Fast Noise Variance Estimation by Principal Component Analysis

Stanislav Pyatykh[a], Lei Zheng[a], Jürgen Hesser[a]

[a]University Medical Center Mannheim, Heidelberg University, Theodor-Kutzer-Ufer 1-3, D-68167 Mannheim, Germany

ABSTRACT

Noise variance estimation is required in many image denoising, compression, and segmentation applications. In this work, we propose a fast noise variance estimation algorithm based on principal component analysis of image blocks. First, we rearrange image blocks into vectors and compute the covariance matrix of these vectors. Then, we use Bartlett's test in order to select the covariance matrix eigenvalues, which correspond only to noise. This allows estimating the noise variance as the average of these eigenvalues. Since the maximum possible number of eigenvalues corresponding to noise is utilized, it is enough to process only a small number of image blocks, which allows reduction of the execution time. The blocks to process are selected from image regions with the smallest variance. During our experiments involving seven state of the art methods, the proposed approach was significantly faster than the methods with similar or higher accuracy. Meanwhile, the relative error of our estimator was always less than 15%. We also show that the proposed method can process images without homogeneous areas.

Keywords: Noise variance estimation, principal component analysis, image denoising, additive white Gaussian noise

1. INTRODUCTION

We consider images corrupted with additive white Gaussian noise (AWGN), which is the most common noise model. The noise is characterized by its variance, which is an important parameter for the majority of image denoising algorithms, because it controls the strength of the filtering. However, the noise variance is often unknown and should be estimated.

The problem of noise variance estimation was being extensively studied during the last two decades. The most algorithms assume that the input image contains a sufficient amount of homogeneous areas, which allows estimating the noise variance as the variance of the input image in these areas. In order to find the homogeneous areas, high-pass filtering[1–8] or preclassification of image blocks[9–12] can be applied. Since blocks, which are classified as homogeneous, can still be affected by image structures, the distribution of the block variances can contain outliers. Hence robust estimators of the noise variance are utilized,[2,3,13–16] such as the median of block variances,[2] the mode of block variances,[14] and the average of several smallest block variances.[15,16]

There are also methods, which are not based on the analysis of image block variances. In particular, the noise variance can be estimated using a measure of image bit-plane randomness.[17] Under the assumption that the kurtosis of marginal bandpass filter response distributions is constant for noise-free images, a kurtosis model for noisy images can be constructed; and the noise variance can be assessed by finding the best parameters of

Further author information: (Send correspondence to Stanislav Pyatykh)
Stanislav Pyatykh: E-mail: stanislav.pyatykh@medma.uni-heidelberg.de, Telephone: +49 621 383 6912
Lei Zheng: E-mail: lei.zheng@medma.uni-heidelberg.de, Telephone: +49 621 383 6912
Jürgen Hesser: E-mail: juergen.hesser@medma.uni-heidelberg.de, Telephone: +49 621 383 5109

this model.[18] A Bayesian framework with a learned Markov random field prior can be used for noise variance estimation as well.[19]

The signal and the noise can also be separated using the wavelet transform.[20–23] When the wavelet coefficients at the finest decomposition level are used directly,[22] the noise variance is significantly overestimated,[23] because image structures affect these coefficients as well. One way to overcome this problem is to estimate the noise level from the wavelet coefficients smaller than a threshold.[20]

Recently, more powerful techniques for the signal and noise separation were proposed. All of them assume that some transform of the noise-free image is sparse. As the noise is uncorrelated and does not allow a sparse representation, it can be easily separated from the noise-free image in the transform domain. For example, when discrete cosine transform (DCT) of image blocks[24] is applied, image structures define only low frequency transform coefficients, whereas the noise affects all coefficients. Therefore, the noise variance can be estimated from high frequency transform coefficients. Another transform is based on self-similarity of image blocks: when similar blocks are combined in 3D stacks and 3D DCT is applied to each stack, image structures occupy only low frequency coefficients, so that the noise variance can be estimated from high frequency coefficients.[25] Furthermore, one can treat image blocks as vectors and make a principal component analysis (PCA) of these vectors.[26] In this case, image structures affect only the largest eigenvalues of the covariance matrix, and the smallest eigenvalue of the covariance matrix can be used for noise variance estimation.

The common problem of the transform-based methods is low computational efficiency[26] compared with the methods, which use block variances in order to estimate the noise variance. In this work, we propose a fast noise variance estimation algorithm. Similar to the method,[26] our approach utilizes PCA of image blocks, but instead of using only the smallest covariance matrix eigenvalue, it automatically determines the eigenvalues, which can be utilized for noise variance estimation. As a result, much smaller amount of image blocks can be used without significant loss of the accuracy, which leads to a dramatic decrease of the execution time.

The rest of the article is organized as follows. We present the theoretical framework in Subsections 2.1 and 2.2; and the fast implementation is described in Subsections 2.3 and 2.4. The results and the discussion are given in Sections 3 and 4 respectively. We conclude in Section 5.

