

# Noise Level Estimation for Natural Images Based on Scale-Invariant Kurtosis and Piecewise Stationarity

Li Dong, *Student Member, IEEE*, Jiantao Zhou, *Member, IEEE*, and Yuan Yan Tang, *Fellow, IEEE*

**Abstract**—Noise level estimation is crucial in many image processing applications, such as blind image denoising. In this paper, we propose a novel noise level estimation approach for natural images by jointly exploiting the piecewise stationarity and a regular property of the kurtosis in bandpass domains. We design a  $K$ -means-based algorithm to adaptively partition an image into a series of non-overlapping regions, each of whose clean versions is assumed to be associated with a constant, but unknown kurtosis throughout scales. The noise level estimation is then cast into a problem to optimally fit this new kurtosis model. In addition, we develop a rectification scheme to further reduce the estimation bias through noise injection mechanism. Extensive experimental results show that our method can reliably estimate the noise level for a variety of noise types, and outperforms some state-of-the-art techniques, especially for non-Gaussian noises.

**Index Terms**—Noise level estimation, scale invariant feature, kurtosis, piecewise stationarity.

## I. INTRODUCTION

NOISE is inevitably incurred during visual data acquisition, transmission, processing and storage [1]. Noise removing is therefore important in many practical scenarios. Most of the existing denoising schemes assumed the availability of the noise level [2], [3], and the denoising performance highly depends on the accuracy of the noise level fed into the denoiser [4]. Unfortunately, in practice, the noise level is often unknown beforehand, which has triggered the research on developing effective noise level estimation techniques. It is

Manuscript received July 22, 2016; revised October 17, 2016 and November 23, 2016; accepted December 9, 2016. Date of publication December 13, 2016; date of current version January 5, 2017. This work was supported in part by the Macau Science and Technology Development Fund under Grant FDCT/046/2014/A1, Grant FDCT/100/2012/A3, and Grant FDCT/026/2013/A, in part by the Research Committee at the University of Macau under Grant MYRG2014-00031-FST, Grant MYRG2015-00056-FST, Grant MYRG2016-00137-FST, Grant MYRG2015-00049-FST, Grant MYRG2015-00050-FST, and Grant RDG009/FST-TYY/2012, in part by the National Science Foundation of China under Grant 61402547 and Grant 61672114, and in part by the Macau-China join Project under Grant 008-2014-AMJ. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Alin M. Achim. (*Corresponding author:* Jiantao Zhou)

L. Dong and Y. Y. Tang are with the Department of Computer and Information Science, Faculty of Science and Technology, University of Macau, Macau 999078, China (e-mail: yb47452@umac.mo; yytang@umac.mo).

J. Zhou is with the Department of Computer and Information Science, Faculty of Science and Technology, University of Macau (UM), Macau 999078, China, and also with the UMacau Zhuhai Research Institute, Zhuhai 519080, China (e-mail: jtzhou@umac.mo).

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.2016.2639447

worth noting that the noise level estimation has been utilized in many other areas as well, e.g., motion deblurring [5], image segmentation [6], super-resolution [7], image restoration [8] and object recognition [9]. Recently, there are some emerging applications in which noise level estimation is used as a tool for detecting image forgery [10], [11].

Generally, estimating noise level from a single image is an ill-posed problem due to the insufficient prior information of noise. Out of analytical tractability, the noise is commonly assumed with certain parametric model. The noise level estimation problem then becomes to blindly infer the parameters associated with the predetermined model. For instance, by modeling the noise with zero-mean additive white Gaussian noise (AWGN), the goal turns to estimate the standard deviation. However, even when noise can be characterized with simplified model, noise level estimation for natural images is still challenging, where the main difficulty lies in effectively distinguishing the image signal from noise.

The existing methods for estimating the noise level can be roughly categorized into three classes: filter-based, patch-based and statistics-based. Filter-based methods apply high-pass filtering on the noisy images to suppress the image signals, and then estimate the noise level from the residual images. In early attempts, Immerkaer [12] designed a filter which is insensitive to image structures, and computed the noise variance by averaging the convolved images. To further reduce the negative effect caused by image edges, Tai and Yang [13] and Yang and Tai [14] suggested to apply edge detection before filtering, and improved the estimation accuracy at low noise levels. In [15], Rank *et al.* proposed to filter the noisy images by a cascade of two 1-D difference operators, and estimated the noise variance as the weighted sum of local variances. As pointed out in [16], the filter-based schemes tend to make overestimates, especially for highly textured images at low noise levels.

Instead of using the whole image, patch-based methods estimate the noise level from carefully selected image patches. Early works attempted to collect homogeneous patches that are the most suitable for noise level estimation. Amer and Dubois [16] designed a series of high-pass operators and masks to select the homogeneous patches for noise level estimation. Similarly, Shin *et al.* [17] selected the smooth patches based on the standard deviation of signal intensities. More recently, by considering the non-local self-similarity

property of natural images, Pyatykh *et al.* [18] applied principal component analysis (PCA) on selected patches, and estimated noise level from the smallest eigenvalue of the covariance matrix. In [4], Liu *et al.* designed an iterative procedure to adaptively select patches for noise level estimation. Wu and Chang [19] recently proposed a superpixel-based noise variance estimation scheme, where irregular shaped regions were exploited.

Meanwhile, statistics-based methods estimate the noise level by establishing a deterministic relationship between noise level with some observable, statistical measurements. Early works focused on low-order statistics. Donoho and Johnstone [20] proposed a robust noise level estimator based on median absolute deviation of wavelet coefficients at the finest scale. This estimator has been widely adopted in the wavelet soft-thresholding denoisers [21], [22]. To further reduce the negative impact resulting from strong image structures, Li *et al.* [23] suggested to compute the wavelet coefficient of the Canny edge map from the noisy image, which is used to steer the exclusion of wavelet coefficient dominated by image signals. In [24], Stefano *et al.* developed a training based approach in wavelet domain by defining a set of non-linear statistical noise estimation functions. Liu and Lin [25] devised an efficient noise level estimation method based on singular value decomposition (SVD) domain analysis. Recently, significant improvement has been achieved by resorting to high-order statistics, among which the kurtosis-based methods were extensively studied. Along this line, Zoran and Weiss [26] examined the kurtosis statistic of DCT marginal filter responses. By assuming that the kurtosis values are invariant (constant) throughout all scales for clean natural images, they linked the noise variance with kurtosis values. The noise level estimation was finally cast into a nonlinear optimization problem. Following Zoran's work, Zhai and Wu [27] studied the skewness and kurtosis of the transformed coefficients affected by the spatial additive i.i.d. noise, and formulated two nonlinear constrained optimization problems for efficient noise level estimation. Lyu *et al.* [10] thoroughly investigated the kurtosis of natural images in various linear transform domains, and provided empirical evidences and theoretical justifications for the scale-invariant kurtosis assumption made by Zoran. The noise variance estimation was eventually formulated as an optimization problem with closed-form solution. To eliminate the kurtosis outliers caused by highly directional edges, Tang *et al.* [28] proposed a dual-transform based estimation method, in which directional DCT bands were intentionally excluded.

In this work, we stick to kurtosis-based noise level estimation, and propose a generalized assumption on scale-invariant kurtosis by taking the piecewise stationarity of image signal into account. Specifically, we suggest that the kurtosis scale-invariant property only holds in band-pass domains for those image regions exhibiting similar statistical behavior. We design a  $K$ -means based algorithm to adaptively partition the whole images into non-overlapping regions, each of whose clean versions is assumed to be associated with a constant, but unknown kurtosis throughout all scales. The noise level estimation is then formulated as a problem to optimally fit

this new kurtosis model. It is shown that such optimization problem, though non-convex, can be solved efficiently by resorting to an alternating iterative optimization procedure. In addition, we develop a rectification scheme to further reduce the estimation bias through noise injection mechanism. Extensive experimental results show that our method can reliably estimate the noise level for a variety of noise types, and outperforms some state-of-the-art techniques, especially for non-Gaussian noises. The main contributions of this work can be summarized as follows:

- Propose a generalized assumption on scale-invariant kurtosis by considering the piecewise stationarity of images.
- Suggest to adaptively partition an image into a series of non-overlapping regions. A noise level estimation approach can then be developed to optimally fit this generalized kurtosis model.
- Design a simple yet effective noise injection-based scheme to further boost the estimation accuracy of the noise level.

The rest of this paper is organized as follows. In Section II, we briefly review the scale-invariant kurtosis assumption. We propose the generalized kurtosis model, and apply it to noise level estimation in Section III. We investigate the noise level estimation for non-Gaussian noises in Section IV. Experimental results on synthetic and real noisy images are given in Section V. Finally, we conclude in Section VI.

## II. NOISE IN NATURAL IMAGES AND SCALE-INVARIANT KURTOSIS ASSUMPTION

An important topic in natural image statistical analysis is to study the coefficient distribution in the transform domain, where the cumulants and moments are fundamental statistical tools [29], [30]. Since low-order statistics such as mean and variance cannot provide adequate information regarding the distribution, high-order statistical quantities such as kurtosis has been widely employed [31]. Specifically, for a random variable  $X$ , its kurtosis is defined as [32]

$$\kappa(X) \triangleq \frac{C_4(X)}{C_2^2(X)} \quad (1)$$

where  $C_k(\cdot)$  is the  $k$ -th cumulant function. Essentially, kurtosis measures the peakedness of a distribution. For Gaussian distribution, the kurtosis value is 0, while for a distribution more (less) concentrated than Gaussian distribution, its kurtosis is positive (negative). One notable property of kurtosis is scale-invariant, namely,  $\kappa(aX) = \kappa(X)$  for any  $a > 0$ . Lam and Goodman [33] suggested that the kurtosis would be constant throughout scales for DCT filter marginal response distributions, while Bethge [34] argued that high frequency bands tend to have less kurtotic distributions than the lower frequency ones. To resolve this controversy, Zoran and Weiss [26] examined the kurtosis of DCT marginal filter response distributions, and made the following assumption: **kurtosis should be constant throughout scales in clean images, and any systematic changes are due to the added white noise**. Based on this assumption, Zoran proposed an effective noise level estimation approach

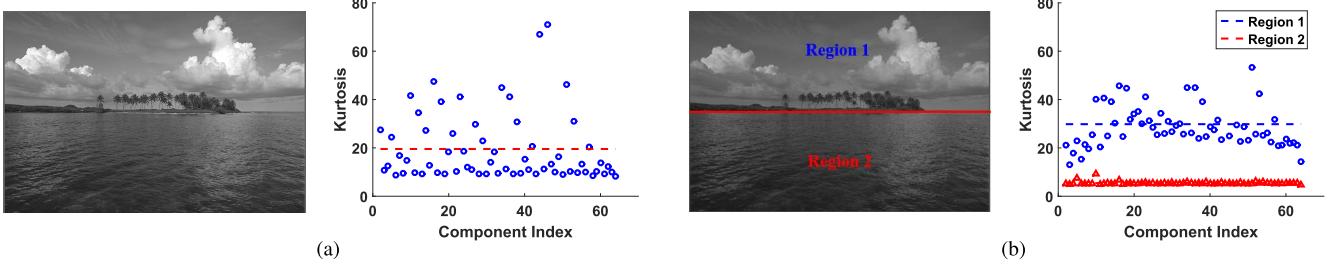


Fig. 1. (a) original clean image together with its kurtosis distribution calculated over the whole image; (b) partitioned images and their kurtosis distributions for different regions. Dash line represents the mean of kurtosis.

