Nonparametric noise estimation method for raw images

Miguel Colom,[1,2,*] Antoni Buades,[1,2] and Jean-Michel Morel[1,2]

[1]Departament de Ciències Matemàtiques i Informàtica, Universitat de les Illes Balears,
Palma de Mallorca, Illes Balears, Spain
[2]Centre de Mathématiques et de Leurs Applications (CMLA), École Normale Supérieure de Cachan, Cachan, France
*Corresponding author: miguel.colom@cmla.ens-cachan.fr

Optimal denoising works at best on raw images (the image formed at the output of the focal plane, at the CCD or CMOS detector), which display a white signal-dependent noise. The noise model of the raw image is characterized by a function that given the intensity of a pixel in the noisy image returns the corresponding standard deviation; the plot of this function is the *noise curve*. This paper develops a nonparametric approach estimating the noise curve directly from a single raw image. An extensive cross-validation procedure is described to compare this new method and with state-of-the-art parametric methods and with laboratory calibration methods giving a reliable ground truth, even for nonlinear detectors.   © 2014 Optical Society of America

   OCIS codes: (040.1520) CCD, charge-coupled device; (100.2960) Image analysis; (110.4280) Noise in imaging systems; (040.0040) Detectors; (040.3780) Low light level; (100.2980) Image enhancement.

1. INTRODUCTION

Most denoising methods assume that the noise in the image is additive, homoscedastic, white, and Gaussian. *Homoscedastic* means that the variance of the Gaussian noise is fixed and does not depend on the pixel position or value. By "white" noise, we mean that the noise pixel values are independent. We shall retain this terminology throughout the paper. The homoscedasticity assumption is not realistic. The photon emission by a body follows a Poisson distribution, which can be approximated by a Gaussian distribution when the number of photons is large enough. But the variance of this Gaussian is signal dependent. In the Poisson model [1–7], an image value $\tilde{\mathbf{U}}(x, y)$ at pixel (x, y) is a Poisson variable with variance and mean equal to $\mathbf{U}(x, y)$, where $\mathbf{U}$ is the ideal noise-free image. The Poisson noise has therefore a standard deviation (STD) equal to $[\mathbf{U}(x, y)]^{1/2}$. Thus, an ideal raw image is a white Poisson noise whose mean at each pixel is the noiseless value. Note that this is related to the quantum nature of light and the probability of emitting a photon, independently of the technology used at the CFA (CCD, CMOS). This Poisson noise adds up to a thermal noise and to an electronic noise, which are approximately additive and white, making the final noise model not necessarily Poisson distributed, but still white and signal dependent.

Noise estimation is a necessary preliminary step for most image processing and computer vision algorithms [8]. Nevertheless, several other denoising methods propose to deal directly with Poisson noise. Wavelet-based denoising methods [9,10] propose to adapt the transform threshold to the local noise level of the Poisson process. Lefkimmiatis *et al.* [11] have explored a Bayesian approach, and Deledalle *et al.* [12] have adapted the nonlocal means algorithm [13] to Poisson noise. These papers assume that no variance stabilizing transform (VST) transforming the signal-dependent noise into a nearly homoscedastic noise is accurate enough to transform the Poisson noise into homoscedastic noise. The advantage of VSTs is that they permit the application of a classic denoising algorithm. The VST associated with Poisson noise is often called Anscombe transform [14], but one can attach a VST to any signal-dependent noise model [8]. As a matter of fact, papers on the Anscombe transform [15] (for low count Poisson noise) and [16] (for Rician noise) argue that, when combined with suitable forward and inverse variance stabilizing transformations, algorithms designed for signal independent Gaussian noise work just as well as *ad hoc* algorithms for Poisson noise models. These considerations confirm the importance of estimating as accurately as possible the noise curves of raw images, since their accurate knowledge is required to compute the VST. In most CCDs and CMOS detectors, the variance of the noise at a pixel is approximated (assuming that all detectors at the CFA are equivalent and thus neglecting the fixed pattern noise) by a simple linear model $\sigma^2 = A + B\mathbf{U}$, where $\mathbf{U}$ is the expectation of the intensity of this pixel in the noisy image. This model is valid under the assumption mentioned above of a combination of a Poisson with a thermal noise. Yet, this assumption holds only if the signal is not saturated and the photon count large enough. At the darkest pixels, the Poisson distribution of the noise cannot be approximated by a Gaussian and it becomes a *shot* noise. In short, the noise variance does not necessarily follow the linear model in the darkest and brightest image regions. An accurate estimation of the noise at the darkest zones is crucial since subsequent processes in the camera chain (specially, the gamma correction step) are designed to increase the dynamics in the dark zones. If the noise is not removed at the raw image stage, it might end up really augmented at the final stage.

Parametric noise estimation methods try to obtain the parameters that control a noise model (for example, the A and B parameters of the linear model). Yet, to get a realistic estimation, they have to take into account the effect of the

saturation in the darkest and brightest pixels in the final noise curve. To validate the estimation of a noise estimation method, its noise curve must be compared to a ground-truth curve. Such a ground truth for a particular camera and settings can be obtained by taking a series of photographs of a pattern, which is mostly flat and contains a wide range of gray levels. The temporal variation of the gray level at a given pixel gives an estimate of the noise STD associated with this gray level. However, the series of photographs must be taken under controlled conditions, to ensure that any variation of the intensity of a pixel can be only explained by the noise. In short, it is a heavy procedure (that is, it requires constant lighting, a camera stabilizer to fix its position, and isolation from any kind of an electromagnetic source that may introduce electronic noise into the camera), which also needs access to the camera that took the photographs. It also requires *a priori* knowledge of the form of the camera noise model, which is not granted. This explains why the establishment of a method able to estimate automatically the noise model from a single snapshot is a valid question. Furthermore, if the method can be shown to be reliable even without any *a priori* model guess, its credibility will be somewhat augmented. In this paper, we show that it is indeed possible to use a nonparametric estimator to get an accurate noise curve from the noisy image itself, by measuring the variance locally with patch-based methods [17–22]. This eliminates the need for a lab calibration procedure. Indeed, the procedure described uses one or several photographs taken in arbitrary environment and yields a nonparametric noise model as good (for those images) as the one obtained by the heavier ground-truth procedure (laboratory calibration). We also examine the question of whether it is better to use a parametric or a nonparametric model when dealing with a single or a few photographs. Our conclusion is that the nonparametric method gives results comparable to the parametric method, but is somewhat less risky as it does not propagate local estimation errors caused by the presence of texture in the image.

