

# Estimation of the noise level function for color images using mathematical morphology and non-parametric statistics

Baptiste Esteban, Guillaume Tochon, Edwin Carlinet and Didier Verna  
EPITA Research and Development Laboratory (LRDE)  
Le Kremlin-Bicêtre, France  
Email: firstname.lastname@lrde.epita.fr

**Abstract**—Noise level information is crucial for many image processing tasks, such as image denoising. To estimate it, it is necessary to find homogeneous areas within the image which contain only noise. Rank-based methods have proven to be efficient to achieve such a task. In the past, we proposed a method to estimate the noise level function (NLF) of grayscale images using the tree of shapes (ToS). This method, relying on the connected components extracted from the ToS computed on the noisy image, had the advantage of being adapted to the image content, which is not the case when using square blocks, but is still restricted to grayscale images. In this paper, we extend our ToS-based method to color images. Unlike grayscale images, the pixel values in multivariate images do not have a natural order relationship, which is a well-known issue when working with mathematical morphology and rank statistics. We propose to use the multivariate ToS to retrieve homogeneous regions. We derive an order relationship for the multivariate pixel values thanks to a complete lattice learning strategy and use it to compute the rank statistics. The obtained multivariate NLF is composed of one NLF per channel. The performance of the proposed method is compared with the one obtained using square blocks, and validates the soundness of the multivariate ToS structure for this task.

## I. INTRODUCTION

During acquisition, natural images may be corrupted by some noise due to physical phenomena. Knowledge of the parameters of this noise related to the image pixels intensity, called noise level, is important for image processing applications such as denoising [1], segmentation [2], super-resolution [3] or image compression [4] and can greatly improve the performance of these algorithms. Furthermore, knowledge of the statistical nature of the noise is also important as it changes the way these applications are processed.

In the past, several methods to estimate the noise level have been proposed. Some of them separate the noise from the image content, either in the case of a signal-independent noise [5] or a signal-dependent noise such as the Poisson-Gaussian noise [6], using a variance stabilization method. With the increasing usage of deep learning based methods for image processing algorithms, some noise level estimations relying on convolutional neural networks were recently proposed, usually in the context of noise removal, for additive Gaussian noise either independent and identically distributed [7], [8] or not [9], [10]. Even though these methods are very effective, they still require knowledge of the statistical nature of the

noise beforehand. Beaurepaire *et al.* [11] propose to first identify the nature of the noise and then estimate its level by analyzing statistics on patches of image, but it did not allow to represent a mixed noise. Finally, Sutour *et al.* [12] propose a method to estimate a noise level function (NLF) [13], represented as a quadratic function, from statistics computed from homogeneous square patches detected by a non-parametric statistical test, which is extended by Esteban *et al.* [14] to homogeneous regions to fit the image content. However, this method is still limited to univariate images.

In this paper, we propose to extend our previous method to color images, solving the issues raised by the multivariate pixel values such as the lack of natural order relationship using a rank map resulting from a learned complete lattice [15] and the usage of the multivariate tree of shapes [16] for the image simplification. Furthermore, we improve the detection of homogeneous regions, using the adjacency relationship as in [12]. Then, we define a noise level function for color images constrained to the case where the noise is not correlated between the channels. Finally, we evaluate the performance of our proposed method on a benchmark of 150 natural color images.

Thus, this paper is structured as follow: first, we recall the estimation of the NLF for univariate images in Section II. Then, its extension to color images is described in Section III. We measure the performance of this extension by comparing it with different methods in Section IV. Finally, we conclude in Section V.

## II. ESTIMATION OF THE NLF FOR UNIVARIATE IMAGES

This section recalls the main steps of [14] for the region-based NLF estimation using the tree of shapes. The noise level of an image represents the variance of the noise related to the pixel intensities of this image. It may be represented as a function, the noise level function (NLF) [13]. This function is modeled as a quadratic function by Sutour *et al.* [12], where each coefficient corresponds to the variance of one type of statistical noise.

Let  $\theta = (a, b, c) \in (\mathbb{R}^+)^3$ . The noise level function of an image is defined as

