Chapter 1

Introduction

Gestalt theory sets a basis for vision perception [254, 130] but as it is non-computational it doesn't provide the required practical setting as well. David Marr, a British neuroscientist and psychologist, published the famous book known as the 'Vision of Marr' [138] in 1982 providing a modern theory for visual perception. Image processing is not a purely algorithmic field. The connection with artificial intelligence is undeniable and methods investigated inevitably require thought into how the human vision system operates and how it can be replicated. Marr realized this connection when writing his book and a number of earlier works [136, 137, 140, 139]. He defines vision as the 'process of discovering from images what is present in the world, and where it is', and defines three levels of the computer vision process (1) computational theory, (2) representation and algorithm, and (3) hardware implementation.

In [52] the computational theory of the LULU operators and the Discrete Pulse Transform (DPT) were presented in detail. Therein the algorithm was also presented in which the DPT is the successive slicing off of local maximum and minimum sets of increasing support size. Software implementation of the DPT has also progressed from MATLAB to Python, one dimension to two dimensions, and further implementation in three dimensions is envisioned. In [138, Chapter 1] it is described how advanced human vision uses both sides of the brain, the left side for shape vision and the right side to interpret the purpose of what is seen, dealing with these separately and then combining the interpretations. The Discrete Pulse Transform extracts discrete pulses of the image of every possible shape and size thereby setting the stage for an effective computer vision method.

As we move from the human vision system to develop a computer vision model, the first requirement is an operational definition of an image. Florack [61] provides this is some detail in the setting of algebras and groups, but the basic definition for practical purposes is the projection of a three dimensional environment into a two dimensional digitized (discrete) space. The process of acquiring an image is done by the capture of light onto a photographic film or more modernly as the conversion of a similarly obtained electric signal into a digital image. The most obvious acquirement process is with a normal digital camera but others such as infrared (long-wave), laser, night vision, and satellite cameras or capturing devices exist as well, including video capture which is simply sequences of images with the third dimension as time. The applications presented in this work are based on grey-scale images but can quite easily be extended to colour images as well as other forms. Methods presented would, however, need to be altered for the intrinsic manner in which different image types represent the image content but the basic ideas would still hold.

The main contributions of this thesis to the field of image analysis are as follows.

Development of the theory of Discrete Pulse Transform (DPT) for images

This follows the extension of the LULU operators to multiple dimensions which is presented in [52]. We prove that the properties if the DPT in one dimension can be generalised almost unchanged for the DPT on multidimensional arrays and in particular for images. The method of proof relies on the properties of morphological connectivity rather than that of finite sets of consecutive integers which makes proofs fundamentally different from the one dimension case. Further, they are applicable for any kind of connectivity since the only structure assumed for the domain is a morphological connection. The established properties of consistent decomposition and total variation preservation are applicable to any hierarchical decomposition and can be considered as a step towards a general formulation-independent theory of nonlinear hierarchical decompositions. Here we need to acknowledge that the property of strong consistency or the so-called highlight conjecture, formulated as an open problem for the one-dimensional DPT in 2005, was first proved by Dirk Laurie [112]. However, the basic consistency property stated here in Theorem 23 was proved first in our paper [8]. Moreover, the method of analysis of the DPT in this paper also led to the Highlight theorem thus providing an alternative proof of the result based on a morphological approach rather than one based on graphs as in [112].

DPT Scale-Space

Scale-space theory is an important approach to most of the problems in image analysis, e.g. feature detection, segmentation, noise removal, etc. We show that the DPT has a natural scale-space associated with it, namely the space of discrete pulses, which can be used successfully in addressing the mentioned problems. We further derived an axiomatic definition of a scale-space, which is an attempt to develop a unified scale-space theory encompassing those scale-spaces defined by integral operators like the ever popular Gaussian scale-space, and the morphological scale-spaces of which the DPT scale-space is a part. This work was published in [7].

Implementation and Application

The presented theory is developed ultimately for practical goals, thus its implementation and demonstration of usefulness in applications are its essential partner. This is particularly important for the DPT since it is very computationally intensive. Using an efficient computer algorithm for derivation and storage of the DPT pulses, the practical soundness of the DPT is investigated in image sharpening, best approximation of an image, noise removal in signals as well as images, feature point detection with ideas to extending work to object tracking in videos, and lastly image segmentation.