2. METHOD

Let $S_1 \times S_2$ be the image size, where S_1 is the number of columns and S_2 is the number of rows, $\mathbf{x}$ be a noise-free image, and $\mathbf{y} = \mathbf{x} + \mathbf{n}$ be an image corrupted with AWGN $\mathbf{n}$. The problem is to estimate the noise variance σ^2.

Suppose that we have N blocks of size $B \times B$ in each of images $\mathbf{x}$, $\mathbf{n}$, and $\mathbf{y}$. (The strategy of image block selection is described in detail in Subsection 2.3.) We treat these blocks as vectors $\mathbf{x}_i$, $\mathbf{n}_i$, $\mathbf{y}_i$, $i = 1, \ldots, N$ of size $M = B^2$, which are realizations of some random variables $\mathbf{X}$, $\mathbf{N}$, and $\mathbf{Y}$ respectively.

2.1 Population PCA

Let $\mathbf{\Sigma_X}$, $\mathbf{\Sigma_N}$, and $\mathbf{\Sigma_Y}$ be the population covariance matrices of $\mathbf{X}$, $\mathbf{N}$, and $\mathbf{Y}$ respectively, and $\lambda_{\mathbf{X},1} \geq \cdots \geq \lambda_{\mathbf{X},M}$ be the eigenvalues of $\mathbf{\Sigma_X}$ with corresponding normalized eigenvectors $\mathbf{v}_{\mathbf{X},1}, \ldots, \mathbf{v}_{\mathbf{X},M}$. Let $q_1, \ldots, q_p$ be the multiplicities of the eigenvalues of $\mathbf{\Sigma_X}$ and

$$\begin{aligned}
\delta_{\mathbf{X},1} &= \lambda_{\mathbf{X},1} = \cdots = \lambda_{\mathbf{X},q_1} \\
\delta_{\mathbf{X},2} &= \lambda_{\mathbf{X},q_1+1} = \cdots = \lambda_{\mathbf{X},q_1+q_2} \\
&\vdots \\
\delta_{\mathbf{X},p} &= \lambda_{\mathbf{X},M-q_p+1} = \cdots = \lambda_{\mathbf{X},M}
\end{aligned}$$

where $\delta_{\mathbf{X},1} > \cdots > \delta_{\mathbf{X},p}$. Since the noise-free image and the noise are independent, $\operatorname{cov}(\mathbf{X}, \mathbf{N}) = 0$ and hence $\mathbf{\Sigma_Y} = \mathbf{\Sigma_X} + \mathbf{\Sigma_N}$. Furthermore, $\mathbf{\Sigma_N} = \sigma^2 I$, because $\mathbf{n}$ is AWGN and elements of $\mathbf{N}$ are independent. For $i = 1, \ldots, M$,

$$\mathbf{\Sigma_Y} \mathbf{v}_{\mathbf{X},i} = \mathbf{\Sigma_X} \mathbf{v}_{\mathbf{X},i} + \sigma^2 \mathbf{v}_{\mathbf{X},i} = \lambda_{\mathbf{X},i} \mathbf{v}_{\mathbf{X},i} + \sigma^2 \mathbf{v}_{\mathbf{X},i} = (\lambda_{\mathbf{X},i} + \sigma^2) \mathbf{v}_{\mathbf{X},i}. \quad (1)$$

Hence $\lambda_{\mathbf{Y},i} = \lambda_{\mathbf{X},i} + \sigma^2$ are the eigenvalues of $\mathbf{\Sigma_Y}$ with corresponding eigenvectors $\mathbf{v}_{\mathbf{X},i}$, $i = 1, \ldots, M$. The multiplicities of the eigenvalues of $\mathbf{\Sigma_X}$ and $\mathbf{\Sigma_Y}$ are the same and we denote the distinct values of the eigenvalues of $\mathbf{\Sigma_Y}$ by $\delta_{\mathbf{Y},1}, \ldots, \delta_{\mathbf{Y},p}$ so that

$$\delta_{\mathbf{Y},i} = \delta_{\mathbf{X},i} + \sigma^2, \quad i = 1, \ldots, p. \tag{2}$$

$\mathbf{v}_{\mathbf{X},i}^{\mathrm{T}}\mathbf{Y}$ are the population principal components of $\mathbf{Y}$; and their variances $\mathrm{var}(\mathbf{v}_{\mathbf{X},i}^{\mathrm{T}}\mathbf{Y})$ equal $\lambda_{\mathbf{Y},i}$.[27]

When $\delta_{\mathbf{X},p} = 0$, $\delta_{\mathbf{Y},p}$ equals σ^2, which provides a way for noise variance estimation by estimating $\delta_{\mathbf{Y},p}$ from the realizations of $\mathbf{Y}$. Therefore, we are interested in a class of noise-free images, for which $\delta_{\mathbf{X},p} = 0$. This condition means that $\mathbf{X}$ takes its values almost surely in some proper subspace of $\mathbb{R}^M$, i.e. in some subspace of $\mathbb{R}^M$ whose dimension is smaller than M. It implies that components of $\mathbf{X}$ are almost surely linearly dependent, i.e. there is a linear dependence between pixels of $\mathbf{x}$ in the image blocks.

