

# LULU Operators and Discrete Pulse Transform for Multidimensional Arrays

Roumen Anguelov and Inger Fabris-Rotelli

**Abstract**—The LULU operators for sequences, that is  $L_n, U_n$  and their compositions, are extended to multidimensional arrays in a way which preserves their essential properties, e.g., consistent separation, total variation and shape preservation. The power of the operators is demonstrated by deriving the discrete pulse transform (DPT), which is a hierarchical decomposition of the input array into pulses. Similar to its 1-D counterpart this transform satisfies a basic consistency property.

**Index Terms**—Area closing, area opening, connection, discrete pulse transform (DPT), LULU, separator, total variation.

## I. INTRODUCTION

THE LULU operators and the associated discrete pulse transform (DPT) developed during the last three decades or so are an important contribution to the theory of the nonlinear multiresolution analysis of sequences. The basics of the theory as well as the most significant results until 2005 are published in the monograph [29]. For more recent developments and applications see [2], [8], [14], [26], and [30].

The LULU operators are morphological filters. However, unlike mainstream mathematical morphology, the emphasis here is on what one may call structure preserving properties, like: consistent separation (e.g., noise from signal), total variation and shape preservation, and consistent hierarchical decomposition, which are discussed in the following.

### A. Consistent Separation

The issue of consistency of nonlinear filters is not easy to address in a straightforward manner. In fact, one may note that there is no established approach to this issue, with some authors only providing empirical evidence on the quality of their considered filters. Characterization of the quality of nonlinear filters is discussed at length in [18]. The concept of a smoother introduced there is based upon preserving some linearity, namely, these are operators which are shift, location and scale invariant.

A common requirement for a filter  $P$ , linear or nonlinear, is its idempotence, i.e.,  $P \circ P = P$ . For example, a morphological filter is by definition an increasing and idempotent operator. For linear operators the idempotence of  $P$  implies the idempotence of the complementary operator  $id - P$ , where  $id$  denotes

the identity operator. For nonlinear filters this implication generally does not hold so the idempotence of  $id - P$ , also called co-idempotence [39] can be considered as an essential measure of consistency.

The previously mentioned properties are all discussed in [29] where they are considered to collectively constitute what the author calls a *consistent separation* and are absorbed into the concept of a *separator*. We will give the definition of a separator for operators on real functions defined on a domain with a group structure. Let  $\Omega$  be an abelian group, so that commutativity always holds. Denote by  $\mathcal{A}(\Omega)$  the vector lattice of all real functions defined on  $\Omega$  with respect to the usual point-wise defined addition, scalar multiplication and partial order. Let us recall that

*Definition 1:* A partially ordered set  $L$  is a **lattice** if any  $\ell_1, \ell_2 \in L$  admit a least upper bound  $\ell_1 \vee \ell_2$  and a largest lower bound  $\ell_1 \wedge \ell_2$ . For a **vector lattice** we have that for two sequences  $x = (x_n), y = (y_n)$  that  $x \leq y \iff x_n \leq y_n \forall n \in \mathbb{Z}$ .

For every  $a \in \Omega$  the operator  $E_a : \mathcal{A}(\Omega) \rightarrow \mathcal{A}(\Omega)$  given by  $E_a(f)(x) = f(x - a), x \in \Omega$ , is called a shift operator.

*Definition 2:* An operator  $P : \mathcal{A}(\Omega) \rightarrow \mathcal{A}(\Omega)$  is called a **separator** if

- $P \circ E_a = E_a \circ P, a \in \Omega$  (horizontal shift invariance);
- $P(f + c) = P(f) + c, f, c \in \mathcal{A}(\Omega), c$ -constant function (vertical shift invariance);
- $P(\alpha f) = \alpha P(f), \alpha \in \mathbb{R}^+, f \in \mathcal{A}(\Omega)$  (scale invariance);
- $P \circ P = P$  (idempotence);
- $(id - P) \circ (id - P) = id - P$  (co-idempotence).

### B. Total Variation and Shape Preservation

The total variation  $TV(\cdot)$  is a seminorm on  $\mathcal{A}(\Omega)$ , that is, it satisfies only three of the four axioms of a norm, namely (1)  $TV(f) \geq 0$ , (2)  $TV(\alpha f) = |\alpha|TV(f)$  and (3)  $TV(f + g) \leq TV(f) + TV(g)$ . It does not satisfy the axiom  $TV(f) = 0 \iff f = 0$ . In the practically significant case of 1-D sequences ( $\Omega = \mathbb{Z}$ ), which is the one considered by Rohwer and his collaborators, the total variation is a generally accepted measure for the amount of information present. Note that any separation of sequences may only increase the total variation. More precisely, for any operator  $P : \mathcal{A}(\mathbb{Z}) \rightarrow \mathcal{A}(\mathbb{Z})$  we have

$$TV(f) \leq TV(P(f)) + TV((id - P)(f)). \quad (1)$$

Hence, it is natural to expect that a good separator  $P$  should not create new variation, that is we have

$$TV(f) = TV(P(f)) + TV((id - P)(f)). \quad (2)$$

Manuscript received June 04, 2009; revised October 26, 2009; accepted April 20, 2010. Date of publication May 18, 2010; date of current version October 15, 2010. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Scott T. Acton.

The authors are with the Department of Mathematics and Applied Mathematics, University of Pretoria, Pretoria, South Africa 0083 (e-mail: roumen.anguelov@up.ac.za; inger.fabris-rotelli@up.ac.za).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TIP.2010.2050639

An operator  $P$  satisfying property (2) is called *total variation preserving* [25].

Shape preservation generally refers to the preservation of edges in the input. Total variation preservation is closely linked to the shape preservation properties of the filter. In the case of sequences, the preservation of shape is actually preservation of trend. It is shown in [29] that a fully trend preserving operator on sequences is also total variation preserving.

### C. Consistent Hierarchical Decomposition

Similarly to the consistency of separation, characterizing the quality of hierarchical decompositions by nonlinear filters is also problematic. See [38] for an example of a hierarchical decomposition derived via functional minimization. In [17], a linear hierarchical decomposition via wavelets is investigated. As a side note, the DPT hierarchical decomposition derived here is in fact closely related to that of stack filters.

As mentioned in [26], a decomposition by linear operators should typically recover the coordinates of any given linear combination of basis vectors, a property not at all applicable to the nonlinear case. A measure of the quality of any hierarchical decompositions is introduced in [26] in the following way:

A nonnegative linear combination of the output of the decomposition is decomposed into the same components. (3)

As mentioned, the LULU theory was developed for sequences, that is, the case  $\Omega = \mathbb{Z}$ . Given a bi-infinite sequence  $\xi = (\xi_i)_{i \in \mathbb{Z}}$  and  $n \in \mathbb{N}$  the basic LULU operators  $L_n$  and  $U_n$  are defined as follows:

$$(L_n \xi)_i = \max\{\min\{\xi_{i-n}, \dots, \xi_i\}, \dots, \min\{\xi_i, \dots, \xi_{i+n}\}\} \quad (4)$$

$$(U_n \xi)_i = \min\{\max\{\xi_{i-n}, \dots, \xi_i\}, \dots, \max\{\xi_i, \dots, \xi_{i+n}\}\}. \quad (5)$$

It is shown in [29] that for every  $n \in \mathbb{N}$  the operators  $L_n$  and  $U_n$  as well as their compositions are increasing separators. Hence, they are an appropriate tool for signal extraction. Furthermore they are fully trend preserving in the sense that both the operator and its complement preserve the monotonicity between consecutive terms in the input sequence. As mentioned this implies that these operators are total variation preserving.

Using the LULU operators, a hierarchical decomposition called DPT of a finite sequence  $\xi$  of length  $N$  is derived in the form

$$DPT(\xi) = (D_1(\xi), D_2(\xi), \dots, D_N(\xi)) \quad (6)$$

with the components of (6) being obtained through

$$D_1(\xi) = (id - P_1)(\xi) \quad (7)$$

$$D_n(\xi) = (id - P_n) \circ Q_{n-1}(\xi), \quad n = 2, \dots, N \quad (8)$$

where  $P_n = L_n \circ U_n$  or  $P_n = U_n \circ L_n$  and  $Q_n = P_n \circ \dots \circ P_1$ ,  $n \in \mathbb{N}$ . In (6), each component  $D_n$  is a sum of discrete pulses with pairwise disjoint supports of size  $n$ . Moreover

$$\xi = \sum_{n=1}^N D_n(\xi). \quad (9)$$

The DPT for LULU operators is defined in [29] but its properties are particularly discussed in [26]. It is proved in [26] that the DPT satisfies the consistency property (3).

The operators  $L_n$  and  $U_n$  are built up from rank selectors since they select the minimum and maximum value in the specified window. It is, thus, natural to compare them to the median operator which is also a rank selector since it selects the middle element in the ranked window. It is given by

$$M_n(x)_i = \text{median}\{x_{i-n}, \dots, x_i, \dots, x_{i+n}\}. \quad (10)$$

A comprehensive discussion can be found in [24], [27], and [29]. In [23], the total variation of the sequence  $x$  is shown to decrease with each application of the median operator. This is indicative of its smoothing ability. However the median operator is not idempotent. This is the main disadvantage of the median operator, and calls for recursive application which introduces additional computational complexity, see [36] and [10] for work dedicated to solving this complication. The median operator is a smoother however, that is, it satisfies the first three axioms in Definition 2. Thus, although the median operator also preserves edges and constant regions, and removes impulsive noise effectively [23] its lack of idempotence results in complications whereas the LULU operators do not.

The aim of this paper is to generalize the LULU operators to functions on  $\Omega = \mathbb{Z}^d$  in such a way that their essential properties, mentioned previously, are preserved. In Section II, the definitions of the basic operators  $L_n$  and  $U_n$  on  $\mathcal{A}(\mathbb{Z}^d)$  are derived by using a strengthened form of the morphological concept of connection. Then we show that indeed these operators replicate the properties of the LULU operators for sequences. Section III deals with their shape preservation properties while in Section IV we discuss the total variation preservation. The developed theory is applied to deriving the DPT for functions on  $\mathbb{Z}^d$  (Section V) which satisfies the consistency property (3). Illustrative examples for the application of the theory are given in Section VI.

