CrossMark

Superpixel-based image noise variance estimation with local statistical assessment

Cheng-Ho Wu[1] and Herng-Hua Chang[2*]

Abstract

Noise estimation is fundamental and essential in a wide variety of computer vision, image, and video processing applications. It provides an adaptive mechanism for many restoration algorithms instead of using fixed values for the setting of noise levels. This paper proposes a new superpixel-based framework associated with statistical analysis for estimating the variance of additive Gaussian noise in digital images. The proposed approach consists of three major phases: superpixel classification, local variance computation, and statistical determination. The normalized cut algorithm is first adopted to effectively divide the image into a set of superpixel regions, from which the noise variance is computed and estimated. Subsequently, the Jarque–Bera test is used to exclude regions that are not normally distributed. The smallest standard deviation in the remaining regions is finally selected as the estimation result. A wide variety of noisy images with various scenarios were used to evaluate this new noise estimation algorithm. Experimental results indicated that the proposed framework provides accurate estimations across various noise levels. Comparing with many state-of-the-art methods, our algorithm strikes a good compromise between low-level and high-level noise estimations. It is suggested that the proposed method is of potential in many computer vision, image, and video processing applications that require automation.

Keywords: Gaussian noise, Noise estimation, Image denoising, Superpixel, Jarque–Bera test

1 Introduction

In the field of computer vision, signal, image, and video processing, noise is unfortunately inevitable during data acquisition and transmission. The accuracy of many algorithms significantly relies on well hand-tuned parameter adjustments to account for variations in noise [1–3]. To automate the process and achieve reliable procedures, the capability for accurate noise estimation is essential to motion estimation, edge detection, super-resolution, restoration, shape-from-shading, feature extraction, and object recognition [4–9]. In particular, image noise having a Gaussian-like distribution is quite often encountered, and it is characterized by adding to each pixel a random value obtained from a zero-mean Gaussian distribution, whose variance determines the magnitude of the corrupting noise. This zero-mean property enables such noise to be removed by locally averaging neighboring pixel values [10, 11].

Indeed, many noise reduction algorithms incorporate the knowledge of the noise level in the denoising process and assume that it is known a priori [12–15]. Accordingly, estimation for the amount of noise is critical in these methods, because it enables the process to adapt to the level of noise rather than using fixed values and thresholds. The challenge of noise estimation is to determine whether local image variations are due to color, texture, and lighting changes of images themselves, or caused by the noise. Nevertheless, existing noise estimation algorithms can be broadly classified into three major categories: filtered-based, block-based, and transform-based approaches [4, 5, 11, 16, 17].

In filtered-based methods, an input image is first filtered by a low-pass filter to smooth the structures and suppress the noise in the image [4]. The noise variance is then estimated from the difference between the noisy image and the filtered image. One fundamental problem of filtered-based methods is that the difference image is assumed to be the noise, but this assumption is not always true in general. This is because the low-pass filtered image is not equivalent to the original noise-free image, particularly

* Correspondence: herbertchang@ntu.edu.tw
[2]Computational Biomedical Engineering Laboratory (CBEL), Department of Engineering Science and Ocean Engineering, National Taiwan University, 1 Sec. 4 Roosevelt Road, Da'an District, 10617 Taipei City, Taiwan
Full list of author information is available at the end of the article

when the image is with strong structures and complicated details. To minimize the influence and obtain a realistic basis for noise level estimation, Rank et al. [18] proposed to use the vertical and horizontal information of an image to extract the noise detail and histogram information in the corresponding components. However, it has a relatively higher computation load and many user-defined parameters to be set.

For block-based algorithms, an image is tessellated into a number of blocks followed by noise variance computation in a set of homogeneous blocks [5, 17, 19]. The philosophy underlying this approach is that a homogeneous block in an image is treated as a perfectly smooth image block with added noise, which has a relatively higher chance to contain useful visual activities. Consequently, the block with a smaller standard deviation has a weaker variation in intensity, leading to a smoother block. One main difficulty of block-based approaches is how to efficiently identify the homogeneous blocks. Lee and Hoppel [20] estimated noise level by assuming that the smallest standard deviation of a block is equivalent to additive white Gaussian noise. This method is simple but tends to produce overestimation results for small noise cases. Shin et al. [5] split an image into a number of blocks, which were further classified by the standard deviation in intensity. An adaptive Gaussian filtering process was then applied to relatively flat blocks, where the noise was estimated from the difference of the selected blocks between the noisy image and its filtered image.

While noise estimation methods in the first two categories work directly on the pixel intensity in the spatial domain, transform-based methods seek particular features in the transformed domain [21]. For example, the median absolute deviation method [22–24] used wavelet coefficients to estimate noise standard deviation based on the assumption that wavelet coefficients in the diagonal subband HH_1 are dominated by noise. This approach provides good estimations for large noise cases, but it can overestimate the noise in small noise cases. The reason for overestimation is that wavelet coefficients in the diagonal subband contain not only added noise but also image details. Subsequently, Li et al. [25] proposed a modified noise estimation algorithm based on the wavelet coefficients in the HH_1 subband. Better results were obtained by reducing the estimated original image contribution from HH_1 comparing to the traditional methods. Liu and Lin [17] investigated the possibility to estimate noise in the singular value decomposition (SVD) domain. The authors used the tail of singular values to alleviate the influence of the signal in the noise estimation process and demonstrated the effectiveness of their method over wavelet-based approaches. However, due to the use of SVD twice in the estimation procedure, the computation time is more expensive.

Alternatively, there are other methods that estimate noise in various manners [26]. Immerkaer [27] proposed a Laplacian-based noise estimation algorithm, which computes the noise variance by convolving the image with a Laplacian-like mask with zero mean. This approach is fast and performs well on images that are corrupted by high level noise. However, for highly textured images, it perceives thin lines as noise, leading to overestimation. Tai and Yang [28] extended Immerkaer's work by introducing the Sobel operator for edge detection to exclude the edge pixels. Salmeri et al. [29] introduced different weights to various subregions based on a similarity measure followed by a fuzzy procedure to estimate the variance of noise. Zoran and Weiss [30] proposed a statistical model to estimate the variance of noise and showed the effectiveness on images with low-level noise. Their assumption is that adding noise to images results in changes to kurtosis values throughout the scales. In their approach, the image was first convolved by the DCT filter to produce a response image, from which the variance and kurtosis were estimated. Aja-Fernández et al. [16] presented a noise estimation method based on the mode of local statistics (MLS). The authors demonstrated the efficiency of using the mode of the local sample statistical distribution for the variance estimation of additive noise provided that a great amount of low-variability areas exist in the image.

