

# Indirect Estimation of Signal-Dependent Noise With Nonadaptive Heterogeneous Samples

Lucio Azzari and Alessandro Foi

**Abstract**—We consider the estimation of signal-dependent noise from a single image. Unlike conventional algorithms that build a scatterplot of local mean-variance pairs from either small or adaptively selected homogeneous data samples, our proposed approach relies on arbitrarily large patches of heterogeneous data extracted at random from the image. We demonstrate the feasibility of our approach through an extensive theoretical analysis based on mixture of Gaussian distributions. A prototype algorithm is also developed in order to validate the approach on simulated data as well as on real camera raw images.

**Index Terms**—Noise estimation, signal-dependent noise, Poisson noise.

## I. INTRODUCTION

THE popularity of signal-dependent noise models, in which the variance of the noise affecting the signal depends on the mean of the signal, is based on the fact that they well approximate noise affecting data of several kinds of acquisition devices, e.g., raw data from a CCD camera. Figure 1 illustrates how the signal-dependent noise differently affects bright and dark regions of an image, and shows a curve that describes the typical mean-variance relation of imaging sensors.

Conventional methods [1], [6]–[13] estimate points of such mean-variance curve isolating and separately processing segments or patches of the signal with common mean and noise variance, so that on each segment or patch simple sample estimators of mean and variance can be applied. In this way, a scatterplot in the mean-variance plane is produced. Then, a curve is fitted to the scatterplot, yielding an estimate of the relation for the whole range of the signal. In this paper we show that, contrary to common belief, the estimation can be accurate even if each scatterplot point is estimated from a heterogeneous sample (e.g., a patch whose pixels can have very different mean values). We justify this result through a mathematical modeling based on mixtures of normal distributions.

Thus, unlike conventional signal-dependent noise estimation techniques that preprocess the image in order to work with homogeneous samples, our approach applies robust estimators

Manuscript received December 19, 2012; revised November 11, 2013 and March 19, 2014; accepted April 13, 2014. Date of publication May 2, 2014; date of current version July 1, 2014. This work was supported by the Academy of Finland under Project 252547. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Jose M. Bioucas-Dias.

The authors are with the Department of Signal Processing, Tampere University of Technology, Tampere, FI-33101, Finland (e-mail: lucio.azzari@tut.fi; alessandro.foi@tut.fi).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TIP.2014.2321504

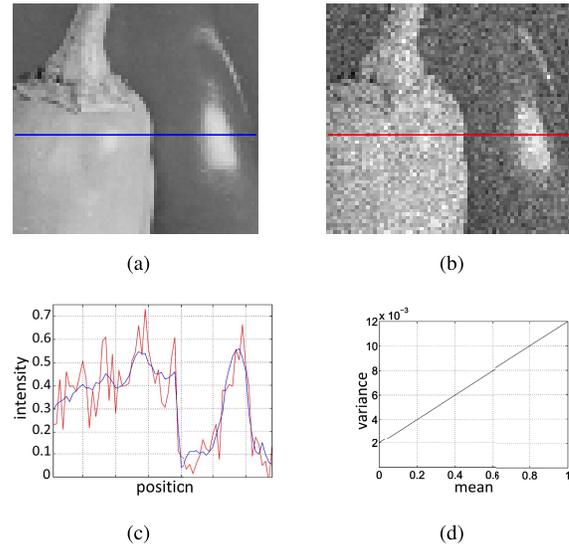


Fig. 1. Detail of the “Peppers” image corrupted by signal-dependent noise with affine variance (2), with parameters  $a = 0.01$  and  $b = 0.002$ . (a) Noise-free image. (b) Noisy realization. (c) Cross-section. (d) Mean  $\mapsto$  variance relation.

to arbitrarily large patches of heterogeneous data extracted at random from the image.

Our analysis is focused on the camera noise models such as the affine-variance model depicted in Figure 1. For the sake of clarity and due to length limitation, we restrict the presentation to the 2D image case; nevertheless, the introduced concepts and the proposed approach apply universally to 1D signals as well as to multidimensional data.

The paper is organized as follows. In Section II we introduce the considered signal-dependent noise model and we describe the conventional approach for its estimation. Next, we present our novel noise estimation technique and a prototype algorithm that exploits it, discussing its difference w.r.t. conventional methods. In Section III we study the main factors contributing to estimation errors, through a theoretical analysis and a Monte Carlo simulation. In Section IV we show the effectiveness of the method in real applications by estimating noise affecting raw data from a CCD camera, and a comparison with a state-of-the-art algorithm. Finally, in Section V and Section VI we provide discussions and conclusions.

## II. METHOD

### A. Problem Statement

Let us consider a noisy observation  $z$  of a deterministic noise-free signal  $y$ , corrupted by additive spatially uncorre-

lated noise with signal-dependent variance:

$$z(\mathbf{x}) = y(\mathbf{x}) + \sigma(y(\mathbf{x}))\zeta(\mathbf{x}), \quad (1)$$

where  $\sigma : \mathbb{R} \rightarrow \mathbb{R}^+$  is a function giving the signal-dependent standard deviation of the noise,  $\mathbf{x} \in \mathbf{X} \subset \mathbb{Z}^2$  is the pixel coordinate, and  $\zeta : \mathbf{X} \rightarrow \mathbb{R}$  is a zero-mean independent random noise with standard deviation equal to 1. Our goal is to estimate the function  $\sigma$ .

The expectation of  $z(\mathbf{x})$ , denoted as  $E\{z(\mathbf{x})\}$ , is the noise-free signal  $y(\mathbf{x})$ ; at the same time, the variance  $\text{var}\{z(\mathbf{x})\}$  and the standard deviation  $\text{std}\{z(\mathbf{x})\}$  of  $z(\mathbf{x})$  are, respectively,  $\sigma^2(y(\mathbf{x}))$  and  $\sigma(y(\mathbf{x}))$ , because  $\text{var}\{y(\mathbf{x})\} = 0$ .

As discussed in [2], the term  $\zeta(\mathbf{x})$  can generally have a different probability distribution for each different coordinate  $\mathbf{x}$ , i.e.  $\zeta(\mathbf{x}_1) \approx \zeta(\mathbf{x}_2)$  if  $\mathbf{x}_1 \neq \mathbf{x}_2$ ; in order to simplify the mathematical model, we approximate  $\zeta(\mathbf{x})$  as a normal distribution  $\mathcal{N}(0, 1)$ . In this way the noise can be considered heteroskedastic Gaussian, with zero mean and signal-dependent variance  $\sigma^2(y(\mathbf{x}))$ , i.e.  $\sigma(y(\mathbf{x}))\zeta(\mathbf{x}) \sim \mathcal{N}(0, \sigma^2(y(\mathbf{x})))$ .

