

# Natural Scene Statistics for Noise Estimation

Praful Gupta, Christos G. Bampis, Yize Jin and Alan C. Bovik

Department of Electrical and Computer Engineering, University of Texas at Austin, Austin, USA

**Abstract**—We investigate the scale-invariant properties of divisively normalized bandpass responses of natural images in the DCT-filtered domain. We found that the variance of the normalized DCT filtered responses of a pristine natural image is scale invariant. This scale invariance property does not hold in the presence of noise and thus it can be used to devise an efficient blind image noise estimator. The proposed noise estimation approach outperforms other statistics-based methods especially for higher noise levels and competes well with patch-based and filter-based approaches. Moreover, the new variance estimation approach is also effective in the case of non-Gaussian noise. The research code of the proposed algorithm can be found at [https://github.com/guptapraful/Noise\\_Estimation](https://github.com/guptapraful/Noise_Estimation).

**Index Terms**—scale invariance, normalized bandpass responses, noise estimation

## I. INTRODUCTION

Images can be corrupted by noise during image acquisition, transmission and storage. Blind image noise estimation, i.e., estimating the noise variance when the reference image is not known, is crucial for various image processing applications including image denoising [1], image segmentation [2], object recognition [3] and image restoration [4].

The statistics of natural images have been well studied in the literature over the past many years [5]. One of the most remarkable property of natural scenes is their scale invariance property, which dates back to Field *et al.* [6] and Burton *et al.* [7]. They discovered that the power spectrum of bandpass-filtered natural images decays as  $\frac{A}{|k|^{2-\eta}}$ , where  $|k|$  is the magnitude of the spatial frequency and  $\eta$  takes a small value that varies depending on the image content. Other scale-invariance properties of natural scenes have also been studied in [8], [9].

The higher order statistics of natural images have also been investigated extensively. It is well known that natural scenes are highly non-Gaussian and the marginal distribution of their bandpass responses is highly kurtotic, i.e., they exhibit sharp peaks and heavy tails [8]. This non-Gaussianity was demonstrated to be modeled well by a generalized Laplacian density [10] and generalized Gaussian density [11]. Other works have described the relationships between subband coefficients from adjacent scales and orientations of natural images as a Gaussian Scale Mixture (GSM) model, and successfully applied the GSM model in various imaging applications including image denoising [1], image restoration [4], full-reference [12], reduced reference [13], [14] and no-reference [15]–[17] image and video quality assessment in both the spatial and wavelet transform domains.

More recently, Zoran *et al.* [18] observed the scale invariance property of kurtosis, the fourth order moment, of

natural images. Specifically, the kurtosis of marginal bandpass coefficients obtained after filtering the image using the DCT basis were demonstrated to remain constant over different scales. This scale-invariance property was successfully utilized to estimate the image noise variance in the DCT [18], random unitary [19] and various other linear transform domains. Further, the scale-invariance of higher-order statistics was used along with the piecewise stationarity of natural scenes in spatial domain for more robust estimation of noise variance [20].

Apart from statistics-based noise estimation methods, there exist filter-based methods which rely on noise estimation of residual images obtained after suppressing the image signals using highpass filtering [21], and patch-based methods that estimate noise levels from a handful of carefully selected homogeneous image patches [22]. Although these filter-based and patch-based approaches estimate the noise variance quite accurately, they lack the conceptual elegance of statistics-based methods that utilize the regular properties of pristine natural images which are affected in the presence of distortions.

We propose a noise estimation approach by exploiting the scale-invariance properties of *normalized, bandpass-filtered* natural image coefficients. We first describe the scale-invariant properties of natural images in Section II, then develop a blind image noise estimation approach using these principles in Section III. Section IV studies the performance of the proposed approach and Section V concludes with future work.