$$\sigma^2 = \text{NLF}_\theta(x) = ax^2 + bx + c \quad (1)$$

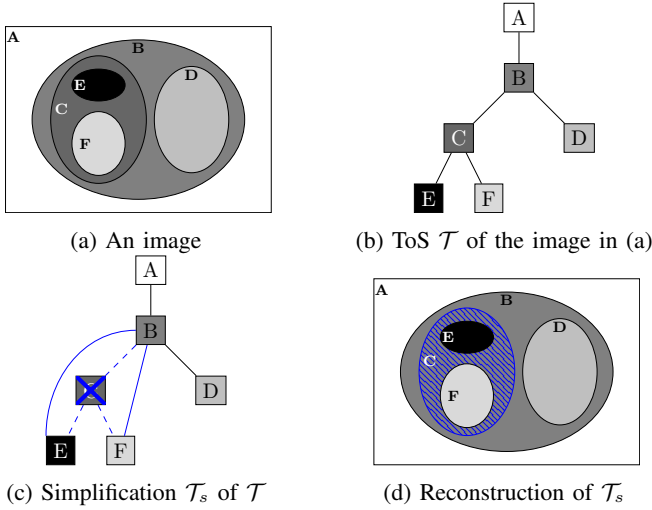


Fig. 1: Illustration of the simplification of the ToS

where  $\sigma^2$  is the variance of the noise for a pixel with an intensity  $x$ . The coefficient  $a$  is related to a multiplicative noise,  $b$  to a Poisson noise and  $c$  to an additive gaussian noise. Thanks to this representation, it is possible to represent a single noise (for an additive gaussian noise,  $(a, b) = (0, 0)$  and  $c > 0$ ) or a mixed one.

#### A. The tree of shapes

The tree of shapes (ToS) [17], denoted by  $\mathcal{T}$ , is a hierarchical representation encoding the inclusion relationship of the connected components of an image. Let  $f : \Omega \rightarrow \mathbb{R}$  be an image defined on a domain  $\Omega \subset \mathbb{Z}^2$ . The upper level set is defined as  $[f \geq \lambda] = \{x \in \Omega | f(x) \geq \lambda\}$  and the lower level set as  $[f < \lambda] = \{x \in \Omega | f(x) < \lambda\}$ . A shape  $\mathcal{C} \subset \Omega$  is a connected component belonging to one of these sets with its holes filled. Fig. 1b illustrates the ToS of the image Fig. 1a.

Furthermore, the contours of the shapes, denoted by  $\partial\mathcal{C}$ , represent the level lines in the image. Thus, the ToS also encodes the inclusion relationship of the level lines of an image. Finally, the ToS is computed with a linear complexity [18].

#### B. Simplification of the ToS

The meaningfulness of a shape can be its importance in the description of the image. A meaningfulness measure is the average gradient magnitude along the contour of a shape, which represents a level line of the image. Thus, removing meaningless shapes from the ToS is a simplification process.

A mean to simplify the ToS rely on the piecewise-constant Mumford-Shah functional [19]. It is defined as below:

$$E(R) = \sum_{R_i \in R} \sum_{x \in R_i} \|\tilde{f}(R_i) - f(x)\|_2^2 + \lambda |\partial R| \quad (2)$$

with  $R = \sqcup R_i$  for  $R_i$  a connected component in the domain of  $f$ ,  $\tilde{f}(R_i)$  is the mean of the values inside a region  $R_i$ ,  $\partial R = \cup \partial R_i$  is the union of the boundaries of each region  $R_i$  and  $\lambda$  is a regularization parameter.

When the minimization of this functional is subordinated to the ToS, the process of simplification consists in selecting a subset of level lines. Thus, the regions boundaries are then fixed to the level lines  $\partial\mathcal{C}_i$ . Formally, the minimization process is given by:

$$\min_{\mathcal{T}_s} E(\mathcal{T}_s) \quad (3)$$

where  $\mathcal{T}_s$  is the simplified ToS. This simplification process is illustrated Fig. 1c, where the blue dashed lines are related to the parenthesis relationship before the simplification and the plain blue line to this relationship after. The resulting reconstructed image Fig. 1d illustrates the impact of the removal of the node C in  $\mathcal{T}$ . An algorithm has been proposed by Xu *et al.* [20] to simplify the ToS, sorting the level lines in a meaningful order and iteratively removing the shapes to decrease the energy of the functional.

#### C. Detection of homogeneous regions

Once the ToS has been simplified, the connected components whose size is above a given threshold are extracted and their homogeneity is tested. A region is homogeneous if it contains only noise, meaning that pixel values are not correlated between them. This correlation is measured using the Kendall  $\tau$  rank correlation coefficient [21].

Let  $(x, y) \in (\mathbb{R}^n)^2$  be two sequences of observations of size  $n$  of two random variables  $X$  and  $Y$  and  $x_i \in x$  is the  $i^{\text{th}}$  value of  $x$  and  $y_i \in y$  be the  $i^{\text{th}}$  value of  $y$ . The Kendall  $\tau$  rank correlation coefficient is defined on the interval  $[-1, 1]$  by:

$$\tau(x, y) = \frac{1}{n-1} \sum_{1 \leq i, j \leq n} \text{sign}(x_i - x_j) \text{sign}(y_i - y_j) \quad (4)$$

for  $x_i \neq x_j$  and  $y_i \neq y_j$ . Furthermore, the definition of the Kendall  $\tau$  has been extended to take into account tied pairs ( $x_i = x_j$  or  $y_i = y_j$ ) [22].

The detection of homogeneous regions relies on a statistical test associated with the Kendall  $\tau$ . Under the null hypothesis  $H_0$ , stating the absence of correlation between the two sequences  $x$  and  $y$  ( $\tau(x, y) = 0$ ), the  $z$ -score associated to the Kendall  $\tau$  follows a standard normal distribution  $\mathcal{N}(0, 1)$ . In contrast with  $H_0$ , the alternative hypothesis  $H_1$  is declared when  $\tau(x, y) \neq 0$ , indicating the fluctuation of the values cannot be explained by the noise only.

