LOCAL DENOISING BASED ON CURVATURE SMOOTHING CAN VISUALLY OUTPERFORM NON-LOCAL METHODS ON PHOTOGRAPHS WITH ACTUAL NOISE

Gabriela Ghimpeteanu[*], *David Kane*[*], *Thomas Batard*[*], *Stacey Levine*[†] *and Marcelo Bertalmío*[*]

[*] Dept. de Tecnologies de la Informació i les Comunicacions, Universitat Pompeu Fabra, Spain
[†] Department of Mathematics and Computer Science, Duquesne University, USA

ABSTRACT

We propose a fast, local denoising method where the Euclidean curvature of the noisy image is approximated in a regularizing manner and a clean image is reconstructed from this smoothed curvature. User preference tests show that when denoising real photographs with actual noise our method produces results with the same visual quality as the more sophisticated, non-local algorithms Non-local Means and BM3D, but at a fraction of their computational cost. These tests also highlight the limitations of objective image quality metrics like PSNR and SSIM, which correlate poorly with user preference.

Index Terms— Image denoising, camera pipeline, image quality metrics, perceptual metrics, psychophysical experiments.

1. INTRODUCTION

Noise in photographs is unavoidable and causes a loss of visual quality and dynamic range. All cameras perform some form of denoising, but this is limited by the computational power of the camera. Techniques such as spectral analysis, variational methods and PDEs have all been successfully used for denoising [1] and while some are fast to compute, they typically do not perform as well as newer, non-local, but computationally intensive, patch based methods [2–6]. These latter algorithms can produce high quality results, but their computational complexity makes in-camera implementation impractical.

Another line of denoising algorithms have suggested denoising some geometric feature of an image, instead of denoising the image directly [7–10]. The authors in [7] proposed the following approach for removing noise from (grayscale) image $I_0 = a + n$, where a is a 'clean' image and n is Gaussian noise of mean zero: given a denoising method $\mathcal{F}$, instead of applying it directly on I_0 it's better to use it to denoise the curvature of level lines of I_0, $\kappa(I_0) = div(\nabla I_0/|\nabla I_0|)$, resulting in $\kappa_d(I_0)$, and then reconstruct a 'clean' image whose curvature matches $\kappa_d(I_0)$. In [7] it is shown that this framework improves on a variety of denoising methods $\mathcal{F}$ including

local and patch-based techniques (e.g. [3, 4, 11]), but the evaluation is only performed in terms of objective measures such as PSNR. Several works have demonstrated that while state of the art patch-based methods are nearing optimality with respect to mean squared error, and thus PSNR, there is some room for improvement [12–14]. Still, while PSNR accounts for the total pixel error across the image, it is widely accepted that it is not a good indicator of perceived image quality. And while SSIM [15] aims to estimate perceived errors, it still does not consistently match human preference [16]. Thus the study of image quality metrics is an active research area, with the main challenge being the lack of accurate models of visual perception and subjective quality.

In this work we propose a local, low complexity denoising method. From [7] we take the notion of reconstructing the denoised image from the denoised curvature, but instead of applying a denoising method $\mathcal{F}$ to the original (noisy) curvature we directly compute a smooth approximation to the curvature for each color channel. A subjective evaluation was performed by asking subjects to choose their preferred image from three on real photographs with visible noise. We find that our method yields solutions that visually match (on average) or surpass (for some images) results from two more sophisticated algorithms (Non-local Means [3] and BM3D [4]), but at a fraction of their computational cost. This suggests the possibility of incorporating the proposed method into the camera processing pipeline. We also conduct a subjective evaluation using clean images with three levels of added synthetic (Gaussian) noise. We test whether the metrics PSNR and SSIM can predict the subjective data. We find that SSIM can reasonably approximate the averaged preference of subjects as a function of the noise level, but that neither metric can reliably predict the preferred denoising algorithm on an image-by-image basis. This observation highlights the limitations of the quantitative metrics PSNR and SSIM when evaluating denoising results.