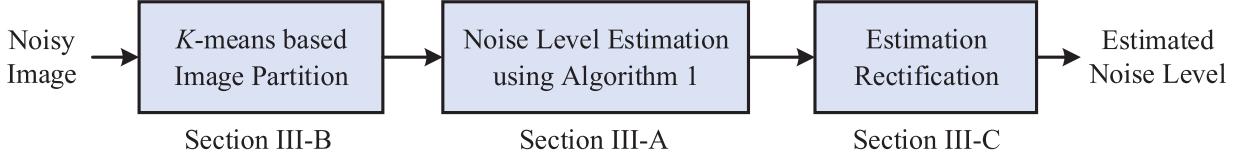


Fig. 2. Block diagram of our proposed method.

by explicitly establishing the relationship between kurtosis and noise variance. Lyu *et al.* [10] validated this assumption in various linear transform domains such as DCT and PCA. They asserted that the kurtosis of transformed coefficients should be constant throughout scales regardless the types of linear transform.

Note that in the above assumption on scale-invariant kurtosis, the variety in spatial/pixel domain has been totally neglected. The fact that semantically meaningful image constructs, such as edges and surface textures, are formed by spatially coherent contiguous pixels, suggests *piecewise*, rather than *global*, statistical stationarity of the image signal. An inspiring example is given in Fig. 1. As can be observed from Fig. 1(a), the kurtosis values generally fluctuate around the mean, and the magnitude of deviation is non-negligible, if the kurtosis values are computed from the whole image. However, when we roughly partition the image into two regions with different characteristics, as shown in Fig. 1(b), and calculate the kurtosis for these two regions individually, we find that the fluctuation can be significantly suppressed. Therefore, for a given clean image, it would be more reasonable to assume that the kurtosis values remain constant or near constant in a small locality, although they may and often do vary significantly in different segments of a scene. This observation motivates us to generalize the previous kurtosis-invariant assumption by taking the piecewise stationarity inherent to image signal into account. As will become clear shortly, such generalized new kurtosis model is capable of benefiting the noise level estimation as well.

### III. PROPOSED NOISE LEVEL ESTIMATION METHOD

An overview of the proposed noise level estimation method is illustrated in Fig. 2. The input noisy image is first partitioned into a series of non-overlapping blocks by using the method to be presented in Section III-B. We then apply the proposed noise level estimation algorithm to the partitioned blocks,

as will be detailed in Section III-A. The estimation performance is finally boosted via the estimation rectification module, as will be described in Section III-C.

We first give the details on the core part, i.e., the noise level estimation based on scale-invariant kurtosis and piecewise stationarity.

#### A. Noise Level Estimation Based on Scale-Invariant Kurtosis and Piecewise Stationarity

In this work, we first assume the spatial domain additive i.i.d. noise model

$$\mathbf{y} = \mathbf{x} + \mathbf{n} \quad (2)$$

where  $\mathbf{x}$  and  $\mathbf{y}$  denote the vectorized clean and noisy image patches, respectively. Here,  $\mathbf{n}$  is zero-mean, signal-independent AWGN with variance  $\sigma_n^2$ , which is unknown and shall be estimated. Instead of performing the noise level estimation over spatial domain, we consider the following noise model in the linear transform domain

$$\underbrace{\mathbf{d}_j' \mathbf{y}}_{\mathbf{y}_j} = \underbrace{\mathbf{d}_j' \mathbf{x}}_{\mathbf{x}_j} + \underbrace{\mathbf{d}_j' \mathbf{n}}_{n_j} \quad (3)$$

where  $\mathbf{d}_j$  denotes the  $j$ th basis vector from a complete orthonormal base matrix  $\mathbf{D}$ , and  $\mathbf{y}_j$ ,  $\mathbf{x}_j$ ,  $n_j$  are the responses of  $\mathbf{y}$ ,  $\mathbf{x}$  and  $\mathbf{n}$  in this  $j$ th band-pass channel, respectively. In the sequel, we also call  $j$  the scale index, and its value ranges from 1 to a predetermined constant  $M$ . The key advantage of the noise level estimation methods operated in the transform domain is that they provide a unified way to handle a variety of noise types, including both Gaussian and many non-Gaussian cases [26]. This is because linear transforms could mix the Gaussian or non-Gaussian noises in the spatial domain into Gaussian noise in the transform domain. For clarity, we currently focus on the Gaussian noise case. The non-Gaussian noise cases will be discussed in Section IV.

Due to the orthonormality of  $\mathbf{D}$ , the noise term  $n_j$  can be treated as a random variable following the zero-mean Gaussian

distribution  $n_j \sim \mathcal{N}(0, \sigma_n^2)$ . As  $x_j$  and  $n_j$  are independent, from (3) we have

$$\sigma_{y_j}^2 = \sigma_{x_j}^2 + \sigma_n^2 \quad (4)$$

$$C_4(y_j) = C_4(x_j) + C_4(n_j) \quad (5)$$

Noticing  $C_2(x_j) = \sigma_{x_j}^2$ , we can expand (5) as

$$\sigma_{y_j}^4 \kappa(y_j) = \sigma_{x_j}^4 \kappa(x_j) + \sigma_{n_j}^4 \kappa(n_j) \quad (6)$$

The above (6) shows the fundamental relationship between kurtosis and noise variance, namely,  $\kappa(y_j)$  can be represented by the linear combination of  $\kappa(x_j)$  and  $\kappa(n_j)$ . Since  $n_j$  follows a Gaussian distribution,  $\kappa(n_j) = 0$  holds, which eliminates the last term of (6). Considering the fact that the distribution of the transformed coefficients usually tends to be leptokurtic (more concentrate than Gaussian distribution) for natural images [34], we have  $\kappa(x_j), \kappa(y_j) \geq 0$ . Replacing  $\sigma_{x_j}^2$  with  $\sigma_{y_j}^2 - \sigma_n^2$ , (6) can be rewritten as

$$\sqrt{\kappa(y_j)} = \sqrt{\kappa(x_j)} - \frac{\sigma_n^2}{\sigma_{y_j}^2} \sqrt{\kappa(x_j)} \quad (7)$$

There are two unknowns in (7): the noise variance  $\sigma_n^2$  to be estimated, and the kurtosis  $\kappa(x_j)$  associated with the clean image. The remaining  $\kappa(y_j)$  and  $\sigma_{y_j}^2$  can be calculated from the noisy image and hence are available. Essentially, this equality links the unknown  $\sigma_n^2$  with the observable statistical quantities  $\kappa(y_j)$  and  $\sigma_{y_j}^2$ . However, as the number of unknowns is larger than that of the equations, the noise variance  $\sigma_n^2$  cannot be directly determined. To resolve this challenge, we propose to jointly consider the scale-invariance of kurtosis in bandpass domain and piecewise stationarity in spatial domain. Specifically, for a given noisy image, we divide it into  $S$  disjoint regions, each of whose clean versions is assumed to be associated with a constant, but unknown kurtosis throughout all scales. The detailed discussion on how to determine these  $S$  regions is deferred to the next subsection. This generalizes Zoran's assumption on scale-invariant kurtosis [26], in which the piecewise stationarity inherent to image signals has totally been neglected.

Let  $\kappa(y_j^i)$  and  $\sigma_{y_j^i}^2$  be the observed kurtosis and variance computed from the  $j$ th scale and the  $i$ th region of the noisy image, respectively. The kurtosis model described in (7) for the  $i$ th region can be expressed as

$$\sqrt{\kappa(y_j^i)} = \sqrt{\kappa(x^i)} - \frac{\sigma_n^2}{\sigma_{y_j^i}^2} \sqrt{\kappa(x^i)} \quad (8)$$

where the scale index  $j$  in  $\kappa(x_j^i)$  has been dropped because it is assumed to be constant for all  $j$ 's with a given  $i$ . The estimation of  $\sigma_n^2$  is then cast into the following problem to optimally fit the kurtosis model (8) across all regions and

throughout all scales

$$\begin{aligned} & \left\{ \hat{\sigma}_n^2, \{\hat{\kappa}(x^i)\}_{i=1}^S \right\} \\ &= \arg \min_{\sigma_n^2, \{\kappa(x^i)\}_{i=1}^S} \\ & \left\{ \sum_{i=1}^S \alpha_i \sum_{j=1}^M \left( \sqrt{\kappa(y_j^i)} - \sqrt{\kappa(x^i)} + \frac{\sigma_n^2}{\sigma_{y_j^i}^2} \sqrt{\kappa(x^i)} \right)^2 \right. \\ & \left. - \lambda \sum_{k=1}^S \sum_{l=1}^S \left( \sqrt{\kappa(x^k)} - \sqrt{\kappa(x^l)} \right)^2 \right\} \end{aligned}$$

$$\text{Subject to: } \kappa(x^i) \geq \frac{1}{M} \sum_{j=1}^M \kappa(y_j^i), \quad \text{for } i = 1, 2, \dots, S \quad (9)$$

where  $\kappa(x^i)$  denotes the constant yet unknown kurtosis associated with the  $i$ th region. The objective function consists of two terms, in which the first one is the kurtosis model fitting error term, and the second one represents the regularization term. The weighting factor  $\alpha_i$  in the first term of the objective function is specified as

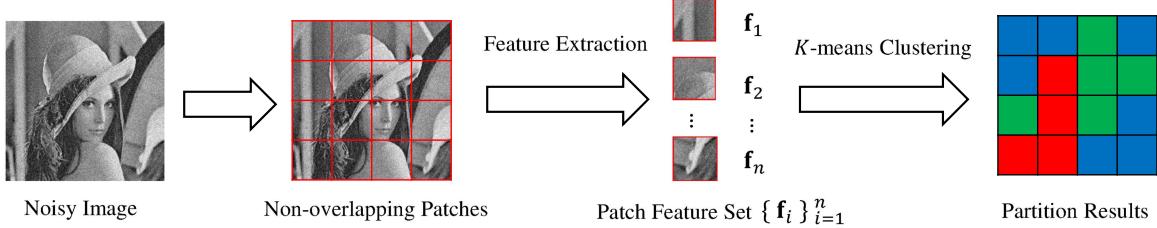
$$\alpha_i = \frac{\sum_j \kappa(y_j^i)}{\sum_{ij} \kappa(y_j^i)} \quad (10)$$

It can be seen that the regions with larger kurtosis values are assigned with larger weighting factors. This is because these regions generally contain more meaningful information, as also pointed out in [27]. The regularization term is adopted as different regions are expected to have rather different kurtosis values. Otherwise, these regions should be merged. The regularization parameter  $\lambda$  is used to control the relative importance of regularization term compared to the fidelity error term, and is empirically set as 0.01 in our experiments. The set of constraints adopted in the optimization problem (9) is to bound the feasible space of optimization variable  $\kappa(x^i)$ . Those inequalities are designed to reflect the phenomenon that noise addition would generally decrease kurtosis [27].

