

# Image Sensor Noise Parameter Estimation by Variance Stabilization and Normality Assessment

Stanislav Pyatykh and Jürgen Hesser

**Abstract**—High-quality image denoising requires taking into account the dependence of the noise distribution on the original image. The parameters of this dependence are often unknown and we propose a new method to estimate them here. Using an optimization procedure, we find a variance-stabilizing transformation, which transforms the input image into an image with signal-independent noise. Principal component analysis of blocks of the transformed image allows estimation of the variance of the signal-independent noise so that the parameters of the original noise model can be computed. The image blocks for processing are selected in such a way that they have low stochastic texture strength but preserve the noise distribution. The algorithm does not require the original image to have homogeneous areas and can accurately process images with regular textures. It has high computational efficiency and smaller maximum estimation error compared with the state of the art. Our experiments have also shown that denoising with the noise parameters estimated by this method leads to the same results as denoising with the true noise parameters.

**Index Terms**—Estimation, image processing, principal component analysis.

## I. INTRODUCTION

**M**ANY image denoising algorithms [1]–[3] take the noise parameters as input values. However, these parameters may not be known beforehand because they may depend on sensor’s operational conditions or the calibration data may not be available. As a result, blind noise parameter estimation is often necessary. Its accuracy heavily affects the performance of denoising algorithms, since some noise remains in the output image when the noise level is underestimated, whereas overestimating the noise level results in oversmoothing the output image.

Signal-independent additive white Gaussian noise (AWGN) is the most widely used noise model. It is easy to analyze and has only one parameter: the noise variance. However, this model does not accurately describe the behavior of digital image sensors. Due to improvements in the sensor technology,

the effect of signal-independent electric and thermal noise is decreasing [4], and photon noise is becoming the main component of sensor noise. Therefore, sensor noise should be considered as signal-dependent.

Compared with the large literature on signal-independent noise level estimation (see [5]–[8] and references therein), there are not so many methods for sensor noise parameter estimation; and these methods include the following steps [4], [9]–[18]:

- 1) Detection of homogeneous areas, for which the original intensity is approximately constant. Edges and regions covered by texture are detected and excluded from the further processing.
- 2) Computation of estimates of the mean and variance for each homogeneous area or for a group of homogeneous areas. Since the original image is considered as constant in each homogeneous area, the noise variance can be estimated as the variance of the input image intensities. Wavelet decomposition [4] or block discrete cosine transform [18] can be used here in order to remove possible image structures and get more accurate estimates.
- 3) Fitting the noise model to the mean and variance estimates. Various methods of cluster center determination and least squares fitting are utilized here.

These steps can be graphically represented as the construction of the scatter-plot for the mean and variance estimates, hence they are often referred to as the scatter-plot approach.

All mentioned methods require a sufficient amount of homogeneous areas in the original image. These areas should have various intensities so that the noise variance as a function of the original pixel value can be accurately estimated. Therefore, images containing mostly textures are not always accurately processed.

In this work, we propose a new parameter estimation method for digital image sensor noise, which is based on principal component analysis (PCA) of image blocks. The advantages of the proposed method are the following:

- 1) The method has smaller maximum estimation error compared with the state of the art. At the same time, it has high computational efficiency.
- 2) Images without homogeneous areas can be processed. To the best of our knowledge, this is the first algorithm designed for digital image sensor noise, which does not require the existence of homogeneous areas.

Manuscript received November 6, 2012; revised April 3, 2014; accepted July 10, 2014. Date of publication July 14, 2014; date of current version July 28, 2014. This work was supported in part by the Atlantic Innovation Fund under Grant KF2769301FRQ and in part by the German Research Foundation under Grant HE 3011/23-1 and Grant HE 3011/14-1. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Adrian G. Bors.

The authors are with the University Medical Center Mannheim, Heidelberg University, Mannheim D-68167, Germany (e-mail: stanislav.pyatykh@medma.uni-heidelberg.de; juergen.hesser@medma.uni-heidelberg.de).

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.2014.2339194

The idea of the proposed method is to apply a variance-stabilizing transformation (VST) in order to transform the noise to AWGN, and then to estimate the noise variance as the variance of the last principal component of image blocks. However, the VST depends on the noise parameters, which are unknown, therefore, a noise normality measure is utilized to select the right VST.

This article is organized as follows. We start with a brief description of digital image sensor noise model in Section II. The construction of a VST is described in Section III; and the algorithm is explained in Section IV. The results and the discussion are represented in Sections V and VI respectively. We conclude in Section VII.

## II. NOISE MODEL

Further in this work, the expected value, the standard deviation, and the variance of a random variable are denoted by  $\mathbf{E}(\cdot)$ ,  $\text{std}(\cdot)$ , and  $\text{var}(\cdot)$  respectively; and  $s^2(\cdot)$  denotes the sample variance. The entry in the  $i$ -th row and  $j$ -th column of a matrix  $\mathbf{A}$  is referred to as  $\mathbf{A}(i, j)$ ; and  $\mathbf{v}^T$  denotes the transpose of a vector  $\mathbf{v}$ .

Let  $\mathbf{p} \in \mathbb{Z}^2$  be a pixel location,  $x = x(\mathbf{p})$  be the original (noise-free and unknown) image, and  $y = y(\mathbf{p})$  be the corresponding noisy image, which is acquired by the sensor. For a function  $f$  of a real variable,  $f(y)$  denotes the image with value  $f(y(\mathbf{p}))$  at position  $\mathbf{p}$ .

Let  $\omega(\mathbf{p}) \sim \text{Poisson}(\lambda(\mathbf{p}))$  be the number of detected photons,  $\sqrt{b}\xi(\mathbf{p}) \sim \mathcal{N}(0, b)$  be signal-independent electric and thermal noise added to the signal by the sensor's hardware, and  $a > 0$  be a multiplicative factor which depends on the sensor's quantum efficiency and the analog gain [4]. Then,

$$x(\mathbf{p}) = a\lambda(\mathbf{p}) \quad (1)$$

$$y(\mathbf{p}) = a\omega(\mathbf{p}) + \sqrt{b}\xi(\mathbf{p}). \quad (2)$$

Since  $\xi$  is signal-independent,

$$\text{var}(y(\mathbf{p})) = a^2\lambda(\mathbf{p}) + b \stackrel{(1)}{=} ax(\mathbf{p}) + b. \quad (3)$$

Therefore, the noise variance linearly depends on the original pixel value.

For sufficiently large  $\lambda(\mathbf{p})$ ,  $\omega(\mathbf{p})$  can be approximated by a normal random variable with mean  $\lambda(\mathbf{p})$  and variance  $\lambda(\mathbf{p})$  [4]. In this case, from (1) and (2),  $y(\mathbf{p})$  is approximately normally distributed with mean  $x(\mathbf{p})$  and variance  $ax(\mathbf{p}) + b$ . Hence the sensor noise can be approximated by additive Gaussian noise with signal-dependent variance:

$$y(\mathbf{p}) \approx x(\mathbf{p}) + \sqrt{ax(\mathbf{p}) + b} \xi(\mathbf{p}). \quad (4)$$

The image sensor noise parameter estimation problem is then the problem of estimation parameters  $a$  and  $b$  given noisy image  $y$ , which is assumed to follow model (2).