## II. BASIC OPERATORS $L_n$ AND $U_n$

The definition of the operators  $L_n$  and  $U_n$  for sequences involves maxima and minima over sets of consecutive terms, thus, making essential use of the fact that  $\mathbb{Z}$  is totally ordered. Since  $\mathbb{Z}^d$ ,  $d > 1$ , is only partially ordered the concept of ‘‘consecutive’’ does not make sense in this setting. Instead, we use the morphological concept of set connection [34].

*Definition 3:* Let  $B$  be an arbitrary nonempty set. A family  $\mathcal{C}$  of subsets of  $B$  is called a connected class or a **connection** on  $B$  if:

- i)  $\emptyset \in \mathcal{C}$ ;
- ii)  $\{x\} \in \mathcal{C}$  for all  $x \in B$ ;
- iii)  $\{C_i : i \in I\} \subseteq \mathcal{C}, \bigcap_{i \in I} C_i \neq \emptyset \implies \bigcup_{i \in I} C_i \in \mathcal{C}$ .

If a set  $C$  belongs to a connection  $\mathcal{C}$  then  $C$  is called **connected**.

This definition generalizes the topological concept of connectivity (i.e., a set is connected if it cannot be partitioned into two open disjoint sets) to arbitrary sets including discrete sets like  $\mathbb{Z}^d$ . It generalizes the concept of graph connectivity. If the underlying set  $B$  is a graph, then the graph connectivity also defines a connectivity.

It is clear from Definition 3 that a connection on  $\mathbb{Z}^d$  does not necessarily contain sets of every size. For example,  $\{\emptyset\} \cup \{\{x\} : x \in \mathbb{Z}^d\}$  and  $\{\emptyset\} \cup \{\{x\} : x \in \mathbb{Z}^d\} \cup \{\mathbb{Z}^d\}$  are connections on  $\mathbb{Z}^d$  but neither of them contain sets of finite size other than 0 and 1. In the definition of the operators  $L_n$  and  $U_n$ , we need sets of every size. We assume that the set  $\mathbb{Z}^d$  is equipped with a connection  $\mathcal{C}$  which satisfies the following conditions:

- $\mathbb{Z}^d \in \mathcal{C}$  (11)

- For any  $a \in \mathbb{Z}^d$ ,  $E_a(C) \in \mathcal{C}$  whenever  $C \in \mathcal{C}$  so that  $\mathcal{C}$  is translation invariant (12)

- If  $V \subsetneq W$ ,  $V, W \in \mathcal{C}$ , then there exists  $x \in W \setminus V$  such that  $V \cup \{x\} \in \mathcal{C}$ . (13)

The aim of the conditions (11)–(13) is to define a connection which is sufficiently rich in connected sets. This is demonstrated by the following property, which is obtained via iterative application of the property (13):

Let  $V \subsetneq W$ ,  $V, W \in \mathcal{C}$ . For every  $k \in \mathbb{N}$  such that  $\text{card}(V) < k < \text{card}(W)$  there exists  $S \in \mathcal{C}$  such that  $V \subseteq S \subseteq W$  and  $\text{card}(S) = k$ . (14)

As usual,  $\text{card}(V)$  is the number of the elements in the set  $V$ .

*Definition 4:* Given a point  $x \in \mathbb{Z}^d$  and  $n \in \mathbb{N}$  we denote by  $\mathcal{N}_n(x)$  the set of all connected sets of size  $n + 1$  that contain point  $x$ , that is

$$\mathcal{N}_n(x) = \{V \in \mathcal{C} : x \in V, \text{card}(V) = n + 1\}.$$

In addition to (11)–(13) we assume that the connection  $\mathcal{C}$  is such that

$$\text{card}(\mathcal{N}_n(x)) < \infty, \quad \forall n \in \mathbb{N}, \quad \forall x \in \Omega. \quad (15)$$

A connection satisfying the properties (11)–(13) and (15) we shortly call an *Array Connection*. We should note that if a connection on  $\mathbb{Z}^d$  is defined via graph connectivity, where all vertices are of finite degree, then properties (13)–(15) hold automatically. In image analysis ( $d = 2$ ), the connectivity is a graph connectivity defined via a neighbor relation, e.g., 4-connectivity, 8-connectivity. In this case, all properties (11)–(15) hold. However, in order to have maximum generality in this section, and the next two, we adopt the present axiomatic approach. Let us also mention that LULU operators on a continuous domain ( $\Omega = \mathbb{R}$ ) are discussed in [2] and [3]. Now the operators  $L_n$  and  $U_n$  are defined on  $\mathcal{A}(\mathbb{Z}^d)$  as follows.

*Definition 5:* Let  $f \in \mathcal{A}(\mathbb{Z}^d)$  and  $n \in \mathbb{N}$ . Then for  $x \in \mathbb{Z}^d$

$$L_n(f)(x) = \max_{V \in \mathcal{N}_n(x)} \min_{y \in V} f(y) \quad (16)$$

$$U_n(f)(x) = \min_{V \in \mathcal{N}_n(x)} \max_{y \in V} f(y). \quad (17)$$

Let us confirm that Definition 5 generalizes the definition of  $L_n$  and  $U_n$  for sequences. Suppose  $d = 1$  and let  $\mathcal{C}$  be the connection on  $\mathbb{Z}$  generated by the pairs of consecutive numbers.

Then all connected sets on  $\mathbb{Z}$  are sequences of consecutive integers and for any  $i \in \mathbb{Z}$  we have

$$\mathcal{N}_n(i) = \left\{ \{i - n, \dots, i\}, \{i - n + 1, \dots, i + 1\}, \dots, \{i, \dots, i + n\} \right\}.$$

Hence, for an arbitrary sequence  $\xi$  considered as a function on  $\mathbb{Z}$  the formulas (16) and (17) are reduced to (4) and (5), respectively.

The median filter, which has been the standard point of comparison for the LULU operators in one dimension, has been extended into two and more dimensions as well. However, the fact that  $\mathbb{Z}^d$ ,  $d > 1$ , in general, and the more common case of  $\mathbb{Z}^2$  in particular, is only partially ordered, leads to problems as those encountered in the extension of  $L_n$  and  $U_n$  to multiple dimensions (see the paragraph before Definition 3). The freedom that is allowed in the shape of the window in multiple dimensions, has allowed for various definitions to be adopted by authors. For example, in [11] we have

$$M(f)(x_{ij}) = \text{med} \{f(x_{i,j-1}), f(x_{i,j+1}), f(x_{ij}), f(x_{i-1,j}), f(x_{i+1,j})\}.$$

Another alternative is presented in [21]

$$M(f)(x_{ij}) = \text{med} \{f(x_{i+r,j+s}) : (r, s) \in I\} \quad (18)$$

where  $I = \{(r, s) \in \mathbb{Z}^2\} \subset \mathbb{Z}^2$ . The formulation in (18) does not specify the window over which the median is taken, that is the index set  $I$ , though. A possible window is investigated in [13]. Here they use a rectangular window of size  $n \times m$  (where  $m, n$  are odd integers) and the median operation is the median of the gray levels of the picture elements lying in this  $m \times n$  window with the window centered at the element  $x_{ij}$ . We, thus, propose the following formulation for the median operator in  $d$  dimensions, which incorporates the requirement that the window should contain every possible connected set in  $\mathcal{N}_n(x)$ , as is the case in one dimension. Following the approach adopted here for the extension of  $L_n$  and  $U_n$ , the median operator on  $\mathbb{Z}^d$  can be defined as follows:

$$M_n(f)(x) = \text{median} \{f(y) : y \in B_n\}$$

where  $B_n = \cup_V \{V \in \mathcal{N}_n(x)\}$ . It is easy to see that this definition, similarly to the ones stated previously, is equivalent to (10) when  $d = 1$ . Indeed, we have

$$\begin{aligned} B_n &= \bigcup_V \{V \in \mathcal{N}_n(x)\} \\ &= \bigcup \left\{ \{x_{i-n}, \dots, x_i\}, \{x_{i-n+1}, \dots, x_{i+1}\}, \dots, \{x_i, \dots, x_{i+n}\} \right\} \\ &= \{x_{i-n}, \dots, x_i, \dots, x_{i+n}\}. \end{aligned}$$

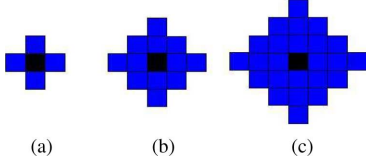


Fig. 1. neighborhoods for the median operator on  $\mathbb{Z}^d$ : (a)  $n = 1$ ; (b)  $n = 2$ ; and (c)  $n = 3$ .

In two dimensions, using 4-connectivity, the neighborhoods for  $n = 1, 2, 3$  are given in Fig. 1.

The operators  $L_n$  and  $U_n$  in Definition 5 as well as their 1-D counterparts (4)–(5) can also be presented in the general setting of Mathematical Morphology. Within this theory  $L_n$  is an area opening and  $U_n$  is an area closing, where the area of a set refers to the number of points in it. Let us recall that a morphological opening (closing) is a composition of an erosion and a dilation (dilation and erosion) with a specified structuring element. Due to the simple structure of  $\mathbb{Z}$  the 1-D  $L_n$  and  $U_n$  are morphological opening and closing, respectively, where the structuring element is a segment of length  $n+1$ . However, the segments are the only connected sets under the considered connectivity on  $\mathbb{Z}$ . In this sense, all connected sets have the same shape. One may consider morphological opening and closing as a generalization of the 1-D  $L_n$  and  $U_n$  to operators on  $\mathcal{A}(\mathbb{Z}^d)$ . However, an essential property of  $L_n$  and  $U_n$  for sequences is that they form a Matheron pair [30], that is

$$L_n \circ U_n \circ L_n = U_n \circ L_n \text{ and } U_n \circ L_n \circ U_n = L_n \circ U_n. \quad (19)$$

It is easy to see by examples that for a general structuring element on  $\mathbb{Z}^d$ ,  $d > 1$ , one can find  $f \in \mathcal{A}(\mathbb{Z}^d)$  such that (19) is violated by the respective morphological opening and closing. This motivates the proposed definition of  $L_n$  and  $U_n$ , when  $d > 1$ , which is independent of shape. One can obtain the proof of (19) from known results on attribute filters of which  $L_n$  and  $U_n$  are particular cases, [9, Property 4]. More precisely,  $L_n$  and  $U_n$  are an area opening and area closing, respectively. However, this property is also an easy consequence from the results that follow as well as those in Section III. A short proof will be given in Section III.