We consider an image with $\delta_{\mathbf{X},p} = 0$ in the following example. Let $\mathbf{x}$ be a checkerboard image:

0	1	0	1
1	0	1	0
0	1	0	1
1	0	1	0

(3)

Taking $B = 2$, there are only two possible blocks:

0	1
1	0

and

1	0
0	1

(4)

and two possible values of $\mathbf{x}_i$: $(0,1,1,0)^{\mathrm{T}}$ and $(1,0,0,1)^{\mathrm{T}}$. Therefore, we can assume that $\mathbf{X} \in \mathbb{R}^4$ takes only these values. That means, $\mathbf{X}$ always lies on the line passing through points $(0,1,1,0)^{\mathrm{T}}$ and $(1,0,0,1)^{\mathrm{T}}$, which is a one-dimensional subspace of $\mathbb{R}^4$. Hence $\mathbf{\Sigma_X}$ has only one nonzero eigenvalue, $p = 2$, $q_2 = 3$, and $\delta_{\mathbf{X},2} = 0$. This example also shows that a noise-free image with zero $\delta_{\mathbf{X},p}$ does not necessarily contain homogeneous areas or only low frequencies.

2.2 Sample PCA

In practice, we have only realizations of $\mathbf{Y}$; and population covariance matrix $\mathbf{\Sigma_Y}$, as well as its eigenvalues, cannot be computed. Therefore, analysis of the corresponding sample quantities is required. Let $S_{\mathbf{Y}}$ be the sample covariance matrix of $\mathbf{Y}$:

$$S_{\mathbf{Y}} = \frac{1}{N-1}\Big(\sum_{i=1}^{N}\mathbf{y}_i\mathbf{y}_i^{\mathrm{T}} - \frac{1}{N}\sum_{i=1}^{N}\mathbf{y}_i\sum_{i=1}^{N}\mathbf{y}_i^{\mathrm{T}}\Big). \tag{5}$$

and $\tilde{\lambda}_{\mathbf{Y},1} \geq \ldots \geq \tilde{\lambda}_{\mathbf{Y},M}$ be the eigenvalues of $S_{\mathbf{Y}}$. As discussed in the previous subsection, the objective is to estimate q_p and $\delta_{\mathbf{Y},p}$ from $\tilde{\lambda}_{\mathbf{Y},1}, \ldots, \tilde{\lambda}_{\mathbf{Y},M}$.

First, the estimation of q_p is required. Let

$$H_{0k} : \lambda_{\mathbf{Y},M-k+1} = \cdots = \lambda_{\mathbf{Y},M} \tag{6}$$

be the hypothesis that the last k eigenvalues of $\mathbf{\Sigma_Y}$ are equal, and

$$H_{1k} : \lambda_{\mathbf{Y},M-k+1} > \lambda_{\mathbf{Y},M} \tag{7}$$

be the alternative hypothesis. H_{0k} can be tested against H_{1k} using Bartlett's test,[27,28] in which H_{0k} is rejected at significance level α if

$$N'\Big(k\ln\big(\frac{1}{k}\sum_{i=M-k+1}^{M}\tilde{\lambda}_{\mathbf{Y},i}\big) - \sum_{i=M-k+1}^{M}\ln(\tilde{\lambda}_{\mathbf{Y},i})\Big) \geq \chi^2_{\nu,\alpha} \tag{8}$$

Figure 1. 512×512 'Cameraman' image blurred with Gaussian kernel with $\sigma = 2$.

where $N' = N - (2M + 11)/6$, $\chi^2_{\nu,\alpha}$ is the value of the inverse cumulative distribution function of the chi-squared distribution with ν degrees of freedom at point $1 - \alpha$, and $\nu = (k + 2)(k - 1)/2$. Repeating this test for $k = 2, 3, 4, \ldots$ until H_{0k} is rejected allows estimating q_p as the maximal k, for which H_{0k} is accepted. We use $\alpha = 0.01$.

Second, when q_p is known, $\delta_{\mathbf{Y},p}$ can be estimated as the average of the last q_p sample eigenvalues:

$$\frac{\tilde{\lambda}_{\mathbf{Y},M-q_p+1} + \cdots + \tilde{\lambda}_{\mathbf{Y},M}}{q_p}. \tag{9}$$

Wielandt's eigenvalue inequality[29] implies that (9) converges to $\delta_{\mathbf{Y},p}$ in probability so that it is a consistent estimator of $\delta_{\mathbf{Y},p}$.

Since Bartlett's test can be used to test the equality of any consecutive eigenvalues,[27] the estimation of q_i and $\delta_{\mathbf{Y},i}$ can be continued in the similar way for $i = p-1, \ldots, 1$. For example, we have computed estimates of q_i and $\delta_{\mathbf{Y},i}$ using 5×5 blocks for the image shown in Fig. 1. This is the standard test image 'Cameraman', which has been blurred in order to remove possible noise. The estimates of q_i and $\delta_{\mathbf{Y},i}$ are shown in Table 1; and the nonzero estimates of $\delta_{\mathbf{Y},i}$ are plotted in Fig. 2. As one can see, all nonzero eigenvalues are distinct and have approximately exponential decay. The last five eigenvalues equal zero, which means that PCA can be used for noise variance estimation for this image.