Our plan follows from the above discussion. Since noise estimation is a well-known procedure for white homoscedastic noise, Section 2 will review the literature on white homoscedastic noise estimation and will point out competitive algorithms. Section 3 explains the procedure that should be followed to get a reliable nonparametric noise curve from a series of images, under controlled conditions. Section 4 discusses how homoscedastic white noise estimation algorithms can be adapted to estimate an arbitrary signal-dependent noise curve. Section 5 compares the root-mean-squared errors (RMSE) between the nonparametric ground truth, the STDs from the series of images, and two state-of-the art parametric methods. Finally, Section 6 presents the conclusions, which validate our proposed nonparametric method, but also the use of two state-of-the-art parametric methods.

2. STATE OF THE ART IN WHITE NOISE ESTIMATION

Many noise estimation methods share the following features, which can be summarized in two sentences:

• estimate noise in high frequencies, where noise dominates over signal;

• estimate noise in image regions with the least variation, typically the blocks with the smallest STDs.

Thus, these numerous methods [18,22–33] proceed roughly as follows:

• they start by applying some high-pass filter, which concentrates the image energy on its edges, while the noise remains spatially homogeneous;

• they compute the energy of many blocks extracted from this high-passed image;

• they estimate the STDs of the blocks;

• to avoid blocks whose STD is mostly explained by the underlying ideal image, a statistic robust to (many) outliers must be applied. The methods therefore prefer the flattest blocks, which belong to a (low) percentile of the STDs of all the blocks.

Note that the power spectral density of a natural image is not homogeneous. Most of the energy corresponding to its geometry is located at the low and medium frequencies [see Eq. (1), for example], whereas high-frequency coefficients bring little visual information (with the exception of the edges). Conversely, an image can be considered "highly textured" if the energy at the high-frequency coefficients is as high as the energy observed at edges. Thus, high-passing the image before estimating the noise spatially (or equivalently, estimating the noise only at the high-frequency discrete cosine transform (DCT) coefficients) is an initial step for many noise estimation algorithms. This enhances the contribution of the noise. Yet, avoiding edges and textures in the estimation remains necessary.

We shall limit ourselves to discussing the method acknowledged as the best estimator for homoscedastic noise in the review [8], the Ponomarenko *et al.* method [31], along with the two of the most competitive parametric methods for noise estimation in raw images [34,35]. We briefly describe these competitors in the next paragraphs. For a complete review on noise estimation methods, we refer the reader to [8] and [30].

A. Ponomarenko *et al.* Approach

The Ponomarenko *et al.* [31] method is an extension of the previous method [18], based on the analysis of the DCT coefficients. The orthonormal 2D DCT-II of each block is computed and denoted by $\mathbf{D}_m(i,j)$, where m is the index of the block, w is its size, and $0 \le i, j < w$ is the frequency pair associated with that coefficient. The algorithm uses two labels: "low frequency" and "medium/high frequencies" according to the value of the function

$$\delta(i,j) := \begin{cases} 1, & (i+j < T) \text{ and } (i+j \neq 0) \rightarrow \text{low freq;} \\ 0, & (i+j \ge T) \text{ or } (i+j = 0) \rightarrow \text{med./high freq.} \end{cases} \quad (1)$$

T is a given threshold. Note that this function does not label the mean of the block term $(i + j = 0)$ as a low frequency. For each block, an (empirical) variance associated to the low-frequency coefficients of the block m is defined as

$$\mathbf{V}_m^L := \frac{1}{\theta} \sum_{i=0}^{w-1} \sum_{j=0}^{w-1} [\mathbf{D}_m(i,j)]^2 \delta(i,j), \quad (2)$$

where

$$\theta = \sum_{i=0}^{w-1} \sum_{j=0}^{w-1} \delta(i,j) \qquad (3)$$

is the adequate normalization factor to get a mean. The set of transformed blocks is rewritten with respect to the corresponding value of $\mathbf{V}_m^L$ in ascending order with m. Given the list of sorted blocks $\{\mathbf{D}_{(m)}\}$, the noise variance estimate associated with the high-frequency coefficient at (i,j) is defined by

$$\mathbf{V}^H(i,j) := \frac{1}{K} \sum_{k=0}^{K-1} [\mathbf{D}_{(k)}(i,j)]^2, \qquad (4)$$

where $i + j \geq T$ and $K = \lfloor pM \rfloor$, $p < 1$ is the position of the p quantile in the list $\{\mathbf{D}_{(m)}\}_{m \in [0, M-1]}$. Note that this empirical variance estimate is made with the list of the K transformed blocks whose empirical variance as measured in their low frequencies is the lowest. It is understood that these blocks are likely to belong to flat or smooth zones. Thus their high frequencies are good candidates to estimate the noise with. In fact, if this is noise, the high and low frequencies are uncorrelated and one can assume that $\mathbf{V}^H(i,j)$ gives an accurate estimation of the noise variance. However, if the image is highly textured, those high-frequency coefficients might give a variance that is explained by the textures of the image and not by the noise.

The final noise estimation is given by the median of the variance estimates $\mathbf{V}^H(i,j)$,

$$\hat{\sigma} := [\text{median}_{i,j}(\mathbf{V}^H(i,j)|i+j \geq T)]^{1/2}. \qquad (5)$$

The authors of [31] proposed an adaptive strategy to find its best value. Nevertheless, we adopted a different strategy based on the ratio between the variances measured at the median and at given p quantile, as explained in Section 4. We now discuss two parametric methods that will be compared here.

B. Practical Poissonian–Gaussian Noise Modeling and Fitting for Single-Image Raw Data

Foi *et al.* proposed a simple parametric noise model [34] that takes into account the nonlinear response of the CDD due to the saturation of the signal and noise at the darkest and brightest pixels of the image. The model assumes the well-known normal approximation, for which the Poisson distribution $\mathcal{P}(\lambda)$ of the noise can in practice be approximated by the normal distribution: $\mathcal{P}(\lambda) \rightarrow \mathcal{N}(\mu = \lambda, \sigma = \lambda)$. The method has two stages. In the first step it estimates several pairs of intensity/STD that form a scatterplot. In the second step, the observed pairs are used to fit a global parametric model. Before applying these two steps, the image is preprocessed. First, the 2D-wavelet transform of the image is computed and the wavelet detail coefficients are stored. The 1D Daubechies wavelet and scaling functions are used to create the 2D kernels of the transform. The STD of the noise is obtained from the detail coefficients of the transformed signal. In order to be robust against edges, the image is segmented into level sets according to the intensity. Since the image to be segmented is noisy, the segmentation is done in a low-pass filtered version. With the selected regions of the image, the

intensity of each pair is obtained as the sample mean of the approximation wavelet coefficients and the estimated variance with the unbiased sample variance estimator. The last step of the method is to fit the A and B parameters of the linear model of the variance, for which a maximum-likelihood (ML) fitting is performed. However, since saturation makes the response of the CCD nonlinear, the method needs to modify the expectation and variance estimators to take saturation into account. The authors calculated the new estimators from the distribution of the nonsaturated signal and gave the explicit expression for the expectation and variance estimators under saturation. Finally, these new pairs are incorporated into the ML fitting in order to get the A and B parameters of the linear model despite the presence of saturation. The model is able to predict the nonlinear response of the CCD under saturation, giving explicitly the variance of the clipped noise for any intensity. Therefore, this method will be used as an example where parametric and nonparametric methods are cross-validated (Section 5).