2. A NEW LOCAL DENOISING METHOD

2.1. Foundations of the model

The origin of the denoising method that we propose in this work is related to the following energy model, introduced

in [17]:

$$E(I) := \int_{\Omega} \|\nabla I\| - \nabla I \cdot \theta(I_0) \; d\Omega \tag{1}$$

where $\theta(I_0)$ is a unit vector field indicating the normals to the level lines of a grey-level image $I_0 : \Omega \subset \mathbb{R}^2 \longrightarrow \mathbb{R}$. In [18] the authors propose a variational method to fuse a short-exposure image, dark and noisy but sharp, with a long-exposure image, bright and clean but blurry. Their technique involves minimizing an energy functional with two terms, where one term locally matches the color histogram of the result to that of the long-exposure image; the other term is in fact the energy (1), with the purpose of matching the level-lines of the solution with the level lines of the short-exposure image I_0. The numerical solution is achieved through a gradient descent algorithm associated to a differentiable approximation of the whole energy. In particular, the gradient descent corresponding to the term (1) is

$$I_t = \kappa_{\epsilon_1}(I) - \kappa_{\epsilon_2}(I_0) \tag{2}$$

$$= \nabla \cdot \left(\frac{\nabla I}{\sqrt{\|\nabla I\|^2 + \epsilon_1}} \right) - \nabla \cdot \left(\frac{\nabla I_0}{\sqrt{\|\nabla I_0\|^2 + \epsilon_2}} \right) \tag{3}$$

where $\kappa_{\epsilon_1}(I)$ is the Euclidean curvature of the level lines of I, $\kappa_{\epsilon_2}(I_0)$ is the Euclidean curvature of the level lines of I_0, and $\epsilon_1, \epsilon_2 > 0$ are introduced to avoid division by zero, a common practice in this sort of scheme [11]. The authors in [18] observed that the term $\kappa_{\epsilon_1}(I) - \kappa_{\epsilon_2}(I_0)$ may have a noticeable and good-quality denoising effect if ϵ_1 is kept very small but ϵ_2 is large: the final fusion result has substantially less noise than the short-exposure image, and the edges are still kept sharp. As the value of ϵ_2 increases, the denoising is more pronounced.

In [7] it is shown that, given any denoising method $\mathcal{F}$, rather than applying $\mathcal{F}$ directly to a noisy image I_0 it's better (in terms of objective metrics PSNR and SSIM [15]) to apply $\mathcal{F}$ to denoise the curvature $\kappa(I_0)$, and then reconstruct a clean image by running to steady state the following evolution equation

$$I_t = \kappa_\epsilon(I) - \mathcal{F}(\kappa(I_0)) + \lambda(I - I_0), \tag{4}$$

where $0 < \epsilon << 1$ such that κ_ϵ approximates $\kappa(I)$, and $\lambda > 0$ is a Lagrange multiplier, based on the noise variance in a Gaussian noise scenario, that automatically stops the evolution.

2.2. The proposed approach

Our aim is to design a computationally fast denoising method that produces good visual quality results on photos with real noise that is generated during the acquisition process of the camera. From [18] we take the idea of using a large value ϵ_2 to compute the curvature of the noisy data I_0, and use $\kappa_{\epsilon_2}(I_0)$ as a quick and easy way to obtain a regularized version of

the curvature. We follow [7] in reconstructing a denoised image from its denoised curvature, but we replace $\mathcal{F}(\kappa(I_0))$ with $\kappa_{\epsilon_2}(I_0)$ and remove from Eq. (4) the term $\lambda(I - I_0)$ since it privileges noise with a Gaussian distribution, whereas actual noise in photographs isn't well described by an additive Gaussian model. Our proposed method is then the following. For each color channel (R,G,B) we take the original noisy data I_0, compute its regularized curvature $\kappa_{\epsilon_2}(I_0)$, and starting from $I^0 = I_0$ iterate for some number N steps of this equation:

$$I^{n+1} = I^n + \Delta t \left[\nabla^- \cdot \left(\frac{\nabla^+ I^n}{\sqrt{\|\nabla^+ I^n\|^2 + \epsilon_1}} \right) - \kappa_{\epsilon_2}(I_0) \right], \tag{5}$$

where ∇^+ (resp. ∇^-) denotes the forward (resp. backward) spatial difference operator. This is clearly a local method since the curvature at each pixel location is estimated using just a 3×3 stencil around it. The parameter ϵ_1 is fixed and is chosen very small in order to approximate $\kappa(I)$, and the time step Δt and the number of iterations N are also fixed. The values we used are: $\epsilon_1 = 10^{-6}$, $\Delta t = 0.002$ and $N = 30$. Therefore, our denoising approach has only one parameter: the regularizing value ϵ_2. We will specify the way it is determined in the next Section.

3. EXPERIMENTS

We compare the proposed curvature smoothing (CS) algorithm against two established, non-local, patch based algorithms: NLM [3] and BM3D [4]. The comparison is performed on two image databases: images from the Kodak database [19] with added Gaussian noise and photographs taken by us with real noise (see Fig.1). We evaluate the three denoising methods using subjective testing, and in the case of the Kodak images where we have an established ground truth we also perform an evaluation using the metrics PSNR and SSIM.