2.3 Image Region Selection

Computation of the estimate of $\delta_{\mathbf{Y},p}$ consists of the following parts:

1. Computation of the sample covariance matrix using (5). The execution time is proportional to NM^2.
2. Computation of the eigenvalues of the sample covariance matrix. The execution time is $O(M^3)$.[30]
3. Sequence of Bartlett's tests (8). The execution time is $O(M^2)$.
4. Computation of the estimate of $\delta_{\mathbf{Y},p}$ using (9). The execution time is $O(M)$.

Table 1. The estimates of q_i and $\delta_{\mathbf{Y},i}$ for the image shown in Fig. 1. The estimates of $\delta_{\mathbf{Y},i}$ have been computed with the accuracy 10^{-4}.

i	q_i	$\delta_{\mathbf{Y},i}$
1	1	87541.0010
2	1	1115.0010
3	1	612.8945
4	1	61.4036
5	1	46.7131
6	1	25.0071
7	1	3.5403
8	1	2.3485
9	1	1.2356
10	1	0.7212
11	1	0.1731
12	1	0.1080
13	1	0.0571
14	1	0.0157
15	1	0.0085
16	1	0.0051
17	1	0.0047
18	1	0.0014
19	1	0.0007
20	1	0.0002
21	5	0.0000

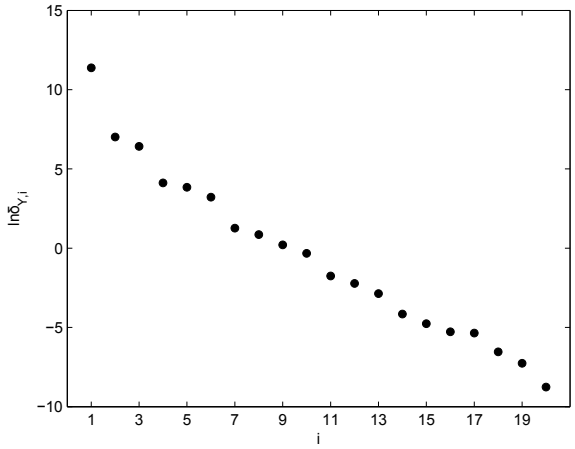

Figure 2. Semi-logarithmic plot of the nonzero estimates of $\delta_{\mathbf{Y},i}$ for the image shown in Fig. 1.

Typically, $N \gg M$, hence the computation of the sample covariance matrix is the most expensive operation.

When one chooses number of image blocks N to be used in sample PCA, there is a trade-off between the execution time and the accuracy:

1. A large number of blocks leads to accurate estimates, but large execution time, because the execution time of the sample covariance matrix computation is proportional to the number of blocks.
2. A small number of blocks leads to small execution time, but the results are not statistically significant.

Therefore, we partition the image domain into square regions $\{R_k\}$ of predefined size $T \times T$ and use image blocks only from several of these regions. The predefined region size guarantees that there is a sufficient number of image blocks for accurate noise variance estimation, whereas processing of only a small number of the regions allows reduction of the execution time. The regions are defined as

$$R_k = R_{i+jK_1} = \big\{(i-1)T+1,\dots,iT\big\} \times \big\{(j-1)T+1,\dots,jT\big\}, \quad i = 1,\dots,K_1,\ j = 1,\dots,K_2 \tag{10}$$

where $K_1 = \lfloor S_1/T \rfloor$ and $K_2 = \lfloor S_2/T \rfloor$ (recall that $S_1 \times S_2$ is the image size). Each region R_{i+jK_1} contains $(T-B+1)^2$ blocks of size $B \times B$ with left-top corner positions in set

$$\big\{(i-1)T+1,\dots,iT-B+1\big\} \times \big\{(j-1)T+1,\dots,jT-B+1\big\}. \tag{11}$$

We use $T = 64$ in our implementation.

In the ideal case, regions with $\delta_{\mathbf{X},p} = 0$ should be selected for processing. However, no information about $\delta_{\mathbf{X},p}$ is available before applying PCA, hence some heuristic should be utilized for the region selection. Intuitively, regions containing simple structures should be taken, which can be described by the region variance, entropy, or autocorrelation. In our implementation, we use the region variance, because it is fast to compute.

In 8-bit images with gray value range $[0, 255]$, noise clipping usually occurs. Therefore, we skip regions, in which more than 5% of pixels have the value 0 or 255.

2.4 Algorithm

In our algorithm, we assume that nonzero eigenvalues of $\mathbf{\Sigma_X}$ are distinct. This assumption is often used to select principal components, which are affected only by noise.[27] It can be explained by the fact that vector $\mathbf{X}$ represents image structures, and it is very unlikely to have the same variance in different directions, which can also be seen in Table 1. Therefore, if $q_p > 1$, $\delta_{\mathbf{X},p}$ is assumed to be zero, $\delta_{\mathbf{Y},p} = \sigma^2$, and (9) can be used as a noise variance estimate. If $q_p = 1$, nothing can be said about $\delta_{\mathbf{X},p}$.