Among existing noise estimation methods, block-based algorithms are relatively simple and straightforward. Nonetheless, one main issue of this approach is how to effectively identify the homogeneous regions while alleviating the dependence of various noise levels. To address this major challenge and overcome the drawbacks in the existing methods, this paper proposes a new noise estimation algorithm that automatically and efficiently divides an image into a number of homogeneous subregions, which are called superpixels. To reduce noise influences, a statistical decision is then made to select the best superpixel, from which the noise variance is estimated. The ambition is to improve the estimation accuracy in low level noise while maintaining precision for higher level noise comparing to existing methods. The remainder of this paper is organized as follows. In Section 2, the noise model along with the probability density function of the Gaussian noise is described. Section 3 introduces the proposed algorithm that consists of three major phases: superpixel classification, local variance computation, and statistical determination. In Section 4, the performance of this new noise estimation method using a wide variety of images with various scenarios is evaluated and compared with many state-of-the-art methods. Finally, Section 5 discusses the results and Section 6 summarizes the contributions of the current work.

2 Noise model

The fundamental assumption of the noise model is that the image is corrupted by additive, zero-mean white Gaussian noise with an unknown variance given by

$$I(x, y) = f(x, y) + n(x, y) \qquad (1)$$

where (x, y) represents the coordinates of a pixel under consideration, $I(x, y)$ is the observed image, $f(x, y)$ is the intact image, and $n(x, y)$ is the Gaussian noise, whose probability density function (PDF) can be written as follows:

$$p(z) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(z-\bar{z})^2/2\sigma^2} \qquad (2)$$

where z represents the intensity, $\bar{z}$ is the mean of z, and σ is the standard deviation used to control the shape of the distribution.

Figure 1 illustrates two eight-bit images corrupted by additive Gaussian noise with $\sigma = 10$ and the corresponding histogram maps. The original image in Fig. 1a has two uniform subregions with intensity values equal to 50

and 200, respectively. It is observed that the histogram distribution has two similar shapes correspondingly centered at the original intensity values after corruption. This is because the Gaussian noise model (Eq. (2)) is actually a normal distribution so that the histogram follows the normal distribution with the same standard deviation in each individual region provided that no influences occur. If the image is further divided into more subregions and the process is repeated as shown in Fig. 1b, the same observation will be obtained as illustrated in the region enclosed by the red box. Rather than estimating the noise level globally in the entire image, this paper proposes to classify the image into several subregions and compute the noise variance locally in each individual region to minimize the influence caused by color, texture, and lighting changes [1, 16, 17, 24].

3 Methods

The proposed noise estimation algorithm can be divided into three major phases as shown in Fig. 2 and described as follows.
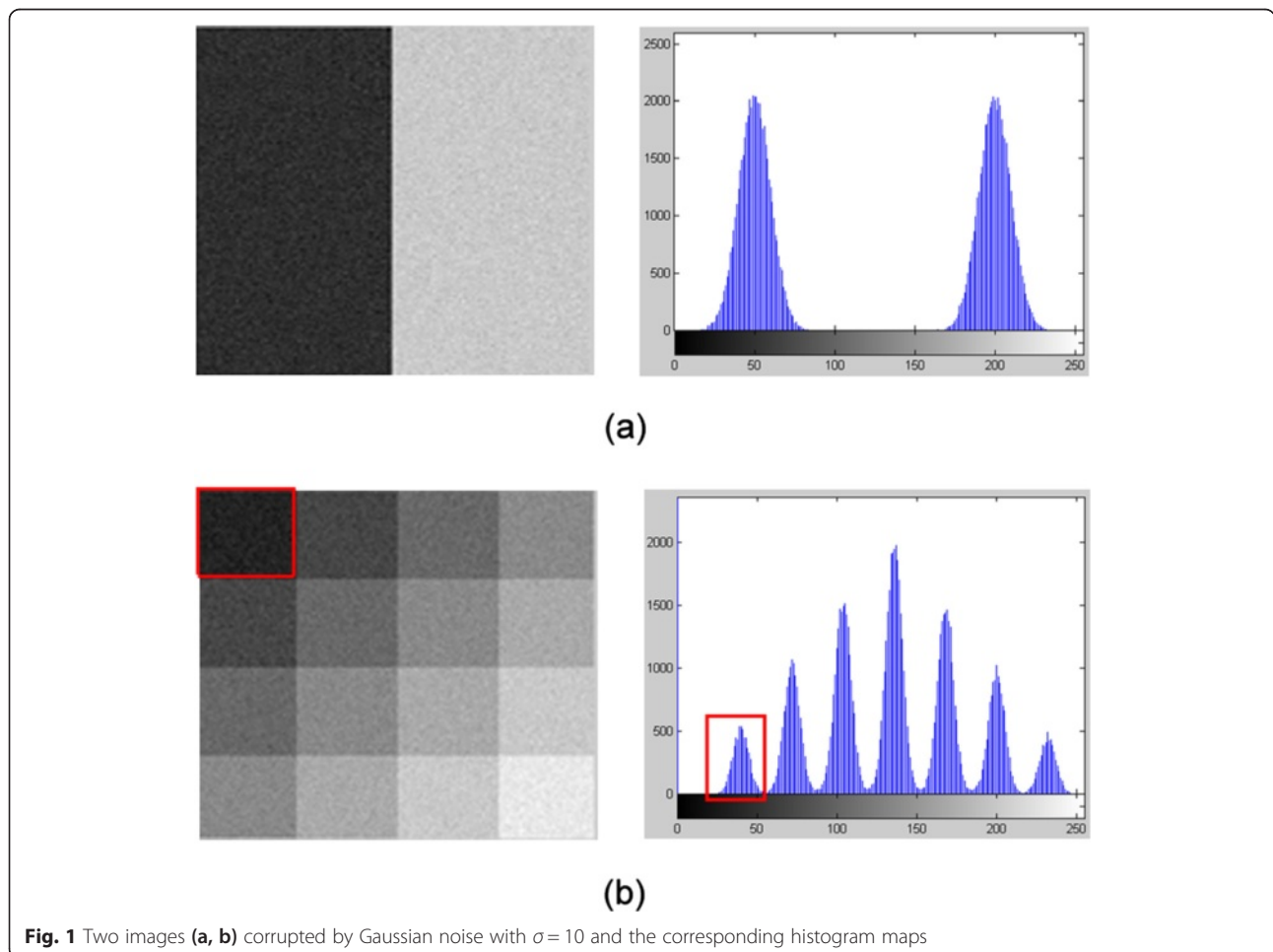
Fig. 1 Two images **(a, b)** corrupted by Gaussian noise with $\sigma = 10$ and the corresponding histogram maps
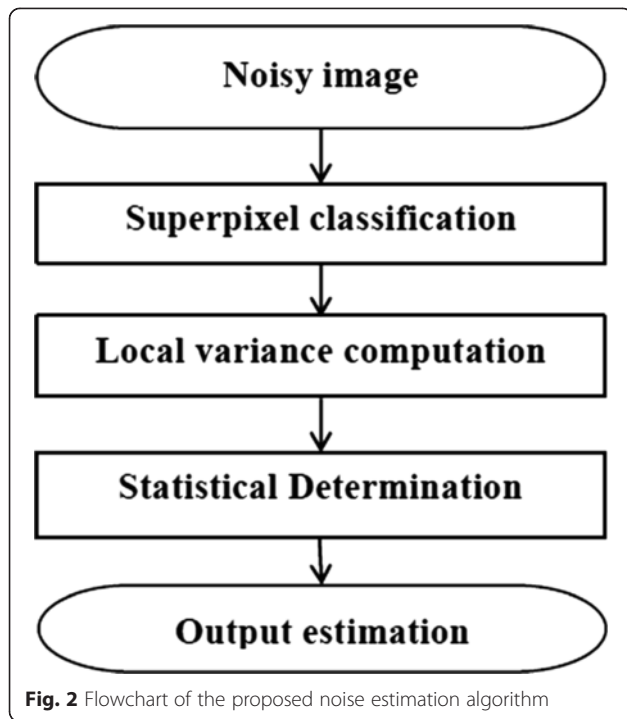
Fig. 2 Flowchart of the proposed noise estimation algorithm