## II. SCALE INVARIANCE PROPERTY OF NORMALIZED DCT FILTERED RESPONSES

In this work, we study the statistics of the *normalized* bandpass responses of the natural images and leverage these to estimate the image noise variance effectively. In particular, we investigate the scale invariance of locally, contrast-normalized coefficients distribution of DCT filtered responses. The motivation behind this study stems from the seminal work of Ruderman [9] on natural images, who observed the scale invariance of the pixel contrasts  $\log(I(x)/I_o)$  in the spatial domain. Divisive normalization using the local energies has a decorrelating and Gaussianizing effect on the bandpass statistics of natural images. It is a perceptually significant step that accounts for gain control – the non-linear behavior of the neurons in the human visual cortex [23]. Since divisive normalization decorrelates the image structures, it is suitable for image noise estimation. We utilize this observation together with the reasonable assumption of scale invariance of normalized bandpass responses to develop an accurate blind image noise estimator.

Before describing the noise estimation steps in detail, we analyze the statistics of normalized bandpass coefficients. First the image,  $I(i, j)$ , is filtered using a DCT basis filter,  $h(i, j)$ , to obtain the bandpass response  $\hat{I}(i, j)$ :

$$\hat{I}(i, j) = h(i, j) * I(i, j)$$

Let  $\bar{I}(i, j)$  be the normalized response, i.e., compute  $\bar{I}(i, j) = \hat{I}(i, j)/z(i, j)$ , where  $z(i, j)$  is the variance field and is determined using the neighboring coefficients as:

$$z(i, j) = \sqrt{\sum_{k=-K}^K \sum_{l=-L}^L w_{k,l} (I_{k,l}(i, j) - \mu(i, j))^2}$$

where  $w = \{w_{k,l}, k = -K, \dots, K \text{ and } l = -L, \dots, L\}$  is a unit volume 2D Gaussian weighting function. Since the DCT filter,  $h(i, j)$  is zero-mean, the mean field,  $\mu(i, j)$  simplifies to 0. We note that the above operation is very similar to contrast normalization used in no-reference IQA in the spatial [15], [16] and wavelet domain [17]. Fig. 1 shows the histograms of before and after normalization of the marginal DCT filtered responses.

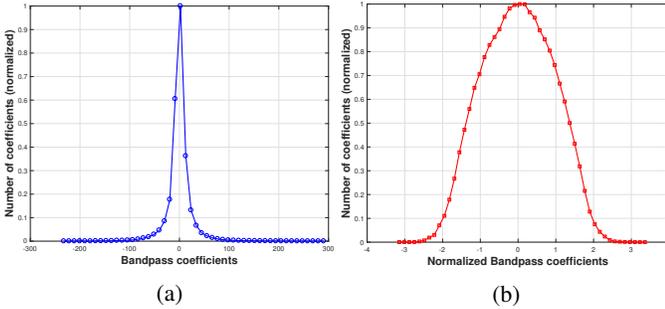


Fig. 1: Distribution of the DCT filtered coefficients (a) before and (b) after normalization.

Next, we consider the scale invariance of the distribution of normalized bandpass coefficients. Fig. 2 indicates that the variance of the normalized responses is roughly constant over different scales. This scale invariance principle does not hold for images afflicted with noise. As one would expect, the variance of the distribution increases as the standard deviation of the noise is increased – the effect becoming more prominent for higher frequencies as shown in Fig. 2.

### III. NOISE ESTIMATION

We begin by assuming that the image is distorted solely by additive white Gaussian noise (AWGN) with unknown variance  $\sigma_n^2$ , that is to be estimated. We denote the noisy image by a random variable  $y$ , AWGN by  $n$  and the underlying pristine image by  $x$ . Since the noise is additive we have,

$$y = x + n \quad (1)$$

An example of an image afflicted by AWGN is shown in Fig. 3. We substitute  $y$  and  $x$  by the normalized bandpass coefficients,  $\bar{y}$  and  $\bar{x}$ , and their corresponding normalizers,  $z_y$  and  $z_x$ , and then apply variance on both sides to get,