It is easy to see that the objective function in (9) is non-convex as we need to optimize  $\{\kappa(x^i)\}_{i=1}^S$  and  $\sigma_n^2$  simultaneously. However, it can be cast into two consecutive convex optimization sub-problems once fixing one variable and optimize the other one. This iterative alternating optimization procedure is given in Algorithm 1. Specifically, we first update kurtosis values by solving sub-problem (11), which is a standard constrained quadratic programming problem [35]. As shown in Appendix A, this problem is convex when the regularization parameter  $\lambda$  is appropriately selected, and thus it can be efficiently solved by utilizing existing toolboxes, e.g., cvx [36]. Then the estimated variance is updated in subproblem (12), which permits a closed-form solution given by

$$\hat{\sigma}_n^2 = \frac{\sum_{ij} \alpha_i \left( \sqrt{\hat{\kappa}(x^i)} - \sqrt{\kappa(y_j^i)} \right)}{\sum_{ij} \frac{\alpha_i}{\sigma_{y_j^i}^2} \sqrt{\hat{\kappa}(x^i)}} \quad (13)$$

The detailed derivation of (13) is provided in Appendix B. Note that in sub-problem (12), the second term in objective

Fig. 3. The workflow of the proposed  $K$ -means based image partition.**Algorithm 1** Estimating  $\sigma_n$  and  $\{\kappa(x^i)\}_{i=1}^S$ 

**Input:** Observed kurtosis  $\kappa(y_j^i)$  and variance  $\sigma_{y_j^i}^2$  for all  $i, j$ .

**Initialization:**  $\hat{\kappa}(x^i) \leftarrow 0$  and  $\hat{\sigma}_n^2 \leftarrow \frac{1}{SM} \sum_{ij} \sigma_{y_j^i}^2$

1: **repeat**

2: Fix  $\hat{\sigma}_n^2$ , update  $\{\hat{\kappa}(x^i)\}_{i=1}^S$  by solving

$$\arg \min_{\{\kappa(x^i)\}_{i=1}^S} \left\{ \sum_{i=1}^S \alpha_i \sum_{j=1}^M \left( \sqrt{\kappa(y_j^i)} - \sqrt{\kappa(x^i)} + \frac{\hat{\sigma}_n^2}{\sigma_{y_j^i}^2} \sqrt{\kappa(x^i)} \right)^2 - \lambda \sum_{k=1}^S \sum_{l=1}^S \left( \sqrt{\kappa(x^k)} - \sqrt{\kappa(x^l)} \right)^2 \right\}$$

$$\text{Subject to : } \kappa(x^i) \geq \frac{1}{M} \sum_{j=1}^M \kappa(y_j^i), \text{ for } i = 1, 2, \dots, S \quad (11)$$

3: Fix  $\{\hat{\kappa}(x^i)\}_{i=1}^S$ , update  $\hat{\sigma}_n^2$  by solving

$$\arg \min_{\sigma_n^2} \sum_{i=1}^S \alpha_i \sum_{j=1}^M \left( \sqrt{\kappa(y_j^i)} - \sqrt{\hat{\kappa}(x^i)} + \frac{\sigma_n^2}{\sigma_{y_j^i}^2} \sqrt{\hat{\kappa}(x^i)} \right)^2 \quad (12)$$

4: **until** Objective function converges

5: **return** Estimated  $\hat{\sigma}_n$  and  $\{\hat{\kappa}(x^i)\}_{i=1}^S$

function and constraints are dropped because optimization is only operated with respect to  $\sigma_n^2$ . The alternating update procedure is repeated until the convergence of objective function. In our experiments, we observe that the objective function converges after around 10 iterations.

**B. K-Means Based Image Partition**

In this subsection, we discuss how to determine the  $S$  disjoint regions, each of whose clean versions is assumed to be associated with a constant, but unknown kurtosis value in band-pass domain. This is an important pre-step of our noise level estimation procedure presented in Algorithm 1. For image partition, one straightforward way is to use existing over-segmentation/superpixel methods such as [37] and [38]. However, the purpose of performing the image partition in our problem is to cluster image segments exhibiting similar kurtosis values into a single region, which is different from the targets of the general over-segmentation/superpixel methods. Furthermore, the kurtosis is computed from marginal filter response, and the filtering is typically applied to regular square patches. Therefore, it is a natural choice to extract the kurtosis features from fixed square patches, rather than from

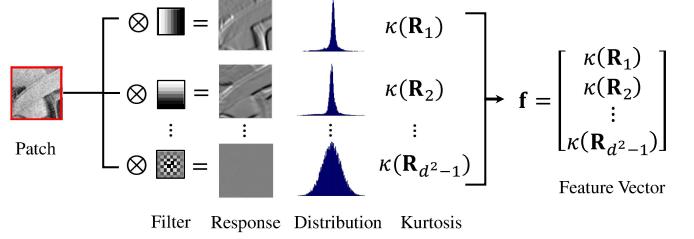


Fig. 4. The workflow of the proposed feature extraction.

the segments with irregular shape yielded by conventional over-segmentation/superpixel methods.

The schematic diagram of the proposed  $K$ -means based image partition algorithm is depicted in Fig. 3. Specifically, we first uniformly partition the available noisy image into  $p \times p$  non-overlapping square patches, where  $p = 16$  in our experiments. Each patch is then convolved with a 2D band-pass filter  $\mathbf{B}_k$  selected from  $d \times d$  linear transform base. Here  $k$  is the index for band-pass filters, and  $0 \leq k \leq d^2 - 1$ . The response image patch can then be given by

$$\mathbf{R}_k = \mathbf{P} \otimes \mathbf{B}_k, \quad (14)$$

where  $\mathbf{R}_k$  and  $\mathbf{P}$  denote the response image patch and the noisy image patch, respectively. The kurtosis value associated with  $\mathbf{R}_k$  is denoted by  $\kappa(\mathbf{R}_k)$ , which can be readily computed using (1). Upon applying all filters and computing the kurtosis values, we form a feature vector for image patch  $\mathbf{P}$ :

$$\mathbf{f} = [\kappa(\mathbf{R}_1), \kappa(\mathbf{R}_2), \dots, \kappa(\mathbf{R}_{d^2-1})]' \quad (15)$$

Note that  $\mathbf{R}_0$  is not included into the above feature vector. This is because the distribution of  $\mathbf{R}_0$  is highly skewed, and kurtosis cannot capture this skewness. The workflow of the proposed feature extraction can also be found in Fig. 4.

We then use the feature vector  $\mathbf{f}$  to trigger a clustering algorithm for determining these  $S$  regions. In this work, we adopt the well-known  $K$ -means algorithm, where the number of cluster  $S$  is given as a prior. In Fig. 5, we show the partition results of four representative images, where  $8 \times 8$  PCA bases are used. One can observe that the images are partitioned into coherent regions accompanied with minor scattered areas that exhibit image structures. In this example and all the forthcoming experiments, we set  $S = 3$ . The thorough justification for this configuration will be provided in the Section V.

**C. Estimation Correction via Noise Injection**

Natural images typically contain prominent structures and complex textures, making it difficult to distinguish image

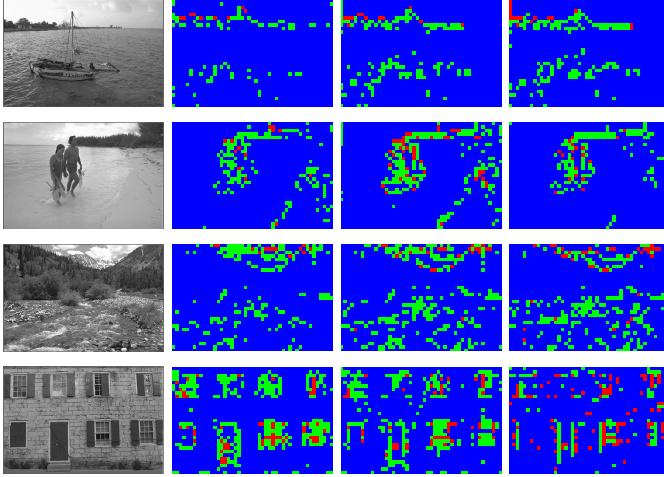


Fig. 5. Examples of image partition using the proposed  $K$ -means based algorithm. From left to right: original images, partition results when  $\sigma = 0$  (clean image),  $\sigma = 5$  and  $\sigma = 15$  for Gaussian noise. Each color represents a cluster.

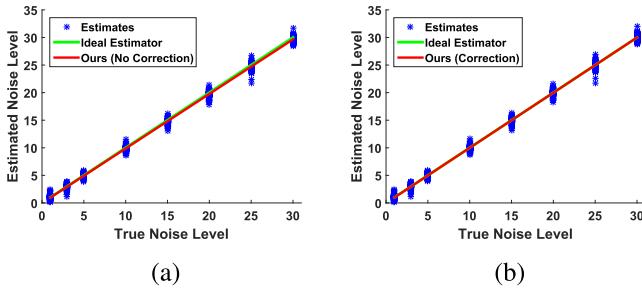


Fig. 6. Noise level estimation bias rectification on 200 synthetic noisy images. (a) our estimator without rectification (red line) slightly deviates from that of ideal estimator (green line),  $MSE = 0.205$ ; (b) our estimator with rectification (red line) almost coincides with that of ideal estimator (green line),  $MSE = 0.168$ .

signal from noise effectively. As a result, image content might magnify or attenuate the noise effects to some extent. It is therefore reasonable to assume that there exists estimation bias caused by image content.

To validate this assumption, we test our algorithm on 200 synthetic noisy images, where the noise levels range from 1 to 30. The estimation results are shown in Fig. 6(a). One can observe that our estimates are around the true noise levels, and the relationship between our estimate and the true noise level is approximately linear. To further reduce estimation bias, we suggest a rectification procedure based on noise injection.

Specifically, we formulate the relationship between estimated noise level  $\hat{\sigma}$  and the underlying true noise level  $\sigma$  using the following linear model

$$\hat{\sigma} = \rho\sigma \quad (16)$$

Clearly, there are two unknown variables  $\rho$  and  $\sigma$  in (16). Uncovering these two unknowns from a single equation is an ill-posed problem. To tackle this challenge, we employ noise injection strategy to generate an additional equation. We inject the same type of i.i.d. noise with variance  $\sigma_t^2$  into the noisy image, and make an additional round of noise estimation. Due to the independence between the original noise and the injected

noise, we have

$$\begin{cases} \hat{\sigma}_1^2 = \rho^2\sigma^2 \\ \hat{\sigma}_2^2 = \rho^2(\sigma^2 + \sigma_t^2) \end{cases} \quad (17)$$

where  $\hat{\sigma}_1^2$  and  $\hat{\sigma}_2^2$  denote the estimated noise variances before and after injection, respectively. Here the value of  $\rho$  is assumed to be invariant in those two rounds of noise level estimation. This assumption is reasonable because  $\rho$  only depends on the original image content, which is fixed. From (17), it is easy to solve  $\sigma^2$ , which reads

$$\sigma^2 = \frac{\hat{\sigma}_1^2\sigma_t^2}{\hat{\sigma}_2^2 - \hat{\sigma}_1^2} \quad (18)$$

Conceptually, the variance of injected noise can be set arbitrarily. In practice, however, too heavy injection tends to wipe out the effects resulting from image content. One reasonable choice is to set the injected noise variance as the first estimated one, i.e.,  $\sigma_t^2 = \hat{\sigma}_1^2$ . Note the noise level becomes larger after the noise injection, which implies  $\hat{\sigma}_2^2 > \hat{\sigma}_1^2$ . Then the underlying true noise variance can be estimated by  $\tilde{\sigma}^2 = \hat{\sigma}_1^4 / (\hat{\sigma}_2^2 - \hat{\sigma}_1^2)$ .