It is well known that the area opening (closing) is an algebraic opening (closing). We may recall that a map is called an algebraic opening (closing) if it is increasing, idempotent and anti-extensive (extensive). Then the following holds:

$$f \leq g \Rightarrow (L_n(f) \leq L_n(g), U_n(f) \leq U_n(g)) \quad (20)$$

$$L_n \circ L_n = L_n, \quad U_n \circ U_n = U_n \quad (21)$$

$$L_n(f) \leq f \leq U_n(f). \quad (22)$$

Furthermore, it is easy to see that these operators are monotone with respect to  $n$  in the following sense:

$$n_1 < n_2 \Rightarrow (L_{n_1} \geq L_{n_2}, U_{n_1} \leq U_{n_2}). \quad (23)$$

Indeed, it follows from (14) that for every  $x \in \mathbb{Z}^d$  and  $V \in \mathcal{N}_{n_2}(x)$  there exists a set  $W \in \mathcal{N}_{n_1}(x)$  such that  $W \subseteq V$ . Therefore

$$\min_{y \in V} f(y) \leq \min_{y \in W} f(y) \leq \max_{S \in \mathcal{N}_{n_1}(x)} \min_{y \in S} f(y) = L_{n_1}(f)(x).$$

Hence

$$L_{n_2}(f)(x) = \max_{V \in \mathcal{N}_{n_2}(x)} \min_{y \in V} f(y) \leq L_{n_1}(f)(x), \quad x \in \mathbb{Z}^d.$$

The inequality for  $U_n$  in (23) is proved in a similar way.

The main consequence of  $L_n$  and  $U_n$  comprising a Matheron pair, as originally shown by Matheron [19], is that  $L_n$ ,  $U_n$ , and all their compositions form a four element semigroup with respect to composition. Moreover, this semigroup is fully ordered as follows:

$$L_n \leq U_n \circ L_n \leq L_n \circ U_n \leq U_n. \quad (24)$$

The semigroup is also a *band* which means that all elements are idempotent.

Next we relate  $U_n$  and  $L_n$  to the concept of a separator given in Definition 2. Indeed, conditions i), ii), and iii) of Definition 2 hold for all openings and closings, [37, Ch. 2]. The idempotence was given in (21). Thus, only the co-idempotence remains. We can remark that co-idempotence is seldom discussed in the standard literature on Mathematical Morphology. However,  $L_n$  and  $U_n$  are also min-max operators as defined by Wild [39]. As such, their co-idempotence follows from [39, Corollary 11]. Therefore

$$L_n, U_n \text{ are separators for every } n \in \mathbb{N}. \quad (25)$$

### III. SHAPE PRESERVATION PROPERTIES OF $L_n$ AND $U_n$

Similar to their counterparts for sequences the operators  $L_n$  and  $U_n$  defined in Section II smooth the input function by removing sharp peaks (the application of  $L_n$ ) and deep pits (the application of  $U_n$ ). The smoothing effect of these operators is made more precise by using the concepts of a local maximum set and a local minimum set defined in the following.

*Definition 6:* Let  $V \in \mathcal{C}$ . A point  $x \notin V$  is called **adjacent** to  $V$  if  $V \cup \{x\} \in \mathcal{C}$ . The set of all points adjacent to  $V$  is denoted by  $\text{adj}(V)$ , that is

$$\text{adj}(V) = \{x \in \mathbb{Z}^d : x \notin V, V \cup \{x\} \in \mathcal{C}\}.$$

An equivalent formulation of the property (13) of the connection  $\mathcal{C}$  is as follows:

$$V, W \in \mathcal{C}, \quad W \subsetneq V \Rightarrow \text{adj}(W) \cap V \neq \emptyset. \quad (26)$$

*Lemma 7:*

- Given  $V, W \in \mathcal{C}$  with  $W \subset V$ . Then for  $x \notin V$  but  $x \in \text{adj}(W)$ , we have  $x \in \text{adj}(V)$ ;
- given distinct  $V, W \in \mathcal{C}$  with  $V \cap W \neq \emptyset$ , there exists an  $x \in V \setminus W$  such that  $x \in \text{adj}(W)$ ;

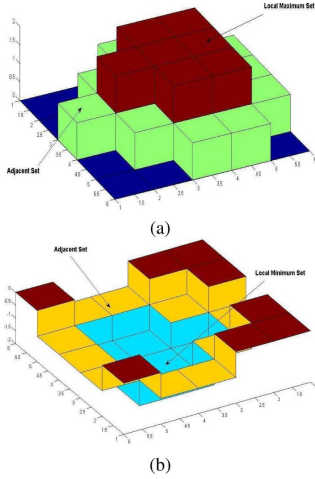


Fig. 2. (a) Local maximum set; (b) local minimum set.

- c) for any  $V \in \mathcal{C}$ , we have  $\text{card}(V) < \infty \Rightarrow \text{card}(\text{adj}(V)) < \infty$ .

*Proof:*

- a)  $V$  and  $W + \{x\}$  are connected and have a nonempty intersection, thus, their union  $V \cup \{x\}$  is also connected. Then by Definition 6,  $x \in \text{adj}(V)$ .
- b) Applying (26) to  $V \cup W$ , since  $V \subsetneq V \cup W$ , we get  $\text{adj}(W) \cap V = \text{adj}(W) \cap (V \cup W) \neq \emptyset$ . Thus, there exists  $x \in V \setminus W$  such that  $x \in \text{adj}(W)$ .
- c) This follows from condition (15). Let  $\text{card}(V) = n$  then for an arbitrary  $x \in V$  we have  $\{\{a\} \cup V : a \in \text{adj}(V)\} \subset \mathcal{N}_{n+1}(x)$ , so that  $\text{card}(\text{adj}(V)) \leq \text{card}(\mathcal{N}_{n+1}(x)) < \infty$  ■

**Definition 8:** A connected subset  $V$  of  $\mathbb{Z}^d$  is called a **local maximum set** of  $f \in \mathcal{A}(\mathbb{Z}^d)$  if  $\sup_{y \in \text{adj}(V)} f(y) < \inf_{x \in V} f(x)$ . Similarly  $V$  is a **local minimum set** if  $\inf_{y \in \text{adj}(V)} f(y) > \sup_{x \in V} f(x)$ , see Fig. 2.

The next four theorems deal with different aspects of the application of  $L_n$  and  $U_n$  to functions in  $\mathcal{A}(\mathbb{Z}^d)$ . They are followed by a discussion on their cumulative effect. All theorems contain statements a) and b). Due to the duality we present only the proofs of a).

**Theorem 9:** Let  $f \in \mathcal{A}(\mathbb{Z}^d)$  and  $x \in \mathbb{Z}^d$ . Then we have

- a)  $L_n(f)(x) < f(x)$  if and only if there exists a local maximum set  $V$  of  $f$  such that  $x \in V$  and  $\text{card}(V) \leq n$ ;
- b)  $U_n(f)(x) > f(x)$  if and only if there exists local minimum set  $V$  of  $f$  such that  $x \in V$  and  $\text{card}(V) \leq n$ .

*Proof:* a) *Implication to the left.* Suppose that there exists a local maximum set  $V \in \mathcal{N}_k(x)$ ,  $k < n$ . Consider an arbitrary  $W \in \mathcal{N}_n(x)$  and let  $S$  be a connected component of  $W \cap V$ . Then  $W \setminus V \neq \emptyset$  since  $\text{card}(S) < \text{card}(W)$  by (14) and by (26) we have  $\text{adj}(S) \cap W \neq \emptyset$ . Let  $z \in \text{adj}(S) \cap W$ . If  $z \in V$  then this means  $z \in V \cup W$ , and so  $S \cup \{z\}$  is connected. This is a contradiction on  $S$  being a connected component of  $V \cap W$ . So  $z \notin V$ . Then using also that  $V$  is a local maximum set we obtain  $\min_{y \in W} f(y) \leq f(z) < \min_{t \in V} f(t) \leq f(x)$ . Since the set  $W \in \mathcal{N}_n(x)$  is arbitrary, this inequality implies that  $L_n(f)(x) < f(x)$ .

*Implication to the right.* Suppose  $L_n(f)(x) < f(x)$ . Let  $V$  be the greatest (in terms of  $\subseteq$ ) connected set containing  $x$  with

$$f(y) \geq f(x) \forall y \in V. \quad (27)$$

The set  $V$  is obviously unique and can be constructed as  $V = \gamma_x(Y)$ , where  $\gamma_x$  is the morphological point connected opening generated by  $x$ , see [33] or [35], and  $Y = \{y \in \mathbb{Z}^d : f(y) \geq f(x)\}$ . Assume that  $\text{card}(V) > n$ . It follows from (14) that there exists  $W \in \mathcal{N}_n(x)$  such that  $W \subset V$ . Then

$$\begin{aligned} L_n(f)(x) &= \max_{S \in \mathcal{N}_n(x)} \min_{y \in S} f(y) \\ &\geq \min_{y \in W} f(y) \\ &\geq \min_{y \in V} f(y) \\ &= f(x). \end{aligned}$$

This contradicts the assumption  $L_n(f)(x) < f(x)$ . Therefore,  $\text{card}(V) \leq n$ . We have  $f(z) < f(x)$ , for all  $z \in \text{adj}(V)$ , because otherwise (27) is satisfied on the larger connected set  $\{z\} \cup V$ . Then, also using Lemma 7c), we obtain  $\max_{z \in \text{adj}(V)} f(z) < f(x) = \min_{y \in V} f(y)$ . Hence,  $V$  is a local maximum set. ■

**Theorem 10:** Let  $f \in \mathcal{A}(\mathbb{Z}^d)$ . Then

- a) the size of any local maximum set of the function  $L_n(f)$  is larger than  $n$ ;
- b) the size of any local minimum set of the function  $U_n(f)$  is larger than  $n$ .