$$\text{Var}(z_y \bar{y} + \mu_y) = \text{Var}(z_x \bar{x} + \mu_x + n)$$

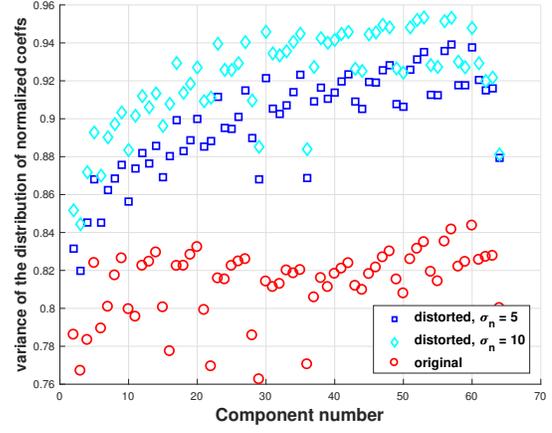


Fig. 2: The scale-invariance property of natural images in the variance of contrast normalized bandpass coefficients. Images afflicted by AWGN do not obey this property.

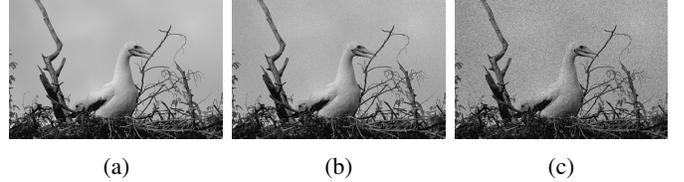


Fig. 3: (a) Original image from BSD database [24], (b) image with  $\sigma_n = 5$ , (c) image with  $\sigma_n = 15$ .

Given that  $n$  is zero-mean and independent of the original signal implies that the mean field of  $y$  and  $x$  are equal ( $\mu_y = \mu_x$ ) and that  $\text{Cov}(x, n) = 0$ , i.e.,

$$\text{Var}(z_y \bar{y}) = \text{Var}(z_x \bar{x}) + \sigma_n^2 \quad (2)$$

using the linearity property of the variance operator under independence. As in the GSM model in the wavelet domain [1], we make a reasonable assumption that the DCT bandpass coefficients  $\hat{x}$  of a pristine natural image can be written as a GSM vector and can be expressed as a product between the variance field  $z_x$  and the zero-mean Gaussian vector  $\bar{x}$ :

$$\hat{x} = z_x \bar{x} \quad (3)$$

where  $\bar{x}$  is independent of  $z_x$ . Using this independence assumption from (3) and the zero-mean property of  $\bar{x}$ , we have  $\text{Var}(z_x \bar{x}) = \mathbb{E}[z_x^2] \mathbb{E}[\bar{x}^2]$ . Plugging this back in (2) yields:

$$\text{Var}(z_y \bar{y}) = \mathbb{E}[z_x^2] \mathbb{E}[\bar{x}^2] + \sigma_n^2 \quad (4)$$

Using the spatial stationarity of noise, we can express  $\mathbb{E}[z_x^2] = \mathbb{E}[z_y^2] - \sigma_n^2$ . Similar notion was employed in [25] in which the image noise variance was estimated using average of the local variances of unique set of homogeneous patches. Substituting for  $\mathbb{E}[z_x^2]$  in (5), we get:

$$\text{Var}(z_y \bar{y}) = (\mathbb{E}[z_y^2] - \sigma_n^2) \mathbb{E}[\bar{x}^2] + \sigma_n^2 \quad (5)$$

The above equation can be used to solve for  $\sigma_n^2$ . However, since the underlying pristine image  $x$  is not known,  $\mathbb{E}[\bar{x}^2]$  is also an unknown quantity. The hypothesis we developed in the previous section that the variance of the distribution of

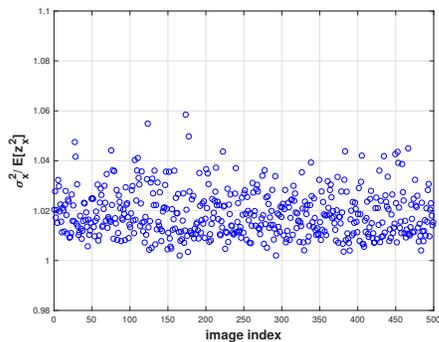


Fig. 4: Plot of  $\sigma_x^2 / \mathbb{E}[z_x^2]$  for all images from the BSD database.