To provide practical experimental results of our method, we shall refer to the affine noise variance model [5], which is one of the most suitable for modeling the noise in digital image sensors. According to this model, the noise variance is approximated as

$$\sigma^2(y(\mathbf{x})) = ay(\mathbf{x}) + b, \quad (2)$$

where  $ay(\mathbf{x})$  and  $b$  are, respectively, the variances of the signal-dependent and signal-independent parts of the noise. The former part is due to a photon-counting process (Poisson distribution), while the latter is caused by a combination of dark noise (Poisson distribution) and thermal-electronic noise (normal distribution). Because of a central-limit theorem argument and because of the good approximation of the Poisson by a Gaussian, the normal approximation of  $\zeta(\mathbf{x})$  is valid. For (2), the problem of estimating  $\sigma^2$  can be reduced to the estimation of the two constants  $a$  and  $b$ .

### B. Conventional Approach

The conventional approach for the estimation of signal-dependent noise is to segment the image into regions where pixels have constant intensity, and hence, because of (1), constant noise; then, the mean and noise variance are estimated for each region independently. In this way it is possible to create a scatterplot that relates the noise-free intensity values of  $y$  (abscissa) with the respective noise variances (ordinate), that, finally, is used to approximate the function  $\sigma(y)$  in (1) (or equivalently the function  $\sigma^2(y)$ ).

There are different methods for partitioning the image, with different complexity and accuracy. The partition can be made, e.g., by simply using pixels extracted from a sufficiently small window from the noisy image [9], with the constraint that the intensity does not change much within the window [7], [8], or by segmenting the image into level sets (bins) with individual intensity values [1], [6], [10], [11], [13]. More sophisticated techniques, such as DCT-based estimators [12], have been also proposed. However, the backbone idea is still to exploit homogeneous samples for the actual noise estimation.

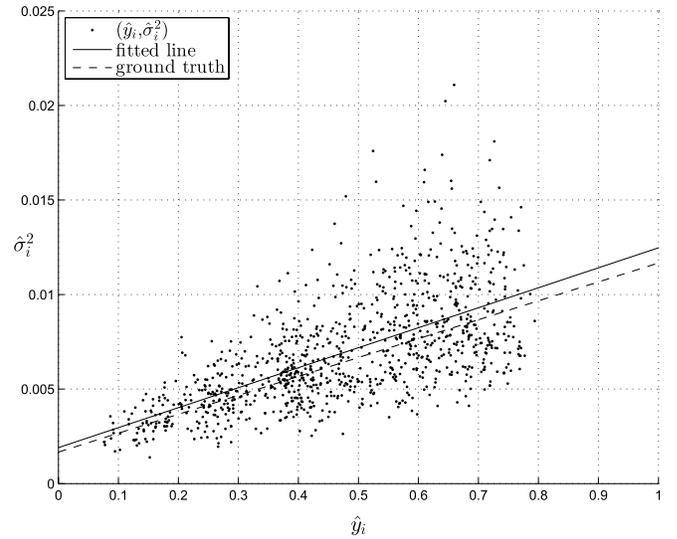


Fig. 2. Scatterplot of the mean-variance pairs  $(\hat{y}_i, \hat{\sigma}_i^2)$ , fitted line  $\hat{\sigma}(y) = \hat{a}y + \hat{b}$ , and ground truth line  $\sigma(y) = ay + b$  from the “Peppers” image, corrupted with noise parameters  $a = 0.01$  and  $b = 0.0017$ . We use 1000 blocks of size  $16 \times 16$  each yielding a point in the scatterplot.

The rationale of these techniques is that, being the segments homogeneous, also the noise variance is homogeneous, as can be trivially concluded from (1). Hence, standard estimators of the sample mean and sample variance can be directly applied to the segments, yielding unbiased estimates of the mean and noise variance. In other words, the resulting scatterplot points are distributed about the noise variance curve  $\sigma^2(y)$ .

### C. Main Idea

In contrast with the common procedure based on relatively small homogeneous segments, we show that the estimation of each scatterplot point can be performed processing large heterogeneous samples. As we shall demonstrate, considering a heterogeneous group of elements taken from  $z$ , the expectations of the estimators of its mean and noise variance are still a coordinate of a point that belongs to the function  $\sigma^2(y)$ . Consequently, it is not necessary to partition the image into segments of constant intensity levels and noise variances, but it is possible to process together parts of the image corrupted by noise with various variance values, without compromising the estimation. In particular, adaptive segmentation is no longer required in order to estimate signal-dependent noise, but its only advantage consists in limiting the positive bias due to outliers that could occur when estimating the variance. In this way we can avoid the segmentation step and, consequently, simplify the entire process.

We define our approach *indirect* because the pair estimated from one block does not represent directly a single relation mean-noise variance, like for the conventional methods, but it represents the mean and the variance of an heterogeneous group of elements, i.e. a mixture of distributions.

An example of the scatterplot computed from the blocks taken at random positions from the whole noisy image in Figure 1(a) is shown in Figure 2 (black dots), with its estimation of  $\sigma^2(y)$  and the ground truth.

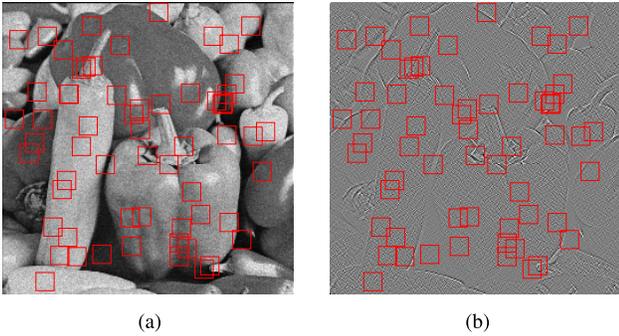


Fig. 3. Example of  $16 \times 16$  windows at random position in  $z$  and at corresponding positions in  $z^H$ . (a) Windowing in  $z$ . (b) Windowing in  $z^H$ .

#### D. Prototype Algorithm

The simplest algorithm that can leverage the above idea can be divided in three basic steps:

- (a) *High-Pass Filtering*: most of the energy of the noise-free signal  $y$  is usually confined to the lower frequencies of  $z$ , thus, applying an high-pass filter to  $z$  permits to extract the zero mean noise from it [3]. We obtain the high-frequency part of  $z$ , referred to as  $z^H$ , by convolving  $z$  against a 2D high-pass function  $\psi$  (e.g., a wavelet)

$$z^H = z \otimes \psi, \quad (3)$$

where  $\psi$  has zero mean, i.e.  $\sum_i \psi(i) = 0$ , and  $\ell^2$ -norm equal to one, i.e.  $\sum_i \psi^2(i) = 1$ .

- (b) *Local Estimation*: once the detail image  $z^H$  is computed, we randomly choose  $N$  coordinates within the image  $z$ , like in Figure 3; then, from these locations,  $N$  square blocks  $W_i^z$ ,  $i = 1, \dots, N$ , of size  $\sqrt{n} \times \sqrt{n}$  are extracted from  $z$ . Similarly,  $N$  blocks  $W_i^H$ ,  $i = 1, \dots, N$ , of the same size and from the same positions of  $W_i^z$ , are extracted from  $z^H$ . We estimate the means  $y_i$  from the blocks  $W_i^z$ , while from  $W_i^H$  we estimate the corresponding noise variances  $\sigma_i^2$ . In this way, for each block  $W_i^z$ , we obtain a pair  $(\hat{y}_i, \hat{\sigma}_i^2)$  which can be represented by a point in the scatterplot. The pairs  $(\hat{y}_i, \hat{\sigma}_i^2)$  are, therefore, the estimates of the blocks means and noise variances  $(y_i, \sigma_i^2)$ .