## III. VARIANCE-STABILIZING TRANSFORMATION

A VST of random variable  $y(\mathbf{p})$ , whose standard deviation depends on  $x(\mathbf{p})$ , is function  $f(\cdot; a, b)$  such that the standard deviation of transformed random variable  $f(y(\mathbf{p}); a, b)$  is independent of  $x(\mathbf{p})$ , i.e.

$$\text{std}(f(y(\mathbf{p}); a, b)) = \sigma \quad (5)$$

where  $\sigma$  is independent of  $x(\mathbf{p})$ .

Using the first-order Taylor expansion of  $f$  around the mean value

$$f(y(\mathbf{p}); a, b) \approx f(x(\mathbf{p}); a, b) + f'(x(\mathbf{p}); a, b)(y(\mathbf{p}) - x(\mathbf{p})) \quad (6)$$

(5) can be approximated as

$$f'(x(\mathbf{p}); a, b) \cdot \text{std}(y(\mathbf{p})) = \sigma \quad (7)$$

or

$$f'(x(\mathbf{p}); a, b) = \frac{\sigma}{\text{std}(y(\mathbf{p}))} = \frac{\sigma}{\sqrt{ax(\mathbf{p}) + b}} \quad (8)$$

Integrating the last equality, we have

$$f(t; a, b) = \frac{2\sigma}{a} \sqrt{at + b}. \quad (9)$$

The VST (9) is a smooth function and  $y(\mathbf{p})$  is approximately normally distributed; hence, if  $y(\mathbf{p})$  has a sufficiently small variance, the transformed pixel intensity  $f(y(\mathbf{p}); a, b)$  is approximately normally distributed as well. Additionally,  $\text{std}(f(y(\mathbf{p}); a, b)) \approx \sigma$  for all  $\mathbf{p}$ . Therefore, the transformed pixel intensities  $f(y(\mathbf{p}); a, b)$  can be considered as independent identically distributed random variables drawn from a normal distribution with variance  $\sigma^2$ , i.e. the noise in the transformed image  $f(y; a, b)$  can be approximated by AWGN.

In order to validate this conclusion, we utilized Q-Q plots [19]. A Q-Q plot can be used to assess normality of the noise by making a plot of the noise distribution against the normal distribution with the same mean and variance. Denoting the cumulative distribution functions (CDFs) of the noise distribution and the normal distribution by  $\hat{F}$  and  $F$  respectively, the Q-Q plot can be constructed as the parametric curve

$$(F^{-1}(t), \hat{F}^{-1}(t)) \quad t \in [0, 1]. \quad (10)$$

The closer this curve lies to the 45° line, the closer the noise distribution is to the normal distribution.

In order to construct the Q-Q plots for images  $y$  and  $f(y; a, b)$ , we took a  $2048 \times 2048$  image, whose grayvalues were uniformly distributed in the range  $0, \dots, 4095$ , as the original image  $x$ . (The range  $0, \dots, 4095$  corresponds to the grayvalue range of a 12-bit sensor.) Then, the noisy image  $y$  was generated according to Poisson-Gaussian model (2); and the Q-Q plots for the noise distribution in images  $y$  and  $f(y; a, b)$  were constructed. For high-quality sensors, the level  $b$  of signal-independent electric and thermal noise remains low. At the same time, the level  $a$  of signal-dependent noise component depends on the imaging conditions and can be arbitrary large. Therefore, we fixed  $b$  to 25 and varied  $a$  from 10 to 400.

The Q-Q plots are presented in Fig. 1. The noise in image  $y$  always has longer tails compared with the normal distribution. Regarding the noise in image  $f(y; a, b)$ , it can be approximated by the normal distribution very well if  $a$  is not greater than 20. For larger  $a$ , the tails of the Q-Q plot lie considerably below the 45° line, which indicates that the noise distribution is skewed to the left. This can be explained by the fact that linear approximation around  $x(\mathbf{p})$  (6) is not

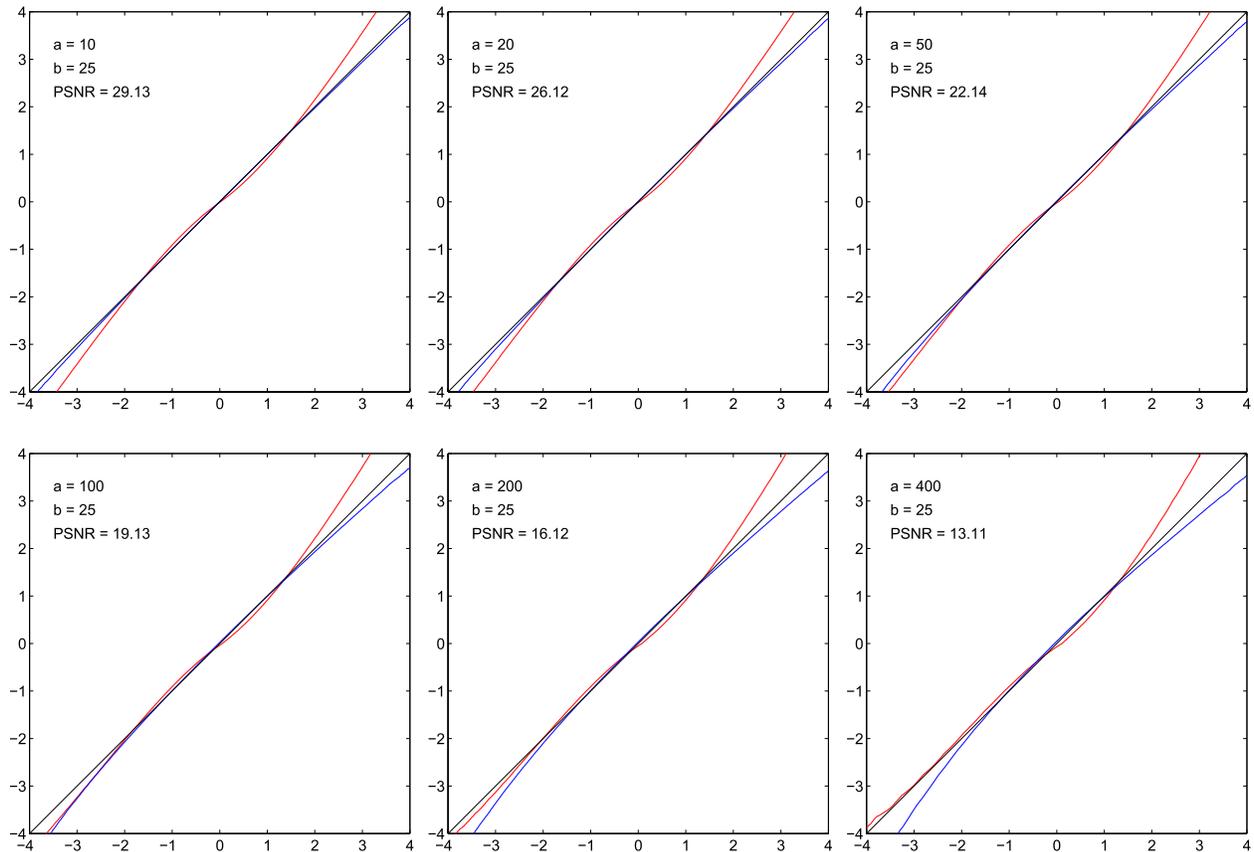


Fig. 1. The Q-Q plots for the noise distribution in image  $y$  (red curve) and in image  $f(y; a, b)$  (blue curve). The PSNR is given in decibels. The black line is the  $45^\circ$  line.