the normalized bandpass coefficients,  $\mathbb{E}[\bar{x}^2]$ , is scale-invariant, and that the presence of noise violates the scale-invariance principle is utilized to address for the unknown  $\mathbb{E}[\bar{x}^2]$ . In practice, the noisy image is the observed quantity and we can recover both  $\mathbb{E}[\bar{x}^2]$  and  $\sigma_n^2$  by minimizing:

$$\hat{\sigma}_n^2, \hat{\mathbb{E}}[\bar{x}^2] = \arg \min_{\sigma_n^2, \mathbb{E}[\bar{x}^2]} \sum_{i=2}^{N^2} \left( \text{Var}(z_{y,i} \bar{y}_i) - (\mathbb{E}[z_{y,i}^2] - \sigma_n^2) \mathbb{E}[\bar{x}^2] - \sigma_n^2 \right)^2 \quad (6)$$

where  $i$  is the scale index ranging from 2,3,...,  $N^2$  ignoring the DC component ( $i = 1$ ) and  $z_{y,i}$  denotes the variance map of the noisy image at scale  $i$ . Taking the derivative of the objective to be minimized in (6) with respect to  $\sigma_n^2$  and setting it to 0 yields:

$$2 \sum_{i=2}^{N^2} \left( \text{Var}(z_{y,i} \bar{y}_i) - (\mathbb{E}[z_{y,i}^2] - \sigma_n^2) \mathbb{E}[\bar{x}^2] - \sigma_n^2 \right) \left( \mathbb{E}[\bar{x}^2] - 1 \right) = 0$$

It is clear from the above equation that one of the extrema is at  $\mathbb{E}[\bar{x}^2] = 1$ , which needs to be avoided. We utilize the other term in the above equation to get a necessary bound on  $\mathbb{E}[\bar{x}^2]$ . Equating the first term to 0 for any scale,  $k$ , followed by straightforward simplification yields:

$$\mathbb{E}[\bar{x}^2] = \frac{\text{Var}(z_{y,k} \bar{y}_k) - \sigma_n^2}{\mathbb{E}[z_{y,k}^2] - \sigma_n^2} = \frac{\sigma_{x,k}^2}{\mathbb{E}[z_{x,k}^2]}$$

For any given natural image, the ratio of the global variance at scale  $k$ ,  $\sigma_{x,k}^2$  and the average of the local variances,  $\mathbb{E}[z_{x,k}^2]$  can be assumed to be greater than 1 since images are generally piecewise smooth. To further strengthen this argument, we selected the BSD database [24] and computed the ratio  $\sigma_{x,k}^2 / \mathbb{E}[z_{x,k}^2]$  for all the images at different scales. Indeed, as shown in Fig. 4, the ratio  $\sigma_x^2 / \mathbb{E}[z_x^2]$  is larger than 1 in all cases. Integrating the lower bound with (6) yields the following optimization:

$$\hat{\sigma}_n^2, \hat{\mathbb{E}}[\bar{x}^2] = \arg \min_{\sigma_n^2, \mathbb{E}[\bar{x}^2]} \sum_{i=2}^{N^2} \left( \text{Var}(z_{y,i} \bar{y}_i) - (\mathbb{E}[z_{y,i}^2] - \sigma_n^2) \mathbb{E}[\bar{x}^2] - \sigma_n^2 \right)^2 \quad (7)$$

subject to  $\mathbb{E}[\bar{x}^2] > 1$ , which can be solved for  $\sigma_n^2$  using any non-linear optimization routine. Here we used Matlab's `fminsearchbnd` function to perform the constrained non-linear optimization.

## IV. PERFORMANCE EVALUATION

### A. Additive White Gaussian Noise

We evaluated the performance of the proposed noise estimator on two databases: Berkeley Segmentation Database (BSD) [24] (100 images) and LIVE database [26] (29 images). To simulate the effects of noise, we added AWGN to every high-quality image using a  $\sigma_n$  from 5 to 50 in steps of 5.