Our noise variance estimation method is presented in Algorithm 1. First, sample variance s^2 of image pixels is computed for each region R_k. The sample variances are sorted in ascending order, and the sorting result is stored in permutation π.

In lines 3–20, we try to compute a noise variance estimate using blocks of size $B \times B$. Since the sample covariance matrix computation time is proportional to $NM^2 = NB^4$, we start with small $B = B_{min}$. If our assumption is satisfied for some B, the algorithm stops (line 19). If the assumption is not satisfied, $\delta_{\mathbf{X},p}$ may be nonzero, i.e. $\mathbf{X}$ may not lie in some proper subspace of $\mathbb{R}^M$. We suppose that if the vector dimension is increased from $M = B^2$ to $M = (B+1)^2$, $\mathbf{X}$ will lie in some proper subspace of $\mathbb{R}^M$ and $\delta_{\mathbf{X},p}$ will be zero. Hence we increase B in the loop from B_{min} to B_{max}. If the assumption is not satisfied for any B, we take the estimate of $\delta_{\mathbf{Y},p}$ computed using the largest $B = B_{max}$ as the final estimate (line 13). In this case, the highest number of dimensions $M = B_{max}^2$ is used to represent image structures so that the estimation error, which is $\delta_{\mathbf{X},p}$, is expected to be the smallest.

In lines 6–16, we process the image blocks from k regions with the smallest variance, namely $R_{\pi(1)},\dots,R_{\pi(k)}$. Final estimate σ^2_{est} is updated if $q_p > 1$ or if its previous value was not valid (lines 8–16). We start with $k = 1$, for which the number of processed image blocks is minimal and, therefore, the chance that $\delta_{\mathbf{X},p} = 0$ is the highest. Then, we increase k in order to process a larger number of blocks and increase the accuracy. k is increased until

Algorithm 1

Input: image $\mathbf{y}$ corrupted with AWGN
Output: noise variance estimate σ^2_{est}

1: $\pi \leftarrow \mathrm{sort}\big(\{s^2(R_k)\}\big)$
2: **for** $B = B_{min}$ **to** B_{max} **do**
3: $\quad\sigma^2_{est} \leftarrow$ InvalidValue
4: $\quad$assumptionIsSatisfied $\leftarrow$ **false**
5: $\quad$**for** $k = 1$ **to** k_{max} **do**
6: $\quad\quad\tilde{\lambda}_{\mathbf{Y},1}, \ldots, \tilde{\lambda}_{\mathbf{Y},M} \leftarrow \mathrm{PCA}(R_{\pi(1)}, \ldots, R_{\pi(k)})$
7: $\quad\quad$Estimate q_p using sequence of tests (8)
8: $\quad\quad$**if** $q_p > 1$ **then**
9: $\quad\quad\quad\sigma^2_{est} \leftarrow (9)$
10: $\quad\quad\quad$assumptionIsSatisfied $\leftarrow$ **true**
11: $\quad\quad$**else**
12: $\quad\quad\quad$**if** $\sigma^2_{est} =$ InvalidValue **then**
13: $\quad\quad\quad\quad\sigma^2_{est} \leftarrow (9)$
14: $\quad\quad\quad$**end if**
15: $\quad\quad\quad$**break**
16: $\quad\quad$**end if**
17: $\quad$**end for**
18: $\quad$**if** assumptionIsSatisfied **then**
19: $\quad\quad$**break**
20: $\quad$**end if**
21: **end for**
22: **return** σ^2_{est}

$q_p = 1$ (line 15), i.e. until we cannot assume that $\delta_{\mathbf{X},p} = 0$. As a result, the enumeration of k allows computation of the estimate, for which our assumption is satisfied, with the highest possible accuracy.

Note that when calculating (5) for regions $R_{\pi(1)}, \ldots, R_{\pi(k)}, R_{\pi(k+1)}$, we reuse values $\sum \mathbf{y}_i \mathbf{y}_i^{\mathrm{T}}$ and $\sum \mathbf{y}_i$ computed for regions $R_{\pi(1)}, \ldots, R_{\pi(k)}$ in order to avoid redundant computations.

We use the following algorithm parameters: $B_{min} = 5$, $B_{max} = 8$, $k_{max} = 2$.

3. EXPERIMENTS

The proposed method has been evaluated on the TID2008 database,[31] which has been already used for testing noise variance estimation methods.[12,25,26] It contains 25 RGB images: 24 real-world scenes and one artificial image. The noise variance has been estimated in each color component independently, that means, 75 grayscale images have been used for noise level estimation. The database includes noisy images with the noise variance 65 and 130; and we have additionally generated noisy images with the noise variance 25 and 100. Several images from this database are shown in Fig. 3.

The noisy images have been computed as the sums of the reference images from the database and realizations of Gaussian noise with the variance given above (25, 65, 100, or 130). However, the reference images contain a small level of noise, which should be considered in order to get the true noise variance in the noisy images. We used estimates σ^2_{ref} of the noise variance in each of the reference images computed by a semi-automatic procedure;[26] and the estimates of the noise variance in the noisy images have been corrected as follows:

$$\sigma^2_{corr} = \sigma^2_{est} - \sigma^2_{ref} \tag{12}$$

where σ^2_{est} is the algorithm output.

