalent results are obtained with all 15 coefficients. The average of the estimations along all coefficients $[i, j] \in [0, w-1]^2, [i, j] \neq [0, 0]$ is also given. The comparison results are similar for 8×8 blocks, although we do not have space to show them in this short paper.

Despite small oscillation in the spatial estimation, there is an accurate match between both the spatial and temporal estimations in the case of raw and JPEG images. It can be concluded that the method is able to estimate reliably signal-dependent noise at each frequency.

Note that this test was performed with snapshots of the calibration pattern (see Fig. 1), which is not textured and contains large flat areas whose spatial variations are caused mainly by the noise. Thus, the final validation must use real natural images compressed with JPEG. Since a proper noise model for JPEG encoding has not been already described, a visual comparison of the quality of the images before and after denoising using the frequency-by-frequency estimation given by the proposed method is needed. This comparison is performed in Sec. 3.1.

3.1. Denoising results

Old photographs are particularly adapted to evaluate our noise estimation method. Indeed for such images that involve two different successive acquisition systems, one chemical and one digital, there is no way to evaluate a parametric noise model. And there is of course no ground truth. Yet the visual inspection of the noise gives a very good hint at its independence from the (recovered) signal. To denoise JPEG digital images of old photographs, we used a modified version of the NL-Bayes algorithm [9] using the noise DCT coefficients estimated by our algorithm in Sec. 2. Of course, other block-based denoisers [12, 20, 21, 22] may be used instead. The details of the denoiser can be found in [8].

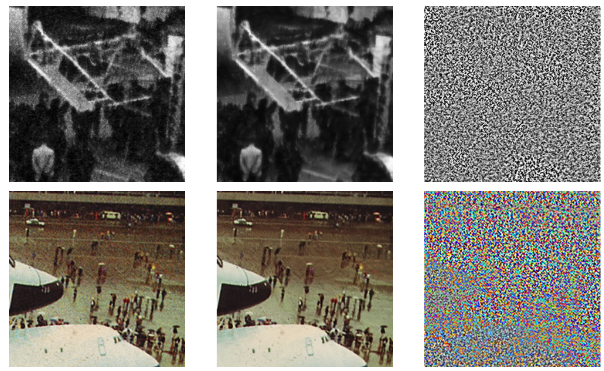
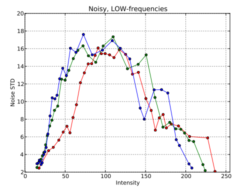
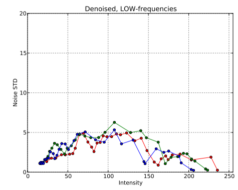


Fig. 2. Denoising results of real images with unknown noise model encoded with JPEG with unknown quality factor parameters. Left: detail of the noisy image. Middle: detail of the denoised image. Right: difference image (removed noise). The noise estimation method is validated, since the colored noise is removed (see the color spots in the difference image and its random geometry at zones with the same intensity) whereas signal detail is kept.

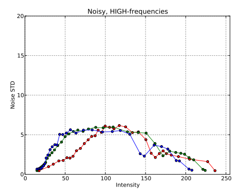
Fig. 3 presents the noise curves corresponding to the low and high frequencies of the JPEG image whose detail is shown at the bottom of Fig. 2 using DCT blocks of 4×4 coefficients. A coefficient at frequency $[i, j] \in [0, 3]^2$ is assumed to belong to a "low-frequency" if $i + j \leq 2$ and to a "high-frequency" otherwise. The image is a scan from a 1983 postcard that was afterwards compressed with JPEG. It suffered not only the degradation of JPEG lossy compression, but also other unknown digital and chemical acquisition distortions. We show the mean of the noise curves from the low-frequencies before (a) and after (b) denoising, where it can be observed that most of the noise remains at the low-frequencies of the image and that is strongly reduced after denoising. We also show the means for the high-frequencies before (c) and after (d) denoising. Since JPEG quantizes the value of the DCT coefficients at the high-frequencies (thus cancelling most of them), the noise is clearly lower that what is observed at the low-frequencies, but nevertheless the noise could also be removed.