sufficiently accurate if  $y(\mathbf{p})$  is far from  $x(\mathbf{p})$ , i.e. when the variance of  $y(\mathbf{p})$  is large. Nevertheless, we assume further in the paper that the transformed image  $f(y; a, b)$  has AWGN, which reflects the behavior for sufficiently small  $a$ .

The Matlab code to generate the plots from Fig. 1 is available at <http://physics.medma.uni-heidelberg.de/cms/projects/132-pcanle>.

#### IV. NOISE PARAMETER ESTIMATION

##### A. Transform of Image Blocks

Since we assume that the transformed image  $f(y; a, b)$  has AWGN with standard deviation  $\sigma$ , this image can be considered as the result of corruption of some noise-free image  $z$  with AWGN. In this case, image  $z$  is the expected value of image  $f(y; a, b)$ :

$$z = \mathbf{E}(f(y; a, b)). \quad (11)$$

Let  $\mathbf{p}_1, \dots, \mathbf{p}_N$  be pixel locations;  $\mathbf{M} \in \{0, 1\}^{B \times B}$  be a  $B \times B$  binary matrix representing a pixel mask; and  $K$  be the number of non-zero elements in this matrix. By shifting the top-left corner of matrix  $\mathbf{M}$  to locations  $\mathbf{p}_1, \dots, \mathbf{p}_N$ , selecting pixel values below the non-zero elements of this matrix, and rearranging these values into vectors of size  $K$ , we can construct  $N$  vectors from image  $z$  and  $N$  vectors from image  $f(y; a, b)$ . Let us denote these vectors  $\mathbf{u}_1, \dots, \mathbf{u}_N$  and  $\mathbf{v}_1, \dots, \mathbf{v}_N$  respectively.

In order to be able to separate the noise from the signal, one has to assume that the information in image  $z$  is redundant, i.e. the intrinsic dimension of vectors  $\mathbf{u}_1, \dots, \mathbf{u}_N$  is smaller than the number of coordinates  $K$ . In order to utilize this redundancy, a transform, which can represent each of these vectors with less than  $K$  coefficients, should be applied. We suggest using PCA as transform, because it is data-adaptive and can separate noise not only in homogeneous areas but also in regular textures.

PCA is applied to the transformed noisy image  $f(y; a, b)$  corrupted with AWGN with variance  $\sigma^2$  as follows [20]:

- 1) The mean of vectors  $\mathbf{v}_1, \dots, \mathbf{v}_N$  is computed:

$$\bar{\mathbf{v}} = \frac{1}{N} \sum_{i=1}^N \mathbf{v}_i. \quad (12)$$

- 2) The sample covariance matrix of  $\mathbf{v}_1, \dots, \mathbf{v}_N$  is constructed:

$$\mathbf{S} = \frac{1}{N-1} \sum_{i=1}^N (\mathbf{v}_i - \bar{\mathbf{v}})(\mathbf{v}_i - \bar{\mathbf{v}})^T. \quad (13)$$

- 3) The normalized eigenvectors  $\mathbf{a}_1, \dots, \mathbf{a}_K$  of matrix  $\mathbf{S}$  are calculated. These vectors form an orthonormal basis and obey the relation

$$s^2(\mathbf{a}_1^T \mathbf{v}_i) \geq s^2(\mathbf{a}_2^T \mathbf{v}_i) \geq \dots \geq s^2(\mathbf{a}_K^T \mathbf{v}_i) \quad (14)$$

where the sample variances  $s^2(\cdot)$  are computed over  $i$ .

4) Scores  $w_{k,i}$ ,  $k = 1, \dots, K$ ,  $i = 1, \dots, N$ , are calculated:

$$w_{k,i} = \mathbf{a}_k^T (\mathbf{v}_i - \bar{\mathbf{v}}) \quad (15)$$

i.e.  $w_{k,i}$  is the  $k$ -th coordinate of the centered vector  $(\mathbf{v}_i - \bar{\mathbf{v}})$  in basis  $\mathbf{a}_1, \dots, \mathbf{a}_K$ . Let us further denote the score set  $w_{k,1}, \dots, w_{k,N}$  by  $w_k$ ,  $k = 1, \dots, K$ ; and the set of all scores  $w_{k,i}$  by  $w$ .

Since the distribution of the noise vectors  $(\mathbf{v}_i - \mathbf{u}_i)$  is multivariate Gaussian with the covariance matrix  $\sigma^2 \mathbf{I}$ , we have<sup>1</sup>:

$$\begin{aligned} s^2(\mathbf{a}_1^T \mathbf{v}_i) &\approx s^2(\mathbf{a}_1^T \mathbf{u}_i) + \sigma^2 \\ &\dots \\ s^2(\mathbf{a}_K^T \mathbf{v}_i) &\approx s^2(\mathbf{a}_K^T \mathbf{u}_i) + \sigma^2 \end{aligned} \quad (16)$$

where the sample variances are computed over  $i$ . We may remark that variances  $s^2(\mathbf{a}_k^T \mathbf{v}_i)$  equal the eigenvalues of the matrix  $\mathbf{S}$ .

If PCA can exploit the redundancy of image  $z$ , vectors  $\mathbf{u}_1, \dots, \mathbf{u}_N$  are represented only by the first  $m < K$  eigenvectors  $\mathbf{a}_1, \dots, \mathbf{a}_m$  and are orthogonal to the last eigenvector  $\mathbf{a}_K$ . Therefore, scores  $w_K$  are not affected by the image content and their distribution equals the noise distribution [8]. This fact provides a way to analyze the noise distribution by analyzing the distribution of scores  $w_K$ . In particular, since  $s^2(\mathbf{a}_K^T \mathbf{u}_i) = 0$ , the noise variance can be estimated as the score variance (see (16)):

$$s^2(w_K) \approx \sigma^2. \quad (17)$$

### B. Measurement of the Stochastic Texture Contribution

The original image can contain stochastic textures, which cannot be distinguished from noise by PCA. Hence it is necessary to select blocks with the smallest stochastic texture strength for processing in order to have minimal influence of the stochastic texture on the result, i.e. all image blocks should be sorted by the stochastic texture strength.