C. Image Informative Maps for Component-Wise Estimating Parameters of Signal-Dependent Noise

In the paper [35] Uss *et al.* propose to adapt the use of disjoint informative maps [36] to estimate a parametric signal-dependent Poisson-like noise model. It discriminates between two kinds of nonoverlapping blocks [scanning window (SW)]: those which belong to textures [texture informative (TI)] and those that are suitable for noise estimation [noise informative (NI)]. To describe the textures of a given SW in the image, the 2D fractal Brownian motion (fBm) model is used, since the model is able to characterize a texture with few parameters. The roughness of the texture is obtained from the Hurst exponent in the fBm model. The estimation of the noise is obtained from a limited set of high-frequency coefficients of the DCT transform of the SWs that belong to the NI map. This idea was introduced in the Ponomarenko *et al.* method [31] and stated as the state-of-the-art technique for noise estimation [8]. The Cramér–Rao lower bound (CRLB) is used to decide if a SW belongs to the TI or NI maps, on the texture parameters and the noise STD of the SW. All the SWs in the image are sorted according to increasing CRLB and then compared to a threshold. The SWs below the threshold have the lowest CRLB and therefore belong to the NI map. The rest are assumed to be textures and assigned to the TI map. Since the criterion based on the CRLB relies on the (unknown) texture and noise parameters, the method begins with an initial guess for the NI and TI maps by fixing a noise STD and texture level to have an initial and rough CRLB criterion. Then, with the available CRLB criterion, better STD and texture levels are computed, allowing for an even better CRLB criterion. The refining loop is iterated until convergence is reached. To estimate signal-dependent noise, the set of SW is partitioned into disjoint intensity sets according to their mean intensity, and the method is applied separately to each set in order to get an (intensity–STD) pair. Therefore, this method is coherent with the claim we make in Section 4, which states that any block-based homoscedastic noise estimation method can be easily adapted to deal with signal-dependent noise, just by splitting the whole set of image blocks of the image into disjoint in intensity sets to apply then the homoscedastic version of the method to each of these sets.

For example, if the input image has size $N_x \times N_y$, there are $M = (N_x - w + 1)(N_x - w - 1)$ overlapping blocks, which may be distributed into a set of M/k bins, where each bin contains k image blocks/bin whose mean intensity is a part of the complete intensity range of the image.

3. NONPARAMETRIC NOISE GROUND-TRUTH CURVE

Parametric methods fix an *a priori* model for the noise. For example, at the output of the CCD a good approximation of the Poisson noise is to use the normal distribution approximation, at least when the number of incoming photons is large enough. Therefore, the variance of the noise is equal to the expectation. The noise at the output of the CCD detector is Poissonian, and therefore its variance is linear with the intensity. Also, thermal and electronic noise are added, and the noisy signal is amplified afterward. Thus, the variance of the noise can be modeled as a function of the intensity of the ideal (noise-free) image: $\sigma^2(\mathbf{U}) = A + B\mathbf{U}$. However, since the dynamics of the digital output from the CCD are limited, the darkest and brightest pixels of the image can get saturated because of the noise, which becomes clipped noise. Because of the saturation, the probability distribution of the noise is no longer a symmetric normal distribution, but a truncated version with different statistics. The variance of the truncated distribution does not coincide with that of the normal distribution. Therefore, any realistic parametric estimation method must take into account that under saturation the simpler linear model is no longer valid. Some methods [34] adapt their expectation and variance estimators in order to take into account the effect of the saturation before fitting the linear function, while others [35] try to fit with polynomials of higher order or transform the image in such a way that the linear model holds.

In any case, the parametric model has to be validated in order to ensure that the curve they provide is indeed a function that accurately relates the intensity of the ideal image with the STD of the added noise. To do it, the estimations of the parametric method must be compared with a ground-truth noise curve. For the construction of the ground truth the constraint of using just a single image is not needed. Indeed, it can be built from a series of snapshots of a calibration pattern taken from a camera in fixed position. The series must be taken under controlled laboratory conditions that ensure that the temperature and lighting remain constant. Ideally, any two images of the series should be exactly equal in absence of noise. Therefore, any variation between the images is only explained by stochastic light fluctuations (photon noise and shot noise) and the noise generated by the camera itself (dark noise, readout noise and electronic noise).

If $\tilde{\mathbf{U}}_i(x, y)$ is a pixel of the noisy image i at position (x, y), the intensity of the ideal image can be approximated by its empirical expectation $\hat{\mu}(x, y) = E[\{\tilde{\mathbf{U}}_i(x, y)\}]$ for $i = 1, ..., N$, where N is the number of snapshots in the series. The empirical variance associated to intensity $\hat{\mu}(x, y)$ is $\hat{\sigma}^2(x, y) = \text{Var}[\{\tilde{\mathbf{U}}_i(x, y)\}]$.

The calibration pattern must be mostly flat and represent a wide range of gray levels. Since the noise curve mainly depends on the ISO sensitivity, a different noise curve is estimated for each ISO level. Series of different exposure times were taken for each ISO in order to get representative information in the whole gray level range. The noise curves for

different times of exposure were combined to obtain a single curve. In order to get a ground-truth noise curve, for each exposure time (1/30 s, 1/250 s, 1/400 s, 1/640 s) about 200 pictures of the calibration pattern were taken. Since each 2×2 block of the color filter array (CFA) contains one sample of the red channel, two samples of the green channel, and one sample of the blue channel, the raw image was resampled as an image with four different color channels of half-width and height. Thus, four different noise ground-truth curves were obtained from the series, each corresponding to one of the four channels of the CFA. By splitting the color range into *bins* (disjoint in intensity intervals) and computing the median value of the STDs at each bin, a ground truth is obtained for the camera noise curve given the ISO and exposure times.