However, due to the simplicity of this linear model, the rectified noise level may still slightly deviate from the true noise level. To make a more robust estimation, we fuse the estimated results before and after the noise injection through a linear convex combination

$$\hat{\sigma}_F^2 = \beta_0\tilde{\sigma}^2 + \beta_1\hat{\sigma}_1^2 \quad (19)$$

where  $\hat{\sigma}_F^2$  denotes the final estimation, and  $\beta_0$  and  $\beta_1$  are weighting factors satisfying  $\beta_0 + \beta_1 = 1$ . In our proposed scheme,  $\beta_0 = 0.606$  and  $\beta_1 = 0.394$  which are obtained from an off-line training procedure with over 200 natural images. The estimation results with rectification are presented in Fig. 6(b). As can be seen, the estimation bias is further reduced.

#### IV. HANDLING NOISE LEVEL ESTIMATION FOR NON-GAUSSIAN NOISE

In this section, we show that our proposed noise level estimation approach can be readily extended to handle the non-Gaussian cases. In the noise model described in (2), the noise term  $\mathbf{n}$  is assumed to be zero-mean, signal-independent AWGN. In fact, the restrictions on the noise Gaussianity in the spatial domain can be relaxed. As long as the noise in the linear transform domain is i.i.d. Gaussian, the proposed framework for noise level estimation still works. This implies that our method can accommodate a variety of noise types in the spatial domain. This is because i.i.d. noise, though non-Gaussian in the pixel domain, will be mixed into Gaussian noise in the linear transform domain, due to the noise independence and central limit theorem [32].

Specifically, we first estimate the noise variance in the linear transform domain, using our method presented in Section III. Noticing that the employed linear transform in (3) is unitary, the noise variance is invariant across the spatial domain and the transform domain [27]. The estimated noise level can then be used to determine the parameters in the model characterizing

the noise. In the following, we demonstrate the effectiveness of our proposed method by considering three non-Gaussian noise cases: uniform noise, Laplacian noise and multiplicative Gamma noise.

### A. Uniform Noise

The uniform noise model can characterize the quantization noise which, e.g., incurs during quantizing a sensed image [1]. Mathematically, for the additive noise  $Z$  following uniform distribution:  $Z \sim \mathcal{U}(a, b)$ , the mean and variance are given by

$$\mu_Z = \frac{1}{2}(b - a), \quad \sigma_Z^2 = \frac{1}{12}(b - a)^2 \quad (20)$$

We particularly focus on the widely used zero-mean uniform noise model, which can be obtained by setting  $b = -a > 0$ . In this case, the probability density function of original uniform noise is only parameterized by the parameter  $b$ . The estimate of  $b$  can then be derived in a straightforward way

$$\hat{b} = \sqrt{3} \hat{\sigma}_Z = \sqrt{3} \hat{\sigma}_n \quad (21)$$

### B. Laplacian Noise

The Laplacian noise is a kind of heavy-tailed noise. It is often used to model some impulsive noise, which is caused by analog-to-digital converter errors, bit errors in transmission [1], etc. For the additive noise following Laplacian distribution  $Z \sim \mathcal{L}(0, v)$ , its probability density function is defined as

$$f_Z(z) = \frac{1}{2v} \exp\left\{-\frac{|z|}{v}\right\} \quad (22)$$

where  $u$  and  $v$  are the location and scale parameters. Its variance is given by  $\sigma_Z^2 = 2v^2$ . The estimate of parameter  $v$  can then be obtained by

$$\hat{v} = \frac{\hat{\sigma}_Z}{\sqrt{2}} = \frac{\hat{\sigma}_n}{\sqrt{2}} \quad (23)$$

### C. Multiplicative Gamma Noise

The multiplicative noise is a widely adopted noise model in addition to additive noise. It could well model the noises incurred in some remote sensing systems such as ultrasound, synthetic aperture radar (SAR) and laser imaging [39]. More precisely, the independent multiplicative noise corruption model is given by  $\mathbf{y} = \mathbf{x} \circ \mathbf{z}$ , where  $\mathbf{x}$  and  $\mathbf{y}$  are the original and noisy image, respectively, and  $\mathbf{z}$  denotes the positive i.i.d. noise. Here,  $\circ$  represents component-wise multiplication operation. One widely adopted multiplicative noise model is Gamma model [1]

$$f_Z(z) = \frac{\beta^\alpha}{\Gamma(\alpha)} z^{\alpha-1} e^{-\beta z}, \quad \text{for } z \in (0, +\infty) \quad (24)$$

where  $\alpha, \beta > 0$  are the shape and rate parameters, respectively, and  $\Gamma(\cdot)$  is the standard Gamma function. It was shown that the variance of this multiplicative Gamma noise  $\sigma_z^2 = \alpha/\beta^2$  [40]. In practice, it is often assumed that the expectation of natural logarithm on noise is zero, i.e.,  $\mathbb{E}[\ln(z)] = 0$  [10]. In general, Gamma noise is signal-dependent because large signals will be corrupted by noise with large variances. To eliminate such

dependence, one common strategy is to convert multiplicative noise into additive model by taking logarithm of (24) [41], yielding

$$\underbrace{\ln(\mathbf{y})}_{\tilde{\mathbf{y}}} = \underbrace{\ln(\mathbf{x})}_{\tilde{\mathbf{x}}} + \underbrace{\ln(\mathbf{z})}_{\tilde{\mathbf{z}}} \quad (25)$$

The resultant noise  $\tilde{\mathbf{z}}$  turns to be signal-independent and is still i.i.d.. As illustrated in [10] and [34], log transformed natural images preserve the scale-invariant kurtosis property. Owing to these desirable properties in the transform domain, we can apply the proposed method to estimate noise level of  $\tilde{\mathbf{z}}$ , denoted by  $\hat{\sigma}_{\tilde{z}}$ . To determine the parameters  $\alpha$  and  $\beta$ , we express the mean and variance of noise  $\tilde{\mathbf{z}}$  using  $\alpha$  and  $\beta$  [42]:

$$\mathbb{E}(\tilde{z}) = \psi(\alpha) - \ln(\beta) = 0 \quad (26)$$

$$\sigma_{\tilde{z}}^2 = \psi_1(\alpha) \quad (27)$$

Here  $\psi(x) = \frac{d}{dx} \Gamma(x)$  and  $\psi_1(x) = \frac{d^2}{dx^2} \Gamma(x)$  are the first and second derivative of gamma function, which are called digamma and trigamma functions, respectively. Then shape parameter  $\alpha$  is estimated using (27), which gives  $\hat{\alpha} = \psi_1^{-1}(\hat{\sigma}_{\tilde{z}}^2) = \psi_1^{-1}(\hat{\sigma}_n^2)$ . The corresponding rate parameter  $\beta$  can be estimated using the fact  $\beta = e^{\psi(\alpha)}$ , which holds from (26).

## V. EXPERIMENTAL RESULTS

### A. Experiment Setup

The proposed method is evaluated and compared with the state-of-the-art noise level estimation techniques on images from the following test sets:

- All the 24 images of size  $768 \times 512$  from Kodak image database [43].
- 100 images of size  $512 \times 384$  randomly selected from UCID-v2 corpus [44].

In our experiments, we choose PCA linear transform. It should be pointed out that using DCT as linear transform leads to similar estimation performance, and hence, the results are not given here. To compare the estimation performance over different image sets, we report the mean and standard deviation (STD) of estimation for noise levels ranging from 1 to 30 for each method. Additionally, the overall estimation performance is evaluated using the following three metrics: mean squared error (MSE), mean absolute deviation (MAD) and averaged relative estimation error rate  $\bar{e}_r = \frac{|\hat{\sigma}_n - \sigma_n|}{\sigma_n} \times 100\%$ .

### B. Noise Level Estimation on Synthetic Noisy Images

1) *AWGN Noise Level Estimation*: We compare our noise level estimation method with six works: filter-based method [12], statistics-based method [21], kurtosis-based methods [10], [26], patch-based methods [4], [18].

The comparison results are summarized in Table I. Compared with the filter-based method [12] and the statistics-based method [21], our approach constantly achieves better performance for all noise levels on all test image sets. This superiority is more evident at low noise levels, e.g.,  $\sigma_n \leq 5$ , where [12] and [21] generate severe overestimates. In terms of the overall evaluation metrics, the improvement over those two

TABLE I  
COMPARISON OF ESTIMATION RESULTS FOR AWGN. THE BEST RESULTS ARE HIGHLIGHTED WITH BOLDFACE

Noise Level	[12]		[21]		[26]		[10]		[18]		[4]		Proposed		
	Mean	STD	Mean	STD	Mean	STD	Mean	STD	Mean	STD	Mean	STD	Mean	STD	
Kodak	1	3.59	1.52	2.86	1.26	0.86	0.87	0.75	0.74	1.37	0.27	1.25	0.17	1.34	0.28
	3	4.87	1.30	4.45	1.06	2.23	1.35	2.30	1.27	3.16	0.12	3.07	0.12	3.13	0.16
	5	6.49	1.11	6.25	0.94	4.53	1.04	4.58	0.82	5.11	0.07	4.97	0.12	5.06	0.18
	10	11.04	0.80	10.95	0.77	9.52	0.84	9.74	0.46	10.05	0.07	9.76	0.10	10.03	0.16
	15	15.83	0.63	15.77	0.64	14.37	0.87	14.76	0.40	15.01	0.15	14.60	0.12	15.04	0.19
	20	20.73	0.50	20.65	0.57	19.24	0.87	19.76	0.35	19.91	0.22	19.47	0.13	20.08	0.18
	25	25.68	0.46	25.56	0.49	24.12	0.85	24.78	0.34	24.81	0.31	24.35	0.18	25.08	0.20
	30	30.64	0.39	30.45	0.44	29.01	1.05	29.78	0.34	29.69	0.42	29.22	0.20	30.05	0.21
	MSE	1.953		1.194		0.475		0.126		0.040		0.199		<b>0.019</b>	
UCID	MAD	1.234		0.992		0.640		0.319		0.162		0.368		<b>0.101</b>	
	$\bar{e}_r$	46.99%		35.12%		8.57%		8.00%		5.91%		5.12%		<b>5.11%</b>	
	1	3.75	2.02	2.62	1.38	0.85	1.15	0.68	0.79	1.23	0.24	1.18	0.16	1.19	0.26
	3	5.11	1.75	4.37	1.20	2.29	1.18	2.35	1.11	3.13	0.12	3.07	0.16	3.04	0.21
	5	6.76	1.51	6.26	1.12	4.34	1.00	4.54	0.82	5.12	0.09	5.03	0.19	5.04	0.15
	10	11.30	1.12	11.04	0.96	9.22	0.91	9.62	0.56	10.11	0.14	10.01	0.13	10.03	0.20
	15	16.07	0.87	15.88	0.83	14.08	0.94	14.62	0.50	15.07	0.24	14.97	0.16	14.99	0.29
	20	20.97	0.71	20.75	0.72	18.98	0.98	19.62	0.47	19.99	0.32	19.95	0.20	19.97	0.33
	25	25.93	0.61	25.65	0.65	23.85	1.01	24.63	0.45	24.93	0.42	24.91	0.23	25.01	0.33
	30	30.88	0.53	30.59	0.59	29.02	1.09	29.63	0.42	29.74	0.56	29.86	0.24	30.05	0.42
	MSE	2.567		1.162		0.715		0.180		0.022		0.009		<b>0.006</b>	
	MAD	1.471		1.021		0.795		0.413		0.126		0.073		<b>0.050</b>	
	$\bar{e}_r$	51.54%		32.24%		9.79%		9.21%		4.12%		2.79%		<b>2.75%</b>	

