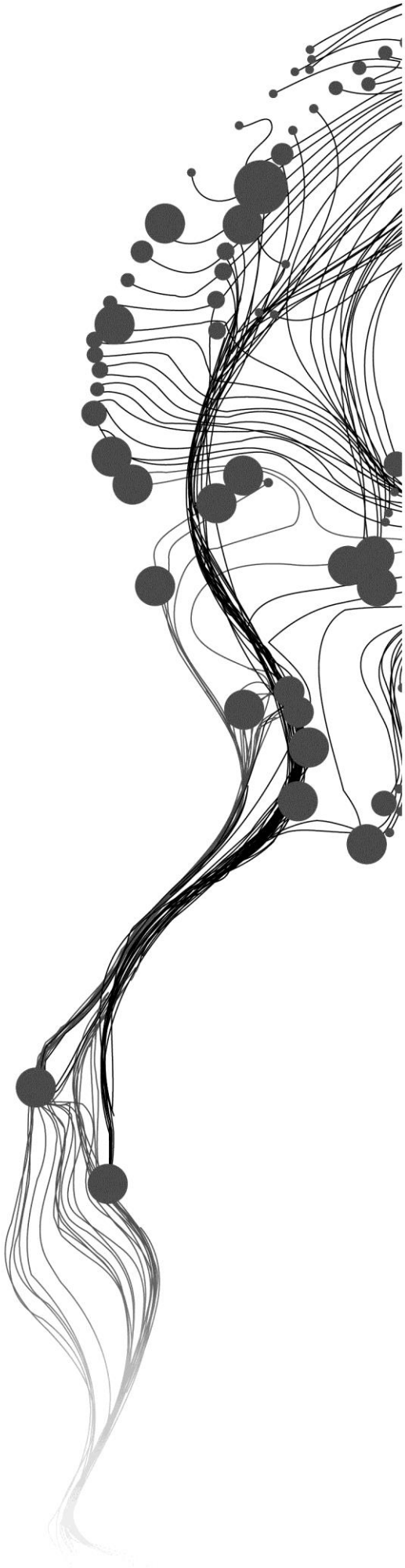
LULU OPERATORS FOR IMAGE SEGMENTATION AND OBJECT DETECTION

AMARE DEGEFAW NIGATU
February, 2013

SUPERVISORS:
Dr. V. A. Tolpekin
Prof. Dr. Ir. A. Stein



LULU OPERATORS FOR IMAGE SEGMENTATION AND OBJECT DETECTION

AMARE DEGEFAW NIGATU

Enschede, The Netherlands, February, 2013

Thesis submitted to the Faculty of Geo-Information Science and Earth Observation of the University of Twente in partial fulfilment of the requirements for the degree of Master of Science in Geo-information Science and Earth Observation.
Specialization: Geoinformatics (GFM)

SUPERVISORS:
Dr. V. A. Tolpekin
Prof. Dr. Ir. A. Stein

THESIS ASSESSMENT BOARD:
Dr. R. Zurita-Milla (Chair)
Dr.Ir. B.G.H. Gorte (External Examiner, Delft University of Technology)

DISCLAIMER

This document describes work undertaken as part of a programme of study at the Faculty of Geo-Information Science and Earth Observation of the University of Twente. All views and opinions expressed therein remain the sole responsibility of the author, and do not necessarily represent those of the Faculty.

ABSTRACT

Image segmentation remains a challenge task; most of the segmentation approaches may not be even applicable in remotely sensed image because of the complexity due to multi-spectral, multi-scale and heterogeneity properties. The aim of this study is to segment and automatically identify objects based on LULU operators recursive application, the Discrete Pulse Transform and the scale space analysis. The LULU-DPT algorithm, The Discrete Pulse Transform 2D algorithm and the scale-space automatic object identification algorithm are applied on a very high resolution image on urban areas of Cairo city, Egypt. Six areas are selected from different parts of the city. The first principal component analysis is applied and the subset images are decomposed by recursive application into one dimension and two dimensions discreet pulses. From man-made features, object of interest such as buildings and roads are automatically identified from pulses based on the scale space image analysis that generates the breaking points at which objects are drastically changed in size and shape, change in mean and standard deviation along the scale function. In multi resolution analysis, segments and objects are identified from the decomposed input image using DPT either partial reconstruction from selected pulses by identifying feature of interest and removing back ground features or fully reconstruction from the sum of all non-zero DPT levels. Identified objects are compared with outputs of thresholding segmentation method, complex urban features such as buildings and apartment blocks are easily identified by pulses. Segmentation accuracy assessment is resulted a high accuracy output; with area fit index for buildings 0.10 to 0.11 and for road 0.19 to 0.21. The maximum overlap between the reference data and segmented output is 90% for building and 81% for roads. The object boundary fit between the segmented object and reference is 0.7 after the correction factor 0.90 obtained and applied.

Keywords: LULU operators, DPT, MRA, scale-space, Area Fit Index

ACKNOWLEDGEMENTS

The Amazing Blessing of Almighty God; who made it possible to begin and finish.

I would like to express my most sincere gratitude to my first supervisors Dr. Valentyn Tolpekin, for his patience, encouragement, guidance, unreserved support and sharing of his great experience. You gave me strength, courage, and confidence throughout the research period. I would like to thank my second supervisor Prof. Alfred Stein for his motivation and inspiration, sharing kindly his knowledge and challenging me to go forward in this research.

Special thanks to Stefan van der Walt and his colleagues for making the Discrete Pulse Transform 2D source code available online as open source.

I would like to thank the Netherlands Fellowship Programme (NFP) for giving me the opportunity to study at ITC by support the financially matter.

I would like to thank all the academic staff members of GFM including the supporting staffs, especially my instructors for their guidance throughout the study program, inspiring lectures, sharing experience in research skills.

Special thanks to my friends Metadel Fentahun, Bezaye Tesfaye and Gezahegn Aweke for their inspiration and making my stay in the Netherlands grateful.

Manuel Garcia, Arun Poudyal, Federico Figueroa, Bayarmaa Enkhtur, Albert Anning, Parya Pasha, Fikerte Abebe, Saron Araya thanks for the great time I had in ITC. Studying in cluster, coffee breaks and lunches were fun. I wish you all the best for the rest of your life.

Studying at ITC gave me an opportunity to meet different people around the world, with different back ground and experience. I have learned a lot from this experience and I would like to thank ITC students and staff members.

Special thanks to my dearest friends Moges Asemamaw and Nishan Tegegn who are always been a great support for me, helping me in various aspects including moral support in my weakest moments.

Finally, I want to thank my parents Degefaw Nigatu and Bayushe Melaku. I am always grateful; they are always supportive and helpful in every aspect of my life. I would not have come this far without your love, support and help. I would like to thank my sister Hiwot Degefaw, my brothers Leulesegede Degefaw and Sileshe Degefaw. I would like to dedicate this thesis to my parents, sister and brothers. Thanks for all your love and encouragement.

God Bless You All.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTODUCTION

1.1. Motivation and problem statement

Currently technology advancements provided as tremendous volume of data that need to be processed to extract information for different purpose and applications. In line with information usage and voluminous of data, data manipulation is becoming more important and automated data analysis tools are increasing in quantity. Different data analysis tools are useful under different sets of assumptions for different purposes and applications.

Remotely sensed images are rich in information content. Currently, with the advancement of technology remote sensing images of different spatial and spectral resolutions are available for different applications. Very high resolution (VHR) multispectral satellite images such as RapidEye, IKONOS, QuickBird, GeoEye-1 and WorldView-2 are available for extraction of information and the spatial and temporal coverage of these satellites image tend to increase. VHR satellites image consists of visible, near infrared and panchromatic bands. The spatial resolution at nadir may range from 1.65 m to 5 m in multi spectral bands and 0.82 m to 0.41 m in panchromatic band. This provides enhanced capabilities in earth surface monitoring specially in urban applications, such as detailed urban land cover mapping

Remote sensing images have to be analysed to extract meaningful information relevant to the users. Richness of information and complexity of analysis questions, the analysis of remote sensing images is a challenging task. One of the basic and useful image analysis methods to extract information is image segmentation.

Segments are blocks of pixels that have similar characteristics. Such segments are generated by one or more criteria of homogeneity. Segmentation reduces complexity and removes superfluous details from images depending on the context definition of the segmentation (Stein and De Beurs, 2005). Many image segmentation algorithms exist and there is no proof for a single superior segmentation algorithm that performs better than other segmentation algorithm across a wide range of images and applications.

Segmentation is an active area of on-going research. Wang et al. (2009) discovered that there is lack of research on image segmentation algorithms for analysis of remotely sensed images. Despite the fact that recent efforts have been made in developing image segmentation methods appropriate for remote sensing image, new, more sophisticated and effective segmentation methods are required for more effective and accurate segmentation result on fine resolution image (Peijun et al., 2011). For example object oriented image analysis relies on accurate multi-resolution segmentation, segments of different size might be required to create objects.

An interesting, relatively novel approach to segmentation of remotely sensed image could be served by the LULU operators. LULU is LnUn (lower upper) and UnLn (upper lower) operators (Rohwer, 1989), also known as MaxMin and MinMax filters (Kao, 2001) introduced by Carl Rohwer in 1989.

LULU theory developed a nonlinear multi-resolution analysis based on the heuristic ideas from Fourier analysis and wavelet (linear) analysis that resulted in discrete pulse transforms (DPT). DPT is a sum of

pulses of a hierarchical decomposed input signals in which signals have a constant value in a connected set and zero elsewhere (Anguelov and Fabris-Rotelli, 2008; Rohwer and Wild, 2007). DPT based on LULU operators achieves all advantages of wavelets properties such as multi-resolution analysis, predictability, efficiency and locality. The application of multidimensional DPT based on LULU operators in remote sensing image analysis is not yet explored. In image analysis field, publications by (Anguelov and Fabris-Rotelli, 2010; Laurie, 2011) is shows that promising results.

Different domains of study deal with image segmentation in different approaches. Successful image segmentation remains a challenge task; most of these approaches may not be even applicable in remotely sensed images because of their complexity due to multi-spectral, multi-scale and different information content. In this research by considering properties of remotely sensed image, problems on image segmentation will be addressed based on the DPT method on VHR image emphasized the man-made features in urban area.

1.2. Research identification

Availability of huge amount of remote sensing data and the tendency to increase a VHR image requires an intelligent image segmentation method, which automatically extracts useful information for various applications. Therefore, an alternative relatively novel approach to segmentation of remotely sensed image could be served by the LULU operators.

1.2.1. Research objectives

The main objective of the proposed research is to explore LULU operators and associated DPT (LULU-DPT) for automatic object identification from the remotely sensed image.

This main objective can be reached by defining following sub-objectives:

1. To apply a segmentation method for remotely sensed image based on the LULU-DPT.
2. To apply LULU-DPT method for identification of objects from VHR multi-spectral remotely sensed image.
3. To apply LULU-DPT segmentation method for identification of objects based on multi-resolution analysis.
4. To apply scale space analysis for automatic identification of objects.
5. To evaluate the outputs of the LULU-DPT segmentation method.

1.2.2. Research questions

1. What are the essential properties of LULU operators and associated DPT for image segmentation?
2. How to apply the LULU-DPT segmentation method on multi-spectral remotely sensed image?
3. How to apply the multi-resolution analysis in the LULU-DPT segmentation method?
4. How the scale of DPT be used in the scale space analysis?
5. Which accuracy measures could assesses the segmentation result?

1.3. Workflow

The workflow of the study is presented in Figure 1.1; it briefly shows how the study is designed to address the objectives of the study based on the formulated research questions.

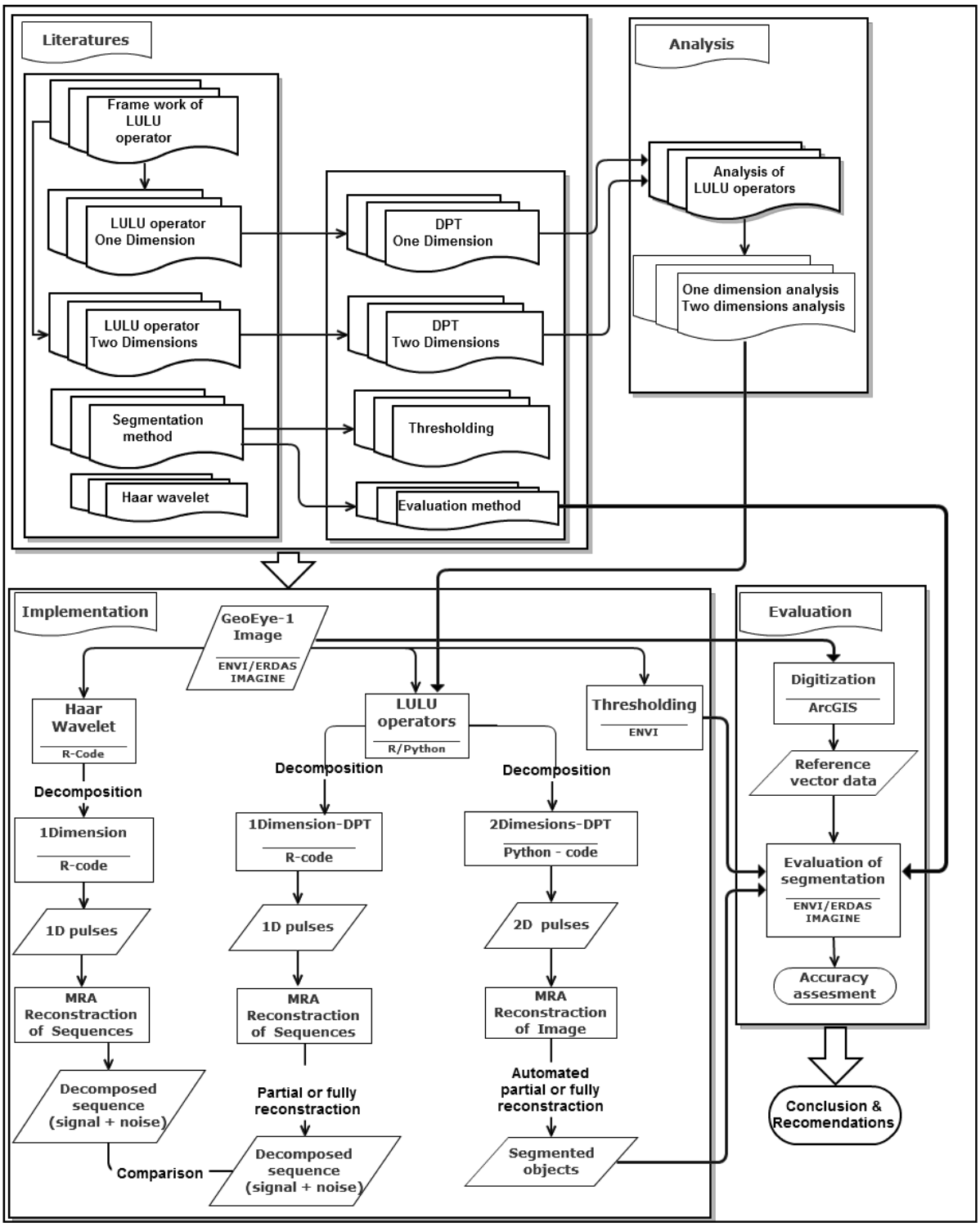


Figure 1. 1 The workflow diagram

2. LULU THEORY, IMAGE SEGMENTATION AND WAVELET

This chapter consists of four sections. The first section covers the theory and application of LULU operator for image segmentation and object detection as non-linear multi-resolution, multi-dimension, computationally efficient alternative decompositions method and its usefulness in Image analysis. The second section briefly reviews of image segmentation, the different types of segmentation method and its purpose. The third section reviewed about the wavelets in Multi-resolution analysis and the Haar wavelet in image processing. And the fourth section contains the related works.

2.1. LULU theory

2.1.1. Basis for LULU operators

As a background, smoothing and separation is explained by defining their properties and describing the concepts involved in the process of smoothing and separation. These properties of a smoother and separator are used to develop criteria to evaluate the smothers and separators such as LULU operators and they will have direct influence on the process of smoothing and separating a sequence of data. In section 2.1.1.1 definitions of sequences is given to relate with LULU operators. Definition of operators and their properties is presented in section 2.1.1.2. Criteria's for the design and comparison of smoothers and separators such as the concepts of effectiveness, consistency, stability, efficiency, idempotency, and co-idempotency are explained in detail in section 2.1.1.3 based on Rohwer suggestions for LULU operators (Rohwer, 2005).

2.1.1.1. Sequences

LULU theory is based on two operators L_n "lower" and U_n "upper"(Rohwer, 1989). These operators are smoothers and are applied to a bi-infinite sequence of real numbers.

Definition: Let X be the set of all bi-infinite and the sequences is defined as:

$$X = \{x = \{x_i\}: i \in \mathbb{Z}, x_i \in \mathbb{R}\}$$

In real problems these sequences are generally finite, but may be extended with zeros on both sides to make them infinite in length with the assumption that ℓ_1 is bounded, i.e. $\left|\left|x\right|\right|_1 = \sum_i |x_i| < \infty$. This sequence is more commonly referred to as a signal (du Toit, 2007; Fabris-Rotelli, 2009; Rohwer, 2005).

2.1.1.2. The Operators

Definition: Let $G(X)$ is the set of all operators on X:

$$G(X) = \{A: X \to X\}$$

Usual notations and operators on X are the following (du Toit, 2007; Rohwer, 2005).

Definition: Let $F(X)$ is the set of operators on X. Then:

(a) $(A+B)x = Ax + Bx,\ \forall x\ \in X$	Sum of operators
(b) $Ix = x,\ \forall x\ \in X$	Identity operators
(c) $Ox = 0,\ \forall x\ \in X$ and 0 the null sequence	Zero operators
(d) $(\alpha A)x = \alpha(Ax),\ \forall x\ \in X\ ,\ \alpha\ \in \mathbb{R}$	Scalar associativity
(e) $(AB)x = A(Bx),\ \forall x\ \in X$	Operator composition
(f) $(Ex)i = x_{i+1},$ for each i and $x\ \in X$	Shift operator
(g) $Nx = -X,\ \forall x\ \in X$	Negative operator
(h) $(A+B)C = AC + BC,\ \forall x\ \in X$	Right distributive
(i) $A^0 = I,\ \ A^{n+1} = AA^n,\ n\ \in \mathbb{Z}$	Operator power

Definition: Any operator,P, is called variation diminishing if:

$$T(Px) \leq T(x)$$

Where T is total variation(du Toit, 2007)

2.1.1.3. Smoothers and Separators

Rohwer (2005) described some of the appropriate simple criteria's for design and comparison of smoothing operators for separating some of signal from a sequence where the signal is contaminated by noise. These are:

Effectiveness - For each x, Px should be a signal and $(I-P)x$ noise, *I* is the identity operator. In other words, the effectiveness of the smoothers is the ability to reduce the noise present while maintaining the true signal.

Consistency - Signals should be preserved and noise mapped onto 0. After the signal and noise are extracted, each of them should be preserved i.e. $P(Px) = Px$ (idempotent) and $(I-P)\big((I-P)x\big) = (I-P)x$ (co-idempotent). If the signals and noise are preserved, then the smoother separates x consistently.

Stability - Small input perturbations in x should not distort output excessively, so that the signal is recovered well.

Efficiency – during computation the smoothers and /or the separator should be economical.

The axioms of Mallows can be taken as initial criteria for evaluation of smother operators to the principal smoother axioms and to yield the axioms of separators, an elite type of axioms can be added such as idempotence and co-idempotence (Rohwer, 1989, 2005).

The concepts of signal and noise should be translation independent on both axes for consistency; it is reasonable and necessary to demand these properties (Rohwer, 2005). This leads to the yield the first two essentially axioms of Mallows. It is also reasonable to demand the third axiom scale independence. This requires that $P(\alpha x) = \alpha P(x),$ for $\alpha \geq 0$. Mallows omits the restriction to $\alpha\ \geq\ 0$, restricting α to non-negative values allows scale independency for operators like LULU(Rohwer, 2002b, 2005). Moreover, if 0 is a signal it is also noise, since$(I-P)0 = 0$.

Axioms

Mallows Smoother Axioms (Rohwer, 2005): An operator P on X is a smoother if:

1. $PE = EP$ (Horizontal shift Invariance)

2. $P(x + c) = P(x) + c$, for each $x,\ c \in X$ such that c is a constant sequence. (Vertical shift Invariance)
3. $P(\alpha x) = \alpha P(x)$, for each $x \in X$ and scalar $\alpha \geq 0$ (Scale Invariance)

Separator Axioms: A smoother *P* is a separator if it also satisfies the additional axioms:

4. $P^2 = P$ (Idempotence).
5. $(I - P)^2 = I - P$ (Co-idempotence).

Rohwer (2005) defines a separator as operator when it holds the axioms of idempotent, co-idempotent and a smoother. LULU operators holds all the above criteria's and axioms that makes it powerful nonlinear smoother and separator (Rohwer, 1989, 2005; Rohwer and Wild, 2007).

A smoother can be regarded as device that separates signal and noise. For better result one can re-smooth the signal part for additional noise removal if there is any remaining noise during the first removal. It is also possible to put the noise removed in the first attempt to the smoother to see if it yields more signals. This process yields a signal $\big(P^2 + P(I - P)\big)x$ and noise$((I - P)^2 + (I - P)P)x$.

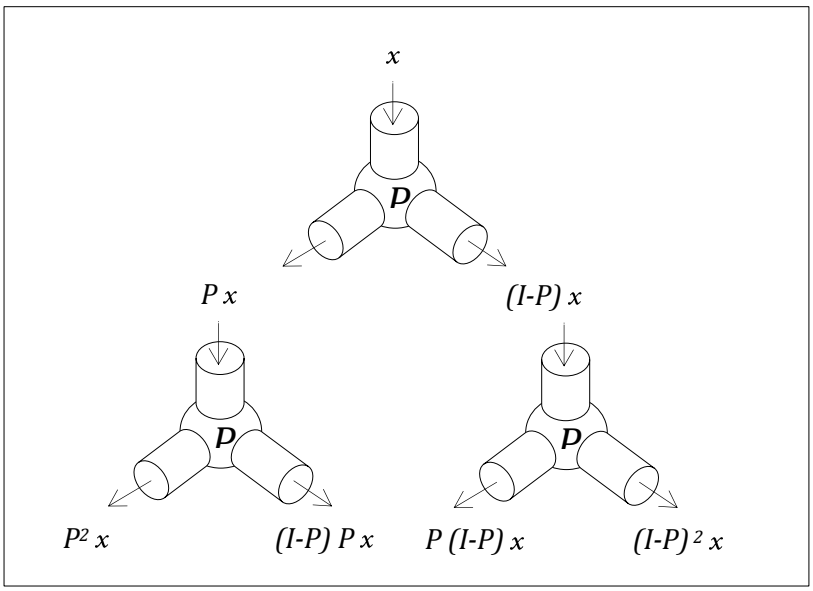


Figure 2. 1 A two-stage separator (source: (Rohwer, 2005))

2.1.2. LULU operators

To discuss the basics framework of LULU operator, the elementary building blocks such as protons and electrons (primitive smoothers) can be combined to build the basic operator that is the neutrally charged computational atom (Rohwer and Wild, 2007).

Definition: The elementary operators ∧ and ∨ on X are:

$$(\wedge x)_i = \min\{x_{i-1}, x_i\}$$
$$(\wedge x)_i = \max\{x_i, x_{i+1}\}$$

These elementary are selectors, therefore only values in the input sequence are allowed in the output sequence. The dilation operator ∨ is related to a union of sets because it widens a collection of sets that represents the maximum set covered by the collection. Similarly, the erosion operator ∧ can relate to an intersection of sets, which results in the minimum set covered by the collection. Particular interest is a

composition of the neutrally charged computational atom that consists of composition of equally many ∨ and ∧ (Rohwer and Wild, 2007), since ∨ will remove downward impulses and ∧ will remove upward impulses and its combination become less destructive than individual operators. In the sense of mathematical morphology, the operator yields the most relevant dilation-erosion pair.

Definition: for $n \geq 0$ put $L_n = \vee^n \wedge^n$ and $U_n = \wedge^n \vee^n$. Any finite composition of these will be called LULU operator (Rohwer and Wild, 2007).

2.1.2.1. The ranges of LULU operator

Definition: Let $\mathcal{M}_0 = \mathbb{R}^{\mathbb{Z}}$ be the setoff all sequences that can appear in the analysis. A sequence x is n-monotone if $\{x_i, x_{i+1}, \dots, x_{i+n+1}\}$ is monotone for each i.

Definition $\mathcal{M}_n$ is the set of all sequences x that are n-monotone.

In real analysis for a sequence the idea of local monotonicity has been a classical concept of smoothness in the interpretation of the popular median smoother (Rohwer and Laurie, 2006). Rohwer (1989) proved that the popular median smoother$(M_n x)_i = median\{x_{i-n}, \dots, x_i, \dots, x_{i+n}\}$ is included in the LULU interval $[U_n L_n,\ L_n U_n]$ since $U_n L_n \leq M_n \leq\ L_n U_n$ and this is crucial in the analysis and comparisons of large classes of smoothers. Obviously any sequence is 0-monotone i.e. belongs to $\mathcal{M}_n$, and then $\mathcal{M}_0 \supset \mathcal{M}_1 \supset \mathcal{M}_2 \supset \cdots$, and the set $\mathcal{M}_n$ is not a vector space except when n=0 (Rohwer and Laurie, 2006).

The following figure shows the key order relationship among the LULU operators.

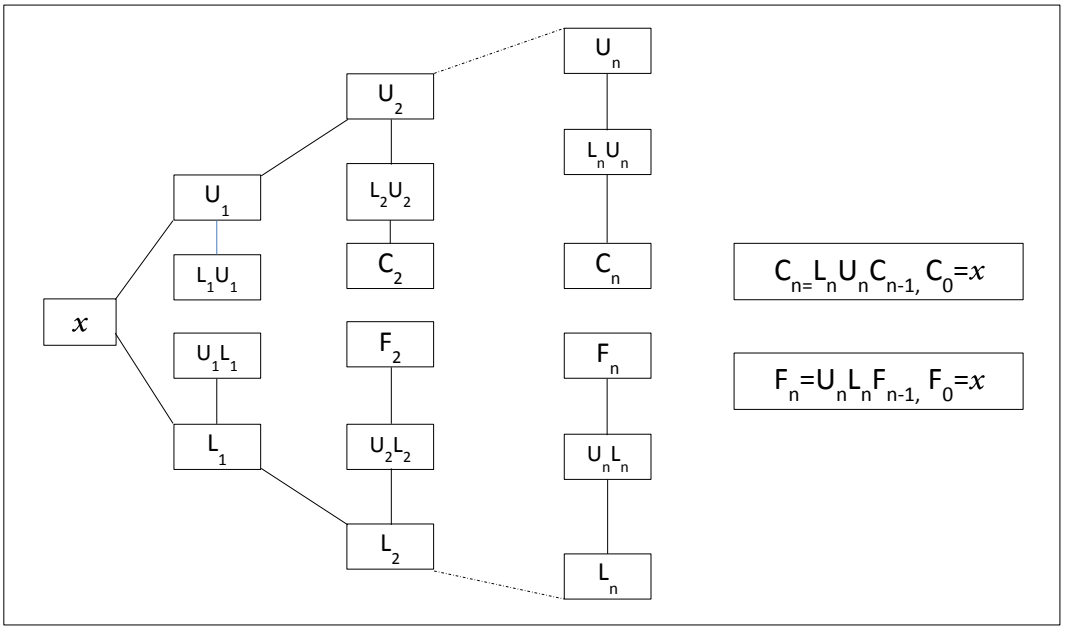


Figure 2. 2 The LULU operators order and components (Source: (Rohwer and Wild, 2007))

2.1.2.2. Shape preservation

Median smoothing is known for its good edge preservation with comparisons to linear smoothers in image processing. The report from the image processing expert showed that LULU decompositions do well, even better than the median smoothing and progressively more local types of trend preservation observed (Rohwer and Wild, 2007).