Figure 1 shows the noise curves obtained with a Nikon D80 camera with fixed ISO sensitivity of 1250 and 1600 and four exposure times, $t \in \{1/30 \text{ s}, 1/250 \text{ s}, 1/400 \text{ s}, 1/640 \text{ s}\}$. The obtained curves overlap perfectly. Each one treats a different color interval, thus permitting us to fuse them into a single noise curve. This fused curve can be observed in the same figure. For each color value, the fused noise estimation is obtained by the median of the available estimations obtained for the different exposure time. The value for each curve is linearly interpolated using the two closest neighbors. Since the noise curve does not depend on the exposure time, these curves overlap (hence the double values). However, this overlap is not perfect because the STD is computed with a finite number of samples and therefore the estimation has some variance that causes a small error centered at the theoretical value. Curve (b) is the mean of all four curves at different exposure times, which cancels their variation around the theoretical value, and therefore it can be used finally as a ground truth for evaluating noise estimation algorithms. Figure 2 displays the approximation of the computed ground-truth values by a linear model with the Nikon D80
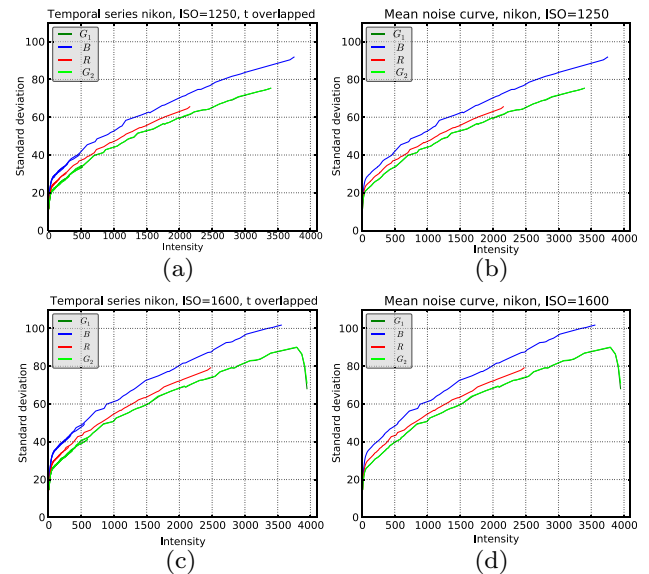

Fig. 1. Noise ground-truth curves obtained for a Nikon D80 camera with fixed ISOs of (a) 1250 and (c) 1600 and four exposure times, $t \in \{1/30 \text{ s}, 1/250 \text{ s}, 1/400 \text{ s}, 1/640 \text{ s}\}$, using laboratory calibration. Channels G_1 and G_2 give the same STD. The obtained curves overlap perfectly. Since they cover different color intervals, their fusion yields a complete noise curve (b), (d).
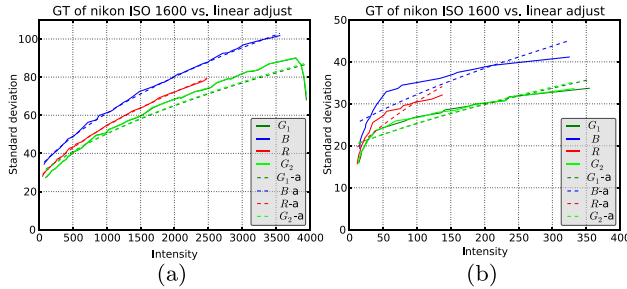
Fig. 2.   Linear approximation of the variance ground truth in Fig. 1 with the Nikon D80 with ISO 1600 (solid lines, original values; dashed lines, approximation). Exposure time (a) 1/30 s and (b) 1/640 s. Because of the saturation at the darkest zones, the estimated noise in the dark gray level does not follow a linear model.

camera. Because of the saturation at the darkest zones, the estimated noise in the dark gray level does not follow a linear model. However, using the partial linear model splitting the curve into three parts might be useful to model this kind of curve if the noise model is known in advance. The ground-truth curve obtained with the procedure presented here describes accurately the characteristics of the noise without depending only on the minimal assumption that the noise depends only on the intensity. Parametric methods assume priors for a particular noise model, and therefore their accuracy depends on how realistic these assumptions are. Section 4 shows that it is possible to get a reliable noise curve that matches with negligible error the nonparametric ground truth, and Section 5 shows that indeed it is possible to validate parametric methods with the nonparametric ground-truth curve.

4. NONPARAMETRIC SIGNAL-DEPENDENT NOISE ESTIMATION

Parametric models are accurate under the condition of prior knowledge about the noise model. For example, the Foi *et al.* [34] method assumes the linear model $\sigma^2 = A + B\mathbf{U}$ for the variance, but with a saturation effect. On the other hand, Uss *et al.* showed that the measured noise variance cannot always be fitted with a linear function, but with a polynomial of at least second order [35]. However, we were unable to fit a second-, third-, or fourth-order polynomial to the saturated noise curve in Fig. 2. In order to use a linear function, these authors modify the intensity of the pixels at each SW of the image by a function that nullifies quadratic and higher terms of the noise variance model. After this transformation, the estimation is accurate.

Parametric methods require a validation, by a comparison to ground-truth noise curves. The data in the ground truth must be empirical, in the sense that it does not assume any prior (with the exception that the variance of the noise is a function of the expectation) and simply measures the variance of the noise *as-is*. As discussed in Section 3, the major problem of the comparison against the ground truth is that it is different for each camera model and it must be obtained under controlled laboratory conditions.

Our goal here is precisely to show that the laboratory calibration method used to obtain the ground truth can be replaced by a nonparametric method, estimating directly on the image the signal-dependent noise. We adapted the Ponomarenko *et al.* method [31], since it is scored as the best

method in a previous review [8]. Other nonparametric estimation methods could be used as well. For example, in the paper of Liu *et al.* [37], a flexible eigenfunction representation of the noise level curves was proposed, but it requires *a priori* segmentation of the noisy image.

We extended the Ponomarenko *et al.* method [31] to be able to estimate signal-dependent noise. To this aim, the means of the blocks are classified into a disjoint union of variable intervals (bins), in such a way that each interval contains a fixed and large enough number of elements. Thus, these intervals are automatically adapted to the image itself instead of prefixed, since the intensity range of each bin depends on the mean intensity of the $w \times w$ blocks in the image. We found that each bin should contain at least 42,000 samples, which seems to be the lowest number permitting a reliable estimation. The value of p for the p quantile of block variances must be small to avoid blocks with large variance, corresponding to edges and textures. In general, if a bin is made of blocks that belong to a flat or smooth zone, we found experimentally that 210 per bin are enough to estimate the variance (using $p = 0.5$, the median). However, with a smaller $p = 0.005$ percentile value, we can discard 99.5% of the blocks with a higher variance and therefore in general all of the blocks affected by edges and textures. By choosing a large value for the bin cardinality, namely 42,000, we ensure that the blocks below this low quantile are still numerous enough, $42000 \times 0.005 = 210$, so that they ensure a reliable estimate of the noise variance.

To each bin a list of image blocks is associated, each of them being endowed with a list of STDs. Notice that a bin does not correspond to a spatial region of the image, but only to a set of blocks with similar means.