Because the blocks are taken from random positions within the image, each block may contain pixels having various expected intensity levels. Therefore, the distribution of noise in a single block  $W_i^z$  can be considered as a mixture of normal distributions with different variances. This marks a principal difference with the conventional methods that look for uniform blocks (or regions) for the estimation, and that model the noise within a single block as realization of a single normal distribution with given mean and variance.

In the next section we investigate the effects of exploiting elements taken from a mixture instead of from a single normal distribution.

- (c) *Fitting*: in order to estimate the parameters that describe the curve  $\sigma^2(y)$ , we fit the pairs  $(\hat{y}_i, \hat{\sigma}_i^2)$ ,  $i = 1, \dots, N$ ,

using a least squares (LS) method, which is the simplest fitting technique at our disposal.

### III. ESTIMATION ERROR

#### A. Noise Analysis

Let us model image blocks as composed by  $R_i$  regions (piecewise modeling), with  $R_i \leq n$ , and let  $W_i^y$  denote the noise-free block corresponding to  $W_i^z$ .

We shall refer as *ideal* the case in which, in  $W_i^H$ , the amount of energy due to  $y$  is negligible with respect to the noise energy. For example, this is the case when  $W_i^y$  can be treated as piecewise constant with edges having small excursions with respect to the noise standard deviation, or, equivalently, when the high-pass filter perfectly extract the noise component from  $z$ . In this case, the elements of  $W_i^z$  and  $W_i^H$  are, respectively, realization of two mixtures of  $R_i$  normal distributions with probability density functions (*p.d.f.*'s):

$$f_i^z(x) = \sum_{k=1}^{R_i} \lambda_k^{(i)} p_k^z(x), \quad p_k^z \sim \mathcal{N}(m_k, s_k^2), \quad (4)$$

$$f_i^H(x) = \sum_{k=1}^{R_i} \lambda_k^{(i)} p_k^H(x), \quad p_k^H \sim \mathcal{N}(0, s_k^2), \quad (5)$$

where  $p_k^z$  and  $p_k^H$  are, respectively, the *p.d.f.*'s of the  $k$ -th normal distributions of  $f_i^z$  and  $f_i^H$ ,  $\lambda_k^{(i)}$  is the proportion of the elements of the  $k$ -th population respect to the total number of elements  $n$ ,  $m_k$  is the mean of the  $k$ -th normal function in  $f_i^z$ , i.e. the  $k$ -th intensity value in  $W_i^y$ , and  $s_k^2$  is the variance of both  $p_k^z$  and  $p_k^H$ . It is important to notice that the ideality of this case relies mainly on the fact that the variances of the  $k$ -th distributions are equal.

Trivially we have

$$y_i = \sum_{k=1}^{R_i} \lambda_k^{(i)} m_k. \quad (6)$$

Exploiting the moments of a general mixture of normal distributions,<sup>1</sup> and the fact that all the  $p_k^H$  have zero mean, we obtain

$$\sigma_i^2 = \sum_{k=1}^{R_i} \lambda_k^{(i)} s_k^2. \quad (7)$$

Considering now the particular Poisson-Gaussian noise, it follows that the elements of  $W_i^H$  can be individually modeled as realizations of independent normal random variables with variances defined by the affine transformation (2) of  $W_i^y$ :

$$s_k^2 = am_k + b.$$

<sup>1</sup>The expectation  $m$  and the variance  $s^2$  of a mixture of  $G$  normal distributions are

$$m = \sum_{k=1}^G v_k m_k, \\ s^2 = \sum_{k=1}^G v_k [(m_k - m)^2 + s_k^2],$$

where  $m_k$ ,  $s_k^2$  and  $v_k$  are, respectively, the expectation, the variance and the proportion of the  $k$ -th distribution [4].

Consequently, noting that  $\sum_{k=1}^{R_i} \lambda_k^{(i)} = 1$ ,

$$\begin{aligned} \sigma_i^2 &= \sum_{k=1}^{R_i} \lambda_k^{(i)} a m_k + \sum_{k=1}^{R_i} \lambda_k^{(i)} b \\ &= a \sum_{k=1}^{R_i} \lambda_k^{(i)} m_k + b = a y_i + b. \end{aligned} \quad (8)$$

This means that the point  $(y_i, \sigma_i^2)$  belongs to the line (2). Therefore, if  $\hat{y}_i$  and  $\hat{\sigma}_i^2$  are computed, respectively, with unbiased estimators of the population mean and variance of a mixture of normal distributions, the points  $(\hat{y}_i, \hat{\sigma}_i^2)$  will yield a cloud scattered about the line (2), and the only error occurring in the computation of the pair  $(\hat{y}_i, \hat{\sigma}_i^2)$  is the one due to the variances of the estimators.

The above proof shows that, in ideal conditions, the presented algorithm ensures correct estimation even using blocks affected by different noise levels.

Let us now consider a more *practical* scenario where the presence of the noise-free signal is still appreciable in  $W_i^y$ , due to strong edges and texture in  $W_i^y$ . In this case, the noise distribution in  $W_i^H$  can no longer be approximated as a mixture of zero-mean normal distributions. In practice, this means that  $W_i^H$  does not contain only noise, and that, among its detail coefficients, there could be elements that introduce a bias in the estimation of  $\sigma_i^2$ . Consequently, the estimation error does not depend only on the variance of the estimator, but it is also influenced by the presence of edges in  $W_i^y$ .

To reduce the effect of these outliers, we use the median of absolute deviation (MAD) [15], [16] as robust estimator of  $\sigma_i^2$  and, for coherence, the median (med) as estimator of the mean:

$$\hat{y}_i = \text{med} \{W_i^z\}, \quad (9)$$

$$\hat{\sigma}_i^2 = \left[ \frac{\text{MAD} \{W_i^H\}}{\Phi^{-1}(\frac{3}{4})} \right]^2. \quad (10)$$

Here,  $\text{MAD} \{W_i^H\} = \text{med} \{|W_i^H - \text{med} \{W_i^H\}|\}$ , and  $\Phi^{-1}$  denotes the inverse cumulative distribution function (*c.d.f.*) of the standard normal distribution, and the constant factor  $1/\Phi^{-1}(\frac{3}{4}) = 1.4826$  makes the estimator asymptotically unbiased in case of *i.i.d.* normal samples.

When using MAD, it is important to consider that the relation (8) may fail, because of the potential discrepancy between the mean and the median of distributions that are not *i.i.d.* normal. Nevertheless, the use of the MAD estimator on  $W_i^H$  can be justified because of the *Gaussianization* of the coefficients resulting by a transformation of the type (3) [2]. We support this thesis providing, in the next section, an accurate study of the robust estimators errors in practical applications.

