



Nonlinear Filters and Characterization of the Discrete Pulse Transform of Images

Inger Fabris-Rotelli

Department of Statistics, University of Pretoria
 Pretoria, South Africa
 inger.fabris-rotelli@up.ac.za

Received: 15 July 2012, accepted: 11 November 2012, published: 29 December 2012

Abstract—The Discrete Pulse Transform (DPT) for images and videos has been developed over the past few years and provides a theoretically sound setting for a nonlinear decomposition of an image or video. In [1] the theoretical basis of the DPT was presented. In this paper we now present a sound characterization of this useful nonlinear hierarchical decomposition by referring to its ability as a separator, the consistency of the decomposition, as well as the smoothing ability of the decomposition.

Keywords—LULU; Discrete Pulse Transform; DPT; Nonlinear Decomposition; Feature Detection

I. INTRODUCTION

The Discrete Pulse Transform (DPT) is a nonlinear hierarchical decomposition obtained by the successive application of the LULU operators L_n and U_n where n increases from 1 to N , where N is the number of data points in the signal. For a concise overview of the one-dimensional LULU operators see [2] as well as further collaboration with Laurie and Wild. The LULU operators were extended in detail to multidimensional arrays in [1], which provides a framework for the obvious areas of image processing in two dimensions, as well as video processing in three dimensions. We provide a short overview of these operators and the DPT here for completeness.

A. LULU Operators

The concept of morphological connectivity was introduced by J. Serra and G. Matheron in the 1980's. They recognised the need for the concept of an axiomatic

connectivity and thus the axiomatic approach to connectivity was introduced. In 1988 Serra and Matheron, [3], introduced the concept of a connectivity class, for use in Mathematical Morphology.

Definition 1: \mathcal{C} is a **connectivity class** or a **connection** on the power set $\mathcal{P}(E)$ if the following axioms hold: (i) $\emptyset \in \mathcal{C}$ (ii) $\{x\} \in \mathcal{C}$ for each $x \in E$ (iii) For each family $\{C_i\}$ in \mathcal{C} such that $\bigcap C_i \neq \emptyset$, we have $\bigcup C_i \in \mathcal{C}$. A set $C \in \mathcal{C}$ is called **connected**.

We define the operators L_n and U_n on $\mathcal{A}(\mathbb{Z}^d)$, where $\mathcal{A}(\mathbb{Z}^d)$ is the vector lattice of all real functions defined on \mathbb{Z}^d with respect to the usual point-wise defined addition, scalar multiplication and partial order.

Definition 2: Let $f \in \mathcal{A}(\mathbb{Z}^d)$ and $n \in \mathbb{N}$. Then

$$L_n(f)(x) = \max_{V \in \mathcal{N}_n(x)} \min_{y \in V} f(y), \quad x \in \mathbb{Z}^d,$$

$$U_n(f)(x) = \min_{V \in \mathcal{N}_n(x)} \max_{y \in V} f(y), \quad x \in \mathbb{Z}^d,$$

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

It is important to notice that here the collection of n -neighbourhoods, $\mathcal{N}_n(x)$, the LULU operators act on, can take on any shape as the only restriction is on their size. This is the important advantage of the LULU operators, which are only concerned about size, and morphological filters, which operate in conjunction with a specified structuring element with a specified size. This allows for open investigation of the image structures instead of searching for structures of a specific shape.

B. The Discrete Pulse Transform

The operators L_n and U_n smooth the input signal by removing peaks (the application of L_n) and pits (the application of U_n). These peaks and pits are defined mathematically in Definition 3 and 4.

Definition 3: 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, $\text{adj}(V) = \{x \in \mathbb{Z}^d : x \notin V, V \cup \{x\} \in \mathcal{C}\}$.