The subjective evaluation involved 17 participants (all with normal or corrected to normal vision). Subjects sat in a well-lit office environment at approximately 64 cm from the display and were presented with four versions of an image: the original at the top, and the three denoising results (CS, NLM and BM3D) in some random order at the bottom. The observer was asked to look at the original image and then indicate which of the three denoised images provided they preferred.

In terms of computational cost, our method, being local, has linear complexity (as we mentioned, curvature is computed on 3×3 stencils) whereas NLM and BM3D involve comparing patches over an area so their complexity is quite higher; in practice, NLM takes 7 times longer and BM3D 10 times longer than CS, even though our implementation is not optimized while for NLM and BM3D we use the optimized code from [20] and [21], respectively.

Table 1: Optimized parameter value ϵ_2 as a function of σ.

σ	3	6	9
ϵ_2	0.00032	0.003	0.00608

Fig. 1: Test images. Top two rows: crops from Kodak images. Rows 3 to 7: crops from photographs with real noise.

3.1. Kodak database images

We randomly picked three images from the Kodak database ("kodim1", "kodim3" and "kodim13") and added three levels of Gaussian noise $\sigma = 3, 6, 9$; these values may seem low, but they correspond to normal noise present in properly exposed pictures, as the real examples in the next subsection will attest. The implementations of denoising algorithms NLM [20] and BM3D [21] take the value of σ as input and require no further configuration. For our method we find the value of ϵ_2 via a subjective methodology; subjects are asked to adjust ϵ_2 via key presses until they find the most pleasing result. For each noise level, we average across subjects and images to obtain a single value of ϵ_2 as shown in Table 1.

Having the ground truth we can compare the values of PSNR and SSIM for each denoising method. We also perform a user preference test (with the procedure described above). We take crops from the full size images to allow the simultaneous presentation of the 4 images at their native resolution (thus avoiding resizing). The results are summarised in Fig.2. We add 95% error bars (estimated via bootstrapping), thus significance can be inferred from visual inspection. We find that the SSIM metric provides a reasonable approximation of the subjective results, correctly predicting that the differences between the algorithms should be small at low noise levels and that the proposed CS method performs poorly at high noise levels. Despite this, both the PSNR and SSIM metrics are poor at predicting which algorithm is preferred on an image-by-image basis. To assess the metric performance we first compute an upper bound by randomly splitting the subjective data into two subject groups (A and B). For each image we then compute a percentage correct score, the score is 100% if the order is

entirely correct, 33% for only getting the order of one correct or 0% for a complete failure. We find that on average group A predicts the data from group B 64% of the time. In contrast, both the SSIM and PSNR achieve a score of less than 46%. Note the baseline score is 33%.

3.2. Real noise images

Now we compare user preference on the results from CS, NLM and BM3D over 30 images cropped from 5 real noise photos. Since the noise standard deviation is not known, we find the values for σ (the parameter for NLM and BM3D) and ϵ_2 (parameter for CS) through user tests in the same manner as above. The results of the psychophysical experiment are shown in Fig.4. On the left hand side of the figure we plot the user preference average across all thirty images. The results show that the performance of the model is statistically identical for the three algorithms. The remaining graphs show the results for five individual images as denoted above each graph. Notice how some of these images are quite noisy. For the same image, the optimal value σ chosen by user tests differs for NLM (values up to 5.10) and BM3D (values up to 8.49). If the noise were Gaussian, according to PSNR and SSIM there should be a clear difference in performance of algorithms. Despite the image-dependent variability in the preferred algorithms, the overall results show no one algorithm wins overall. Again, this contradicts the result that *for additive Gaussian noise*, PSNR and SSIM predict that the visual quality of BM3D is always clearly superior to that of NLM, which is always supposed to outperform CS. Fig.5 shows individual examples where CS does well in terms of visual quality.