One way to carry it out is to separate all scores into two groups: the scores used for noise distribution analysis and the scores used to measure the texture strength. For example, since scores  $w_K$  are utilized for noise distribution analysis, scores  $w_{K'}, \dots, w_{K-1}$  can be used to measure the texture strength for some  $K' < K$ . However, scores  $w_{K',i}, \dots, w_{K,i}$  are computed from the same vector  $\mathbf{v}_i$  and are statistically dependent. Therefore, skipping blocks with high texture strength using scores  $w_{K'}, \dots, w_{K-1}$  changes the distribution of  $w_K$ , which makes this strategy incorrect.

For this reason, we use different pixel masks for noise distribution analysis and for texture strength measurement. Let  $\mathbf{M}_{tex}$  and  $\mathbf{M}_{dist}$  be binary pixel mask matrices of size  $B \times B$ , which form complimentary stripe patterns:

$$\begin{aligned} \mathbf{M}_{tex}(p, q) &= ((-1)^q + 1)/2 \\ \mathbf{M}_{dist}(p, q) &= 1 - \mathbf{M}_{tex}(p, q) \end{aligned} \quad (18)$$

<sup>1</sup>The equalities are approximate because of the finite number of vectors  $N$ . The variance of the difference between the left and the right side is inverse proportional to  $N$  so that it does not influence the results when  $N$  is sufficiently large.

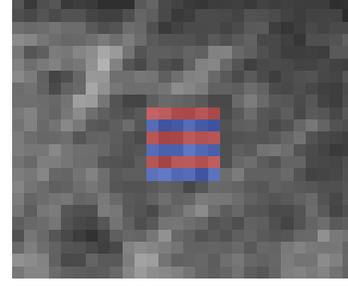


Fig. 2.  $6 \times 6$  pixel masks  $\mathbf{M}_{tex}$  (blue) and  $\mathbf{M}_{dist}$  (red) overlaid with a part of the top-left image in Fig. 4. Since the texture properties change only slightly within the block, the scores  $w^{(tex)}$  computed from the blue pixels are used to estimate the stochastic texture strength in the red pixels.

where  $p, q = 1, \dots, B$ . Then, applying the block selection and PCA as described in Section IV-A for each of masks  $\mathbf{M}_{tex}$  and  $\mathbf{M}_{dist}$ , we get two score sets  $w^{(tex)}$  and  $w^{(dist)}$  respectively. Scores  $w^{(tex)}$  are utilized to measure the stochastic texture strength, whereas scores  $w^{(dist)}$  are used for noise distribution analysis. For each  $i$ , scores  $w_{1,i}^{(tex)}, \dots, w_{K,i}^{(tex)}$  and scores  $w_{1,i}^{(dist)}, \dots, w_{K,i}^{(dist)}$  are independent, so that skipping blocks using scores  $w^{(tex)}$  does not influence the distribution of  $w_K^{(dist)}$ . At the same time, the strength of stochastic texture in scores  $w_K^{(dist)}$  can be estimated using scores  $w^{(tex)}$  very well, because the stochastic texture strength changes gradually and matrices  $\mathbf{M}_{tex}$  and  $\mathbf{M}_{dist}$  have the interleaving pattern (see Fig. 2).

Regular structures, which can be compactly represented by PCA, are stored in the first several score sets  $w_1^{(tex)}, \dots, w_{K'}^{(tex)}$ ,  $K' \leq K$ . On the other hand, stochastic textures cannot be compactly represented and affect all scores  $w_1^{(tex)}, \dots, w_K^{(tex)}$ . Hence we use the last several score sets in order to sort blocks by the stochastic texture strength. More specifically, image blocks are sorted by the following measure:

$$T(w^{(tex)}, i) = \sum_{k=K'}^K (w_{k,i}^{(tex)})^2 \quad (19)$$

where  $K' \leq K$  is a predefined parameter.  $T(w^{(tex)}, i)$  contains the contribution of noise as well but, since the noise in image  $f(y; a, b)$  is assumed to be AWGN, this contribution is approximately the same for all blocks so that it has almost no influence on the order of sorted blocks.

### C. Noise Normality Assessment

Since the true noise parameters are unknown, the VST parameters used in (9) are selected by the algorithm, and they can differ from the true noise parameters. In this case, VST (9) may not stabilize the variance and the noise distribution may deviate from a normal distribution. Therefore, we measure the normality of the noise distribution in order to assess the accuracy of the selected VST parameters.

Specifically, we utilize the excess kurtosis to measure the noise normality. The excess kurtosis of random variable  $X$  is

defined as

$$\gamma_X = \frac{\mathbf{E}((X - \mathbf{E}(X))^4)}{\mathbf{E}((X - \mathbf{E}(X))^2)^2} - 3. \quad (20)$$

$\gamma_X$  is zero when  $X$  is normal, i.e. vanishing excess kurtosis is a necessary condition for the noise normality. Let us now show that, for noise model (4), it is also an approximate sufficient condition for the noise normality. Let  $x_1 < \dots < x_M$  be pixel values in the original image  $x$ , and  $h_1, \dots, h_M$  be their probabilities. Suppose that VST (9) is applied with some parameters  $a'$  and  $b'$ , which are not necessary equal to the true noise parameters  $a$  and  $b$ . Using the first-order Taylor expansion of  $f$  around  $x(\mathbf{p})$  as in Section III,

$$\begin{aligned} \text{std}(f(y(\mathbf{p}); a', b')) &\approx f'(y(\mathbf{p}); a', b') \cdot \text{std}(y(\mathbf{p})) \\ &= \sigma \frac{\sqrt{ax(\mathbf{p}) + b}}{\sqrt{a'x(\mathbf{p}) + b'}}. \end{aligned} \quad (21)$$

Let us denote

$$\sigma_i^2 = \sigma^2 \frac{ax_i + b}{a'x_i + b'} \quad i = 1, \dots, M. \quad (22)$$

As follows from (21), the intensity of a pixel, which has value  $x_i$  in the original image, has normal distribution with variance  $\sigma_i^2$  in image  $f(y; a', b')$ . Therefore, the noise distribution in image  $f(y; a', b')$  can be represented as the mixture of the normal distributions  $\mathcal{N}(0, \sigma_i^2)$  with weights  $h_i$ . The excess kurtosis of the noise distribution can be then written as [21, p. 11]:

$$\gamma = 3 \frac{\sum_{i=1}^M h_i \sigma_i^4}{(\sum_{i=1}^M h_i \sigma_i^2)^2} - 3. \quad (23)$$

As follows from Jensen's inequality<sup>2</sup>, the excess kurtosis (23) is always non-negative, and it equals zero if and only if all  $\sigma_i$  are equal to each other, i.e. if parameters  $a'$  and  $b'$  are proportional to  $a$  and  $b$ :

$$\frac{a}{a'} = \frac{b}{b'}. \quad (24)$$

But in this case,  $\text{std}(f(y(\mathbf{p}); a', b'))$  (21) is independent of  $x(\mathbf{p})$ , i.e. the noise in the transformed image  $f(y; a', b')$  is AWGN. As a result, vanishing  $\gamma$  is a necessary and approximate sufficient condition<sup>3</sup> for the noise to be AWGN; and the deviation of  $\gamma$  from zero can be used to assess the deviation of the noise from normality.