### B. Error Analysis

As described in the previous section, the estimation error is composed by two parts: one due to the variance of the estimators (the only one in the ideal case), and one due to the presence of outliers (e.g., edges). In this section we analyze

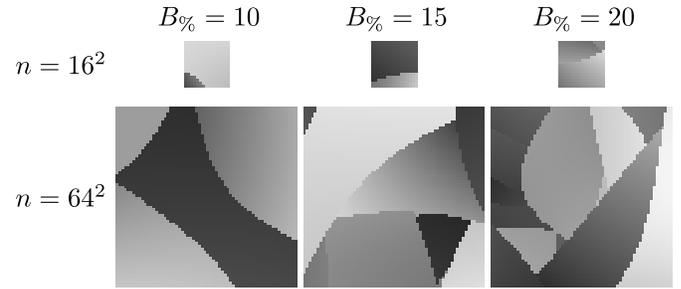


Fig. 4. Examples of the patches  $W_i^y$  used in the Monte Carlo simulation, with different block sizes  $n$  and percentages of boundaries  $B\%$ .

quantitatively how these outliers affect the computation of the pairs  $(\hat{y}_i, \hat{\sigma}_i^2)$ .

For this purpose we performed a Monte Carlo simulation where we compute the average estimation error on a pair  $(y_i, \sigma_i^2)$  from a block containing a certain amount of edges:

- for each task, a patch containing a random number of regions and corrupted by affine signal-dependent noise is created;
- the patches are then grouped depending on the percentage of edges  $B\%$  within them;
- the mean-noise variance pairs are then estimated;
- the estimation errors are computed for each block;
- finally, the errors are averaged, separately, for each group.

In this way, we compute the average estimation error in function of the amount of edges in the block.

We now describe more accurately the entire process.

1) *Patch Generation and Grouping*: we generate patches  $W_i^y$  containing a random number of regions; each region of each patch is piecewise smooth with piecewise smooth boundaries (examples are shown in Figure 4). The minimum and maximum intensity values of each region are realization of random variables uniformly distributed in  $[0, 1]$ . The patches are then grouped depending on the percentage of edges  $B\%$  within them. Every patch is corrupted by the noise defined in (2), and filtered as described in (3).<sup>2</sup> In this way we create  $W_i^z$  and  $W_i^H$ , which are used for computing  $\hat{y}_i$  and  $\hat{\sigma}_i^2$ , respectively. The noise parameters  $a$  and  $b$  are chosen, for each patch, as realization of random variables uniformly distributed respectively in  $[0, 0.002]$  and  $[0, 0.0006]$ , in order to operate on noise ranges comparable to those considered in, see [1], [2], which are representative of typical consumer camera sensors.

2) *Error Computation and Normalization*: for every patch, the estimation error  $e_i$  is computed as the distance between the point  $(\hat{y}_i, \hat{\sigma}_i^2)$  estimated with (9) and (10), and the ground-truth line  $ay + b$ , i.e. the distance between  $(\hat{y}_i, \hat{\sigma}_i^2)$  and its orthogonal projection  $(\hat{y}_{i\perp}, \hat{\sigma}_{i\perp}^2)$  on the line  $ay + b$ .

Intuitively, the estimation errors of the mean and variance are function of the noise variance that we are estimating, i.e. larger noise variance implies larger estimation error. Consequently, estimation errors on patches having the same amount of edges, but affected by different noise levels, can

<sup>2</sup>To eliminate the boundary artifacts in the computation of  $W_i^H$ , we create a bigger patch (padding) in order to discard the boundaries once the filtering is performed.

be significantly different. We normalize the square estimation error  $e_i^2$  by dividing it by the mean square error (MSE)  $\bar{e}^2(\hat{\sigma}_{i\perp}^2)$  that we would have had if we were performing the estimation on a flat patch containing only one region, and affected by constant noise variance  $\hat{\sigma}_{i\perp}^2$ . In this way, the normalized error becomes an index of the goodness of the estimation with respect to the simplest possible case, i.e. a single flat region.

Let us now show how the MSE  $\bar{e}^2(\cdot)$  depends on the noise variance  $\bar{\sigma}^2$  of a generic flat patch  $\bar{W}^z$ , denoting  $\bar{W}^H$  its filtered version:

$$\begin{aligned} \text{MSE} \left\{ \text{med} \left\{ \bar{W}^z \right\} \right\} &= \text{var} \left\{ \text{med} \left\{ \bar{W}^z \right\} \right\} \\ &= v^z \left( \bar{\sigma}^2 \right) = \frac{\pi}{2n} \bar{\sigma}^2, \end{aligned} \quad (11)$$

$$\begin{aligned} \text{MSE} \left\{ \text{MAD} \left\{ \bar{W}^H \right\} \right\} &= \text{var} \left\{ \text{MAD} \left\{ \bar{W}^H \right\} \right\} \\ &= v^H \left( \bar{\sigma}^2 \right) = \frac{\alpha}{n} \bar{\sigma}^4, \end{aligned} \quad (12)$$

$$\bar{e}^2 \left( \bar{\sigma}^2 \right) = v^z \left( \bar{\sigma}^2 \right) \sin^2 \left( \theta \right) + v^H \left( \bar{\sigma}^2 \right) \cos^2 \left( \theta \right), \quad (13)$$

where  $v^z(\bar{\sigma}^2)$  and  $v^H(\bar{\sigma}^2)$  are, respectively, the variances of the median and MAD estimators applied to the patches  $\bar{W}^z$  and  $\bar{W}^H$ , and  $\alpha$  is a constant that depends on the function<sup>3</sup>  $\psi$  that we use to filter  $\bar{W}^z$  in order to obtain  $\bar{W}^H$ . The MSEs of the estimators coincide with their variances because the patches are flat and the estimation errors have zero mean, i.e. the samples are unbiased because there are no outliers.

In (13), the terms  $\sin^2(\theta)$  and  $\cos^2(\theta)$  are used to compute the orthogonal components of (11) and (12) to the line  $ay + b$ , the only components of the variances that mislead the estimation, with  $\theta$  being the angle between the line  $ay + b$  and the horizontal axes, i.e.  $\theta = \arctan(a)$ .