4. CONCLUSION

We have introduced a local denoising method based on reconstructing a clean image from a smoothed version of the Euclidean curvature of the original noisy input. The method was compared with the non-local, computationally much more intensive algorithms of NLM and BM3D. Our results indicate that in terms of user preference, our proposed method, NLM and BM3D all have the same average performance on real-noise images, that PSNR and SSIM do not correlate with user preference at low, but still quite noticeable, noise levels common in many properly-exposed photographs, and finally, that the outcomes of user preference tests on the results of denoising algorithms applied to images with added Gaussian noise can not be extrapolated to the real case of noisy images where the noise is due to the acquisition process. We are currently building a larger image database on which to perform tests, considering also images coming from digital cinema cameras where the bit depth is higher and gamma correction is replaced with logarithmic encoding, factors which may affect the performance of our method in comparison with the non-local approaches.
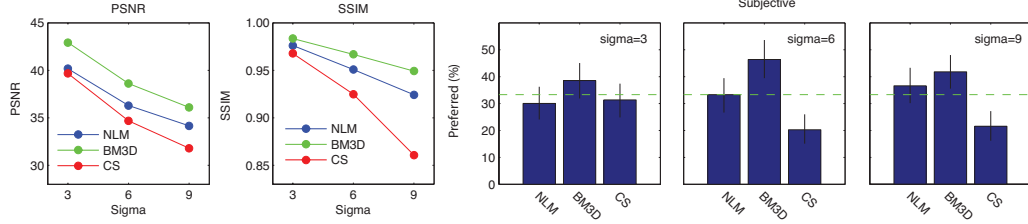
Fig. 2: From left to right: Average PSNR and SSIM computed for 3 images from the Kodak data base, and results of psychophysical experiment for comparing our proposed local denoising method to BM3D and NLM.

Fig. 3: Visual comparison for one test crop from image "kodim3" and user preferences.
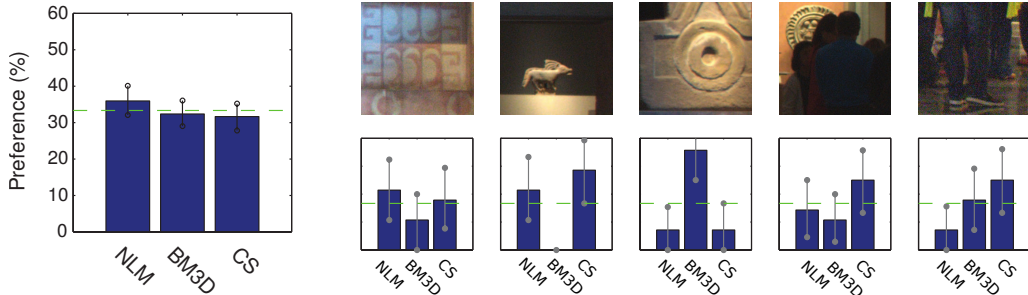

Fig. 4: Left: Average visual comparison results for 30 real-noise images, for comparing our denoising method to BM3D and NLM. Right: Image example crops and user preferences. Notice the case of the second image example, where the noise level is high and users may prefer CS: user estimation of $\sigma = 4.73$ for NLM and $\sigma = 7.15$ for BM3D.
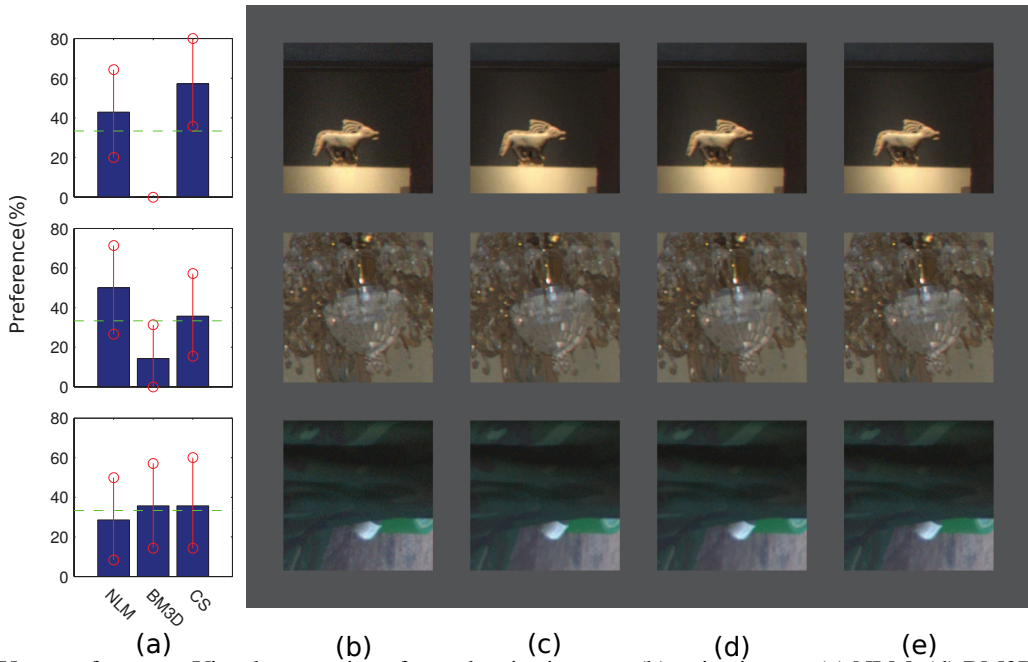

Fig. 5: (a) User preferences. Visual comparison for real-noise images: (b) noisy image, (c) NLM, (d) BM3D and (e) CS.