3.1 Superpixel classification

The first step in our noise estimation framework is to divide a noisy image into several subregions. Unlike conventional block-based methods, each subregion is not necessary to be a rectangular block and it is usually not. In essence, each region is expected to have similar gray-level, color, and texture characteristics regardless of its geometry. To do this, the normalized cut algorithm [31] is adopted to achieve this goal. The basic idea is to use the theoretic criteria of graph to measure the goodness of an image partition. More specifically, it measures both the total dissimilarity between different groups as well as the total similarity within groups. The optimization of this criterion can be formulated as a generalized eigenvalue problem that can be efficiently solved. The concept of this perceptual grouping technique is briefly described as follows.

Given an image of N pixels, the set of pixels can be represented as a weighted undirected graph $G = (V, E)$, where V represents nodes of the graph corresponding to the pixels in the feature space, E represents edges that are formed between every pair of nodes. A weight $w(i, j)$ is assigned to each edge that captures the similarity between nodes i and j. In grouping, the goal is to partition the set of vertices into m disjoint sets $V_1, V_2,..., V_m$, where, by some measure, the similarity among the vertices in a set V_i is high while it is low across different sets.

For simplicity, a graph partitioned into two disjoint sets, A and B, is considered by simply removing edges connecting these two parts that satisfies $A \cup B = V, A \cap B = \Phi$. The degree of dissimilarity between these two sets can be computed as a total weight of the edges that have been removed, which is called the cut:

$$\text{cut}(A, B) = \Sigma_{u \in A, \, v \in B} w(u, v), \tag{3}$$

where the graph edge weight connecting two nodes i and j is defined as follows:

$$w(i, j) = exp\left(\frac{-\|I(i) - I(j)\|_2^2}{\sigma_I^2}\right)$$

$$\times \begin{cases} exp\left(\frac{-\|X(i) - X(j)\|_2^2}{\sigma_X^2}\right) & , \text{if } \|X(i) - X(j)\|_2^2 < r \\ 0, & \text{otherwise} \end{cases}$$

$$\tag{4}$$

where $X(i)$ and $X(j)$ are the spatial coordinates of nodes i and j, respectively. In Eq. (4), r is a prescribed threshold, $I(i)$ and $I(j)$ are the intensity values at the corresponding locations, and σ_I and σ_X are the standard deviations
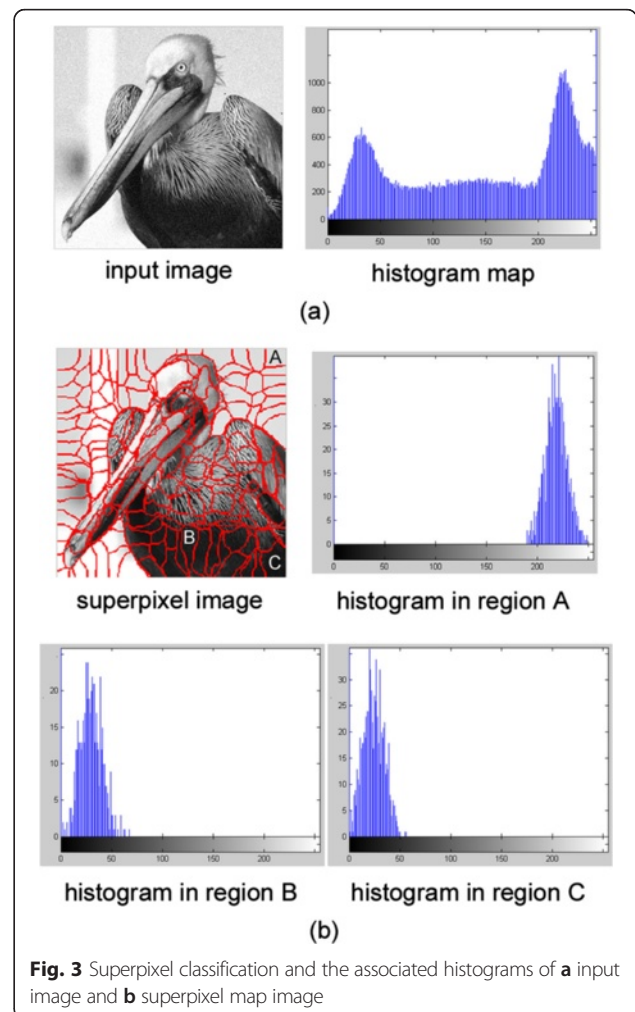
Fig. 3 Superpixel classification and the associated histograms of **a** input image and **b** superpixel map image

for the intensity component and the spatial component, respectively.