We can finally define the normalized square estimation error  $\bar{e}_i^2$  as

$$\bar{e}_i^2 = \frac{e_i^2}{\bar{e}^2(\hat{\sigma}_{i\perp}^2)}. \quad (14)$$

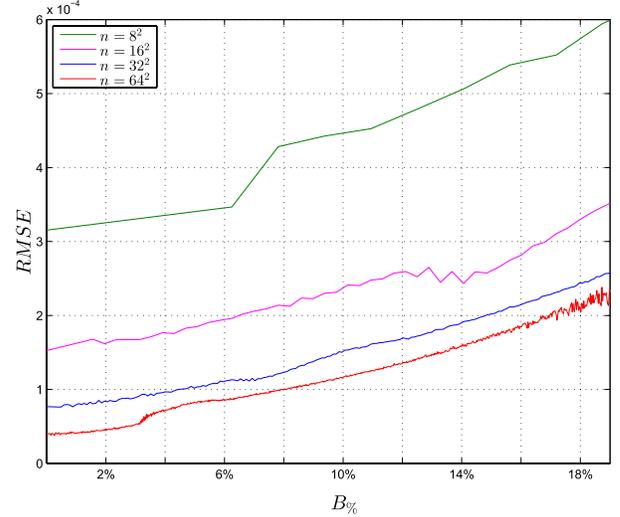
3) *Averaging and Error Trend*: Figure 5 shows the root mean square error (RMSE) and the root mean normalized square error (RMNSE) resulting from respectively averaging the estimation errors  $e_i^2$  and  $\bar{e}_i^2$  over groups of patches having the same percentage of edges  $B_{\%}$ . We separately consider four different window sizes  $n$ .

The RMNSE curves in Figure 5 are approximately monotonically increasing with common minimum 1 at  $B_{\%} = 0$ , where patches are composed of a single region and have no internal edges. Note that the patches  $W_i^y$  are piecewise smooth, and not perfectly flat as in the ideal case; nevertheless at  $B_{\%} = 0$  the RMNSE is practically 1. This means that, when  $B_{\%} = 0$ , the RMSE essentially coincides with the standard deviation of the estimators and, when  $B_{\%} > 0$ , the estimation errors increase almost entirely due to the presence of edges.

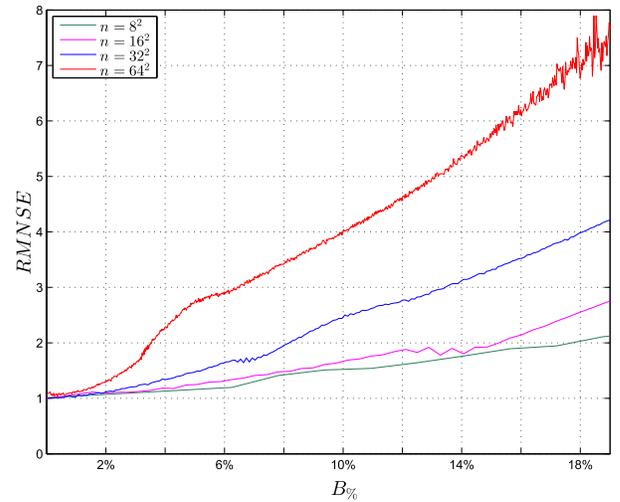
<sup>3</sup>In our experiments  $\psi$  is generated by separable convolution of one 1D Daubechies wavelet kernel,

$$\psi = \psi_{1D} \otimes \psi_{1D}^T,$$

where  $\psi_{1D} = [-0.333, 0.807, -0.460, -0.135, 0.085, 0.035]$ . For this  $\psi$ , we empirically computed  $\alpha = 9.9076$ .



(a)



(b)

Fig. 5. RMSE and RMNSE as function of the percentage of edges  $B_{\%}$  within each block, for block size  $n = 8^2, 16^2, 32^2, 64^2$ . The estimations have been performed using the robust estimators in (9) and (10). (a) Root mean square error (RMSE). (b) Root mean normalized square error (RMNSE).

#### IV. EXPERIMENTS ON CAMERA RAW IMAGES

To validate the proposed algorithm in a practical context, we apply it to raw images from a digital camera. The images are shown in the left and center columns of Figure 6 and were taken using a *Canon PowerShot S90* 10-Megapixel camera. We adjusted the exposure times in order to avoid clipping (e.g., overexposure). The pictures were acquired with various ISOs and exposure times, so to have realizations of different noise levels [14].

In the rightmost column of Figure 6, the lines estimated by the proposed prototype algorithm (continuous lines) are compared against those estimated by a state-of-the-art algorithm [1] (dashed lines), here used as reference method. This algorithm first preprocesses the image in order to detect and exclude edges and texture from the noise estimation; it then partitions the remaining image into segments of constant