Another modification with respect to the original method is the procedure to find the best p quantile. The values in the list $\{\mathbf{V}^H(i,j)\}$ [see Eq. (4)] depend on the choice of p quantile. If p is small, the method becomes more robust to the influence of the textures and geometry of the image, but the accuracy of the estimation also decreases with p. Our assumption is that the variance measured using $p = 0.5$ (the median) should not be significantly different from the variance obtained with lower values of p, unless the image is composed mainly of textures. The proposed iteration to get a robust estimation of the variance, adapting p, is as follows. At each bin,

1. Initially, $p = 0.5$ (the median) and $\Delta p = 0.005$.
2. Set $S = \{\mathbf{V}^H(i,j) | i + j \geq T\}$.
3. Set V_s to the median of the values of S under the Δp quantile.
4. Set V_p to the median of the values of S under the p quantile.
5. If $p \geq \Delta p$ and $V_p \leq V_s$, then [set $p = p - \Delta p$ and go to step 3]; else END.

This procedure decreases the initial p from the median to a lower value that makes the estimation robust to textures and geometry, if needed.

About the $w \times w$ size of the SW, we use the same value that the authors of the original algorithm proposed, and we found that indeed the best results are obtained with $w = 8$ in most natural images. Since the optimal size of the window depends on the density of edges and level of texturization of the noisy image, if it is *a priori* known that the image is mainly composed of large flat or smooth areas, it is better to use a larger

window (up to 21×21) and to choose a smaller size in the opposite case (but at least 3×3) to obtain a reliable variance noise estimation. A larger window estimates the noise more accurately (since the sample variance estimator has itself a variance that depends on the number of available samples), but is less likely to contain only data from flat or smooth zones, and more likely to capture edges and geometry. Nevertheless, the proposed method is "blind," in the sense that no prior information about the characteristics of the image or the noise is available. Adapting the window size is beyond the scope of this paper, and it is left as future work.

To avoid outliers in the estimation, we systematically discard completely saturated blocks. Indeed, when the number of photons counted by the CCD during the exposure time is too high, its output may get saturated, and therefore underestimated. When the signal saturates the output of the CCD, the measured variance in the saturated areas of the image is zero. Indeed, the effect of the saturation must be measured and given by the nonparametric method in the produced noise curve, but the completely saturated pixels have outlier intensities. Figure 3 shows a noise curve obtained by using or avoiding the saturated blocks, where the modified Ponomarenko *et al.* [31] algorithm was performed with 49 bins. Since the intensity of the saturated pixels is much higher (outlier) than the intensity of the rest of the points, the noise curve is linearly interpolated along the gap in between. This natural image is a normal scene that is useful to illustrate the problem of the saturation. The bike is not illuminated directly by any source of light (only ambient light), and therefore it does not reflect much light, with the exception of a few points at the handlebar that reflect light with enough power to saturate the detectors. The STD equal to zero (measured near intensity 4000) is indeed correct, but all the interpolated points in between are definitely not. The strategy we adopted was to discard the blocks that contain a subgroup of 2×2 pixels sharing the same intensity, in any of the channels. It must be noted that this only removes the blocks containing pixels that are completely saturated, but keeps the rest of the blocks, including those where the noise distribution is truncated, but not absolutely saturated. This permits us to measure and observe the saturation in the curve, as shown in Fig. 2.

5. CROSS-VALIDATION OF SEVERAL METHODS. DISCUSSION

In order to compare ground truth, parametric methods, and our nonparametric method, we used a dataset of 20 images obtained with a Nikon D80 camera using ISO 1250 and exposure time 1/640 s. In these images the darkest pixels are saturated, and therefore the noise curve does not follow the linear variance model. Our dataset contains some views of a room with objects over a table, and images of corridors, bookshelves, (Fig. 6) stairs, and classrooms inside a building, with different lighting levels. Also, two outdoor images of highly textured images (Fig. 5). For each test image, we computed the RMSE between the STDs given by the method and by the ground truth. The control points are given by the method, and the STD of the ground truth corresponding to that intensity is obtained by linear interpolation between the two nearest intensity control points of the ground-truth curve.

Figure 4 shows the obtained results. In general, the RMSE of the modified Ponomarenko *et al.* (red curve) method is close to zero, which means that it could be used to establish a (nonparametric) ground truth. The estimations given by Foi *et al.* (green curve) and Uss *et al.* (blue curve) are really close to the nonparametric ground truth, and are therefore also validated by our approach. The Foi *et al.* method failed to measure the noise correctly when the images were composed mainly of textures (image nos. 19 and 20; see Fig. 5), whereas the Uss *et al.* and the proposed method gave good results in that case. Note that textures cause a localized error in the nonparametric curve (middle), whereas they cause a global error in the parametric curves. Figure 6 shows three examples of images in our dataset (images nos. 8 and 12), where all algorithms estimated the noise correctly. The Foi *et al.* method (red curve) matches accurately the ground-truth curves (green and blue), since it is designed to predict the shape of the curve under saturation conditions, whereas the Uss *et al.* estimation is overall correct, except in the saturation zone, as expected. As explained in Section 2, the original Ponomarenko method is only able to estimate homoscedastic noise, that is, a value of STD that does not depend on the intensity. Equation (5) shows how this STD, which does
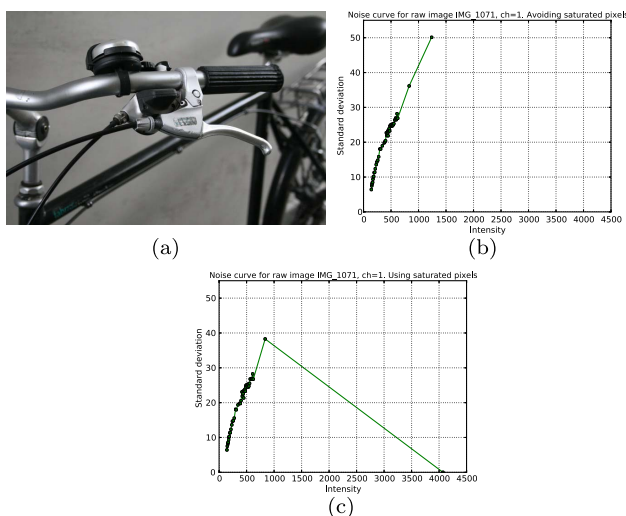

Fig. 3. Noise curve obtained (b) when the saturated pixels are avoided in the noise estimation and (c) when they are taken into account by using the Ponomarenko *et al.* method [31] with 49 bins in an image with saturated pixels (a).
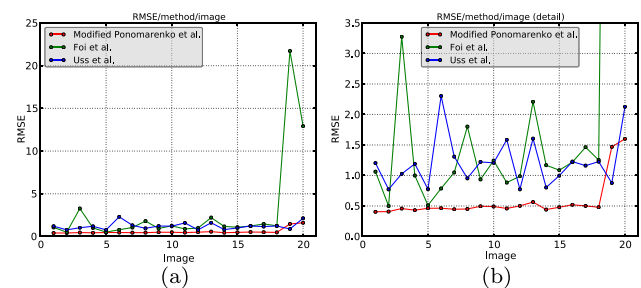

Fig. 4. RMSE between the methods and the ground truth for all 20 images in our dataset. In general, the RMSE of the modified Ponomarenko *et al.* (red curve) method is close to zero, which means that indeed it can be considered a nonparametric ground-truth curve. The estimations given by Foi *et al.* (green curve) and Uss *et al.* (blue curve) are really close to the nonparametric ground truth, and therefore they are also validated by our approach. (a) Obtained RMSEs/image, (b) detailed view.