Definition 4: A connected subset V of \mathbb{Z}^d is called a **local maximum (minimum) set** of $f \in \mathcal{A}(\mathbb{Z}^d)$ if

$$\sup(\inf)_{y \in \text{adj}(V)} f(y) < (>) \inf(\sup)_{x \in V} f(x).$$

The results proved in [1] and [4] provide the following summary of the effect of the operators L_n and U_n on a function $f \in \mathcal{A}(\mathbb{Z}^d)$:

- (1) The application of L_n (U_n) removes local maximum (minimum) sets of size smaller or equal to n .
- (2) No new local minimum (maximum) sets are created where there were none, however the action of L_n (U_n) may enlarge existing local minimum (maximum) sets or join two or more local minimum (maximum) sets of f into one local minimum (maximum) set of $L_n(f)$ ($U_n(f)$).
- (3) $L_n(f) = f$ ($U_n(f) = f$) if and only if f does not have local maximum (minimum) sets of size n or less.

The DPT provides a representation of an image (when $d = 2$) and higher dimensional arrays at all the scale levels. We obtain a decomposition of a function $f \in \mathcal{A}(\mathbb{Z}^d)$, with finite support. Let $N = \text{card}(\text{supp}(f))$. 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

$$DPT(f) = (D_1(f), D_2(f), \dots, D_N(f)), \quad (1)$$

where the components of (1) are obtained through $D_1(f) = (I - P_1)(f)$, $D_n(f) = (I - P_n) \circ Q_{n-1}(f)$, $n = 2, \dots, N$, $P_n = L_n \circ U_n$ or $P_n = U_n \circ L_n$ and $Q_n = P_n \circ \dots \circ P_1$, $n \in \mathbb{N}$. This decomposition results in components D_n which are each a sum of discrete pulses ϕ_{ns} , $s = 1, 2, \dots, \gamma(n)$ with disjoint supports of size n , where in this setting a discrete pulse is defined as follows,

Definition 5: A function $\phi \in \mathcal{A}(\mathbb{Z}^d)$ is called a **pulse** if there exists a connected set V and a real number α 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 ϕ , that is $\text{supp}(\phi) = V$.

We can then reconstruct the original signal as

$$f = \sum_{n=1}^N D_n(f) = \sum_{n=1}^N \sum_{s=1}^{\gamma(n)} \phi_{ns}.$$

II. CHARACTERIZATION OF THE DPT

A. Linear versus Nonlinear

As discussed in [4], the nonlinearity of the LULU smoothers makes theoretical development more complicated than for linear operators. However, taking on the additional complexity is justified since in two dimensions an image is basically the transformation of data by a human eye or measuring instrument. This transformation is significantly complicated to be considered nonlinear [5]. Thus taking this stance the analysis of images via nonlinear operators is more logical than that of linear. Linear operators are also notorious for blurring edges.

Linear processing techniques are however a natural starting point for analysis due to the simplicity of their application and theoretical backbone available. Examples of linear filters are the Fourier transform, Hadamard transform, the discrete cosine transform, and wavelets. They also provide sufficient results in most applications, but there are problems in which a nonlinear process would prove more viable and efficient. Pitas and Venetsanopoulos [6] provide examples of such cases, such as signal dependent noise filtering e.g. photoelectron noise of photosensing devices; multiplicative noise appearing as speckle noise in ultrasonic imaging and laser imaging; and nonlinear image degradations e.g. when transmission occurs through nonlinear channels. Advantages of nonlinear filters are 1) the ability to handle various noise types, 2) edge preservation, 3) fine detail preservation, 4) unbiasedness (directional and illumination based) or invariance, and 5) computational complexity [6].

Nonlinear filtering techniques can be broadly classified accordingly in the following areas: order statistic filters, homomorphic filters, polynomial filters, mathematical morphology, neural networks, and nonlinear image restoration [6]. The LULU operators fit nicely into the areas of mathematical morphology, due to their similarity to area operators therein, as well as order statistics, two areas which have been integrated quite efficiently in literature [6]. Examples of order statistics, discussed in detail in [6], are the median, rank-order filters, max-min filters, L_p -mean filters, and α -trimmed mean filters. The LULU operators are examples of max-min filters but with the disadvantages listed in [6] improved upon.

The basic filters of mathematical morphology are the erosion and dilation, and subsequently the morphological opening and closing, to which the LULU filters are again closely related.