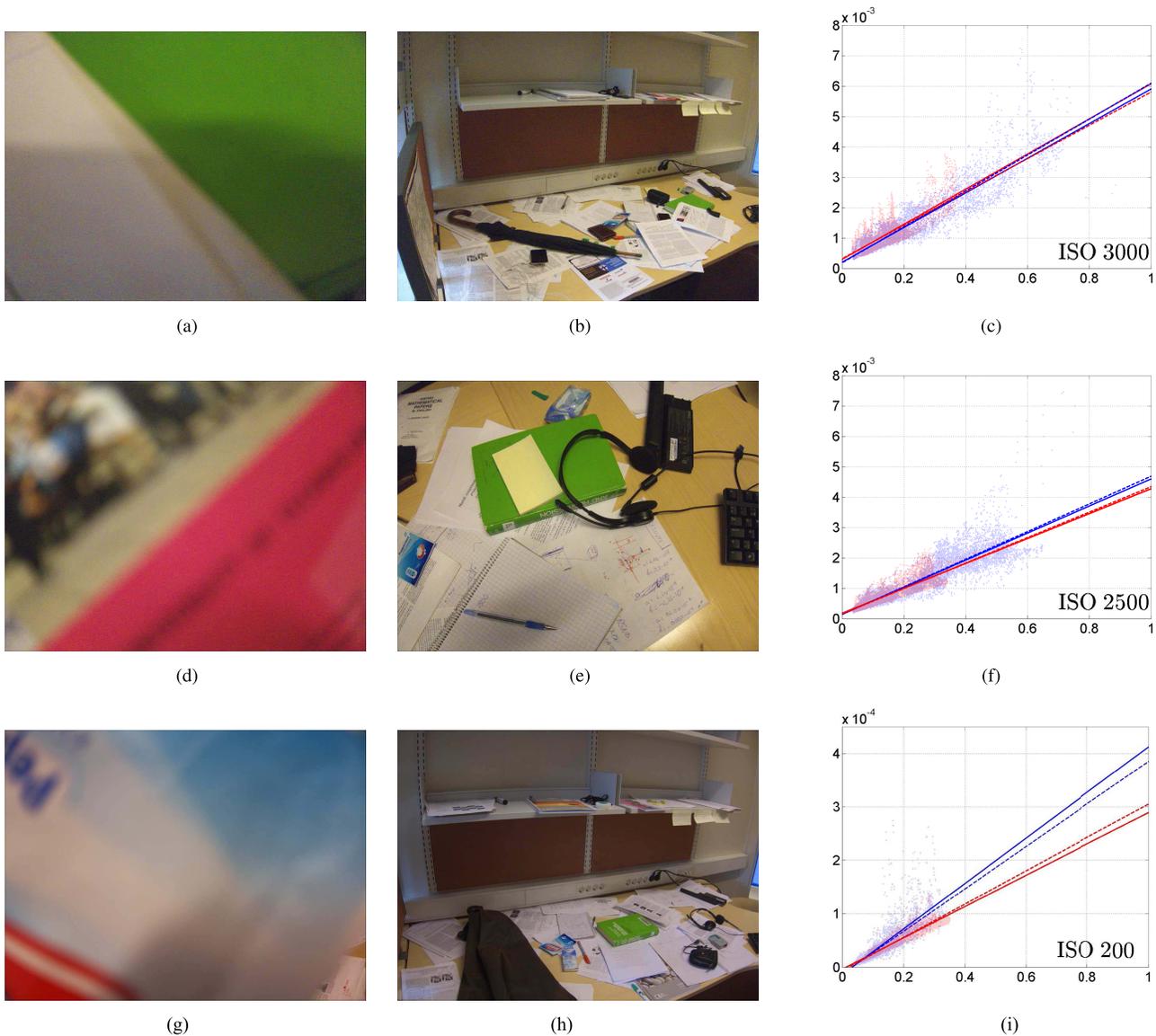


Fig. 6. Scatterplots and estimated functions for out-of-focus (red clouds and red continuous lines) and complex natural (blue clouds and blue continuous lines) images. The images have been taken with a *Canon PowerShot S90*, ISO 3200 (top row), ISO 2500 (middle row), and ISO 200 (bottom row) using exposure times respectively equal to 1/1000, 1/600, and 1/125. The estimation is performed using 2000 patches for each channel ( $[R, B; G1, G2]$ ) of size  $64 \times 64$ . The dashed lines show the functions estimated by the ref. [1]. (a) Out-of-focus image. (b) Natural image. (c) Scatterplot and fitted lines. (d) Out-of-focus image. (e) Natural image. (f) Scatterplot and fitted lines. (g) Out-of-focus image. (h) Natural image. (i) Scatterplot and fitted lines.

intensity level; a scatterplot is thus obtained by applying a robust unbiased estimator of the variance on each segment, with each point of the scatterplot being modeled according to a bivariate normal distribution; the noise model parameters  $a$  and  $b$  are finally estimated through a maximum *a posteriori* fitting. For these experiments, our prototype algorithm uses blocks of size  $64 \times 64$ , and, in order to reduce the variability of the results on the particular random choice of the block positions, 2000 patches are extracted from each color channel of the images.

In Section III-A we discussed the theoretical behavior of our method in the ideal conditions where the extracted patches are free of edges ( $B_{\%} = 0$  in Section III-B). In order to reproduce these assumptions, the raw images include 3 out-of-focus (OoF) pictures, shown in the leftmost column of Figure 6.

The lines estimated by the two algorithms (red continuous and dashed lines) are always close to each other, confirming that, in the ideal case, the proposed algorithm gives results congruent to those of the reference algorithm.

The 3 pictures of a complex natural scene, shown in the center column of Figure 6, are used to investigate the practical case. The lines estimated with the proposed algorithm (blue continuous lines) are again close to the reference ones (blue dashed lines), confirming that the proposed algorithm performs similar to the reference algorithm also on complex images.

In Figure 6, the OoF and natural pictures that are on the same row were acquired under the same operating conditions (ISO, exposure time, ambient temperature) and are hence corrupted by noise with the same parameters [14]. Therefore, the blue solid and dashed lines in each subplot may be

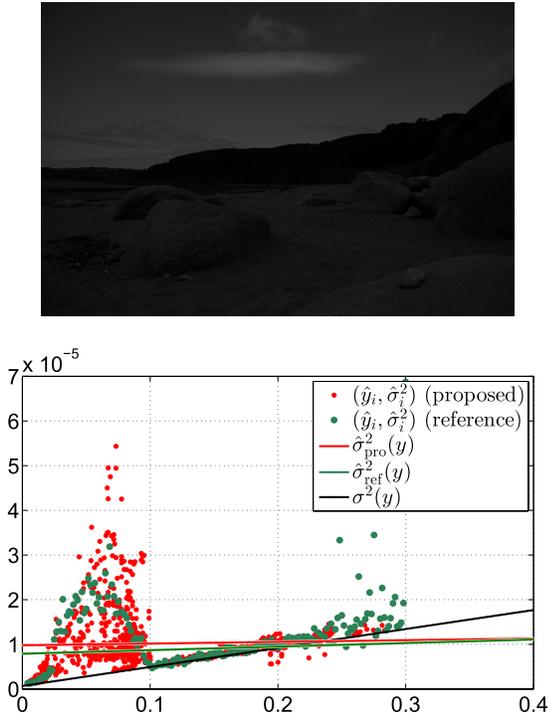


Fig. 7. Example in which both proposed and reference algorithm fail the estimation due to the presence of several outliers. Top: underexposed image with large highly textured areas from the NED dataset [19]. Bottom: scatterplot of mean-variance pairs with corresponding noise line  $\hat{\sigma}_{\text{pro}}^2(y)$  estimated by the proposed prototype algorithm (red). The result is compared with the line  $\hat{\sigma}_{\text{ref}}^2(y)$  estimated using the reference algorithm (green) and the ground-truth  $\sigma^2(y)$  (black). Due to the overwhelming presence of outliers in the scatterplot, both the proposed and the reference algorithm fail to correctly estimate the noise line.

expected to coincide with the respective red lines. Indeed, for large ISO (top and middle rows of Figure 6), the lines estimated from OoF and natural images are very close to each other, because the large noise variance makes easier for the algorithms to separate the noise from the noise-free signal. In case of small ISO (bottom row), instead, the estimation from the natural image diverges from the OoF ones, for both proposed and reference algorithms, since the variance of the noise is small with respect to the signal. The degradation of accuracy of the proposed algorithm is comparable to that of the reference one.

In Figure 7 we report the result  $\hat{\sigma}_{\text{pro}}^2(y)$  of the proposed prototype algorithm applied to an image that contains large highly textured areas. The image belongs to the NED dataset [19] of raw images with large areas of high-frequency texture, which makes noise estimation particularly challenging. The image has been captured with a *Nikon D80* at ISO 125, and the response of the sensor has been linearized by a calibrated nonlinear correction function. In the same scatterplot we also present the mean-variance pairs and the line  $\hat{\sigma}_{\text{ref}}^2(y)$  estimated with algorithm [1], and the ground-truth line  $\sigma^2(y)$  too.