Speaking of the calculation of the excess kurtosis for finite samples, the sample excess kurtosis  $G(X_i)$  for a sample  $X_1, \dots, X_N$  is computed by using the sample mean instead of the expected value in (20). If  $X_1, \dots, X_N$  have a normal distribution, the sample excess kurtosis multiplied by factor  $\sqrt{N/24}$  is asymptotically normally distributed [22]:

$$G(X_i)\sqrt{N/24} \xrightarrow{D} \mathcal{N}(0, 1) \quad (25)$$

<sup>2</sup>Jensen's inequality states that, for a strictly convex function  $\psi$ ,  $\psi(\sum_{i=1}^M h_i t_i) \leq \sum_{i=1}^M h_i \psi(t_i)$  and equality occurs only for  $t_1 = t_2 = \dots = t_M$ . It can be applied to (23) by substituting  $\psi(t) = t^2$  and  $t_i = \sigma_i^2$ .

<sup>3</sup>It is only an approximate sufficient condition because it is based on the first-order Taylor expansion of  $f$ .

which means that the condition

$$G(X_i)\sqrt{N/24} < T_\gamma \quad (26)$$

for some threshold  $T_\gamma > 0$  can be utilized to test that  $\gamma_X$  is zero provided that  $\gamma_X$  is always non-negative.

#### D. Parameter Transform

As follows from the previous section, the VST parameters can be found by minimizing the excess kurtosis of the noise distribution. In order to make the minimization efficient, the noise parameters  $(a, b)$  are transformed into parameters  $(\sigma, \phi)$  in such a way that

- 1) The noise in image  $f(y; a(\sigma, \phi), b(\sigma, \phi))$  is AWGN only for one value of  $\phi$ , i.e. the objective function has a single global minimum.
- 2) Parameter  $\phi$  has a bounded domain. This allows application of efficient numerical minimization methods.

Let us consider the polar coordinates:

$$\begin{aligned} a &= \sigma^2 \cos \phi \\ b &= \sigma^2 \sin \phi \end{aligned} \quad \sigma > 0, \quad \phi \in (0, \pi/2). \quad (27)$$

Then, for

$$\begin{aligned} a' &= (\sigma')^2 \cos \phi' \\ b' &= (\sigma')^2 \sin \phi' \end{aligned} \quad \sigma' > 0, \quad \phi' \in (0, \pi/2) \quad (28)$$

we have that

$$\begin{aligned} (24) &\Leftrightarrow \frac{\cos \phi}{\cos \phi'} = \frac{\sin \phi}{\sin \phi'} \\ &\Leftrightarrow \sin(\phi' - \phi) = 0 \\ &\Leftrightarrow \phi' = \phi. \end{aligned} \quad (29)$$

Therefore, the noise in image  $f(y; \sigma^2 \cos \phi, \sigma^2 \sin \phi)$  is AWGN only for a single value of  $\phi$ ; and, as the domain of  $\phi$  is bounded, the transform to polar coordinates is a suitable parameter transform according to the conditions above.

Substituting the polar coordinates into VST (9), we have that the VST is independent of  $\sigma$ , so we denote it further as function  $g(\cdot; \phi)$ :

$$\begin{aligned} f(t; \sigma^2 \cos \phi, \sigma^2 \sin \phi) &= \frac{2\sigma}{\sigma^2 \cos \phi} \sqrt{t\sigma^2 \cos \phi + \sigma^2 \sin \phi} \\ &= \frac{2}{\cos \phi} \sqrt{t \cos \phi + \sin \phi} \\ &\stackrel{\text{def}}{=} g(t; \phi). \end{aligned} \quad (30)$$

#### E. Algorithm

Our noise parameter estimation method is presented in Algorithm 1. The set  $\{\mathbf{p}_i\}_{i=1}^N$  denotes the set of all possible overlapping block positions. The score sets are presented as functions of the image and the block position set, i.e. notation  $w_k^{(dist)}(g(y; \phi), P)$  means that scores  $w_k^{(dist)}$  are computed from image  $g(y; \phi)$  using block position set  $P$ .

The noise parameters should be estimated from those blocks, which do not contain stochastic textures. On the other hand, we need the noise parameters in order to compute the stochastic texture strength in each block. Therefore, we apply an iterative procedure, in which noise parameter estimation

**Algorithm 1** Estimate Noise Parameters**Input:** image  $y$ **Output:** noise parameter estimates  $(a_{est}, b_{est})$ 

```

1:  $n \leftarrow 0$ 
2:  $(\sigma_0, \phi_0) \leftarrow (0, 0)$ 
3: while  $n < 2$  or  $|\phi_n - \phi_k| > 10^{-3}$ ,  $k = 1, \dots, n - 1$  do
4:    $a_{est} = \sigma_n^2 \cos \phi_n$ 
5:    $b_{est} = \sigma_n^2 \sin \phi_n$ 
6:    $w^{(tex)} \leftarrow w^{(tex)}(g(y; \phi_n), \{\mathbf{p}_i\}_{i=1}^N)$ 
7:    $\tau \leftarrow \text{Sort}(T(w^{(tex)}, 1), \dots, T(w^{(tex)}, N))$ 
8:    $n \leftarrow n + 1$ 
9:    $N' \leftarrow N_{min}$ 
10:  while  $N' \leq N$  do
11:     $P \leftarrow \{\mathbf{p}_{\tau(i)}\}_{i=1}^{N'}$ 
12:     $\phi^*, G^* \leftarrow \min_{\phi} G(w_K^{(dist)}(g(y; \phi), P))$ 
13:    if  $G^* \sqrt{N'/24} < T_{\gamma}$  or  $N' = N_{min}$  then
14:       $\phi_n \leftarrow \phi^*$ 
15:       $\sigma_n^2 \leftarrow s^2(w_K^{(dist)}(g(y; \phi^*), P))$ 
16:    else
17:      break
18:    end if
19:     $N' \leftarrow N' + N_{step}$ 
20:  end while
21: end while

```

and the computation of the stochastic texture strength are interleaved. The algorithm starts with  $\phi_0 = 0$ , which corresponds to pure Poisson noise; and a new estimate  $(\sigma_n, \phi_n)$  is computed at each iteration of the outer loop (lines 3–21). The outer loop stops when convergence is reached, namely, when  $\phi_n$  equals one of the previous estimates of  $\phi$  with tolerance  $10^{-3}$ . During our experiments with real images, at most 9 iterations were required for convergence.

Output estimates  $a_{est}$  and  $b_{est}$  are updated in lines 4–5 according to (27). In lines 6–7, the stochastic texture strength for all blocks is computed as described in Section IV-B. In line 7, the blocks are sorted in ascending order by the stochastic texture strength; and the permutation containing the new order of the blocks is stored in  $\tau$ . Then,  $N'$  blocks with the smallest stochastic texture strength are selected in line 11.