B. Separators

A common requirement for a filter P , linear or non-linear, is its idempotence, that is $P \circ P = P$. For linear operators the idempotence of P implies the idempotence of the complementary operator $I - P$, where I denotes the identity operator. For nonlinear filters this implication generally does not hold, so the idempotence of $I - P$, also called co-idempotence, [7], can be considered as an essential measure of consistency.

For every $a \in \mathbb{Z}^d$ the operator $E_a : \mathcal{A}(\mathbb{Z}^d) \rightarrow \mathcal{A}(\mathbb{Z}^d)$ given by $E_a(f)(x) = f(x - a)$, $x \in \mathbb{Z}^d$, is called a shift operator. We now define a separator which mimics the actions required of an operator P . The first three properties in Definition 6 define a **smoother**. More detail on smoothers can be found in [4].

Definition 6: An operator $P : \mathcal{A}(\mathbb{Z}^d) \rightarrow \mathcal{A}(\mathbb{Z}^d)$ is called a **separator** if

- (i) $P \circ E_a = E_a \circ P$, $a \in \mathbb{Z}^d$ (Horizontal shift invariance)
- (ii) $P(f + c) = P(f) + c$, $f, c \in \mathcal{A}(\mathbb{Z}^d)$, c a constant function (Vertical shift invariance)
- (iii) $P(\alpha f) = \alpha P(f)$, $\alpha \in \mathbb{R}$, $\alpha \geq 0$, $f \in \mathcal{A}(\mathbb{Z}^d)$ (Scale invariance)
- (iv) $P \circ P = P$ (Idempotence)
- (v) $(I - P) \circ (I - P) = I - P$ (Co-idempotence)

Figure 1 illustrates the action of a separator P . It shows how a separator will separate the signal into noise and the true signal without the need for recursive smoothing, that is, it does the separation on the first filter application so that there is no ‘signal’ left in the ‘noise’ nor any ‘noise’ left in the ‘signal’. The median smoother is an example of a filter which requires recursive application. The LULU operators L_n and U_n and all their compositions are separators thus providing a strong separating capability of a signal.

C. Nonlinear Decompositions

The structure of a hierarchical decomposition is as follows in general. The operator \mathcal{F}_1 is applied to the input image f to obtain a decomposition of f into f_1 , the smoother image, and D_1 , the noise component removed. This process is repeated with $\mathcal{F}_2, \mathcal{F}_3, \dots, \mathcal{F}_N$ until there is nothing left to remove except the constant image D_N . The decomposition then has the form $f = \beta_1 D_1 + \beta_2 D_2 + \dots + \beta_N D_N$, for some $\beta_i, i = 1, \dots, N$. Such a

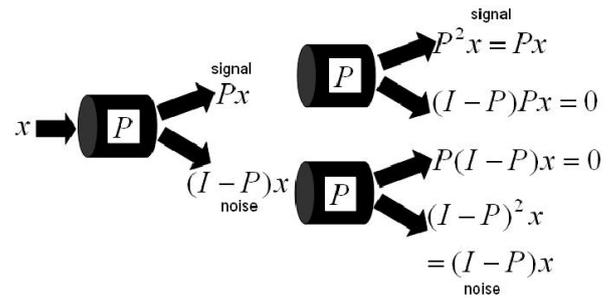


Fig. 1. The action of a separator P

hierarchical decomposition has been investigated intensively in literature, see [8], [9], [10] for some nonlinear cases. However, in no literature have we found a unified theoretical backbone to connect such nonlinear hierarchical decompositions and provide methods of comparison nor methods of testing the capability of the structure of the decomposition. In Tadmor et al [8], for example, a decomposition $f = \sum_{j=1}^k u_j + v_k$ is obtained, where v_k is the noise component and the u_j 's the decomposition components, by functional minimization. Tadmor et al discuss convergence of the minimizer, localization and adaptability, but nothing to indicate the strength of the decomposition save numerical visual examples. Similarly Wong et al [10] do not provide a theoretical indication of the strength of their decomposition obtained as a probabilistic scale-space derived from the nonlinear diffusion equation in [11]. In [9], the authors even state that comparisons with their proposed nonlinear scale-space and other nonlinear hierarchical decompositions ‘are to be made with care’.