In Chapter 2 we summarize the LULU theory thoroughly presented in [52]. A look at characterizing the resulting Discrete Pulse Transform amongst nonlinear decompositions is presented Chapter 3.4. The Discrete Pulse Transform is connected to the frequently investigated scale-space theory in Chapter 4. A thorough review of the original Gaussian scale-space, a linear scale-space, is also provided as well as alternative scale-spaces and the progress of scale-spaces after the Gaussian scale-space. A formal, much needed, definition of a general scale-space is introduced together with a proof for the connection of the LULU scale-space via the Discrete Pulse Transform to this definition. In Sections 4.8.2 and 4.8.3 as well as Chapter 5 the basic image analysis techniques, namely feature detection, image segmentation and cleaning up of an image (sharpening and noise removal) are investigated. These provide an indication of the ability of the LULU operators and the Discrete Pulse Transform to perform the basics in image analysis. As a completely new theory, the foundation methods need to be investigated before more precise and more detailed methods are progressed towards. This work does just this.

Chapter 2

LULU Theory Background

The LULU and Discrete Pulse Transform (DPT) theory on sequences and multidimensional arrays was presented in detail in [52]. We repeat the work here in a summarized manner for completeness.

2.1 Setting

Let Ω be an abelian group, so that commutativity always holds. Recall that an Abelian group is an algebraic structure with a set, say G, and set operation, say $*$, satisfying the five axioms of closure (for all $a, b \in G \Rightarrow a * b \in G$), associativity ($(a * b) * c = a * (b * c)$), identity element ($\exists\ e \in G$ such that $a * e = e = e * a$ for all $a \in G$; e is called the identity element), inverse element (for each $a \in G$ there exists $b \in G$ such that $a * b = e = b * a$) and commutativity ($a * b = b * a$ for all $a, b \in G$) [79]. Denote by $\mathcal{A}(\Omega)$ the vector lattice of all real functions defined on Ω 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 $x \leq y \iff x_n \leq y_n\ \forall\ n \in \mathbb{Z}$.*

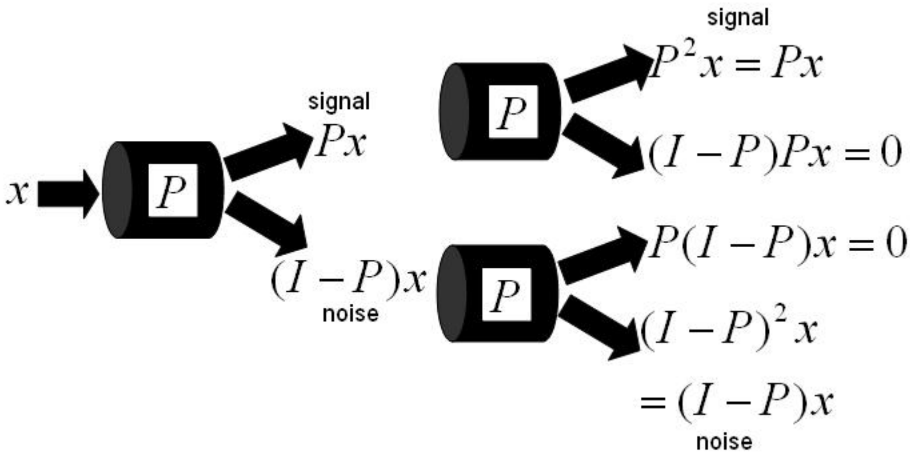


Figure 2.1: *The action of a separator P*

2.2 Separators

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 denoted the identity operator. For nonlinear filters this implication generally does not hold so the idempotence of $id - P$, also called co-idempotence, [243], can be considered as an essential measure of consistency.

For every $a \in \Omega$ the operator $E_a : \mathcal{A}(\Omega) \to \mathcal{A}(\Omega)$ given by $E_a(f)(x) = f(x-a)$, $x \in \Omega$, 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 2 define a **smoother**. More detail on smoothers can be found in [52].

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

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

Figure 2.1 illustrates the action of a separator P. It illustrates 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 the first time completely so that there is no 'signal' left in the 'noise' nor any 'noise' left in the 'signal'. The median, for example, smoother requires recursive application and thus does not possess this desirable property.

2.3 One Dimensional LULU

The LULU operators and the associated Discrete Pulse Transform developed during the last three decades or so are an important contribution to the theory of the nonlinear multi-scale analysis of sequences. The basics of the theory as well as the most significant results until 2005 are published in the monograph [183]. For more recent developments and applications see [5], [38], [106], [113], [184]. This 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

$$(L_n\xi)_i = \max\{\min\{\xi_{i-n}, ..., \xi_i\}, ..., \min\{\xi_i, ..., \xi_{i+n}\}\},\ i \in \mathbb{Z}. \tag{2.1}$$

$$(U_n\xi)_i = \min\{\max\{\xi_{i-n}, ..., \xi_i\}, ..., \max\{\xi_i, ..., \xi_{i+n}\}\},\ i \in \mathbb{Z}. \tag{2.2}$$

Figure 2.2[1] illustrates how the operators L_1 and U_1 affect a sequence x, by respectively lowering or raising a local maximum or minimum point to the value of its nearest neighbour.