The normalized cut (Ncut) between two sets, A and B, is proposed to solve the problem of unnatural bias based on Eq. (3) in such a way to partition out small sets of pixels using the following:

$$\text{Ncut}(A,B) = \frac{\text{cut}(A,B)}{\text{assoc}(A,V)} + \frac{\text{cut}(A,B)}{\text{assoc}(B,V)}, \qquad (5)$$

where $\text{cut}(A,B)$ is the total weight of the edges that have been removed after a graph is partitioned into two disjoint sets A and B, $\text{assoc}(A,V) = \sum_{u \in A, t \in V} w(u,t)$ is the total connection from nodes in A to all nodes in the graph, and $\text{assoc}(B,V)$ is similarly defined. The challenge is to find optimal sets A and B such that $\text{Ncut}(A,B)$ in Eq. (5) is minimized. Unfortunately, minimizing the normalized cut is exactly NP-complete, even for the special case of graphs on grids. Based on the spectral graph
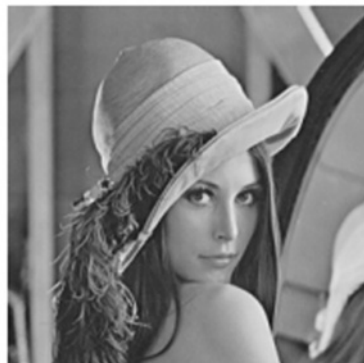

Fig. 4 Representative images for the experiments. The image dimension is 512×512

theory, an approximately discrete solution can be efficiently obtained by thresholding the eigenvector corresponding to the second smallest eigenvalue λ_2 of the generalized eigenvalue system with

$$(D-W)y = \lambda Dy, \tag{6}$$

where D is a diagonal matrix with entries D_{ii} given as

$$D_{ii} = d(i) = \sum_{j \in V} w(i, j) \tag{7}$$

which is the total connection from node i to all other nodes in the graph.

As illustrated in Fig. 3, the input image in Fig. 3a is classified into several subregions using the normalized cut algorithm as shown in Fig. 3b. Herein, each subregion is referred to as a "superpixel." The concept of superpixel is based on over-segmentation results, and a superpixel is local and coherent that preserves most of the structure at the scale of interest [32]. After the classification procedure, a set of regions R representing the superpixel map is obtained. Note that the histogram distribution of intensity in each superpixel is approximately a normal distribution centered at different intensity values as illustrated in Fig. 3b, comparing to the overall histogram distribution shown in Fig. 3a.

3.2 Local variance computation

After the superpixel classification procedure and obtaining R, local variance computation is performed inside each superpixel using

$$\mu_{Ri} = \sum_{x=1}^{n_i} I(R_i, x)/n_i, \quad i = 1, 2, ..., N \tag{8}$$

and

$$\sigma_{R_i}^2 = \sum_{x=1}^{n_i} \left(I(R_i, x) - \mu_{Ri}\right)^2/n_i, \quad i = 1, 2, ..., N, \tag{9}$$

where $I(R_i, x)$ is the intensity of pixel x in superpixel R_i, μ_{Ri} is the mean intensity in R_i, n_i is the number of all pixels in R_i, σ_{Ri}^2 is the variance in R_i, and N is the total number of superpixel regions in R.

Fig. 5 Some of the 100 images obtained from the Berkeley image database for the experiments. The image dimension is 481 × 321 or 321 × 481

3.3 Statistical determination

By now, the noise variance values for all superpixel regions in R are obtained. Intuitively, the smallest variance value should be selected for the noise variance estimation result. Practically, however, the variance is somewhat affected by the size, detail, and texture of each individual superpixel so that underestimation could occur. Since the probability distribution of the Gaussian noise is normal, the region that is most close to normal distribution is chosen as an estimation candidate. The Jarque–Bera (JB) test [33, 34],

which is a goodness-of-fit test, is used to decide whether sample data match a normal distribution based on the skewness and kurtosis. The statistical JB test is defined as follows:

$$\mathrm{JB} = \frac{n}{6}\left(S^2 + \frac{1}{4}(K-3)^2\right) \tag{10}$$

where n is the number of observations (or degree of freedom in general).
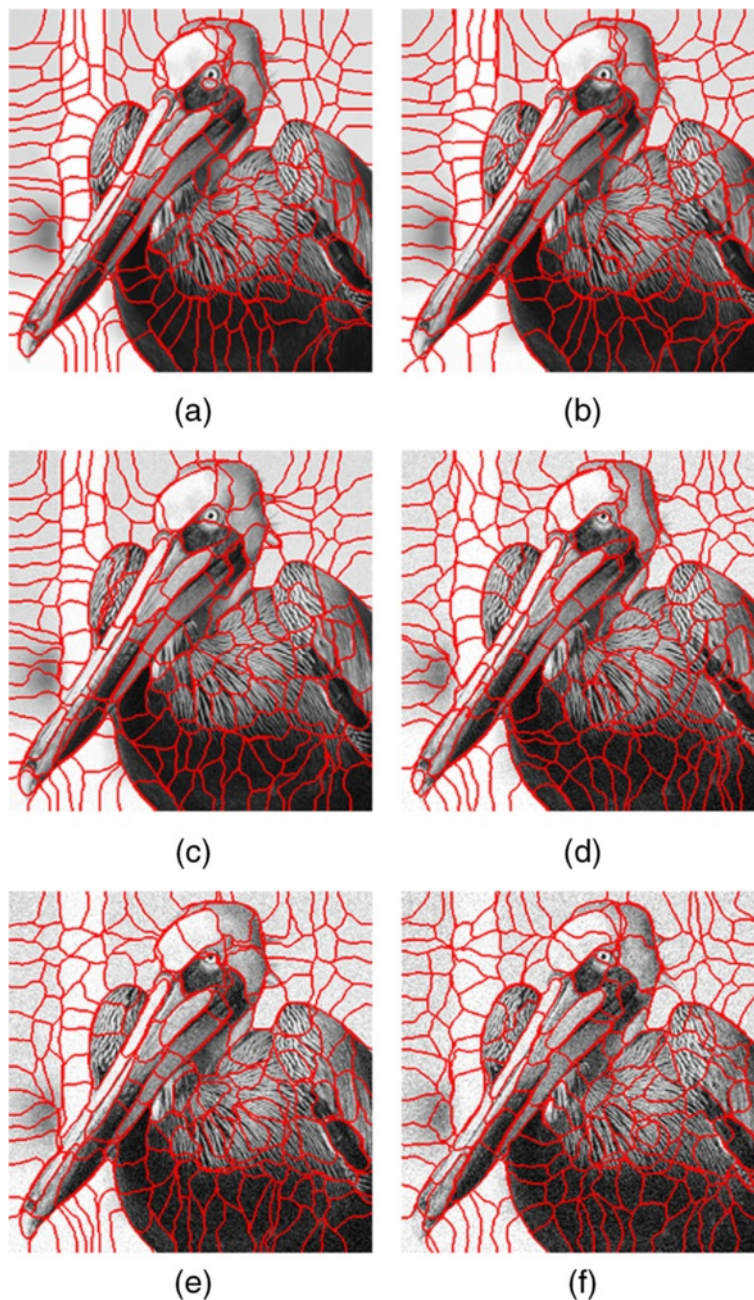
Fig. 6 Superpixel maps of Bird with different levels of Gaussian noise. **a** $\sigma_a = 1$. **b** $\sigma_a = 5$. **c** $\sigma_a = 10$. **d** $\sigma_a = 20$. **e** $\sigma_a = 30$. **f** $\sigma_a = 40$