Tables I and II show the performance in root-mean-squared-error (RMSE) of the proposed method against that of other well-known and state-of-the-art noise estimators for each noise level averaged across all the images in BSD and LIVE database respectively. We infer from these tables that the performance of the proposed noise estimator is competitive to PCA based estimation [22] for mid-level noise standard deviations and outperforms other methods for higher noise levels. We note that the PCA based noise estimation approach makes an initial guess of the noise's standard deviation upper bound and significantly depends on the initialization. By contrast, our approach is very simple, is based on natural scene statistics models, requires no tuning and is highly robust to the initialization point.

It is important to observe that patch-based approaches [22], [27] are significantly better for larger noise levels since they adaptively consider homogeneous patches for noise estimation neglecting patches where the contribution from image structure is significant. Statistics-based approaches as in [18], [20]<sup>1</sup> are sensitive to higher noise levels since the contribution of variance from the edges is relatively small to the noise variance, affecting the kurtosis in an unpredictable manner. By contrast, our approach competes well with patch-based methods even for higher noise levels, since it decorrelates the image structures and considers second-order statistics, enabling the optimizer to recover the noise standard deviation more precisely.

### B. Non-Gaussian Noise

While our approach assumes that the noise added is Gaussian-distributed, similar principles hold in the non-Gaussian case. Zoran *et al.* [18] showed that non-Gaussian independent noise in the pixel domain results in Gaussian noise in the transform domain due to the central limit theorem and the noise independence. Therefore, we can apply similar ideas in the case of non-Gaussian noise, including Laplacian and Uniform-distributed, as shown in Fig. 5. It can be seen that the approach in [18] performs poorly for higher noise levels, while the proposed method maintains a much better performance level.

## V. CONCLUSION

We described the scale invariance of contrast-normalized bandpass responses of natural images in the DCT-filtered domain. We demonstrated that in the presence of additive Gaussian noise this principle is violated and we used this

<sup>1</sup>Since we were not able to find a publicly available implementation of [20], we did not compare its performance in Tables I and II.

TABLE I: Comparison of different noise estimation algorithms in terms of RMSE between the predicted and actual  $\sigma_n$  evaluated at each noise level on BSD database, The best two results are highlighted in boldface.

|                            | $\sigma_n$  |             |             |             |             |             |             |             |             |             |
|----------------------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
|                            | 5           | 10          | 15          | 20          | 25          | 30          | 35          | 40          | 45          | 50          |
| Zoran <i>et al.</i> [18]   | 1.40        | 1.00        | 1.71        | 3.79        | 9.47        | 14.08       | 21.11       | 27.01       | 34.00       | 38.88       |
| Liu <i>et al.</i> [27]     | <b>0.21</b> | <b>0.36</b> | 0.52        | 0.75        | 1.08        | 1.52        | 2.07        | 2.77        | <b>3.61</b> | <b>4.61</b> |
| Pan <i>et al.</i> [28]     | 1.45        | 1.14        | 1.44        | 1.97        | 2.66        | 3.53        | 4.55        | 5.73        | 7.08        | 8.58        |
| Pyatykh <i>et al.</i> [22] | <b>0.31</b> | <b>0.26</b> | <b>0.34</b> | <b>0.43</b> | <b>0.57</b> | <b>0.80</b> | <b>1.17</b> | <b>1.66</b> | <b>2.29</b> | <b>3.11</b> |
| Proposed                   | 1.24        | 0.79        | <b>0.49</b> | <b>0.60</b> | <b>0.92</b> | <b>1.39</b> | <b>1.99</b> | <b>2.74</b> | 3.63        | 4.67        |

TABLE II: Comparison of different noise estimation algorithms in terms of RMSE between the predicted and actual  $\sigma_n$  evaluated at each noise level on LIVE database. The best two results are highlighted in boldface.