*Proof:* a) Assume the opposite, that is, there exists a local maximum set  $V$  of  $L_n(f)$  such that  $\text{card}(V) \leq n$ . By Theorem 9 we have that  $L_n(L_n(f))(x) < L_n(f)(x)$ ,  $x \in V$ . Since  $L_n$  is idempotent, see (21), this implies the impossible inequality  $L_n(f)(x) < L_n(f)(x)$ , which completes the proof. ■

**Theorem 11:** Let  $V \in \mathcal{C}$  and let  $x \in \text{adj}(V)$ .

- a) If  $f(x) \leq \inf_{y \in V} f(y)$  then  $L_n(f)(x) \leq \inf_{y \in V} L_n(f)(y)$ ;
- b) If  $f(x) \geq \sup_{y \in V} f(y)$  then  $U_n(f)(x) \geq \sup_{y \in V} U_n(f)(y)$ .

*Proof:* a) For any  $W \in \mathcal{N}_n(x)$  the set  $W \cup V$  is connected and of size at least  $n + 1$ . Therefore, by (14), for every  $y \in V$  there exists  $S_y \in \mathcal{N}_n(y)$  such that  $S_y \subset W \cup V$ . Then, using also the given inequality and since  $\inf_{z \in V \cup \{x\}} f(z) = f(x)$ , for every  $y \in V$  and  $W \in \mathcal{N}_n(x)$  we have

$$\min_{z \in W} f(z) = \inf_{z \in W \cup V} f(z) \leq \min_{z \in S_y} f(z) \leq L_n(f)(y).$$

Hence,  $L_n(f)(x) = \max_{W \in \mathcal{N}_n(x)} \min_{z \in W} f(z) \leq \inf_{y \in V} L_n(f)(y)$ . ■

**Theorem 12:** Let  $f \in \mathcal{A}(\mathbb{Z}^d)$  and  $V$  be a finite connected set.

- a) If  $V$  is a local minimum set of  $L_n(f)$  then there exists a local minimum set  $W$  of  $f$  such that  $W \subseteq V$ .
- b) If  $V$  is a local maximum set of  $U_n(f)$  then there exists a local maximum set  $W$  of  $f$  such that  $W \subseteq V$ .

*Proof:*

- a) Let  $V$  be a finite connected set which is a local minimum set of  $L_n(f)$ . Then by Lemma 7c) the set  $\text{adj}(V)$  is finite and we have

$$\min_{y \in \text{adj}(V)} f(y) \geq \min_{y \in \text{adj}(V)} L_n(f)(y) > L_n(f)(x) \quad \forall x \in V.$$

Let  $q \in \text{adj}(V)$  be such that  $f(q) = \min_{y \in \text{adj}(V)} f(y)$  and let  $Y = \{y \in V : f(y) < f(q)\}$ . Note that  $Y \neq \emptyset$ , since if  $Y = \emptyset$ , then  $f(q) \leq \inf_{y \in V} f(y)$  and by Theorem 11 we have  $L_n(f)(q) \leq \inf_{y \in V} L_n(f)(y)$ . This is a contradiction since  $V$  is a local minimum set of  $L_n(f)$ . Nevertheless, this result is an essential ingredient of the proof not least due to the fact that this is the only point where we use that  $V$  is a local minimum set of  $L_n(f)$ . Let  $t \in Y$  and let  $W$  be the largest connected component of  $Y$  containing  $t$  so that  $W = \gamma_t(Y)$  as in the proof of Theorem 9. For every  $z \in \text{adj}(W)$ , due to Lemma 7 we have either  $z \in V$  or  $z \in \text{adj}(V)$ . In both cases, it easy to see that  $f(z) \geq f(q) > \max_{y \in W} f(y)$ . Therefore  $W$  is a local minimum set of  $f$ . ■

Theorems 9–12 provide the following characterization of the effect of the operators  $L_n$  and  $U_n$  on a function  $f \in \mathcal{A}(\mathbb{Z}^d)$ .

- The application of  $L_n(U_n)$  removes local maximum (minimum) sets of size smaller or equal to  $n$ .
- The operator  $L_n(U_n)$  does not affect the local minimum (maximum) sets in the sense that such sets may be affected only as a result of the removal of local maximum (minimum) sets. However, no new local minimum (maximum) sets are created where there were none. This does not exclude the possibility that the action of  $L_n(U_n)$  may enlarge existing local minimum (maximum) sets or join two or more of them into one local minimum (maximum) set of  $L_n(f)(U_n(f))$ .
- $L_n(f) = f(U_n(f) = f)$  if and only if  $f$  does not have local maximum (minimum) sets of size  $n$  or less.

Furthermore, as an immediate consequence of Theorem 10 and Theorem 12 we obtain the following corollary.

*Corollary 13:* For every  $f \in \mathcal{A}(\mathbb{Z}^d)$  the functions  $(L_n \circ U_n)(f)$  and  $(U_n \circ L_n)(f)$  have neither local maximum sets nor local minimum sets of size  $n$  or less. Furthermore

$$(L_n \circ U_n)(f) = (U_n \circ L_n)(f) = f$$

if and only if  $f$  does not have local maximum sets or local minimum sets of size less than or equal to  $n$ .

*Theorem 14:* For  $f \in \mathcal{A}(\mathbb{Z}^d)$

- a)  $L_n(f)$  is constant on any local maximum set  $W$  of  $f$  with  $\text{card}(W) \leq n + 1$ ;
- b)  $U_n(f)$  is constant on any local minimum set  $W$  of  $f$  with  $\text{card}(W) \leq n + 1$ .

*Proof:* We only prove a). Result b) is proven by duality. Let  $W$  be a local maximum set of  $f$  with  $\text{card}(W) \leq n + 1$  and take arbitrary  $p, q \in W$ . Consider  $V \in \mathcal{N}_n(p)$  such that  $V \neq W$ . Then  $V \cup W$  is connected and so by Lemma 7b) there exists  $x \in V$  with  $x \in \text{adj}(W)$ . Since  $W$  is a local maximum set of  $f$  we have  $f(x) \leq \inf_{z \in W} f(z)$  and, hence,  $\inf_{z \in V} f(z) = \inf_{z \in V \cup W} f(z)$ . By (14) there exists  $U \in \mathcal{N}_n(q)$  such that  $U \subset$

$V \cup W$ . Thus,  $\inf_{z \in V} f(z) = \inf_{z \in V \cup W} f(z) \leq \inf_{z \in U} f(z)$ , and since  $U \in \mathcal{N}_n(q)$  we have

$$\inf_{z \in V} f(z) \leq \inf_{z \in U} f(z) \leq L_n(f)(q). \tag{28}$$

We consider  $V \in \mathcal{N}_n(p)$  such that  $V = W$ , we have

$$\inf_{z \in V} f(z) = \inf_{z \in W} f(z) \leq L_n(f)(q)$$

since  $q \in W$  means that  $W \in \mathcal{N}_n(q)$ . Thus, since (28) holds for all  $V \in \mathcal{N}_n(p)$  we have  $L_n(f)(p) \leq L_n(f)(q)$ . By interchanging the role of  $p$  and  $q$  we obtain the other inequality and, thus, have equality. ■

We should remark that in the 1-D setting, the sequences without local maximum sets or local minimum sets of size less than or equal to  $n$  are exactly the so-called  $n$ -monotone sequences [29]. Hence, Corollary 13 generalizes the respective results in the LULU theory of sequences, [29, Th. 3.3].

As mentioned earlier, using Theorem 9 to Theorem 12, it is easy to show that  $L_n$  and  $U_n$  are a Matheron pair, that is, they satisfy (19). Indeed, if we assume that first identity in (19) is violated, in view of the inequality

$$L_n \circ U_n \circ L_n \leq id \circ U_n \circ L_n = U_n \circ L_n$$

we obtain that there exists  $f \in \mathcal{A}(\mathbb{Z}^d)$  and  $z \in \mathbb{Z}^d$  such that

$$L_n(U_n(L_n(f)))(z) < U_n(L_n(f))(z).$$

It follows from Theorem 9 that there exists  $k \leq n$  and  $V \in \mathcal{N}_k(z)$  such that  $V$  is a local maximum set for  $U_n(L_n(f))(z)$ . Then, by Theorem 12, there exists  $W \subseteq V$  such that  $W$  is a local maximum set of the function  $L_n(f)$ . We have  $\text{card}(W) \leq k \leq n$ . However,  $L_n(f)$  does not have any local maximum sets of size less than or equal to  $n$ , see Theorem 10. This contradiction shows the first identity in (19) holds. The second one is obtained by the duality.

The preservation of shape presented in Theorem 9 to Theorem 12 in this section can be made more precise by generalising to  $\mathcal{A}(\mathbb{Z}^d)$  the concepts of neighbor trend preserving and fully trend preserving introduced in [29, Ch. 6] for sequences.

*Definition 15:* An operator  $P$  is **neighbor trend preserving** if for any points  $p, q \in \Omega$ , such that  $\{p, q\} \in \mathcal{C}$ , and for  $f \in \mathcal{A}(\mathbb{Z}^d)$  we have  $f(p) \leq f(q) \implies P(f)(p) \leq P(f)(q)$ . The operator  $P$  is **fully trend preserving** if both  $P$  and  $id - P$  are neighbor trend preserving.

In Definition 15, for  $P$  to be fully trend preserving the requirement on  $id - P$ , that is the neighbor trend preserving property, can be equivalently formulated as

$$|P(f)(p) - P(f)(q)| \leq |f(p) - f(q)|. \tag{29}$$

In the context of sequences the property (29) is called *difference reducing*.

*Theorem 16:* The operators  $L_n, U_n, n = 1, 2, \dots$ , and their compositions, are all fully trend preserving.

*Proof:* We prove the result for  $L_n$ . The case for  $U_n$  follows by duality. Furthermore, it is easy to obtain that compositions

of fully trend preserving operators are fully trend preserving, which proves the rest of the theorem.