D. Consistent Decompositions

As stated in [6] and above in Section II-C, the main limitations of nonlinear decompositions is the lack of a theory with which to compare the ability of various decompositions. The Highlight theorem first conjectured in 2007 [12], and later proved in 2010 [13], provides this much needed backbone. The quality of a nonlinear hierarchical decomposition, such as the Discrete Pulse Transform given in (1), can be characterized through the concept of *consistent decomposition* (also called *strong consistency*, [12]). For a linear decomposition of two signals f and g , namely $D(f) = \sum_i D_i(f)$ and $D(g) = \sum_j D_j(g)$, the equality

$$D(\alpha f + \beta g) = \alpha \sum_i D_i(f) + \beta \sum_j D_j(g) \quad \forall \alpha, \beta \in \mathbb{R}$$

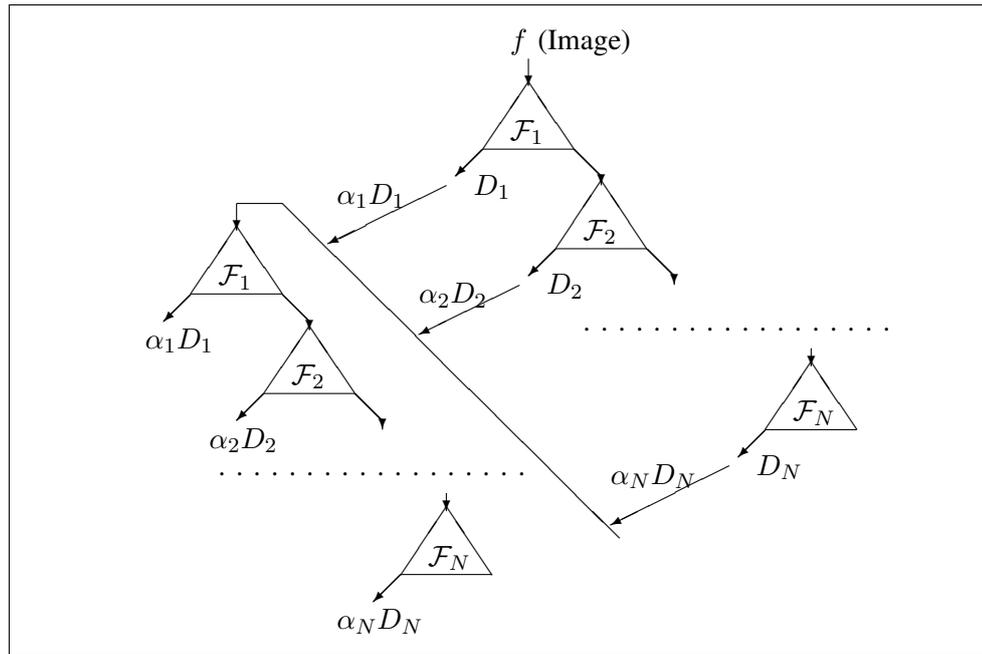


Fig. 2. Illustration of the Highlight Theorem

always holds. This is clearly a desirable property but for a nonlinear decomposition this will in general not be satisfied. However, a weaker form of linearity is provided by the Highlight Theorem. It shows that the multidimensional Discrete Pulse Transform in (1) is strongly consistent, in the sense that the above holds for $\alpha, \beta > 0$. The Highlight Theorem [13] states the following: *Given a basis of pulses identified by the DPT for a signal f , form a new function g as any linear combination of the pulses of f with heights the same sign as before. Then the DPT of g will identify the same basis of pulses, and recover the new heights.*

Figure 2 provides an illustration of the Highlight theorem. The name of theorem indicates its usefulness. Besides the ‘weak’ linearity it presents for a nonlinear decomposition, which we shall define as *highlight-linearity*, it allows for the highlighting or emphasizing of specific pulses deemed to be important, without destroying the structure of the decomposition. We state the theorem more precisely in the following form.