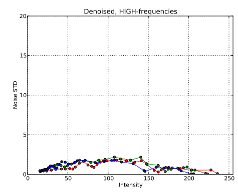
Low-freqs., noisy (a)



Low-freqs., denoised (b)

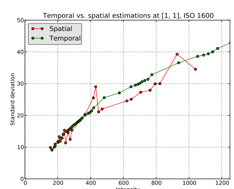


High-freqs., noisy (c)

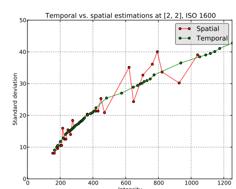


High-freqs, denoised (d)

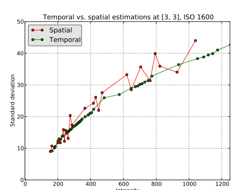
Fig. 3. Noise curves corresponding to the low and high frequencies of the JPEG image whose detail is shown at the bottom of Fig. 2 using DCT blocks of 4×4 coefficients. (a) and (b): mean noise curve at the low-frequencies before (a) and after (b) denoising. (c) and (d): mean noise curves at the high-frequencies before (c) and after (d) denoising. Most of the noise is at the low-frequencies. The color of each of the curves corresponds to each color channel of the image (red, green, and blue).



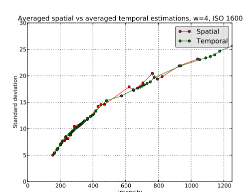
Frequency $[1, 1]$



Frequency $[2, 2]$

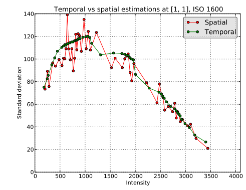


Frequency $[3, 3]$

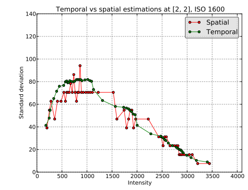


Average comparison

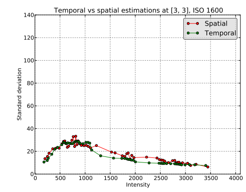
Fig. 4. Comparison of the temporal (GT, in green) and spatial STD (in red) for the Canon EOS 30D in raw images for ISO speed 1600 using blocks of 4×4 DCT coefficients. The temporal and spatial STD match despite some oscillation in the spatial estimation. The curve at the bottom right is the comparison between the averaged mean temporal STDs and the averaged mean spatial STDs (along all frequencies except DC), showing that in average both estimations match accurately.



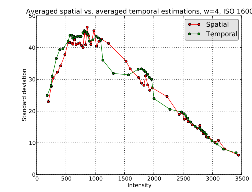
Frequency $[1, 1]$



Frequency $[2, 2]$



Frequency $[3, 3]$



Average comparison

Fig. 5. Comparison of the temporal (GT, in green) and spatial STD (in red) for the Canon EOS 30D in JPEG-encoded images with quality factor $Q = 92$ for ISO speed 1600 using blocks of 4×4 DCT coefficients. The curve at the bottom right is the comparison between the averaged temporal STDs and the averaged mean spatial STDs (along all frequencies except DC), showing that in average both estimations match.

4. CONCLUSION

We presented a non-parametric noise estimation method for intensity-frequency dependent noise. It can be applied to images where the noise model is not available [2], as in the case of JPEG images. Instead of assuming a prefixed noise model and then obtaining the parameters that control it (as parametric models do), our non-parametric method obtains at the same time both the noise model for the patches and its characteristics, that is, the noise estimation according to the discovered model. The method was validated by showing that the STD obtained at the temporal series (the GT) coincides with the spatial STD given by the proposed algorithm, for both raw and JPEG images. The denoising results show that indeed the noise estimator is able to give an accurate estimation, since low frequency noise is removed and most of the fine details are kept. Our next endeavour would be to include an impulse noise estimator to the non-parametric noise estimation model. Old photographs can indeed present this sort of noise. Nevertheless our estimation algorithm cannot be applied to *any* noisy image. For example, it does not apply if the noise is space dependent (and not only signal dependent), as can be observed in some synthetic images.

Acknowledgement: work partially supported by the Centre National d'Etudes Spatiales (CNES, MISS Project), the European Research Council (Advanced Grant Twelve Labours), the Office of Naval Research (Grant N00014-97-1-0839), Direction Générale de l'Armement (DGA), Fondation Mathématique Jacques Hadamard, and Agence Nationale de la Recherche (Stereo project).

