MORPHOLOGICAL SIZE DENSITIES: APPLICATION TO OPTIMAL TEXTURE CLASSIFICATION AND NOISE FILTERING

K. Sivakumar and J. Goutsias

Department of Electrical and Computer Engineering
Image Analysis and Communications Laboratory
The Johns Hopkins University
Baltimore, MD 21218
siva@shape.ece.jhu.edu, goutsias@mycenae.ece.jhu.edu

ABSTRACT

Morphological size densities are frequently used as descriptors of granularity or texture within an image. They have been successfully applied in many image processing and analysis tasks. It is extremely difficult however to analytically calculate the size density. In a previous work, we studied the problem of estimating the (discrete) morphological size density of random images, by means of empirical as well as Monte Carlo estimators. In this paper, we present applications of size density estimators to the problems of *texture classification* and *morphological filtering*. Experimental results demonstrate that texture classification, by means of morphological size densities, produces highly accurate results, even when the texture classes are chosen to be visually similar. For morphological filtering, we have obtained, by means of morphological size densities, two useful classes of filters, namely the class of *generalized alternating filters* and the class of *generalized alternating sequential filters*, which generalize the well known *alternating filters* and *alternating sequential filters*, respectively.

1. INTRODUCTION

The notion of morphological size density was first conceived by Matheron [1]. Morphological size densities provide a wealth of information about image structure and, as such, they have been extensively used in a number of image processing and analysis tasks, such as shape and texture analysis [2], multiscale shape representation [3], morphological shape filtering [4, 5], and the analysis, segmentation, and classification of shape and texture [6]. Given a grayscale random field $\mathbf{X}$ on $\mathbb{Z}^2$, of G gray levels, the function

$$s_X(s) = \frac{1}{G|W|}\begin{cases} \mathrm{E}[\|\gamma_s(\mathbf{X}) - \gamma_{s+1}(\mathbf{X})\|_W], & \text{for } s \geq 0 \\ \mathrm{E}[\|\phi_{|s|}(\mathbf{X}) - \phi_{|s|-1}(\mathbf{X})\|_W], & \text{for } s \leq -1 \end{cases}, \quad (1)$$

where $\mathrm{E}[\ ]$ denotes expectation, $|A|$ denotes the area (cardinality) of a set A, $\|X\|_W = \sum_{w\in W} |X(w)|$ and W is a data observation window, is known as the *morphological size density* of $\mathbf{X}$ (s in (1) is an integer). Mostly, we consider either $\gamma_s(X) = X \ominus sB$, $\phi_s(X) = X \oplus s\check{B}$, where $\ominus$ and $\oplus$ denote morphological *erosion* and *dilation*, respectively, with B being a flat finite structuring element that contains the origin, $\check{B} = \{-b \mid b \in B\}$, and $sB = (s-1)B \oplus B$, for $s \geq 1$, or $\gamma_s(X) = X \bigcirc sB$, $\phi_s(X) = X \bullet s\check{B}$, where $\bigcirc$ and $\bullet$ denote morphological *opening* and *closing*, respectively.

As noted by Matheron, it is extremely difficult to analytically calculate the size density of a given random field. In order to effectively employ size densities in practice, it is imperative to obtain reliable estimates of these quantities. In a previous work [7], we studied the problem of estimating the (discrete) morphological size density of random images, by means of empirical as well as Monte Carlo estimators. The superiority of the Monte Carlo estimation approach was evident from theoretical analysis and experimental observations.

In this paper, we apply size density estimators to two traditional image processing and analysis problems. Our presentation is brief due to space constraints. For further details however, the reader is referred to [8].