In line 12, noise parameter  $\phi^*$  and corresponding excess kurtosis  $G^*$  are found by minimization of the excess kurtosis of  $w_K^{(dist)}$ . The minimization is done by the golden section search in the interval  $(0, \pi/2)$ . An example of the objective function plot is shown in Fig. 3. As one can see, the function is unimodal, so that the golden section search is guaranteed to find the global minimum.

In line 13, we test if the noise in the image  $g(y; \phi^*)$  is AWGN by checking condition (26) for  $G^*$ . If (26) holds, we estimate the noise variance according to (17) in line 15. The same is done when  $N = N'$  in order to guarantee that  $\phi_n$  and  $\sigma_n$  are valid. If (26) does not hold, we break the loop (line 17), because if texture and noise separation is not possible for the current block set, it will be also impossible for a larger block set.

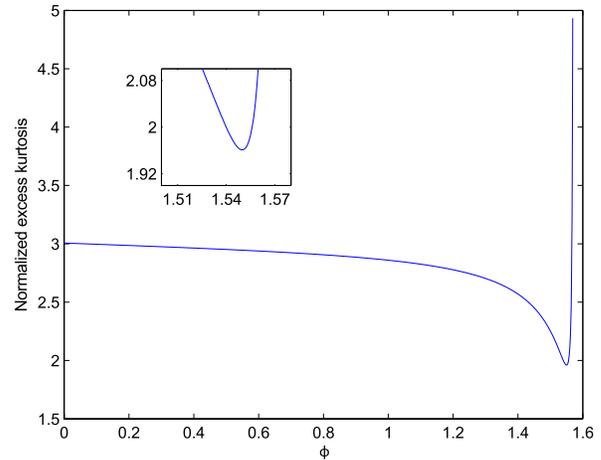


Fig. 3. Normalized excess kurtosis  $G(w_K^{(dist)}(g(y; \phi), P))\sqrt{N'/24}$  for the bottom-right image in Fig. 4 at the first iteration of the outer loop. The number of blocks  $N'$  equals 30000. The minimum is achieved at  $\phi = 1.550$ . The true value of  $\phi$  is 1.544.

TABLE I  
ALGORITHM PARAMETERS

Parameter	Denotation	Value
Block size	$B$	6
Excess kurtosis threshold in (26)	$T_{\gamma}$	3
Minimum block count	$N_{min}$	5000
Block count step	$N_{step}$	25000
Starting index in (19)	$K'$	4

The suggested values of the algorithm parameters  $B$ ,  $T_{\gamma}$ ,  $N_{min}$ ,  $N_{step}$ ,  $K'$  are listed in Table I. Vector dimension  $K$  is not a free parameter and equals  $(B \times B)/2$ , because each of masks  $\mathbf{M}_{tex}$  and  $\mathbf{M}_{dist}$  contains a half of pixels of a  $B \times B$  block. Block size  $B$  should be as large as the textural pattern in order to find correlations between pixels of the image texture and separate it from the noise. On the contrary,  $s^2(w_K^{(dist)})$  has a negative bias, which increases with  $K$ , because scores  $w_K^{(dist)}$  have the smallest sample variance by construction of PCA. Therefore, the block size should be kept small. We have found  $B = 6$  to be a good compromise. Threshold  $T_{\gamma}$  equals to three standard deviations of the normalized sample excess kurtosis (25). Parameter  $N_{min}$  corresponds to the minimal number of blocks, for which the results are statistically significant. Score set index  $K'$  should be small enough to make the sum (19) contain a sufficient number of items and be stable to noise. At the same time, the first several score sets  $w_k^{(tex)}$ , which contain the most of the image content, should not be considered.  $K' = 4$  provides a good trade-off.

The Matlab and C++ implementations of the algorithm are available at <http://physics.medma.uni-heidelberg.de/cms/projects/132-panle>.

## V. EXPERIMENTS

### A. Measurement of the Estimation Accuracy

Further in the experiment descriptions, let us denote the true noise parameters by  $a$  and  $b$ , and their estimates by  $a_{est}$  and  $b_{est}$ .



Fig. 4. Images from the NED2012 database.

The relative errors of  $a_{est}$  and  $b_{est}$  do not always provide a good measure of the accuracy, because the estimation errors for  $a$  and  $b$  can compensate each other. For example, for constant image  $x(\mathbf{p}) = x_0$ , all estimates satisfying the equality

$$a_{est}x_0 + b_{est} = ax_0 + b \quad (31)$$

are accurate.

For this reason, we define seminorm  $\rho_x$  of the standard deviation function  $h(t; a, b) = \sqrt{at + b}$ :

$$\rho_x(h(\cdot; a, b)) = \sqrt{\sum_{\mathbf{p}} h^2(x(\mathbf{p}); a, b)}. \quad (32)$$

This seminorm puts more weight to the grayvalues which occur often in image  $x$ . Then, we compute the estimation error using the following relation of seminorms, which mimics the relative error:

$$\begin{aligned} \delta &= \frac{\rho_x(h(\cdot; a_{est}, b_{est}) - h(\cdot; a, b))}{\rho_x(h(\cdot; a, b))} \\ &= \left( \frac{\sum_{\mathbf{p}} (\sqrt{a_{est}x(\mathbf{p}) + b_{est}} - \sqrt{ax(\mathbf{p}) + b})^2}{\sum_{\mathbf{p}} (ax(\mathbf{p}) + b)} \right)^{1/2}. \end{aligned} \quad (33)$$

Since original images  $x$  are not available for real noisy images, the images produced by the denoising algorithm [3] with the true noise parameters were used instead of  $x$  to compute  $\delta$ . Note that  $\delta$  is insensitive to this substitution due to the summation over all pixels in (33).

### B. Experiments With NED2012

In order to test our method on real noisy images, we utilized the NED2012 image database [18], [23]. The database contains 25 real-world RGB images of size  $1950 \times 1305$  formed from the raw data of a CCD camera. Some of the images are shown in Fig. 4. As one can see, there are images almost completely covered by textures, so the database is challenging for noise parameter estimation algorithms. The database also contains the true noise parameters for each color channel obtained by a calibration procedure (see Table II).

TABLE II  
THE TRUE NOISE PARAMETERS FOR NED2012

Channel	$a$	$b$
R	0.1460	7.6876
G	0.1352	5.0834
B	0.1709	12.3381

TABLE III  
THE ESTIMATION ERRORS FOR NED2012. THE SMALLEST ERROR FOR EACH CHANNEL IS SELECTED WITH THE BOLD FONT

Channel	Average $\delta$			Maximal $\delta$		
	[4]	[18]	proposed	[4]	[18]	proposed
R	0.592	<b>0.023</b>	0.025	7.746	0.076	<b>0.073</b>
G	0.651	0.041	<b>0.035</b>	5.952	0.214	<b>0.096</b>
B	0.337	0.056	<b>0.054</b>	3.769	0.146	<b>0.095</b>



Fig. 5.  $512 \times 512$  test images: "Barbara", "Couple", "Fingerprint", "Hill", "Mandrill", and "Pirate".