Since, the neighbor trend preserving property of  $L_n$  follows directly from Theorem 11, we only need to prove the neighbor trend preserving property of  $id - L_n$  or equivalently, the inequality (29). Consider  $p, q \in \mathbb{Z}^d$  such that  $\{p, q\} \in \mathcal{C}$ . We may assume without loss of generality that  $f(p) \geq f(q)$ . Then  $L_n(f)(p) \geq L_n(f)(q)$  by the neighbor trend preservation. By (22) we have either (i)  $L_n(f)(q) = f(q)$  or (ii)  $L_n(f)(q) < f(q)$ . If (i) holds then,  $L_n(f)(p) - L_n(f)(q) = L_n(f)(p) - f(q) \leq f(p) - f(q)$ , again due to (22), so that  $|L_n(f)(p) - L_n(f)(q)| \leq |f(p) - f(q)|$ . If (ii) holds, by Theorem 9,  $q$  must belong to the support of a local maximum set, say  $W$ , of size at most  $n$  of  $f$ . Since  $f(p) \geq f(q)$  and  $\{p, q\} \in \mathcal{C}$ ,  $q$  must also belong to the support of  $W$ . This means that by Theorem 14 we have  $L_n(f)(p) = L_n(f)(q)$  so that  $|L_n(f)(p) - L_n(f)(q)| = 0 \leq |f(p) - f(q)|$ . So  $L_n$  is difference reducing and, thus, fully trend preserving. ■

The next theorem generalizes the properties of  $L_n$  and  $U_n$  in Theorem 12 to arbitrary neighbor trend preserving operators.

*Theorem 17:* Let  $A : \mathcal{A}(\mathbb{Z}^d) \rightarrow \mathcal{A}(\mathbb{Z}^d)$  be a neighbor trend preserving operator and let  $f \in \mathcal{A}(\mathbb{Z}^d)$ . For every finite local minimum (maximum) set  $V$  of  $A(f)$  there exists a local minimum (maximum) set  $W$  of  $f$  such that  $W \subseteq V$ .

*Proof:* Let  $V$  be a finite local minimum set of  $A(f)$ . Let  $q \in \text{adj}(V)$  be such that  $f(q) = \min_{y \in \text{adj}(V)} f(y)$ . It follows from (13) that there exists  $p \in V$  such that  $\{p, q\}$  is connected. By the local minimality of  $V$  we have

$$A(f)(p) < A(f)(q). \quad (30)$$

Due to the fact that  $A$  is neighbor trend preserving the inequality  $f(p) \geq f(q)$  implies that  $A(f)(p) \geq A(f)(q)$  which contradicts (30). Therefore,  $f(p) < f(q)$ . Let  $Y = \{x \in V : f(x) < f(q)\}$ . Clearly,  $p \in Y$ . Denote by  $W$  the largest connected component of  $Y$  which contains  $p$ . In terms of the notations used in the proof of Theorem 9, we have  $W = \gamma_p(Y)$ . We will show that  $W$  is a local minimum set of  $f$ . Let  $z \in \text{adj}(W)$ . It follows from the construction of  $W$  that  $f(q) > \max_{x \in W} f(x)$ . Hence, it is enough to show that  $f(z) \geq f(q)$ . By the Lemma 7a) we have two possibilities:

i)  $z \in V$ . Then  $z \notin Y$ . Indeed, if  $z \in Y$  then  $W \cup \{z\}$  is a connected component of  $Y$  which contains  $W$ . This is impossible since  $W$  is the largest such component of  $Y$ . Since  $z \notin Y$ , then it violates the defining inequality of  $Y$ , that is we have  $f(z) \geq f(q)$ .

ii)  $z \in \text{adj}(V)$ . Then  $f(z) \geq \min_{y \in \text{adj}(V)} f(y) = f(q)$ .

The proof for local maximum sets is carried out similarly. ■

#### IV. TOTAL VARIATION PRESERVATION

In this section, we assume that the connection  $\mathcal{C}$  on  $\mathbb{Z}^d$  is defined via the so-called graph connectivity. More precisely, the points of  $\mathbb{Z}^d$  are considered as vertices of a graph with edges connecting some of them. Equivalently, the connectivity of such a graph can be defined via a relation  $r \subset \mathbb{Z}^d \times \mathbb{Z}^d$ , where  $p \in \mathbb{Z}^d$  is connected (by an edge) to  $q \in \mathbb{Z}^d \iff (p, q) \in r$ . The relation  $r$  reflects what we consider neighbors of a point in the given context. For example, in image analysis ( $d = 2$ ), it

TABLE I  
STANDARDIZED TOTAL VARIATION OF SOME SAMPLE IMAGES

Image in Figure 3	Total Variation (standardized)
(a)	109173
(b)	132527
(c)	167011
(d)	193650
(e)	235908
(f)	245480
(g)	386408
(h)	703707

is common to use 4-connectivity (neighbors left, right, up and down) and 8-connectivity (in addition, the diagonal neighbors are considered). Let  $r$  be a relation on  $\mathbb{Z}^d$ . We call a set  $C \subseteq \mathbb{Z}^d$  *connected*, with respect to the graph connectivity defined by  $r$ , if for any two points  $p, q \in C$  there exists a set of points  $\{p_1, p_2, \dots, p_k\} \subseteq C$  such that each point is neighbor to the next one,  $p$  is neighbor to  $p_1$  and  $p_k$  is neighbor to  $q$ . Here we assume that:

- $r$  is reflexive, symmetric, shift invariant;
- $(p, p + e_k) \in r$ , for all  $k = 1, 2, \dots, d$  and  $p \in \mathbb{Z}^d$ , where  $e_k \in \mathbb{Z}^d$  is defined by  $(e_k)_i = \begin{cases} 0 & \text{if } i \neq k \\ 1 & \text{if } i = k \end{cases}$ .

These two assumptions shown previously ensure that the set of connected sets  $\mathcal{C}$  defined through this relation is a connection in terms of Definition 3 and satisfies the conditions (11)–(13). The second assumption is essential to the definition of total variation as will be seen in the sequel.

Since the information in an image is in the contrast, the total variation of the luminosity function is an important measure of the quantity of this information. Image recovery and noise removal via total variation minimization are discussed in [6] and [31]. It should be noted that there are several definitions of total variation for functions of multidimensional argument (Arzelá variation, Vitali variation, Pierpont variation, Hardy variation, etc., see [1], [7], [20]). In the applications cited previously, the total variation is the  $L^1$  norm of a vector norm of the gradient of the function. Here we consider a discrete analogue of this concept. Namely, the *Total Variation* of  $f \in \mathcal{A}(\mathbb{Z}^d)$  if given by

$$TV(f) = \sum_{p \in \mathbb{Z}^d} \sum_{i=1}^d |f(p + (e_k)_i) - f(p)|. \quad (31)$$

If  $TV(f) < \infty$ , then  $f$  is said to be of *bounded variation*. Table I gives the total variation of a few sample images seen in Fig. 3. Notice that the pure noise image has the highest total variation and as the images contain more homogenous areas their total variation reduces.

As mentioned in the introduction, the LULU operators for sequences are total variation preserving. We show that their  $d$ -dimensional counterparts considered here have the same property with respect to the total variation as given in (31).

Let us denote by  $BV(\mathbb{Z}^d)$  the set of all functions of bounded variation in  $\mathcal{A}(\mathbb{Z}^d)$ . Clearly, all functions of finite support are in  $BV(\mathbb{Z}^d)$ . For example, the luminosity functions of images are in  $BV(\mathbb{Z}^2)$ . Note that when  $d = 1$  (31) gives the total variation of sequences as discussed in [29, Ch. 6]. Similar to sequences the total variation in (31) is a seminorm so that (1) holds. An

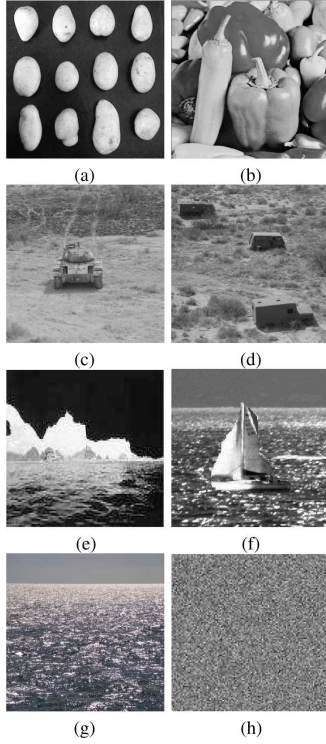


Fig. 3. Sample images.

operator  $P$  on  $BV(\mathbb{Z}^d)$  is called *total variation preserving* if (2) holds.

*Theorem 18:* If an operator  $P : BV(\mathbb{Z}^d) \rightarrow BV(\mathbb{Z}^d)$  is fully trend preserving then it is also total variation preserving.

*Proof:* For  $p, q \in \Omega$  such that  $\{p, q\} \in \mathcal{C}$ , we have

$$|f(p) - f(q)| \leq |P(f)(p) - P(f)(q)| + |f(p) - P(f)(p) - (f(q) - P(f)(q))|. \quad (32)$$

If  $f(p) \geq f(q)$  then  $P(f)(p) - P(f)(q) \geq 0$  since  $P$  is neighbor trend preserving. Then  $|f(p) - P(f)(p) - (f(q) - P(f)(q))| = (f(p) - f(q)) - (P(f)(p) - P(f)(q)) \geq 0$  since  $P$  is difference reducing, see (29). Thus, (32) holds as an equality. Hence

$$\begin{aligned} TV(f) &= \sum_{p \in \mathbb{Z}^d} \sum_{i=1}^d |f(p + (e_k)_i) - f(p)| \\ &= \sum_{p \in \mathbb{Z}^d} \sum_{i=1}^d \left( |P(f)(p + (e_k)_i) - P(f)(p)| \right. \\ &\quad \left. + |f(p + (e_k)_i) - P(f)(p + (e_k)_i) - (f(p) - P(f)(p))| \right) \\ &= TV(P(f)) + TV((id - P)(f)). \end{aligned}$$

As an easy consequence of Theorems 16 and 18 we have

*Theorem 19:* The operators  $L_n, U_n, n = 1, 2, \dots$ , and all their compositions, are total variation preserving.

Note that if an operator  $P$  is total variation preserving then the complementary operator  $id - P$  is also total variation preserving by (2). Similarly to the case for compositions of fully trend preserving operators, it is easy to show that compositions of opera-

tors which individually preserve the total variation maintain the preservation as well.