2. TEXTURE CLASSIFICATION

Texture classification is a key problem in image analysis. The purpose of this section is to demonstrate the power of morphological size densities, and the associated estimators, for classifying random textures. We assume here that textures are to be classified into K different classes. We follow a model–based approach, with each class k characterized by means of a *Markov random field* (MRF) model $\mathbf{X}_k$. The class conditional probability density function is assumed to be multivariate Gaussian with mean vector m_k and covariance matrix C_k, where $\mathsf{m}_k(i) = \mathrm{E}[\mathsf{v}_k(i)]$ and $\mathsf{C}_k(i,j) = \mathrm{E}[\mathsf{v}_k(i)\,\mathsf{v}_k(j)]$, with

$$\mathsf{v}_k(i) = \begin{cases} \|\gamma_{s_0-i+1}(\mathbf{X}_k) - \gamma_{s_o-i+2}(\mathbf{X}_k)\|_W \\ \|\phi_{|s_o-i+1|}(\mathbf{X}_k) - \phi_{|s_o-i+1|-1}(\mathbf{X}_k)\|_W \end{cases}, \quad (2)$$

for $1 \leq i \leq s_o+1$ and $s_o+2 \leq i \leq 2(s_o+1)$, respectively, and some maximal size $s_o > 0$. If the prior class probability is uniform, the minimum error Bayes classifier is equivalent to assigning a texture sample X with *feature vector* v, given by

This work was supported by the Office of Naval Research, Mathematical, Computer, and Information Sciences Division, under ONR Grant N00014–90–1345.

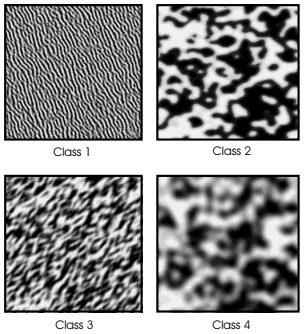
(2) with $\mathbf{X}_k$ being replaced by X, to the class of minimum distance value

$$d_k(\mathsf{v}) = (\mathsf{v} - \mathsf{m}_k)' \, \mathsf{C}_k^{-1} \, (\mathsf{v} - \mathsf{m}_k) + \log(\det(\mathsf{C}_k)), \qquad (3)$$

where $(\,)'$ denotes vector transpose and $\det(\mathsf{A})$ is the determinant of matrix A.

Implementation of the this approach requires a training and a classification step. During training, a MRF model is fit to class k and m_k, C_k are calculated for this model. Model fit may be achieved by means of a MRF statistical inference technique, like Monte Carlo maximum likelihood [9], that determines the energy function $U_k(X)$ of MRF $\mathbf{X}_k$ from texture data representative of class k. After $U_k(X)$ is determined, the Monte Carlo size density estimators, discussed in [8], are employed in order to estimate vector m_k. Matrix C_k is estimated by means of a *Markov chain Monte Carlo* (MCMC) approach similar to the one used for m_k. The training step thus produces a collection $\{(\mathsf{m}_k, \mathsf{C}_k), k = 1, 2, \ldots, K\}$ of mean/covariance pairs that characterize the K classes under consideration. During classification, the feature vector v is calculated from a given image X, by means of the empirical estimators discussed in [8], and X is classified to be in class l if $d_l(\mathsf{v}) = \min\{d_k(\mathsf{v}), k = 1, 2, \ldots, K\}$, with $d_k(\mathsf{v})$ given by (3).

The use of erosion/dilation and opening/closing size densities as features for texture classification was tested in two experiments. Figure 1 depicts the results for grayscale (for $G = 15$) texture classification. Four classes were assumed with each class characterized by a 256×256 pixel auto–binomial model with given parameters [8].



Classification Results
(Erosion/Dilation Size Density)

	Class 1	Class 2	Class 3	Class 4
Class 1	19	0	0	0
Class 2	0	19	0	0
Class 3	0	0	19	0
Class 4	0	0	0	19

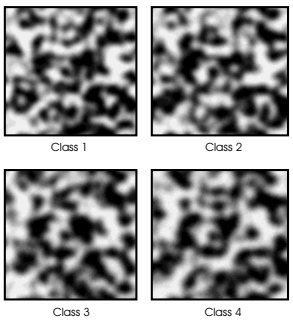
Classification Results
(Opening/Closing Size Density)

	Class 1	Class 2	Class 3	Class 4
Class 1	19	0	0	0
Class 2	0	19	0	0
Class 3	0	0	19	0
Class 4	0	0	0	19

Figure 1: Opening/Closing and Erosion/Dilation size density based texture classification results in 100% classification accuracy.