(a)   (b)
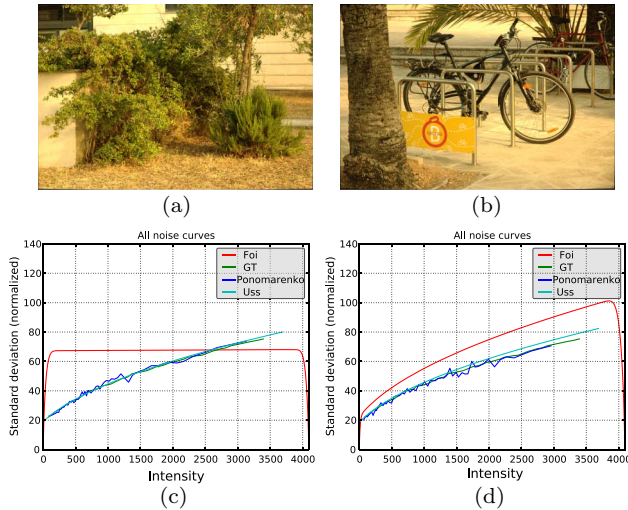

(c)   (d)

Fig. 5. Highly textured images that caused small oscillations in the noise curves with the proposed method and wrong results with Foi *et al.*: images (a) no. 19 and (b) no. 20 (see the obtained RMSEs in Fig. 4). (c), (d) Ground truth obtained with the series (green), the non-parametric ground truth (darker blue), the Uss *et al.* method (brighter blue), and the Foi *et al.* method (red).

not depend on the intensity, is computed. However, in Figs. 5 and 6 we show noise curves that correspond to the modified Ponomarenko method: added bins to get control points in the

(a)   (b)

(c)   (d)
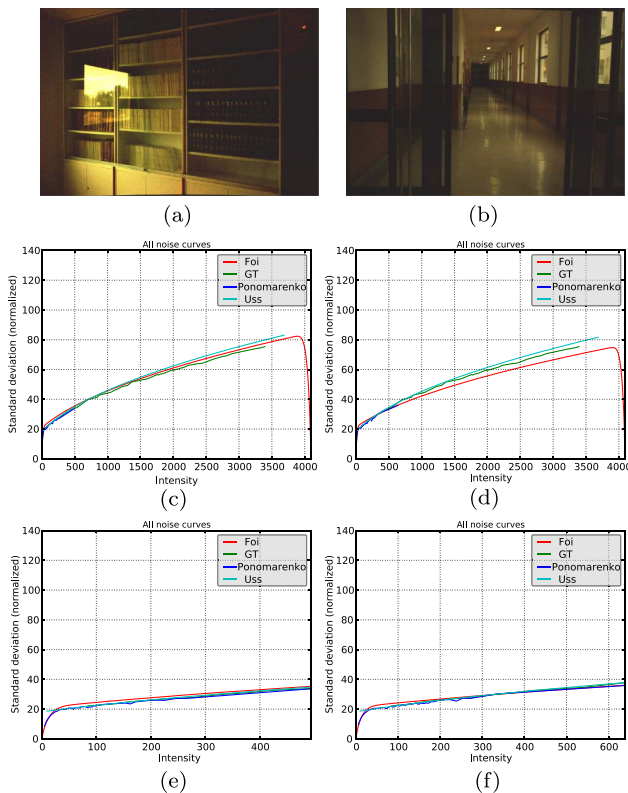

(e)   (f)

Fig. 6. Examples of images [(a) no. 8 and (b) no. 12 in our dataset] in which all algorithms estimated the noise correctly. (c), (d) Their noise curves along all the intensity range. (e), (f) Detail of the noise curves only within the range of the estimation given by the modified Ponomarenko *et al.* method (nonparametric ground truth, green curve). Note that the Foi *et al.* method (red curve) matches accurately the ground-truth curves (green and blue), since it is designed to predict the shape of the curve under saturation conditions, whereas the Uss *et al.* estimation is overall correct, except in the saturation zone.

curve for different intensities and avoid using completely saturated points before the estimation. Of course, the overestimation caused by a bin where all samples belong to textures can be avoided if more than a single image is available, by estimating the noise in the mosaic made of several different input images.

As shown in Fig. 2, the linear model does not hold when the image is saturated. Uss *et al.* tried to use a second-order polynomial to fit the saturated noise curve. However, we found that a second-order polynomial was not general enough to fit the saturated curves. Foi *et al.* assumed the linear model, but taking into account the effect of the saturation. However, both methods assume that the noise can be modeled with a linear function when there is no saturation. This is true for most CCDs, but the output recorded in the raw file given by the camera might not be a linear function of the intensity [38]. This makes clear the necessity of validating parametric methods, which assume an *a priori* model for the noise. In contrast, the estimates of a nonparametric method rely on the minimal assumption that the signal is a function of the expectation. In general, the best results are obtained with the modified Ponomarenko *et al.* method.

To decide if a method is valid or not, its RMSE with respect to the nonparametric ground truth has to be compared to a threshold. We consider that a method is valid if the RMSE between the measured STD and the ground truth is less than or equal to $\Delta\hat{\sigma}_8 = 0.15$ (assuming that the images are encoded with 8 bits). This value was chosen to be as low as possible and, at the same time, consistent with the accuracy of state-of-the-art noise estimation methods. Since the raw images are encoded with 12 bits, the threshold is $\gamma = \Delta\hat{\sigma}_8 \times 16 = 2.4$. The estimation of the Foi *et al.* method is considered valid in 17 of 20 images, whereas the Uss *et al.* method is validated with all the images in our dataset.