## V. DPT OF IMAGES

The DPT based upon the LULU operators for sequences was derived in [26], [29], and [30]. Using the extension of the LULU operators to functions on  $\mathbb{Z}^d$  in the preceding sections we derive the DPT for functions in  $\mathcal{A}(\mathbb{Z}^d)$  now. Following the success of the DPT for sequences in signal processing one may expect that the DPT on  $\mathcal{A}(\mathbb{Z}^d)$  can play an important role in the analysis of these functions. Some illustrative applications in the case  $d = 2$ , that is image analysis, are given in the Section VI.

Similar to the case of sequences we obtain a decomposition of a function  $f \in \mathcal{A}(\mathbb{Z}^d)$ , with finite support. As usual  $\text{supp}(f) = \{p \in \mathbb{Z}^d : f(p) \neq 0\}$ . Let  $N = \text{card}(\text{supp}(f))$ . Following (6)–(8) we derive the DPT of  $f \in \mathcal{A}(\mathbb{Z}^d)$  by applying iteratively the operators  $L_n, U_n$  with  $n$  increasing from 1 to  $N$  as follows:

$$DPT(f) = (D_1(f), D_2(f), \dots, D_N(f)) \quad (33)$$

where the components of (33) are obtained through

$$D_1(f) = (id - P_1)(f) \quad (34)$$

$$D_n(f) = (id - P_n) \circ Q_{n-1}(f), \quad n = 2, \dots, N \quad (35)$$

and  $P_n = L_n \circ U_n$  or  $P_n = U_n \circ L_n$  and  $Q_n = P_n \circ \dots \circ P_1, n \in \mathbb{N}$ . We will show that this decomposition retains the properties of the decomposition (6) in the sense that each component  $D_n$  in (33) is a sum of discrete pulses with pairwise disjoint supports of size  $n$ , where in this setting a discrete pulse is defined as follows.

*Definition 20:* A function  $\phi \in \mathcal{A}(\mathbb{Z}^d)$  is called **apulse** if there exists a connected set  $V$  and a nonzero real number  $\alpha$  such that

$$\phi(x) = \begin{cases} \alpha, & \text{if } x \in V \\ 0, & \text{if } x \in \mathbb{Z}^d \setminus V. \end{cases}$$

The set  $V$  is the support of the pulse  $\phi$ , that is  $\text{supp}(\phi) = V$ . The concept of a pulse as defined in Definition 20 is similar to the idea of a flat zone from mathematical morphology. It should be remarked that the support of a pulse may generally have any shape, the only restriction being that it is connected. Note that the smoothing process ultimately results in the last component  $D_N(f)$  being a constant image, that is, one pulse the size of the entire image. The remaining image is  $Q_N(f) = 0$ . It follows from (34)–(35) that:

$$f = \sum_{n=1}^N D_n(f). \quad (36)$$

The usefulness of the representation (36) of a function  $f \in \mathcal{A}(\mathbb{Z}^d)$  is in the fact that all terms are sums of pulses as stated in the sequel. We use the next technical lemma proved in [4].

*Lemma 21:* Let  $f \in \mathcal{A}(\mathbb{Z}^d)$ ,  $\text{supp}(f) < \infty$ , be such that  $f$  does not have local minimum sets or local maximum sets of size smaller than  $n$ , for some  $n \in \mathbb{N}$ . Then we have the following two results:

a)

$$(id - P_n)f = \sum_{i=1}^{\gamma^-(n)} \phi_{ni} + \sum_{j=1}^{\gamma^+(n)} \varphi_{nj} \quad (37)$$



where  $V_{ni} = \text{supp}(\phi_{ni}), i = 1, 2, \dots, \gamma^-(n)$ , are local minimum sets of  $f$  of size  $n$ ,  $W_{nj} = \text{supp}(\varphi_{nj}), j = 1, 2, \dots, \gamma^+(n)$ , are local maximum sets of  $f$  of size  $n$ ,  $\phi_{ni}$  and  $\varphi_{nj}$  are negative and positive discrete pulses, respectively, and we also have that

$$\bullet V_{ni} \cap V_{nj} = \emptyset \text{ and } \text{adj}(V_{ni}) \cap V_{nj} = \emptyset \\ i, j = 1, \dots, \gamma^-(n), \quad i \neq j \quad (38)$$

$$\bullet W_{ni} \cap W_{nj} = \emptyset \text{ and } \text{adj}(W_{ni}) \cap W_{nj} = \emptyset \\ i, j = 1, \dots, \gamma^+(n), \quad i \neq j \quad (39)$$

$$\bullet V_{ni} \cap W_{nj} = \emptyset \\ i = 1, \dots, \gamma^-(n), \quad j = 1, \dots, \gamma^+(n). \quad (40)$$

b) For every fully trend preserving operator  $A$

$$U_n(id - AU_n) = U_n - AU_n \\ L_n(id - AL_n) = L_n - AL_n.$$

*Theorem 22:* Let  $f \in \mathcal{A}(\mathbb{Z}^d)$ . For every  $n \in \mathbb{N}$  the function  $D_n(f)$  derived through (34)–(35) is a sum of discrete pulses with pairwise disjoint support, that is, there exist  $\gamma(n) \in \mathbb{N}$  and discrete pulses  $\psi_{ns}, s = 1, \dots, \gamma(n)$ , such that

$$D_n(f) = \sum_{s=1}^{\gamma(n)} \psi_{ns} \quad (41)$$

and

$$\text{supp}(\psi_{ns_1}) \cap \text{supp}(\psi_{ns_2}) = \emptyset \text{ for } s_1 \neq s_2. \quad (42)$$

Moreover, if  $n_1, n_2, s_1, s_2 \in \mathbb{N}$  are such that  $n_1 < n_2, 1 \leq s_1 \leq \gamma(n_1)$  and  $1 \leq s_2 \leq \gamma(n_2)$ . Then

$$\text{supp}(\psi_{n_1 s_1}) \cap \text{supp}(\psi_{n_2 s_2}) \neq \emptyset \\ \implies \text{supp}(\psi_{n_1 s_1}) \subset \text{supp}(\psi_{n_2 s_2}). \quad (43)$$

*Proof:* According to (35),  $D_n$  is obtained by applying  $id - P_n$  to the function  $Q_{n-1}(f)$  which, by Corollary 13, does not have local maximum or minimum sets of size less than  $n$ . Thus, by Lemma 21a) we have that  $D_n(f) = (id - P_n)Q_{n-1}(f)$  is a sum of pairwise disjoint discrete pulses as given in (37). More precisely

$$D_n(f) = \sum_{s=1}^{\gamma(n)} \psi_{ns} = \sum_{i=1}^{\gamma^-(n)} \phi_{ni} + \sum_{j=1}^{\gamma^+(n)} \varphi_{nj}$$

where  $\gamma(n) = \gamma^-(n) + \gamma^+(n)$ . Property (42) follows from (38)–(40).

Let  $\text{supp}(\psi_{n_1 s_1}) \cap \text{supp}(\psi_{n_2 s_2}) \neq \emptyset$ . It follows from the construction of (41) derived above that the functions  $Q_n(f)$  and  $L_{n+1}(Q_n(f)), n \geq n_1$ , are constants on the set  $\text{supp}(\psi_{n_1 s_1})$ . Furthermore, the set  $\text{supp}(\psi_{n_2 s_2})$  is a local maximum set of  $Q_{n_2-1}(f)$  or a local minimum set of  $L_{n_2}(Q_{n_2-1}(f))$ . From the definition of local maximum set and local minimum set, it follows that  $\text{supp}(\psi_{n_1 s_1}) \subset \text{supp}(\psi_{n_2 s_2})$ . ■

Using Theorem 22, the identity (36) can be written as

$$f = \sum_{n=1}^N \sum_{s=1}^{\gamma(n)} \psi_{ns}. \quad (44)$$

The equality (44) is a discrete pulse decomposition of  $f$ , where the pulses  $\psi_{ns}$  have the properties (42)–(43).

Although the importance of total variation preservation for separators cannot be doubted, it is even more so for hierarchical decompositions like the DPT, due to the fact that they involve iterative applications of separators. Using Theorem 19 it is easy to obtain the statement of the following theorem, which shows that, irrespective of the length of the vector in (33) or the number of terms in the sum (44), no additional total variation, or noise, is created via the decomposition.

*Theorem 23:* The discrete pulse decomposition (33) is total variation preserving, that is

$$TV(f) = \sum_{n=1}^N \sum_{s=1}^{\gamma(n)} TV(\psi_{ns}). \quad (45)$$

We should remark that representing a function as a sum of pulses can be done in many different ways. However, in general, such decompositions increase the total variation, that is, we might have strict inequality in (45) instead of equality. Based upon Theorem 23 we can construct the total variation distribution of images. More precisely, this is the distribution of the total variation of an image among the different layers of the DPT. That is, essentially the plot of  $TV(D_n(f))$  versus  $n$ . In Fig. 4, we present the total variation distributions of some of the images in Fig. 3, where one can observe how the total variation of each image as given in Table I is distributed over the pulse size. A log scale is used on the vertical axis and the pulse size values are grouped as on a histogram. The different character of the images naturally manifests through different kinds of total variation distributions.

As discussed in the introduction, the quality of a nonlinear hierarchical decomposition, such as the DPT given in (33), can be characterized through the concept of *consistent decomposition* (also called *strong consistency* [26]) given in (3). Whether or not the multidimensional DPT in (33) is strongly consistent is still an open problem. In the 1-D case, it is stated in [15] to have been proven and the proof is presented at a later stage in [26]. However, the next theorem shows that the DPT in (33) satisfies a weaker kind of consistency, referred to in [26] as *basic consistency*, involving only the sums of the output layers.

*Theorem 24:* Let  $f \in \mathcal{A}(\mathbb{Z}^d)$ . For any two integers  $m$  and  $n$  such that  $m < n$  the function  $g = D_m(f) + D_{m+1}(f) + \dots + D_n(f)$  decomposes consistently, that is

$$D_j(g) = \begin{cases} D_j(f), & \text{for } m \leq j \leq n \\ 0, & \text{otherwise.} \end{cases}$$

The proof uses the following Lemma.

*Lemma 25:* Let  $Q_n = P_n P_{n-1} \dots P_1$  where  $P_k = L_k U_k$  or  $P_k = U_k L_k$ . We have the following:

- $Q_n Q_m = Q_{\max\{n,m\}}$ ;
- $Q_m(id - Q_n) = Q_m - Q_n = (id - Q_n)Q_m$  for all integer  $m, n$  such that  $m \leq n$ .