Theorem 7: For a DPT decomposition of f , let $g = \sum_{n=1}^N \sum_{s=1}^{\gamma(n)} \alpha_{ns} \phi_{ns}$ where the constants α_{ns} are positive. Then the DPT decomposition of g is obtained as $DPT(g) = \sum_{n=1}^N \sum_{s=1}^{\gamma(n)} \alpha_{ns} \phi_{ns}$, so that the pulses of g are obtained as $\alpha_{ns} \phi_{ns}$. If $\alpha_{ns} = \alpha_n$ for each n then $DPT(g) = \sum_{n=1}^N \alpha_n D_n(f)$, so that $D_n(g) = \alpha_n D_n(f)$.

The proof of Theorem 7 by Laurie for any dimension

can be found in [13]. Therein the theory is described by considering the signal as a graph where the data points are the graph vertices and with the edges between the vertices based on the geometry of the of the signal (i.e. connectivity). Laurie also provides an implementation algorithm for the DPT which is $O(m)$ where m is the number of edges in the graph.

The DPT decomposition is also total variation preserving. We assume \mathcal{C} on \mathbb{Z}^d is a graph connectivity, for example for images the pulses of the DPT are based on 4- or 8-connectivity, the individual pulses can be viewed as a graph $\mathcal{G} = (V_n, E_m)$ with the data points as the n vertices V_n and the neighbour relation between data points as the m edges E_m . The connectivity of such a graph \mathcal{G} 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 neighbours of a point in the given context. For example, 4-connectivity and 8-connectivity. 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 neighbour to the next one, p is neighbour to p_1 and p_k is neighbour to q . Here we assume that r is reflexive, symmetric and shift invariant, $(p, p + e_k) \in r$ for all $k = 1, 2, \dots, d$ and $p \in \mathbb{Z}^d$, where

$$e_k \in \mathbb{Z}^d \text{ is defined by } (e_k)_i = \begin{cases} 0 & \text{if } i \neq k \\ 1 & \text{if } i = k \end{cases} .$$

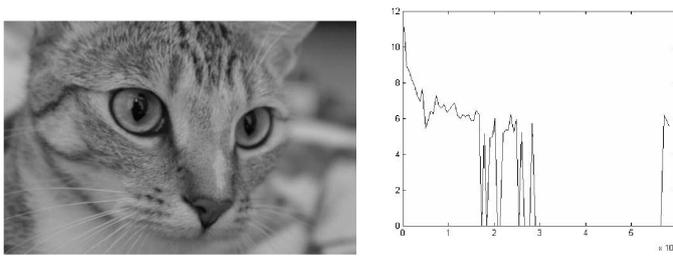


Fig. 3. Full total variation distribution and log histogram total variation distribution of Cat Image

The *Total Variation* of $f \in \mathcal{A}(\mathbb{Z}^d)$ is given by

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

Although the importance of total variation preservation for separators cannot be doubted, it is even more so for hierarchical decompositions like the Discrete Pulse Transform, due to the fact that they involve iterative applications of separators. Since the operators $L_n, U_n, n = 1, 2, \dots, N$, and all their compositions, are total variation preserving, it is easy to obtain the following result, which shows that, irrespective of the length of the vector or the number of terms in the sum, no additional total variation, or noise, is created via the decomposition, namely

$$TV(f) = \sum_{n=1}^N \sum_{s=1}^{\gamma(n)} TV(\phi_{ns}). \tag{2}$$

The proof of this result can be found in [1]. 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 (2) instead of equality. Based on this result 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 $TV(D_n(f))$ vs. n . In Figure 3 we present the total variation distribution of an image, where one can observe how the total variation is distributed over the pulse size. In the graph a log scale is used on the vertical axis and the pulse size values are grouped to form a histogram. The different character of images naturally manifests through different forms of total variation distributions. Using the total variation distribution as a guide for the content of the image we can determine, firstly, that the image is relatively ‘denoised’ when the total variation graph stabilizes, that is, the very little information is

removed after this scale by the DPT. In Figure 4 various scales of the DPT are picked out based on the total variation graph. Scales 1 to 17859 indicate the smoothed image, that is, when the total variation removal stabilizes. This smoothed image appears very similar to the original (with an MSSIM value of 0.9736 (see section II-E)) but has been smoothed by removing the remaining scale levels, 17860 to 58571, which contain the large undetailed pulses. Scales 1 to 880 provide the texture or detail of the image, scales 4385 to 4395 represent the large eye of the cat and scale 11420 the eyes, nose and mouth of the cat.