The four depicted images are representative realizations from each class obtained by means of a MCMC interchange algorithm, where the histogram of the initial image was chosen to be uniform. The feature vector v was obtained with $s_o = 9$ and γ, ϕ being either erosion/dilation or opening/closing with respect to the RHOMBUS structuring element. A total of 4×19 images were employed (19 images per class) for testing classification performance. 4×7 images were obtained from the training step (these images were used in the MCMC estimation of m_k and C_k, for $k = 1, 2, 3, 4$) with the other 4×12 images obtained by means of MCMC independently of the training step. Classification based on the erosion/dilation and opening/closing size densities were exceptional, producing no misclassifications; i.e., 100% accuracy.

The second experiment, depicted in Figure 2, was designed to classify 4×19 grayscale images (with $G = 15$) into four classes as well. In this case however, the classes are characterized by very similar size densities. Class 1 is characterized by the same auto–binomial model as Class 4 in Figure 1. Classes 2, 3, and 4 are characterized by auto–binomial models whose parameters are the ones of Class 1 divided by 2, 3, and 4, respectively. Notice the visual similarity between the four realizations.



Classification Results
(Erosion/Dilation Size Density)

	Class 1	Class 2	Class 3	Class 4
Class 1	19	0	0	0
Class 2	7	12	0	0
Class 3	0	0	19	0
Class 4	0	0	0	19

Classification Results
(Opening/Closing Size Density)

	Class 1	Class 2	Class 3	Class 4
Class 1	14	5	0	0
Class 2	0	19	0	0
Class 3	0	0	19	0
Class 4	0	0	0	19

Figure 2: Erosion/Dilation size density based texture classification resulted in 91% accuracy, whereas Opening/Closing based classification resulted in 93% accuracy.

Classification based on the erosion/dilation size density produced 7 errors, resulting in 91% classification accuracy, whereas classification based on the opening/closing size density produced 5 errors, resulting in 93% accuracy. All errors occurred in classifying Class 2 textures as Class 1 (in the case of the erosion/dilation size density) or Class 1 textures as Class 2 (in the case of the opening/closing size density).

Texture classification by means of size densities produces highly accurate results even in cases when textures are visually indistinguishable. The training step is computationally intensive due to the need of determining the proper MRF model for each class and the extensive MCMC simulations required for obtaining estimates of $(\mathsf{m}_k, \mathsf{C}_k)$, for $k = 1, 2, \ldots, K$. This step however needs to be done once and off–line. The classification step is easy to implement, due to the relatively simple form of the feature vector v. A classification technique based on the erosion/dilation size density should be considered first, since this technique requires substantially less computations than a technique based on opening/closing size densities and may provide acceptable classification results in certain cases.

3. MORPHOLOGICAL FILTERING

In many image processing applications, image data Y are corrupted by noise and clutter. It is therefore of great interest to design an operator Ψ that, when applied on Y, *optimally* removes noise and clutter. Solutions to this problem, known as *image restoration*, have been recently obtained by means of *morphological filtering* [4, 5]. Image data Y are

considered to be realizations of a random field $\mathbf{Y}$, given by means of a *degradation equation* of the form

$$\mathbf{Y} = (\mathbf{X} \wedge \mathbf{N}_1^*) \vee \mathbf{N}_2, \tag{4}$$

where $\wedge$, $\vee$ denote minimum and maximum, respectively, $\mathbf{X}$ is a random field that models the image under consideration, $\mathbf{N}_1$, $\mathbf{N}_2$ are two random fields that model degradation, and $N_1^*(w) = G - N_1(w)$, for every w in an observation window W. In particular, $\mathbf{N}_1$, $\mathbf{N}_2$ may model degradation due to clutter, sensor noise, incomplete data collection, or occlusion. This type of noise is frequently referred to as a *min/max noise* or "two–sided" noise. The problem of optimal image restoration consists of designing a set operator Ψ such that

$$\hat{\mathbf{X}} = \Psi(\mathbf{Y}) = \Psi((\mathbf{X} \wedge \mathbf{N}_1^*) \vee \mathbf{N}_2), \tag{5}$$

is "optimally close" to $\mathbf{X}$, in the sense of minimizing a *distance metric* $D_{\hat{X},X}$. It is this problem that we are considering in this section. Before we proceed however we need the following definition (see also Property 4 in [4]).