Both scatterplots reveal the presence of several outliers in the intensity range  $y \in [0, 0.1]$ , mostly generated by textures present on the mountains. These outliers cause the misestimation of the lines fitted by either the proposed and the reference

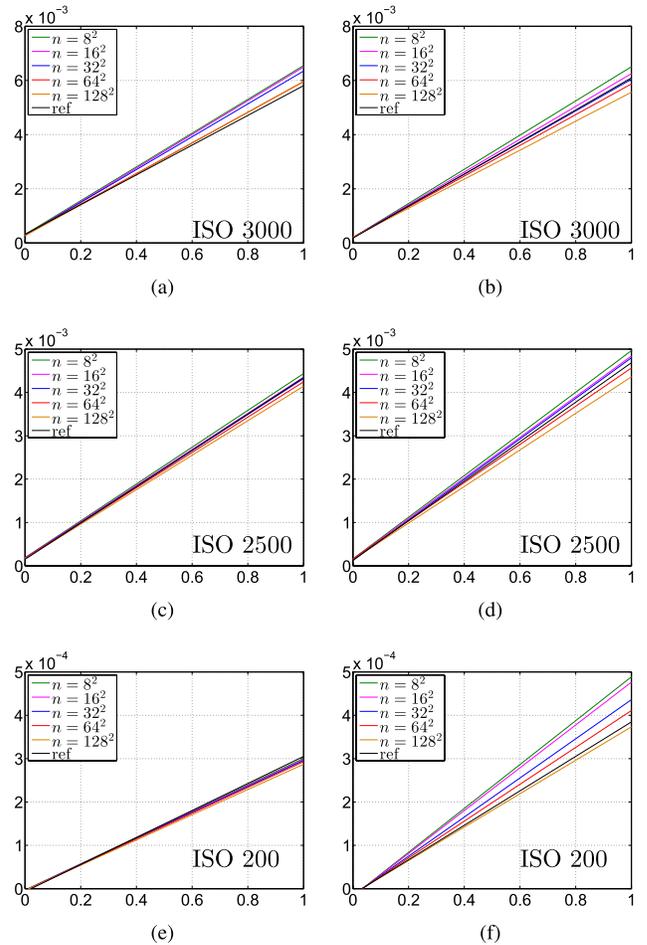


Fig. 8. Lines estimated from the images in Figure 6 separately using 2000 patches for each channel ( $R, B; G1, G2$ ) of size  $n = 8^2, 16^2, 32^2, 64^2, 128^2$ . The results are compared with the estimates of the reference algorithm. (a) Out-of-focus image. (b) Natural image. (c) Out-of-focus image. (d) Natural image. (e) Out-of-focus image. (f) Natural image.

algorithm.<sup>4</sup> This result confirms that textures and edges are the main cause of misestimation, since they affect similarly proposed and reference algorithm, and that the scatterplot points can be estimated using heterogeneous samples.

To evaluate the impact of the block size, we repeat the experiment presented in Figure 6 separately using patches of size  $n = 8^2, 16^2, 32^2, 64^2, 128^2$ ; the results are reported in Figure 8. The lines estimated from the OoF images, showed on the left column of the figure, are relatively close to each other independent of the block size. Observing the results from natural images, showed on the right column of the figure, we can notice, especially at low ISO, that estimates from larger blocks are less affected by overestimation bias.

Finally, to illustrate the essential role of the robust estimators in alleviating the bias effect due to outliers from edges and texture, we performed the same Monte Carlo simulation described in Section III-B, using the sample mean and sample variance to estimate the scatterplot points. In Figure 9 we show the RMNSE of these non-robust estimators in comparison to

<sup>4</sup>A robust variant of [1] was recently published [20] while the present article was already in press. The variant models the scatterplot points as an adaptive mixture of Gaussian and Cauchy distributions, and thus yields more accurate results in cases with outliers such as that illustrated in Figure 7.

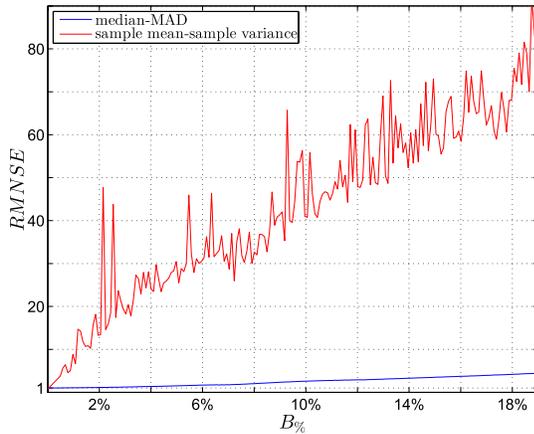


Fig. 9. Root mean normalized square error (RMNSE) of the pairs *median-MAD* and *sample mean-sample variance* for blocks of size  $n = 32^2$ . Robust estimators lead to a comparable reduction of error also for the other block sizes.

that of the med-MAD (9)–(10). Note how the robust estimators, and in particular the MAD, drastically reduce the error.

## V. DISCUSSION

In Section III-B and Section IV, we demonstrated that the average estimation error of the points  $(y_i, \sigma_i^2)$  may get smaller if larger blocks are used in conjunction with robust estimators, in spite of the fact that the samples get more heterogeneous. However, there is also an inevitable trade-off in the choice of the block size: when using large patches it is unlikely that the mean (or median)  $y_i$  reaches the extremes of the distribution of the image intensity values  $y$ . As a consequence, the scatterplot may cluster about the point  $(c, ac + b)$ ,  $c$  being the mean (or median) of  $y$  over the whole image, and, thus, the accuracy of the estimated line may be degraded. On the other hand, smaller patches allow the scatterplot points to distribute on a wider interval, at the expense of higher estimation variance for each point, and risk of larger bias on some of them. While the variance errors may cancel out through the line fitting, the bias errors will eventually corrupt the final estimate unless a robust fitting is utilized.

Our analysis and algorithm are developed and validated on the specific affine-variance model (2), and may fail for a generic non-affine  $\sigma^2(y)$ . On the other hand, the proposed algorithm still gives accurate results if  $\sigma^2$  is well approximated by a locally (i.e. separately on each block) affine function of  $y$ . On the other hand, the proposed algorithm still gives accurate results if  $\sigma^2$  is well approximated by a locally (i.e. separately on each block) affine function of  $y$ . However, in many cases (e.g., in the case of clipping) it can be difficult to verify the local affinity of  $\sigma^2$  without any strong assumptions on the image  $y$ .

Let us discuss also about ways how to possibly improve the estimation accuracy. In its prototype implementation, our algorithm is limited by the accuracy of the MAD estimator and thus cannot reach the accuracy of algorithms (e.g., [17]) that adopt more sophisticated estimators for the estimation of the variance. Likewise, the simplest LS fitting method is not robust to outliers in the scatterplot. Therefore, the use of a better

variance estimator and a better (e.g., robust) fitting algorithm [20] could further improve the estimation, so to possibly deal with highly textured images such as the example in Figure 7.

Adaptive procedures such as segmentation may be crucial for alleviating the impact of high-frequency texture on the variance estimation, but we especially emphasize that this is not a peculiarity of signal-dependent noise models, and it applies also to constant-variance (homoskedastic) noise models, including additive white Gaussian noise (AWGN). In fact, the advanced methods [17], [18] are developed for AWGN estimation. As shown in our theoretical and experimental analysis, the fact that the variance of the noise is not constant (heteroskedasticity), and depends instead on the signal, does not *per se* imply an additional need for adaptive segmentation.

Finally, let us note that the proposed model deals with the estimation of signal-dependent noise that is spatially uncorrelated, i.e. noise with diagonal covariance matrix. It is nevertheless possible to extend the proposed approach also to the correlated-noise case. If the correlation model (i.e. the shape of the noise power spectral density (PSD)) is known one can compute the noise energy in the high-pass image  $z^H$  from which the blocks  $W_i^H$  are extracted, and hence normalize the output of the variance estimator based on the product of the PSD with the spectrum of  $\psi$ . This product can be preconditioned by suitably downsampling the data prior to analyzing the noise; downsampling may be also desirable, as a means to reduce the amount of data to be processed.