We have compared our approach with the following recent methods:

Figure 3. Images from the TID2008 database.

1. Pyatykh et al.,[26] which takes the smallest eigenvalue of the image block covariance matrix as the noise variance estimate.
2. Danielyan et al.,[25] where self-similarity of image blocks is utilized in order to separate the noise from the signal.
3. Ponomarenko et al.,[24] which applies DCT for signal and noise separation.
4. Yang et al.,[6] where a Sobel edge detection operator is used in order to exclude the noise-free image content and image blocks with the smallest variance are utilized for noise variance estimation.
5. van Zyl Marais et al.,[15] where the input image is divided into blocks and the block variances are computed.
6. Barducci et al.,[17] which utilizes a measure of bit-plane randomness.
7. Starck et al.,[20] where the multiresolution support data structure is used under the assumption that small wavelet transform coefficients correspond to the noise.

For all methods, we have used only machine code implementations, which allows comparison of the execution time.

The results of our noise level estimation experiments are presented in Tables 2 and 3. Because the execution time of the method[25] is not provided by the author, we have measured the execution time of the BM3D filter,[32] which is a part of the method.[25]

Compared with the approaches,[24–26] the proposed algorithm has approximately the same bias. The estimate standard deviation and the maximum error of the proposed method are comparable with those of the methods,[24,25] but they are larger than those of the algorithm.[26] However, the proposed solution is more than 30 times faster* than the algorithms.[24–26]

Compared with the algorithms,[6,15,17,20] our method has more than 2 times smaller bias. The estimate standard deviation and the maximum error of the proposed algorithm are more than 2 times smaller for the noise variances 25 and 65 and more than 1.3 times smaller for the noise variances 100 and 130.

*We assume that the CPU Intel Celeron 1.4 GHz used to measure the performance of the method[24] is two times slower than the CPU Intel i7 920 2.67 GHz used to measure the performance of the other methods.

Table 2. The accuracy of the considered methods for TID2008. σ are the true noise standard deviation values; and σ_{corr} are the noise standard deviation estimates. $\overline{\sigma}_{corr} - \sigma$ is the bias of corrected estimates, $s(\sigma_{corr})$ is the standard deviation of corrected estimates, $\max|\sigma_{corr} - \sigma|$ is the maximum difference between a corrected estimate and the true value. The last column is the percentage of the images, for which a method cannot estimate the noise level. For the methods,[24,25] the values of σ_{est} have been provided by the authors.

Method	$\overline{\sigma}_{corr} - \sigma$	$s(\sigma_{corr})$	$\max\|\sigma_{corr} - \sigma\|$	% of failures
$\sigma^2 = 25$ ($\sigma = 5$)				
proposed	0.026	0.223	0.685	0
Pyatykh et al.[26]	-0.027	0.147	0.500	0
Danielyan et al.[25]	-0.039	0.158	0.525	0
Ponomarenko et al.[24]	–	–	–	–
Yang et al.[6]	0.322	0.547	2.859	0
van Zyl Marais et al.[15]	-1.499	1.822	5.000	57.3
Barducci et al.[17]	3.227	2.266	9.158	0
Starck et al.[20]	2.144	2.224	8.903	0
$\sigma^2 = 65$ ($\sigma \approx 8.062$)				
proposed	0.021	0.214	0.822	0
Pyatykh et al.[26]	-0.043	0.103	0.486	0
Danielyan et al.[25]	–	–	–	–
Ponomarenko et al.[24]	0.001	0.209	1.078	0
Yang et al.[6]	0.228	0.430	2.093	0
van Zyl Marais et al.[15]	-1.467	2.044	8.062	45.3
Barducci et al.[17]	3.173	1.671	8.968	0
Starck et al.[20]	2.067	2.325	10.160	0
$\sigma^2 = 100$ ($\sigma = 10$)				
proposed	0.052	0.293	1.066	0
Pyatykh et al.[26]	0.009	0.125	0.307	0
Danielyan et al.[25]	0.040	0.175	0.717	0
Ponomarenko et al.[24]	–	–	–	–
Yang et al.[6]	0.232	0.412	1.935	0
van Zyl Marais et al.[15]	-1.517	2.145	10.000	42.7
Barducci et al.[17]	2.204	2.519	9.551	0
Starck et al.[20]	2.248	1.868	7.281	0
$\sigma^2 = 130$ ($\sigma \approx 11.402$)				
proposed	0.083	0.291	1.062	0
Pyatykh et al.[26]	0.014	0.110	0.386	0
Danielyan et al.[25]	–	–	–	–
Ponomarenko et al.[24]	0.094	0.228	1.170	0
Yang et al.[6]	0.224	0.390	1.943	0
van Zyl Marais et al.[15]	-1.467	2.086	6.811	41.3
Barducci et al.[17]	2.700	2.510	9.604	0
Starck et al.[20]	2.132	2.279	10.197	0

Table 3. The average execution time of the considered methods. All implementations are single-threaded. For the method,[24] the value have been provided by the authors.