A. Complexity

The Uss *et al.* method follows four steps: (1) initialization of the TI map and the polynomial function for the variance; (2) estimate texture and noise variance for each TI and NI SW and label the SW into NI or TI; (3) update the CRLB; and (4) apply the noise estimator to the samples associated with each bin and update the variance polynomial. Steps 2, 3, and 4 are iterated until convergence is reached. The complexity of the Uss *et al.* and the modified Ponomarenko *et al.* methods is similar, and their complexity is linear with the number of pixels in the image. Both imply an estimation of the noise variance at the DCT coefficients in small patches of the image after classifying them according to their intensity. For its part, the Foi *et al.* method follows these steps in order to obtain the final parametric model: compute the detail wavelet coefficients of the image, segment the image to find homogeneous zones, estimate locally pairs of intensity/variance, and finally the ML fitting of the global parametric model. All steps can be computed quickly, but unlike Uss *et al.* and the modified Ponomarenko *et al.* method, it requires a previous segmentation of the image.

B. Denoising Results

We used the noise curves obtained with the Uss *et al.*, the Foi *et al.*, and the modified Ponomarenko *et al.* methods as the input of the NL-Bayes [39] denoising algorithm, after applying

a VST to the noisy image. Only the green channel was used. Note that according to our threshold criterion, both the Uss *et al.* and Foi *et al.* methods are validated, and therefore their denoising results in almost all images in the dataset being very similar. Figure 7(a) shows details of the results obtained for image no. 3 of our dataset, where the Foi *et al.* method failed to estimate the noise correctly. While the Uss *et al.* and the modified Ponomarenko *et al.* methods denoise the image properly, the noise at the dark zone (the bag over the table) remains visible. Image (e) is test image no. 20 of our dataset, where the Uss *et al.* and the modified Ponomarenko *et al.* methods give an valid estimation, whereas the Foi *et al.* method overestimates. All methods gave an increased RMSE for that particular image, which causes blurred denoised images and loss of fine details. Both the Uss *et al.* and the modified Ponomarenko *et al.* methods give similar visual
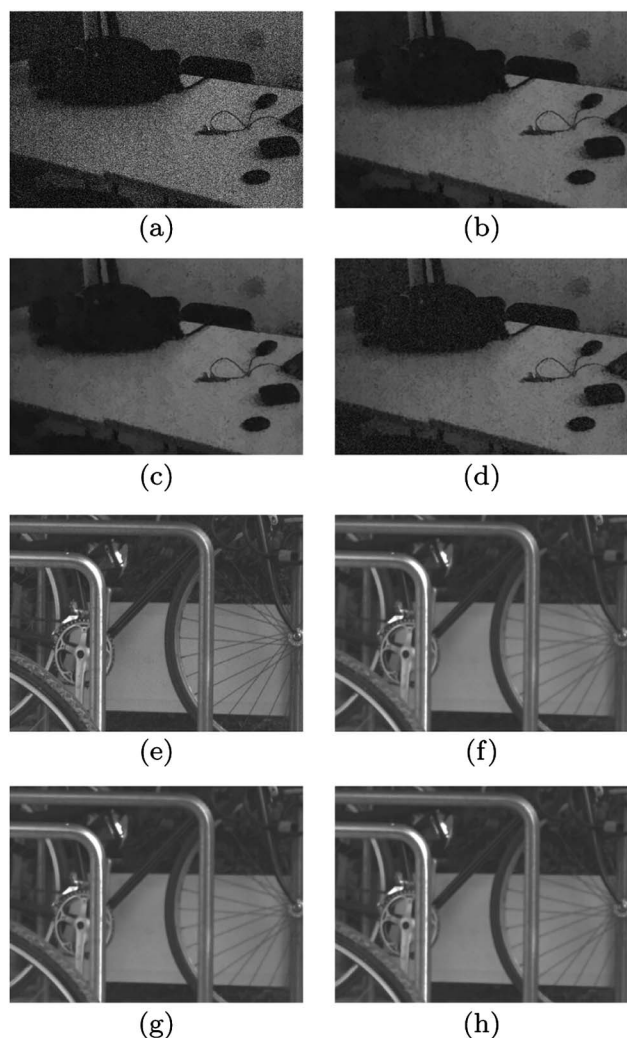

Fig. 7. Details of the denoising results with the NL-Bayes algorithm using the noise curves obtained from the noisy images (a), (e) with (b), (f) modified Ponomarenko *et al.*, (c), (g) Uss *et al.*, and (d), (h) Foi *et al.* methods. Image (a) is test image no. 3 of our dataset (it is very dark, so we increased the brightness for visualization purposes), where the Foi *et al.* method was unable to give a reliable estimation and thus the noise is not removed at the dark zones and remains visible (the bag over the table). Image (e) is a detail from test image no. 20 of our dataset, where the Uss *et al.* and the modified Ponomarenko *et al.* methods give a valid estimation, whereas the Foi method overestimates.

results, whereas the overestimation in the method blurs the image even more.

6. CONCLUSIONS

We showed that estimating an accurate noise curve from a single raw image is possible and can be done by an adaptation of a nonparametric noise estimator [31]. The only minimal assumption is that the noise STD is a function of the expected signal. Being able to apply a noise estimator (with relatively low complexity compared to denoising algorithms) to each raw image frees the users of a tedious and sometimes impossible camera calibration task. Indeed, noise curves obtained in an optical lab require measurements for each ISO and each optical setup, a heavy and costly procedure. By estimating the noise directly on the raw image, there is no risk of model error or accuracy loss caused by a noise parameter estimation on another camera.

According to the provided RMSE results (see Fig. 4), the nonparametric method proposed here exhibits a very stable error (close to RMSE = 0.5) when the image is not composed mainly of textures. However, even if the image is highly textured (see images nos. 19 and 20 in Fig. 5), the error is small and similar to the RMSE obtained with the compared state-of-the-art methods.

In general, the estimation given by the proposed method is as reliable as the actual ground truth obtained from the temporal series of the series of images of a calibration pattern in the laboratory, and matches the best parametric methods.

ACKNOWLEDGMENTS

REFERENCES

1. P. Milanfar, "A tour of modern image filtering: new insights and methods, both practical and theoretical," IEEE Signal Process. Mag. **30**(1), 106–128 (2013).
2. H. Rabbani, M. Sonka, and M. D. Abramoff, "Optical coherence tomography noise reduction using anisotropic local bivariate," Int. J. Biomed. Imag. **3**, 417491 (2011).
3. J. Schmitt, J. L. Starck, J. M. Casandjian, J. Fadili, and I. Grenier, "Multichannel Poisson denoising and deconvolution on the sphere: application to the Fermi gamma ray space telescope," Astron. Astrophys. **546**, A114 (2012).
4. F. Luisier, T. Blu, and M. Unser, "Image denoising in mixed Poisson–Gaussian noise," IEEE Trans. Image Process. **20**, 696–708 (2011).
5. F.-X. Dupé, J. M. Fadili, and J.-L. Starck, "A proximal iteration for deconvolving Poisson noisy images using sparse representations," IEEE Trans. Image Process. **18**, 310–321 (2009).
6. H. Rabbani, R. Nezafat, and S. Gazor, "Wavelet-domain medical image denoising using bivariate laplacian mixture model," IEEE Trans. Biomed. Eng. **56**, 2826–2837 (2009).
7. B. Zhang, J. M. Fadili, and J.-L. Starck, "Wavelets, ridgelets, and curvelets for poisson noise removal," IEEE Trans. Image Process. **17**, 1093–1108 (2008).
8. M. Lebrun, M. Colom, A. Buades, and J. M. Morel, "Secrets of image denoising cuisine," Acta Numerica **21**, 475–576 (2012).