To perform the detection, a region is divided randomly into two sequences  $x$  and  $y$  of  $n$  pixels several times. The  $p$ -value  $p = \mathbb{P}(\tau(X, Y) > \tau(x, y) | H_0)$  is computed for all divisions and if the second smallest  $p$ -value is superior to a predefined level of detection  $\alpha$ , the region is considered homogeneous. In practice, each region is divided randomly 10 times.

#### D. Estimation of the NLF

The NLF estimation is a 3 steps process: the segmentation of the image, the statistical test on the connected components extracted from these segmentations and the estimation of the NLF using statistics computed from the homogeneous regions.

The segmentation process is as follow: first, the image is blurred with a gaussian kernel, reducing the amount of level lines in the image. Then, a ToS is built on this blurred image. This ToS is simplified using the Mumford-Shah functional with various values of the regularization parameter  $\lambda$ . This allows to obtain image segmentations of various simplification scales, with regions of different size. The number of removed shapes in the tree increases with the value of the regularization parameter  $\lambda$ , resulting in a more simplified reconstructed image.

Then, the homogeneity of the regions extracted from the simplified ToS is tested using the Kendall  $\tau$ , but using the pixel values of the noisy image to conduct the rank-based test. The first and second order statistics, denoted respectively by  $\hat{\mu}$  and  $\hat{\sigma}^2$  of those homogeneous regions are computed, leading to information about the noise. From Eq. 1, the NLF is the variance of the noise related to pixel intensity. Thus, the NLF is estimated by minimizing the residual error between the empirical variance of a region  $\hat{\sigma}_i^2$  and the variance predicted by the relation  $\text{NLF}_\theta(\hat{\mu}_i)$  for all homogeneous region:

$$\hat{\theta} = (\hat{a}, \hat{b}, \hat{c}) = \underset{\theta \in (\mathbb{R}^+)^3}{\text{argmin}} \sum_i \|\text{NLF}_\theta(\hat{\mu}_i) - \hat{\sigma}_i^2\|_1 \quad (5)$$

The minimization is performed using the preconditionned primal-dual algorithm of Chambolle-Pock [23], relying on a  $L_1$  minimization since it is more robust to outliers than a  $L_2$  minimization.

Comparative results between the block-based estimation from Sutour *et al.* and the estimation described in this section were presented in [14] and demonstrated the pertinence in using regions adapted to the image content instead of square blocks.

### III. EXTENSION OF THE ESTIMATION TO COLOR IMAGES

Color images suffer from the multivariate nature of their values. Rank-based statistics and mathematical morphology are two fields strongly relying on the rank of the considered values. However, multivariate values do not have a natural ordering, which is a well known and studied issue in mathematical morphology [24]. In order to make it possible to estimate the NLF for color images, we use tools from multivariate mathematical morphology: first, the ToS has been extended to multivariate images, resulting in the multivariate ToS [16], its construction depending on the inclusion relationship of the shapes obtained from the channelwise ToSs. Then, we propose to construct a total ordering on the pixel color values using a complete lattice learning approach [15], which allows us to extend the computation of the Kendall  $\tau$  rank correlation coefficient.

#### A. Rank transform based on complete lattice learning

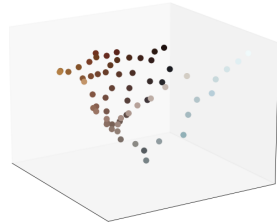
Let  $f : \Omega \rightarrow I \subset \mathbb{R}^n$  be a multivariate image defined on a domain  $\Omega$  and  $I$  its pixel values set containing  $m$  vectors of  $n$  element in  $\mathbb{R}$ . A rank transform  $\mathcal{R} : I \rightarrow [1, m]$  is a mapping from an element of  $I$  to its rank according to the order relationship endowing  $I$ . This operation requires  $I$  to



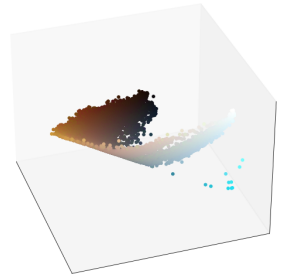
(a) An image  $f : \Omega \rightarrow I \subset \mathbb{R}^3$



(b) Rank map  $\mathcal{R}_f$



(c) Projection  $h_D$  for the values in the dictionary  $D$



(d) Projection  $h$  for the values of  $f$

Fig. 2: Illustration of the complete lattice learning to obtain the rank map  $\mathcal{R}_f$

be a complete lattice, denoted by  $(I, \leq)$ , where  $\leq$  is the total ordering relationship between two elements of  $I$ . This ordering does not exist naturally for multivariate values.