It is shown in [183] 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

[1] Graphs are from collaborative work done with PJ van Staden and K van Oldenmark [56]

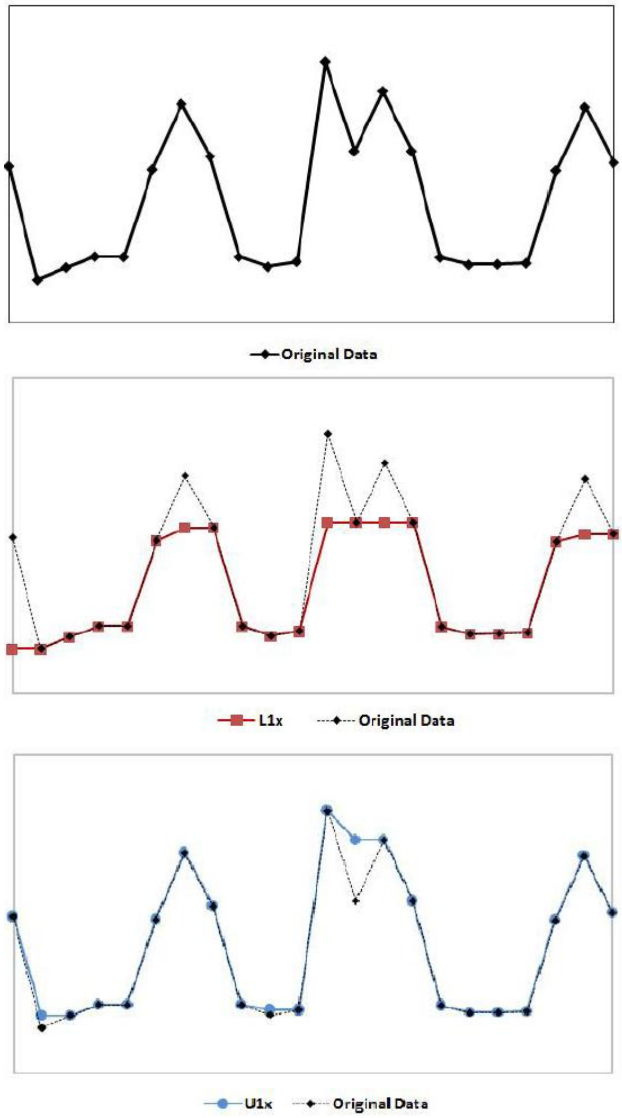


Figure 2.2: *An illustration of the effect of L_1 and U_1 on a sequence respectively*

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. This implies that these operators are total variation preserving.

2.4 One Dimensions to n Dimensions

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, [207].

Definition 3 *Let B be an arbitrary non-empty 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) for any family $\{C_i\} \subseteq \mathcal{C}$ *we have* $\bigcap_{i\in I} C_i \neq \emptyset \Longrightarrow \bigcup_{i\in I} C_i \in \mathcal{C}$.
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.

2.5 n Dimensional LULU

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 conditions assumed for the connection $\mathcal{C}$ we also assume that

$$\text{card}(\mathcal{N}_n(x)) < \infty,\ \forall n \in \mathbb{N},\ \forall x \in \Omega. \tag{2.3}$$

In image analysis ($d = 2$) the simplest and most obvious connectivity to make use of is a graph connectivity defined via a neighbour relation, e.g. 4-connectivity, 8-connectivity. However, in order to have maximum generality, we adopt the present axiomatic approach. Let us also mention that LULU operators on a continuous domain ($\Omega = \mathbb{R}$) are discussed in [9] and [5].

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*

$$L_n(f)(x) = \max_{V \in \mathcal{N}_n(x)} \min_{y \in V} f(y), \quad x \in \mathbb{Z}^d, \tag{2.4}$$

$$U_n(f)(x) = \min_{V \in \mathcal{N}_n(x)} \max_{y \in V} f(y), \quad x \in \mathbb{Z}^d. \tag{2.5}$$

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

$$\mathcal{N}_n(i) = \{\{i-n, i-n+1, ..., i\}, \{i-n+1, i-n+2, ..., i+1\}, ..., \{i, i+1, ..., i+n\}\}.$$

Hence for an arbitrary sequence ξ considered as a function on $\mathbb{Z}$ the formulas (2.4) and (2.5) are reduced to (2.1) and (2.2), respectively.