5. REFERENCES

[1] A. Foi, M. Trimeche, V. Katkovnik, and K. Egiazarian, "Practical Poissonian-Gaussian noise modeling and fitting for single-image raw-data," *IEEE Transactions On Image Processing*, vol. 17, no. 10, pp. 1737–1754, 2008.

[2] R. A. Boie and I. J. Cox, "An analysis of camera noise," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 6, pp. 671–674, 1992.

[3] J. F. Hamilton Jr. and J. E. Adams Jr., "Adaptive color plan interpolation in single sensor color electronic camera," May 13 1997, US Patent 5,629,734.

[4] A. Buades, B. Coll, J. M. Morel, and C. Sbert, "Self-similarity driven demosaicking," *Image Processing On Line*, vol. 2011, 2011.

[5] N. N. Ponomarenko, V. V. Lukin, K. O. Egiazarian, and J. T. Astola, "A method for blind estimation of spatially correlated noise characteristics," in *IS&T/SPIE Electronic Imaging*. International Society for Optics and Photonics, 2010, pp. 753208–753208.

[6] Anil Kokaram, Damien Kelly, Hugh Denman, and Andrew Crawford, "Measuring noise correlation for improved video denoising," in *Image Processing, 2012. ICIP'12. 2012 International Conference on*. IEEE, 2012, pp. 1201–1204.

[7] G. K. Wallace, "The JPEG still picture compression standard," *Communications of the ACM*, vol. 34, no. 4, pp. 30–44, 1991.

[8] M. Lebrun, M. Colom, and J.M. Morel, "The Noise Clinic : A universal blind denoising algorithm," *Image Processing, 2014. ICIP'14. 2014 International Conference on*, 2014.

[9] M. Lebrun, A. Buades, and J.M. Morel, "Implementation of the "non-local Bayes" (NL-Bayes) image denoising algorithm," *Image Processing On Line*, vol. 2013, pp. 1–42, 2013.

[10] R. Neelamani, R. De Queiroz, Z. Fan, S. Dash, and R. G. Baraniuk, "JPEG compression history estimation for color images," *Image Processing, IEEE Transactions on*, vol. 15, no. 6, pp. 1365–1378, 2006.

[11] C. Liu, W. Freeman, R. Szeliski, and S. Kang, "Automatic estimation and removal of noise from a single image," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 2, pp. 299–314, February 2008.

[12] J. Portilla, "Full blind denoising through noise covariance estimation using Gaussian scale mixtures in the wavelet domain," *Image Processing, 2004. ICIP'04. 2004 International Conference on*, vol. 2, pp. 1217–1220, 2004.

[13] S.I. Olsen, "Estimation of noise in images: An evaluation," *CVGIP: Graphical Models and Image Processing*, vol. 55, no. 4, pp. 319–323, 1993.

[14] S. Pyatykh, J. Hesser, and L. Zheng, "Image noise level estimation by principal component analysis," *IEEE Transactions on Image Processing*, 2012.

[15] J. Immerkaer, "Fast noise variance estimation," *Computer Vision and Image Understanding*, vol. 64, no. 2, pp. 300–302, 1996.

[16] K. Rank, M. Lendl, and R. Unbehauen, "Estimation of image noise variance," in *Vision, Image and Signal Processing, IEE Proceedings-*. IET, 1999, vol. 146, pp. 80–84.

[17] P. Meer, J. M. Jolion, and A. Rosenfeld, "A fast parallel algorithm for blind estimation of noise variance," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 12, no. 2, pp. 216–223, 1990.

[18] M. Lebrun, M. Colom, A. Buades, and J.M. Morel, "Secrets of image denoising cuisine," *Acta Numerica*, vol. 21, pp. 475–576, 2012.

[19] M.L. Uss, B. Vozel, V. V. Lukin, and K. Chehdi, "Image informative maps for component-wise estimating parameters of signal-dependent noise," *Journal of Electronic Imaging*, vol. 22, no. 1, pp. 013019–013019, 2013.

[20] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3D transform-domain collaborative filtering," *IEEE Transactions on Image Processing*, vol. 16, no. 82, pp. 3736–3745, 2007.

[21] A. Buades, B. Coll, and J.M. Morel, "A non local algorithm for image denoising," *IEEE Computer Vision and Pattern Recognition*, vol. 2, pp. 60–65, 2005.

[22] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *Signal Processing, IEEE Transactions on*, vol. 54, no. 11, pp. 4311–4322, 2006.