Lezoray [15] proposes to solve this issue by using an unsupervised  $h$ -ordering. A  $h$ -ordering, denoted by  $\leq_h$ , is the application of a bijective projection  $h : I \rightarrow L$  such that, for any two elements  $u, v \in I, u \leq_h v \Leftrightarrow h(u) \leq h(v)$ . Thus,  $I$  does not need to be a complete lattice anymore, since the ordering relationship depends on the elements of  $L$ .

The construction of the projection  $h$  is a three steps process. First, a dictionary  $D$  of  $p \ll m$  elements is extracted out of the whole amount of data using a vector quantization algorithm (such as the  $k$ -means clustering algorithm for instance). This quantization is necessary due to the complexity of the manifold learning algorithm, which is quadratic with the number of input samples.

The new representation is constructed through dimensionality reduction thanks to the Laplacian Eigenmaps [25]. To this aim, the pairwise similarity between each element of the dictionary is computed using a Gaussian kernel, resulting in a similarity matrix  $S_D$ . The similarity matrix can be seen as the adjacency matrix of a graph, where each node is an element of  $D$ . The diagonal degree matrix, denoted by  $D_D$ , is the sum of the weights of the edges connected to each node such that  $(D_D)_{ii} = \sum_j (S_D)_{ij}$ . Using these matrices, the Laplacian matrix is obtained computing  $L_D = D_D - S_D$ . The new representation is finally obtained by computing the eigenvectors of the normalized Laplacian matrix defined as follows:

$$\overline{L}_D = D_D^{-\frac{1}{2}} L_D D_D^{-\frac{1}{2}} \quad (6)$$

The obtained representation, denoted by  $h_D$ , is defined by  $h_D(c_i) = (\phi_D^1(c_i), \dots, \phi_D^p(c_i))$  where  $c_i$  is the  $i^{\text{th}}$  element

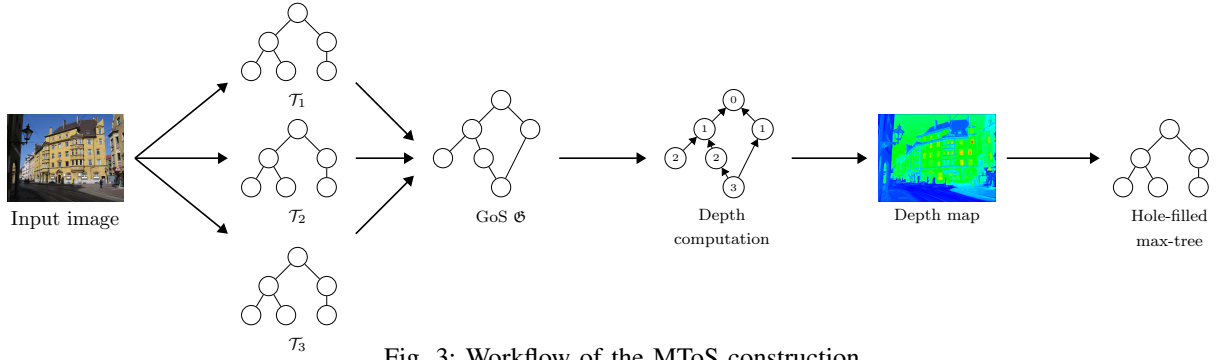


Fig. 3: Workflow of the MToS construction

of  $D$  and  $\phi_D^j$  is the  $j^{\text{th}}$  eigenvector of  $\overline{L}_D$ , with  $\phi_D^j(c_i)$  its value at index  $i$ . However, it is only valid for the dictionary values resulting from the vector quantization. Consequently, it needs to be extended to all image values. This is done using the Nyström extension [26]. For  $\Phi_D$  the matrix of the eigenvectors,  $\Lambda_D$  the diagonal matrix of eigenvalues obtained by the eigen-decomposition of the normalized Laplacian matrix and  $D_I$  is the diagonal degree matrix computed from  $I$ , the Nyström extension is defined by:

$$\Phi_I = D_I^{-\frac{1}{2}} S_D^T D_D^{-\frac{1}{2}} \Phi_D \Lambda_D^{-1} \quad (7)$$

The projection  $h$  is defined as  $h(v) = (\phi_I^1(v), \dots, \phi_I^p(v))$  where  $v \in I$  and  $\phi_I^i$  is the  $i^{\text{th}}$  extrapolated eigenvector of the normalized Laplacian matrix for all the values in  $I$ . Finally, the complete lattice  $(I, \leq)$  is obtained by using the lexicographic ordering on this new representation and the rank map  $\mathcal{R}_f$  resulting from the application of the rank transform on the image  $f$  is computed. The results of each step of the complete lattice learning to obtain the rank map are illustrated in Fig. 2. Instead of computing the Kendall  $\tau$  directly on the image value, this rank map is used.

### B. Multivariate Tree of Shapes

Instead of using the total ordering provided by the rank map, the ToS is extended to multivariate images by using the inclusion relationship between the shapes provided by the channelwise ToSs. This new representation, proposed by Carlinet *et al.* [16], is the multivariate tree of shapes (MToS). Its computation, described below, is composed of two major steps: the computation of a graph of inclusion and the extraction of a new tree from this graph.

