



EE 638: ESTIMATION & IDENTIFICATION

PROJECT REPORT

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

*Tanya Choudhary*

November 28, 2018

## 1 Introduction

Most of the existing denoising schemes assume the availability of the noise level beforehand. Unfortunately in practice it is not known. Thus estimation of it is an important step in all applications. Estimating noise level from a single image is an ill-posed problem due to the insufficient prior information of noise. 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. The paper [1] implemented uses kurtosis to estimate noise level. It takes into account the piecewise-stationarity of natural images. And based on that derives an algorithm to estimate the noise level which optimally fits the generalized kurtosis model.

## 2 Kurtosis

Kurtosis is a high-order statistical quantity employed for statistical analysis of natural images. For a random variable  $X$  it is defined as,

$$\kappa(X) = \frac{C_4(X)}{C_2^2(X)}$$

where  $C_k$  is the  $k^{th}$  cumulant function. Kurtosis measures the peakedness of a distribution. For Gaussian distribution, the kurtosis value is 0, while for a distribution more concentrated than Gaussian distribution, its kurtosis is positive and for less its negative. Kurtosis value is assumed to be constant in small non-overlapping regions of clean image, throughout all scales. That means if we separate the regions of a clean and pass them through different bandpass filters, kurtosis value remains unchanged.

## 3 Method

Noise is chosen to be i.i.d and additive in spatial domain, i.e.

$$\mathbf{y} = \mathbf{x} + \mathbf{n} \quad (1)$$

where  $\mathbf{x}$  and  $\mathbf{y}$  are clean and noisy vectorized image patches respectively.  $\mathbf{n}$  is zero-mean, signal-independent AWGN with variance  $\sigma_n^2$ , which is to be estimated. To handle different noise models, estimation is performed in linear transform domain. This is because linear transforms can mix non-Gaussian noise in spatial domain to Gaussian noise in transform domain. Transformed relation is

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

where  $d_j$  denotes the  $j^{th}$  basis vector of a complete orthonormal basis matrix  $\mathbf{D}$ . We take only the Gaussian case here. Due to orthonormality of  $\mathbf{D}$ , noise distribution in the

transformed domain too is Gaussian,  $n_j \sim \mathcal{N}(0, \sigma^2)$ . As  $x_j$  and  $n_j$  are independent, we have,

$$\begin{aligned}\sigma_{y_j}^2 &= \sigma_{x_j}^2 + \sigma_{n_j}^2 \\ C_4(y_j) &= C_4(x_j) + C_4(n_j)\end{aligned}$$

Since  $C_2(x_j) = \sigma_{x_j}^2$ , we have

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

This shows the relationship between noise variance and kurtosis,  $\kappa(y_j)$  can be represented as the linear combination of  $\kappa(x_j)$  and  $\kappa(n_j)$ . Since  $n_j$  follows a Gaussian distribution,  $\kappa(n_j) = 0$ . Also replacing  $\sigma_{x_j}^2$  with  $\sigma_{y_j}^2 - \sigma_{n_j}^2$ , we have

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

There are two unknowns in this equation, the noise variance  $\sigma_n^2$  to be estimated and the kurtosis  $\kappa(x_j)$  associated with the clean image. To solve them the noisy image is divided into  $S$  disjoint regions, each of whose clean version is associated with a constant but unknown kurtosis throughout all scales. Thus the above equation holds for each region, irrespective of scale. This it can be written as

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

Hence our estimation problem can be formulated as

$$\begin{aligned}\arg \min_{\sigma_n^2, \{\kappa(x^i)\}_{i=1}^S} & \left\{ \sum_{i=1}^S \alpha_i \sum_{j=1}^M \left( \sqrt{\kappa(y_j^i)} - \sqrt{\kappa(x^i)} + \frac{\sigma_n^2}{\sigma_{y_j^i}^2} \sqrt{\kappa(x^i)} \right)^2 \right. \\ & \left. - \lambda \sum_{i=1}^S \sum_{j=1}^S \left( \sqrt{\kappa(x^k)} - \sqrt{\kappa(x^l)} \right)^2 \right\}\end{aligned}\quad (3)$$

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

The weighing factor  $\alpha_i$  is specified as

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

Weighing factor is used to assign larger weight to regions with larger kurtosis values. This is because these regions generally contain more meaningful information. The regularization parameter  $\lambda$  is used to control the relative importance of regularization term compared to the error term, and is empirically set as 0.01 in experiments.

The problem is non convex but it can be broken down into two convex sub-problems by fixing one variable and optimizing the other. This gives an iterative alternating minimization procedure. First kurtosis values are updated by solving

$$\arg \min_{\{\kappa(x^i)\}_{i=1}^S} \left\{ \sum_{i=1}^S \alpha_i \sum_{j=1}^M \left( \sqrt{\kappa(y_j^i)} - \sqrt{\kappa(x^i)} + \frac{\sigma_n^2}{\sigma_{y_j^i}^2} \sqrt{\kappa(x^i)} \right)^2 - \lambda \sum_{i=1}^S \sum_{j=1}^S \left( \sqrt{\kappa(x^k)} - \sqrt{\kappa(x^l)} \right)^2 \right\} \quad (4)$$

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

Further the estimate of noise variance is updated by solving the sub-problem

$$\arg \min_{\sigma_n^2} \sum_{i=1}^S \alpha_i \sum_{j=1}^M \left( \sqrt{\kappa(y_j^i)} - \sqrt{\hat{\kappa}(x^i)} + \frac{\sigma_n^2}{\sigma_{y_j^i}^2} \sqrt{\hat{\kappa}(x^i)} \right)^2 \quad (5)$$

First sub-problem is a constrained quadratic programming problem. It is solved by `quadprog` in MATLAB. The second subproblem can be solved to give a closed form solution

$$\hat{\sigma}_n^2 = \frac{\sum_{ij} \alpha_i \left( \sqrt{\hat{\kappa}(x^i)} - \sqrt{\kappa(y_j^i)} \right)}{\sum_{ij} \frac{\alpha_i}{\sigma_{y_j^i}^2} \sqrt{\hat{\kappa}(x^i)}} \quad (6)$$

This alternating update procedure is repeated until convergence of objective function.