Trend preservation

Rohwer (2002a) investigate that the LULU operators exhibit strong trend preservation properties and this property defined as follow.

Definition: An operator P on X is neighbor trend preserving (NTP) if, for each sequence $x \in X$:

$$x_{i+1} \leq x_i \Rightarrow Ax_{i+1} \leq Ax_i, \qquad \text{and}$$

$$x_{i+1} \geq x_i \Rightarrow Ax_{i+1} \geq Ax_i$$

Local ordering cannot be inverted by neighbour trend preserving operator at any place in a signal. Rohwer and Wild (2007) shown the atom turned out to be easily NTP, in which no smoother involved in the decomposition ever change the order of neighbours. This indicate that after application of a neighbour trend preserving operator, any constant region in a signal will stay constant possibly with another value (du Toit, 2007). The NTP Therefore this is an important shape preserving property and this operator cannot map a sequence into a rougher (lower) smoothness class (Rohwer, 2002a).

An operator A is a dual of an operator B if $AN = NB$, the verification by Rohwer and Wild (2007) showed that L_n and U_n are dual to each other in the following equality equation.

$$L_n(-x) = -U_n x, \qquad \text{for all } x \in \mathbb{R}^{\mathbb{Z}}$$

Furthermore Rohwer and Wild (2007) noted that all the composition of LULU atoms and their complements are inherent to variation preservation. Their investigation shows that they are all NTP and they defined a strong property as follows.

Definitions: A separators S is fully trend preserving (FTP) if both S and $I - S$ are NTP.

Such separators trivially have that S and $I - S$ are difference reducing, and that

$$|x_{i+1} - x_i| = |Sx_{i+1} - Sx_i| + |x_{i+1} - Sx_{i+1} - x_i + Sx_i|. \text{ Therefore}$$

$$\sum|x_{i+1} - x_i| = \sum|Sx_{i+1} - Sx_i| + \sum|x_{i+1} - S_{i+1} - (x_i + Sx_i)|,$$

for any subset chosen for the summation have variation preserving (VP) (Rohwer and Wild, 2007).

Eventually, the two properties of FTP and VP are actually equivalent that one implies the other in the LULU theory(Rohwer and Wild, 2007).

2.1.2.3. Full Trend Preservation and Consistency

The NTP and FTP operators abounded in the LULU decomposition inherited by composition are the key ideas emerged to a simple proof structure for the observed consistency. For this the first fundamental consistency theorem and the second fundamental consistency theorem are proved by Rohwer (2005). In the decomposition of the primary separator, after a refined investigation into the consistency, the author convinced that the two above fundamental theorems are crucial in a theory of consistency (Rohwer and Wild, 2007).

For nonlinear operators consistency is explained in consistently separation of both signal and noise. For consistent separation the operator must be separator, that means both idempotence (for signal separation) and co-idempotence (the idempotence of $I - S$ or noise) are required(Rohwer and Laurie, 2006).

2.1.2.4. Total variation

Smoothing of a sequence x can be achieved by improving the continuity of sequence by two methods; one is that by increasing the monotonicity of the sequence x or equivalently by decreasing the variation with in the sequence. Monotonicity and total variation are used to measure the level of continuity of the original sequence as well as the resulting sequences in the LULU operator.

Definition: The total variation for a sequence x is given by,

$$T(x) = \left|\left|\Delta x\right|\right|_1 = \sum_{i=-N}^{N} |x_{i+1} - x_i|$$

Where the sum is from $-N$ to N since $x_i = 0$ for $i < -N$ and $i > N$ (du Toit, 2007; Fabris-Rotelli, 2009).

2.1.3. The Discrete Pulse Transform (DPT)

DPT is the result of the decomposition of a sequence x in the application of Multi-resolution Analysis (MRA) (Fabris-Rotelli, 2009). In MRA reconstruction or partial reconstruction of sequence can be done for the subset or for the entire decomposed sequence from the additive part of the decomposed sequence (Fabris-Rotelli, 2009).

A DPT represents a sequence as a sum of pulses, where a pulse is a sequence that is zero everywhere except for a certain number of consecutive elements that have a constant nonzero value (Rohwer and Laurie, 2006).

Definition: The DPT of a sequence x is given by:

$$D(x) = DPT(x) = [D_1(x), D_2(x), D_3(x), \dots, D_N(x), D_\alpha(x)]$$

$\mathrm{DPT(x)}$ is a decomposition of the input sequence x using by recursive application of either $L_n U_n$ or $U_n L_n$.

The DPT from LULU operator is based on the order relations between elements of the sequence and showed that it is comparable and computationally more efficient than the median transformation and it is more amendable to theoretical analysis(Rohwer and Laurie, 2006).

DPT based on LULU operators achieves all advantageous properties of wavelets such as: 1) Multi-resolution analysis at which signals are decompose as a sum of separate signals, each of which is meaningful at its own resolution level, 2) Predictability that makes it theoretically possible to predict its behaviour, the DPT is consistent to recover the same components of the original sequence 3) Efficiency: computationally very efficient procedures and 4) Locality: the component sequences are local in a precisely definable sense(Rohwer, 2002a, 2005; Rohwer and Laurie, 2006). In addition DPT based on LULU operators have properties like Incisiveness which makes it to behave at least as well as the median transforms in the presence of discontinuities(Rohwer and Laurie, 2006).

There are two main types of operators with in LULU *(ceiling and floor)*, when LULU applies iteratively on a sequence for $n = 1,2, \dots, n$ (Rohwer, 1989).

Definition: The C_n and F_n operators are given by

$$C_0 = L_0U_0 = x = U_0L_0 = F_0,$$

$$C_n = L_nU_nC_{n-1}, \quad F_n = U_nL_nF_{n-1}$$

A sequence decomposed by L_nU_n into the resolution sequences by $r^n = (I - L_nU_n)C_{n-1}x,$ for $n = 1,2, \dots$ and for a sequence decomposed by U_nL_n is $r^n = (I - U_nL_n)F_{n-1}x,$ for $n = 1,2, \dots$

The following figure shows the sequential separation in the DPT.

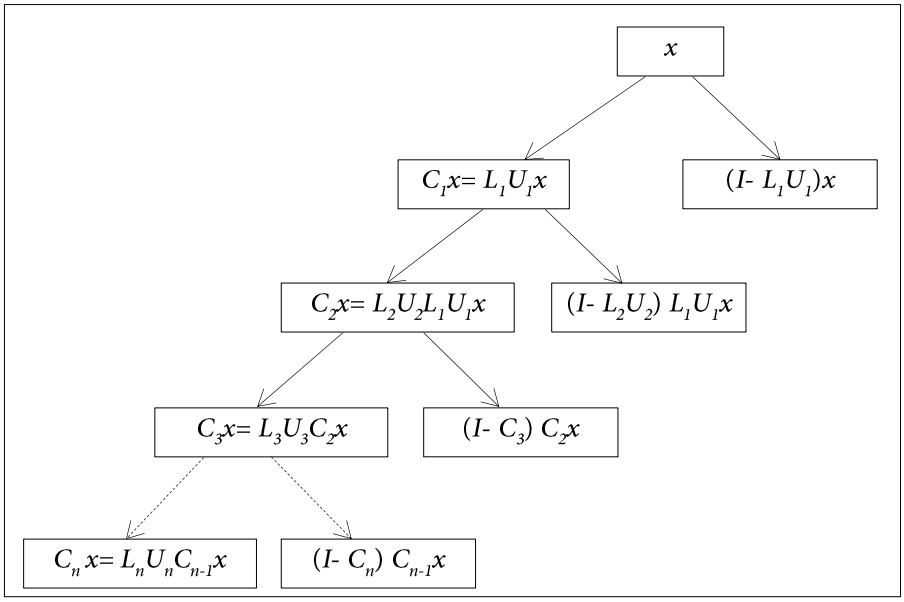


Figure 2. 3 A sequence separation and the DPT (source:(Rohwer and Laurie, 2006)).

2.1.4. LULU theory on Multi-dimension analysis

For the analysis of multi-dimension data LULU operator precisely extended including all the properties from one dimensional operator such as consistent separation (noise from signal), i.e., for operator P the separator has horizontal and vertical shift invariance, scale invariance, idempotent property (P *o* P=P), and co-idempotent ($(I - P)^2 = I - P$ where I is identity); total variation and shape preservation (Anguelov and Fabris-Rotelli, 2010). The one-dimensional LULU operators are applied to an image by Rohwer and Wild (2007) by decomposing each row and columns of the image separately. After processing done on the rows, construction of the image from these rows was performed. The same can be done for all columns. The final image is created by calculating the mean image of the horizontal and vertically analysed image. This is not a complete extension into two dimensions analysis (Rohwer and Wild, 2007). Anguelov and Fabris-Rotelli (2010) present the extension of LULU operator in two dimension by giving emphasis on the structure preserving properties of LULU theory such as consistent separation (i.e. noise from signal), consistent hierarchical decomposition, total variation and shape preservation.

Consistent separation are absorbed into the concept of a separator, hence the idempotent $(P^2 = P)$ and co-idempotent $(I - P)^2 = (I - P)$, *I is identity operator)* properties are hold by the operator in addition to the horizontal shift invariance, vertical shift invariance and the scale invariance (Anguelov and Fabris-Rotelli, 2010).

The total variation only satisfied three axioms from the four axioms of a norm (Anguelov and Fabris-Rotelli, 2008; Rohwer, 2004; Rohwer and Wild, 2007):
(1) $TV(f) \geq 0$,
(2) $TV(\alpha f) = \alpha\, TV(f)$ and
(3) $TV(f + g) \leq TV(f) + TV(g)$ but not the axiom $TV(f) = 0 \iff f = 0$.
TV is a unit for a quantity of information present (Rohwer, 2005).

So for any operator $P: A(\mathbb{Z}) \longrightarrow A(\mathbb{Z})$, $TV(f) \leq TV(P(f)) + TV(I - P)(f))$, a good separator has $TV(f) = TV(P(f)) + TV(I - P)(f))$ (Anguelov and Fabris-Rotelli, 2010).

An operator *P* when it satisfies $TV(\alpha f) = \alpha\, TV(f)$, it is called total variation preserving (Rohwer, 2002b). Shape preservation property closely linked to the total variation preservation and generally shape preservation refers to the preservation of edge in the input (Anguelov and Fabris-Rotelli, 2010). Preservation of trend in one dimension sequence is actually preservation of shape and a fully trend preserving operator on sequence is also total variation preserving.

The consistence hierarchical decomposition of nonlinear filters characterization is problematic for the quality of hierarchical decomposition. The derived LULU-DPT hierarchical decomposition is in fact closely related to that of stack filters (Anguelov and Fabris-Rotelli, 2010).

2.2. Analysis of LULU operators and DPT

2.2.1. LULU operators and DPT in one dimension

The LULU operator for one dimension signals (sequences) have been developed over the last two decades by Rohwer and his collaborators (Rohwer, 1989, 2005; Rohwer and Wild, 2007). For a signal $x = (x_i)_{i=-N}^{N}, \mathrm{i} \in \mathbb{Z}$, the LULU operators L_n and U_n at position i in the signal and for $n = 1, 2, 3 \ldots$ is:

$$(L_n x)_i = \max\{\min\{x_{i-n}, \ldots, x_i\}, \ldots, \min\{x_i, \ldots, x_{i+n}\}\},\ i \in \mathbb{Z}$$

$$(U_n x)_i = \min\{\max\{x_{i-n}, \ldots, x_i\}, \ldots, \max\{x_i, \ldots, x_{i+n}\}\},\ i \in \mathbb{Z}$$

Based on the definition Rohwer (2005) of the LULU operators are nonlinear, multi-resolution analysis separator which having a very useful properties such as consistent separation, fully trend preserving and total variation preserving. Since $L_n(x) \leq x \leq U_n(x)$, when the two operators used individually L_n smooth the signal from above and U_n smooth the signal from below produce slightly different results(Rohwer, 1989). Thus use of the two compositions together as either $L_n U_n$ or $U_n L_n$ minimizes the difference of the output (Rohwer, 2005).

The Discrete Pulse Transform (DPT) is an application in Multi-resolution analysis (MRA) and obtained by iterative application of $L_n U_n$ or $U_n L_n$ operator on a sequence x. The Discrete Pulse Transform is a mapping of a sequence x into a vector. The aim is to decompose a sequence into a sum of component sequences (Rohwer, 2005; Rohwer and Laurie, 2006).

Each component sequences representing the different resolution levels, $r^{(n)} = D_n\ (x)$, such that, for a finite sequence of length N,

$$x = (x_{i,}\ i = 1,2,3, \dots, N)$$

$$x = D_1(x) + D_2(x) + D_3(x) + \cdots + D_N(x) + D_\alpha(x)$$

Which can be write as

$$D(x) = DPT(x) = [D_1(x), D_2(x), D_3(x), \dots, D_N(x), D_\alpha(x)]$$

If $x = \sum_{i=0}^{N} \alpha_i\, D_i(x)$, then $D_n(x) = \alpha_n D_n(x)$ for each $n = 0,1, \dots, N$

Since the final sequence (resolution level) is constant and equal to zero if $x \in \ell_1$ space, the final resolution label is given as $D_\alpha(x)$ rather than $D_{N+1}(x)$. $DPT(x)$ is a decomposition of x obtained using either $L_n U_N$ or $U_n L_n$ (Fabris-Rotelli, 2009).

The LULU smoothers in a natural further extension, when iteratively apply to $n = 1, 2, \dots, n$ it leads to two main type of smoothers - ceiling and flooring (Rohwer, 2005), and described as follows:

$$C = (C_1, C_2, C_3, \dots), \qquad C_n = L_n U_n C_{n-1}, \qquad C_0 = x, \qquad n = 1,2, \dots, N$$

$$F = (F_1, F_2, F_3, \dots), \qquad F_n = U_n L_n F_{n-1}, \qquad F_0 = x, \qquad n = 1,2, \dots, N$$

In general, $F_n x \leq C_n x$ in applications and the information given by the pulse decomposition ($DPT(x)$) is obtained from the differences of $C_{n-1}x - C_n x, \quad F_{n-1} - F_n$.

Apply $L_1 U_1$ to x. Then

$$x = (L_1 U_1)x + (I - L_1 U_1)x = S_1(x) + D_1(x)$$

$S_1(x)$ is the 'smoother' sequence and $D_1(x)$ is the noise removed by $L_1 U_1$. The first component of the DPT is then D_1.

Apply $L_1 U_1$ to $S_1(x)$. Then

$$S_1(x) = (L_2 U_2)x + (I - L_2 U_2)S_1(x) = S_2(x) + D_2(x)$$

$S_2(x)$ is the 'smoother' sequence (even smoother than $S_1(x)$ and $D_2(x)$ is the noise removed by $L_2 U_2$. The second component of the DPT is then D_2 (Fabris-Rotelli, 2009).

Continue this decomposition until $L_n U_n$ is applied where N is the size of the signal x. This last application will result in

$$S_{N-1}(x) = (L_N U_N)x + (I - L_N U_N)S_{N-1}(x) = S_N(x) + D_N(x)$$

S_N is a constant sequence and denoted by $D_\alpha(x)$ and D_N is then the second to last member of the *DPT*. This is so that, the Rohwer and Laurie (2006) decomposition contains all the information in the original sequence and the original sequence can be reconstructed using the *DPT*.

The original sequence x can then be reconstructed using the *DPT* as follows(Rohwer and Laurie, 2006):

$$x = \sum_{n=1}^{N} D_n(x) + D_\alpha(x)$$

2.2.2. LULU operators and DPT in two dimensions

The one-dimensional LULU operators act on sequences in the space ℓ_1. In this space we have an ordering of the elements, namely x_{i+1} follow x_i and x_{i-1} proceeded by x_i. It is then normal to consider the elements x_{i+1} and x_{i-1} as the neighbours of x_i. The sequences involve maxima and minima over sets of consecutive terms (Fabris-Rotelli, 2009). The one-dimensional LULU operators act specifically on the n-neighbourhoods of each x_i is,

$$\{x_{i-n}, x_{i-n+1}, \dots, x_i\}, \dots, \{x_i, x_{i+1}, \dots, x_{i+n}\}$$

In two dimensions we would like to have operators, which logically extend from one dimension by preserving all the properties present in one dimension and having the equivalent higher dimensional concepts. Since $\mathbb{Z}^d, d > 1$, is only partially ordered the concept of consecutive does not make sense in this setting. To achieve this (Anguelov and Fabris-Rotelli, 2008, 2010; Fabris-Rotelliy and van der Walt, 2009) make use of the concept of connectivity in the extension using the concept of morphological set connection that was introduced by Serra (Serra, 1982, 1988).

To analyse the connectivity concept first let describe the axiom for connectivity class.

Let $\boldsymbol{C}$ is a connectivity class or a connection on $P(E)$ if the following axioms hold:

i. $\emptyset \in \boldsymbol{C}$
ii. $\{x\} \in \boldsymbol{C}$ for each $x \in E$
iii. For each family $\{\mathsf{C}_i\}$ in $\boldsymbol{C}$ such that $\cap\, \mathsf{C}_i \neq \emptyset$, we have $\cup\, \mathsf{C}_i \in \boldsymbol{C}$

A set $\mathsf{C} \in \boldsymbol{C}$ then C is the called connected (Anguelov and Fabris-Rotelli, 2010; Fabris-Rotelliy and van der Walt, 2009).

A set of every size is needed in the definition of operators by assuming that the set $\mathbb{Z}^d$ is equipped with the connection $\boldsymbol{C}$ satisfies (1) $\mathbb{Z}^d \in \boldsymbol{C}$, (2) for any $b \in \mathbb{Z}^d, E_b(\mathsf{C}) \in \boldsymbol{C}$ whenever $\mathsf{C} \in \boldsymbol{C}$ and (3) If $V \subsetneq W, V, W \in \boldsymbol{C}$, then there exist $x \in W \backslash V$ such that $V\{x\} \cup \in \boldsymbol{C}$ (Anguelov and Fabris-Rotelli, 2010).

Given a point $x \in \mathbb{Z}^2$ and $n \in \mathbb{N}$ we denote by $K_n(x)$ the set of all connected sets of size $n+1$ that contain point x is,

$$K_n(x) = \{V \in \boldsymbol{C} : x \in V, card(V) = n+1\}$$

Where $card(V)$ is the number of elements in the set V. In image analysis, the connectivity is defined on a graph via a neighbour relation (Figure 3.1), e.g., 4-connectivity or 8-connectivity (Anguelov and Fabris-Rotelli, 2008; Fabris-Rotelliy and van der Walt, 2009). Now the operators L_n and $U_n$are defined on $A(\mathbb{Z}^d)$ for an image $f(x)$ on the domain $\mathbb{Z}^d$ is as follows. $A(\mathbb{Z}^d)$ is the vector lattice of all real functions defined on $\mathbb{Z}^d$.

L_n and U_n on two dimension is a function of $f(x)$ at $f \in A(\mathbb{Z}^2)$ and $n \in \mathbb{N}$. Then

$$L_n(f)(x) = \max_{V \epsilon K_n(x)} \min_{y \epsilon V} f(y), \qquad x \in \mathbb{Z}^d$$

$$U_n(f)(x) = \min_{V \epsilon K_n(x)} \max_{y \epsilon V} f(y), x \in \mathbb{Z}^d$$

The operators L_n and U_n defined for smoothing the input function by removing sharp peaks by applying of L_n and deep valleys by applying of U_n (Anguelov and Fabris-Rotelli, 2008).

Thus the Discrete Pulse Transform is obtained by the successive removal of peaks (local maximum sets) and valleys (local minimum sets) from the image by applying L_n and U_n respectively. The smoothing and/or separating effect of these operations is made more precise by using the concepts of a local maximum set and a local minimum set is given below (Anguelov and Fabris-Rotelli, 2008).

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

$$adj(V) = \{x \in \mathbb{Z}^2 : x \notin V, V \cup \{x\} \in \boldsymbol{C}\}$$

A connected subset V of $\mathbb{Z}^2$ is called a local maximum set of $f \in A(\mathbb{Z}^2)$ if the least upper bound (supremum) set of $f(y)$ is less than the greatest lower bound (infimum) of set $f(x)$.

$$\sup_{y \, \epsilon \, adj(V)} f(y) < \inf_{x \, \epsilon \, adj(V)} f(x)$$

Similarly V is a local minimum set if

$$\inf_{y \, \epsilon \, adj(V)} f(y) > \sup_{x \, \epsilon \, adj(V)} f(x)$$

To derive the DPT of image from the above definitions of LULU operator can be done by defining a grayscale image that is given through a function f on a rectangular domain $\Omega \subset \mathbb{Z}^2$, with the value of f is the brightness (DN values) at the respective pixel (Anguelov and Fabris-Rotelli, 2008). For more convenient assumption the function can be defined in the whole space $\mathbb{Z}^2$ by defining f on the set $\mathbb{Z}^2 \backslash \Omega$ as a constant, example 1, by considering the set $A(\mathbb{Z}^2)$. Hence neighbours of a pixel are appropriate connections of image that are defined through a relation r on $\mathbb{Z}^2$in the given context. Figure 3.1 shows some of the neighbours of the pixel (i, j).

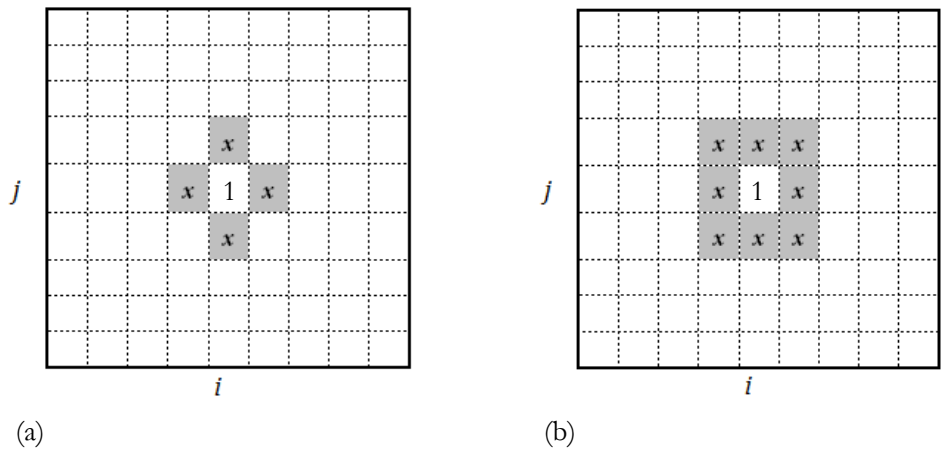


Figure 2. 4 The pixel neighbours of (i, j), (a) 4 - connected pixels (b) 8 - connected pixel

Anguelov and Fabris-Rotelli (2010) describe a neighbourhood in the connected set as: if for any two pixels $p, q \in \mathsf{C}$ in a set $\mathsf{C} \subseteq \mathbb{R}^2$ connected; there exists a set of pixels such that each pixel is neighbours to the next one. The assumption is that the neighbour relation r on $\mathbb{Z}^2$, such that r is reflexive, symmetric and shift invariant and $\big((i,j),(i \pm 1,j)\big) \in r$ and$((i,j),(i,j \pm 1)) \in r$, for all $i,j \in \mathbb{Z}$.

If $N = card\big(supp(f)\big)$, that is, the size of the image. Let us denote by $A(\mathbb{Z}^2)$ the set of all functions defined on $\mathbb{Z}^2$ which have finite support. The DPT of a grey scale image is a function $f \in A(\mathbb{Z}^2)$ derived by applying iteratively the operators L_n and U_n with n increasing from 1 to N. Here N is the number of pixels in the image (Anguelov and Fabris-Rotelli, 2008; Rohwer and Laurie, 2006).

$$DPT(f) = [D_1(f), D_2(f), D_3(f), \dots, D_N(f)]$$

Where the components of (1) are obtained through

$$D_1(f) = (\mathrm{I} - \mathrm{P}_1)(\mathrm{f})$$

$$D_n(f) = (\mathrm{I} - \mathrm{P}_n)Q_{n-1}(\mathrm{f}), \quad n = 2, \dots, N$$

And $\mathrm{P}_n = \mathrm{L}_n\mathrm{U}_n$ or $\mathrm{P}_n = \mathrm{U}_n\mathrm{L}_n$ and $\mathrm{Q}_n = \mathrm{P}_n\mathrm{P}_{\mathrm{n}-1}, \dots, \mathrm{P}_2\mathrm{P}_1, \quad \mathrm{n} \in \mathbb{N}$

A function $\Phi \in A(\mathbb{Z}^2)$ is called *a pulse* if there exists a connected set V and a real number β such that

$$\Phi(\mathrm{x}) = \begin{cases} \beta & if \quad x \in V \\ 0 & if \quad x \in \mathbb{Z}^2 \setminus V \end{cases}$$

The set V is the support of the pulse Φ that is $supp\,(\Phi) = V$.

$$D_n(f) = \sum_{s=1}^{\gamma(n)} \Phi_{\mathrm{ns}},$$

The function $\Phi_{\mathrm{ns}}, \ \mathrm{s} = 1, \dots, \gamma(n)$ is being discrete pulses with support of size $n, \ n = 1, \dots N$. $\gamma(n)$ is the number of pulses at $D_n(f)$. The discrete pulse is a function a function $f \in A(\mathbb{Z}^2)$ which is zero everywhere except on a connected set V (Anguelov and Fabris-Rotelli, 2008). We thus obtain the following decomposition of the image $f \in A(\mathbb{Z}^2)$,

$$f = \sum_{n=1}^{N} D_n\,(f) = \sum_{n=1}^{N} \sum_{s=1}^{\gamma(n)} \Phi_{\mathrm{ns}}$$

Where each Φ_{ns} is a discrete pulse of size *n,* which can be positive or negative, and all discrete pulses of fixed size n have disjoint supports. The value of Φ on V is the value of the pulse. For $n_1, n_2, s_1, s_2 \in \mathbb{N}$ such that $n_1 < n_2, 1 \leq s_1 \leq \gamma(n_1)$ and $1 \leq s_2 \leq \gamma(n_2)$ (Fabris-Rotelliy and van der Walt, 2009; Rohwer and Laurie, 2006), it has that

$$card(supp(\big(\Phi_{n_1 s_1}\big) = n,$$

$$supp(\Phi_{n_1 s_1}) \cap supp(\Phi_{n_2 s_2}) \neq \emptyset$$
$$\Rightarrow supp(\Phi_{n_1 s_1}) \subseteq supp(\Phi_{n_2 s_2})$$

The Discrete Pulse Transform decomposes a signal into a collection of pulses. In one dimension, a pulse is characterized by its start and end position, as well as by its amplitude. In two dimensions, a pulse describes a connected region over which function values are constant i.e., 8-connection, where two function values are equal in the upper-lower or left-right directions.(Anguelov and Fabris-Rotelli, 2010; Fabris-Rotelliy and van der Walt, 2009).