E. Measuring the Smoothing Ability of the LULU Operators

The ability of an operator to effectively remove noise and smooth the signal is usually measured by its output variance or the rate of success in the noise removal [6]. Other measures used to assess the performance are the mean square error (MSE) and signal-noise-ratio (SNR) [6]. In this section we shall present a method in which to measure the quality of the smoothed image as the question of the smoothing ability of the DPT arises. The aim of a smoother is of course to remove the noise element present. The noise can be due to a number of factors, for example, acquisition, processing, compression, storage, transmission and reproduction of the image [14]. The easiest method of evaluation is purely subjective - namely, human visual investigation. In order for evaluation to be objective, quantitative methods need to be used instead.

Quantitative methods can be divided into three categories [14]. First, *full-reference*, where the complete reference (undistorted) image is known with certainty, secondly, *no-reference*, where this reference image is not known at all, and third, *reduced-reference*, where only part of the original reference image is known, for example, a set of extracted features. We measure the similarity of the smoothed images $P_n(f)$ to the original unsmoothed image f with the structural similarity index [14],

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)},$$

for two corresponding sets of pixels, x and y , in each image, where $\mu_i, i = x, y$ is the mean of the pixel values in i , $\sigma_i^2, i = x, y$ is the variance of the pixel values in i , σ_{xy} the covariance between x and y , $c_j = (k_j L)^2, j = 1, 2$ constants to stabilize the division by the weak denominator, $L = 255$ for greyscale images and

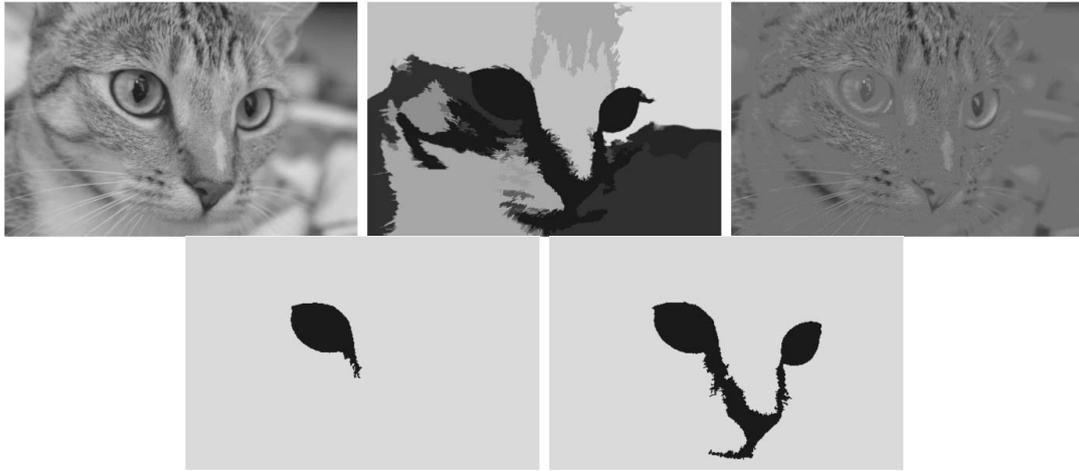


Fig. 4. Chelsea image using only scales 1 to 17859, 17860 to 58571, and 1 to 880 in the first row, and 4385 to 4395 and 11420 in the second row, respectively