5. REFERENCES

[1] G. Aubert and P. Kornprobst, *Mathematical problems in image processing*, vol. 147 of *Applied Mathematical Sciences*, Springer-Verlag, New York, 2002, Partial differential equations and the calculus of variations, With a foreword by Olivier Faugeras.

[2] S.P. Awate and R. T. Whitaker, "Unsupervised, information-theoretic, adaptive image filtering for image restoration," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 3, pp. 364–376, 2006.

[3] A. Buades, B. Coll, and J. M. Morel, "A review of image denoising algorithms, with a new one," *Multiscale Model. Simul.*, vol. 4, no. 2, pp. 490–530 (electronic), 2005.

[4] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-d transform-domain collaborative filtering," *Image Processing, IEEE Transactions on*, vol. 16, no. 8, pp. 2080–2095, 2007.

[5] M. Elad, *Sparse and redundant representations*, Springer, New York, 2010, From theory to applications in signal and image processing, With a foreword by Alfred M. Bruckstein.

[6] M. Lebrun, A. Buades, and J. M. Morel, "A nonlocal bayesian image denoising algorithm," *SIAM Journal on Imaging Sciences*, vol. 6, no. 3, pp. 1665–1688, 2013.

[7] M. Bertalmío and S. Levine, "Denoising an image by denoising its curvature image," *SIAM J. Imaging Sci.*, vol. 7, no. 1, pp. 187–211, 2014.

[8] G. Ghimpeteanu, T. Batard, M. Bertalmio, and S. Levine, "A decomposition framework for image denoising algorithms," *Image Processing, IEEE Transactions on*, vol. 25, no. 1, pp. 388–399, Jan 2016.

[9] M. Lysaker, S. Osher, and X.-C. Tai, "Noise removal using smoothed normals and surface fitting," *IEEE Trans. Image Process.*, vol. 13, no. 10, pp. 1345–1357, 2004.

[10] S. Osher, M. Burger, D. Goldfarb, J. Xu, and W. Yin, "An iterative regularization method for total variation-based image restoration," *Multiscale Model. Simul.*, vol. 4, no. 2, pp. 460–489 (electronic), 2005.

[11] L.I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D: Nonlinear Phenomena*, vol. 60, no. 1-4, pp. 259–268, 1992.

[12] P. Chatterjee and P. Milanfar, "Is denoising dead?," *IEEE Transactions on Image Processing*, vol. 19, no. 4, pp. 895–911, 2010.

[13] A. Levin and B. Nadler, "Natural image denoising: Optimality and inherent bounds," in *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*. IEEE, 2011, pp. 2833–2840.

[14] A. Levin, B. Nadler, F. Durand, and W.T. Freeman, "Patch complexity, finite pixel correlations and optimal denoising," Tech. Rep., MIT - Computer Science and Artificial Intelligence Laboratory, 2012.

[15] Z. Wang and A.C. Bovik, "A universal image quality index," *Signal Processing Letters, IEEE*, vol. 9, no. 3, pp. 81–84, 2002.

[16] J.-F. Pambrun and R. Noumeir, "Limitations of the ssim quality metric in the context of diagnostic imaging.," in *ICIP*. 2015, pp. 2960–2963, IEEE.

[17] C. Ballester, M. Bertalmío, V. Caselles, G. Sapiro, and J. Verdera, "Filling-in by joint interpolation of vector fields and gray levels," *IEEE Trans. Image Process.*, vol. 10, no. 8, pp. 1200–1211, 2001.

[18] M. Bertalmío and S. Levine, "Variational approach for the fusion of exposure bracketed pairs," *IEEE Trans. Image Process.*, vol. 22, no. 2, pp. 712–723, 2013.

[19] "Kodak," http://r0k.us/graphics/kodak/, [Online; accessed January-2016].

[20] A. Buades, B. Coll, and J. M. Morel, "Non-local means denoising," *Image Processing Online*, vol. 1, 2011.

[21] A. Buades, B. Coll, and J. M. Morel, "An analysis and implementation of the bm3d image denoising method," *Image Processing Online*, vol. 2, pp. 175–213, 2011.