In Eq. (10), S is the sample skewness and K is the sample kurtosis respectively defined as follows:

$$S = \frac{\hat{\mu}_3}{\hat{\sigma}_3} = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^3}{\left(\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2\right)^{3/2}} \quad (11)$$

and

$$K = \frac{\hat{\mu}_4}{\hat{\sigma}_4} = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^4}{\left(\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2\right)^2}, \quad (12)$$

where $\hat{\mu}_3$ and $\hat{\mu}_4$ are, respectively, the estimates of the third and fourth central moments; $\bar{x}$ is the sample mean; and $\hat{\sigma}^2$ is the estimate of the second central moment, i.e., the variance. If the data present a normal distribution, the JB statistic will have a chi-squared distribution with 2 degrees of freedom asymptotically. The null hypothesis is a joint hypothesis with both the skewness

Fig. 7 Noise estimation results with various levels: **a** Bird. **b** Countryroad

and the excess kurtosis being 0. Any deviation from this condition will increase the JB statistic. Accordingly, the null hypothesis based on the JB test is defined as follows:

$$JB_{Ri} = \begin{cases} 0, & \text{accept null hypothesis} \\ 1, & \text{reject null hypothesis} \end{cases}. \quad (13)$$

In other words, the JB value equals to 0 if the corresponding superpixel region is examined as a normal distribution; otherwise, it is set to 1.

After the JB test procedure in each individual superpixel region in R, the final noise estimation result is produced based on the following rules:

1. Sort R_i based on the standard deviation σ_{R_i} in ascending order.
2. Exclude the superpixel region whose JB value equals to 1.
3. Exclude the superpixel region whose pixel number is less than $10 \times \min(\sigma_{R_i})$, where $\min(\sigma_{R_i})$ is the smallest value of σ_{R_i} in region R.
4. Choose the smallest value of σ_{R_i} from the remaining regions as the noise estimation result.

The reason for excluding the regions with a small pixel number in rule 3 is due to the fact that these regions may not have an enough sample quantity to reflect the real noise distribution, leading to poor estimations. As the region size is somewhat related to the value of σ_{R_i}, the threshold is thus defined as it is and scaled by an experimental constant.

4 Experimental results

To evaluate the proposed algorithm, a wide variety of photographic images (512×512) were tested, and some representative images are shown in Fig. 4. In addition, the Berkeley image database [35] was adopted to evaluate the performance of the algorithm. The Berkeley database is a public domain that contains hundreds of images (481×321 or 321×481) of plants, animals, persons, landscapes, and architectures as partly illustrated in Fig. 5. Different levels of additive Gaussian noise were superimposed on those images to generate various noisy images for experiments. To assess the performance quantitatively, the relative error in terms of the standard deviation was computed as given in the following equation:

$$\varepsilon_r = \frac{|\sigma_e - \sigma_a|}{\sigma_a} \times 100 \%, \quad (14)$$

where σ_e represents the standard deviation of the estimated noise, σ_a represents the standard deviation of the added noise, and ε_r represents the relative percentage error between the added and estimated noise levels.
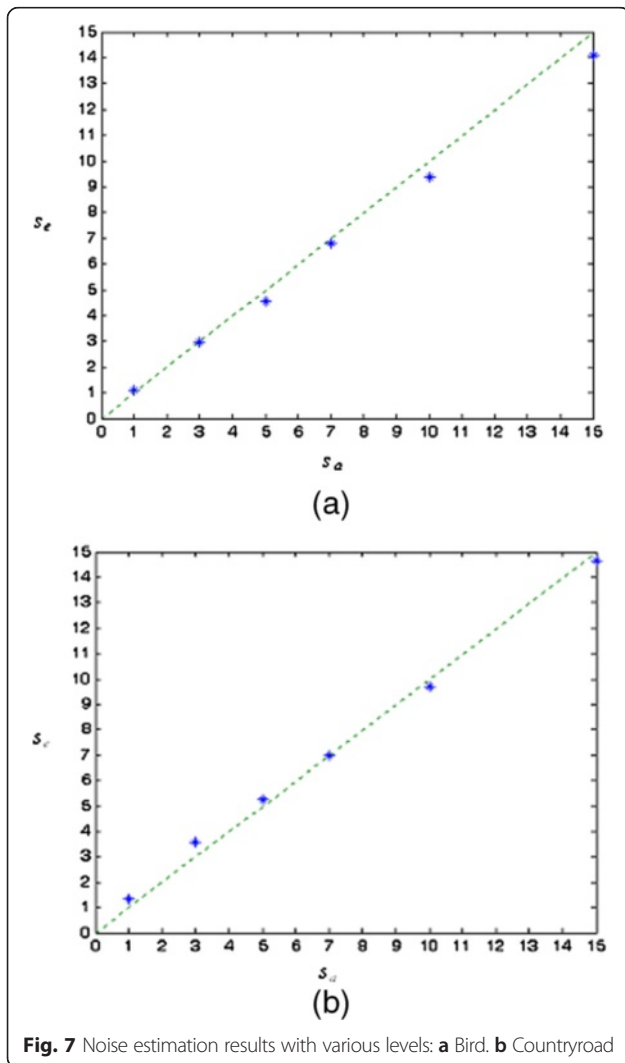
Table 1 Quantitative analysis in noise level estimation with σ_a = 1, 3, 5, 7, 10, 15, 20 on the Bird image using different methods

	JI96		Z&W09		MLS09		SVD13		Ours	
σ_a	σ_e	ε_r (%)	σ_e	ε_r (%)	σ_e	ε_r (%)	σ_e	ε_r (%)	σ_e	ε_r (%)
1	7.43	642.71	0.00	100.00	1.39	38.73	2.96	195.76	1.10	9.71
3	7.91	163.80	0.00	100.00	2.66	11.23	4.02	33.90	2.85	4.93
5	8.82	76.38	2.68	46.40	4.90	1.98	5.46	9.14	4.81	3.72
7	10.00	42.89	4.43	36.67	6.84	2.31	7.11	1.55	6.47	7.58
10	12.10	20.97	6.96	30.44	9.65	3.5	9.88	1.17	9.63	3.73
15	16.13	7.55	11.24	25.05	14.66	2.24	14.85	0.97	14.49	3.43
20	20.15	0.07	15.32	23.40	19.03	4.83	19.73	1.35	17.92	10.42
Average $\bar{\varepsilon}_r$		136.34		51.71		9.26		34.83		6.22