## 4 K-means based image partition

An important pre-step to the algorithm is to determine the  $S$  disjoint regions, whose clean versions is assumed to be associated with constant kurtosis. It is done as follows:

- The noisy image is uniformly partitioned in  $p \times p$  non-overlapping patches, with  $p = 16$ .
- Each patch is convolved with a 2D band-pass filter  $\mathbf{b}_k$  selected from  $d \times d$  linear transform base. Here  $k$  is the index of band-pass filters,  $0 \leq k \leq d^2 - 1$ . The response image patch can be given by

$$\mathbf{R}_k = \mathbf{P} \otimes \mathbf{B}_k$$

where  $\mathbf{R}_k$  and  $\mathbf{P}$  denote the response image patch and the noisy image patch, respectively.

- Kurtosis value,  $\kappa(\mathbf{R}_k)$  associated with each response is computed.
- A feature vector for image patch  $\mathbf{P}$  is made:

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

- The feature vector  $\mathbf{f}$  is then used for  $K$ -means algorithm to determine  $S$  regions. The number of clusters,  $S$ , is given as a prior to the algorithm, equal to 3.

## 5 Estimation of non-Gaussian noise

We can relax the restriction on noise being Gaussian. Only requirement is noise to be i.i.d Gaussian in linear transform domain. This will hold, as has been proved in [2], non-Gaussian i.i.d noise is transformed into Gaussian by linear transform domain. Once we estimate the variance in that domain, we can compute the noise parameters in spatial domain. I tested the algorithm for uniform and Laplacian noise.

### 5.1 Uniform Noise

Uniform noise model is usually encountered in quantization noise of sensed image. It follows the uniform distribution:  $Z \sim \mathcal{U}(a, b)$ . For the zero-mean uniform noise we have  $b = -a > 0$ . Thus the variance is given by

$$\sigma_Z^2 = \frac{1}{3} (b)^2$$

The probability density function is parameterized by  $b$  alone, and its estimate can be derived as

$$\hat{b} = \sqrt{3}\hat{\sigma}_Z = \sqrt{3}\hat{\sigma}_n \quad (7)$$

### 5.2 Laplacian Noise

The Laplacian noise is a kind of heavy-tailed noise used to model some impulsive noise. For Laplacian distribution  $Z \sim \mathcal{L}(0, v)$ , its probability density function is defined as

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

The estimate of parameter  $v$  can then be obtained by

$$\hat{v} = \frac{\hat{\sigma}_Z}{\sqrt{2}} = \frac{\hat{\sigma}_n}{\sqrt{2}}$$

## 6 Experiments

I tested the algorithm on `lena_gray` image. Zero-mean i.i.d Gaussian noise with different variances was added to create noisy images. For estimation, firstly mean is subtracted from the noisy image to remove DC component. Then as described in section 4,  $S$  regions are obtained and patches belonging to each region are clustered together. I have chosen the PCA basis, same as authors for linear transformation of domain. Then for each cluster, kurtosis and variance is computed, which are fed as inputs to alternating convex minimization routine along with the weighing factors.

First sub-problem is solved using `quadprog` with interior point method. Second sub-problem is computed directly using the closed form expression (6). Estimation performance was evaluated using mean squared error (MSE) as metric. Estimated noise level is tabulated in 1. The noise added is image of range  $[0,255]$ .

Noise level	Gaussian	Uniform	Laplacian
<b>5</b>	4.77	4.57	4.98
<b>10</b>	9.81	9.69	9.79
<b>15</b>	14.87	14.81	14.76
<b>20</b>	19.98	20.18	19.41
<b>25</b>	25.28	25.31	24.30
<b>30</b>	30.08	30.14	28.69

Table 1: Results of estimated noise-levels

	Gaussian	Uniform	Laplacian
<b>MSE</b>	0.032	0.077	0.442

Table 2: Mean Squared error for each noise distribution

## 7 Conclusion and discussion

The assumption made by paper, that kurtosis is scale-invariant in band-pass domains and piecewise stationary in spatial domain holds true for natural images. K-means based approach adaptively partitions the image into a series of disjoint regions, each with same kurtosis. The problem of estimating noise level is efficiently cast into convex sub-problems, which are solvable iteratively. The key feature of paper is ability to handle a variety of noise types while estimating them reliably.

## References

- [1] L. Dong, J. Zhou, and Y. Y. Tang. Noise level estimation for natural images based on scale-invariant kurtosis and piecewise stationarity. *IEEE Transactions on Image Processing*, 26(2):1017–1030, Feb 2017.
- [2] W. Feller. An introduction to probability theory and its applications. *Hoboken, NJ, USA: Wiley*, 3rd ed., vol. 1, Jan. 1968.