The proposed algorithm has been compared with methods [4] and [18]. The average and maximum values of  $\delta$  over all images are presented in Table III. The average error of the proposed method is comparable to that of [18]. At the same time, the maximum error of our algorithm is significantly smaller than that of [18] for the green and blue channels.

For the single-threaded C++ implementation of the proposed method, the average time to process one  $1950 \times 1305$  image was 37 seconds on a PC with CPU Intel i7 920 2.67 GHz and 6 GB RAM.

### C. Experiments With Synthetic Noise

In order to examine the behavior of our algorithm on a wider range of noise levels, we made noise parameter estimation and denoising experiments with images corrupted with synthetic noise. The standard test images shown in Fig. 5 were used as the original images, and the noisy images were generated according to the Poisson-Gaussian model (2). The original images were rescaled to the 12-bit range  $0, \dots, 4095$  in order to keep the noise parameters in the same scale as in the NED2012 database. As in the Section III, the level  $b$  of signal-independent noise was fixed to 25; and the level  $a$  of signal-dependent noise was varied from 10 to 400.

TABLE IV  
THE ESTIMATION ERROR  $\delta$  FOR SYNTHETIC NOISE

Image	$a$					
	10	20	50	100	200	400
Barbara	0.028	0.024	0.015	0.013	0.015	0.035
Couple	0.072	0.048	0.024	0.025	0.033	0.046
Fingerprint	0.065	0.039	0.023	0.018	0.022	0.026
Hill	0.054	0.040	0.021	0.016	0.021	0.034
Mandrill	0.012	0.015	0.008	0.013	0.019	0.028
Pirate	0.028	0.023	0.019	0.017	0.018	0.027

TABLE V  
THE DENOISING RESULTS. TOP ROW: THE NOISY IMAGE PSNR. MIDDLE  
ROW: THE DENOISED IMAGE PSNR FOR THE TRUE NOISE  
PARAMETERS. BOTTOM ROW: THE DENOISED IMAGE  
PSNR FOR THE ESTIMATED NOISE PARAMETERS.  
THE PSNR IS GIVEN IN DECIBELS

Image	$a$					
	10	20	50	100	200	400
Barbara	29.48	26.48	22.50	19.49	16.48	13.48
	35.61	33.99	31.87	30.20	28.42	26.21
	35.64	34.00	31.88	30.22	28.47	26.59
Couple	29.36	26.35	22.37	19.36	16.35	13.34
	34.87	33.22	31.10	29.48	27.85	25.97
	34.78	33.18	31.08	29.46	27.84	26.15
Fingerprint	28.67	25.66	21.69	18.68	15.67	12.66
	32.94	31.09	28.76	27.07	25.42	23.69
	33.03	31.12	28.75	27.06	25.43	23.82
Hill	29.69	26.67	22.70	19.68	16.68	13.67
	34.70	33.08	31.13	29.71	28.30	26.54
	34.57	32.98	31.09	29.69	28.31	26.83
Mandrill	29.10	26.08	22.12	19.10	16.10	13.08
	33.98	31.93	29.31	27.40	25.61	23.95
	33.98	31.91	29.30	27.39	25.58	24.01
Pirate	29.70	26.70	22.72	19.70	16.70	13.69
	34.89	33.08	30.86	29.36	27.92	26.22
	34.85	33.04	30.82	29.34	27.93	26.52

This corresponds to varying the peak signal-to-noise ratio (PSNR) approximately in the range from 29 dB to 13 dB. Denoising was performed using the algorithm [3], which represents the state of the art in signal-dependent noise removal.

As shown in Table IV, the estimation error  $\delta$  remains small in the whole range of parameter  $a$ . That means, the estimator is insensitive to the negative skewness of the noise distribution in image  $f(y; a, b)$ , which takes place for large  $a$  (see Fig. 1).

Table V presents the results of denoising with the true and estimated noise parameters. As one can see, the denoised image PSNR for the estimated noise parameters is approximately the same as that for the true ones. That means the estimation error does not make the denoising results worse.

## VI. DISCUSSION

### A. Choice of the Algorithm Components

Since our approach is based on the VST, its idea is independent of the noise model. In this paper, we demonstrate the applicability of this approach for digital image sensor noise, which is characterized by variance (3) and VST (9). Nevertheless, other signal-dependent noise models can be handled in the same way as well.

Regarding the transform of image blocks, PCA is utilized in this work, because it is data-adaptive and can separate noise

not only in homogeneous areas but also in regular textures. However, some other orthogonal transform, which is able to separate signal and noise, e.g. the discrete cosine transform, can be applied as well. In this case,  $w_{ki}$  are the transform coefficients and they can be used in the other parts of the algorithm without changes.

Speaking of the noise normality assessment, it can be carried out by measuring the difference between the noise CDF and the normal CDF, or by measuring the difference between the standardized moments of the noise and the moments of the standard normal distribution.

The difference between the noise CDF and the normal CDF, which is the basis of the Lilliefors test [24], the Anderson-Darling test [25], and the Cramér-von Mises criterion [26], is very sensitive to outliers and local deviations from normality. For this reason, when it is used as the objective function instead of the excess kurtosis, it often misleads the selection of  $\phi$  and results in a relatively low accuracy of the noise parameter estimates. Moreover, sensitivity to local deviations from normality is not required, since only the variance is utilized to compute  $\sigma_n^2$  (see line 15 of Algorithm 1).

The standardized moments of the noise are relatively insensitive to outliers and provide a suitable way for normality assessment. Since the first two moments, which correspond to the mean and the variance, are free parameters of a normal distribution, the 3rd, the 4th, and high-order moments can be utilized here. As shown in Section IV-C, the 4th standardized moment, which corresponds to the excess kurtosis, is an efficient solution.

### B. Comparison With the State of the Art

Compared with the scatter-plot approach, which is the state of the art in signal-dependent noise parameter estimation, the proposed method utilizes a different assumption about the original image. Instead of looking for homogeneous areas, it selects a block subset, whose intrinsic dimension is smaller than the number of block pixels. This allows efficient processing of highly textured images.

Compared with the PCA-based method [8], which is designed for AWGN variance estimation, the proposed method utilizes a more sophisticated block selection procedure. The block selection method in [8] cuts the tails of the noise distribution so that it cannot be applied in the presented algorithm, which is based on the noise normality assessment. On the contrary, the proposed block selection approach preserves the noise distribution properties. Moreover, it considers only stochastic textures and is insensitive to image structures. Additionally, the noise excess kurtosis is studied in order to measure the noise normality, since the proposed approach is designed for signal-dependent noise.

## VII. CONCLUSION

In this work, we propose a new image sensor noise parameter estimation algorithm. It is qualitatively different from the state of the art scatter-plot method and allows processing of images without homogeneous areas.

According to our noise parameter estimation experiments, the average error of the proposed algorithm is comparable with the state of the art. However, the maximum error is smaller than that of the existing approaches, which demonstrates a higher robustness of our method.

The denoising experiments show that the proposed noise parameter estimator can be utilized together with the state of the art denoising methods, leading to the same results as noise removal with the true noise parameters.