As the proposed algorithm was insensitive to parameter settings, all experiments were conducted using the same fixed parameters in the process. All compared methods were executed with appropriate parameter settings as suggested by the corresponding authors, if any. The robustness of the superpixel classification procedure based on the normalized cut algorithm with respect to different levels of additive Gaussian noise was first investigated. As illustrated in Fig. 6, the superpixel maps with different noise standard deviations varying from $\sigma_a = 1$ to $\sigma_a = 40$ were approximately similar from the perspective of smooth regions. Figure 7 shows the estimation results of using the proposed algorithm on the Bird and Country-road images, whose noise levels were quite widespread with $\sigma_a = 1$, 3, 5, 7, 10, and 15. It is obvious that the estimated values of σ_e are quite accurate and close to the corresponding values of σ_a in all levels of noise in both images. Table 1 summarizes the quantitative analysis in estimating various noise levels on the Bird image using the JI96 [27], Z&W09 [30], MLS09 [16], SVD13 [17], and the proposed algorithm. For $\sigma_a = 1$ and $\sigma_a = 3$, the proposed framework produced much higher accuracy than all other methods. For $\sigma_a = 5$ and $\sigma_a \geq 7$, the MLS09 and SVD13 methods outperformed all other methods, respectively.

However, both MLS09 and SVD13 methods severely overestimated small noise levels that resulted in the augmentation of the overall error. While the JI96 method overestimated and the Z&W09 method underestimated all noise levels, the proposed method provided consistent accuracy that achieved a small average error of 6.22 %.

Table 2 presents the statistical comparison in noise level estimation on all six representative images in Fig. 4 between JI96, Z&W09, MLS09, SVD13, and the proposed algorithm. The JI96 continued to overestimate the noise levels with $\sigma_a \leq 20$, which were particularly worse in low-level noise cases. Not only did the Z&W09 method underestimate all noise levels with $\sigma_a \leq 25$, but it also generated no results in larger noise levels with $\sigma_a \geq 30$ (not shown). Other three methods alternately produced the best estimation results on different noise levels. The proposed method and MLS09 achieved similar accuracy with the average error less than 12 %. Although MLS09 had slightly smaller error than our framework, it was not statistically significant based on only six images. For completeness, Table 3 shows the quantitative results of noise estimation using the JI96, Z&W09, MLS09, SVD13, and the proposed algorithm on 100 images, which were randomly selected

Table 2 Statistical comparison of noise level estimation on the six representative images in Fig. 4 between different methods

	JI96		Z&W09		MLS09		SVD13		Ours	
σ_a	$\sigma_e \pm$ std.	ε_r (%)	$\sigma_e \pm$ std.	ε_r (%)	$\sigma_e \pm$ std.	ε_r (%)	$\sigma_e \pm$ std.	ε_r (%)	$\sigma_e \pm$ std.	ε_r (%)
1	5.15 ± 2.15	415.23	0.92 ± 0.71	58.62	1.60 ± 0.36	59.91	1.75 ± 0.92	96.18	1.22 ± 0.48	38.12
3	5.97 ± 1.86	99.03	2.16 ± 1.34	35.92	3.18 ± 0.20	6.79	3.24 ± 0.61	17.53	3.20 ± 0.27	10.26
5	7.25 ± 1.56	44.98	4.3 ± 0.88	15.81	5.08 ± 0.22	3.64	5.08 ± 0.43	6.70	4.83 ± 0.37	6.27
7	8.77 ± 1.30	25.34	6.15 ± 0.84	12.25	7.04 ± 0.36	3.84	7.01 ± 0.36	3.76	7.34 ± 1.01	8.65
10	11.28 ± 1.00	12.80	8.89 ± 0.91	11.06	9.95 ± 0.46	3.34	9.97 ± 0.30	2.11	9.72 ± 0.57	5.85
15	15.79 ± 0.69	5.26	13.52 ± 1.06	9.84	14.63 ± 0.61	4.19	15.06 ± 0.34	1.49	14.05 ± 0.53	6.33
20	20.28 ± 0.67	2.28	17.99 ± 1.28	10.04	19.34 ± 0.53	3.29	20.16 ± 0.42	1.02	17.59 ± 1.43	12.03
25	24.76 ± 0.83	3.06	22.51 ± 1.63	9.97	24.10 ± 0.67	3.88	25.27 ± 0.21	1.10	23.44 ± 1.31	6.94
Average $\bar{\varepsilon}_r$		76.00		20.44		11.11		16.24		11.81

std. standard deviation

Table 3 Statistical comparison of noise level estimation on 100 images obtained from the Berkeley image database between different methods

	JI96		Z&W09		MLS09		SVD13		Ours	
σ_a	$\sigma_e \pm$ std.	ε_r (%)	$\sigma_e \pm$ std.	ε_r (%)	$\sigma_e \pm$ std.	ε_r (%)	$\sigma_e \pm$ std.	ε_r (%)	$\sigma_e \pm$ std.	ε_r (%)
1	5.68 ± 2.26	468.22	1.25 ± 1.61	114.72	2.51 ± 2.59	150.99	3.58 ± 3.35	302.81	1.87 ± 1.39	100.62
3	6.41 ± 2.05	113.78	2.75 ± 1.68	40.33	4.35 ± 2.67	46.09	4.40 ± 3.05	76.67	3.61 ± 1.31	32.50
5	7.59 ± 1.78	51.86	4.78 ± 1.59	22.60	6.22 ± 2.52	25.89	5.58 ± 2.76	38.80	5.42 ± 1.19	18.74
7	9.04 ± 1.55	29.14	6.62 ± 1.47	15.91	8.10 ± 2.25	17.43	6.98 ± 2.47	25.75	7.80 ± 1.49	17.37
10	11.44 ± 1.29	14.89	9.45 ± 1.37	12.55	10.91 ± 2.04	11.34	9.29 ± 2.10	18.07	10.57 ± 1.88	13.20
15	15.75 ± 1.09	6.61	14.12 ± 1.49	9.51	15.67 ± 1.62	6.66	13.49 ± 1.69	13.70	15.47 ± 1.83	8.76
20	20.20 ± 1.10	3.75	18.65 ± 1.69	8.95	20.36 ± 1.55	4.66	17.87 ± 1.43	12.01	19.71 ± 1.98	7.90
25	24.63 ± 1.28	3.25	23.04 ± 1.95	9.16	25.12 ± 1.47	3.93	22.26 ± 1.22	11.43	24.36 ± 1.92	6.15
30	29.02 ± 1.53	3.93	N/A	N/A	29.85 ± 1.47	3.46	26.78 ± 1.15	10.88	29.07 ± 1.86	5.87
35	33.30 ± 1.80	5.09	N/A	N/A	34.46 ± 1.54	3.73	31.15 ± 1.19	11.04	33.63 ± 2.24	5.49
40	37.47 ± 2.09	6.37	N/A	N/A	39.41 ± 1.52	3.22	35.65 ± 1.13	10.89	37.91 ± 3.13	6.95
Average $\bar{\varepsilon}_r$		64.26		N/A		25.22		48.37		20.32

N/A not applicable, *std.* standard deviation

from the Berkeley database [35], some of which are shown in Fig. 5. The JI96 method severely overestimated the Gaussian noise levels until the standard deviation $\sigma_a > 15$. On the other hand, the Z&W09 method performed better in lower level noise with $\sigma_a < 5$, but it was unable to handle noise with $\sigma_a > 25$. Both MLS09 and SVD13 methods provided higher accuracy in larger noise levels but achieved worse estimations in small noise levels, particularly for $\sigma_a = 1$. Nevertheless, the proposed scheme produced better accuracy in smaller noise levels and compatible estimations in large noise levels that resulted in a smaller overall error than all other methods.