*Proof:* We consider only  $P_k = L_k U_k$  as the other case is dealt with by symmetry. Let  $f \in \mathcal{A}(\mathbb{Z}^d)$ .

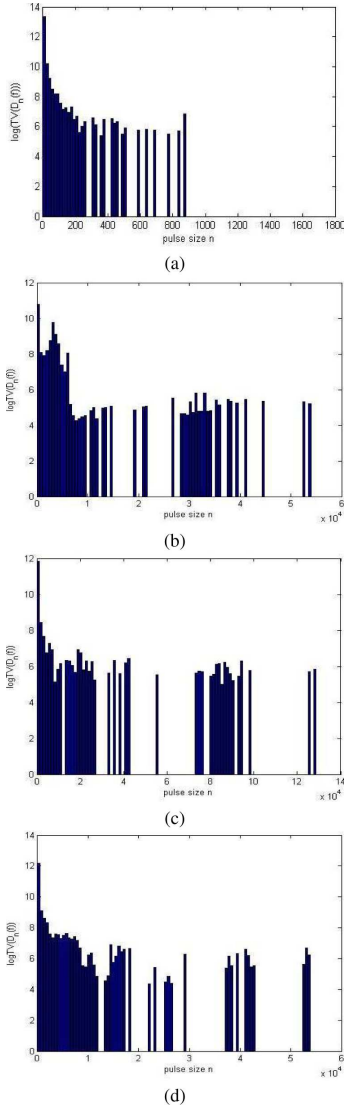


Fig. 4. Total variation distributions of images in: (a) Fig. 3(h)-noise; (b) Fig. 3(a)-potatoes; (c) Fig. 3(c)-tank; and (d) Fig. 3(f)-boat with glint.

a) It follows from Corollary 13 that  $Q_n(f)$  does not have any local minimum or local maximum sets of size  $n$  or less. Hence,  $P_k(Q_n(f)) = Q_n(f)$  for  $k = 1, \dots, n$ . For  $m \leq n$  this implies that  $Q_m(Q_n(f)) = Q_n(f)$ . If  $m > n$  then we have  $(Q_m Q_n)(f) = (P_m \dots P_{n+1} P_n \dots P_1)(Q_n(f)) = (P_m \dots P_{n+1})(Q_n(f)) = C_m(f)$ . b) We use induction on  $j$  as in the proof of this property in the 1-D case, see [28]. Let  $j = 1$ . Using the result in Lemma 21b), the full trend preservation property of the LULU operators established in Theorem 16 and the absorption property in a) we have

$$\begin{aligned} Q_1(id - Q_n) &= L_1(U_1(id - Q_n L_1 U_1)) \\ &= L_1(U_1 - Q_n L_1 U_1) = L_1(id - Q_n L_1) U_1 \\ &= (L_1 - Q_n L_1) U_1 \\ &= Q_1 - Q_n = (id - Q_n) Q_1. \end{aligned}$$

Assume now that the statement is true for some  $m = j < n$ . From the inductive assumption, we have

$$\begin{aligned} Q_{j+1}(id - Q_n) &= P_{j+1} Q_j (id - Q_n) = P_{j+1} (Q_j - Q_n) \\ &= P_{j+1} (Q_j - Q_n Q_j) = P_{j+1} (id - Q_n) Q_j. \end{aligned}$$

Using Lemma 21b), Theorem 16 and a) as for  $j = 1$  we obtain

$$\begin{aligned} P_{j+1}(id - Q_n) Q_j &= L_{j+1}(U_{j+1}(id - Q_n L_{j+1} U_{j+1}) Q_j) \\ &= (L_{j+1} U_{j+1} - Q_n L_{j+1} U_{j+1}) Q_j \\ &= Q_{j+1} - Q_n = (id - Q_n) Q_{j+1}. \end{aligned}$$

■

*Proof of Theorem 24:* Using Lemma 25, function  $g$  can be written in the following equivalent forms:

$$\begin{aligned} g &= ((id - P_m) Q_{m-1} + (id - P_{m+1}) Q_m \\ &\quad + \dots + (id - P_n) Q_{n-1})(f) \\ &= (Q_{m-1} - Q_n)(f) = (id - Q_n) Q_{m-1} \\ &= Q_{m-1} (id - Q_n). \end{aligned} \quad (46)$$

It follows from Corollary 13 and Theorem 17 that  $g$  does not have any local maximum or local minimum sets of size less than  $m$ . Hence,  $P_k(g) = g$  for  $k = 1, \dots, m-1$  and therefore  $Q_k(g) = g$  for  $k = 1, \dots, m-1$ . Then it follows from (46) that  $D_j(g) = (id - P_j)(g) = 0$  for  $j < m$ . Let  $m \leq j \leq n$ . Then using again Lemma 25 we obtain

$$\begin{aligned} D_j(g) &= (Q_{j-1} - Q_j)(g) \\ &= (Q_{j-1}(id - Q_n) Q_{m-1} - Q_j(id - Q_n) Q_{m-1})(f) \\ &= ((id - Q_n) Q_{j-1} Q_{m-1} - (id - Q_n) Q_j Q_{m-1})(f) \\ &= ((id - Q_n) Q_{j-1} - (id - Q_n) Q_j)(f) \\ &= (Q_{j-1} - Q_n - Q_j + Q_n)(f) = D_j(f). \end{aligned}$$

Finally, for  $k \geq n$  we have

$$\begin{aligned} Q_k(g) &= (Q_k(id - Q_n) Q_{m-1})(f) \\ &= (Q_k Q_n (id - Q_n) Q_{m-1})(f) = 0 \end{aligned}$$

which implies that  $D_j(g) = 0$  for  $j > n$ . ■

## VI. SOME APPLICATIONS TO IMAGE ANALYSIS

It is generally accepted that an image is perceived through the contrast, that is, the difference in the luminosity of neighbor pixels. The DPT (33) extracts all such differences as single pulses. Hence, (33) can be a useful tool in the analysis of images.

Random noise has very distinctive discrete pulse decomposition characterized by fast decrease of the number of pulses with the increase of the pulse size. The total variation of the pulses in decomposition (33) versus their size for an image of random noise [in Fig. 3(h)] is plotted in Fig. 4(a). It is apparent that random noise seldom generates pulses of large size. In fact, 90% of the total variation represented on Fig. 4(a) arises from pulses of size less than or equal to 76, only about 8% of the total variation arises from pulses with size greater than 100 and

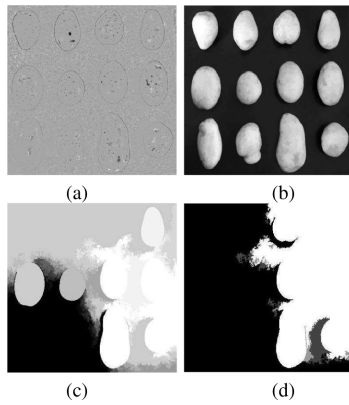


Fig. 5. Reconstruction via DPT of Fig. 4(a) with only pulses of size: (a) 1 to 100; (b) 101 to 22000; (c) 22001 to 45000; and (d) 45001 to 57478.

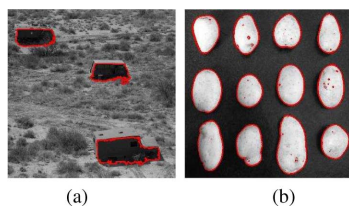


Fig. 6. (a) Detection of the targets in Fig. 3(d). (b) Potato defects detected.

only about 5% of the total variation arises from pulses with size greater than 200. Hence, by removing the pulses of small support we remove a large portion of the impulsive noise.

Possibly the simplest application of the discrete pulse decomposition (33) is via partial reconstructions of images. This can be used for example in removing noise or extracting features of interest. See [12] for additional examples. Fig. 4(b) gives total variation distribution of the image in Fig. 3(a). A large portion of the pulses have small support but, unlike Fig. 4(a), we have also significant proportion of the total variation arising from pulses with relatively larger support. Partial reconstruction of the image by using pulses of selected sizes is given on Fig. 5. We can consider (a) as removal of impulsive noise and the defects on each of the potatoes, (b) as the removal of the background surface illumination and detection of the individual potatoes, (c) as the extracted illumination and (d) as extraction of large features.

Every feature on an image is represented without distortion by a set of pulses in the DPT. This can be used to carry out many of the usual tasks in image processing like segmentation, detection, identification. In the general setting of DPT, the issue of detection or identification of objects means identifying pulses uniquely associated these objects. Hence, a good detection algorithm will necessarily employ a good criterion for selecting pulses in the DPT of the image. Some applications of are presented in [5]. For example, due to the largely homogeneous background the three targets in Fig. 3(d) can be easily detected by selecting pulses of a particular size. By varying the pulse sizes on a sliding scale from 4500 to 5300 pixels at the top to 9000 to 12000 in the bottom, three nested sets of pulses associated with the targets are selected. The pulses with largest support in each set are indicated in Fig. 6(a).

As a further example, one may detect the defects on the potatoes in Fig. 3(a) (possibly for automatic grading) as small pulses

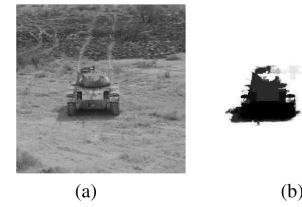


Fig. 7. (a) Original image; (b) pulses with high rectangularity.

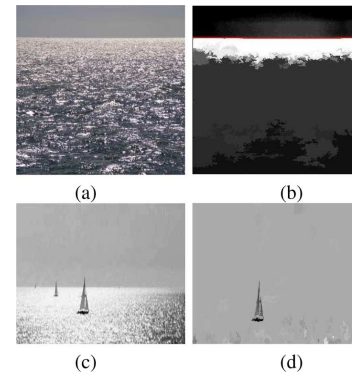


Fig. 8. (a) Original glint image; (b) layers of the ocean image in (a); (c) image with three boats and glint; and (d) detection of the biggest boat within ocean glint by restricting the total variation only.

of sufficiently large value (relative luminosity) appearing within the pulses representing the potatoes, see Fig. 6(b).