2.3. Image Segmentation

Currently, remote sensing images provide a huge amount of data about the surface of the earth that is available for different applications such as earth surface monitoring and management. Simultaneously with the advancement of technology the availability of high resolution satellite image increases by volume. With the fact that, information richness of remotely sensed image and voluminous of data, there is a need to a method for image analysis to extract useful information. In image analysis, one of the basic steps in extraction of information is image segmentation.

At present, many image segmentation algorithms exist and there is no proof for a single superior segmentation algorithm that performs better than other segmentation algorithm across a wide range of images. Almost all methods target at a specific single application and there is no one single method that is suitable to all the application (Bhanu et al., 1995; Thakare, 2011). For example nature-inspired techniques focusing on multi-objective clustering and classification approaches have a number technique for different applications involves in medical-related image, natural, remote sensing, and simulated or handcrafted image (Bong and Rajeswari, 2011). There are also examples from computer vision that are specific to applications; for example, Vision Research laboratory develops many automated image segmentation methods, which are depending on the application requirement (Vision Research lab-UCSB, 2008).

Already 20 years ago, Fu and Mui (1981) categorized the image segmentation method into three categories, (1)characteristic feature thresholding or clustering, (2) edge detection, and (3) region extraction. After a decade Pal and Pal (1993) reviewed and summarized existing methods of segmentation into six categories such as gray level thresholding, iterative pixel (which include relaxation, MRF based approaches and neural network based approaches), surface based segmentation, segmentation of colour images, edge detection methods based on fuzzy set theory (which includes fuzzy clustering, fuzzy thresholding and fuzzy edge detection). Most of image segmentation algorithms types are fallen on those three Fu and Mui (1981) categories. The detail of thresholding is reviewed in the following sections.

2.3.1. Thresholding

In image processing thresholding is the simple and effective method in separating objects from the background (Sezgin and Sankur, 2004). This method is used widely in image segmentation (Fu and Mui, 1981). The gray levels of pixels that belong to the object are different from that of the background. During image processing, the gray levels of pixels are grouped into two classes, the object and the background in binary image. In most of the application the foreground objects are represented by gray-level 1 with white colour, whereas the background image is represented by gray-level 0 with black (Sezgin and Sankur, 2004).

Based on the information they are exploiting Sezgin and Sankur (2004) categorize the thresholding methods into six groups.

1. Histogram shape-based methods- in this method the thresholds are selected based on the shape of the histogram and peaks, valleys and curvatures of the smoothed histogram are analysed.

2. Clustering-based methods - in this method clustering of the gray-level samples is applied to cluster into object and background.

3. Entropy-based methods result in algorithms that use the entropy of the foreground and background regions, the cross-entropy between the original and binarized image,

4. Object attribute-based methods – in this method the segmentation is based on the searched criteria such as fuzzy shape similarity, edge coincidence etc. which are a measure of similarity between the gray-level and the binarized images.

5. The spatial methods use higher-order probability distribution and/or correlation between pixels

6. Local methods - this method adapt the threshold value on each pixel such as range, variance, or fitting parameters to the local image characteristics.

The results of thresholding method depending on the diversity of features present in the image. If two distinct features are present in the mage, the segmentation gives good results. Thresholding method is noise sensitive than other methods(Fu and Mui, 1981)

2.4. Wavelets

Wavelets transform is a transformation of signals into components of the signals, for example in Sub-band filtering in signal processing signals are decomposed into frequency bands, in time-frequency analysis signals are decomposed into time and frequency component and in multi-resolution analysis signals are decomposed into different level of resolution (Sweldens and SchrSder, 2000). Wavelet transformation plays an important role in image processing. Wavelets used in image analysis to denoising of signal(Ogden, 1997).

Multi-resolution analysis

In image processing the analysis and extraction of information content from the image based on the brightness value of pixel at different scale level is the efficient way of extraction of information. During the application of MRA the size of the target object defines the resolution level which measures the local variation of the image. Interpretation of image based on scale- invariant is effective in multi-resolution analysis (Mallat, 1989).

A multi-resolution decomposition provides logically represented structured output information for interpretation of an image. The detail information of the image varies depending on the resolution level and the resolution levels determine the physical structure of the objects in the scene (Mallat, 1989). The highest course resolution level provides large structures which determine the overall context of the image. When the course resolution level down to the fine resolution level detail structures in the scene are extracted.

Multi-resolution analysis (MRA) is a unique mathematical theory that integrates and combines different image processing techniques such as sub-band coding (a set of band-limited components), pyramidal image processing (a collection of decreasing resolution images), and quadrature mirror filtering (Acharya and Ray, 2005; Mallat, 1989). Mainly the MRA is used to obtain different approximations of a function $\mathrm{f(x)}$ at different levels of resolution (Acharya and Ray, 2005; Lee, 2007).

In MRA there are two main functions, the mother wavelet function $\psi(x)$ and the scaling function $\varphi(x)$. In the analysis of the image at different level of decomposition the scaling function is used to approximate the image function in MRA (Acharya and Ray, 2005).

The scaling functions from the integer translations and binary scaling are:

$$\varphi_{j,k}(x) = 2^{j/2}\varphi\left(2^{j}x - k\right), \qquad j \; and \; k \in \mathbb{Z}, \text{ and } \varphi(x) \in L^2(\mathbb{R})$$

The position of $\varphi_{j,k}(x)$ is determined by k parameter and the width of $\varphi_{j,k}(x)$ is determined by j parameter along the x-axis. In the wavelet subspace for a specific j the expansion set is:

$$V_j = \underset{k}{Span}\left\{\varphi_{j,k}(x)\right\}$$

The scaling function must fulfil the following four requirements in MRA: (1) orthogonal to its integer translates, (2) higher resolutions must contain the lower resolution in the subspaces spanned $V_{-\infty} \subset \cdots \subset V_0 \subset \ldots \subset V_\infty = \{0\}$, (3) common function to all V_j is $f(x) = 0$, i.e. $V_{-\infty} = \{0\}$ and (4) when a function satisfies the upward complementary property $V_\infty = \{L^2(\boldsymbol{R})\}$. The Wavelet function is a function that satisfied the four requirements of MRA for scaling functions (Acharya and Ray, 2005; Lee, 2007). The function $\psi(x)$ that combined the integer translation and binary scaling is

$$\psi_{j,k}(x) = 2^{j/2}\psi\left(2^{j}x - k\right), \quad k \in \mathbb{Z}$$

The space W_j span is

$$W_j = \underset{k}{Span}\left\{\psi_{j,k}(x)\right\}$$

Haar wavelet

Haar wavelet is one of the simplest and 'First Generation Wavelets' (Sweldens and SchrSder, 2000) family. The function of Haar wavelet (Haar, 1910) is given by:

$$\psi(x) = \begin{cases} 1, & 0 \le x < \dfrac{1}{2} \\ -1, & \dfrac{1}{2} \le x < 1 \\ 0, & otherwise \end{cases}$$

Haar wavelets has an advantage on filtering or smoothing by taking two neighbouring samples of a sequence, i.e. the average s and the difference d of a neighbouring sample a and b in the following equations (Sweldens and SchrSder, 2000).

$$s = \frac{a+b}{2}$$
$$d = b - a$$

Always recover a and b as:

$$a = s - d/2$$
$$b = s + d/2$$

For a signal $\mathrm{s_n}$ of 2^{n} sample values$\mathrm{s_{n,l}}$:

$$s_n = \{\mathrm{s_{n,l}} | 0 \leq l < 2^{\mathrm{n}}\}$$

Each pair, $a = s_{2l}$ and $b = s_{2l+1}$ for $2^{\mathrm{n}-1}$ pairs ($l = 0 \dots 2^{\mathrm{n}-1}$) apply the average and difference transform gives:

$$s_{n-1,l} = \frac{s_{n,2l} + s_{n,2l+1}}{2}$$
$$d_{n-1,l} = s_{n,2l+1} - s_{n,2l}$$

Haar wavelet decomposition can be applied to a sequence to project onto the smoother sequence and its complement. In the wavelet subspace the projection of x onto the smoother component and the additive component on the first higher level represented by $P_1 x$ as the smoothed sequence and $Wx = (I - P_1), I$ is identity, as the additive component by defining P_1,

$$(P_1 x)_{2i} = \frac{1}{2}[x_{2i-1} + x_{2i}] = (P_1 x)_{2i-1}$$

represented by $\frac{1}{2}N$ coefficients in the wavelet (Rohwer and Wild, 2007).

2.1. Related Work

In image processing, data analysis involves in recognition and extraction of parts of the data such as separation of signals, noise, impulsive noise and outliers, or smoothing of image depends on the purpose of the analysis. Due to a number of reasons such as atmospheric disturbance, strong electromagnetic field, transmission errors, images are often corrupted by impulsive noises which are characterized by short, abrupt alterations of the grey or colour values in the image (Kao, 2001). These noises are significantly different from their local neighbourhood and can be treated by applying some filtering methods.

Linear smoothers traditionally were used for smoothing. In the presence of signal dependent noise, linear smoothers do not remove impulsive noise or outliers effectively and tend to results blurred edge during image processing (Pitas and Venetsanopoulos, 1990). A number of researches on the use and on the

investigation of the behaviour of non-linear filters motivated due to the lack of linear filters in treating adequately the impulsive noise or outliers(Rohwer, 1989). Nonlinear smoothers have been studied extensively in the literature and regularly used in practice to treat impulsive noise or outliers. These smoothers have been designed to meet criteria such as robustness, adaptability to noise probability distributions, preservation of edge information and image details (Jankowitz, 2007). Nonlinear filters are able to reduce noise level without simultaneously blurring edges; however, their theoretical foundations are far less secure (Glasbey and Horgan, 1995).

Carl Rohwer (1989) introduced LULU theory as non-linear operator and investigated surprising properties that performs their task in a prescribed and predictable way. The LULU operators nonlinearity is theoretically secured, the output is not significantly affected by small changes in the input, which preserves the signal carefully and effectively separate the noise and signal (Rohwer and Laurie, 2006). His work on these operators continued developing into an extensive theory in (Rohwer, 1999, 2002a, 2002b, 2004, 2005, 2007). Mainly in year 2005 the author briefly described and discussed the LULU theory and the practical use of LULU smoothers leading up to a full multi-resolution analysis of any finite sequence (Rohwer, 2005). In year 2007 with collaboration to Marcel Wild, Rohwer presented a work on LULU theory and the need of mathematics of vision in the field of image processing viewed from wavelet theory and mathematical morphology (Rohwer and Wild, 2007).

Furthermore, researches related to LULU operators are: locally monotone robust approximation of sequences (Rohwer and Toerien, 1991), the comparison of LULU operator with median filters in one dimension (Rohwer and Wild, 2002), the statistical aspect and distributions (Conradie et al., 2006; Jankowitz, 2007), the LULU operators on a continuous argument and domain (Anguelov, 2006; Anguelov and Rohwer, 2009), the LULU semi group for δ envelopes (Malkowsky and Rohwer, 2004).

Applications in image processing are currently being investigated using LULU operators via the associated DPT (Laurie, 2011). There are some motivations done in image analysis application, Kao (2001) successfully removed impulse noise during the smoothing process of an image by implementing modify LULU operators for preservation of image details. Study by Anguelov and Fabris-Rotelli (2008) showed a decomposition of the original image into pulses and they performed a partial reconstruction and identifying pulses of interest. Anguelov and Fabris-Rotelli (2010) also provided illustrative applications to problems and they proposed to investigate its applicability in a scale space theory, granulometries, image compression, pattern recognition and image segmentation. In addition, the associated Discrete Pulse Transform studied in (du Toit, 2007; Rohwer and Laurie, 2006), application of the associated Discrete Pulse Transform for image (Anguelov, 2008; Fabris-Rotelliy and van der Walt, 2009) and application of LULU operators for images (Fabris-Rotelli, 2009).

Carl Rohwer developed the LULU smoothers for one dimensional sequence over the last two decades, later on these operator extend to multi dimension array such as two dimension image in which all the properties for the one dimensional smoothers are achieved (Anguelov and Fabris-Rotelli, 2008, 2010; Fabris-Rotelliy and van der Walt, 2009), based on morphological concept of set connection (Serra, 1982, 1988). LULU operators are powerful nonlinear operator, which is characterized as smoothers and separators; with properties of consistent separation, total variation preservation, shape preservation and consistent hierarchical decompositions (Anguelov and Fabris-Rotelli, 2010; Fabris-Rotelli, 2009; Rohwer, 2005).

Study by (Rohwer, 2002b) explained the properties of the LULU smoother in the preservation of a norm, that is the total variation of a sequence becomes a natural norm, and is a measure of smoothness. This is

important in image processing, where total variation is used as an appropriate norm (Anguelov and Rohwer, 2009). In addition, (Rohwer and Wild, 2002) proved that LULU operators are computationally convenient, conceptually simpler and provide significant insight into the removal of impulsive noise.

The most important application of LULU operator is the Discrete Pulse Transform which is obtained via recursive application of the operator. It results in a multi-resolution decomposition of the sequence in one dimension and decomposition of image in two dimensions that is into a number of scale/decomposition levels. At each resolution level different structure of the sequence and image represented. Currently some researches are investigating the application of these powerful operators, its properties, as well as its associated Discrete Pulse Transform for simple applications in image processing specifically smoothing, image segmentation and feature detection.

The theoretical foundation of the nonlinearity of LULU operator is more complicated than the linear filters in multi-dimension (Rohwer, 2005); however, by taking into considerations of the complexity of transformation or conversion of data (nonlinear) from two dimensional images to human eye or measuring instrument, it is more logical to use image analysis via nonlinear operators than that of linear Operators(Rohwer and Wild, 2007)

In conclusion, much work has been done on LULU operators and the associated DPT. Image segmentation applications on VHR image is a relevant case study to test and evaluate the potential and efficiency of this operator and then to discover the relationship with other segmentation methods.

3. STUDY AREA AND DATA DESCRIPTION

In this chapter of the study the description of the study area and the material used explained to implement the LULU operators. To achieve the objective of the study and to answer the related research questions, a VHR image is chosen and the LULU operators applied on this datasets. A VHR image such as GeoEye-1 has been chosen for this study by taking into consideration such as the spatial resolution and the multi-spectral bands of the image. In the studying of urban areas focusing on detection and identification of urban features it is important to have an image of high resolution with a multi-spectral band which can explain the variation across the scene.

3.1. The study area: Cairo, Egypt

Cairo is one of the largest cities in Africa and in the Middle East, it a capital city of Egypt. Cairo city is geographically located at a 30^0 03' N $31^0$14' E. The greater Cairo has three municipalities Cairo, Giza and Qaliyabaya which makes the city as typical mega-city. Some of the ancient Egyptian civilization, the Great Sphinx and the Pyramids of Giza around the Nile River are parts of the Great Cairo.

In the long history of the city, numerous kinds of transformation experienced that shape the spatial structure and extent of the city (Gertel and Samir, 2002). Currently the city has a complex urban structure that represents the modern day urban structure and post-modern structure such as high-rise buildings, centuries old buildings and slum areas. The Great Cairo city faces a number of problems, one of the problems is informal settlement areas which keep expanding and becoming denser than the formal settlement areas based on the Egypt's restrictive building codes requirements (the plots minimum areas, heights of buildings, the buildings density, architectural appearances, etc.) (Howeidy et al., 2009). In addition, according to some studies Cairo is one of the highest populated city in the world with 32, 000 people live in 1 km^2 (Gertel and Samir, 2002). In 2006 and 2012 the great Cairo was inhabited by 16 million and 17 million peoples respectively, and in 2020 and 2050 it will be inhabited by 20 million and 30 million peoples respectively (Tarbush, 2012).

In 2008, the Egyptian government began promoting a "Vision of Cairo 2050" that aims to transform the Cairo city into a global city like Paris or Tokyo in terms of social and spatial structure(Tarbush, 2012). The spatial structure aimed in urbanization of the Great Cairo which include improving and building of houses, infrastructures and managing of land uses (Tarbush, 2012). The urbanization of a city in terms of the spatial structure can be studied using remote sensing images.

One of the basic methods in extraction of useful information about urban spatial structure from the VHR remote sensing image is image segmentation. The urban areas of the city characterized by manmade urban features such as various kinds buildings; different types roads, parking lots and other features is vital in studying of urban expansion and rehabilitation, pollution and intra-city variations (neighbourhood delineation for poverty, health and infrastructure study). An accurate and efficient method of image segmentation facilitates the extraction of information for further analysis depending on the purpose of the study. In this study, LULU operators are implemented by considering its multi-resolution analysis property which preserves the edges of the shape of the objects to identify man-made objects in the city.

3.2. Image

GeoEye-1 provides a very high resolution image for earth observation. Images are available at 0.41 meter panchromatic and 1.65 meter multispectral bands at nadir point. The GeoEye-1 multispectral has four bands in Blue, Red, Green and NIR. The availability of GeoEye-1 satellite image in high spatial resolution simplifies the extraction of information especially for urban applications.

The following GeoEye-1 image represents the south-west part of the great Cairo city captured at July 02, 2009. Figure 3.1 shows GeoEye-1 image for the parts of the great Cairo city and the areas of interest for implementation.

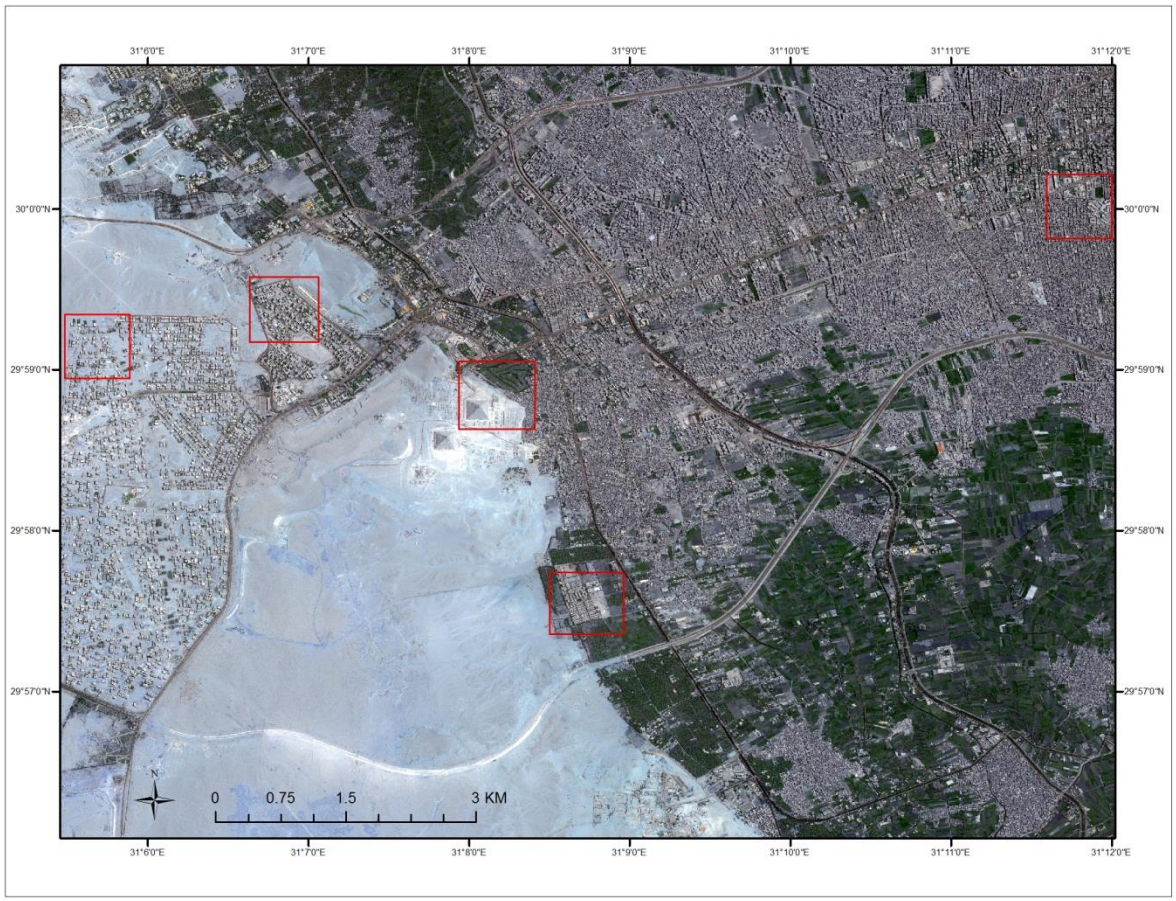


Figure 3. 1 Part of Cairo city: the study area, from GeoEye-1 image.

3.3. Selected areas

To implement the LULU operator for the entire image is extremely large and time consuming. For this reason six areas of interest are selected and subset from the provided GeoEye-1 image. These subset images are extracted from the different parts of the image to represent the urban structure variability. Figure 3.2, Figure 3.3, and Figure 3.4, shows the subset images. In addition, the description of the subset image is presented as follow:

Subset-1: in subset-1 small buildings are sparsely exist, relatively to the other subsets there is a wide open space around the building and the road.
Subset-2: relatively the buildings are large and have complex structure. The open space around the building is very small. The road is relatively wider than the subset-1 road and accesses all the buildings.

Figure 3. 2 Subset-1 (left) and Subset-2 (right) image

Subset-3: subset-3 combines both the simple and complex structure of urban areas such as large buildings with complex structure and medium and small buildings. There is also a major road which splits the scene into parts.
Subset-4: this subset mainly consist the Giza Pyramid. There are also small patches of houses and large green area far from the pyramid.

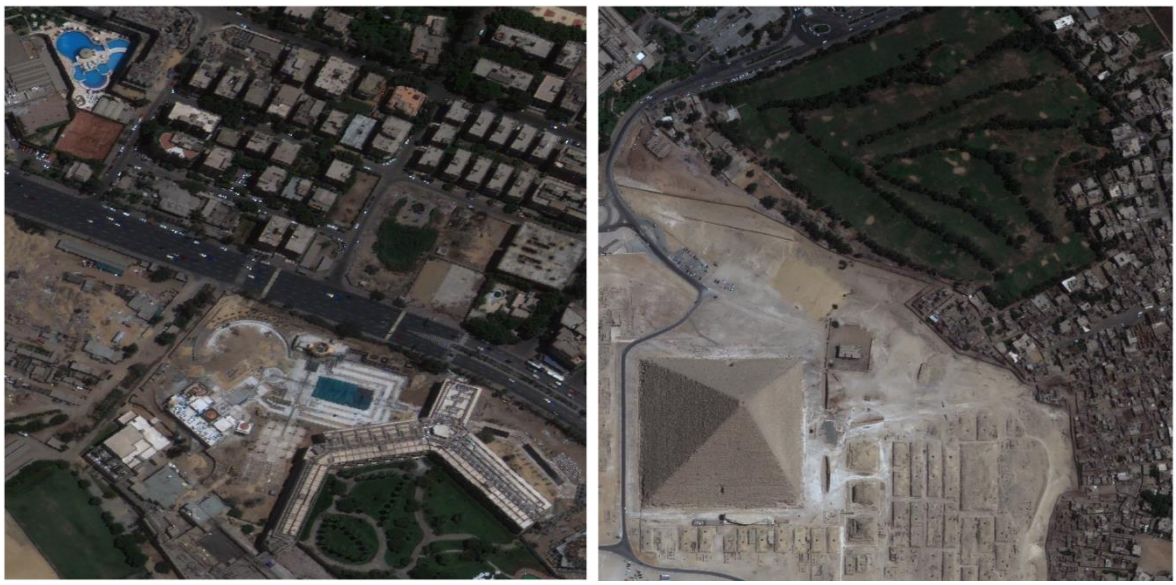


Figure 3. 3 Subset-3 (left) and Subset-4 (right) image

Subset-5: this subset is relatively away from other urban structure and it is surrounded by large green areas. There are also some small and medium complex buildings.
Subset-6: this scene is a combination of complex urban structures such as very small building, medium and large buildings with major and minor roads.

Figure 3. 4 Subset-5 (left) and Subset-6 (right) image

Reference dataset

For evaluation of the accuracy assessment to the implemented segmentation method, the reference data sets such as polygons of buildings and poly-line roads was created by manual digitization from the GeoEye-1 image using ArcGIS software.

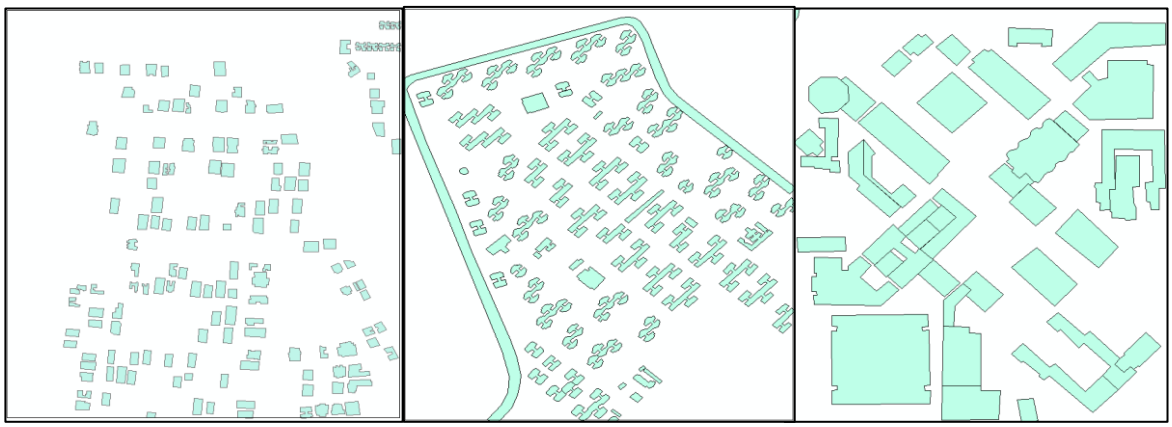


Figure 3. 5 References for Subset-1, Subst-2 and Subset-6 image

4. METHOD

This chapter describes the method used to achieve the objectives and to answer the research questions of the study. It is prepared based on the review of literatures in chapter two with the aim of developing a method for segmentation of image and identification of objects from the remotely sensed image based on the LULU operator.

4.1. The general approach

To achieve the objectives of the study and to answer the corresponding research questions the adopted approach is implemented into four major parts: Literature review, analysis, implementation and evaluation.

Literature review: the relevant literatures are reviewed to understand:

- The definition, properties and characteristics of LULU operators in one and two dimensions.
- The associated Discrete Pulse Transform (DPT).
- The application of LULU-DPT in one and two dimensional arrays.
- The segmentation methods used in different applications.
- The application of LULU-DPT on multi-dimension image for image segmentation.
- The approaches for evaluation of the segmentation output.

To implement the method and to identify objects based on LULU operators; methods that were used by other researchers in one and two dimensional signal analysis are reviewed. Based on these methods the properties of the operator and the associated DPT are studied to understand how the operator works on remotely sensed images. In addition, techniques used for evaluate of segmentation accuracy assessment are reviewed and used to assess the outputs from this method. Overall, the detail of this part of the study is briefly presented in chapter two.