Method	Time, ms	CPU
proposed	4.1	Intel i7 920 2.67 GHz
Pyatykh et al.[26]	159	Intel i7 920 2.67 GHz
Danielyan et al.[25,32]	2628	Intel i7 920 2.67 GHz
Ponomarenko et al.[24]	~ 250	Intel Celeron 1.4 GHz
Yang et al.[6]	3.1	Intel i7 920 2.67 GHz
van Zyl Marais et al.[15]	1.1	Intel i7 920 2.67 GHz
Barducci et al.[17]	123	Intel i7 920 2.67 GHz
Starck et al.[20]	805	Intel i7 920 2.67 GHz

4. DISCUSSION

The proposed method takes advantage of the redundancy of the information in image blocks, which is done also by the methods.[24–26] Instead of the direct search of similar blocks, which is used in the approach,[25] analysis of the block covariance matrix is used in order to find correspondences between the blocks, which is much more computationally efficient. Compared with image block DCT, which is utilized in the algorithm,[24] the transform computed by PCA is not fixed and depends on the input image, which allows processing images containing textures with high frequencies as shown in the example in Subsection 2.1. In contrast to the method,[26] which uses only the smallest covariance matrix eigenvalue for noise variance estimation, the proposed algorithm automatically computes the number of the eigenvalues which correspond only to the noise, and it utilizes all of them in order to compute the noise variance estimate. Therefore, a much smaller number of image blocks can be used, which allows a dramatic reduction of the execution time.

Our algorithm can be extended to 3D images by processing 3D regions of size $T \times T \times T$ and 3D blocks of size $B \times B \times B$. It can also be combined with a variance-stabilizing transformation in order to estimate parameters of signal-dependent noise.[33,34] Besides, it is possible to optimize the proposed algorithm by parallel computation of the elements of the sample covariance matrix.

5. CONCLUSION

In this work, we present a fast noise variance estimation algorithm, which is based on image block PCA. It provides a good compromise between the accuracy and the execution time: it is much faster than the methods[24–26] and considerably more accurate than the methods.[6,15,17,20] Our algorithm does not require the existence of homogeneous areas in the input image; and it can also handle textures containing high frequencies. The proposed approach can be used in image denoising, compression and segmentation applications, which take the noise variance as an input parameter.

ACKNOWLEDGMENT

The authors would like to thank Dr. S. M. Yang, Prof. Dr. S. C. Tai, Dr. I. van Zyl Marais, Dr. A. Barducci, Dr. I. Pippi, Dr. J. L. Starck for providing the implementations of their methods.

The authors would also like to thank Dr. B. Vozel, D.Sc.Tech. A. Foi, Prof. Dr. Techn. Sci. V. Lukin, Prof. Cand. Techn. Sci. S. Abramov, D.Tech. N. Ponomarenko for providing the results of their methods.

This work was supported by grants AIF KF2769301FRQ, DFG HE 3011/23-1, and DFG HE 3011/14-1.

REFERENCES

[1] Liévin, M., Luthon, F., and Keeve, E., "Entropic estimation of noise for medical volume restoration," in [*Pattern Recognition, 2002. Proceedings. 16th International Conference on*], **3**, 871–874, IEEE (2002).

[2] Corner, B., Narayanan, R., and Reichenbach, S., "Noise estimation in remote sensing imagery using data masking," *International Journal of Remote Sensing* **24**(4), 689–702 (2003).

[3] Bilcu, R. and Vehvilainen, M., "New method for noise estimation in images," in [*Nonlinear Signal and Image Processing, 2005. NSIP 2005. Abstracts. IEEE-Eurasip*], 25, IEEE.

[4] Russo, F., "Gaussian noise estimation in digital images using nonlinear sharpening and genetic optimization," in [*Instrumentation and Measurement Technology Conference Proceedings, 2007. IMTC 2007. IEEE*], 1–5, IEEE.

[5] Wyatt, P. and Nakai, H., "Developing nonstationary noise estimation for application in edge and corner detection," *Image Processing, IEEE Transactions on* **16**(7), 1840–1853 (2007).

[6] Yang, S. and Tai, S., "Fast and reliable image-noise estimation using a hybrid approach," *Journal of Electronic Imaging* **19**(3), 3007 (2010).

[7] Liu, A., "A fast method of estimating gaussian noise," in [*Information Science and Engineering (ICISE), 2009 1st International Conference on*], 441–444, IEEE (2009).

[8] Salmeri, M., Mencattini, A., Rabottino, G., and Lojacono, R., "Signaldependent noise characterization for mammographic images denoising," in [*IMEKO TC4 Symposium (IMEKOTC408)*], (2008).

[9] Salmeri, M., Mencattini, A., Ricci, E., and Salsano, A., "Noise estimation in digital images using fuzzy processing," in [*Image Processing, 2001. Proceedings. 2001 International Conference on*], **1**, 517–520, IEEE (2001).