Such a simple approach does not always work. For example the image in Fig. 7(a) has no pulses of a particular size associated only with the tank. In this sense, the tank can be considered to be well camouflaged. However, as it often happens, straight lines and rectangular shapes are associated with man-made objects. In Fig. 7(b), we indicate the pulses of size 8001 to 43000 with rectangularity larger than 0.5. This selection criterion yields only pulses associated with the tank and that can be considered to be specific for its “DPT signature.” Here we use rectangularity, which is a measure of approximation of the support of the pulse by a rectangle and is measured as the ratio between the size of the support and the size of its minimum bounding rectangle.

As already indicated in Fig. 5(c) and (d), large pulses are a result of global features like illumination. In Fig. 8(b), large pulses represent the layers of the sea and the horizon from the image in Fig. 8(a). The horizon has also been highlighted. Due to the angle of the camera the glint on the ocean surface becomes more homogenous out towards the horizon and we are able to pick out three different glint layers, as well as the horizon, separately. It is well known that detection on a sea surface is a difficult problem due to glint on the ocean’s surface producing “noise” of various shapes, sizes and luminosities. The pulses associated with the boats in Fig. 8(c) have lower total variation on their support, thus, restricting the total variation we obtain Fig. 7(d). Here, we see the boat without the glint.

## VII. CONCLUSION

In this paper, we derived an extension of Carl Rohwer’s LULU theory for sequences to higher dimensions, namely to functions on multidimensional arrays. We proved that the

essential properties of the LULU operators (e.g., separator, shape preservation, and total variation preservation) are replicated in the considered multidimensional setting. Hence, we can construct a DPT, where the pulses are functions which are constant on a connected set and zero elsewhere. The quality of a hierarchical decomposition, of which the DPT is a particular case, can be characterized in terms of their consistency. We prove that the DPT satisfies the basic consistency, while its strong consistency remains an open problem. Some motivation for this theory and its further development is provided through applications to problems in image analysis. These applications, although mainly for illustrative purposes, nevertheless indicate a wide range of problems that one might be able to address using DPT. Future work will focus on both extending the theory and strengthening its applications. There are many image processing ideas that have sprung up during this research which will be further investigated in the near future. These include a scale space theory derived via the DPT, granulometries, image compression, pattern recognition and image segmentation, amongst others. The applicability of the extension for images, and more complex domains e.g., video, is also now open for investigation.

#### REFERENCES

- [1] C. R. Adams and J. A. Clarkson, "Properties of functions  $f(x, y)$  of bounded variation," *Trans. AMS*, vol. 36, no. 4, pp. 711–730, 1934.
- [2] R. Anguelov, "LULU operators and locally  $\delta$ -monotone approximations," in *Constructive Theory of Functions*. Sofia, Bulgaria: Marin Drinov, 2006, pp. 22–34.
- [3] R. Anguelov and C. H. Rohwer, "LULU operators for functions of continuous argument," *Quaestiones Math.*, vol. 32, no. 2, pp. 187–202, 2009.
- [4] R. Anguelov and I. N. Fabris-Rotelli, "Properties of the discrete pulse transform for multi-dimensional arrays," [Online]. Available: <http://arxiv.org/abs/1003.0776>
- [5] B. Anguelov, "Discrete pulse transform of images: Algorithm and applications," in *Proc. Int. Conf. Pattern Recognit.*, Tampa, FL, Dec. 2008, pp. 8–11.
- [6] A. Chambolle and P.-L. Lions, "Image recovery via total variation minimization and related problems," *Num. Math.*, vol. 76, no. 1997, pp. 167–188.
- [7] J. A. Clarkson and C. R. Adams, "On definitions of bounded variation for functions of two variables," *Trans. AMS*, vol. 35, no. 4, pp. 824–854, 1933.
- [8] W. J. Conradie, T. de Wet, and M. Jankowitz, "Exact and asymptotic distributions of LULU smoothers," *J. Comput. Appl. Math.*, vol. 186, pp. 253–267, 2006.
- [9] J. Crespo, "Adjacency stable connected operators and set levelings," in *Proc. 8th Int. Symp. Math. Morphology*, Rio de Janeiro, Brazil, Oct. 10–13, 2007, pp. 215–226.
- [10] R. de la Vega and A. Restrepo, "Efficient computation of locally monotonic regression," *IEEE Signal Process. Lett.*, vol. 3, no. 9, pp. 263–265, Sep. 1996.
- [11] E. R. Dougherty and J. Astola, *An Introduction to Nonlinear Image Processing*. Bellingham, WA: SPIE, 1994.
- [12] R. Anguelov and I. N. Fabris-Rotelli, "Discrete pulse transform of images," in *Proc. Int. Conf. Image Signal Process.*, Normandy, France, Jul. 1–3, 2008, vol. 5099, pp. 1–9.
- [13] T. S. Huang, G. J. Yang, and G. Y. Tang, "A fast two-dimensional median filtering algorithm," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-27, no. 1, pp. 13–18, Feb. 1979.
- [14] O. Kao, "Modification of the LULU operators for preservation of critical image details," in *Proc. Int. Conf. Imag. Sci., Syst. Technol.*, Las Vegas, NV, Jun. 2002, pp. 24–27.
- [15] D. P. Laurie and C. H. Rohwer, "Fast implementation of the discrete pulse transform," in *Proc. Int. Conf. Numer. Anal. Appl. Math.*, Crete, Greece, Sep. 15–19, 2006, pp. 484–487.
- [16] D. P. Laurie, "The roadmaker's algorithm for the discrete pulse transform," *IEEE Trans. Image Process.*, 2010, submitted for publication.
- [17] S. G. Mallet, "A theory for multiresolution signal decomposition: The wavelet representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, no. 7, pp. 674–693, Jul. 1989.
- [18] C. L. Mallows, "Some theory of nonlinear smoothers," *Ann. Stat.*, vol. 8, pp. 695–715, 1980.
- [19] G. Matheron, "Filters and lattices," in *Image Analysis and Mathematical Morphology, Volume II: Theoretical Advances*, J. Serra, Ed. New York: Academic, 1988, ch. 6, pp. 115–140.
- [20] A. B. Owen, "Multidimensional variation for quasi-monte carlo," in *Contemporary Multivariate Analysis and Design of Experiments*, J. Fan and G. Li, Eds. Singapore: World Scientific, 2005.
- [21] I. Pitas and A. N. Venetsanopoulos, "Order statistics in digital image processing," *Proc. IEEE*, vol. 80, no. 12, pp. 1893–1921, Dec. 1992.
- [22] I. Fabris-Rotelli and S. Van der Walt, "The discrete pulse transform in two dimensions," in *Proc. 20th Ann. Symp. Pattern Recognit. Assoc. South Africa*, 2009 [Online]. Available: <http://www.prasa.org/proceedings/2009/prasa09-15.pdf>
- [23] A. Restrepo Palacios and L. Chacon, "A smoothing property of the median filter," *IEEE Trans. Signal Process.*, vol. 42, no. 6, pp. 1553–1555, Jun. 1994.
- [24] C. H. Rohwer, "Idempotent one-sided approximation of median smoothers," *J. Approx. Theory*, vol. 58, no. 2, pp. 151–163, 1989.
- [25] C. H. Rohwer, "Variation reduction and LULU—Smoothing," *Quaestiones Mathematicae*, vol. 25, pp. 163–176, 2002.
- [26] C. H. Rohwer and D. P. Laurie, "The discrete pulse transform," *SIAM J. Math. Anal.*, vol. 38, no. 3, pp. 1012–1034, 2006.
- [27] C. H. Rohwer and M. Wild, "Natural alternatives for one dimensional median filtering," *Quaestiones Math.*, vol. 25, pp. 135–162, 2002.
- [28] C. H. Rohwer, "Fully trend preserving operators," *Quaestiones Math.*, vol. 27, pp. 217–230, 2004.
- [29] C. H. Rohwer, *Nonlinear Smoothers and Multiresolution Analysis*. Cambridge, MA: Birkhäuser, 2005.
- [30] C. H. Rohwer and M. Wild, "LULU theory, idempotent stack filters, and the mathematics of vision of marr," *Adv. Imag. Elect. Phys.*, vol. 146, pp. 57–162, 2007.
- [31] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D*, vol. 60, pp. 259–268, 1992.
- [32] J. Serra, *Image Analysis and Mathematical Morphology*. New York: Academic, 1982.
- [33] J. Serra, *Image Analysis and Mathematical Morphology, Volume II: Theoretical Advances*, J. Serra, Ed. New York: Academic, 1988.
- [34] J. Serra, "Mathematical morphology for boolean lattices," in *Image Analysis and Mathematical Morphology, Volume II: Theoretical Advances*, J. Serra, Ed. New York: Academic, 1988, ch. 2, pp. 37–58.
- [35] J. Serra, "A lattice approach to image segmentation," *J. Math. Imag. Vis.*, vol. 24, pp. 83–130, 2006.
- [36] N. D. Sidropoulos, "Fast digital locally monotonic regression," *IEEE Trans. Signal Process.*, vol. 45, no. 2, pp. 389–395, Feb. 1997.
- [37] P. Soille, *Morphological Image Analysis*. New York: Springer-Verlag, 1999.
- [38] E. Tadmor, S. Nezzar, and L. Vese, "Multiscale hierarchical decomposition of images with applications to deblurring," *Denoising Segment., Comm. Math. Sci.*, vol. 6, no. 2, pp. 281–307, 2008.
- [39] M. Wild, "Idempotent and co-idempotent stack filters and min-max operators," *Theor. Comput. Sci.*, vol. 299, pp. 603–631, 2003.



**Roumen Anguelov** received the M.Sc. degree in mathematics from Sofia University, Bulgaria, and the Ph.D. degree in mathematics from the University of South Africa, Pretoria.

He is an Associate Professor at the University of Pretoria, South Africa. His current research interests include hierarchical decomposition of images, mathematical morphology, and PDE-based methods of image analysis.



**Inger Fabris-Rotelli** received the M.Sc. degree in applied mathematics from the University of Pretoria, South Africa, and is pursuing the Ph.D. degree in applied mathematics from the same university.

She is a Lecturer at the University of Pretoria. Her current research interests include feature detection in images, hierarchical decomposition of images, and noise removal in signals and images.