TABLE II  
COMPARISON OF ESTIMATION RESULTS FOR THREE NON-GAUSSIAN NOISE TYPES. THE BEST RESULTS ARE HIGHLIGHTED WITH BOLDFACE

Noise Type	Image Set	Metrics	[12]	[21]	[26]	[10]	[18]	[4]	Proposed
Uniform	Kodak	MSE	8.989	6.075	0.790	0.385	0.176	0.317	<b>0.091</b>
		MAD	2.795	2.414	0.855	0.523	0.353	0.465	<b>0.169</b>
		$\bar{e}_r$	94.16%	69.64%	13.88%	12.46%	14.91%	20.23%	<b>11.47%</b>
	UCID	MSE	11.225	5.709	1.486	0.498	0.122	0.109	<b>0.028</b>
		MAD	3.189	2.372	1.137	0.638	0.328	0.317	<b>0.110</b>
		$\bar{e}_r$	101.1%	62.70%	12.76%	11.64%	11.10%	9.80%	<b>6.35%</b>
Laplacian	Kodak	MSE	1.613	1.291	0.614	0.378	10.710	1.717	<b>0.026</b>
		MAD	0.934	1.003	0.706	0.569	2.456	1.081	<b>0.102</b>
		$\bar{e}_r$	45.02%	32.84%	8.12%	9.22%	16.72%	11.58%	<b>6.02%</b>
	UCID	MSE	2.048	1.109	1.017	0.611	10.161	0.768	<b>0.046</b>
		MAD	1.119	0.933	0.934	0.724	2.413	0.704	<b>0.175</b>
		$\bar{e}_r$	49.35%	29.38%	11.50%	10.81%	15.45%	5.42%	<b>3.25%</b>
Gamma	Kodak	MSE	1.851	0.815	10.283	3.996	0.115	0.069	<b>0.005</b>
		MAD	0.952	0.608	2.364	1.412	0.336	0.226	<b>0.052</b>
		$\bar{e}_r$	4.84%	3.97%	12.38%	6.81%	7.15%	2.17%	<b>1.17%</b>
	UCID	MSE	6.897	2.280	3.006	0.282	0.079	0.239	<b>0.049</b>
		MAD	1.893	1.031	1.522	0.485	0.256	0.403	<b>0.138</b>
		$\bar{e}_r$	9.83%	5.73%	15.61%	8.83%	6.25%	3.33%	<b>0.99%</b>

competitors is still quite significant. For instance, the average relative estimation error rate for UCID-v2 corpus is only 1/18 of that given by [12].

When comparing with the kurtosis-based methods [10], [26], our proposed method produces more accurate and stable estimates for noise level  $\sigma_n > 1$ . This can be observed from the smaller deviations of the mean values away from the true ones, and the lower STD. In the case that  $\sigma_n = 1$ , [26] leads to the best accuracy measured by the mean values of the estimates; but accompanied with rather high STD. By further inspecting the results for individual images, we find that most of the estimates obtained by [26]

are very close to zeros, and the remaining estimates tend to be positively biased from 1 with large magnitudes.

Compared with the patch-based methods [4], [18], our method is slightly inferior when the noise level is low; but gains superiority when the noise level is medium and high. Under the overall metrics MSE, MAD and  $\bar{e}_r$ , the proposed scheme outperforms these two competitors.

2) *Non-Gaussian Noise Level Estimation:* We evaluate the estimation performance for three non-Gaussian cases: additive zero-mean uniform noise, additive Laplacian noise, multiplicative Gamma noise. The experimental results are compiled in Table II. Compared with the filter-based method [12] and

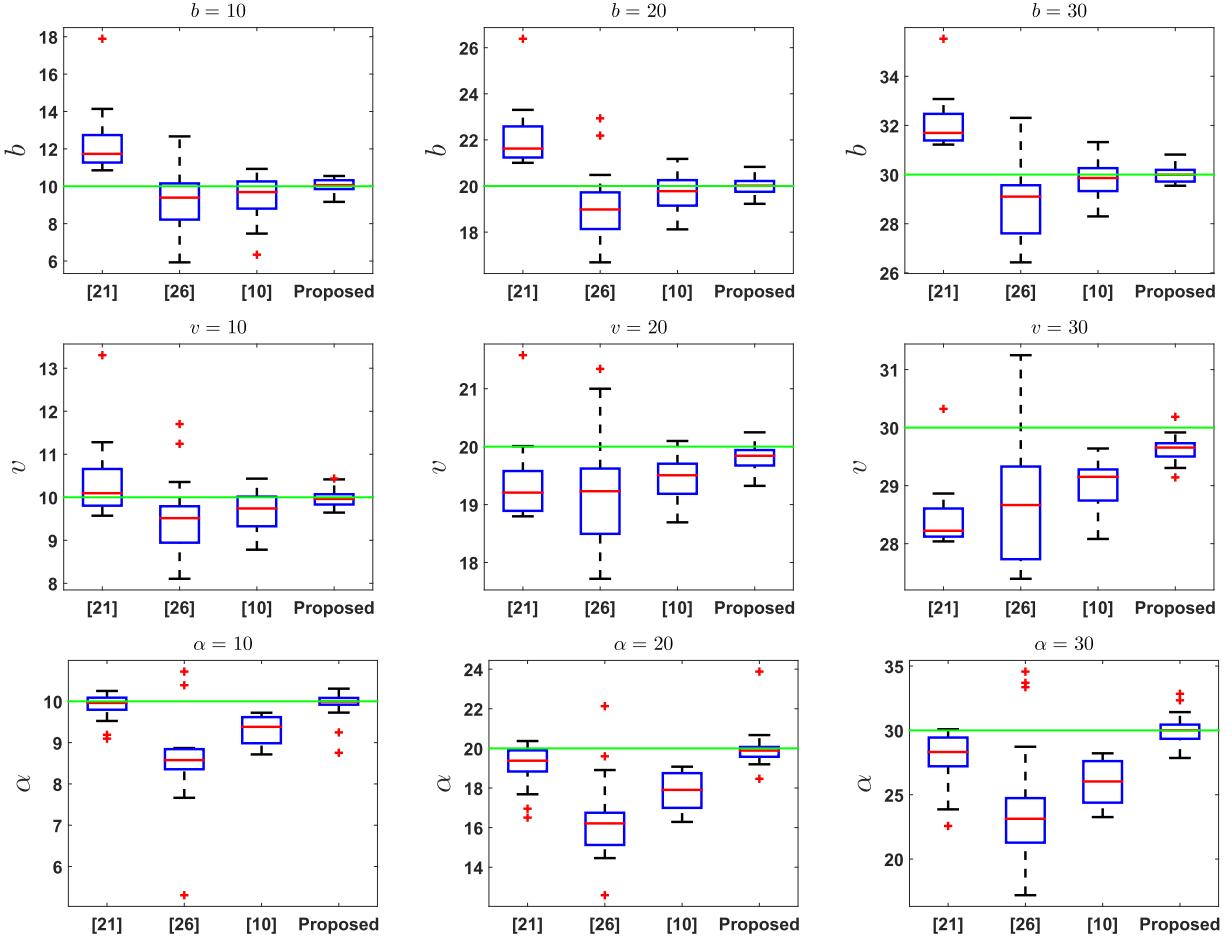


Fig. 7. Box-plot comparison with [10], [21], and [26] for non-Gaussian noise estimation on Kodak image set. Top to bottom: uniform, Laplacian and multiplicative Gamma noise. Left to right: estimation results for three non-Gaussian noise models, in which the ground-truth model parameters are 10, 20, 30, respectively. The green line represents the true parameter; the red line denotes the median of estimation, and the red dots denote outliers.

patch-based methods [4], [18], our approach substantially outperforms those three works on all test image sets. Take uniform noise on Kodak test set as an example. The MSE of our method is 0.091, while [4], [12], and [18] produce MSEs 0.317, 8.989 and 0.176, respectively. Similarly, the experimental results under the other metrics also demonstrate the superiority of our work. Furthermore, compared with the kurtosis-based methods [10], [26], the performance improvement of our approach is pronounced, especially for multiplicative Gamma noise. For instance, our method could attain the relative estimation error rate  $\bar{e}_r$  less than 1% on the UCID-v2 corpus, while the results given by [10] and [26] are all larger than 8%.

To further illustrate the detailed estimation results on image sets, we present the statistical Box-plot comparison with [10], [21], and [26] on Kodak image set in Fig.7. As can be seen, all the other three competitors tend to underestimate the model parameters. In contrast, our method can achieve more accurate estimation robustly, though a few outliers exist.

### C. Noise Level Estimation for Image AWGN Denoising

Over the past decades, huge improvements have been made in the field of image denoising. One well-known denoising

scheme is the adaptive Wiener filter [45], which can be expressed as

$$\hat{x} = \mu + \frac{\sigma_y^2 - \sigma_n^2}{\sigma_y^2}(y - \mu) \quad (28)$$

where  $\hat{x}$  and  $y$  are the denoised and the observed noisy pixels, respectively;  $\mu$  and  $\sigma_y^2$  are the mean and the variance in the local neighborhood of  $y$ , and  $\sigma_n$  is the noise level parameter. In some implementations, such as MATLAB `winer2` default setting,  $\sigma_n^2$  is set as the average of all the local estimated variances when the ground-truth noise level is not provided. Here, the local estimated variance is computed as the variance of all noisy pixels in a local neighborhood. However, this noise parameter setting would severely degrade the final denoising performance. To see this, we conduct adaptive Wiener denoising on Kodak image set with three fed noise level parameters: the true noise level, the `winer2` default setting and our estimated noise level. The test noise levels range from 3 to 15, and the denoising performance comparison is shown in Fig.8. One shall see that the denoising performance with our estimated noise level is almost the same with that of true noise level, while the `winer2` default setting generally leads to much inferior denoising results.

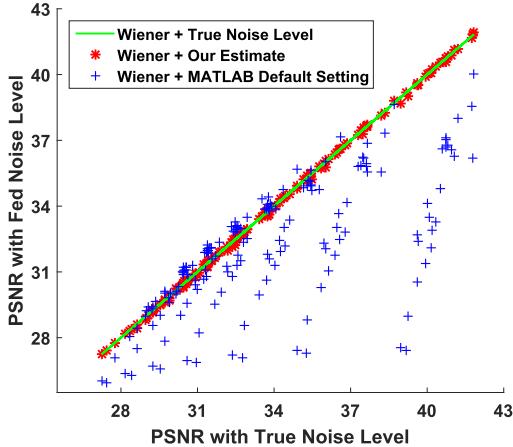


Fig. 8. Denoising performance (in PSNR) comparison of adaptive Wiener filtering with three noise level parameter settings: the true noise level, the MATLAB wiener2 default setting and our estimated noise level.

This experiment shows that our noise level estimation procedure can be incorporated into the existing non-blind image denoising schemes and consequently benefits the denoising.