Analysis: the detail analysis of the LULU operators and the associated DPT for one and two dimensions are briefly presented in chapter two, section 2.2.

The implementation of the method and the evaluation of the outputs are discussed in section 4.2 and 4.3.

4.2. Implementation

A number of researches implement the LULU operator in different algorithms, i.e. du Toit (2007) applied the discrete pulse transform algorithm, Anguelov (2008) applied Fast Pulse Transform algorithm, Fabris-Rotelliy and van der Walt (2009) applied Discrete Pulse Transform in 2D algorithm, Fabris-Rotelli (2009) and Laurie (2011) applied the Roadmaker's algorithm.

In this study the LULU-DPT algorithm is scripted and implemented in one dimension application based on the theory of LULU operators and the associated DPT in R-programming language. In two dimensions implementation, the Discrete Pulse Transform in 2D source code (Fabris-Rotelliy and van der Walt, 2009) is used and modified in Python programming language to decompose the fine resolution remote sensing image in Geo-tiff format. In addition, for automated identification of objects from the decompose image scale space analysis based algorithm is developed based on the mean, standard deviation

to scale function. Since both one dimension and two dimensions algorithms limited to use a single band, the PCA is applied to the input multi-spectral image.

Phase one: one dimension LULU implementation

The LULU-DPT algorithm is scripted in R programing language from the Carl Rohwer's LULU theory. The scripting of the operator is performed in the following steps.

1. Areas of interest are subset from the input GeoEye-1 image and the first principal component analysis PCA is applied to those subset images. Subset images are loaded to R-programing language and processed in the LULU-DPT algorithm either in rows or columns or histogram of the image to generate the discrete pulse transform (DPT). For example, the histogram of the image is read as:

```
h <- hist(as.vector(img), breaks=2000, xlim=c(-1, 3000),  xlab='DN value')
x <- h$mids      # the mid points between the break line
y <- h$counts    # the counts of the mid points
n <- length(y)
```

2. To calculate the minimum and the maximum values from the input sequence, first the input sequence is moved one step to the right by adding the copy of the first element on the left of the sequence for minimum calculation and one step to the left by eliminating the first element of the sequence and by adding the copy of the last element on the right of the sequence for maximum calculation.

```
minop <-function(y)         # Minimum operator
{
       n <- length(y)
       y_ls <- c(y[1],y[-n])
       return(pmin(y,y_ls))
}
maxop <-function(y)         # Maximum operator
{
       n <- length(y)
       yr <- y[n]
       y_rs <- c(y[-1],yr)
       return(pmax(y,y_rs))
}
```

3. To calculate the maximum-minimum (L_n) or the minimum-maximum (U_n) of the sequence the following function is defined based on the minimum and maximum calculations from the first step.

```
Ln <-function(y,n)         # Lower operator
{
       val <- y
       for(i in 1:n)
       {
              val <- minop(val)
       }
       for(i in 1:n)
       {
              val <- maxop(val)
       }
       return(val)
}
```

```
Un <-function(y,n)        # Upper operator
{
       val <- y
       for(i in 1:n)
       {
              val <- maxop(val)
       }
       for(i in 1:n)
       {
              val <- minop(val)
       }
       return(val)
}
```

4. The two main type of recursive smoothers and/or separator, the ceiling and flooring described by $C_n = L_n U_n C_{n-1},\ C_0 = x,$ and $F_n = U_n L_n F_{n-1},\ F_0 = x$ are scripted by the following function

```
Cn <-function(y,n)         # C 'ceiling' operator for L_nU_n
{
       if(n==0)return(y)
       else return(Ln(Un(Cn(y,n-1),n),n))
}
Fn <-function(y,n)          # F 'floor' operator for U_nL_n
{
       if(n==0)return(y)
       else return(Un(Ln(Fn(y,n-1),n),n))
}
```

5. The DPT of the LULU operator is computed from the recursive application of the operator by scripting the following code based on the LULU operator theory.

```
DPT <- array(0,c(length(y),n))
Cyn <- array(0,c(length(y),n))
n.pulses <- array(0,n)
pulses <- array(0,c(n,n))
k<-1
Clow <- y
while(k<=n)
{
       Cup  <- Ln(Un(Clow,k),k)
       Cyn[,k] <- Clow
       tmp <- Clow - Cup
       DPT[,k] <- tmp
       ind <- which(abs(tmp)>0)
       if(length(ind>0))
       {
              n.pulses[k] <- length(ind)/k
              for(l in 1:n.pulses[k])
              {
                     pulses[k,l] <- ind[1]
                     ind <- ind[-(1:k)]
              }
       }
       Clow <- Cup
       k <- k+1
}
```

```
DPT <- array(0,c(length(y),n))
Fyn <- array(0,c(length(y),n))
n.pulses <- array(0,n)
pulses <- array(0,c(n,n))
k<-1
Clow <- y
while(k<=n)
{
       Fup  <- Un(Ln(Clow,k),k)
       Fyn[,k] <- Clow
       tmp <- Clow - Cup
       DPT[,k] <- tmp
       ind <- which(abs(tmp)>0)
       if(length(ind>0))
       {
              n.pulses[k] <- length(ind)/k
              for(l in 1:n.pulses[k])
              {
                     pulses[k,l] <- ind[1]
                     ind <- ind[-(1:k)]
              }
       }
       Clow <- Cup
       k <- k+1
}
```

6. The output DPT is analysed to reconstruct the signal by separating the noise introduced in the image. Since in some of the resolution level, the local maximum-minimum or local minimum-maximum does not exist and the DPT at those levels will have zero values. For efficient use of the algorithm only non-zero DPT levels is used for reconstruction of the signal. Selective or partial reconstruction of sequence is performed by the following script.

```
E <- array(0,n)              # the energy of DPT
for(k in 1:n)E[k]<-sqrt(sum((DPT[,k])^2)/k)
Recon <- array(0,length(y)) # Selective reconstruction from DPT
min_level <- 1
max_level <- length(y)
Recon <- rowSums(DPT[,min_level:max_level])

Recon <- array(0,length(y)) # Only non-empty levels
ind <- which(E>0)
if(length(ind)>0)
Recon <- rowSums(DPT[,ind])
```

To illustrate how the operator smoothed and/or separate signal from noise, the following input sequence are taken from the histogram of the Subset-6 image processing the brightness value of the image from two dimensions arrays of matrix to one dimension row vector by reading all rows of the image in one row vector form and further analysed until generation of discrete pulses. First to apply the minimum and maximum operators the input signal (sequence) is shifted to the right for minimum operator and to the left for maximum operator

The input signal

620	618	567	687	678	629	687	695	703	...	664	604	601	579	570	576	588	574	584

Shift to the right

620	620	618	567	687	678	629	687	695	...	641	664	604	601	579	570	576	588	574

Minimum operator

620	618	567	567	678	629	629	687	695	...	641	604	601	579	570	570	576	574	574

Shift to the left

618	567	687	678	629	687	695	703	710	...	604	601	579	570	576	588	574	584	584

Maximum operator

620	618	687	687	678	687	695	703	710	...	664	604	601	579	576	588	588	584	584

Lower operator (L_1)

620	618	567	678	678	629	687	695	703	...	641	604	601	579	570	576	576	574	574

Upper operator (U_1)

620	618	618	687	678	678	687	695	703	...	664	604	601	579	576	576	588	584	584

C operator (L_1U_1)

620	618	618	678	678	678	687	695	703	...	641	604	601	579	576	576	584	584	584

F operator (U_1L_1)

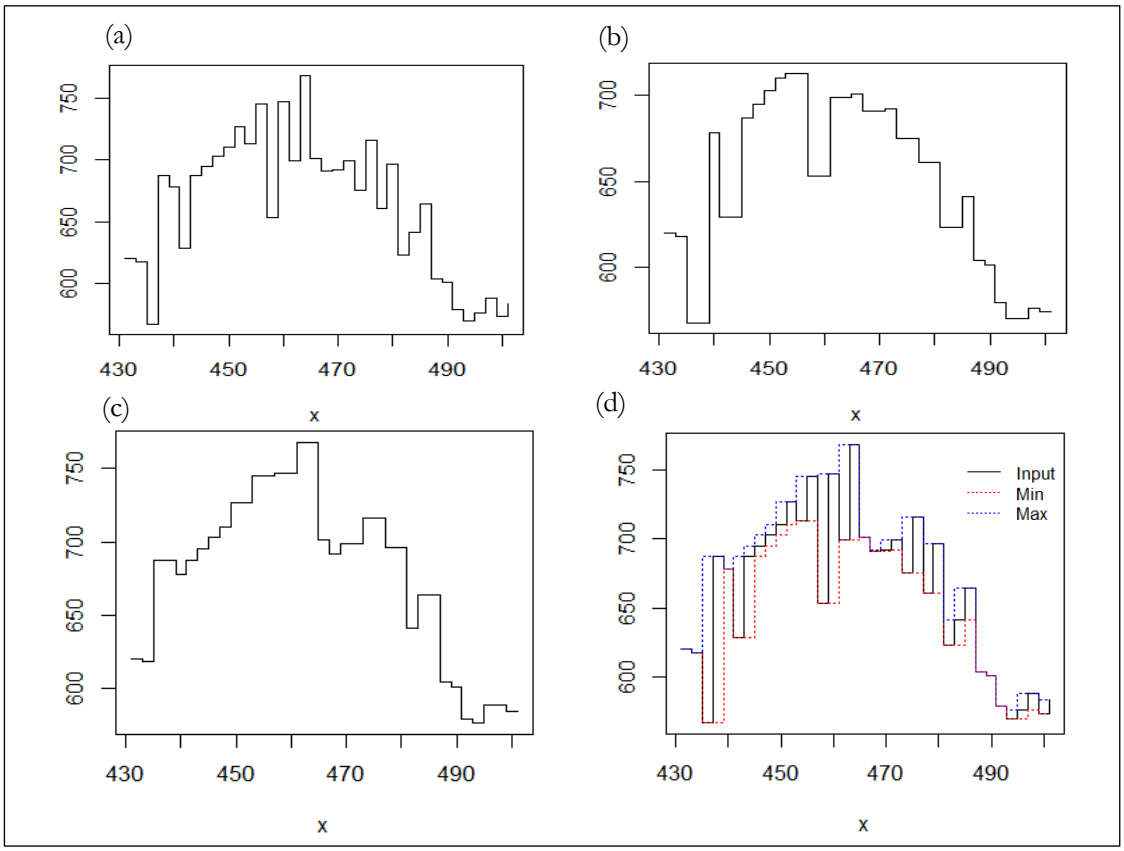
620	618	618	678	678	678	687	695	703	...	641	604	601	579	576	576	576	574	574

DPT Level 1 from *C* operator

0	0	-51	9	0	-49	0	0	0	...	23	0	0	0	-6	0	4	-10	0

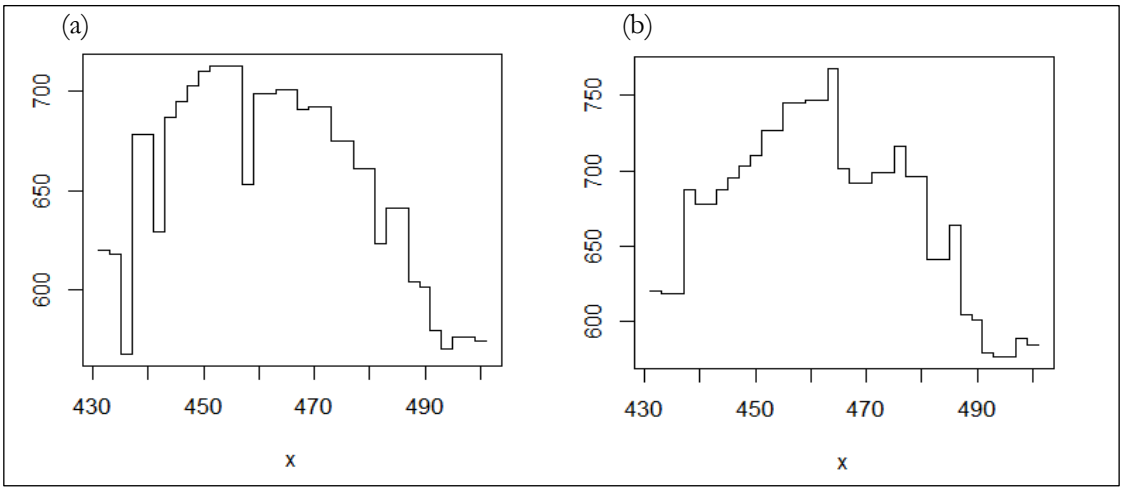
DPT Level 1 from *F* operator

0	0	-51	9	0	-49	0	0	0	...	23	0	0	0	-6	0	12	0	10



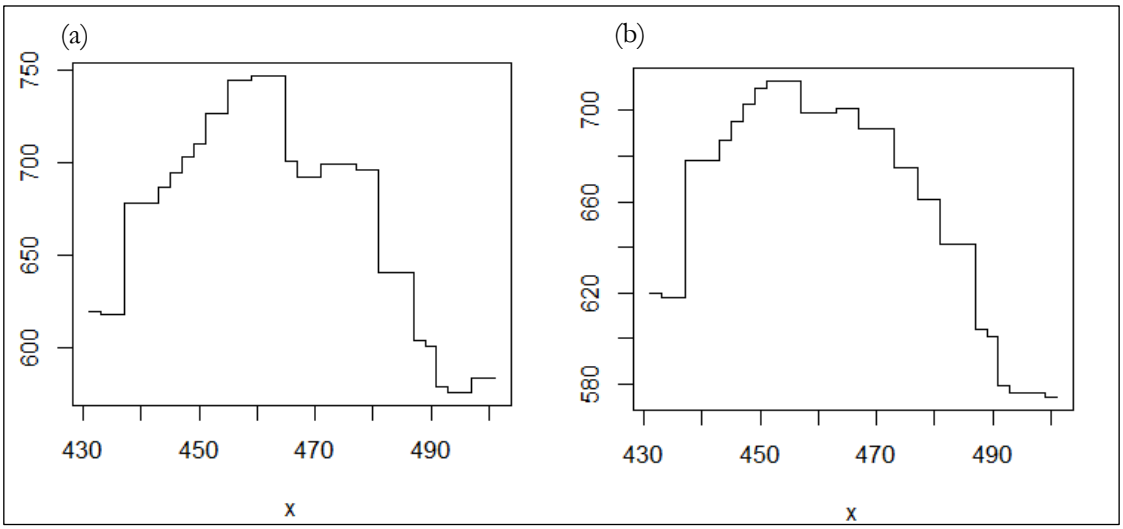
(a) Input signal (b) Minimum operator (c) Maximum operator (d) the combined plot of the input, Maximum and Minimum

Figure 4. 1 The input signal and the Minimum and Maximum operators in one dimension



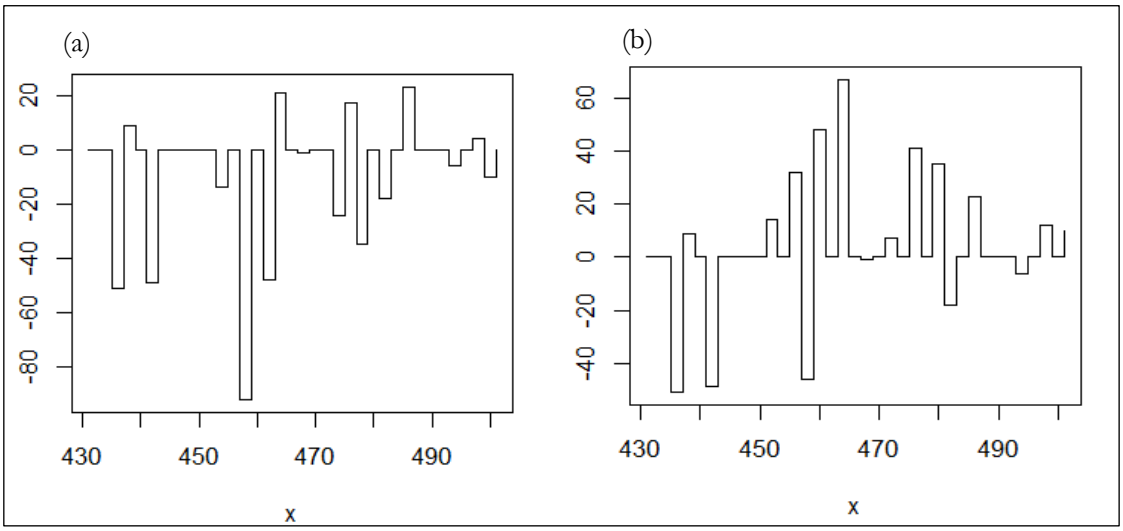
(a) The Lower operator (b) The Upper operator

Figure 4. 2 The Lower and Upper operators in one dimension



(a) The *C* operator (b) The *F* operator

Figure 4. 3 The *C* and *F* operators in one dimension



(a) D_1 from the *C* operator (b) D_1 from the *F* operator

Figure 4. 4 DPT at the first level from *C* operator and *F* operator in one dimension

In one dimension wavelets analysis, the same data used and loaded to R-programming language and the discrete wavelet transform for Haar wavelet is applied. The Haar wavelet analysis is used to denoising the input signal.

The final process of the one dimension analysis is performed by comparing the output from the LULU-DPT algorithm to the Haar wavelet output.

Phase two: two dimensions LULU implementation

The Discrete Pulse Transform in 2D algorithm (Fabris-Rotelliy and van der Walt, 2009) available at http://dip.sun.ac.za/~stefan/dpt., is used to decomposes the input image into a number of pulses. Basically the algorithm works on the concepts of Serra (1982) on connectivity of connected set, which is also developed and tested by Anguelov and Fabris-Rotelli (2008). Each pulse has constant amplitude (height) which is made up of from a connected pixel regions from a local minimum or maximum set. The connection between pixels in a connected region is based on 4-connections (Fabris-Rotelliy and van der Walt, 2009).

Since the process of decomposition produces large number of pulses, for example, 500 rows by 500 columns image may have from 100, 000 to 200,000 pulses, the algorithm used the Compressed Sparse Row (CSR) format storage scheme to represent sparse matrices of pulses (Fabris-Rotelliy and van der Walt, 2009).

The input image accepted by the algorithm is a $M \times N$ matrix of integer values between 0 and 255 in jpg format. Images are converted to multi-dimensional arrays and matrices to perform the decomposition. The outputs of the algorithm are pulses in array and can be exported to NumPy(`.np`) format by Numerical Python package for scientific computing in Python. Since the input image used in this study is a geo-referenced (geo-tiff) unsigned 16 bit image, the algorithm was modified to work on this remotely sensed image. To process the geo-referenced data the GDAL (Geospatial Data Abstraction Library) library used to convert the image from geo-tiff to gdal array of unsigned 16 bit. After the decomposition, during the reconstruction of the image, the gdal array is converted to geo-tiff image.

```
def load_image(fname='.tif'):
        if fname is None:
        fname = 'example.tif'
    fname = os.path.join(os.path.dirname(__file__), 'data/' + fname)
    if len(sys.argv) == 2:
        img = Image.open(sys.argv[1])
    else:
        img = gdal_array.LoadFile(fname)
    return img.astype(int)
```

From the DPT 2D algorithm the script that was developed for decomposition is used to decompose the input subset image. During the process of the decomposition, when the connectivity of pixel exists and if the neighbouring pixels have the same value in 4-conneted pixel, they are assigned to the same label value. The identification of these connected regions is processed to yield the DPT. The python code used to decompose the input image is:

```
fname='.tif'
img=gdal_array.LoadFile(fname)
if __name__ == "__main__":
    import sys
    if len(sys.argv) >= 2 and '-UL' in sys.argv:
        operator = 'UL'
        sys.argv.remove('-UL')
    else:
        operator = 'LU'
    img = load_image(fname)
    print("Decomposing using the %s operator." % operator)
    if operator == 'LU':
        print("Use the '-UL' flag to switch to UL.")

    tstart = time.time()
    pulses = lulu.decompose(img, operator=operator)
    tend = time.time()
    print "Execution time: %.2fs" % (tend - tstart)
```

The input image is decomposed to hierarchal sum of pulses (DPT) in which signal has a constant value on a connected set and zero elsewhere, based on the multi-resolution analysis (MRA) of the LULU operators. MRA is applied to generate DPT's to extract object at different resolution level.

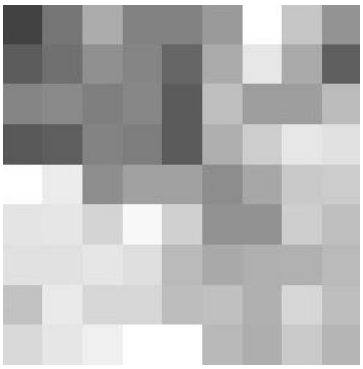
Objects are identified from the decomposed input image using DPT either by partial or full reconstruction of the image. Partial reconstruction is done by extracting feature of interest from non-zero DPT level that is by removing back ground features from discrete pulse decomposed image. For reconstruction of image from n to m DPT scale level, we should provide the value of input n and m. If the area of the pulse is equal to n then pulses are extracted from the connected region (`cr`) by connected region handler (`chr`). This process is continued until the reconstruction reached to the input value m by escaping zero DPT levels.

In the case of fully reconstruction of the image, the reconstruction of the image is done by the summation of all non-zero DPT level from the decomposed signal.

```
n= input("enter n: ")
m= input("enter m: ")

pulses_n = np.zeros(img.shape, dtype=int)
for f in (n,m):
    for area in pulses:
        if area == n:
           for cr in pulses[area]:
            crh.set_array(pulses_n, cr,
                          np.abs(crh.get_value(cr)), 'add')

if sum(sum(abs(pulses_n)))<>0:
    gdal_array.SaveArray(pulses_n, '.tif'%n, 'GTiff',gdal.Open(fname))
```

```
while n < m:
    n+=1
    pulses_n = np.zeros(img.shape, dtype=int)
    for f in (n,m):
        for area in pulses:
            if area == n:
               for cr in pulses[area]:
                crh.set_array(pulses_n, cr,
                              np.abs(crh.get_value(cr)), 'add')

    if sum(sum(abs(pulses_n)))<>0:
        gdal_array.SaveArray(pulses_n,'.tif'%n, 'GTiff', gdal.Open(fname))
```

To illustrate how the 2D operators (the *minimum* operator and *maximum* operator) applied on adjacent set, the following input image is subset (9 × 9 pixels) from the Subset-6 image and further analysed to generation the local minimum and maximum which is used to yield the discrete pulses. To apply the minimum and maximum operators, we used here a 3 ×3 pixels of moving neighborhood. Each pixel is filled with the minimum value from the surrounding 3 ×3 neighborhood pixel for minimum operator and the maximum value from the minimum operator for max-min operator (L_n). For maximum operator, each pixel is filled with the maximum value from the surrounding 3 ×3 neighborhood pixel and the minimum value from the maximum operator for min-max operator (U_n). The illustrations are presented as followed.



727	854	986	882	882	941	1224	1048	925
791	840	917	890	816	983	1129	981	799
891	903	876	894	789	1031	951	951	1023
783	799	885	873	789	991	1068	1128	1114
1236	1140	913	957	957	909	973	1055	1062
1121	1126	1081	1170	1071	922	922	1068	1032
1113	1113	1126	1108	1019	977	994	996	1019
1038	1134	1089	1089	1026	1034	995	1088	1028
1094	1126	1150	1185	1185	1017	989	1058	1006

Figure 4. 5 Input subset image (left) and the corresponding pixel values (right)

(a)

727	727	840	816	816	816	816	799	799
727	727	840	789	789	789	941	799	799
783	783	799	789	789	789	951	951	951
783	783	799	789	789	789	909	951	951
783	783	799	789	789	789	909	922	1032
1113	913	913	913	909	909	909	922	996
1038	1038	1081	1108	922	922	922	922	996
1038	1038	1089	1026	977	977	977	989	996
1038	1038	1089	1026	1017	989	989	989	1006

(b)

727	840	840	840	816	941	941	941	799
783	840	840	840	816	941	941	951	951
783	840	840	840	789	951	951	951	951
783	799	799	799	789	951	951	1032	1032
1113	1113	913	913	913	909	922	1032	1032
1113	1113	1108	1108	1108	922	922	1032	1032
1113	1113	1089	1108	1108	977	989	996	996
1038	1089	1089	1108	1108	1017	989	1006	1006
1038	1089	1089	1089	1026	1017	989	1006	1006

(a) Minimum operator output (b) Maximum-minimum (L_n) operator output

Figure 4. 6 The Minimum and Maximum-minimum operators in two dimensions

(a)

854	986	986	986	983	1224	1224	1224	1048
903	986	986	986	1031	1224	1224	1224	1048
903	903	903	917	1031	1068	1129	1129	1128
1236	1236	1140	957	1031	1068	1128	1128	1128
1236	1236	1140	1170	1170	1071	1128	1128	1128
1236	1236	1140	1170	1170	1071	1068	1068	1068
1134	1134	1134	1170	1170	1071	1088	1088	1088
1126	1150	1185	1185	1185	1185	1088	1088	1088
1134	1150	1185	1185	1185	1185	1088	1088	1088

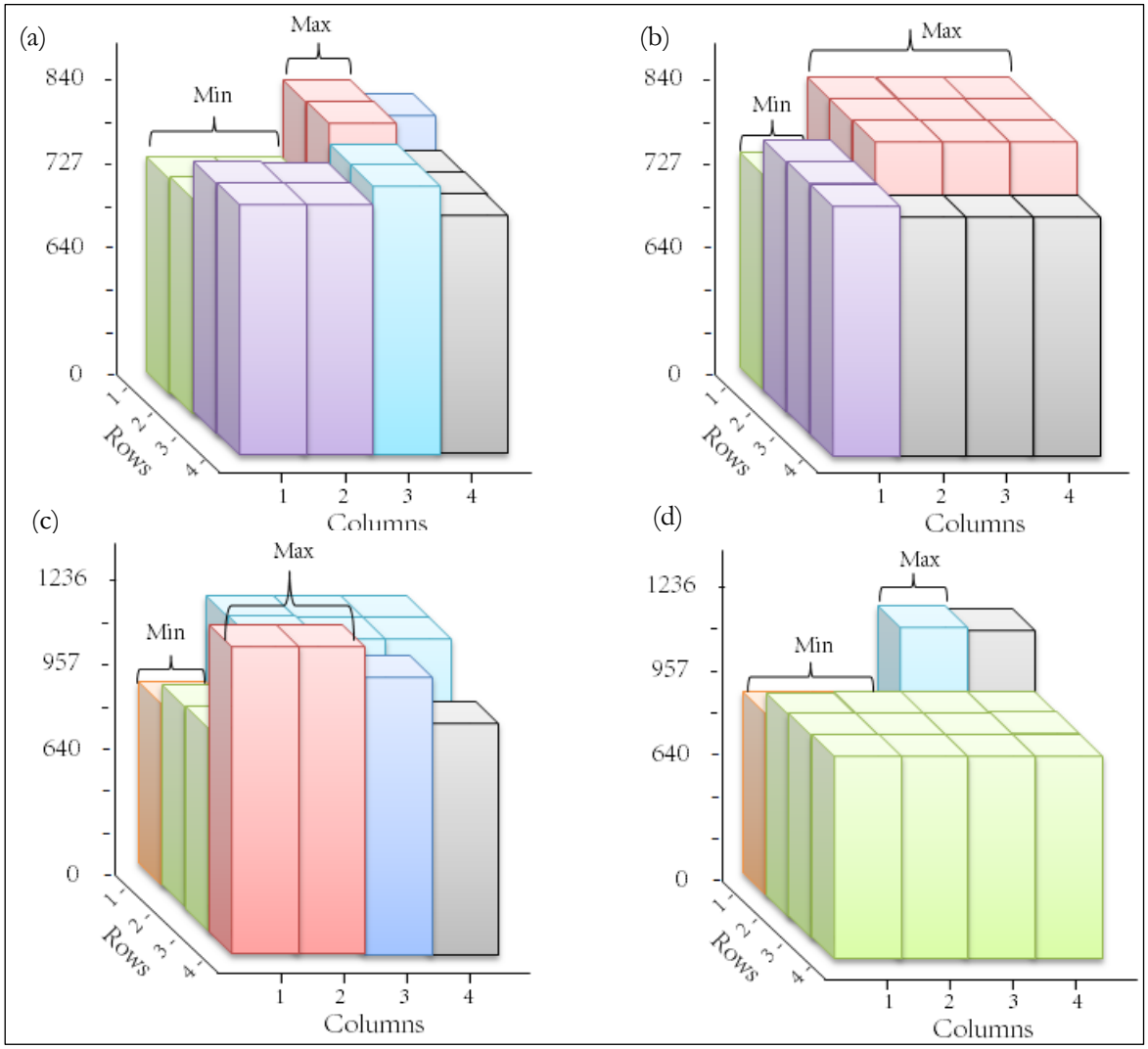
(b)

854	854	986	983	983	983	1048	1048	1048
903	903	903	903	917	1031	1068	1048	1048
903	903	903	903	917	1031	1068	1048	1048
903	903	903	903	917	1031	1068	1128	1128
1236	1140	957	957	957	1031	1068	1068	1068
1134	1134	1134	1134	1071	1068	1068	1068	1068
1126	1126	1134	1134	1071	1068	1068	1068	1068
1126	1126	1134	1134	1071	1088	1088	1088	1088
1126	1126	1150	1185	1185	1088	1088	1088	1088

(a) Maximum operator output (b) Minimum-maximum operator (U_n) output

Figure 4. 7 The Maximum and Minimum-maximum operators in two dimensions

The first upper left 4×4 pixels from the minimum, maximum-minimum (L_n), maximum and maximum-minimum (U_n) operators are plotted in Figure 4.7 to illustrate the minimum and maximum values from the adjacent set of pixel values.