First, the ToS  $\mathcal{T}_i$  of each channel  $f_i$  is computed, with  $\mathcal{C}_i$  its associated set of shapes. Then, these trees are merged into a graph of shapes  $\mathcal{G}$ , which is a directed acyclic graph. Its edges represent the inclusion relationship between the connected components of a given channel as for the ToS, but also the inclusion of the shapes between the channelwise ToSs. The root of the trees, representing the whole image, constitutes the root of this graph. Thus, for  $\mathcal{C} = \bigcup_i \mathcal{C}_i$ , the set  $(\mathcal{C}, \subseteq)$  is the cover of  $\mathcal{G}$ .

The second step consists in extracting a tree from  $\mathcal{G}$ . First, the depth attribute is computed. As  $\mathcal{G}$  is a directed acyclic graph, this attribute corresponds to the longest path from a

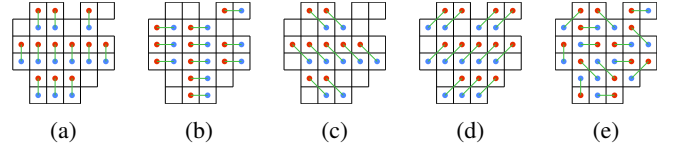


Fig. 4: Region division schemes into two sequences  $x$  and  $y$

node to the root of  $\mathcal{G}$ . Then, a depth map is built, associating to each point of the original image the maximum depth value of the nodes associated to the current point. Finally, a max-tree is built on the depth map, ensuring that the components of the upper level set of the map are valid shapes by applying a hole-filling operator. The resulting tree is MToS. The MToS computation is summarized in Fig. 3.

### C. Estimation of the color noise level function

The NLF defined by (1) is a valid representation for univariate images, but does not allow to represent the noise level for color images. Let  $\text{CNLF}(x) = (\text{NLF}_{\theta_r}(x_r), \text{NLF}_{\theta_g}(x_g), \text{NLF}_{\theta_b}(x_b))$  be the color noise level function (CNLF) where  $x \in \mathbb{R}^3$  is a multivariate pixel value and subscript  $r$ ,  $g$ , and  $b$  correspond respectively to the red, green and blue channels of the color image. This representation assumes that the noise is not correlated between the channels.

The tools described previously are used to extend the estimation of the NLF to color images. First, the MToS is used instead of the ToS. It provides the simplifications to extract the connected components. Then, each connected component is divided into two sequences using the adjacency relationship scheme proposed in Sutour *et al.* [12]. Thus, the division is processed 4 times, using the adjacent relationship as illustrated by Fig. 4a to 4d for the horizontal, vertical and the two diagonal directions, and the Kendall  $\tau$  rank correlation coefficient is computed using the rank map obtained from a rank transform on the original image.

Finally, the empirical mean and variance are computed on each channel for the detected homogeneous regions and the minimization problem expressed by equation (5) is solved channelwise using the Chambolle-Pock algorithm.

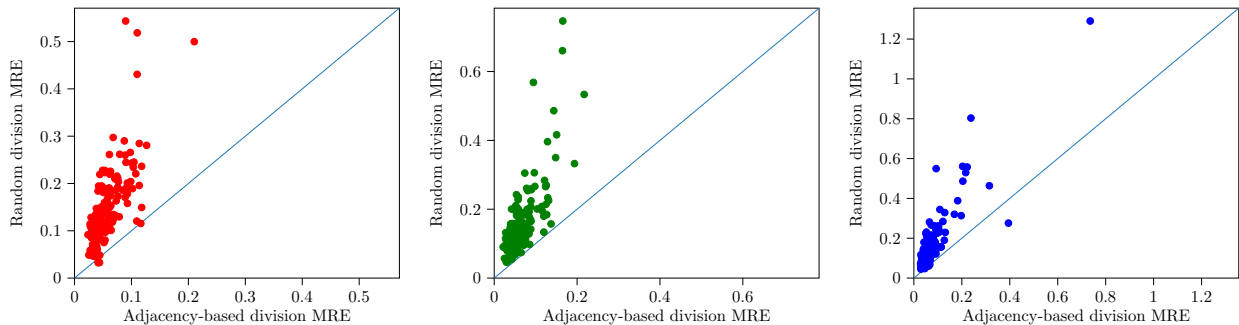


Fig. 5: Comparison of the MRE of the shape-based estimation using a random division and an adjacency-based division For a matter of style, one point has been removed from the plots, respectively at (5.7, 2.2), (7.8, 10.4) and (6.1, 8.0).