Definition 1 *Let X and N be two discrete–valued functions, taking values in $\{0, 1, \ldots, G\}$, defined over a discrete domain W. The* support $\mathrm{supp}(X)$ *of X is defined as* $\mathrm{supp}(X) = \{w \in W \mid X(w) \neq 0\}$. *$X$ and N are said to be* non–interfering, *with respect to a structuring element B, if the supports of X and N are disjoint and $(X \vee N) \bigcirc kB = (X \bigcirc kB) \vee (N \bigcirc kB)$, for every $k = 0, 1, \ldots$.*

Consider now the special case when $\mathbf{N}_1(w) = 0$, for every $w \in W$, *almost surely* (a.s.), and $\mathbf{X}$, $\mathbf{N}_2$ are a.s. non–interfering. By extending the discussion in [5] to the gray-scale case, we consider estimators $\hat{\mathbf{X}}$ of $\mathbf{X}$ of the form

$$\hat{\mathbf{X}} = \Psi_{\bigcirc}(\mathbf{Y}) = \sum_{s \in S_+} [\mathbf{Y} \bigcirc sB - \mathbf{Y} \bigcirc (s+1)B] \tag{6}$$

and use as error criterion the (normalized) *expected absolute difference metric*

$$D_{\hat{X},X} = \frac{1}{G|W|} \mathrm{E}[\, \|\hat{\mathbf{X}} - \mathbf{X}\|_W \,]. \tag{7}$$

The optimal estimator $\hat{\mathbf{X}}$ of $\mathbf{X}$ is then obtained by determining an index set S_+ in (6) that results in minimum expected absolute difference. We now have the following proposition.

Proposition 1 *When $\mathbf{N}_1 = 0$, a.s., and $\mathbf{X}$, $\mathbf{N}_2$ are a.s. non–interfering, the index set $S_+ = \{s \geq 0 \mid s_{N_2}(s) < s_X(s)\}$, where $s_X(s)$ and $s_{N_2}(s)$ are the opening/closing size densities of $\mathbf{X}$ and $\mathbf{N}_2$, respectively, minimizes the expected absolute difference metric $D_{\hat{X},X}$, given by (7), where $\hat{\mathbf{X}}$ is given by (6).*

It is worthwhile noticing here that, if $S_+ = \{s_o, s_o + 1, \ldots\}$, for some $s_o \geq 0$, then $\Psi_{\bigcirc}(\mathbf{Y}) = \mathbf{Y} \bigcirc s_o B$, whereas if $S_+ = \{0, 1, \ldots, s_o - 1\}$, for some $s_o \geq 1$, then $\Psi_{\bigcirc}(\mathbf{Y}) = \mathbf{Y} - \mathbf{Y} \bigcirc s_o B$. Moreover, when $G = 1$ (i.e., in the binary case) operator $\Psi_{\bigcirc}$ reduces to the filter suggested in [5]. Proposition 1 requires only the assumption that $\mathbf{X}$, $\mathbf{N}_2$ are a.s. non–interfering (assuming also that $\mathbf{N}_1 = 0$, a.s.), which is equivalent to assuming that the supports of $\mathbf{X}$ and $\mathbf{N}_2$ are a.s. disjoint and that

$$(\mathbf{X} \vee \mathbf{N}_2) \bigcirc kB = (\mathbf{X} \bigcirc kB) \vee (\mathbf{N}_2 \bigcirc kB), \quad \text{a.s.}, \tag{8}$$

for every $k = 0, 1, \ldots$. This is in sharp contrast to the assumptions made in [5] which are quite restrictive and rather unrealistic.