We also test our noise level estimator with a more sophisticated denoiser BM3D [3]. In our experiments, the noise level is first estimated, and then fed into the BM3D denoiser. Two visual comparison examples are shown in Fig.9 and Fig.10. One can observe that our estimator always generates more accurate noise level estimates than the competing estimators [10], [12], [21], [26]. Consequently, our estimator generally leads to higher PSNR. It should also be mentioned that the highest PSNR is not always achieved by feeding the true noise level to the denoiser. This phenomenon has been investigated in [4], which suggested that the estimated noise level should be further tuned for denoising algorithms. The in-depth discussion of this issue, however, is beyond the scope of this work.

Moreover, we apply the proposed method in real-world image denoising application with BM3D denoiser. Due to the limitation of imaging hardware devices, some images captured by early cameras were inevitably corrupted by small amount of noise. It is therefore useful to restore the underlying “clean” images given those original ones. Here, as examples, we test on three images: Lena, Boat and Airplane. The noise levels are first estimated from the given original images by using our estimator, yielding 2.02, 1.71 and 0.92, respectively. Those estimates are then fed into BM3D denoiser to obtain the restored images. For better visual comparison, part of original and restored image is enlarged. One could see that the noises in plain regions (e.g., cheek of Lena) are greatly suppressed, while the image structures (e.g., eyelash of Lena) are well preserved. We further present the kurtosis distributions of original images accompanied with those of restored ones. As can be observed, the kurtosis distribution of the original images has large fluctuation across the scales. Instead, one can observe that the kurtosis distribution of restored images is, in general, more concentrated than that of their original counterparts, especially for the large scales (e.g., component indexes larger than 40). This observation again validates the plausibility of our proposed generalized kurtosis assumption.

TABLE III  
EXPERIMENTS ON THE CHOICE OF CLUSTER NUMBER  $S$

Metric	Cluster Number $S$					
	1	2	3	4	5	6
MSE	0.0940	0.0202	<b>0.0181</b>	0.0188	0.0215	0.0220
MAD	0.2085	0.1182	<b>0.0862</b>	0.0960	0.1023	0.1003
$\bar{e}_r$	7.50%	3.60%	<b>3.57%</b>	3.72%	4.02%	4.24%

#### D. Noise Level Estimation for Multiplicative Speckle Noise Reduction

To further demonstrate the practical usefulness of our proposed method, we apply it to the denoising of real noisy synthetic aperture radar (SAR) images. The test SAR image Spain, Barcelona-ORI, SE was captured in 2007 under StripMap mode with resolution  $4201 \times 5581$ , comprising of both rural and urban targets, e.g., mountains, roads, buildings and airport. It is available from TerraSAR-X sample imagery, which is provided by Astrium Geo-Information Services website [46]. The noise presented in SAR images can be well modeled as multiplicative Gamma speckle noise, as suggested in many existing literature [41], [47], [48]. We first estimate the noise level by using our method presented in Section IV-C, and then employ the well-known Kuan filter to perform the denoising of speckle noise [49], [50]. Essentially, the Kuan filter is operated in a pixel-by-pixel fashion, and can be mathematically expressed as

$$\hat{x} = \mu + \frac{\sigma_x^2}{\sigma_x^2 + \sigma_z^2(\mu^2 + \sigma_x^2)} \cdot (y - \mu) \quad (29)$$

where  $\hat{x}$  and  $y$  are the restored and observed noisy pixels, respectively;  $\mu$  denotes the mean value of the pixels in a local window centered by  $y$ ;  $\sigma_z^2$  is the noise variance estimated by our method from the noisy image; and  $\sigma_x^2$  is the variance of the unknown clean pixels around  $x$ , which needs to be estimated by

$$\sigma_x^2 = \frac{\sigma_y^2 - \mu^2 \sigma_z^2}{1 + \sigma_z^2} \quad (30)$$

Here,  $\sigma_y^2$  denotes the variance of the noisy pixels in a local neighborhood of  $y$ . For the test image Spain, Barcelona-ORI, SE, the estimated  $\sigma_z^2 = 0.0396$ , which can then be plugged into (29) and (30) for denoising purpose. The experimental results are shown in Fig.12, where two selected regions of the test image before and after the denoising are illustrated. As can be observed, the speckle noise is greatly suppressed, and the fine details are well preserved.

#### E. Performance Analysis

1) *Image Cluster Number  $S$* : Essentially, the choice of cluster number  $S$  in our method shall be dependent with the image scene and the noise level. One could manually set  $S$  for each image; but such tedious manual setting barriers the usage in practice. To facilitate the practical deployment, we set  $S = 3$  to be a fixed parameter. Below we provide the experimental justification for this setting. In Table.III, we show the relationship between the MSE, MAD and  $\bar{e}_r$

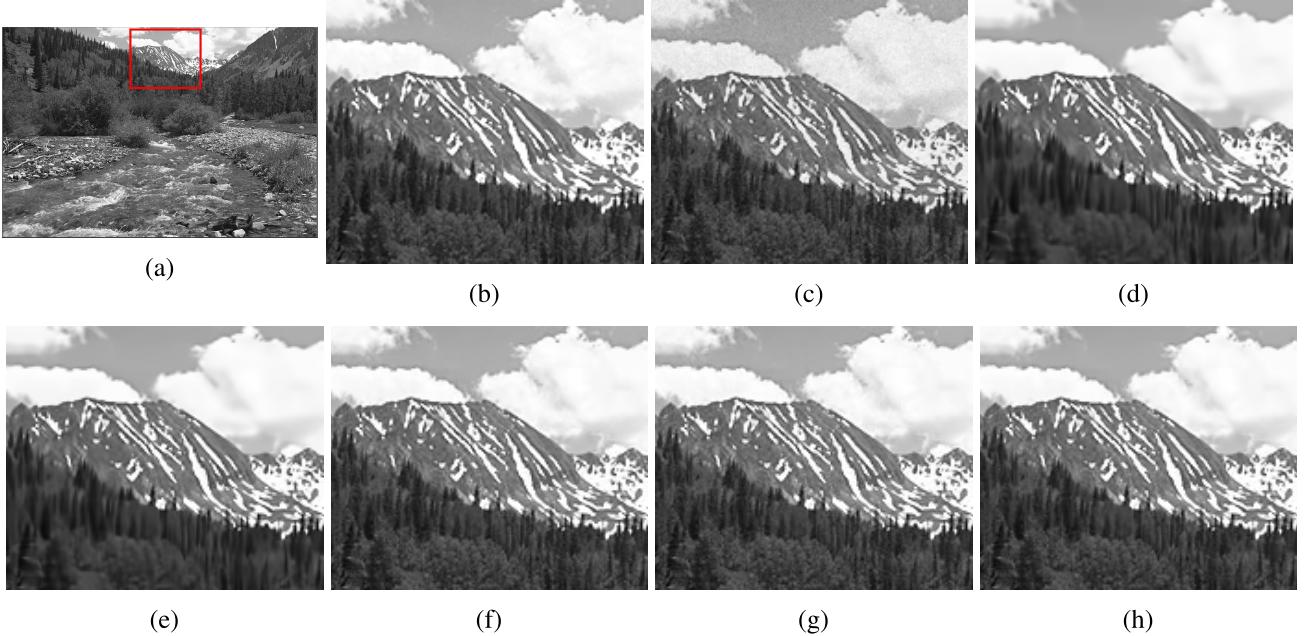


Fig. 9. Comparison of denoising results using BM3D with different noise level parameters. (a) Original image *kodim33* from [43]. (b) Part of image (in red rectangle) enlarged; (c) Noisy version of (b) with AWGN,  $\sigma_n = 5$ , PSNR = 34.14dB. (d) [12] + BM3D,  $\hat{\sigma}_n = 10.59$ , PSNR = 31.23dB. (e) [21] + BM3D,  $\hat{\sigma}_n = 9.82$ , PSNR = 30.18dB. (f) [26] + BM3D,  $\hat{\sigma}_n = 4.47$ , PSNR = 35.19dB. (g) [10] + BM3D,  $\hat{\sigma}_n = 4.32$ , PSNR = 35.17dB. (h) Ours + BM3D,  $\hat{\sigma}_n = \mathbf{5.15}$ , PSNR = 35.25dB.

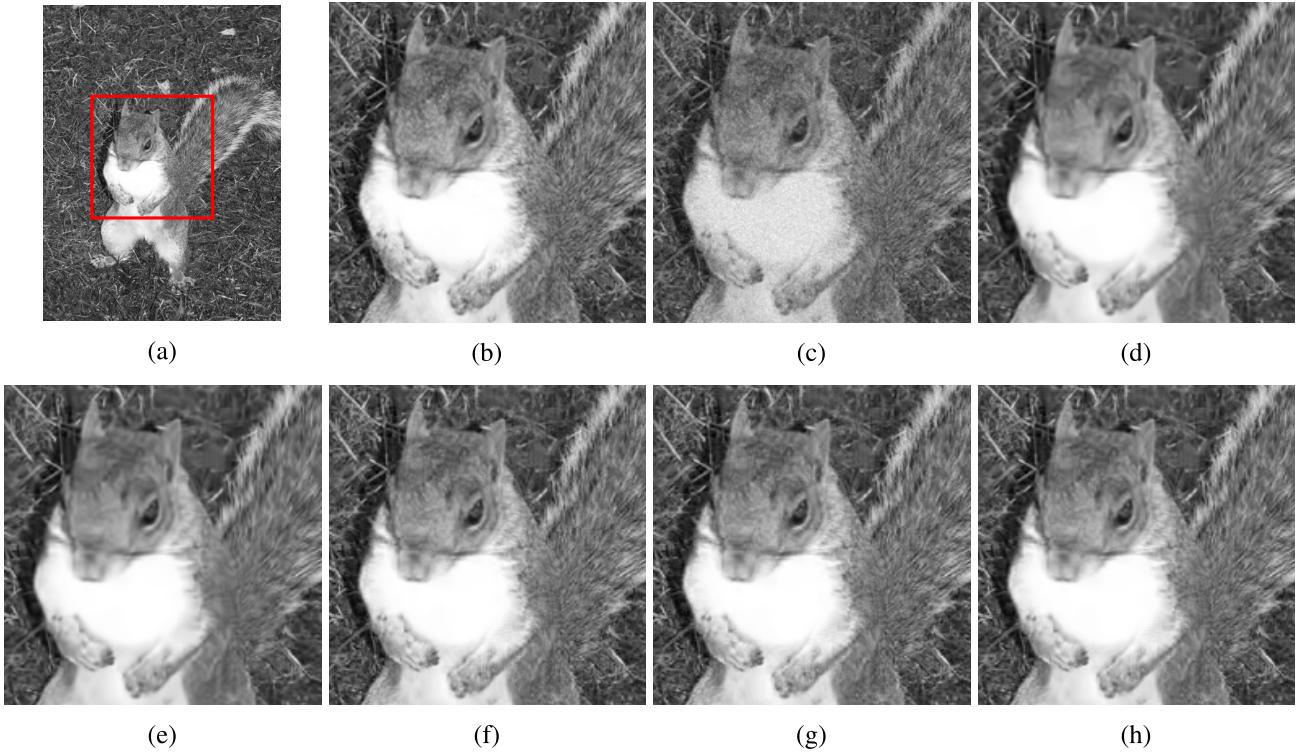


Fig. 10. Comparison of denoising results using BM3D with different noise level parameters. (a) Original image *ucid00510* from [44]. (b) Part of image (in red rectangle) enlarged; (c) Noisy version of (b) with AWGN,  $\sigma_n = 10$ , PSNR = 28.15dB. (d) [12] + BM3D,  $\hat{\sigma}_n = 13.57$ , PSNR = 28.85dB. (e) [21] + BM3D,  $\hat{\sigma}_n = 13.73$ , PSNR = 28.76dB. (f) [26] + BM3D,  $\hat{\sigma}_n = 9.17$ , PSNR = 30.14dB. (g) [10] + BM3D,  $\hat{\sigma}_n = 9.05$ , PSNR = 30.13dB. (h) Ours + BM3D,  $\hat{\sigma}_n = \mathbf{9.90}$ , PSNR = 30.18dB.