2.6 Properties

Matheron Pair

An essential property of L_n and U_n is that they form a Matheron pair [8], that is we have

$$L_n \circ U_n \circ L_n = U_n \circ L_n \text{ and } U_n \circ L_n \circ U_n = L_n \circ U_n. \tag{2.6}$$

Area Opening and Closing

The operators L_n and U_n are an area opening and area closing respectively. 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,

$$f \leq g \Longrightarrow (L_n(f) \leq L_n(g),\ U_n(f) \leq U_n(g)) \tag{2.7}$$
$$L_n \circ L_n = L_n, \quad U_n \circ U_n = U_n \tag{2.8}$$
$$L_n(f) \leq f \leq U_n(f) \tag{2.9}$$

Monotonicity

The operators are monotone with respect to n in the following sense,

$$n_1 < n_2 \Longrightarrow (L_{n_1} \geq L_{n_2},\ U_{n_1} \leq U_{n_2}). \tag{2.10}$$

Semigroup

The operators L_n, U_n and all their compositions form a four element semi-group with respect to composition. Moreover, this semi-group is fully ordered as follows,

$$L_n \leq U_n \circ L_n \leq L_n \circ U_n \leq U_n. \tag{2.11}$$

The semi-group is also a *band* which means that all elements are idempotent.

Separators

The operators L_n, U_n are separators for every $n \in \mathbb{N}$.

Action of the Operators

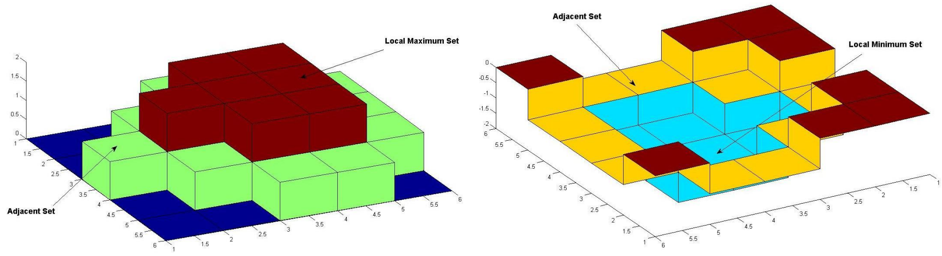
Similar to their counterparts for sequences the operators L_n and U_n defined for multidimensional arrays above smooth the input function by removing peaks (the application of L_n) and pits (the application of U_n). The smoothing effect of these operators is made more precise by using the concepts of a local maximum set and a local minimum set defined below.

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 $\mathrm{adj}(V)$, that is,*

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

Definition 7 *A connected subset V of $\mathbb{Z}^d$ is called a* **local maximum set** *of $f \in \mathcal{A}(\mathbb{Z}^d)$ if*

$$\sup_{y \in \mathrm{adj}(V)} f(y) < \inf_{x \in V} f(x).$$



(a) (b)

Figure 2.3: *(a) A Local Maximum Set (b) A Local Minimum Set*

Similarly V is a **local minimum set** *if*

$$\inf_{y \in \mathrm{adj}(V)} f(y) > \sup_{x \in V} f(x).$$

Figure 2.3 provides an illustration for the concepts introduced in Definition 41. In this figure the case of constant sets is presented. Although this is not required by Definition 41 it is illustrated as such because the Discrete Pulse Transform acts on such sets due to the mechanism employed in its application. The following theorem illustrates this.

Theorem 8 *For $f \in \mathcal{A}(\mathbb{Z}^d)$,*

a) $L_n(f)$ is constant on any local maximum set W of f with card$(W) \leq n+1$

b) $U_n(f)$ is constant on any local minimum set W of f with card$(W) \leq n+1$

We present two more theorems which illustrate the relationship between the LULU operators and local maximum and minimum sets.

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 $\mathrm{card}(V) \leq n$*;*

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 $\operatorname{card}(V) \leq n$.

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.

In summary the theorems 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 local minimum (maximum) sets of f 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 a consequence of the preceding results we obtain the following corollary.

Corollary 11 *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.

Neighbour Trend Preservation

Definition 12 *An operator P is* **neighbour 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 neighbour trend preserving.*

In Definition 12, for P to be fully trend preserving the requirement on $id-P$, that is the neighbour trend preserving property, can be equivalently formulated as

$$|P(f)(p) - P(f)(q)| \leq |f(p) - f(q)|. \tag{2.12}$$

The property (2.12) is called *difference reducing*.