#### ACKNOWLEDGMENT

The authors would like to thank Dr. A. Foi for making his methods publicly available. The authors are also grateful to Dr. M. Uss for providing the results of his method and for making the NED2012 image database publicly available.

#### REFERENCES

- [1] A. Buades, B. Coll, and J. M. Morel, "On image denoising methods," *SIAM Multiscale Model. Simul.*, vol. 4, no. 2, pp. 490–530, 2005.
- [2] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Trans. Image Process.*, vol. 16, no. 8, pp. 2080–2095, Aug. 2007.
- [3] A. Foi, "Clipped noisy images: Heteroskedastic modeling and practical denoising," *Signal Process.*, vol. 89, no. 12, pp. 2609–2629, Dec. 2009.
- [4] 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.
- [5] A. Danielyan and A. Foi, "Noise variance estimation in nonlocal transform domain," in *Proc. Int. Workshop Local Non-Local Approximation Image Process. (LNLA)*, Aug. 2009, pp. 41–45.
- [6] 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.
- [7] N. Ponomarenko, V. Lukin, M. Zriakhov, A. Kaarna, and J. Astola, "An automatic approach to lossy compression of AVIRIS images," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Jul. 2007, pp. 472–475.
- [8] S. Pyatykh, J. Hesser, and L. Zheng, "Image noise level estimation by principal component analysis," *IEEE Trans. Image Process.*, vol. 22, no. 2, pp. 687–699, Feb. 2013.
- [9] G. Torricelli, F. Argenti, and L. Alparone, "Modelling and assessment of signal-dependent noise for image de-noising," in *Proc. 11th Eur. Conf. Signal Process. (EUSIPCO)*, 2002, pp. 287–290.
- [10] B. Aiazzi, L. Alparone, S. Baronti, and A. Garzelli, "Coherence estimation from multilook incoherent SAR imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 11, pp. 2531–2539, Nov. 2003.
- [11] A. Bosco, R. A. Bruna, D. Giacalone, S. Battiato, and R. Rizzo, "Signal-dependent raw image denoising using sensor noise characterization via multiple acquisitions," in *Proc. Soc. Photo-Opt. Instrum. Eng. (SPIE) Conf. Ser.*, vol. 7537, 2010, p. 4.
- [12] W. Förstner, "Image preprocessing for feature extraction in digital intensity, color and range images," in *Geomatic Method for the Analysis of Data in the Earth Sciences (Lecture Notes in Earth Sciences)*, vol. 95, A. Dermanis, A. Grün, and F. Sansò, Eds. Berlin, Germany: Springer-Verlag, 2000, pp. 165–189, doi: 10.1007/3-540-45597-3\_4.
- [13] 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.
- [14] P. Paul, H. Duessmann, T. Bernas, H. Huber, and D. Kalamatianos, "Automatic noise quantification for confocal fluorescence microscopy images," *Comput. Med. Imag. Graph.*, vol. 34, no. 6, pp. 426–434, 2010.
- [15] V. Zabrodina, S. Abramov, V. Lukin, J. Astola, B. Vozel, and K. Chehdi, "Blind estimation of mixed noise parameters in images using robust regression curve fitting," in *Proc. 19th Eur. Signal Process. Conf.*, Barcelona, Spain, Aug./Sep. 2011, pp. 1135–1139.
- [16] L. Alparone, M. Selva, B. Aiazzi, S. Baronti, F. Butera, and L. Chiarantini, "Signal-dependent noise modelling and estimation of new-generation imaging spectrometers," in *Proc. 1st Workshop Hyperspectral Image Signal Process., Evol. Remote Sens. (WHISPERS)*, Aug. 2009, pp. 1–4.
- [17] J.-S. Lee, K. Hoppel, and S. A. Mango, "Unsupervised estimation of speckle noise in radar images," *Int. J. Imag. Syst. Technol.*, vol. 4, no. 4, pp. 298–305, 1992.
- [18] M. L. Uss, B. Vozel, V. V. Lukin, and K. Chehdi, "Image informative maps for component-wise estimating parameters of signal-dependent noise," *J. Electron. Imag.*, vol. 22, no. 1, p. 013019, 2013.
- [19] M. B. Wilk and R. Gnanadesikan, "Probability plotting methods for the analysis for the analysis of data," *Biometrika*, vol. 55, no. 1, pp. 1–17, 1968.
- [20] I. Jolliffe, "Principal component analysis," in *Encyclopedia of Statistics in Behavioral Science*. New York, NY, USA: Wiley, 2002.
- [21] S. Frühwirth-Schnatter, *Finite Mixture and Markov Switching Models*. New York, NY, USA: Springer-Verlag, 2006.
- [22] K. V. Mardia, "Measures of multivariate skewness and kurtosis with applications," *Biometrika*, vol. 57, no. 3, pp. 519–530, 1970.
- [23] M. Uss. (2012). *Image Database for Benchmarking Signal-Dependent Noise Estimation Algorithms: NED2012*. [Online]. Available: <http://rsd.khai.edu/ned2012/ned2012.php>.
- [24] H. W. Lilliefors, "On the Kolmogorov–Smirnov test for normality with mean and variance unknown," *J. Amer. Statist. Assoc.*, vol. 62, no. 318, pp. 399–402, 1967.
- [25] T. W. Anderson and D. A. Darling, "A test of goodness of fit," *J. Amer. Statist. Assoc.*, vol. 49, no. 268, pp. 765–769, Dec. 1954.
- [26] H. Cramér, "On the composition of elementary errors," *Scandin. Actuarial J.*, vol. 1928, no. 1, pp. 13–74, 1928.



**Stanislav Pyatykh** was born in Ekaterinburg, Russia, in 1987. He received the Diploma degree in computer science from Belarusian State University, Minsk, Belarus, in 2009, and the Dr.rer.nat. degree in computer science from Heidelberg University, Heidelberg, Germany, in 2014.

He joined Medical Faculty Mannheim, Heidelberg University, in 2009, where he is currently a Post-Doctoral Fellow. His research interests include image segmentation, noise-level estimation, and image denoising.



**Jürgen Hesser** received the Diplom-Physiker degree and the Dr.rer.nat. degree in physics with a focus on global optimization with genetic algorithms from Heidelberg University, Heidelberg, Germany, in 1988 and 1992, respectively, and the Habilitation degree in computer science from the University of Mannheim, Mannheim, Germany, in 1999.

He has been with the University of Mannheim as a Post-Doctoral Fellow of Computer Graphics, Bioinformatics, and Medical Technology since 1992. He became a Deputy Professor of Medical Technology with the University of Mannheim in 1998. He joined the Medical Faculty Mannheim at Heidelberg University as a Professor of Experimental Radiation Oncology in 2007. He was the Co-Founder of three spinoff companies.

Prof. Hesser was a recipient of several Innovation Awards for an intelligent plaster in 2011, the Innovation Award and Roboters Prize for kV–MV imaging in 2010, the DEGRO Poster Awards in 2009, and several national and local Innovation Awards for excellent research.