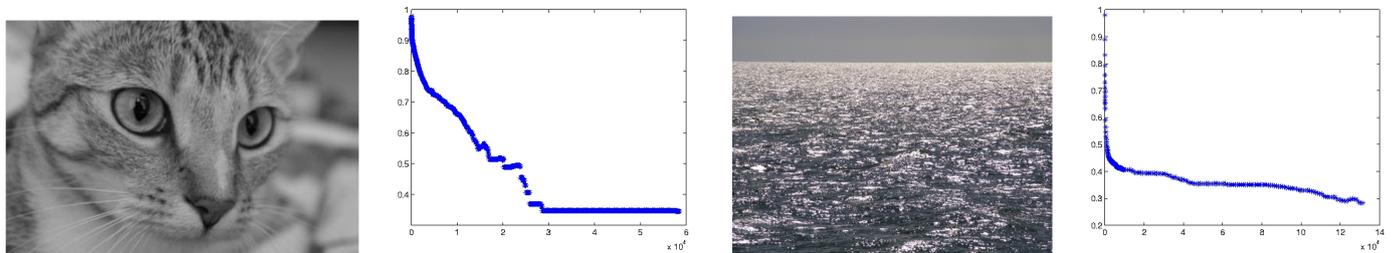


Fig. 5. MSSIM values comparing $P_n(f)$ with f values plotted against scale for the Cat image and an Ocean image

$k_j \ll 1$ constants (we used $k_j = 0.05$). This measure is a full-reference measure which provides a useful framework since we are comparing a smoothed version of the original distorted image with the original distorted image. The most widely used such measures are the mean-square-error (MSE) and the peak signal-noise-ratio (PSNR), but these measures do not compare well with the perceived visual quality of the human visual system. This measure is introduced in order to penalize errors based on their visibility, that is, to simulate the HVS as much as possible. This measure is applied to 8×8 windows in the image for each pixel and a final mean structural similarity index is calculated as the average of these SSIM values, called the MSSIM. An MSSIM value closer to 1 indicates stronger similarity. Wang et al provide MATLAB code for the implementation of the MSSIM as a free download, which was made use of.

Figure 5 provides MSSIM values for various images as the LULU smoothing progresses through the DPT from scale $n = 1$ up to N . Notice how, based on the content of the images, the reduction in the MSSIM values as the DPT progresses varies from image to image. The graphs provide a mechanism to determine where

visual structure is in the image, that is, when the HVS would pick out structures of significance. Notice how the ocean image in Figure 5 contains very little structure and the MSSIM plot decreases gradually through the application of the DPT without any ‘occurrences’. The cat image however presents a number of phenomena in its MSSIM plot which indicate structure. Figure 6 indicates what is present at these scales. Scales 1 to 4030 represent the detail and the remaining scales the large relatively unimportant scales. The eye, face and forehead are represented at scales 4234 to 4235, 4325 to 4335 and 14565 to 14575 respectively.

III. CONCLUSION

We have presented an overview of the LULU operators and the resulting Discrete Pulse Transform for multidimensional arrays. As a new hierarchical decomposition the status of the DPT within the image processing community needs to be justified, thus we provide a characterization of the theoretical backbone of this nonlinear decomposition. This also provides a method of comparison for other nonlinear decompositions, which does not currently exist. The opportunity for further measures

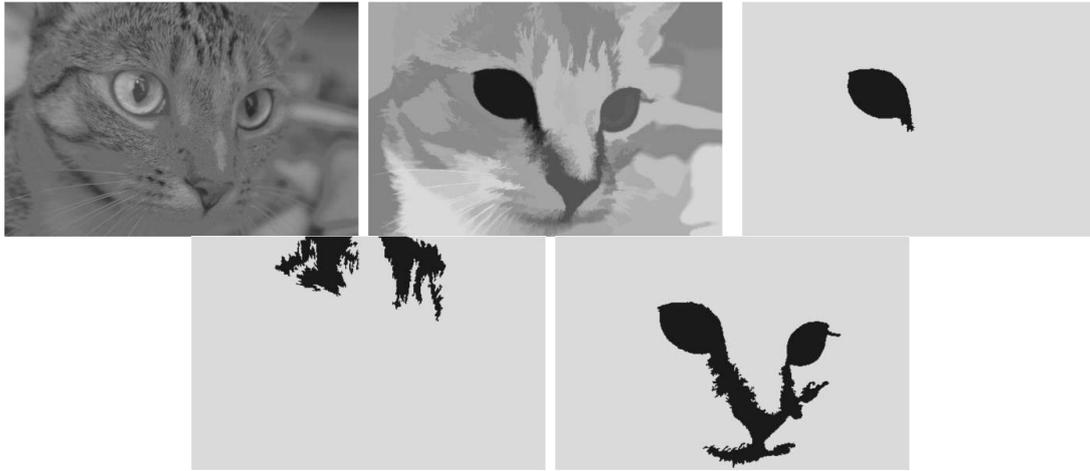


Fig. 6. Cat image at scales 1 to 4030, 4031 to 58571, 4234 to 4235, 4325 to 4335, and 14565 to 14575 respectively