Assuming the supports of $\mathbf{X}$ and $\mathbf{N}_2$ to be a.s. disjoint is natural in certain applications. For example, consider the case when filtering is employed in order to remove clutter from image data so as to restore an object of interest X. In this case, clutter consists of all objects whose support is disjoint to that of $\mathbf{X}$ (assuming no occlusion) and can be modeled by means of a random field $\mathbf{N}_2$ such that $\mathrm{supp}(\mathbf{X}) \cap \mathrm{supp}(\mathbf{N}_2) = \emptyset$, a.s. Condition (8) is a technical condition, required by the proof of Proposition 1, that limits the potential candidates for $\mathbf{X}$ and $\mathbf{N}_2$ in (4). The non–interfering assumption may therefore be viewed as a regularizing condition necessary to obtain a meaningful solution of the inverse (filtering) problem at hand.

By duality (since $X \wedge N_1^* = (X^* \vee N_1)^*$, in which case $(X^* \oplus B)^* = X \ominus \check{B}$ and $(X^* \bigcirc B)^* = X \bullet \check{B}$) and due to the particular form of (7), for which $D_{\hat{X}^*,X^*} = D_{\hat{X},X}$, we have the following proposition.

Proposition 2 *When $\mathbf{N}_2 = 0$, a.s., and $\mathbf{X}^*$, $\mathbf{N}_1$ are a.s. non–interfering, then*

$$\begin{aligned} \hat{\mathbf{X}} &= [\Psi_{\bigcirc}(\mathbf{Y}^*)]^* = \Psi_{\bullet}(\mathbf{Y}) \\ &= \Big(\sum_{s \in S_-} [\mathbf{Y} \bullet |s| \check{B} - \mathbf{Y} \bullet (|s| - 1) \check{B}] \Big)^*, \end{aligned} \tag{9}$$

minimizes the expected absolute difference metric $D_{\hat{X},X}$, given by (7), provided that $S_- = \{s \leq -1 \mid s_{N_1^*}(s) < s_X(s)\}$, where $s_X(s)$ and $s_{N_1^*}(s)$ are the opening/closing size densities of $\mathbf{X}$ and $\mathbf{N}_1^*$, respectively.

Notice that, if $S_- = \{\ldots, -s_o - 2, -s_o - 1\}$, for some $s_o \geq 0$, then $\Psi_{\bullet}(\mathbf{Y}) = \mathbf{Y} \bullet s_o \check{B}$, whereas if $S_- = \{-s_o, -s_o + 1, \ldots, -1\}$, for some $s_o \geq 1$, then $\Psi_{\bullet}(\mathbf{Y}) = [\mathbf{Y} \bullet s_o \check{B} - \mathbf{Y}]^*$.

We may relax the non–interfering assumptions for $\mathbf{X}$, $\mathbf{N}_2$ and $\mathbf{X}^*$, $\mathbf{N}_1$, and we may consider in (5) an operator $\Psi = \Psi_{\bullet} \Psi_{\bigcirc}$. In this case, $\Psi(\mathbf{Y})$ will be a *suboptimal*, but nevertheless useful, solution to the image restoration problem under consideration. If we assume that $\mathbf{X}$ is a random field with size density $s_X(s)$, and if N_1, N_2 are two realizations of a random field $\mathbf{N}$ with size density $s_N(s)$, then (recall (6) and (9))

$$\hat{\mathbf{X}} = \Psi_{\bullet}(\tilde{\mathbf{X}}) = \Big(\sum_{s \in S_-} [\tilde{\mathbf{X}} \bullet |s| \check{B} - \tilde{\mathbf{X}} \bullet (|s| - 1) \check{B}] \Big)^*, \tag{10}$$

where

$$\tilde{\mathbf{X}} = \Psi_{\bigcirc}(\mathbf{Y}) = \sum_{s \in S_+} [\mathbf{Y} \bigcirc sB - \mathbf{Y} \bigcirc (s+1)B]. \tag{11}$$