(a) Minimum (b) Max-min (L_n) (c) Maximum (d) Min-max (U_n) operators output

Figure 4. 8 Two dimensions illustrations of the Minimum, Max-min, Maximum and Min-max operators

For automated identification of objects, scale space analysis is used on the ranges of scale level by linking all the non-zero DPT level objects to the original input subset image based on a given input point or pixel. From the non-zero DPT level; objects are taken as area of interest to link with the PCA 1 image and the multi-spectral image to generate the mean and standard deviation for those areas of interest. Then, based on the mean to scale and the standard deviation to scale plot, the point which shows drastic change on a certain threshold value is considered as a breaking point for identification of objects at that scale level.

To determine the breaking points, the derivatives of the standard deviation and the scale functions to the decomposition level are computed with the following equation.

$$\Delta\sigma_i = \sigma_{i+1} - \sigma_i$$
$$\Delta S_i = S_{i+1} - S_i$$

Where σ is the standard deviation and S is the scale level and i is the decomposition level, $i = 1,2,3, \dots, n$, $n \in \mathrm{M} \times \mathrm{N}$ size of image.

The derivatives of these equations approximate the change in gradient quantities of σ and S with respect to the decomposition level of the image i.

The final process of the two dimension analysis is performed by comparing the output from the LULU-DPT algorithm to the thresholding segmentation output.

To summarize some of the tasks during the implementation are:

- The input image is decomposed to hierarchal sum of pulses (DPT), i.e., to signal that are constant on a connected set and zero elsewhere. These DPTs are used to extract object from the decomposed image based on multi-resolution analysis (MRA).
- Segments and objects are extracted from the decomposed input image using DPT either by partial or full reconstruction of the image. This is done by extracting feature of interest and removing back ground features from discrete pulse decomposed image.
- Automated identification of objects based on scale space analysis at different scale level from non-zero DPT levels.

4.3. Evaluation

After implementation of the segmentation method the accuracy of the result need to be determined by comparing with a reference data. Therefore, evaluation of segmentation is necessary to assess the accuracy of the output. In addition, it is important to notice that data users and researchers should have strong knowledge about the evaluation method because the segmentation accuracy is not only influenced by the segmentation method; it is also influenced by the accuracy assessment method(Congalton, 1991).

4.3.1. Segmentation accuracy assessment

Lucieer et al. (2004) suggested that for quantitative assessment of segmentation result, it can be assessed by quantitative method by identifying the m objects from the reference data and calculating the percentage of overlap of these object from the largest segmented object. The following equation of area fit index (AFI) is used to quantify the fit of each of the reference object with that of the largest segments in the overlapping object.

$$AFI = \frac{A_{reference\ object} - A_{largest\ segment}}{A_{reference\ object}}$$

Where A is the area in pixels. If the overlap is less than one hundred percent, the image is over segmented and AFI is greater than 0.0. If the overlap is more than hundred the image is under segmented and AFI is less than 0.0. For a perfect overlap is 100% and AFI equals to 0.0. In some cases when the object is over segmented but the largest segment is larger than the reference object, the overlap can be less than 100% and AFI is less than 0.0.

Delves et al. (1992) introduce a method for evaluation method for segmentation quantitatively by compare segment boundaries with boundaries on a reference data. For example, let p be a boundary pixel of a region in the reference map and $D(p)$ be the shortest (Euclidean) distance, measured in pixels, between p and any boundary pixel in the segmented image, i.e.,

$$D(b) = \frac{\sum_{boundary\ pixels} - D(p)}{N}$$

Where the sum is taken over all boundary pixels in region b, and N is the number of boundary pixels in the reference data set. The average distance between a segment boundary pixel and the reference boundary is measured by $D(b)$. If a perfect fit exist the value of $D(b)$ is equals 0.0.

The image segmentation accuracy measure when the region b is equals to the whole image is obtained by $D(B)$ without considerations of M which represent the number of boundary pixels in the segmented image. Low $D(B)$ values are obtained for high values of M in which many boundary pixels in the neighbourhood of p occured. In addition image over-segmented may happen for boundary image with high M values and to correct for M the following correction equation is used during calculation of $D(B)$.

$$D(B)_{corr} = \frac{|N - M|}{N} + D(B)$$

Another method for accuracy assessment especially for building identification is presented by Zhan et al. (2005). This object-based accuracy assessment method is based on the matching of objects from the reference data and the segmented output. During overlaying of two datasets the assumption is that if at least fifteen present of the area covered is common area in the overlaid of the two objects from reference data then they are the same object. In areas where relatively large buildings are present this approach works quite well.

In this study one of the object-based approaches for accuracy assessment is used. The method is overlaying the extracted object with the reference data based on their geometric centres, which take into account the shape and size of the object. If the geometric centre of extracted object is located into the reference data then this extracted object is an existing object.

5. RESULT

In this chapter, the obtained results are presented from the implementation of the algorithm which is explained in the previous chapter in order to segment and identify objects from a VHR satellite image. The results are presented in two sections. The first section contains the output from LULU operators Multi-resolution analysis by taking into consideration the dimension analysis of LULU operators. The results are presented in one dimension and more emphasis is given to two dimensions outputs. In the second section the outputs from the first sections two dimensions output are evaluated and presented.

5.1. Multi-resolution Analysis of LULU operators and the associated DPT

This section describes the results obtained after the implementation of the algorithm for multi-resolution analysis explained in chapter two.

5.1.1. LULU operators and DPT output in one dimension Multi-spectral Analysis

For multi-spectral analysis first, the Cairo city GEoEye-1 image is subset into six areas of interest. Second, principal component analysis (PCA) is performed on the subset of multispectral bands of the image mainly for two purposes. (1) To reduce the number of bands comprising a dataset, because bands are correlated (Table 5.1) (2) to identify hidden patterns in the data. Table 5.1 shows the correlation coefficients between bands of GeoEye-1 image. The bands of GeoEye-1 image show high correlation coefficients. Relatively, the Green band and Blue band shows high correlation coefficient (0.99) and the Red band and the NIR band shows the lower correlation coefficient (0.93) in the image. Figure 5.1 shows, some of the PCA 1 of the subset of image and their corresponding histogram plot.

Correlation	Blue Band	Green Band	Red Band	NIR Band
Blue Band	1.00	0.99	0.96	0.96
Green Band	0.99	1.00	0.98	0.96
Red Band	0.96	0.98	1.00	0.93
NIR Band	0.96	0.97	0.93	1.00

Table 5. 1 Correlation coefficient between bands of GeoEye-1 image

The L_nU_n operator was applied to decompose the subset image in one dimension to different levels of decomposition. It is done by transforming the brightness value of the image from two dimensions arrays of matrix to one dimension row vector by reading all rows of the image in one row vector form. Figure 5.2 shows the original signal profile and the reconstructed signal of a subset image after the LULU-DPT decomposition. The reconstruction of signal by LULU-DPT was performed for the whole subset image (full reconstruction) from the non-zero DPT level. The result shows that there are some outliers visible in the output graphs.

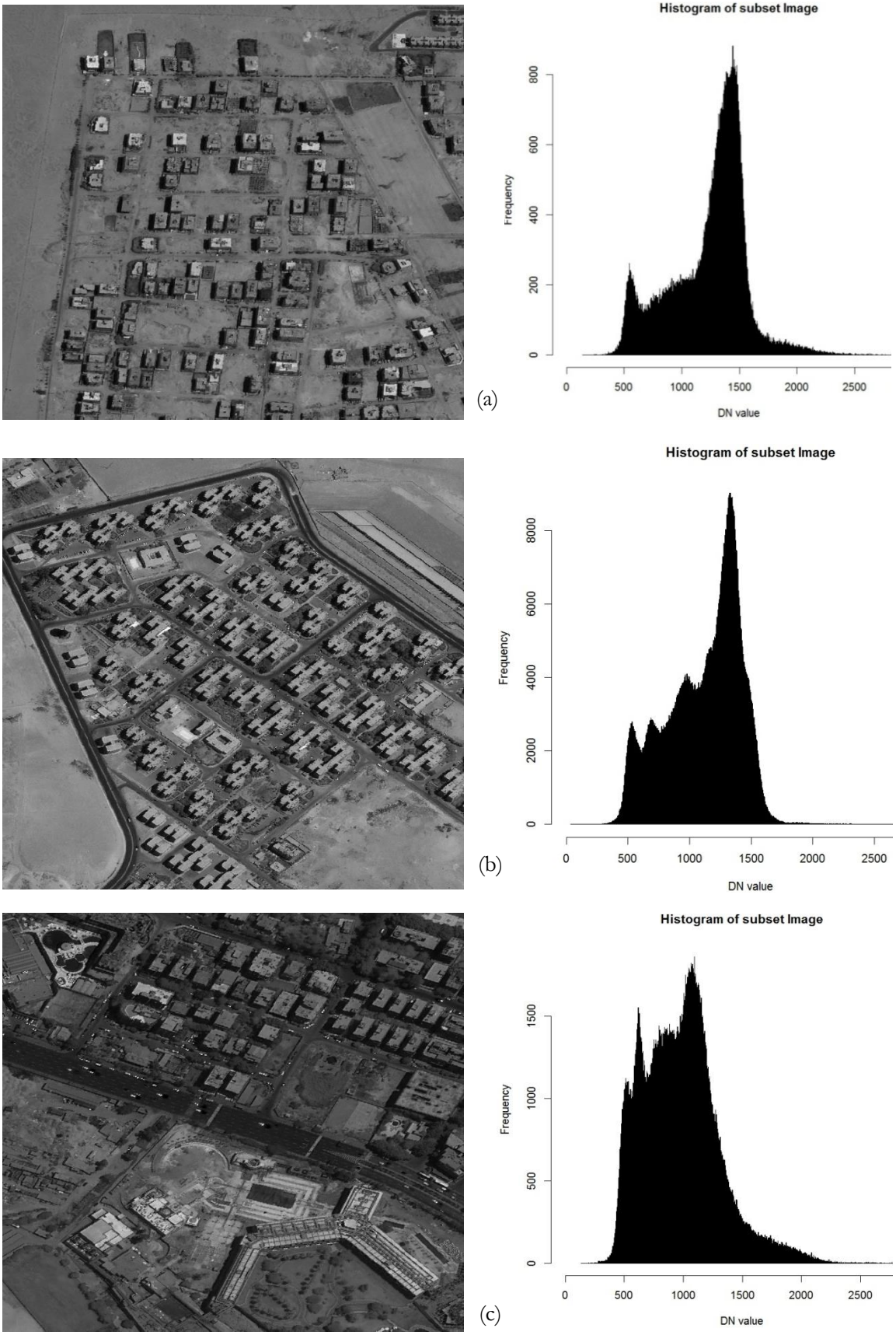


Figure 5. 1 PCA 1 of Subset-1 (a), Subset-2 (b) and Subset-3 (c) image, histogram (right)

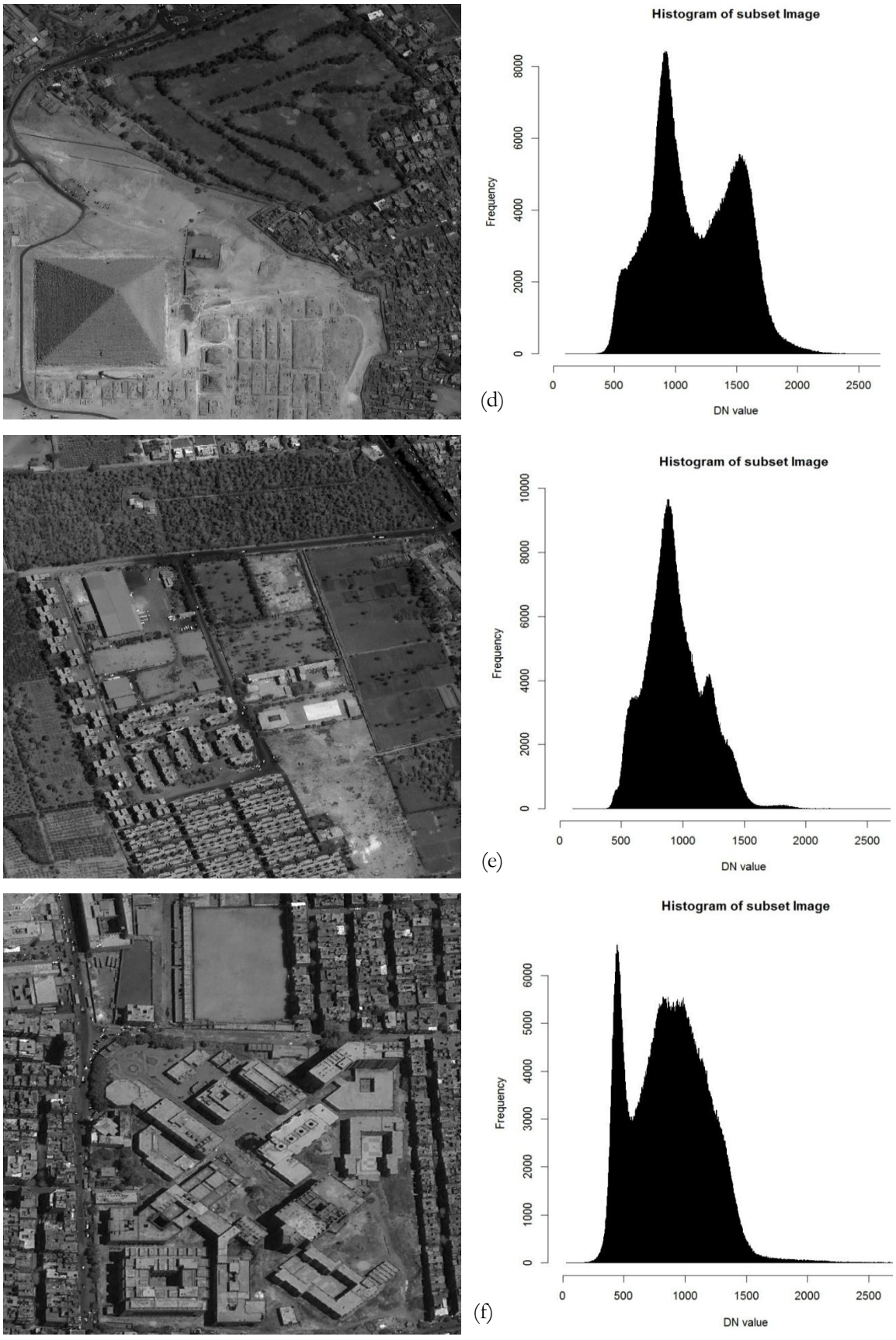
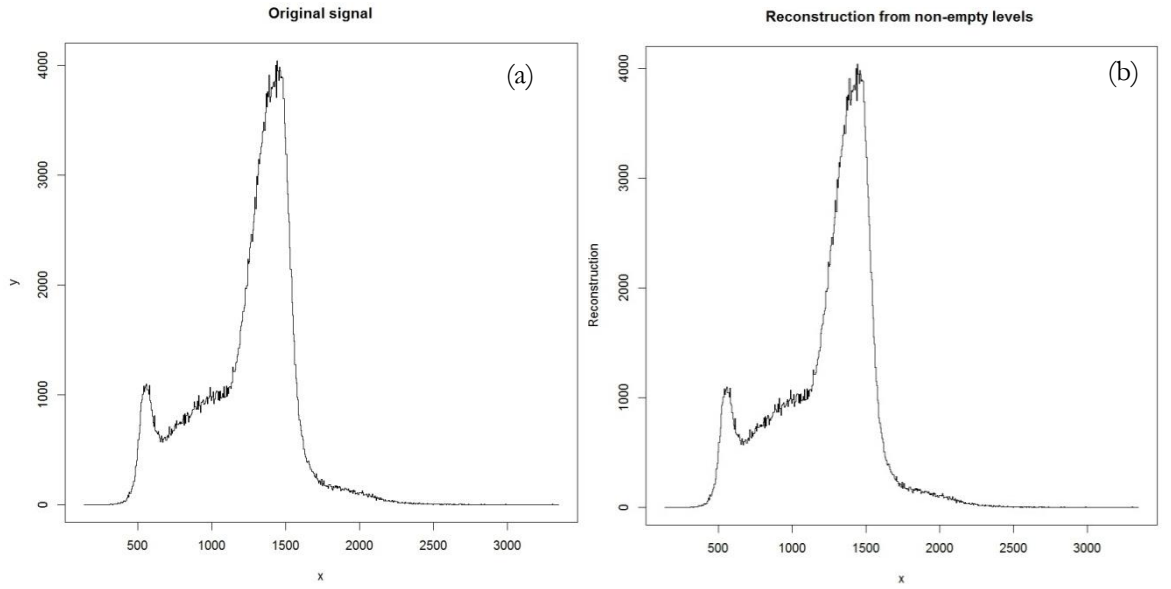
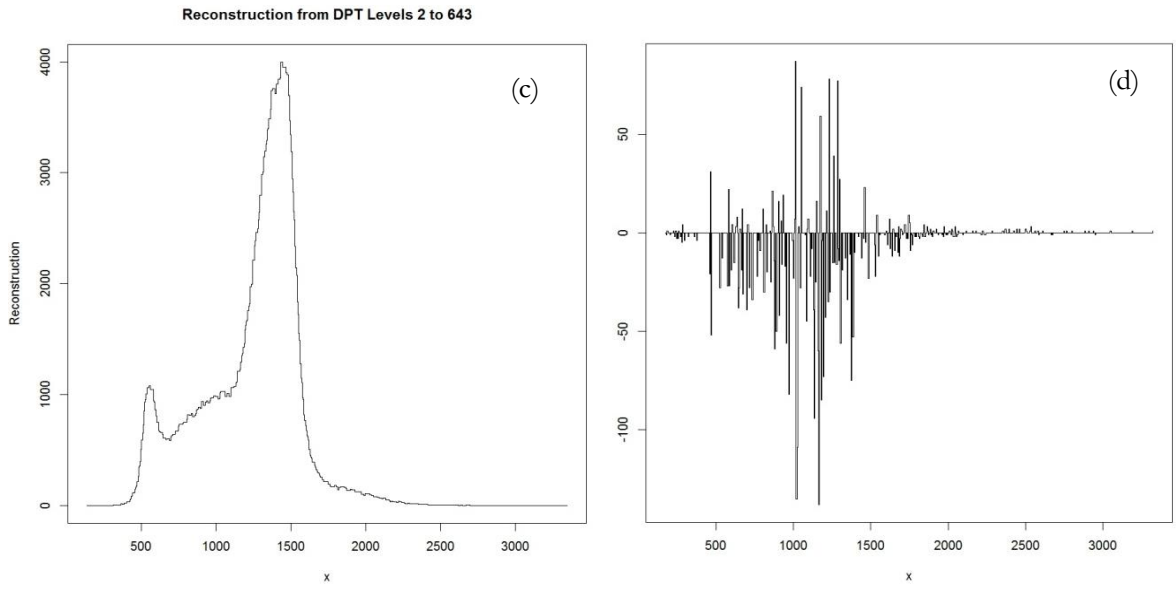


Figure 5. 2 PCA 1 of Subset-4 (d), Subset-5 (e) and Subset-6 image (f), histogram (right)

To illustrate how the LULU-DPT algorithm separate the signal from the noise, the partial reconstruction is performed from the second level of resolution to the last level of decomposition. The result graph in Figure 5.3(c) is shows that a smoother output than the original graph after reconstruction. Figure 5.3 (d) shows the impulsive noise extracted from the subset image at the first decomposition level.



(a) The original signal of Subset-1 image (b) The reconstructed signal from non-zero DPT level



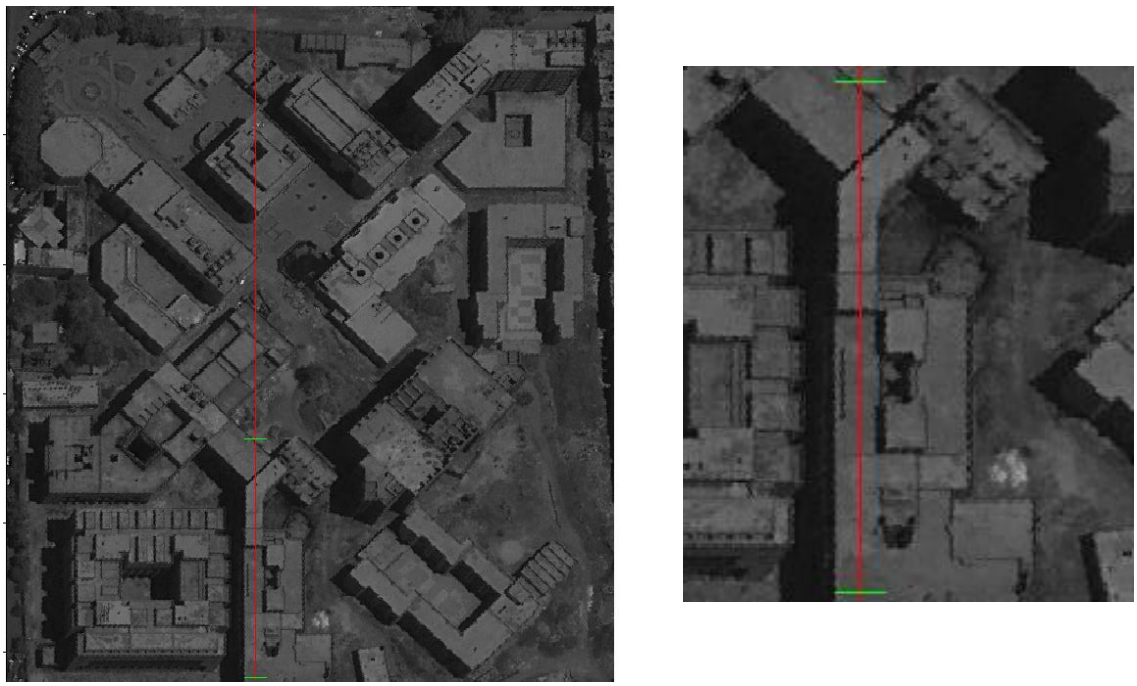
(c) Reconstruction from level 2 to 643 (d) Extracted impulsive noise

Figure 5. 3 One dimension output for Subset-1 image

Analysis of the LULU-DPT resolution component

In order to analyse the resolution components of the image in one dimension during the application of the LULU operator, the brightness value is extracted from the subset image along the red line (Figure 5.6) by considering as a sequence as displayed in the figure. The decomposition of sequence x yields a simple representation of a signal in terms of block pulses of different width. For example the first scale level of the sequence x cannot have two consecutive nonzero elements, eventually the second scale level of the

sequence have two consecutive nonzero elements. Also the third scale level has three consecutive nonzero elements and this result continue until the last component being a constant sequence. To demonstrate the component analysis of the decomposed of sequence, a few quantized components (n = 22, 40, 70 and 162 level of scale) are chosen and presented in Figure 5.5. Based on the resolution of a sequence x given by $\mathrm{r}^{(\mathrm{n})} = \mathrm{D}_{\mathrm{n}}(x)$ from the extracted sequence, the LULU operator was applied to identify the part of the building located in the lower part of the subset image and the result is shown in the Figure 5.6.



(a) Subset image with selected signal at red line (b) zoomed subset image with selected signal

Figure 5. 4 Subset-6 image with selected signal

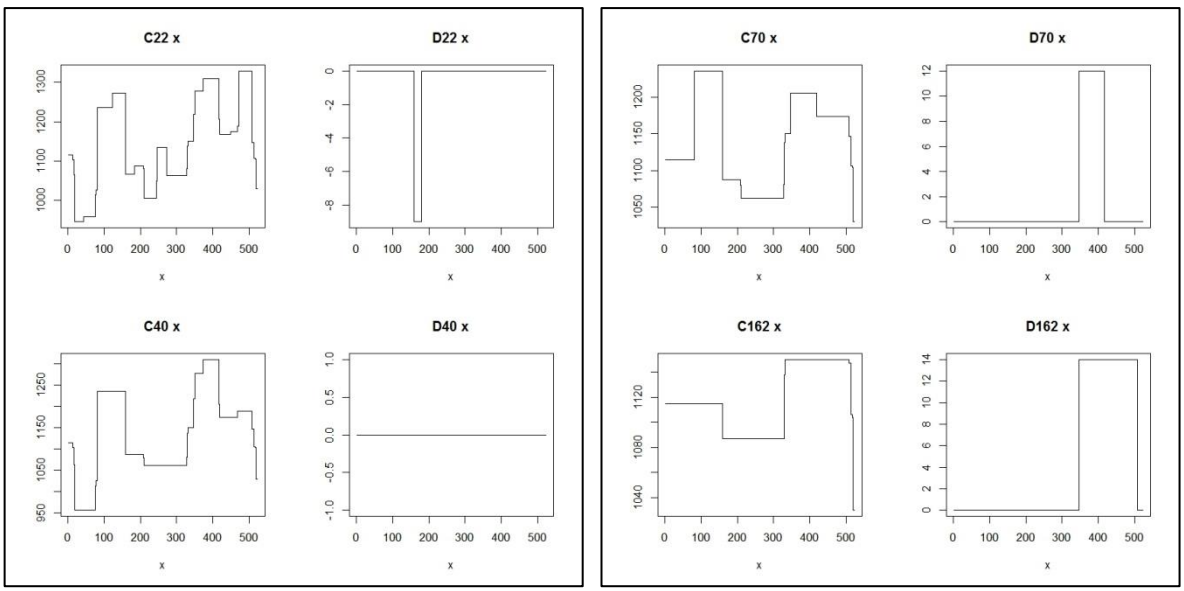


Figure 5. 5 The C operator and the DPT result at the level of 22, 44, 70 and 162.