Theorem 13 *The operators L_n, U_n, $n = 1, 2, ...,$ and their compositions, are all fully trend preserving.*

Total Variation Preserving

We assume for this section 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 neighbours of a point in the given context. For example, in image analysis ($d = 2$), it is common to use 4-connectivity (neighbours left, right, up and down) and 8-connectivity (in addition, the diagonal neighbours 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 $\{p_1, p_2, ..., p_k\} \subseteq C$ 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 (2.13)
- $(p, p + e_k) \in r$, for all $k = 1, 2, ..., d$ and $p \in \mathbb{Z}^d$, (2.14)

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

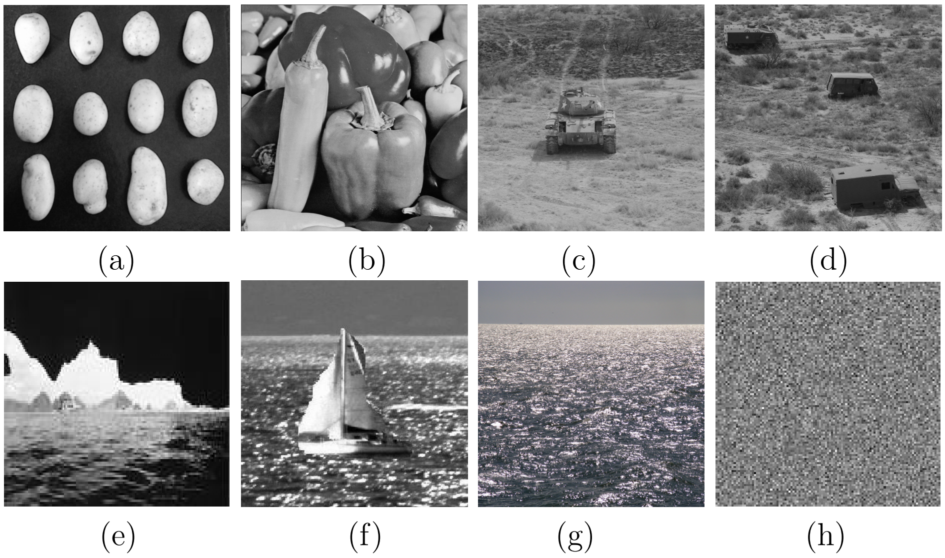


Figure 2.4: *Sample images*

Conditions (2.13) and (2.14) ensure that the set of connected sets $\mathcal{C}$ defined through this relation is a connection in terms of Definition 3. Condition (2.14) 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 [34] and [193]. It should be noted that there are several definitions of total variation for functions of multi-dimensional argument (Arzelá variation, Vitali variation, Pierpont variation, Hardy variation, etc., see [2] [36] [164]). In the applications cited above 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

$$TV(f) = \sum_{p \in \mathbb{Z}^d} \sum_{i=1}^{d} |f(p + (e_k)_i) - f(p)|. \tag{2.15}$$

If $TV(f) < \infty$, then f is said to be of *bounded variation.* Table 2.6 gives the total variation of a few sample images seen in Figure 2.4. Notice that the pure noise image, Figure 2.4(h), has the highest total variation and as the images contain more homogenous areas their total variation reduces.

Let us denote by $BV(\mathbb{Z}^d)$ the set of all functions of bounded variation in

Image in Figure 2.4	Total Variation (standardized)
(a)	109173
(b)	132527
(c)	167011
(d)	193650
(e)	235908
(f)	245480
(g)	386408
(h)	703707

Table 2.1: *Standardized Total Variation of Some Sample Images*

$\mathcal{A}(\mathbb{Z}^d)$. 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 $d = 1$ equation 2.15 gives the total variation of sequences as discussed in [183, Chapter 6]. Similar to sequences the total variation in equation 2.15 is a semi-norm. An operator P on $BV(\mathbb{Z}^d)$ is called *total variation preserving* if

$$TV(f) = TV(P(f)) + TV((id - P)(f)). \tag{2.16}$$

It is natural to expect that a good separator P will not create new variation as this property requires. An operator P satisfying property 2.16 is called *total variation preserving* [185].

Theorem 14 *The operators L_n, U_n, $n = 1, 2, ...,$ and all their compositions, are total variation preserving.*

2.7 Conclusion

This chapter provided a summary on the theory of the LULU operators developed for multidimensional arrays. We now proceed with the resulting Discrete Pulse Transform.