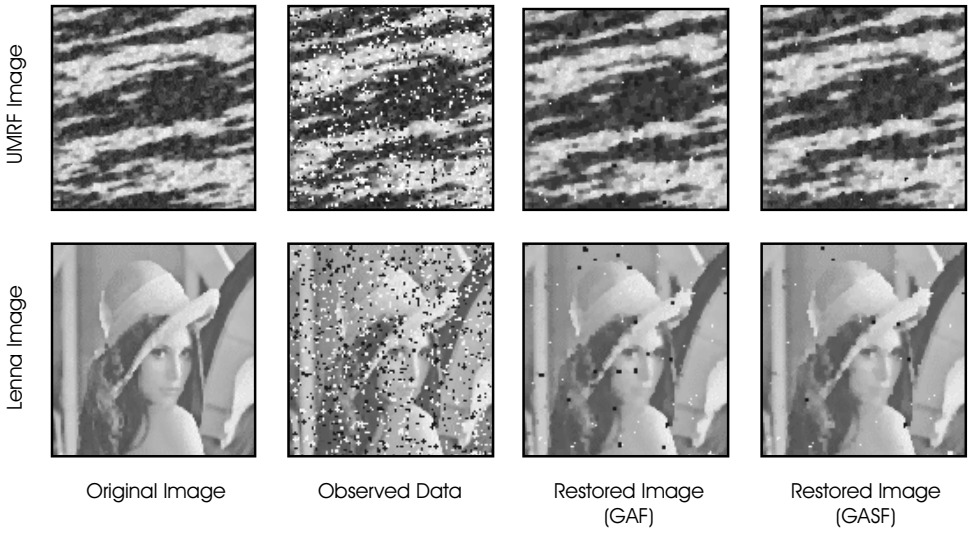


Figure 3: Grayscale image restoration examples for a MRF and the Lenna image corrupted by min/max noise by means of a generalized alternating filter (GAF) and a generalized alternating sequential filter (GASF).

In (10) and (11),

$$S_- = \{s \leq -1 \mid s_{N^*}(s) < s_X(s)\}$$

$$= \{s \leq -1 \mid s_N(|s|-1) < s_X(s)\}, \tag{12}$$

whereas

$$S_+ = \{s \geq 0 \mid s_N(s) < s_{X \wedge N^*}(s)\}. \tag{13}$$

Implementation of operator $\Psi_\bullet \Psi_\circ$, by means of (10)–(13), requires knowledge of the size density $s_N(s)$ of degradation $\mathbf{N}$, for $s \geq 0$, the size density $s_X(s)$ of image $\mathbf{X}$, for $s \leq -1$, as well as the size density $s_{X \wedge N^*}(s)$ of random field $\mathbf{X} \wedge \mathbf{N}^*$, for $s \geq 0$. In the special case when $S_+ = \{s_o, s_o+1, \dots\}$ and $S_- = \{\dots, -s_o-2, -s_o-1\}$, for some $s_o \geq 1$, $\Psi_\bullet \Psi_\circ(\mathbf{Y}) = (\mathbf{Y} \circ s_o B) \bullet s_o \check{B}$ and $\Psi_\bullet \Psi_\circ$ is an *alternating filter* (AF) of order s_o. We refer to operator $\Psi_\bullet \Psi_\circ$ as the *generalized alternating filter* (GAF) of order s_o.

Since *alternating sequential filters* (ASF) usually perform better than AF's (e.g., see [4]), we may modify the implementation of $\Psi_\bullet \Psi_\circ$ so it resembles an ASF. For $k \geq 1$, let us define (see (6) and (9))

$$\Psi_\circ^{(k)}(\mathbf{Y}) = \mathbf{Y} \circ kB \; + \sum_{s \in S_+,\, s \leq k-1} [\mathbf{Y} \circ sB - \mathbf{Y} \circ (s+1)B], \tag{14}$$

and

$$\Psi_\bullet^{(k)}(\mathbf{Y}) = \mathbf{Y} \bullet k\check{B} \; - \sum_{s \in S_-,\, s \geq -k} [\mathbf{Y} \bullet |s|\check{B} - \mathbf{Y} \bullet (|s|-1)\check{B}], \tag{15}$$