## IV. RESULTS

### A. Experimental set-up

Following what is done in [12] and [14], we use the mean relative error (MRE) to compare two CNLF channelwise in the following experiments. Let  $NLF_{\theta}$  be the NLF corresponding to a known set of parameters  $\theta = (a, b, c)$  and  $NLF_{\hat{\theta}}$  be the estimated NLF. The MRE is defined as:

$$\text{MRE}(\hat{\theta}) = \frac{1}{|I|} \sum_{x \in I} \frac{|NLF_{\theta}(x) - NLF_{\hat{\theta}}(x)|}{NLF_{\theta}(x)} \quad (8)$$

where  $I$  is the set of pixel values of the image. The lower the MRE, the better the estimation. The experiments are performed on all 150 images from the Laurent Condat image database<sup>1</sup>, where the acquisition noise in the high definition images can be neglected, making this dataset ideal to experiment on noise estimation tasks.

In the following experiments, the images from the benchmark data set are first polluted by a mixed noise with a given CNLF where the coefficients  $\theta_i = (a_i, b_i, c_i)$  are picked from a standard normal law  $\mathcal{N}(0.01, 0.003)$ . Following what is done for grayscale images in [12], [14], the minimum size of the tested regions is set to 250 and the detection level  $\alpha$  to 0.4. The MToS is simplified with a Mumford-Shah regularization parameter  $\lambda$  empirically set to 200, 500 and 700. To evaluate in practice the robustness of the proposed method, this experiment (adding a mixed noise, estimating the NLF, measuring the MRE) is repeated 20 times per image in the benchmark, for 20 different drawings of the noise parameters  $(a, b, c)$ , and the average MRE for each channel of the image is computed. In order to present concisely and qualitatively the results obtained on the whole benchmark, we use 3 scatterplots (on per channel) where x-axis and y-axis support the average MRE for the two methods being compared. Thus, a point above the plotted blue line corresponds to better results for the method supported by the x-axis in the scatterplot.

### B. Evaluation of the division scheme

In [12], Sutour *et al.* propose to use the adjacency schemes to perform the division of blocks into sequences  $x$  and  $y$ .

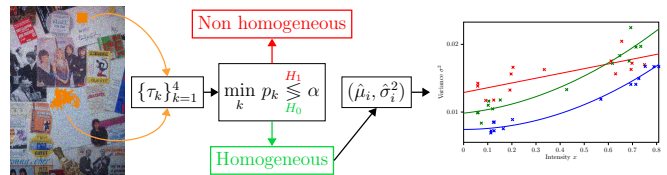


Fig. 6: Summary of the estimation pipeline

Because the regions extracted from the ToS do not have regular shapes allowing these division schemes, we propose in [14] to perform a random splitting instead, illustrated by Fig. 4e. In the current work, we compare the random splitting of [14] with the adjacency-based division of [12], adapted to the irregular morphology of the shapes extracted from the MToS. Even though the random division gave better results than the adjacency-based division of blocks for grayscale images, the adjacency relationship improves the accuracy of the estimation, according to the average channelwise MREs. Thus, for the red, green and blue CNLF components, the average MRE of the adjacency-based division scheme is better for respectively 146, 149 and 147 images.

### C. Comparison with a block-based estimation

In order to be concordant with the experiments performed in [14], we evaluate the accuracy of the proposed CNLF estimation with an extension of the block-based estimation proposed in [12]. The difference between the two compared methods is illustrated in Fig. 6. The block-based estimation divides the image into several square blocks. Each block is then split into two sequences of pixels using an adjacency-based scheme. The statistical test is then performed on these sequences, but instead of using the rank map resulting from a rank transform, the Kendall  $\tau$  rank correlation coefficient is computed on the pixel intensity values of each channel, since it turned out to be more efficient for the block-based estimation. A block is declared homogeneous if all the three channelwise statistical tests are concordant and in favor of homogeneity. In order to attain an overall detection level of  $\alpha$ , this detection level is lowered to  $\alpha/3$  for the three channelwise tests. Furthermore, we extended the automatic

<sup>1</sup><https://lcondat.github.io/imagebase.html>

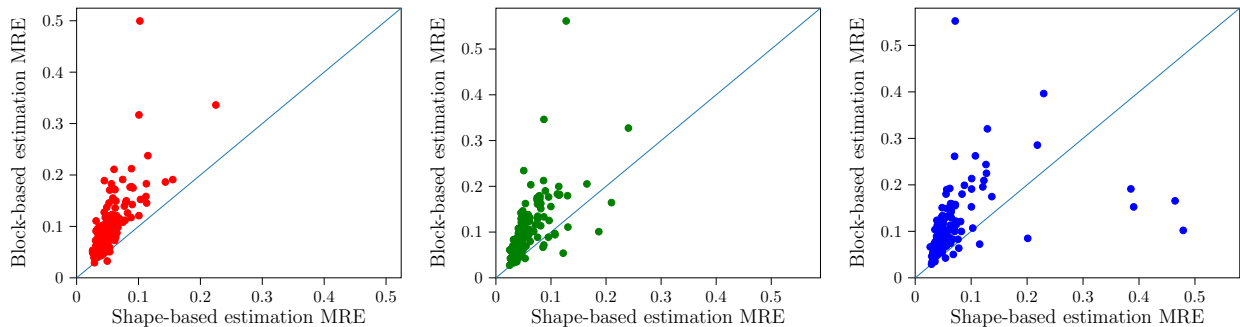


Fig. 7: Comparison of the MRE between the block-based channelwise estimation and the region-based estimation. For a matter of style, one point has been removed from the plots, respectively at (3.9, 0.9), (5.0, 1.0) and (2.1, 1.0).

block size decrease by imposing the same criteria as in [12] for every channel.