During the course of decomposition, as shown in Figure 5.5 the smoothness of the sequence x increases as the resolution level increase during the analysis in addition to separation of the additive component.

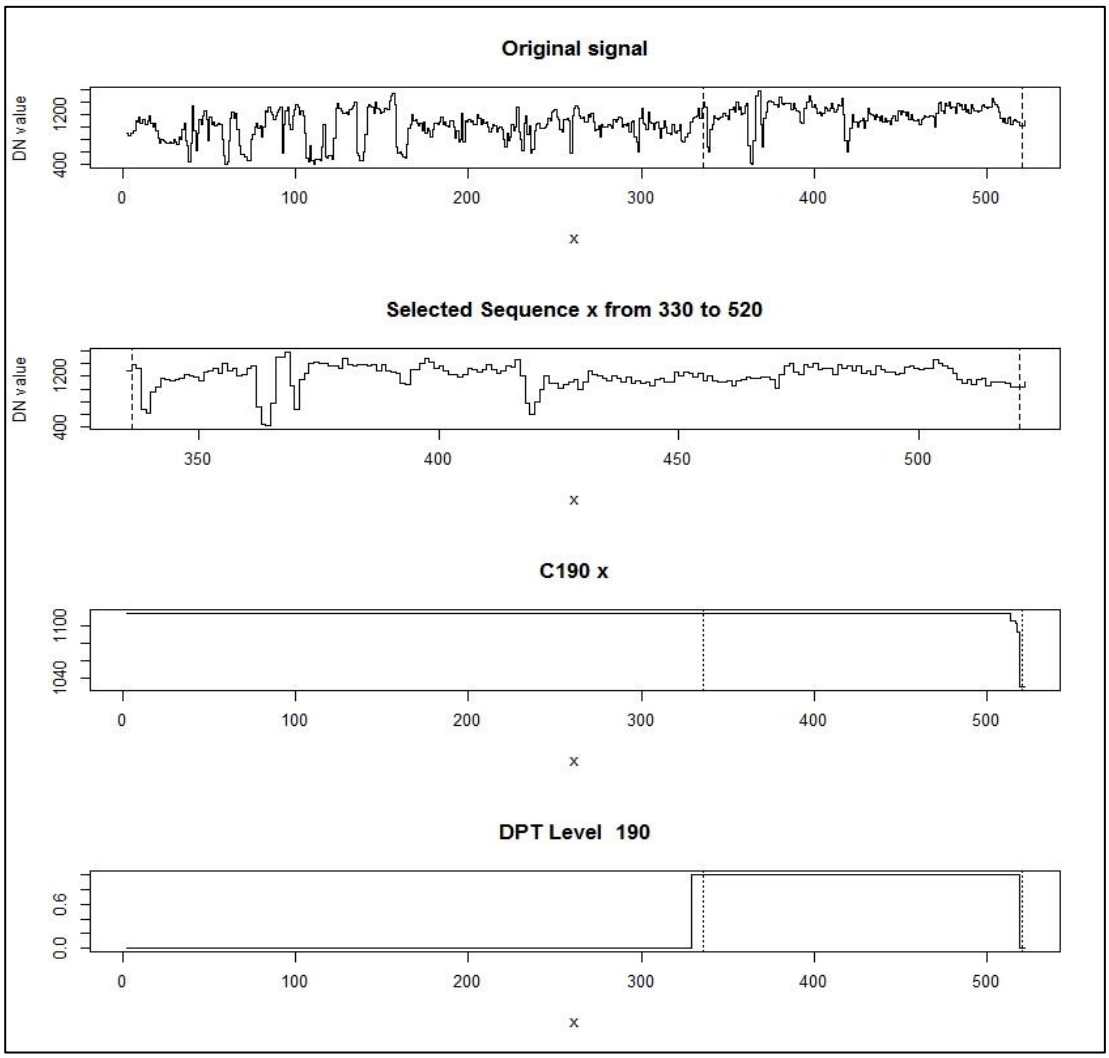


Figure 5. 6 Identification of building part from one dimension analysis

To illustrate the differences between the outputs from the LULU operator and wavelet, the following figure (Figure 5.7) represents the result from the Haar wavelet at the first, second and third level of decompositions. The Haar wavelet result shows that the output signal become smoother and smoother than the original signal and the width between two consecutive signal become wider and wider. The wavelet coefficients $\psi_1(x), \psi_2(x)$ and $\psi_3(x)$ and as well as the scaling coefficients $\varphi_1(x), \varphi_2(x)$ and $\varphi_3(x)$ shows significant differences in the graph when the level of decomposition increases. The qualitative comparison between the first higher resolution between the LULU DPT and the Haar wavelet coefficient shows that the LULU DPT systematically removes the impulsive noise from the original signal (Figure 5.8). However, the Haar wavelet shows some random peaks and valleys in the output. In addition at LULU-DPT level 190 (Figure 5.6), LULU operator identifies a building structure, in contrary, it was impossible to identify the structure of the building from the Haar wavelet output.

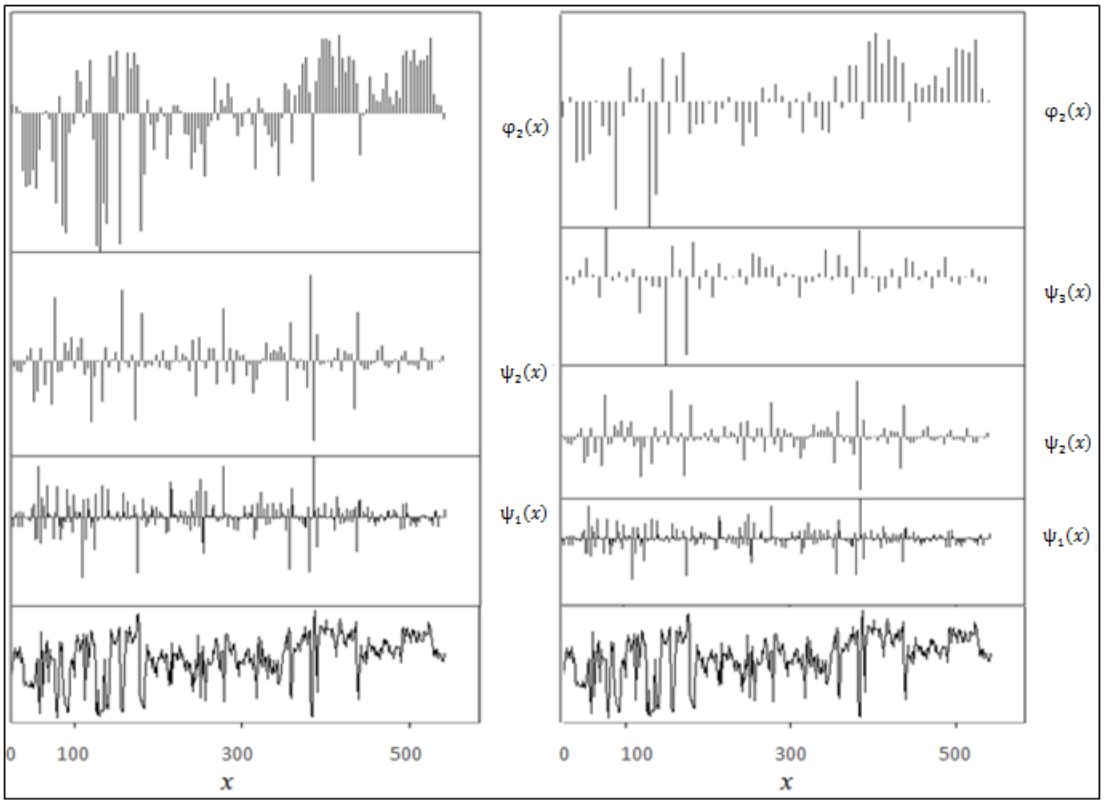


Figure 5. 7 A Haar wavelet decomposition of the subset image

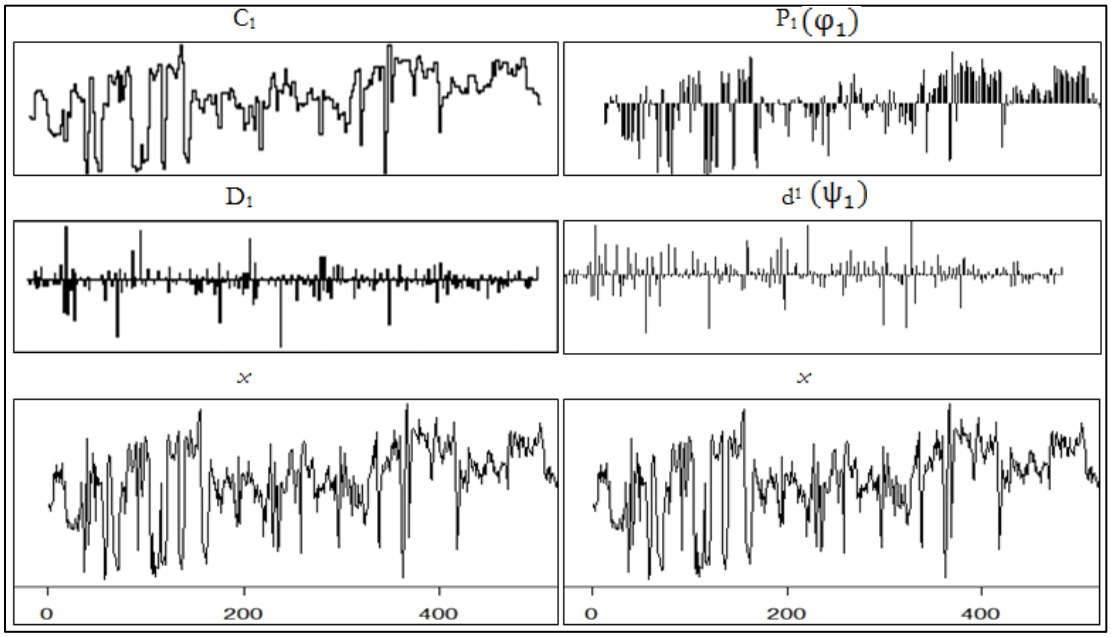


Figure 5. 8 The LULU- DPT (left) and Haar Wavelet (right) result at the level 1 resolution analysis

One dimension multi-resolution analysis based on LULU operator permits the analysis of signals in detail between successive levels of scale or resolution. The LULU-DPT algorithm in one dimension used to separate signals from noise consistently. The results in Figure 5.3(b) shows the reconstruction of signal from non- zero DPT level since in some of the resolution level the local maximum-minim or the local minimum-maximum does not exist and the DPT at that resolution level has zero value as presented in Figure 5.5 at n=40 scale level ($D_{40}x$). In Figure 5.3(c) signals are reconstructed by separating impulsive

noises from the signal that is introduced in the image due to a number of reasons such as atmospheric effect, transmission error or electromagnetic effect of the sensor, etc. Figure 5.3(d) shows the impulsive noise extract at the first level of decomposition from the original signal. Moreover, the DPT can identify the different structures of the objects at different level of reconstruction. This result is revelled by the result obtained in Figure 5.5 at $D_{22}x$, $D_{70}x$ and $D_{162}x$ which extract a sequence width of 22, 70 and 162 respectively. In addition, the smoothness of the sequence increases from the first resolution level to the next resolution level as shown in $C_{162}x$, which is smoother than $C_{40}x$, $C_{70}x$ and $C_{22}x$ (Figure 5.5). LULU operator as a separator extracted building structure in a given sequence presented in the Figure 5.6 at the 190 DPT level, which is equivalent to the length of the building structure in one dimension.

5.1.2. LULU operators and DPT output in two dimensions Multi-spectral Analysis

To illustrate the analysis of LULU operators and the associated DPT in two dimensions, the operators applied to decompose the subset image into $N \times M$ resolution levels where N and M are the width and length (number of rows and columns) of the input image.

The discrete pulse transforms (DPT), which is resulted from LULU is an application for hierarchal decomposition of the input array into a sum of pulses in which signal are constant on a connected set and zero elsewhere. Except for zero DPT level, each level has one or more objects and these objects are detected by keeping the lower abstraction level of pixels. Then, the basis for these object identification is the scale function of the LULU operators. Figure 5.9 shows the Subset-6 image, the initial pixels (or points) for automated identification and the objects that are planned to extract from the image based on automated scale space analysis.

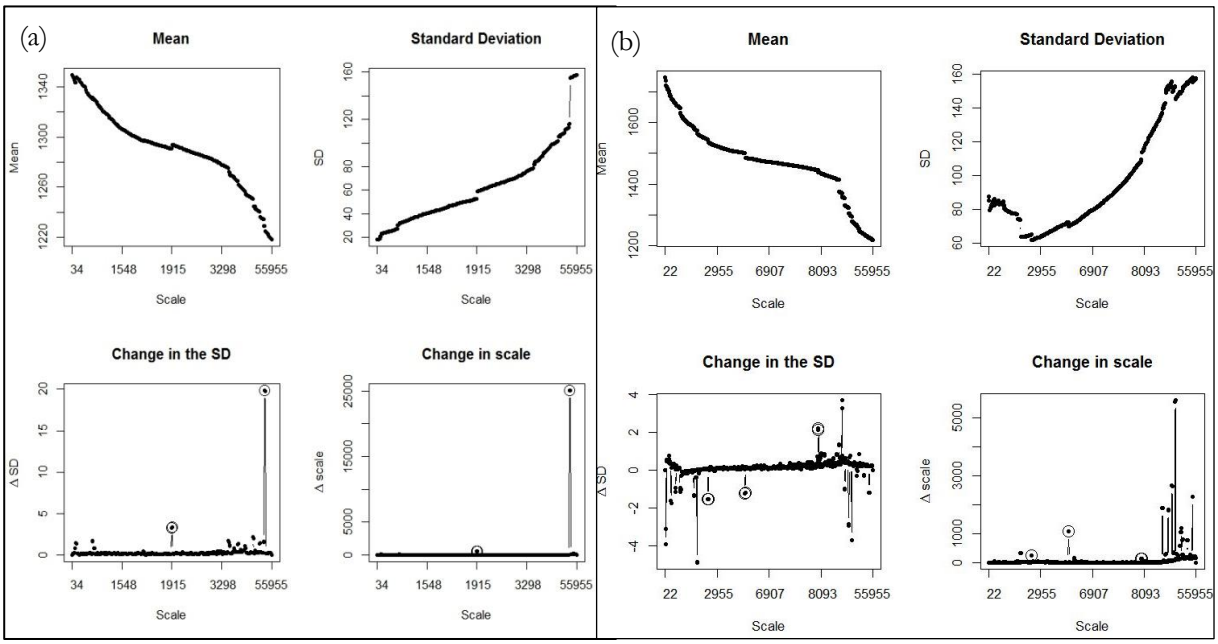
Figure 5. 9 Subset image (left) and Buildings (right)

To illustrate the automated detection of objects of interest, ten pixels (or points) are selected from the buildings, open space and vegetation area. The automated detection of objects of interest is based on

mean and standard deviation to scale function to generate the breaking points in which objects are drastically changed in size and shape. Break points are determined by calculating the change in standard deviation as well as the change in scale by providing a constant threshold to the algorithm. For example, for building detection the threshold values are: ≥ 1.5 for change in standard deviation, ≥ 50 for change in scale and > 500 for scale parameter. For Open space and vegetation: ≥ 0.5 for change in standard deviation and ≥ 5 for change in scale parameter. Objects of interest are identified from those breaking points generated by the algorithm.

Reconstruction of objects by LULU-DPT multi-resolution analysis is mainly constrained by the scale factor which determines the level of detail and the size of the reconstructed object. For example, Figure 5.11 shows the partial reconstruction of the objects by LULU-DPT (non-zero DPT level) from a single pixel to a set of connected pixels in which the pixel of smallest object becomes element of the largest object. As presented in the Figure 5.11 the level of detail and the size of the object increase from the lower scale to the higher scale level.

The mean to scale and the standard deviation to scale plot describe the level of variability as well as the homogeneity of the extracted object at the certain ranges of scale value. For example, Building-1 (Figure 5.11 - left) identified based on the breaking points at the scales of 2997 and 4220 that is obtained based on the change in the standard deviation and the change in scale as shown in Figure 5.10(a). Therefore, the different structure of Building-1 is reconstructed at the pulses of size Φ_{1915} , Φ_{2997} and Φ_{4220}. Different structure of Building-2 (Figure 5.11- right) is also reconstructed at the breaking points or scales or pulses of size 1905, 2403, 7770 and 8014 as shown in Figure 5.10(b).

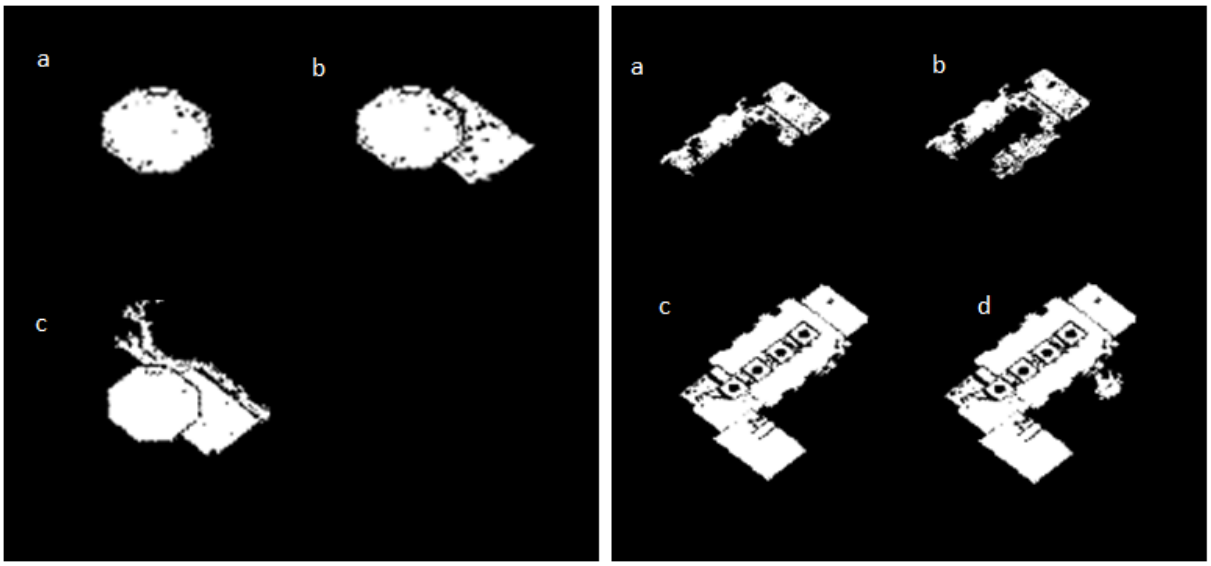


(a) Building-1 mean, standard deviation and scale (b) Building-2 mean, standard deviation and scale

Figure 5. 10 Mean, Standard Deviation and Scale plot for Building-1 and Building-2

The scale of Building-1 (Figure5.10 (a)) from 1 to 4220 is showed that the standard deviation increased while the mean value decreased, since the object is grown from a single pixel to a collection of many pixels. The standard deviation is drastically increased and a high difference is obtained at the scales of 1915, 2997 and 4220. For Building-2 the standard deviation was high at the beginning and drastically

decreased until scale level 1905 (Figure5.10 (b)). Then after, the standard deviation gradually increased to the last level except at the scale level 8014, which shows variation due to the object added heterogeneous pixels. The results of the output (Figure5.11) is showed that at the higher scale level the objects become clearer and preserve their shape and appearances, since they include more pixels than the lower scale levels.

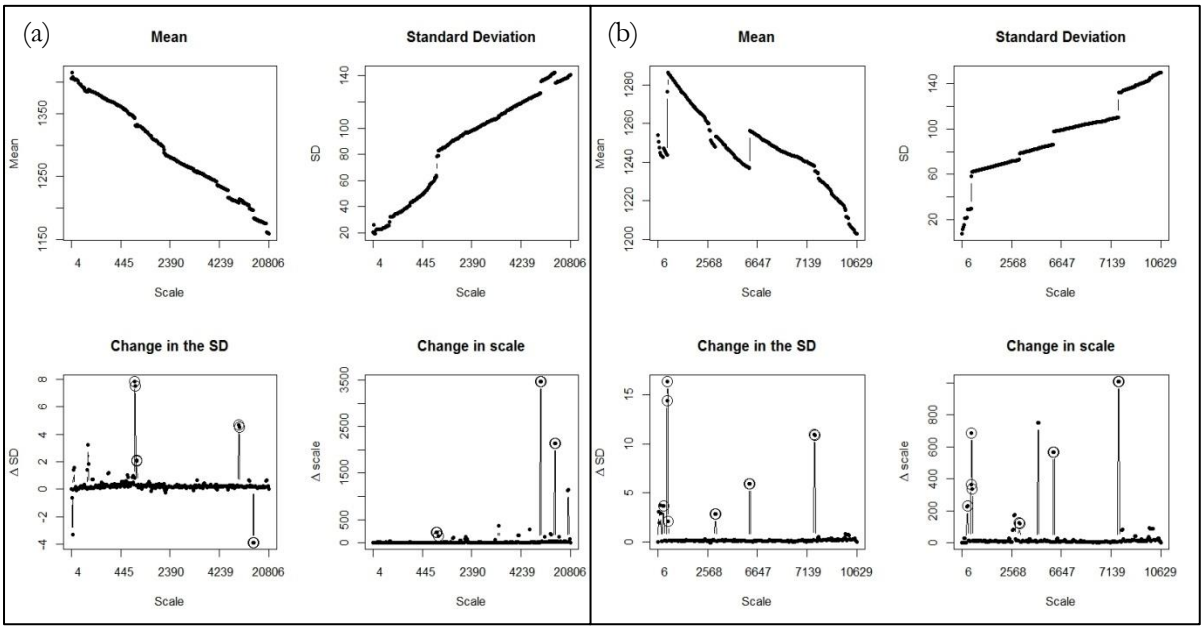


Identified Building-1 (left) based on Point-1 at (a) Φ_{1915} (b) Φ_{2997} (c) Φ_{4220}
Identified Building-2 (right) based on Point-2 at (a) Φ_{1905} (b) Φ_{2403} (c) Φ_{7770} (d) Φ_{8014}

Figure 5. 11 Identified Building-1 and Building-2

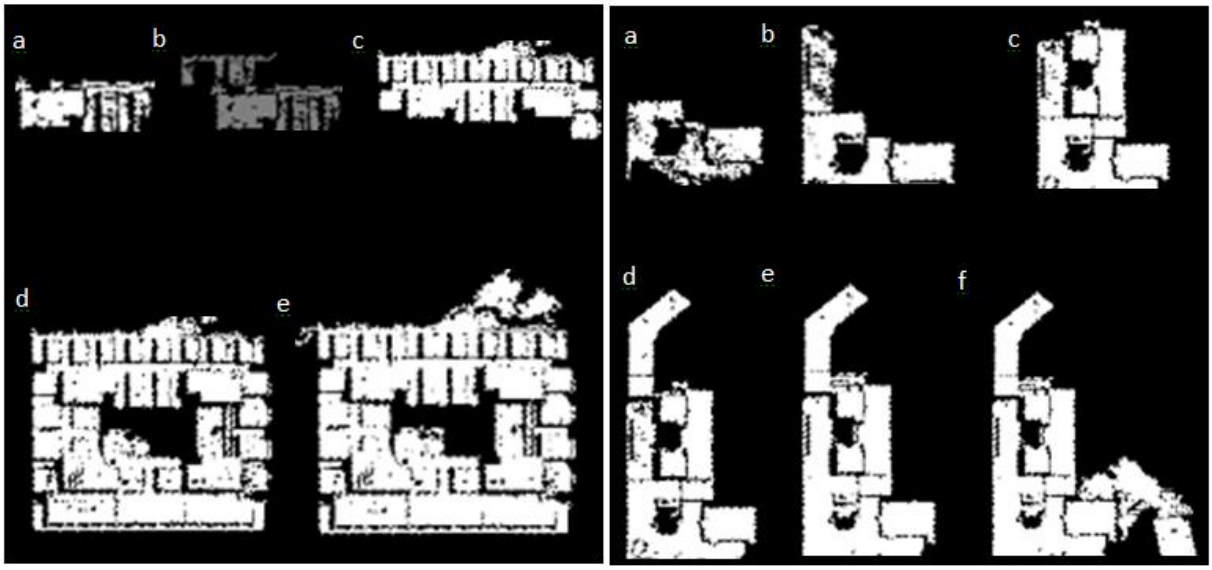
The scale of Building-3 (Figure5.12 (a)) from 1 to 13247 is showed that the standard deviation increased while the mean value decreased. In some points, the standard deviation is drastically increased and a high difference is obtained at 970, 1254, 5142, 12074 and 13247 scale levels. For Building-4 the standard deviation is shown drastic change at the scales or breaking points of 514, 1947, 5411, 6540, 7206 and 9216 (Figure5.12 (b)). These scale values or breaking points used to reconstruct the building from the corresponding pulse.

Different structure of Building-3 (Figure 5.13 - left) is identified at the pulses of size $\Phi_{970}\Phi_{1254}$ Φ_{5142} Φ_{12074} and Φ_{13247} in the scale range from 970 to 13247 with breaking points 970, 1254, 5142, 12074 and 13247. The first three pulses only built the upper side part of the building and the last two pulses built the complete structure of the building. Also the identification of different structures of Building-4 (Figure 5.13 - right) is done at the pulses of size Φ_{514} Φ_{1947} Φ_{5411} and Φ_{6540} at the breaking points 514, 1947, 5411 and 6540 in the scale range of 514 to 6540.