where S_- and S_+ are given by (12) and (13), respectively. We may now consider an operator of the form $\Psi_\bullet^{(s_o)} \Psi_\circ^{(s_o)} \cdots \Psi_\bullet^{(2)} \Psi_\circ^{(2)} \Psi_\bullet^{(1)} \Psi_\circ^{(1)}$, where $s_o \geq 1$ is the smallest non–zero integer such that $S_+ \supseteq \{s_o, s_o+1, \dots\}$ and $S_- \supseteq \{\dots, -s_o-2, -s_o-1\}$. We refer to this operator as the *generalized alternating sequential filter* (GASF) of order s_o. It is easy to see that, if $S_+ = \{s_o, s_o+1, \dots\}$, for some $s_o \geq 0$, then

$$\Psi_\circ^{(k)}(\mathbf{Y}) = \begin{cases} \mathbf{Y} \circ kB, & \text{for } 1 \leq k \leq s_o \\ \mathbf{Y} \circ s_o B, & \text{for } k \geq s_o + 1 \end{cases}.$$

Similarly, if $S_- = \{\dots, -s_o-2, -s_o-1\}$, for some $s_o \geq 0$, then

$$\Psi_\bullet^{(k)}(\mathbf{Y}) = \begin{cases} \mathbf{Y} \bullet k\check{B}, & \text{for } 1 \leq k \leq s_o \\ \mathbf{Y} \bullet s_o \check{B}, & \text{for } k \geq s_o + 1 \end{cases}.$$

Therefore, and in this case, $\Psi_\bullet^{(s_o)} \Psi_\circ^{(s_o)} \cdots \Psi_\bullet^{(2)} \Psi_\circ^{(2)} \Psi_\bullet^{(1)} \Psi_\circ^{(1)}$ is an ASF of order s_o.

We now demonstrate the use of GAF (i.e., of operator $\Psi_\bullet \Psi_\circ$) and of GASF (i.e., of operator $\Psi_\bullet^{(s_o)} \Psi_\circ^{(s_o)} \cdots \Psi_\bullet^{(2)} \Psi_\circ^{(2)} \Psi_\bullet^{(1)} \Psi_\circ^{(1)}$), given by (10)–(13) and (14), (15), respectively, for restoring images corrupted by min/max noise. In this case, we take $\mathbf{Y} = (\mathbf{X} \wedge \mathbf{N}_1^*) \vee \mathbf{N}_2$, where $\mathbf{N}_1$ and $\mathbf{N}_2$ have the same distribution as a random field $\mathbf{N}$. The first row of Figure 3 depicts the result of image restoration of a random grayscale image $\mathbf{X}$ that was taken to be a 256×256 pixel 16 graylevel MRF. The corrupting random field $\mathbf{N}$ was taken to be a *Boolean model* [2]. The size densities $s_X(s)$, for $s \leq 0$, and $s_N(s)$, $s_{X \wedge N^*}(s)$, for $s \geq 0$, of $\mathbf{X}$, $\mathbf{N}$, and $\mathbf{X} \wedge \mathbf{N}^*$, respectively, were estimated by means of the Monte Carlo estimator suggested in [8] (see Equations (38) and (40)). The SQUARE structuring element has been employed. The second row of Figure 3 depicts a similar restoration example with $\mathbf{X}$ being a 256×256 pixel 256 graylevel Lenna image. $\mathbf{N}$ was taken to be the same Boolean model as before. The size density $s_X(s)$, $s \leq 0$, was calculated by means of the empirical estimator suggested in [8] (see Equations (26) and (27)), whereas the size densities $s_N(s)$, $s_{X \wedge N^*}(s)$, $s \geq 0$, were calculated by means

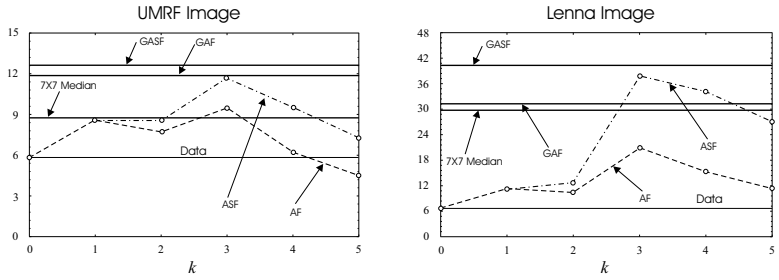


Figure 4: The SNR achieved by restoring X from Y by means of a generalized alternating filter (GAF), a generalized alternating sequential filter (GASF), an alternating filter (AF), an alternating sequential filter (ASF), and a 7×7 median filter for the MRF and Lenna images depicted in Figure 3.