9. E. D. Kolaczyk, "Wavelet shrinkage estimation of certain Poisson intensity signals using corrected thresholds," Statist. Sin. **9**, 119–135 (1999).

10. R. D. Nowak and R. G. Baraniuk, "Wavelet-domain filtering for photon imaging systems," IEEE Trans. Image Process. **8**, 666–678 (1997).

11. S. Lefkimmiatis, P. Maragos, and G. Papandreou, "Bayesian inference on multiscale models for Poisson intensity estimation: application to photo-limited image denoising," IEEE Trans. Image Process. **18**, 1724–1741 (2009).

12. C. A. Deledalle, L. Denis, and F. Tupin, "Nl-insar: nonlocal inter-ferogram estimation," IEEE Trans. Geosci. Remote Sens. **49**, 1441–1452 (2011).

13. A. Buades, B. Coll, and J. M. Morel, "A review of image denoising algorithms, with a new one," Multiscale Model. Simul. **4**, 490–530 (2005).

14. F. J. Anscombe, "The transformation of Poisson, binomial and negative-binomial data," Biometrika **35**, 246–254 (1948).

15. M. Makitalo and A. Foi, "Optimal inversion of the Anscombe transformation in low-count Poisson image denoising," IEEE Trans. Image Process. **20**, 99–109 (2011).

16. A. Foi, "Noise estimation and removal in mr imaging: the variance-stabilization approach," in *2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro* (IEEE, 2011), pp. 1809–1814.

17. J. Salmon, C.-A. Deledalle, and A. Dalalyan, "Image denoising with patch based PCA: local versus global," in *Proceedings of the British Machine Vision Conference* (BMVA, 2011), pp. 25.1–25.10.

18. N. N. Ponomarenko, V. V. Lukin, S. K. Abramov, K. O. Egiazar-ian, and J. T. Astola, "Blind evaluation of additive noise variance in textured images by nonlinear processing of block DCT coefficients," in *Proceedings of the International Society for Optics and Photonics. Electronic Imaging, Image Processing: Algorithms and Systems II* (2003), Vol. **5014**, pp. 178–189.

19. H. Rabbani and S. Gazor, "Local probability distribution of natural signals in sparse domains," Int. J. Adapt. Control Signal Process. **28**, 52–62 (2014).

20. V. I. A. Katkovnik, V. Katkovnik, K. Egiazarian, and J. Astola, *Local Approximation Techniques in Signal and Image Processing* (SPIE, 2006), Vol. **PM157**.

21. S. Pyatykh, J. Hesser, and L. Zheng, "Image noise level estimation by principal component analysis," IEEE Trans. Image Process. **22**, 687–699 (2013).

22. J. S. Lee, "Refined filtering of image noise using local statistics," Comp. Graph. Image Proc. **15**, 380–389 (1981).

23. D. L. Donoho and I. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," Biometrika **81**, 425–455 (1994).

24. D. L. Donoho and I. M. Johnstone, "Adapting to unknown smoothness via wavelet shrinkage," J. Am. Stat. Assoc. **90**, 1200–1224 (1995).

25. R. Bracho and A. C. Sanderson, "Segmentation of images based on intensity gradient information," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (IEEE, 1985), pp. 19–23.

26. J. Immerkaer, "Fast noise variance estimation," Comput. Vis. Image Underst. **64**, 300–302 (1996).

27. J. S. Lee and K. Hoppel, "Noise modelling and estimation of remotely-sensed images," in *Proceedings of the International Geoscience and Remote Sensing Symposium* (1989), Vol. **2**, pp. 1005–1008.

28. G. A. Mastin, "Adaptive filters for digital image noise smoothing: an evaluation," Comput. Vis. Graph. Image Process. **31**, 103–121 (1985).

29. P. Meer, J. M. Jolion, and A. Rosenfeld, "A fast parallel algorithm for blind estimation of noise variance," IEEE Trans. Pattern Anal. Mach. Intell. **12**, 216–223 (1990).

30. S. I. Olsen, "Estimation of noise in images: an evaluation," Graph. Models Image Proc. **55**, 319–323 (1993).

31. N. N. Ponomarenko, V. V. Lukin, M. S. Zriakhov, A. Kaarna, and J. T. Astola, "An automatic approach to lossy compression of AVIRIS images," in *International Geoscience and Remote Sensing Symposium* (IEEE, 2007), pp. 472–475.

32. K. Rank, M. Lendl, and R. Unbehauen, "Estimation of image noise variance," in *IEEE Proceedings on Vision, Image and Signal Processing* (IET, 1999), Vol. **146**, pp. 80–84.

33. H. Voorhees and T. Poggio, "Detecting textons and texture boundaries in natural image," in *Proceedings of the First International Conference on Computer Vision London* (IEEE, 1987), pp. 250–258.

34. A. Foi, M. Trimeche, V. Katkovnik, and K. Egiazarian, "Practical Poissonian-Gaussian noise modeling and fitting for single-image raw-data," IEEE Trans. Image Process. **17**, 1737–1754 (2008).

35. M. L. Uss, B. Vozel, V. V. Lukin, and K. Chehdi, "Image inform-ative maps for component-wise estimating parameters of signal-dependent noise," J. Electron. Imaging **22**, 013019 (2013).

36. M. Uss, B. Vozel, V. Lukin, S. Abramov, I. Baryshev, and K. Chehdi, "Image informative maps for estimating noise standard deviation and texture parameters," EURASIP J. Advances Signal Process. **2011**, 806516 (2011).

37. C. Liu, R. Szeliski, S. B. Kang, C. L. Zitnick, and W. T. Freeman, "Automatic estimation and removal of noise from a single image," IEEE Trans. Pattern Anal. Mach. Intell. **30**, 299–314 (2008).

38. P. L. Vora, J. E. Farrell, J. D. Tietz, and D. H. Brainard, "Linear models for digital cameras," in *IS&T Annual Conference* (The Society for Imaging Science and Technology, 1997), pp. 377–382.

39. M. Lebrun, A. Buades, and J. M. Morel, "A nonlocal Bayesian image denoising algorithm," SIAM J. Imaging Sci. **6**, 1665–1688 (2013).