(a) Building-3 mean, standard deviation and scale (b) Building-4 mean, standard deviation and scale

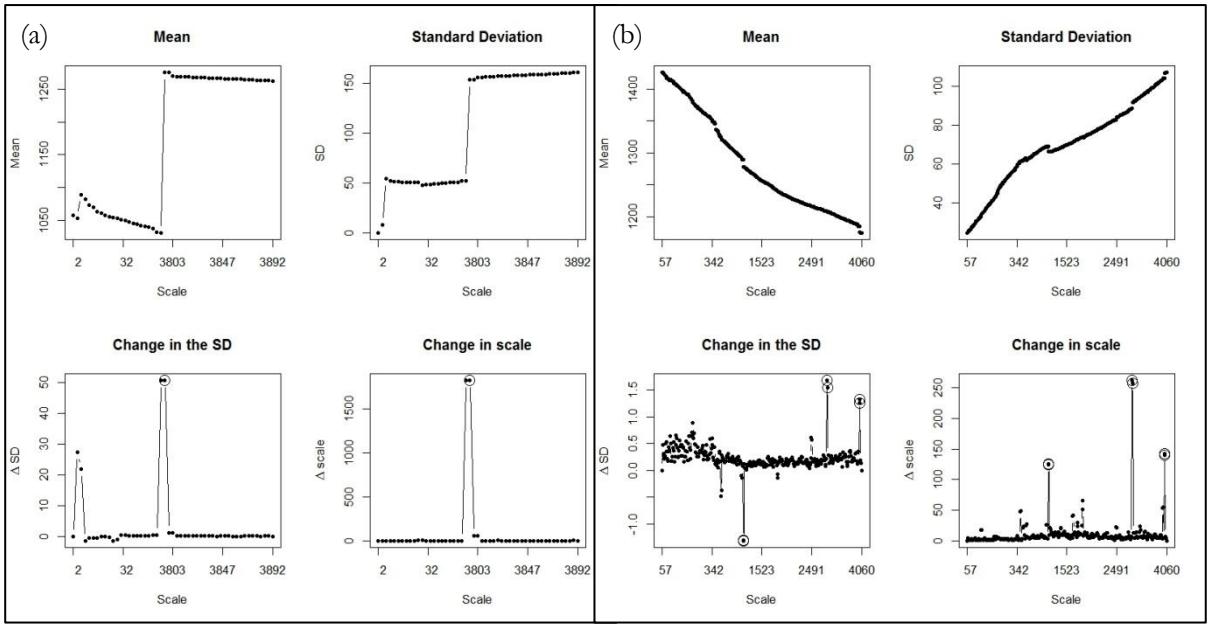
Figure 5. 12 Mean, Standard Deviation and Scale plot for Building-3 and Building-4



Identified Building-3 (left) based on Point-3 at (a) Φ_{970} (b) Φ_{1254} (c) Φ_{5142} (d) Φ_{12074} (e) Φ_{13247}
Identified Building-4 (right) based on Point-4 at (a) Φ_{514} (b) Φ_{1947} (c) Φ_{5411} (d) Φ_{6540}
(e)Φ_{7206} (f) Φ_{9216}

Figure 5. 13 Identified Building-3 and Building-4

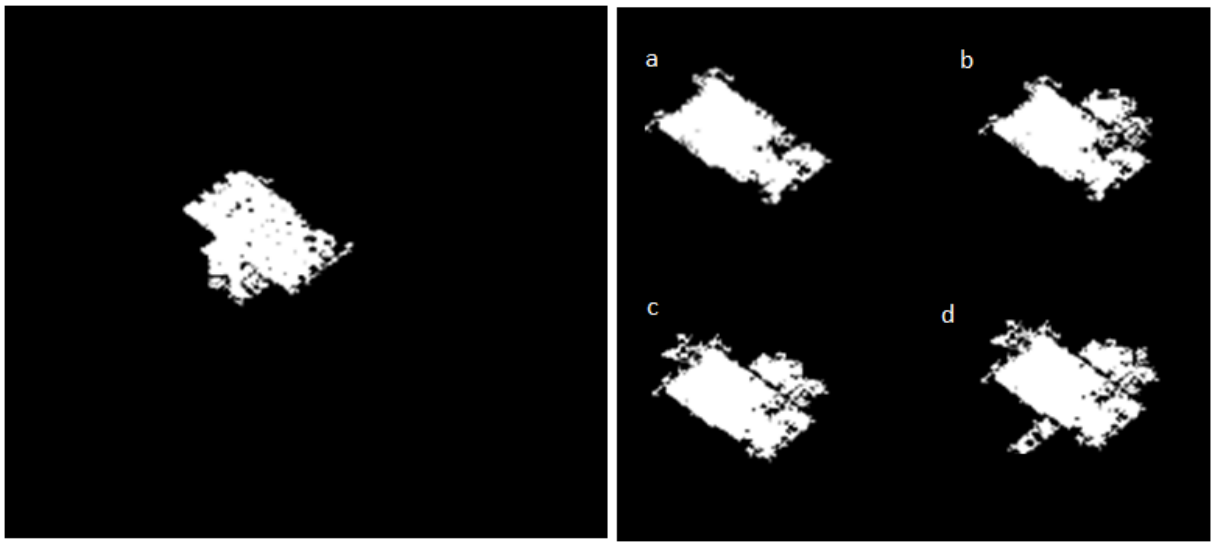
The scale of Building-5 (Figure5.14 (a)) from 10 to 3688 is showed that the standard deviation is almost constant while the mean value is decreased and only one breaking point is obtained at the scale level of 3688. The standard deviation is drastically changed at the breaking point 3688. For Building-6 the standard deviation is increased starting from the scale level 1 to 4040 and some breaking points obtained at 2692, 3204, 3764, and 4040 (Figure5.14 (b)).



(a) Building-5 mean, standard deviation and scale (b) Building-6 mean, standard deviation and scale

Figure 5. 14 Mean, Standard Deviation and Scale plot for Building-5 and Building-6

The structure of Building-5 (Figure 5.15 - left) is identified at the pulse of size Φ_{3688} with the scale of 3688. Also the identification of different structures of Building-6 (Figure 5.15 - right) is done by pulses of size Φ_{2692} Φ_{3204} Φ_{3764} and Φ_{4040} at the breaking points 2692, 3204, 3764 and 4040 in the scale ranges of 2692 to 4040.



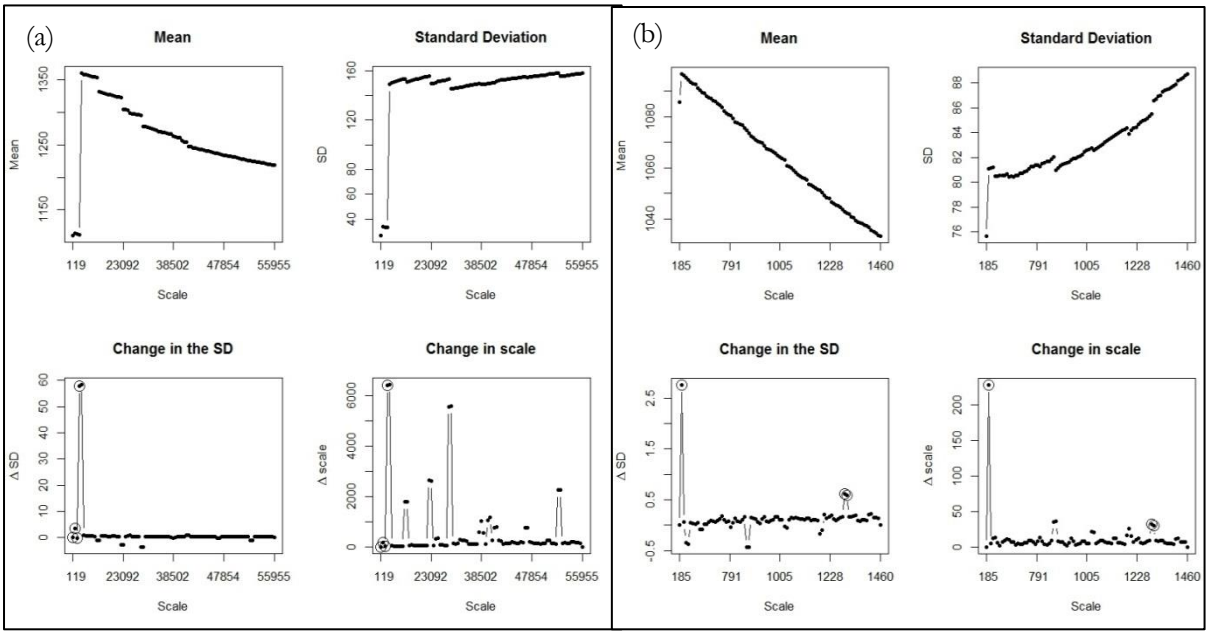
Identified Building-5 (left) based on Point-5 at (a) Φ_{3688}

Identified Building-6 (right) based on Point-6 at (a) Φ_{2692} (b) Φ_{3204} (c) Φ_{3764} (d) Φ_{4040}

Figure 5. 15 Identified Building-5 and Building-6

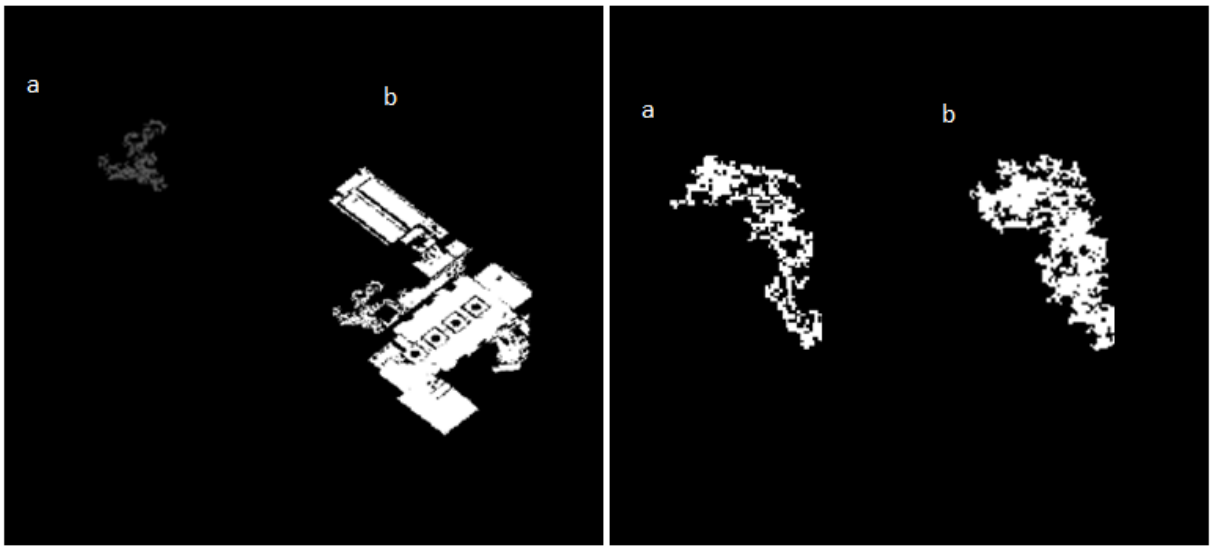
The scale of Open space-1 (Figure5.16 (a)) from 1 to 494 is showed that the standard deviation drastically changed which gives the breaking point 494. At higher scale these open space show high difference in standard deviation. For illustration the higher object which includes the open space is presented in the Figure 5.17 (left). At this point, the building includes the open space at large scale. So, it may not be

important to have such kind of pulse for identification of objects of interest such as buildings. Open space-2 (Figure5.17 (right)) has a standard deviation which increases over the scale (Figure5.16 (b)). Only two breaking points obtained to identify the open space at 636 and 1460 scale levels.



(a) Open space-1 mean, standard deviation and scale (b) Open space-2 mean, standard deviation and scale

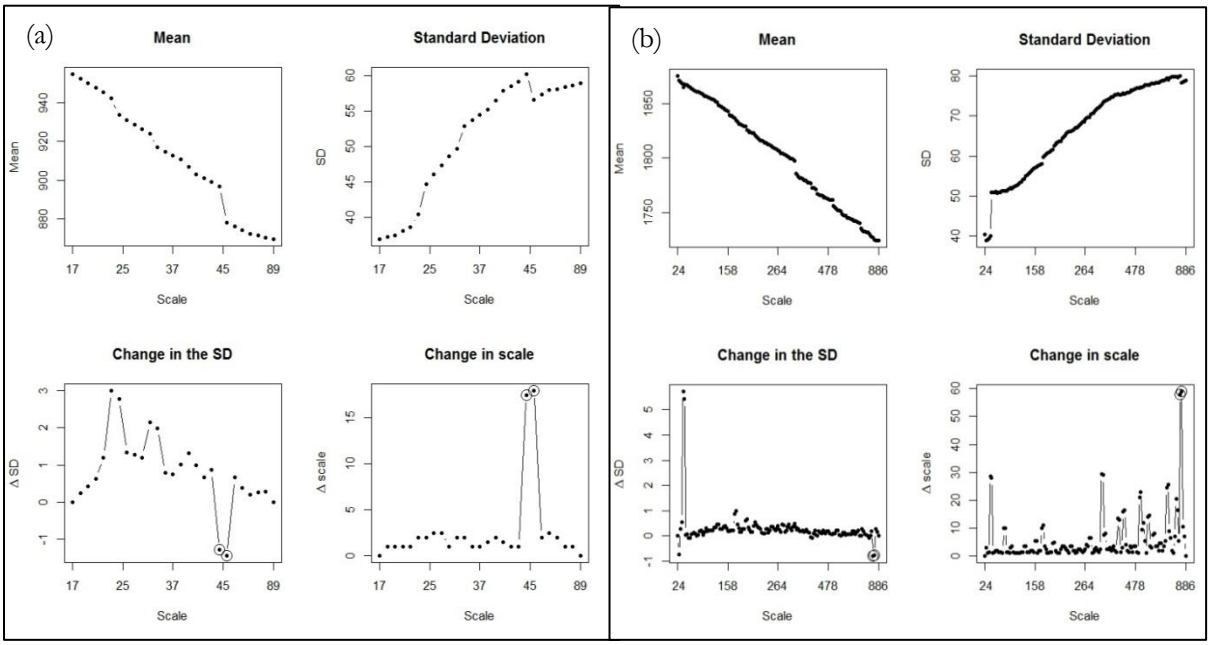
Figure 5. 16 Mean, Standard Deviation and Scale plot for Open space-1 and Open space-2



Identified Open space-1 (left) based on Point-7 at (a) Φ_{494} (b) Φ_{35231}

Identified Open space-2 (upper right) based on Point-8 at (a) Φ_{636} (b) Φ_{1460}

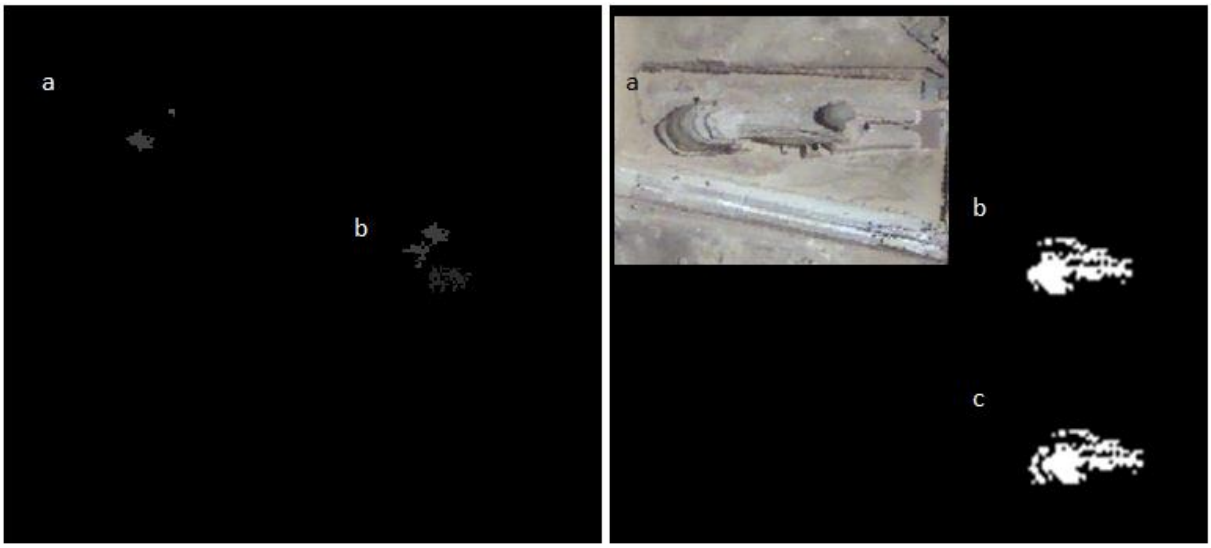
Figure 5. 17 Identified Open space-1 and Open space-2



(a) Vegetation mean, standard deviation and scale (b) The Sphinx mean, standard deviation and scale

Figure 5. 18 Mean, Standard Deviation and Scale plot for Vegetation and The Sphinx

The scale of vegetation (Figure5.18 (a)) from 1 to 45 is showed that the standard deviation drastically changed which gives the breaking point at 45. As shown in the Figure 5.19 (left) the identified vegetation area is small compared to the original image. This is due to high heterogeneity of the area that limits the connectivity of pixel at higher scale levels. In identification of the Sphinx Figure 5.19 (right), the scale plot (Figure 5.18 (b)) is showed that an incremental change in standard deviation is occurred from the scale level 24 to 886. The identified object is small in size and not fully represents the shape of the Sphinx; this is due to the similarity of the brightness value of the surrounding object with the brightness value of the Sphinx.

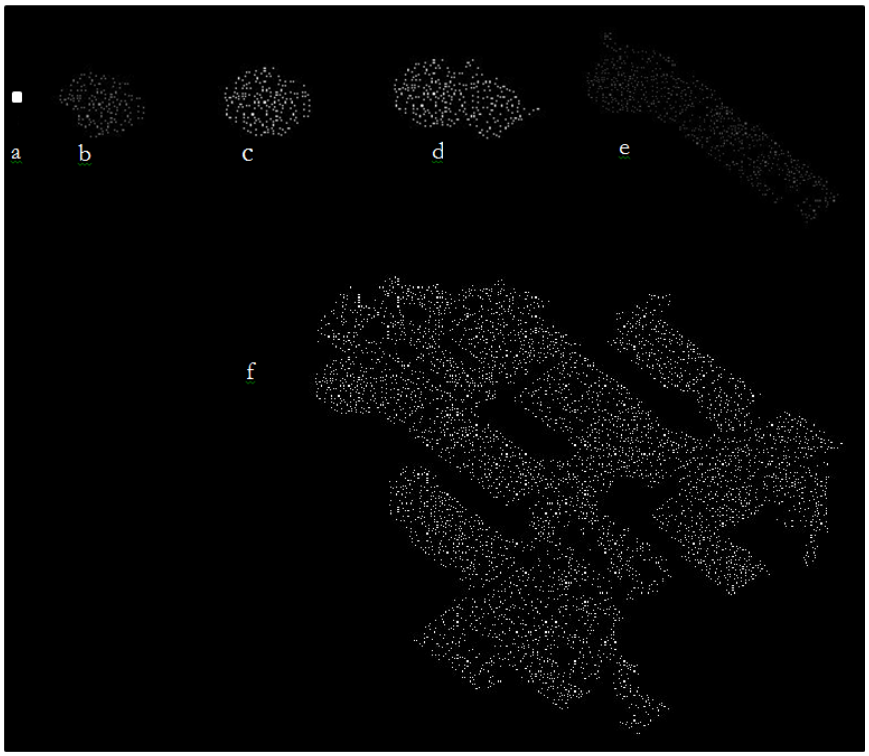


Identified Vegetation (left) based on Point-9 at (a) Φ_{45} (b) Φ_{97}

Original and Identified Sphinx (right) (a) original image (b) Φ_{860} (c) Φ_{886}

Figure 5. 19 Identified Vegetation area and the Sphinx

All objects are made up of from a number of pixels, beginning from a single pixel from the lowest scale level. In LULU-DPT scale space analysis, these objects are represented by pulses of size n and the brightness value of those pixels inside the building determines the height or depth of the pulse(Φ_{ns}) at the local maximum or minimum set. In addition, each pixel is a part of a number of pulses but not for all pulses of size n and the scale determines the life time of each pixel. Figure 5.20 shows the maximum occurred pixels at different resolution level.



(a) Φ_1 (b) Φ_{1122} (c) Φ_{1915} (d) Φ_{2997} (e) Φ_{4220} (f) Φ_{55043}

Figure 5. 20 Life time of pixels

After decomposition of the image, the aim of this study is to reconstruct or identify objects of interest at different scale level without prior knowledge of the object size and characteristics. Objects are appeared in the image in a certain ranges of scale. As well, discrete pulses from MRA also represent an object, which has a similar size to the scale level. This is the task to determine the appropriate scales during identification of objects of interest. It is performed based on the analysis of the mean to scale and the standard deviation to the scale function of the decomposed image. When the scale increases the aggregation of pixels increases, the pixels at the lower level of scale become elements of the object at higher scale level. Meanwhile, the mean of the object decreases from the lower scale level to the higher scale level, the standard deviation of the object increases from the lower scale level to the higher scale level, since the variation (heterogeneity of pixels) inside the object increases. The mean and standard deviation value of an object in a consecutive scale does not show significant change in a certain ranges of scale until the moment the object become part of other object. The mean and the standard deviation of the object in the scale plot have inverse relationship. The mean to scale and standard deviation to scale plot in Figure 5.10, Figure 5.12, Figure 5.14, Figure 5.16 and Figure 5.18 support the output in Figure 5.11, Figure 5.13, Figure 5.15, Figure 5.17 and Figure 5.19.

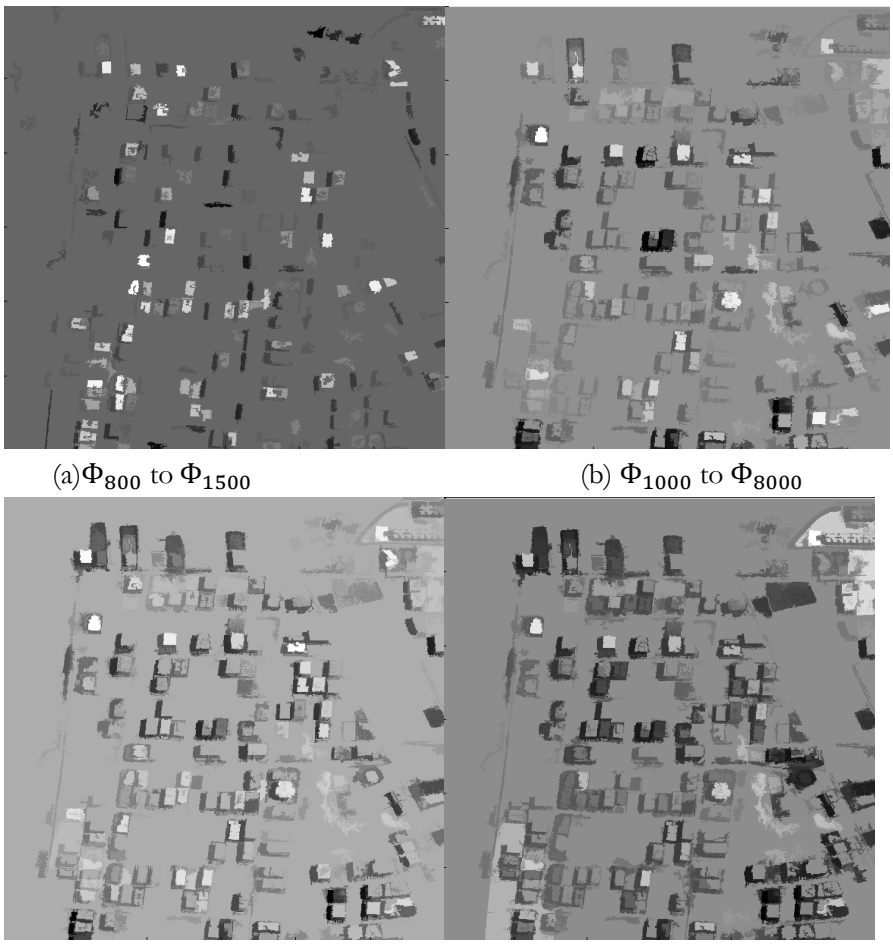
The interpretation of the scale to the standard deviation is that the object grows from a single pixel to a n number of groups of pixels having n size with n scale value. The object keeps growing as the scale increase, meanwhile, the structure and the shape of the object also shows a small change especially when the change in mean and standard deviation of the object is very low.

5.1.3. Object detection

In this section some of the illustrative outputs of DPT application in image processing are presented. After decomposition of the selected Subset image, various objects of the image are associated with a range of pulses and it is possible to identify them.

Result for Subset-1 image

In Subset-1 image, relatively to others subset images, the sizes of buildings are vary from small to medium. At the range of pulses of size Φ_{800} to Φ_{1500} (a) small buildings are detected in from the subset image, in the output (b) at the pulses of size Φ_{1000} to Φ_{8000}, medium size buildings extracted including some of the small size buildings in the subset image. The collection of buildings are extracted in (c) and (d) at the pulses of size Φ_{1500} to Φ_{6000} which is presented in Figure 5.21. Most of the medium size buildings are extracted from the subset image at the pulses of size Φ_{1500} to Φ_{8000} in the output subset image (c).



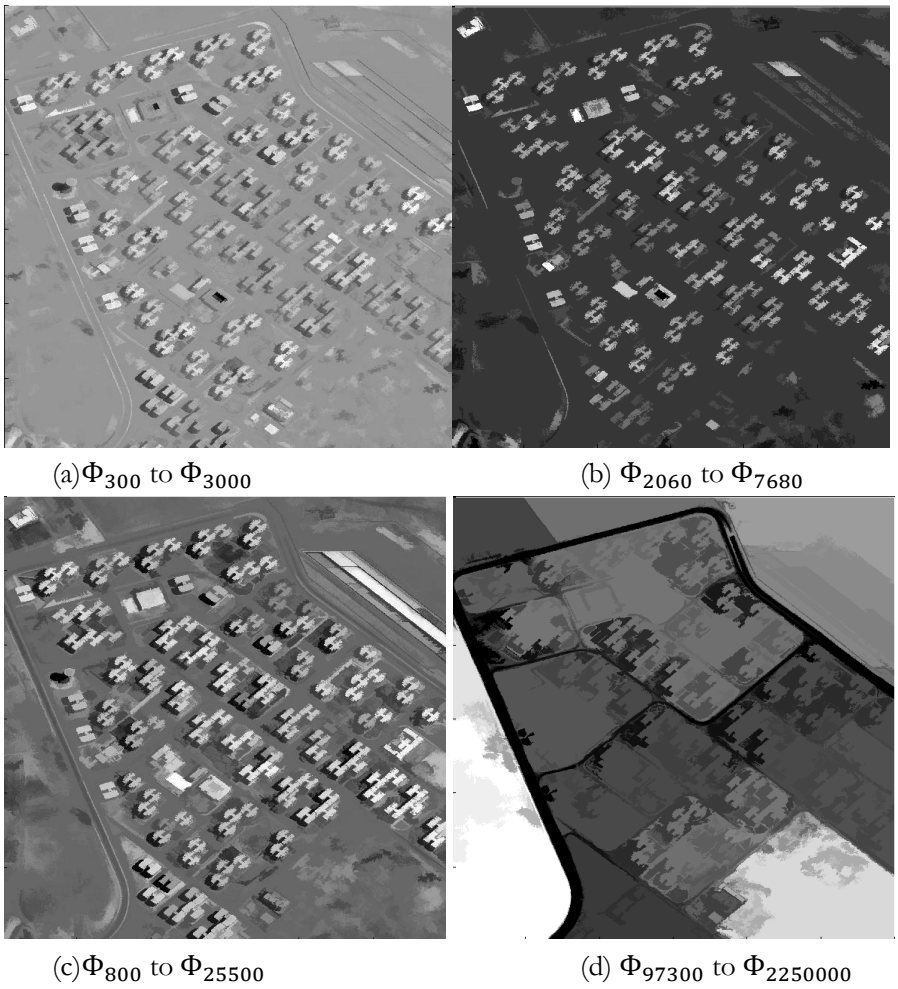
(a)Φ_{800} to Φ_{1500} (b) Φ_{1000} to Φ_{8000}

(c) Φ_{1500} to Φ_{8000} (d) Φ_{1500} to Φ_{60000}

Figure 5. 21 Detected objects from Subset-1 image

Result for Subset-2 image

In Figure 5.22 complex building structures are detected from Subset-2 image at the ranges of Φ_{300} to Φ_{25500}. At the pulses of size Φ_{300} to Φ_{3000} (a) a number of buildings and the edges of the roads are extracted. Most of the buildings are extracted at the pulses of size Φ_{2060} to Φ_{7680} (b). The complex structure of the buildings extracted at the pulses of Φ_{8000} to Φ_{25500} (c). In the output image (d) large areas without the buildings extracted from the subset image in a large size pulses.