of capability should not escape the reader, for example, the ability of noise removal for different types of noise contamination (independent as well as dependent noise structures), total variation preservation, and the nature of the pulse shapes and what the shapes mean relative to the image structure, image patterns and image content. Such work is currently under investigation.

ACKNOWLEDGMENT

The author acknowledges Dr S van der Walt whose collaboration ensured the implementation of LULU and the DPT in Python [15].

REFERENCES

- [1] R. Anguelov and I. N. Fabris-Rotelli, "Lulu operators and discrete pulse transform for multi-dimensional arrays," *IEEE Transactions on Image Processing*, vol. 19, no. 11, pp. 3012–3023, 2010
<http://dx.doi.org/10.1109/TIP.2010.2050639>.
- [2] C. Rohwer, *Nonlinear Smoothers and Multiresolution Analysis*. Birkhäuser, 2005.
- [3] J. Serra, *Image Analysis and Mathematical Morphology, Volume II: Theoretical Advances*, J. Serra, Ed. London: Academic Press, 1988.
- [4] I. Fabris-Rotelli, "Lulu operators on multidimensional arrays and applications," Masters Dissertation, University of Pretoria, November 2009.
- [5] C. Rohwer and M. Wild, "Lulu theory, idempotent stack filters, and the mathematics of vision of marr," *Advances in Imaging and Electron Physics*, vol. 146, pp. 57–162, 2007
[http://dx.doi.org/10.1016/S1076-5670\(06\)46002-X](http://dx.doi.org/10.1016/S1076-5670(06)46002-X).
- [6] I. Pitas and A. N. Venetsanopoulos, "Order statistics in digital image processing," *Proceedings of the IEEE*, vol. 80, no. 12, pp. 1893–1921, 1992
<http://dx.doi.org/10.1109/5.192071>.
- [7] M. Wild, "Idempotent and co-idempotent stack filters and min-max operators," *Theoretical Computer Science*, vol. 299, pp. 603–631, 2003
[http://dx.doi.org/10.1016/S0304-3975\(02\)00540-6](http://dx.doi.org/10.1016/S0304-3975(02)00540-6).
- [8] E. Tadmor, S. Nezzar, and L. Vese, "Multiscale hierarchical decomposition of images with applications to deblurring, denoising and segmentation," *Communications in Mathematical Sciences*, vol. 6, no. 2, pp. 281–307, 2008.
- [9] L. Florack, A. Salden, B. ter Haar Romeny, and J. K. andMA Viergever, "Nonlinear scale-space," *Image and Vision Computing*, vol. 13, no. 4, pp. 279–294, 1995
[http://dx.doi.org/10.1016/0262-8856\(95\)99716-E](http://dx.doi.org/10.1016/0262-8856(95)99716-E).
- [10] A. Wong and A. Mishra, "Generalized probabilistic scale space for image restoration," *IEEE Transactions on Image Processing*, vol. 19, no. 10, pp. 2774–2780, 2010
<http://dx.doi.org/10.1109/TIP.2010.2048973>.
- [11] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, pp. 629–639, 1990
<http://dx.doi.org/10.1109/34.56205>.
- [12] D. Laurie and C. Rohwer, "The discrete pulse transform," *SIAM Journal of Mathematical Analysis*, vol. 38, no. 3, 2007.
- [13] D. Laurie, "The roadmaker's algorithm for the discrete pulse transform," *IEEE Transactions on Image Processing*, vol. 20, no. 2, pp. 361–371, 2011
<http://dx.doi.org/10.1109/TIP.2010.2057255>.
- [14] Z. Wang, A. Bovik, H. Sheik, and E. Somincelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.
- [15] S. J. van der Walt, "Super-resolution imaging," PhD Thesis, Stellenbosch University, December 2010.