|                            | $\sigma_n$  |             |             |             |             |             |             |             |             |             |
|----------------------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
|                            | 5           | 10          | 15          | 20          | 25          | 30          | 35          | 40          | 45          | 50          |
| Zoran <i>et al.</i> [18]   | 0.83        | 0.86        | 0.81        | 4.61        | 6.96        | 13.38       | 18.99       | 25.97       | 32.98       | 38.91       |
| Liu <i>et al.</i> [27]     | <b>0.24</b> | 0.48        | 0.74        | 1.05        | 1.42        | 1.88        | 2.44        | 3.12        | 3.94        | 4.92        |
| Pan <i>et al.</i> [28]     | 0.88        | 0.85        | 1.20        | 1.73        | 2.42        | 3.25        | 4.22        | 5.34        | 6.60        | 8.01        |
| Pyatykh <i>et al.</i> [22] | <b>0.16</b> | <b>0.12</b> | <b>0.18</b> | <b>0.27</b> | <b>0.43</b> | <b>0.68</b> | <b>1.01</b> | <b>1.47</b> | <b>2.11</b> | <b>2.88</b> |
| Proposed                   | 1.28        | <b>0.46</b> | <b>0.52</b> | <b>0.80</b> | <b>1.20</b> | <b>1.70</b> | <b>2.30</b> | <b>3.04</b> | <b>3.90</b> | <b>4.90</b> |

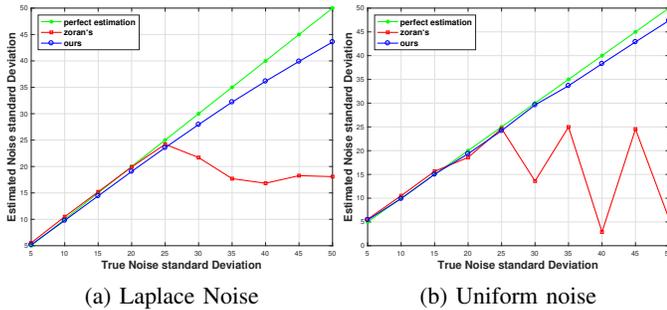


Fig. 5: Noise estimation results on Lena image for Laplace and Uniform noise of different standard deviations

observation to blindly estimate the noise variance. The proposed noise estimator compares well with other state-of-the-art noise estimation approaches on different databases across wide range of standard deviations. Finally, we also showed that our approach can be also be used to estimate additive white non-Gaussian noise with promising performance.

## REFERENCES

- J. Portilla, V. Strela, M. J. Wainwright, and E. P. Simoncelli, "Image denoising using scale mixtures of gaussians in the wavelet domain," *IEEE Trans. on Image Proc.*, vol. 12, no. 11, pp. 1338–1351, 2003.
- P. Rosin, "Thresholding for change detection," in *Intl. Conf. on Comp. Vision*, 1998, pp. 274–279.
- B. J. Kang and K. R. Park, "Real-time image restoration for iris recognition systems," *IEEE Trans. on Systems, Man, and Cybernetics*, vol. 37, no. 6, pp. 1555–1566, 2007.
- Y.-W. Wen, M. K. Ng, and Y.-M. Huang, "Efficient total variation minimization methods for color image restoration," *IEEE Transactions on Image Processing*, vol. 17, no. 11, pp. 2081–2088, 2008.
- G. Burton, N. Haig, and I. Moorhead, "A self-similar stack model for human and machine vision," *Biological Cybernetics*, vol. 53, no. 6, pp. 397–403, 1986.
- D. J. Field, "Relations between the statistics of natural images and the response properties of cortical cells," *Journ. of the Opt. Society of America*, vol. 4, no. 12, pp. 2379–2394, 1987.
- G. Burton and I. R. Moorhead, "Color and spatial structure in natural scenes," *Applied Optics*, vol. 26, no. 1, pp. 157–170, 1987.
- D. L. Ruderman, "Origins of scaling in natural images," *Vision research*, vol. 37, no. 23, pp. 3385–3398, 1997.
- D. L. Ruderman and W. Bialek, "Statistics of natural images: Scaling in the woods," in *Adv. in Neural Information Proc. Sys.*, 1994, pp. 551–558.
- S. G. Mallat, "A theory for multiresolution signal decomposition: the