with respect to  $S$  ranging from 1 to 6. Here the results are obtained by averaging 100 natural images. As can be observed, MSE, MAD and  $\bar{e}_r$  values all drop remarkably when

$S = 2$ , compared with the case of  $S = 1$ . This again explains the importance of taking the image piecewise stationarity into account. The minima of MSE, MAD and  $\bar{e}_r$  values are

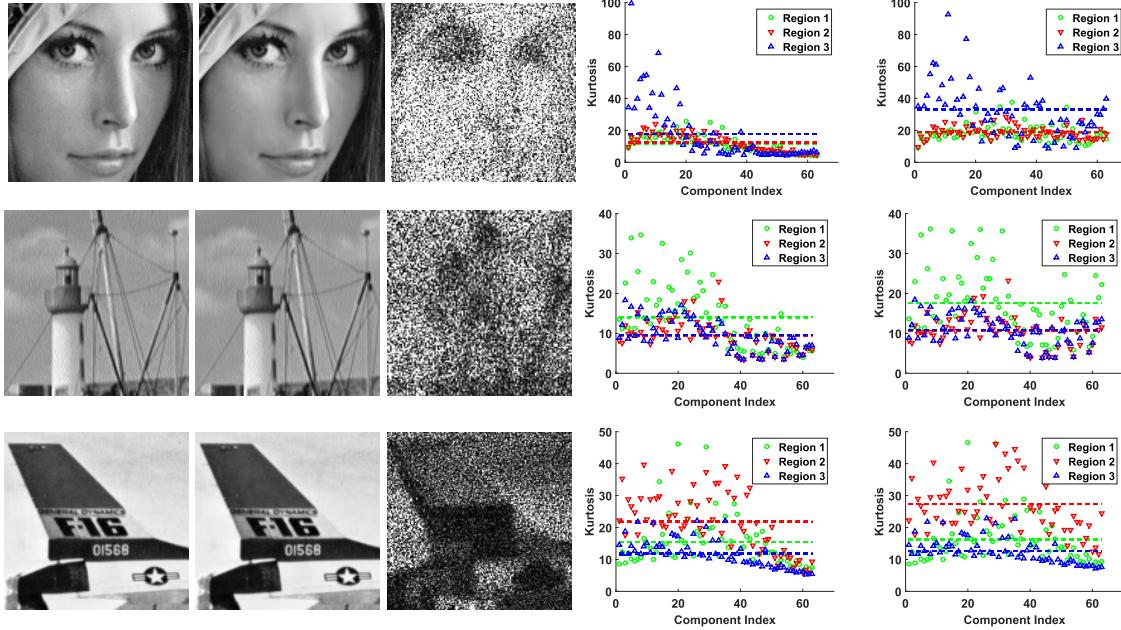


Fig. 11. Uncovering the “clean” images. Part of the image and its denoised result are enlarged for better visual comparison. From left to right: enlarged original image; denoised image; scaled absolute difference between the original and denoised image; kurtosis distributions of the original images; kurtosis distributions of the denoised images. Dash line represents the mean kurtosis value of each region.

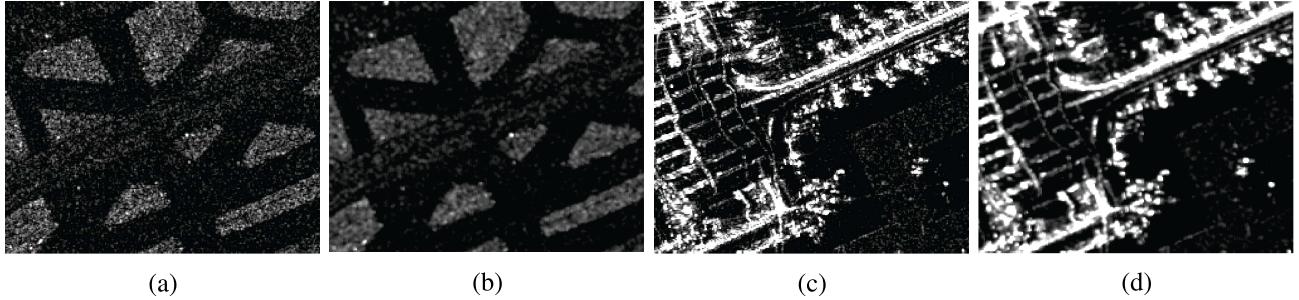


Fig. 12. Removing multiplicative speckle noise from a real noisy SAR image (Spain, Barcelona-ORI, SE) with the Kuan filter [50] and our estimated noise variance. (a) and (c) are two selected regions of the noisy SAR image; (b) and (d) are the corresponding restored versions.

achieved when  $S = 3$ , though the gains over the case of  $S = 2$  are not significant.

**2) Computational Complexity:** To measure the computational complexity, we test the proposed method on UCID-v2 image set. For unoptimized MATLAB implementation of our method, it takes 8.5 seconds on average for one  $512 \times 768$  image on a PC with Intel Core i5-4570 3.2 GHz CPU and 4G RAM.

## VI. CONCLUSION

This paper presents a novel noise level estimation approach by jointly considering scale-invariant kurtosis in bandpass domains and piecewise stationarity in spatial domain. We design a  $K$ -means based algorithm to adaptively partition an image into a series of disjoint regions, each of which is assumed to be associated with a constant kurtosis throughout scales. The noise level estimation is then cast into a problem to optimally fit this new kurtosis model, which can be solved efficiently by resorting to an iterative alternating optimization procedure. Moreover, an estimation rectification procedure based on noise injection is developed, which effectively eliminates the estimation bias. Extensive experimental results

demonstrate that our method can reliably estimate the noise level for a variety of noise types, and outperforms the state-of-the-art techniques, especially for non-Gaussian noises. Our method could be readily adopted into real applications that require the noise level as a crucial parameter.

## VII. APPENDIX A DISCUSSION ON THE CONVEXITY OF SUB-PROBLEM (11)

The objective function in (11) can be written as

$$\begin{aligned} & \sum_{i=1}^S \sum_{j=1}^M \alpha_i \kappa(y_j^i) + 2 \sum_{i=1}^S \sum_{j=1}^M \alpha_i \sqrt{\kappa(y_j^i)} \left( \frac{\sigma_n^2}{\sigma_{y_j^i}^2} - 1 \right) \sqrt{\kappa(x^i)} \\ & + \sum_{i=1}^S \sum_{j=1}^M \alpha_i \left( \frac{\sigma_n^2}{\sigma_{y_j^i}^2} - 1 \right)^2 \kappa(x^i) \\ & - \lambda \sum_{i=1}^S \sum_{j=1}^S \left( \sqrt{\kappa(x^i)} - \sqrt{\kappa(x^j)} \right)^2 \end{aligned} \quad (31)$$

As the first term is irrelevant to the optimization variables, it can be removed from the optimization. The remaining two

terms can then be simplified into the following matrix form

$$\mathbf{k}'(\mathbf{H} - \lambda\mathbf{R})\mathbf{k} + \mathbf{c}'\mathbf{k} \quad (32)$$

where

- $\mathbf{k} = [\sqrt{\kappa(x^1)}, \sqrt{\kappa(x^2)}, \dots, \sqrt{\kappa(x^S)}]'$
- $\mathbf{H}$  is a diagonal matrix of size  $S \times S$  and its diagonal elements are given by

$$H_{ii} = \sum_{j=1}^M \alpha_i \left( \frac{\sigma_n^2}{\sigma_{y_j^i}^2} - 1 \right)^2 \quad (33)$$

- $\mathbf{R}$  is a symmetric matrix given by

$$R_{ij} = \begin{cases} S-1, & i = j \\ -1, & \text{otherwise} \end{cases} \quad (34)$$

- $\mathbf{c}$  is a vector of length  $S$  given by

$$c_i = \sum_{j=1}^M 2\alpha_i \sqrt{\kappa(y_j^i)} \left( \frac{\sigma_n^2}{\sigma_{y_j^i}^2} - 1 \right) \quad (35)$$

If  $(\mathbf{H} - \lambda\mathbf{R})$  is positive definite, the optimization problem (11) is convex. In our experiments, we set  $\lambda = 0.01$ , and we find that  $(\mathbf{H} - \lambda\mathbf{R})$  is always positive definite under this setting for all the test images.

## APPENDIX B

### CLOSED-FORM SOLUTION FOR SUB-PROBLEM (12)

Consider the objective function of the sub-problem (12):

$$L(\sigma_n^2) = \sum_{i=1}^S \alpha_i \sum_{j=1}^M \left( \sqrt{\kappa(y_j^i)} - \sqrt{\hat{\kappa}(x^i)} + \frac{\sigma_n^2}{\sigma_{y_j^i}^2} \sqrt{\hat{\kappa}(x^i)} \right)^2 \quad (36)$$

To minimize  $L(\cdot)$ , we take partial derivative of  $L(\cdot)$  with respect to  $\sigma_n^2$ , and set it to zero, which yields

$$2 \sum_{i=1}^S \sum_{j=1}^M \alpha_i \left( \sqrt{\kappa(y_j^i)} - \sqrt{\hat{\kappa}(x^i)} + \frac{\sigma_n^2}{\sigma_{y_j^i}^2} \sqrt{\hat{\kappa}(x^i)} \right) \frac{\sqrt{\hat{\kappa}(x^i)}}{\sigma_{y_j^i}^2} = 0 \quad (37)$$

We then can obtain the optimal solution  $\hat{\sigma}_n^2$  by

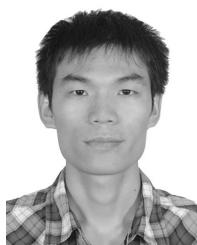
$$\hat{\sigma}_n^2 = \frac{\sum_{ij} \alpha_i \left( \sqrt{\hat{\kappa}(x^i)} - \sqrt{\kappa(y_j^i)} \right)}{\sum_{ij} \frac{\alpha_i}{\sigma_{y_j^i}^2} \sqrt{\hat{\kappa}(x^i)}} \quad (38)$$

We further check the second-order conditions to ensure (38) is the global minima of (36).