5 Discussion

A new superpixel-based algorithm was proposed to estimate the additive Gaussian noise level in images. The approach relied on the normalized cut algorithm to classify the image in order to obtain the superpixel map, from which the local variance was computed and selected. As illustrated in Fig. 6, each superpixel region had similar gray-level, color, and texture characteristics such that it provided great flexibility for subsequent statistical analysis. Since the Gaussian noise obeys a normal distribution, we proposed to use the Jarque–Bera test [33, 34] to separate those superpixel regions that follow a Gaussian distribution from all other regions based on the skewness and kurtosis. After excluding the superpixel regions that had a relatively small number of pixels, the remaining smallest local standard deviation was selected as the noise estimation result.

The proposed framework was applied on a wide variety of images and compared with four state-of-the-art methods, namely JI96 [27], Z&W09 [30], MLS09 [16], and SVD13 [17]. As illustrated in Fig. 7, our technique

provided high accuracy for both simple (Countryroad) and complicated (Bird) texture images, particularly for low-level noise estimations. This great capability of excellent accuracy in low-level noise estimation can also be observed from Tables 1 and 2, which were experimented on the images shown in Fig. 4. In addition, the algorithm was extensively evaluated on 100 randomly selected images from the Berkeley database, which contained a wide diversity of photographic images. In comparison with the JI96, Z&W09, MLS09, and SVD13 methods, the proposed approach provided more accurate estimations for low-level noise images as well as smaller overall estimation errors as presented in Table 3. In general, the JI96 method performs better in images with high-level noise but inadequately for images with low-level noise. In contrast, the Z&W09 method performs better in images with lower level noise, but it fails to produce estimation results for larger noise levels with $\sigma_a > 25$. Both MLS09 and SVD13 methods can

Fig. 8 Image with high details and complicated textures (Berkeley image database: 86016)

Table 4 Performance analysis in noise level estimation on the image in Fig. 8 using different methods

	JI96		Z&W09		MLS09		SVD13		Ours	
σ_a	σ_e	ε_r (%)	σ_e	ε_r (%)	σ_e	ε_r (%)	σ_e	ε_r (%)	σ_e	ε_r (%)
1	10.83	983.20	9.90	890.33	11.61	1061.12	14.23	1323.03	4.48	348.05
3	11.23	274.27	10.59	252.90	11.49	282.95	14.67	388.92	8.58	186.15
5	11.89	137.74	11.59	131.86	12.46	149.16	14.74	194.84	9.51	90.25
7	12.91	84.46	13.18	88.26	13.14	87.74	15.33	118.96	10.79	54.11
10	14.77	47.74	15.70	57.04	15.33	53.27	17.05	70.52	13.23	32.25
15	18.45	23.00	20.04	33.57	18.00	20.03	19.97	33.13	16.70	11.33
20	22.71	13.56	24.61	23.04	23.12	15.59	22.84	14.19	20.88	4.40
25	27.26	9.05	29.23	16.92	25.20	0.81	27.25	9.00	23.92	4.32
30	31.79	5.97	N/A	N/A	30.59	1.95	31.68	5.60	27.28	9.07
35	36.17	3.34	N/A	N/A	34.53	1.35	34.57	1.23	33.67	3.81
40	40.41	1.02	N/A	N/A	34.66	13.35	39.59	1.02	38.42	3.95
Average $\bar{\varepsilon}_r$		142.17		N/A		153.39		196.40		67.97

provide satisfactory results in larger noise level estimation but they may notably overestimate small noise levels.

While our algorithm strikes a good compromise between the JI96, Z&W09, MLS09, and SVD13 methods in providing better estimation results across various noise levels. One limitation of this new noise estimation algorithm is that it could overestimate small noise levels when the image is with high details and complicated textures as illustrated in Fig. 8. Table 4 summarizes the performance of the proposed algorithm along with the JI96, Z&W09, MLS09, and SVD13 methods in estimating the noise levels using the image in Fig. 8. Note that all five methods overestimated the noise levels when $\sigma_a \leq 20$ due to the highly sandy pattern over the entire image. In particular, the poor performance of SVD13 (see Tables 3 and 4) may be due to the singular value decomposition of rectangle images. Lastly, the computation time of our framework was expensive comparing with other tested methods as presented in Table 5. Nevertheless, the proposed algorithm still outperformed all other methods with closer standard deviation and smaller average error across various noise levels.

6 Conclusions

In summary, a new algorithm for additive Gaussian noise level estimation is described, which consists of three major phases: superpixel classification, local variance

Table 5 Comparison of computation time in seconds between tested methods based on different dimensions of images

Size	JI96	SVD13	MLS09	Z&W09	Ours
100 × 100	0.02	0.16	0.32	1.2	3.2
200 × 200	0.02	0.20	0.33	1.2	13.6
300 × 300	0.02	0.26	0.33	1.8	34.8
400 × 400	0.02	0.33	0.35	2.4	85.2
500 × 500	0.02	0.44	0.39	3.6	152.0

computation, and statistical determination. The algorithm strikes a good compromise between low-level and high-level noise estimations. Hundreds of images with various subjects, scenes, textures, and structures were used to evaluate the proposed framework. Experimental results demonstrated the feasibility and effectiveness of the algorithm in providing accurate estimation results across a wide range of noise levels. This robust noise estimation framework is advantageous to automating denoising algorithms that require noise variance information. Moreover, the proposed noise estimation algorithm is of potential and promising in computer vision, image, and video processing applications. Further research is needed to more effectively divide the image into appropriate superpixels, to investigate the incorporation of filtered-based techniques, and to accelerate the computation for real-time applications.