[10] Shin, D., Park, R., Yang, S., and Jung, J., "Block-based noise estimation using adaptive gaussian filtering," *Consumer Electronics, IEEE Transactions on* **51**(1), 218–226 (2005).

[11] Abramov, S., Lukin, V., Vozel, B., Chehdi, K., and Astola, J., "Segmentation-based method for blind evaluation of noise variance in images," *Journal of Applied Remote Sensing* **2**, 023533 (2008).

[12] Uss, M., Vozel, B., Lukin, V., Abramov, S., Baryshev, I., and Chehdi, K., "Image informative maps for estimating noise standard deviation and texture parameters," *EURASIP Journal on Advances in Signal Processing* **2011** (2011).

[13] Alparone, L., Selva, M., Capobianco, L., Moretti, S., Chiarantini, L., and Butera, F., "Quality assessment of data products from a new generation airborne imaging spectrometer," in [*Geoscience and Remote Sensing Symposium, 2009 IEEE International, IGARSS 2009*], **4**, IV–422, IEEE.

[14] Aja-Fernández, S., Vegas-Sánchez-Ferrero, G., Martín-Fernández, M., and Alberola-López, C., "Automatic noise estimation in images using local statistics. additive and multiplicative cases," *Image and Vision Computing* **27**(6), 756–770 (2009).

[15] van Zyl Marais, I. and Steyn, W., "Noise estimation algorithms for onboard image quality assessment," in [*International Conference on Space Technology*], (2009).

[16] Ghazal, M. and Amer, A., "Homogeneity localization using particle filters with application to noise estimation," *Image Processing, IEEE Transactions on* **20**(7), 1788–1796 (2011).

[17] Barducci, A., Guzzi, D., Marcoionni, P., and Pippi, I., "Assessing noise amplitude in remotely sensed images using bit-plane and scatterplot approaches," *Geoscience and Remote Sensing, IEEE Transactions on* **45**(8), 2665–2675 (2007).

[18] Zoran, D. and Weiss, Y., "Scale invariance and noise in natural images," in [*Computer Vision, 2009 IEEE 12th International Conference on*], 2209–2216, IEEE (2009).

[19] Schmidt, U., Schelten, K., and Roth, S., "Bayesian deblurring with integrated noise estimation," in [*Proc. of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*], 2625–2632 (2011).

[20] Starck, J. and Murtagh, F., "Automatic noise estimation from the multiresolution support," *Publications of the Astronomical Society of the Pacific* **110**, 193–199 (1998).

[21] De Stefano, A., White, P., and Collis, W., "Training methods for image noise level estimation on wavelet components," *EURASIP Journal on Applied Signal Processing* **2004**, 2400–2407 (2004).

[22] Donoho, D., "De-noising by soft-thresholding," *Information Theory, IEEE Transactions on* **41**(3), 613–627 (1995).

[23] Hashemi, M. and Beheshti, S., "Adaptive noise variance estimation in bayesshrink," *Signal Processing Letters, IEEE* **17**(1), 12–15 (2010).

[24] Ponomarenko, N., Lukin, V., Zriakhov, M., Kaarna, A., and Astola, J., "An automatic approach to lossy compression of aviris images," in [*Geoscience and Remote Sensing Symposium, 2007. IGARSS 2007. IEEE International*], 472–475, IEEE (2007).

[25] Danielyan, A. and Foi, A., "Noise variance estimation in nonlocal transform domain," in [*Local and Non-Local Approximation in Image Processing, 2009. LNLA 2009. International Workshop on*], 41–45, IEEE (2009).

[26] Pyatykh, S., Hesser, J., and Zheng, L., "Image noise level estimation by principal component analysis," *Image Processing, IEEE Transactions on* **22**, 687–699 (Feb. 2013).

[27] Jolliffe, I., "Principal component analysis," *Encyclopedia of Statistics in Behavioral Science* (2002).

[28] Bartlett, M., "Tests of significance in factor analysis," *British Journal of Statistical Psychology* **3**(2), 77–85 (1950).

[29] Eaton, M. and Tyler, D., "On wielandt's inequality and its application to the asymptotic distribution of the eigenvalues of a random symmetric matrix," *The Annals of Statistics* **19**(1), 260–271 (1991).

[30] Prass, W., [*Numerical recipes: the art of scientific computing*], Cambridge University Press (2007).

[31] Ponomarenko, N., Lukin, V., Zelensky, A., Egiazarian, K., Carli, M., and Battisti, F., "Tid2008 - a database for evaluation of full-reference visual quality assessment metrics," *Advances of Modern Radioelectronics* **10**(4), 30–45 (2009).

[32] Dabov, K., Foi, A., Katkovnik, V., and Egiazarian, K., "Image denoising by sparse 3-d transform-domain collaborative filtering," *Image Processing, IEEE Transactions on* **16**(8), 2080–2095 (2007).

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

[34] Oktem, R., Egiazarian, K., Lukin, V., Ponomarenko, N., and Tsymbal, O., "Locally adaptive dct filtering for signal-dependent noise removal," *EURASIP Journal on Advances in Signal Processing* **2007**, 10 (2007).