## VI. CONCLUSIONS

As opposed to conventional methods that require homogeneous samples for the estimation of mean-variance pairs, our approach to signal-dependent noise estimation utilizes arbitrarily large samples of possibly heterogeneous data. The approach is backed by a Gaussian-mixture modeling, which shows that the individual mean-variance estimates computed from the heterogeneous samples are still representative of the true mean-variance curve. An elementary prototype algorithm based on this modeling is presented for the estimation of signal-dependent noise from a single image. The algorithm extracts large heterogeneous samples from random locations in the image. This corresponds to a fundamental difference versus traditional algorithms, which often involve an adaptive segmentation of the image into narrow homogeneous segments, and it also results in a simplification of the estimation procedure. This approach can be therefore suitable in all applications where a simple noise estimation algorithm is required, and which has to operate on non-intelligent devices. Experiments on real data demonstrate the reliability of the algorithm applied to natural images, showing that its results are comparable with those from a state-of-the-art method.

## REFERENCES

- [1] 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.*, vol. 17, no. 10, pp. 1737–1754, Oct. 2008.
- [2] A. Foi, "Clipped noisy images: Heteroskedastic modeling and practical denoising," *Signal Process.*, vol. 89, no. 12, pp. 2609–2629, Dec. 2009.

- [3] D. L. Donoho and J. M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *Biometrika*, vol. 81, no. 3, pp. 425–455, 1994.
- [4] N. Johnson, S. Kotz, and N. Balakrishnan, *Continuous Univariate Distributions*, vol. 1, 2nd ed. New York, NY, USA: Wiley, 1994, sec. 13.10.
- [5] G. K. Froehlich, J. F. Walkup, and R. B. Asher, "Optimal estimation in signal-dependent noise," *JOSA*, vol. 68, no. 12, pp. 1665–1672, 1978.
- [6] C. Liu, W. T. Freeman, R. Szeliski, and S. B. Kang, "Noise estimation from a single image," in *Proc. IEEE Comput. Soc. Conf. CVPR*, vol. 1, Jun. 2006, pp. 901–908.
- [7] A. Amer and E. Dubois, "Fast and reliable structure-oriented video noise estimation," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 15, no. 1, pp. 113–118, May 2005.
- [8] S. Aja-Fernández, G. Vegas-Sánchez-Ferrero, M. Martín-Fernández, and C. Alberola-López, "Automatic noise estimation in images using local statistics. Additive and multiplicative cases," *Image Vis. Comput.*, vol. 27, no. 6, pp. 756–770, May 2009.
- [9] J. S. Lee and K. Hoppel, "Noise modeling and estimation of remotely-sensed images," in *Proc. 12th Can. Symp. Remote Sensing IGARSS*, vol. 2, Jul. 1989, pp. 1005–1008.
- [10] P. Gravel, G. Beaudoin, and J. A. De Guise, "A method for modeling noise in medical images," *IEEE Trans. Med. Imag.*, vol. 23, no. 10, pp. 1221–1232, Oct. 2004.
- [11] B. Aiazzi, L. Alparone, S. Baronti, M. Selva, and L. Stefani, "Unsupervised estimation of signal-dependent CCD camera noise," *EURASIP J. Adv. Signal Process.*, vol. 2012, no. 1, pp. 1–11, 2012.
- [12] 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. Adv. Signal Process.*, vol. 2011, no. 1, p. 806516, Mar. 2011.
- [13] T. Buades, Y. Lou, J. M. Morel, and Z. Tang, "A note on multi-image denoising," in *Proc. 2009 Int. Workshop Local Non-Local Approx. Image Process.*, pp. 1–15.
- [14] P. Ojala. (2008). *Dependence of the Parameters of Digital Image Noise Model on ISO Number, Temperature and Shutter Time*. TUT/Nokia Mobile Imaging Course [Online]. Available: [http://www.cs.tut.fi/~foi/MobileImagingReport\\_PetteriOjala\\_Dec2008.pdf](http://www.cs.tut.fi/~foi/MobileImagingReport_PetteriOjala_Dec2008.pdf)
- [15] F. R. Hampel, "The influence curve and its role in robust estimation," *J. Amer. Statist. Assoc.*, vol. 69, no. 346, pp. 383–393, 1974.
- [16] F. Mosteller and J. W. Tukey, *Data Analysis and Regression: A Second Course in Statistics*. Reading, MA, USA: Addison-Wesley, 1997.
- [17] A. Danielyan and A. Foi, "Noise variance estimation in nonlocal transform domain," in *Proc. Int. Workshop Local and Non-Local Approx. Image Process.*, Tuusula, Finland, Aug. 2009, pp. 41–45.
- [18] N. N. Ponomarenko, V. V. Lukin, M. S. Zriakhov, A. Kaarna, and J. Astola, "An automatic approach to lossy compression of AVIRIS images," in *Proc. IEEE IGARSS*, Jul. 2007, pp. 472–475.
- [19] (2013, Dec.). *Image Database for Benchmarking Signal-Dependent Noise Estimation Algorithms: Ned2012* [Online]. Available: <http://rsd.khai.edu/ned2012/ned2012.php>
- [20] L. Azzari and A. Foi, "Gaussian-Cauchy mixture modeling for robust signal-dependent noise estimation," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, May 2014, pp. 5394–5398.



**Lucio Azzari** was born in Teramo, Italy, in 1985. He received the B.Sc. and M.Sc. degrees in Electronic and Telecommunication Engineering from Università degli Studi Roma Tre, Rome, Italy, in 2007 and 2010, respectively. He is a researcher at the Department of Signal Processing, Tampere University of Technology, Tampere, Finland. His work regards 3D image and video rendering, noise estimation and denoising techniques, and nonlocal algorithms. He has also collaborated with the Nokia Research Center of Hermia, Tampere, Finland. Currently he is pursuing the Ph.D. degree on Signal Processing in the Tampere University of Technology.



**Alessandro Foi** (SM'–) received the M.Sc. degree in mathematics from the Università degli Studi di Milano, Milan, Italy, in 2001, the Ph.D. degree in mathematics from Politecnico di Milano, Milan, in 2005, and the D.Sc.Tech. degree in Signal Processing from the Tampere University of Technology, Tampere, Finland, in 2007. He is currently an Academy Research Fellow (Akatemiatutkija) with the Academy of Finland, and is with the Department of Signal Processing, Tampere University of Technology, where he is also an Adjunct Professor. His research interests include mathematical and statistical methods for signal processing, functional and harmonic analysis, and computational modeling of the human visual system. His recent work focuses on spatially adaptive (anisotropic, nonlocal) algorithms for the restoration and enhancement of digital images, on noise modeling for imaging devices, and on the optimal design of statistical transformations for the stabilization, normalization, and analysis of random data. He is an Associate Editor of the IEEE TRANSACTIONS ON IMAGE PROCESSING.