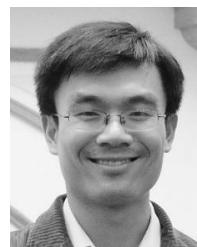
## REFERENCES

- [1] A. C. Bovik, *Handbook of Image and Video Processing (Communications, Networking and Multimedia)*. Orlando, FL, USA: Academic, 2005.
- [2] A. Buades, B. Coll, and J.-M. Morel, "A non-local algorithm for image denoising," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2, Jun. 2005, pp. 60–65.
- [3] 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.
- [4] X. Liu, M. Tanaka, and M. Okutomi, "Single-image noise level estimation for blind denoising," *IEEE Trans. Image Process.*, vol. 22, no. 12, pp. 5226–5237, Dec. 2013.
- [5] Q. Shan, J. Jia, and A. Agarwala, "High-quality motion deblurring from a single image," *ACM Trans. Graph. (SIGGRAPH)*, vol. 27, no. 3, pp. 1–5, Aug. 2008.
- [6] P. L. Rosin, "Thresholding for change detection," in *Proc. Int. Conf. Comput. Vis.*, 1998, pp. 274–279.
- [7] W. Dong, L. Zhang, G. Shi, and X. Wu, "Image deblurring and super-resolution by adaptive sparse domain selection and adaptive regularization," *IEEE Trans. Image Process.*, vol. 20, no. 7, pp. 1838–1857, Jul. 2011.
- [8] Y. W. Wen, M. K. Ng, and Y. M. Huang, "Efficient total variation minimization methods for color image restoration," *IEEE Trans. Image Process.*, vol. 17, no. 11, pp. 2081–2088, Nov. 2008.
- [9] B. J. Kang and K. R. Park, "Real-time image restoration for iris recognition systems," *IEEE Trans. Syst., Man, B, Cybern.*, vol. 37, no. 6, pp. 1555–1566, Dec. 2007.
- [10] S. Lyu, X. Pan, and X. Zhang, "Exposing region splicing forgeries with blind local noise estimation," *Int. J. Comput. Vis.*, vol. 110, no. 2, pp. 202–221, Nov. 2014.
- [11] B. Mahdian and S. Saic, "Using noise inconsistencies for blind image forensics," *Image Vis. Comput.*, vol. 27, no. 10, pp. 1497–1503, 2009.
- [12] J. Immerkær, "Fast noise variance estimation," *Comput. Vis. Image Understand.*, vol. 64, no. 2, pp. 300–302, Sep. 1996.
- [13] S.-C. Tai and S.-M. Yang, "A fast method for image noise estimation using Laplacian operator and adaptive edge detection," in *Proc. Int. Symp. Commun., Control Signal Process.*, Mar. 2008, pp. 1077–1081.
- [14] S.-M. Yang and S.-C. Tai, "Fast and reliable image-noise estimation using a hybrid approach," *J. Electron. Imag.*, vol. 19, no. 3, pp. 033007-1–033007-15, 2010.
- [15] K. Rank, M. Lendl, and R. Unbehauen, "Estimation of image noise variance," *IEE Proc. Vis., Image Signal Process.*, vol. 146, no. 2, pp. 80–84, Aug. 1999.
- [16] A. Amer and E. Dubois, "Fast and reliable structure-oriented video noise estimation," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 15, no. 1, pp. 113–118, Jan. 2005.
- [17] D.-H. Shin, R.-H. Park, S. Yang, and J.-H. Jung, "Block-based noise estimation using adaptive Gaussian filtering," *IEEE Trans. Consum. Electron.*, vol. 51, no. 1, pp. 218–226, Feb. 2005.
- [18] 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.
- [19] C.-H. Wu and H.-H. Chang, "Superpixel-based image noise variance estimation with local statistical assessment," *EURASIP Int. J. Image Video Process.*, vol. 2015, no. 1, pp. 1–12, Dec. 2015.
- [20] D. L. Donoho and I. M. Johnstone, "Adapting to unknown smoothness via wavelet shrinkage," *J. Amer. Statist. Assoc.*, vol. 90, no. 432, pp. 1200–1224, 1995.
- [21] D. L. Donoho, "De-noising by soft-thresholding," *IEEE Trans. Inf. Theory*, vol. 41, no. 3, pp. 613–627, May 1995.
- [22] S. G. Chang, B. Yu, and M. Vetterli, "Adaptive wavelet thresholding for image denoising and compression," *IEEE Trans. Image Process.*, vol. 9, no. 9, pp. 1532–1546, Sep. 2000.
- [23] T. Li, M. Wang, and T. Li, "Estimating noise parameter based on the wavelet coefficients estimation of original image," in *Proc. Int. Conf. Challeng. Environ. Sci. Comput. Eng.*, vol. 1, Mar. 2010, pp. 126–129.
- [24] A. De Stefano, P. White, and W. Collis, "Training methods for image noise level estimation on wavelet components," *EURASIP J. Appl. Signal Process.*, Jan. 2004, pp. 2400–2407.
- [25] W. Liu and W. Lin, "Additive white Gaussian noise level estimation in SVD domain for images," *IEEE Trans. Image Process.*, vol. 22, no. 3, pp. 872–883, Mar. 2013.
- [26] D. Zoran and Y. Weiss, "Scale invariance and noise in natural images," in *Proc. IEEE Int. Conf. Comput. Vis.*, Sep./Oct. 2009, pp. 2209–2216.
- [27] G. Zhai and X. Wu, "Noise estimation using statistics of natural images," in *Proc. IEEE Int. Conf. Image Process.*, Sep. 2011, pp. 1857–1860.
- [28] C. Tang, X. Yang, and G. Zhai, "Noise estimation of natural images via statistical analysis and noise injection," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 25, no. 8, pp. 1283–1294, Aug. 2015.
- [29] A. Srivastava, A. B. Lee, E. P. Simoncelli, and S.-C. Zhu, "On advances in statistical modeling of natural images," *J. Math. Imag. Vis.*, vol. 18, no. 1, pp. 17–33, 2003.
- [30] E. P. Simoncelli and B. A. Olshausen, "Natural image statistics and neural representation," *Annu. Rev. Neurosci.*, vol. 24, no. 1, pp. 1193–1216, 2001.

- [31] S. S. Haykin, J. C. Principe, T. J. Sejnowski, and J. McWhirter, *New Directions in Statistical Signal Processing—From Systems to Brains*. Cambridge, MA, USA: MIT Press, 2007.
- [32] W. Feller, *An Introduction to Probability Theory and Its Applications*, 3rd ed., vol. 1. Hoboken, NJ, USA: Wiley, Jan. 1968.
- [33] E. Y. Lam, “A mathematical analysis of the DCT coefficient distributions for images,” *IEEE Trans. Image Process.*, vol. 9, no. 10, pp. 1661–1666, Oct. 2000.
- [34] M. Bethge, “Factorial coding of natural images: How effective are linear models in removing higher-order dependencies?” *J. Opt. Soc. Amer. A, Opt. Image Sci. Vis.*, vol. 23, no. 6, pp. 1253–1268, Jun. 2006.
- [35] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
- [36] M. Grant and S. Boyd. (Mar. 2014). *CVX: Matlab Software for Disciplined Convex Programming*. [Online]. Available: <http://cvxr.com/cvx>
- [37] X. Ren and J. Malik, “Learning a classification model for segmentation,” in *Proc. IEEE Int. Conf. Comput. Vis.*, Oct. 2003, pp. 10–17.
- [38] G. Mori, “Guiding model search using segmentation,” in *Proc. 10th IEEE Int. Conf. Comput. Vis. (ICCV)*, vol. 2, Oct. 2005, pp. 1417–1423.
- [39] L. Rudin, P.-L. Lions, and S. Osher, “Multiplicative denoising and deblurring: Theory and algorithms,” in *Geometric Level Set Methods in Imaging, Vision, and Graphics*. New York, NY, USA: Springer, 2003, pp. 103–119.
- [40] K. Krishnamoorthy, *Handbook of Statistical Distributions with Applications*. Boca Raton, FL, USA: CRC Press, 2016.
- [41] J. M. Bioucas-Dias and M. A. Figueiredo, “Multiplicative noise removal using variable splitting and constrained optimization,” *IEEE Trans. Image Process.*, vol. 19, no. 7, pp. 1720–1730, Jun. 2010.
- [42] R. Farebrother, “The cumulants of the logarithm of a gamma variable,” *J. Statist. Comput. Simul.*, vol. 36, no. 4, pp. 243–245, 1990.
- [43] R. Franzen. (2004). *Kodak Lossless True Color Image Suite*. [Online]. Available: <http://r0k.us/graphics/kodak>
- [44] G. Schaefer and M. Stich, “UCID: An uncompressed color image database,” *Proc. SPIE*, vol. 5307, pp. 472–480, Dec. 2003.
- [45] J. S. Lim, *Two-Dimensional Signal and Image Processing*, vol. 1. Englewood Cliffs, NJ, USA: Prentice-Hall, 1990.
- [46] Oct. 2016, *Airbus Defence and Space TerraSAR-X—Radar Dataset Sample imagery*. Available: [Online]. <http://www.intelligence-airbuds.com/en/23-sample-imagery>
- [47] V. S. Frost, J. A. Stiles, K. S. Shanmugan, and J. C. Holtzman, “A model for radar images and its application to adaptive digital filtering of multiplicative noise,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-4, no. 2, pp. 157–166, Mar. 1982.
- [48] G. Aubert and J.-F. Aujol, “A variational approach to removing multiplicative noise,” *SIAM J. Appl. Math.*, vol. 68, no. 4, pp. 925–946, Jan. 2008.
- [49] D. T. Kuan, A. A. Sawchuk, T. C. Strand, and P. Chavel, “Adaptive noise smoothing filter for images with signal-dependent noise,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 7, no. 2, pp. 165–177, Mar. 1985.
- [50] D. T. Kuan, A. A. Sawchuk, T. C. Strand, and P. Chavel, “Adaptive restoration of images with speckle,” *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-35, no. 3, pp. 373–383, Mar. 1987.



**Li Dong** (S’14) received the B.Eng. degree in software engineering from Chongqing University, Chongqing, China, in 2012, and the M.S. degree in software engineering from the University of Macau, Macau, China, in 2014, where he is currently pursuing the Ph.D. degree with the Department of Computer and Information Science, Faculty of Science and Technology. His research interests include statistical image modeling and processing, multimedia security and forensic, and machine learning.



**Jiantao Zhou** (M’11) received the B.Eng. degree from the Department of Electronic Engineering, Dalian University of Technology, Dalian, China, in 2002, the M.Phil. degree from the Department of Radio Engineering, Southeast University, Nanjing, China, in 2005, and the Ph.D. degree from the Department of Electronic and Computer Engineering, The Hong Kong University of Science and Technology, Hong Kong, in 2009. He held various research positions with the University of Illinois at Urbana-Champaign, The Hong Kong University of Science and Technology, and McMaster University. He is currently an Assistant Professor with the Department of Computer and Information Science, Faculty of Science and Technology, University of Macau. His research interests include multimedia security and forensics, and high-fidelity image compression. He holds three granted U.S. patents and two granted Chinese patents. He was a co-author of a paper that received the best paper award at the IEEE Pacific-Rim Conference on Multimedia in 2007.



**Yuan Yan Tang** (S’88–M’88–SM’96–F’04) is currently a Chair Professor with the Faculty of Science and Technology, University of Macau, and also a Professor, an Adjunct Professor, and an Honorary Professor with several institutes, including Chongqing University, China, Concordia University, Canada, and Hong Kong Baptist University, Hong Kong. He has authored over 400 academic papers and over 25 monographs/books/book chapters. His current interests include wavelets, pattern recognition, image processing, artificial intelligence. He is the Founder and the Editor-in-Chief of the *International Journal on Wavelets, Multiresolution, and Information Processing*, and an Associate Editor of several international journals. He is the Founder and the Chair of the pattern recognition committee in the IEEE SMC. He has served as the General Chair, the Program Chair, and a Committee Member of many international conferences. He is currently the Founder and the General Chair of the series International Conferences on Wavelets Analysis and Pattern Recognition. He is also the Founder and the Chair of the Macau Branch of International Associate of Pattern Recognition (IAPR). He is a fellow of IAPR.