Quantitative results between the region-based estimation and the block-based estimation are shown in Fig. 7. For the red, green and blue CNLF components, the average MRE is better for respectively 147, 137 and 139 images. These results confirm that using the MToS to adapt the detection of homogeneous areas to the various content of the 150 benchmark images improves the performance of the estimation compared to the usage of blocks.

Furthermore, the difference between the estimations for one image is illustrated in Fig. 8. The homogeneous regions shown in Fig. 8b are extracted from all simplifications. In the plots of the resulting CNLF in Fig. 8c and 8e, there is less statistical samples (19 samples) used to estimate the channelwise NLF for the region-based estimation than for the block-based estimation (46 samples) but they are more precise and the estimated channelwise NLFs, in plain lines in the plot, are closer to the reference NLFs, in dashed lines, used to add noise to the image. The MREs obtained with the proposed estimation are for each channel respectively 0.036, 0.039 and 0.048, which are better than the block-based estimation's MREs, which are 0.118, 0.067 and 0.115. This confirms that using content-adapted homogeneous areas is more efficient than using blocks, even though the number of statistical samples is fewer. The obtained regions are larger than the fixed size blocks, resulting in more precise statistical means and variances used to estimate the CNLF.

The computational time for the block-based estimation ranges between 1.3 and 102.4 seconds, depending whether the block size was automatically decreased or not. For our proposed method, the computational time is around 66.4 seconds in total, decomposed as 10.4 seconds for the MToS computation and simplifications, 41.8 seconds for the complete lattice learning resulting in the rank map, 14.2 seconds for the homogeneous regions detection and 0.007 seconds for the minimization of Eq. 5.

## V. CONCLUSION

In this article, we proposed to extend our previous method to estimate the NLF with the ToS and rank statistics to color images. Issues were raised by the multivariate nature of color

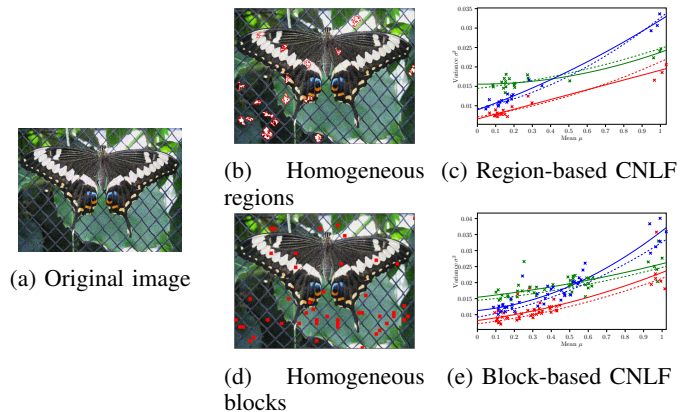


Fig. 8: Illustration of the region-based estimation and the block-based estimation

pixel values. To solve them, we first used the MToS to get the simplifications of an image which minimize the Mumford-Shah functional from which the regions are extracted. The homogeneity of these regions is tested by computing the Kendall  $\tau$  on a rank map resulting from a learned complete lattice. Finally, we have compared the performance of the proposed method with other estimations by comparing their results obtained on a set of 150 benchmark images.

However, this method extracts a partition of the tree which is optimal in the sense of the Mumford-Shah functional, but this partition depends on the value of a regularization parameter which greatly impacts the MToS simplification. Thus, only a subset of the regions contained in the MToS are effectively tested. Consequently, in our future work, we will extend the search for homogeneous areas to all regions contained in the tree structure. Furthermore, we will consider another kind of hierarchical structure being hierarchy of partition, whose construction relies on dissimilarity measures between neighboring regions (and not on the absolute values of their pixels), and which lend themselves better to multivariate data. As the noise is often correlated between channels in real image processing applications, we plan to extend our CNLF model to correlated noise between channel and estimate it using the detected homogeneous regions, these ones handling the values of the three channels.