of the Monte Carlo estimator suggested in [8] (see Equations (38) and (40)). The SQUARE structuring element has been employed here as well. The signal–to–noise ratio $\mathrm{SNR} = \frac{\|X\|_W}{\|X-X\|_W}$, achieved by restoring X from Y by means of a GAF, GASF, AF, ASF (the SQUARE structuring element was used in the implementation of both AF and ASF) and a 7×7 median filter (found to be optimal) is depicted in Figure 4. The GASF outperformed all other operators with GAF producing results very close to that of GASF. In the MRF case, the SNR of the observed data was 5.58 (14.93 dB), whereas, application of GASF on Y resulted in a SNR of 12.60 (22.00 dB). In the case of the Lenna image, the SNR of the observed data was 6.60 (16.39 dB), whereas, the SNR of the restored data by means of GASF was 40.56 (32.16 dB). On the other hand, application of GAF on Y resulted in a SNR of 11.86 (21.48 dB) for the MRF image and 31.66 (30.00 dB) for the Lenna image.

In practice, either GAF or GASF should be preferred over AF or ASF (or even median filtering) since application of AF, ASF, or median filtering requires choosing the "optimal size" of these operators which is not known a–priori. The GAF or GASF can be directly applied on data, as long as the associated size densities are known or have been estimated. If the densities are not known a–priori, they can be estimated from training data or from a statistical model, by means of either the empirical estimators or the Monte Carlo estimators discussed in [8].

To conclude, we should point–out that, in the absence of any a–priori information, the problem of estimating the size densities, required for the implementation of the previously suggested filters, from given data is analogous to the problem of calculating the signal and noise power spectra required for determining an "optimal" Wiener filter, when such filters are used for image restoration. Our problem here however is more complicated due to the non–linear form of the degradation equation (5) and is an exciting topic for further research.

4. CONCLUSION

We have applied previously proposed size density estimators to two traditional image processing and analysis problems: texture classification and image restoration. In texture classification, experiments have shown that our approach is highly accurate even when the underlying classes are close to each other. In image restoration, we have extended the results of Haralick et. al. to the grayscale case and have generalized the well known notions of alternating and alternating sequential filters. The design of such filters requires knowledge of certain size densities which are estimated by means of empirical and Monte Carlo estimators.

5. REFERENCES

[1] G. Matheron. *Random Sets and Integral Geometry*. John Wiley, New York City, New York, 1975.

[2] J. Serra. *Image Analysis and Mathematical Morphology*. Academic Press, London, England, 1982.

[3] P. Maragos. Pattern spectrum and multiscale shape representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11:701–716, 1989.

[4] D. Schonfeld and J. Goutsias. Optimal morphological pattern restoration from noisy binary images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13:14–29, 1991.

[5] R. M. Haralick, P. L. Katz, and E. R. Dougherty. Model–based morphology: The opening spectrum. *Graphical Models and Image Processing*, 57:1–12, 1995.

[6] E. R. Dougherty, J. T. Newell, and J. B. Pelz. Morphological texture–based maximum–likelihood pixel classification based on local granulometric moments. *Pattern Recognition*, 25:1181–1198, 1992.

[7] K. Sivakumar and J. Goutsias. Monte Carlo estimation of morphological granulometric discrete size distributions. In J. Serra and P. Soille, editors, *Mathematical Morphology and Its Applications to Image Processing*, pages 233–240. Kluwer, Dordrecht, The Netherlands, 1994.

[8] K. Sivakumar and J. Goutsias. Discrete morphological size distributions and densities: Estimation techniques and applications. *Journal of Electronic Imaging*, 6:31–53, 1997.

[9] C. J. Geyer. On the convergence of Monte Carlo maximum likelihood calculations. *Journal of the Royal Statistical Society, Series B*, 56:261–274, 1994.