Author details
[1]Graduate Institute of Networking and Multimedia, National Taiwan University, Taipei City, Taiwan. [2]Computational Biomedical Engineering Laboratory (CBEL), Department of Engineering Science and Ocean Engineering, National Taiwan University, 1 Sec. 4 Roosevelt Road, Da'an District, 10617 Taipei City, Taiwan.

References
1. C Liu, WT Freeman, R Szeliski, SB Kang, Noise estimation from a single image. Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on **1**, 901–908 (2006)
2. E Farzana, M Tanzid, KM Mohsin, MIH Bhuiyan, Bilateral filtering with adaptation to phase coherence and noise. Signal, Image and Video Processing **7**, 367–376 (2013)

3. J Kwang Myung, P Nam In, K Hong Kook, C Myung Kyu, H Kwang Il, Mechanical noise suppression based on non-negative matrix factorization and multi-band spectral subtraction for digital cameras. IEEE Trans. Consum. Electron. **59**, 296–302 (2013)

4. SI Olsen, Estimation of noise in images: an evaluation. CVGIP: Graphical Models and Image Processing **55**, 319–323 (1993)

5. D-H Shin, R-H Park, S Yang, J-H Jung, Block-based noise estimation using adaptive Gaussian filtering. IEEE Trans. Consum. Electron. **51**, 218–226 (2005)

6. WT Freeman, EC Pasztor, OT Carmichael, Learning low-level vision. Int. J. Comput. Vision **40**, 25–47 (2000)

7. S Baker, I Matthews, Lucas-Kanade 20 years on: a unifying framework. Int. J. Comput. Vision **56**, 221–255 (2004)

8. G Calvagno, F Fantozzi, R Rinaldo, A Viareggio, Model-based global and local motion estimation for videoconference sequences. IEEE Trans. Circuits Syst. Video Technol. **14**, 1156–1161 (2004)

9. A Bosco, A Bruna, G Messina, G Spampinato, Fast method for noise level estimation and integrated noise reduction. IEEE Trans. Consum. Electron. **51**, 1028–1033 (2005)

10. AC Bovik, *Handbook of Image And Video Processing* (Elsevier Academic Press, San Diego, CA, USA, 2005)

11. F Luisier, T Blu, M Unser, Image denoising in mixed Poisson-Gaussian noise. IEEE Trans. Image Process. **20**, 696–708 (2011)

12. H-H Chang, W-C Chu, Restoration algorithm for image noise removal using double bilateral filtering. J. Electron. Imaging **21**, 023028–1 (2012)

13. HH Chang, Entropy-based trilateral filtering for noise removal in digital images. Image and Signal Processing (CISP), 2010 3rd International Congress on **2**, 673–677 (2010)

14. J Portilla, V Strela, MJ Wainwright, EP Simoncelli, Image denoising using scale mixtures of Gaussians in the wavelet domain. IEEE Trans. Image Process. **12**, 1338–1351 (2003)

15. L Sangkeun, Edge statistics-based image scale ratio and noise strength estimation in DCT-coded images. IEEE Trans. Consum. Electron. **55**, 2139–2144 (2009)

16. S Aja-Fernández, G Vegas-Sánchez-Ferrero, M Martín-Fernández, C Alberola-López, Automatic noise estimation in images using local statistics. Additive and multiplicative cases. Image Vis. Comput. **27**, 756–770 (2009)

17. W Liu, W Lin, Additive white Gaussian noise level estimation in SVD domain for images. IEEE Trans. Image Process. **22**, 872–883 (2013)

18. K Rank, M Lendl, R Unbehauen, Estimation of image noise variance. Vision, Image and Signal Processing, IEE Proceedings **146**, 80–84 (1999)

19. F Shuai, S Qiang, C Yang, Segmentation methods for noise level estimation and adaptive denoising from a single image, in *Control and Decision Conference (CCDC), 2013 25th Chinese*, 2013, pp. 2839–2844

20. JS Lee, K Hoppel, JS Lee, K Hoppel, Noise modeling and estimation of remotely-sensed images. Geoscience and Remote Sensing Symposium, 1989. IGARSS'89. 12th Canadian Symposium on Remote Sensing., 1989 International **2**, 1005–1008 (1989)

21. C-J Zou, Entropy-based estimation of salt-pepper noise in wavelet domain, in *Wavelet Active Media Technology and Information Processing (ICCWAMTIP), 2013 10th International Computer Conference on*, 2013, pp. 366–370

22. DL Donoho, De-noising by soft-thresholding. IEEE Trans. Inf. Theory **41**, 613–627 (1995)

23. DL Donoho, JM Johnstone, Ideal spatial adaptation by wavelet shrinkage. Biometrika **81**, 425–455 (1994)

24. SG Chang, Y Bin, M Vetterli, Adaptive wavelet thresholding for image denoising and compression. IEEE Trans. Image Process. **9**, 1532–1546 (2000)

25. T Li, M Wang, T Li, Estimating noise parameter based on the wavelet coefficients estimation of original image. Challenges in Environmental Science and Computer Engineering (CESCE), 2010 International Conference on **1**, 126–129 (2010)

26. S YiChang, V Kwatra, T Chinen, F Hui, S Ioffe, Joint noise level estimation from personal photo collections, in *Computer Vision (ICCV), 2013 IEEE International Conference on*, 2013, pp. 2896–2903

27. J Immerkar, Fast noise variance estimation. Comput. Vis. Image Underst. **64**, 300–302 (1996)

28. S-C Tai, S-M Yang, *A Fast Method for Image Noise Estimation Using Laplacian Operator and Adaptive Edge Detection*. Communications, Control and Signal Processing, 2008. ISCCSP 2008. 3rd International Symposium on, 2008, pp. 1077–1081

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

30. D Zoran, Y Weiss, *Scale Invariance and Noise in Natural Images. Computer Vision, 2009 IEEE 12th International Conference on*, 2009, pp. 2209–2216

31. J Shi, J Malik, Normalized cuts and image segmentation. IEEE Trans. Pattern Anal. Mach. Intell. **22**, 888–905 (2000)

32. X Ren, J Malik, Learning a classification model for segmentation. Proceedings of the Ninth IEEE International Conference on Computer Vision - Volume **2**, 10 (2003)

33. CM Jarque, AK Bera, A Test for Normality of Observations and Regression Residuals. Longman. International Statistical Review. **55**(2), 163-172 (1987).

34. CM Jarque, AK Bera, Efficient tests for normality, homoscedasticity and serial independence of regression residuals. Ecol. Lett. **6**, 255–259 (1980)

35. DR Martin, C Fowlkes, D Tal, J Malik, *A Database of Human Segmented Natural Images and its Application to Evaluating Segmentation Algorithms and Measuring Ecological Statistics, EECS Department, University of California, Berkeley UCB/CSD-01-1133*, 2001