- wavelet representation," *IEEE Trans. on Pattern Anal. Machine Intell.*, vol. 11, no. 7, pp. 674–693, 1989.
- M. Bethge, "Factorial coding of natural images: how effective are linear models in removing higher-order dependencies?" *JOSA A*, vol. 23, no. 6, pp. 1253–1268, 2006.
- H. R. Sheikh and A. C. Bovik, "Image information and visual quality," *IEEE Trans. Image Process.*, vol. 15, no. 2, pp. 430–444, 2006.
- R. Soundararajan and A. C. Bovik, "RRED indices: Reduced reference entropic differencing for image quality assessment," *IEEE Trans. Image Process.*, vol. 21, no. 2, pp. 517–526, 2012.
- C. G. Bampis, P. Gupta, R. Soundararajan, and A. C. Bovik, "SpEED-QA: Spatial efficient entropic differencing for image and video quality," *IEEE Signal Proc. Letters*, vol. 24, no. 9, pp. 1333–1337, 2017.
- A. Mittal, A. K. Moorthy, and A. C. Bovik, "No-reference image quality assessment in the spatial domain," *IEEE Tran. Image Process.*, vol. 21, no. 12, pp. 4695–4708, 2012.
- A. Mittal, R. Soundararajan, and A. C. Bovik, "Making a "completely blind" image quality analyzer," *IEEE Signal Proc. Letters*, vol. 20, no. 3, pp. 209–212, 2013.
- A. K. Moorthy and A. C. Bovik, "Blind image quality assessment: From natural scene statistics to perceptual quality," *IEEE Trans. Image Process.*, vol. 20, no. 12, pp. 3350–3364, 2011.
- D. Zoran and Y. Weiss, "Scale invariance and noise in natural images," in *Intl. Conf. on Comp. Vision*, 2009., 2009, pp. 2209–2216.
- G. Zhai and X. Wu, "Noise estimation using statistics of natural images," in *Int'l Conf. Image Process*, 2011, 2011, pp. 1857–1860.
- L. Dong, J. Zhou, and Y. Y. Tang, "Noise level estimation for natural images based on scale-invariant kurtosis and piecewise stationarity," *IEEE Trans. on Image Proc.*, vol. 26, no. 2, pp. 1017–1030, 2017.
- J. Immerkaer, "Fast noise variance estimation," *Comp. Vision and image Understanding*, vol. 64, no. 2, pp. 300–302, 1996.
- S. Pyatykh, J. Hesser, and L. Zheng, "Image noise level estimation by principal component analysis," *IEEE Trans. on Image Proc.*, vol. 22, no. 2, pp. 687–699, 2013.
- M. J. Wainwright, O. Schwartz, and E. P. Simoncelli, "Natural image statistics and divisive normalization: modeling nonlinearities and adaptation in cortical neurons," *Stat. Theories of Brain*, pp. 203–222, 2002.
- D. Martin, C. Fowlkes, D. Tal, and J. Malik, "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics," in *Intl. Conf. on Comp. Vision*, 2001., vol. 2, 2001, pp. 416–423.
- P. Jiang and J.-z. Zhang, "Fast and reliable noise level estimation based on local statistic," *Pattern Recognition Letters*, vol. 78, pp. 8–13, 2016.
- H. R. Sheikh, M. F. Sabir, and A. C. Bovik, "A statistical evaluation of recent full reference image quality assessment algorithms," *IEEE Trans. on Image Proc.*, vol. 15, no. 11, pp. 3440–3451, Nov. 2006.
- X. Liu, M. Tanaka, and M. Okutomi, "Noise level estimation using weak textured patches of a single noisy image," in *Int'l Conf. Image Process*, 2012, 2012, pp. 665–668.
- X. Pan, X. Zhang, and S. Lyu, "Exposing image splicing with inconsistent local noise variances," in *Int'l Conf. on Comp. Photography*, 2012, 2012, pp. 1–10.