## REFERENCES

- [1] A. Buades, B. Coll, and J.-M. Morel, "A review of image denoising algorithms, with a new one," *Multiscale modeling & simulation*, vol. 4, no. 2, pp. 490–530, 2005.
- [2] M. Droske and M. Rumpf, "Multiscale joint segmentation and registration of image morphology," *IEEE transactions on pattern analysis and machine intelligence*, vol. 29, no. 12, pp. 2181–2194, 2007.
- [3] W. T. Freeman, T. R. Jones, and E. C. Pasztor, "Example-based super-resolution," *IEEE Computer graphics and Applications*, vol. 22, no. 2, pp. 56–65, 2002.
- [4] J. S. Walker, "Combined image compressor and denoiser based on tree-adapted wavelet shrinkage," *Optical Engineering*, vol. 41, no. 7, pp. 1520–1527, 2002.
- [5] X. Liu, M. Tanaka, and M. Okutomi, "Single-image noise level estimation for blind denoising," *IEEE transactions on image processing*, vol. 22, no. 12, pp. 5226–5237, 2013.
- [6] M. Mäkitalo and A. Foi, "Noise parameter mismatch in variance stabilization, with an application to poisson–gaussian noise estimation," *IEEE Transactions on Image Processing*, vol. 23, no. 12, pp. 5348–5359, 2014.
- [7] K. Zhang, W. Zuo, and L. Zhang, "Ffdnet: Toward a fast and flexible solution for cnn-based image denoising," *IEEE Transactions on Image Processing*, vol. 27, no. 9, pp. 4608–4622, 2018.
- [8] S. Xu, Z. Lin, G. Zhang, T. Liu, and X. Yang, "A fast yet reliable noise level estimation algorithm using shallow cnn-based noise separator and bp network," *Signal, Image and Video Processing*, vol. 14, no. 4, pp. 763–770, 2020.
- [9] Z. Yue, H. Yong, Q. Zhao, D. Meng, and L. Zhang, "Variational denoising network: Toward blind noise modeling and removal," *Advances in neural information processing systems*, vol. 32, 2019.
- [10] H. Tan, H. Xiao, S. Lai, Y. Liu, and M. Zhang, "Pixelwise estimation of signal-dependent image noise using deep residual learning," *Computational intelligence and neuroscience*, vol. 2019, 2019.
- [11] L. Beaurepaire, K. Chehdi, and B. Vozel, "Identification of the nature of local noise and estimation of its statistical parameters by analysis of local histograms," in *1997 IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 4. IEEE, 1997, pp. 2805–2808.
- [12] C. Soutour, C.-A. Deledalle, and J.-F. Aujol, "Estimation of the noise level function based on a nonparametric detection of homogeneous image regions," *SIAM Journal on Imaging Sciences*, vol. 8, no. 4, pp. 2622–2661, 2015.
- [13] C. Liu, W. T. Freeman, R. Szeliski, and S. B. Kang, "Noise estimation from a single image," in *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06)*, vol. 1. IEEE, 2006, pp. 901–908.
- [14] B. Esteban, G. Tochon, and T. Géraud, "Estimating the noise level function with the tree of shapes and non-parametric statistics," in *International Conference on Computer Analysis of Images and Patterns*. Springer, 2019, pp. 377–388.
- [15] O. Lézoray, "Complete lattice learning for multivariate mathematical morphology," *Journal of Visual Communication and Image Representation*, vol. 35, pp. 220–235, 2016.
- [16] E. Carlinet and T. Géraud, "Mtos: A tree of shapes for multivariate images," *IEEE Transactions on Image Processing*, vol. 24, no. 12, pp. 5330–5342, 2015.
- [17] P. Monasse and F. Guichard, "Fast computation of a contrast-invariant image representation," *IEEE Transactions on Image Processing*, vol. 9, no. 5, pp. 860–872, 2000.
- [18] E. Carlinet, S. Crozet, and T. Géraud, "The tree of shapes turned into a max-tree: a simple and efficient linear algorithm," in *2018 25th IEEE International Conference on Image Processing (ICIP)*. IEEE, 2018, pp. 1488–1492.
- [19] D. B. Mumford and J. Shah, "Optimal approximations by piecewise smooth functions and associated variational problems," *Communications on pure and applied mathematics*, 1989.
- [20] Y. Xu, T. Géraud, and L. Najman, "Salient level lines selection using the mumford-shah functional," in *2013 IEEE International Conference on Image Processing*. IEEE, 2013, pp. 1227–1231.
- [21] M. G. Kendall, "A new measure of rank correlation," *Biometrika*, vol. 30, no. 1/2, pp. 81–93, 1938.
- [22] ———, "The treatment of ties in ranking problems," *Biometrika*, vol. 33, no. 3, pp. 239–251, 1945.
- [23] A. Chambolle and T. Pock, "A first-order primal-dual algorithm for convex problems with applications to imaging," *Journal of mathematical imaging and vision*, vol. 40, no. 1, pp. 120–145, 2011.
- [24] E. Aptoula and S. Lefèvre, "A comparative study on multivariate mathematical morphology," *Pattern Recognition*, vol. 40, no. 11, pp. 2914–2929, 2007.
- [25] M. Belkin and P. Niyogi, "Laplacian eigenmaps for dimensionality reduction and data representation," *Neural computation*, vol. 15, no. 6, pp. 1373–1396, 2003.
- [26] A. Talwalkar, S. Kumar, M. Mohri, and H. A. Rowley, "Large scale svd and manifold learning," 2013.