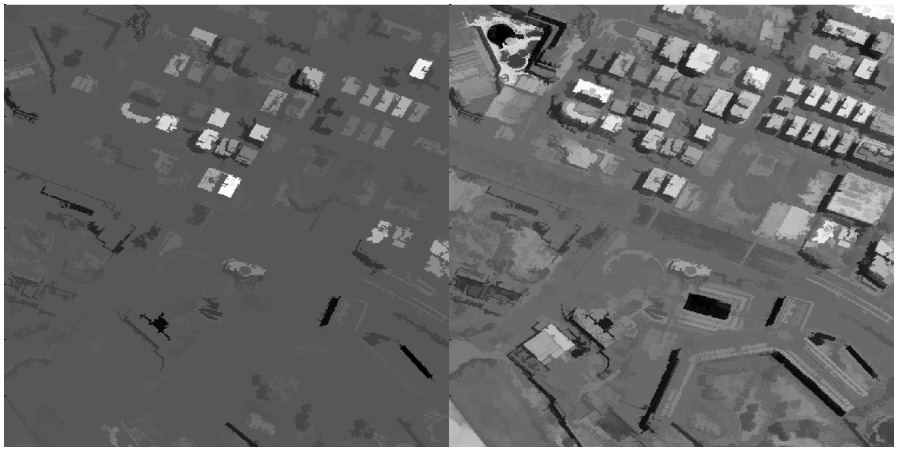
(a)Φ_{300} to Φ_{3000} (b) Φ_{2060} to Φ_{7680}

(c)Φ_{800} to Φ_{25500} (d) Φ_{97300} to $\Phi_{2250000}$

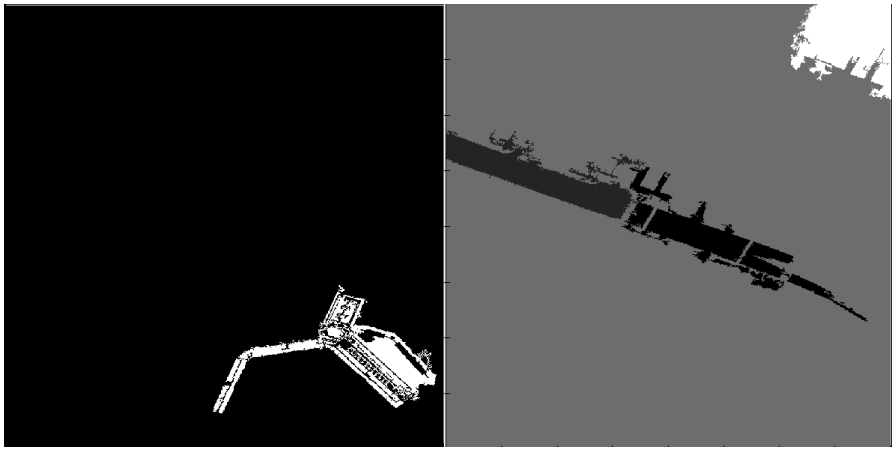
Figure 5. 22 Detected objects from Subset-2 image

Result for Subset-3 image

This subset image consist large, medium, small buildings, major roads and water bodies. In the smallest pulses of sizes from Φ_{650} to Φ_{950} (a) few small buildings are extracted as shown in figure. Both large, medium and small buildings and water bodies are extracted in the wide range of pulses of size Φ_{400} to Φ_{9100} (b). Large building structure (c) and a major road (d) extracted at the pulses of size Φ_{17700} and Φ_{18850} and Φ_{2000} respectively. Figure 5.23 presents the output for Subset-3 image.



(a)Φ_{650} to Φ_{950} (b) Φ_{400} to Φ_{9100}

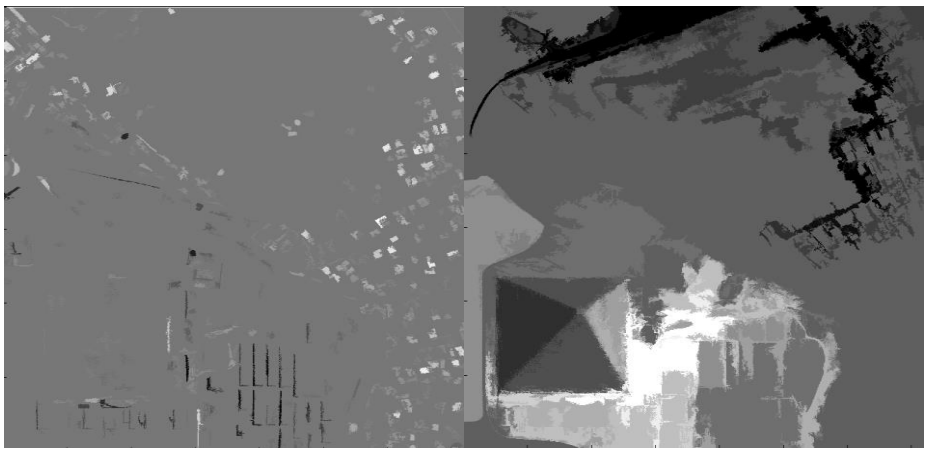


(c)Φ_{17700} (d) Φ_{18850} to Φ_{20000}

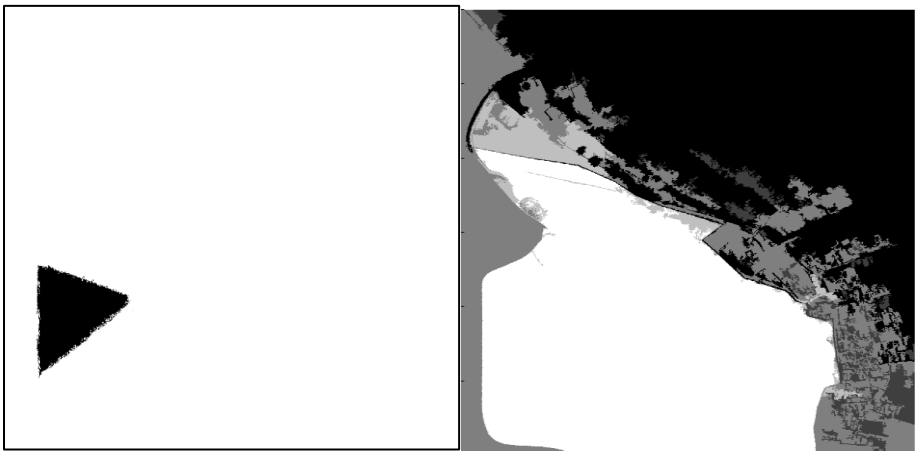
Figure 5. 23 Detected objects from Subset-3 image

Result for Subset-4 image

Pulses of size Φ_{72000} to Φ_{961000} (b) are able to extract the Great Giza Pyramid in subset-4 image along with some objects around the Pyramid due to variability of brightness of each pixel in the image. Small settlement areas are extracted at the pulses of Φ_{300} to Φ_{1000}(a) and the one of the side of the Great Giza Pyramid extracted with pulses of size Φ_{79690} (c).



(a)Φ_{300} to Φ_{1000} (b) Φ_{72000} to Φ_{961000}



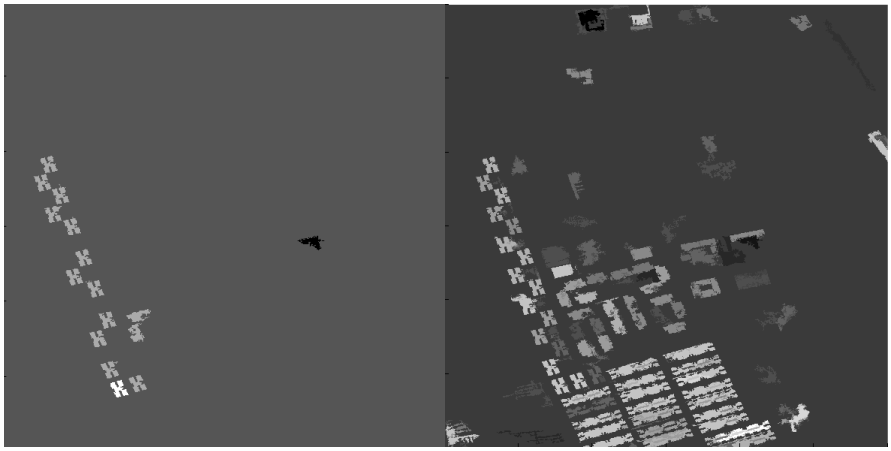
(c)Φ_{79690} (d) Φ_{961000} to $\Phi_{2700000}$

Figure 5. 24 Detected objects from Subset-4 image

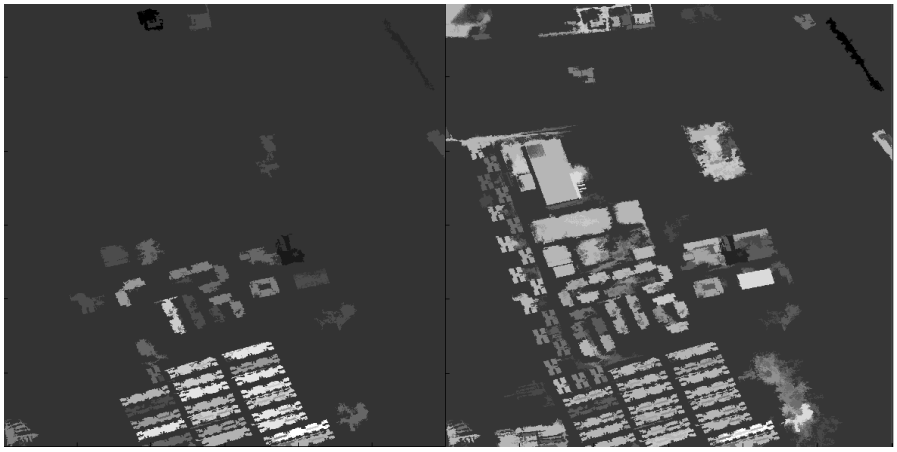
The area inside the boundary of the Pyramid is identified in the pulses of size Φ_{961000} to $\Phi_{2700000}$(d). The results for the Subset-4 are presented in Figure 5.24.

Result for Subset-5 image

In this subset small buildings are extracted in a small range of pulses in the output image (a). The structure of these small building is visible at the pulses of size Φ_{1660} to Φ_{1850}. Complex medium sized buildings and a number of small buildings detect in the output (b) with the pulses of size Φ_{1590} to Φ_{5130}. The complex medium structure buildings extracted at the range of Φ_{3290} to Φ_{5220} (c) in the output. In a wide range of pulses of size Φ_{1660} to Φ_{6870} (d) large, medium and small buildings are extracted. Figure 5.25 presents the output for Subset-5 image.



(a)Φ_{1660} to Φ_{1850} (b) Φ_{1590} to Φ_{5130}

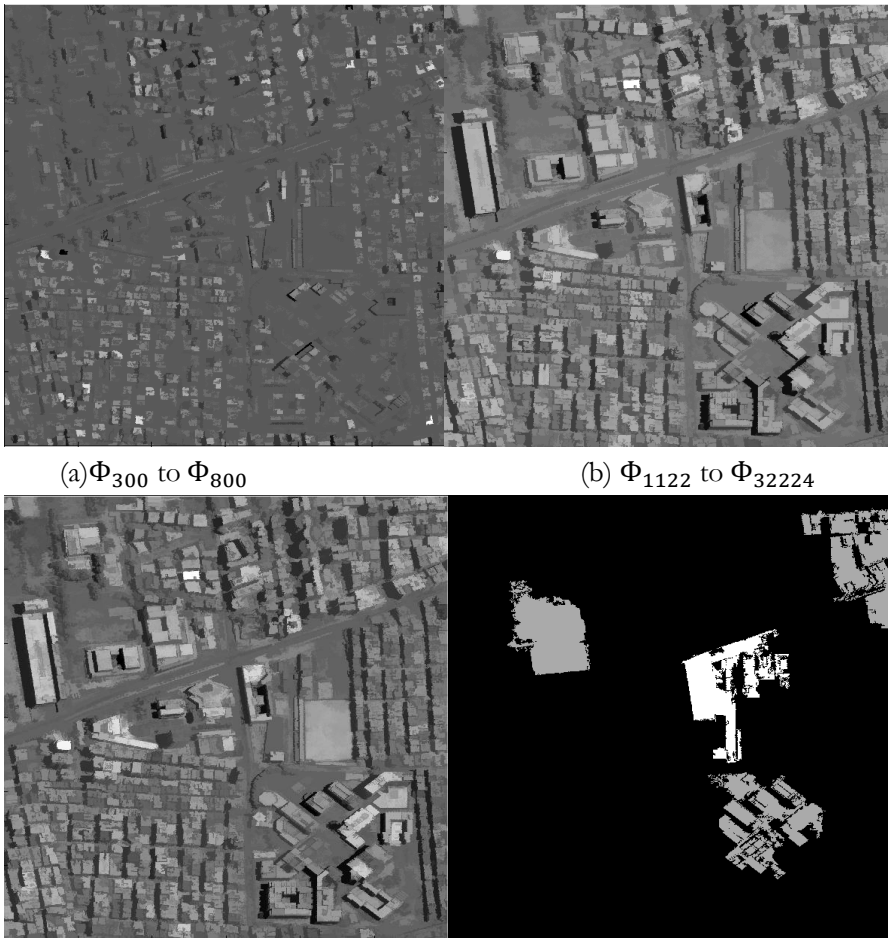


(c)Φ_{3290} to Φ_{5220} (d) Φ_{1660} to Φ_{6870}

Figure 5. 25 Detected objects from Subset-5 image

Result for Subset-6 image

This subset image is characterized by presence of complex urban features. At the small size of pulses Φ_{300} to Φ_{800} (a) small sizes buildings are extracted from the decomposed image.



(a)Φ_{300} to Φ_{800} (b) Φ_{1122} to Φ_{32224}

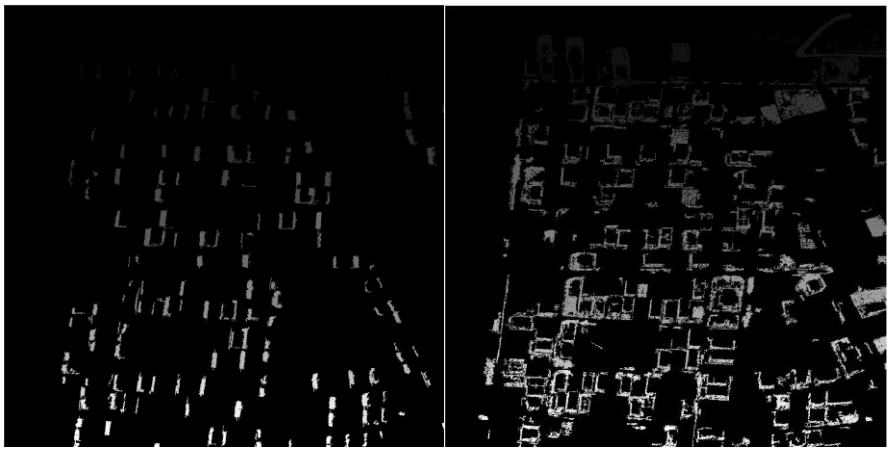
(c)Φ_{800} to Φ_{380000} (d) Φ_{52000} to Φ_{55000}

Figure 5. 26 Detected objects from Subset-6 image

Some of the medium size and the main roads extracted at the pulses of size Φ_{1122} to Φ_{32224}(b)and most of the medium and large buildigs extract in the wide range of pulses of size Φ_{800} to Φ_{380000} (c). Very large objects merged building foot prints are extracted at the pulses of size Φ_{52000} to Φ_{55000} (d). The output for Subset-6 image is presented in Figure 5.26.

Results from image segmentation by thresholding

To perform segmentation based on the brightness value, the analysis of the PCA 1 histogram of all Subset image is done and the thresholding method for image segmentation is applied. The histogram of Subset-1 image (Figure 5.1(a)) has two peaks and one transition in which at the minimum the subset image can have three segmented output. The threshold value from 450 to 670 identifies the shadow of the building where as the threshold value from 670 to 1100 identifies some of the buildings. The bare land is identified in the threshold value of 1100 to 1600. All the output of Subset-1 image is presented in Figure 5.27.

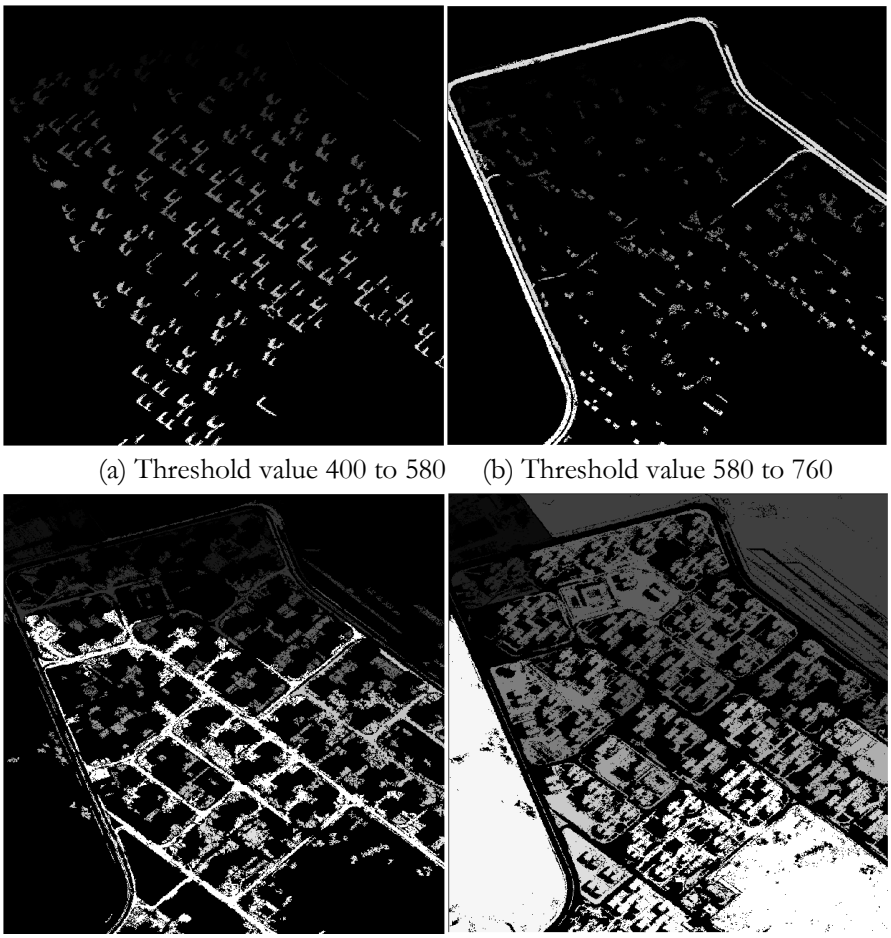


(a) Threshold value 450 to 670 (b) Threshold value 670 to 1100

(c) Threshold value 1100 to 1600

Figure 5. 27 Thresholding output of Subset-1 image

Four peak values are observed in histogram of Subset-2 image (Figure 5.1(b)), accordingly the subset image is segmented based on four threshold range values. These are threshold value from 400 to 580 which identifies the shadows of the building, from 580 to 760 extracts mainly the major road, from 760 to 1170 some of the buildings as well as the roads between the buildings and the threshold value from 1170 to 1590 identifies most of the buildings and the bare land in the subset image. Figure 5.28 presents the output of the thresholding.

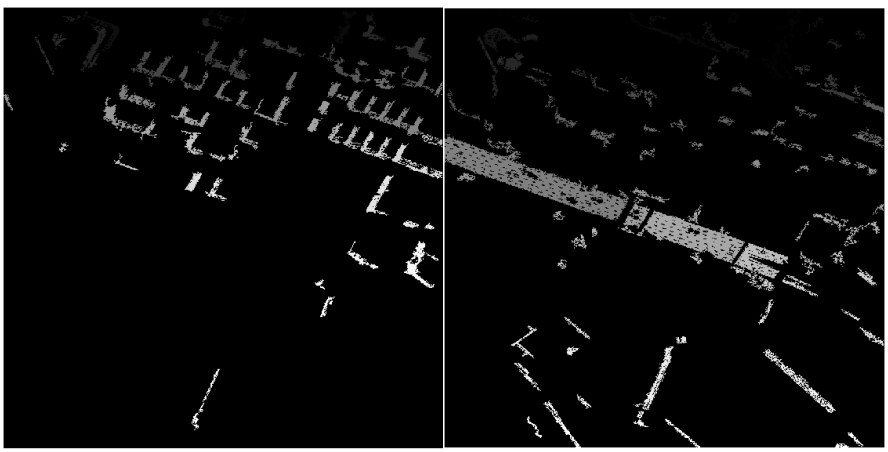


(a) Threshold value 400 to 580 (b) Threshold value 580 to 760

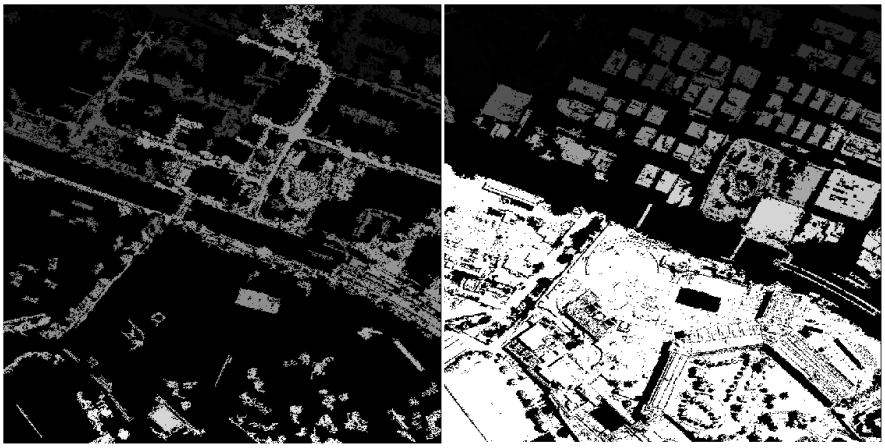
(c) Threshold value 760 to 1058 (d) Threshold value 1170 to 1590

Figure 5. 28 Thresholding output of Subset-2 image

The histogram of Subset-3 image (Figure 5.1(c)) has two distinct peaks and two transitions in which at the minimum the subset image can have four segmented output. The threshold value from 380 to 535 identifies the shadow of the building where as the threshold value from 535 to 680 mainly identifies the major road. The road structure around the parcels is identified in the threshold value from 680 to 925.



(a) Threshold value 380 to 535 (b) Threshold value 535 to 680



(c) Threshold value 680 to 925 (d) Threshold value 925 to 2000

Figure 5. 29 Thresholding output of Subset-3 image

Threshold value from 925 to 2000 identifies the buildings in the upper part and bare land and buildings in the lower part of the subset image. All the output of Subset-3 image is presented in Figure 5.29.

In histogram of Subset-4 image (Figure 5.1(d)) two peak and one transition values are observed, accordingly the subset image is segmented based on three threshold range values. These are threshold value from 400 to 760 which extract small vegetation areas and small houses, from 760 to 1200 extracts mainly the vegetated land, road and the settlement areas and the threshold value from 1200 to 2000 identifies the sides of Giza Pyramid and the surroundings bare land in the subset image. Figure 5.30 presents the output of the thresholding.

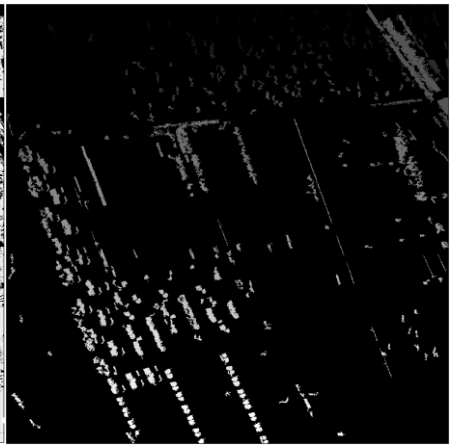
(a) Threshold value 400 to 760 (b) Threshold value 760 to 1200

(c) Threshold value 1200 to 2000

Figure 5. 30 Thresholding output of Subset-4 image

The histogram of Subset-5 image (Figure 5.1(e)) has two distinct peaks and one transition in which at the minimum the subset image can have three segmented output. The threshold value from 400 to 640 identifies the vegetation areas where as the threshold value from 640 to 1100 identifies the major road. Threshold value from 1100 to 1600 mainly identifies the buildings and small areas of bare lands. All the output of Subset-5 image is presented in Figure 5.31.

(a) Threshold value 400 to 640

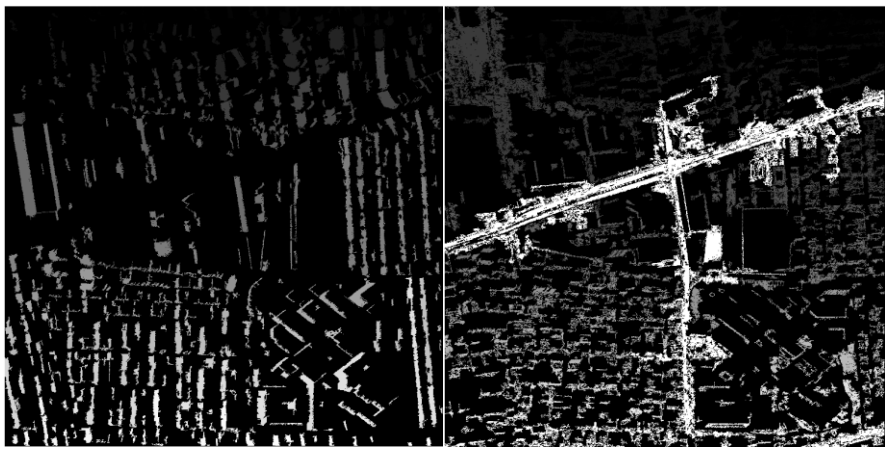


(b) Threshold value 640 to 1100

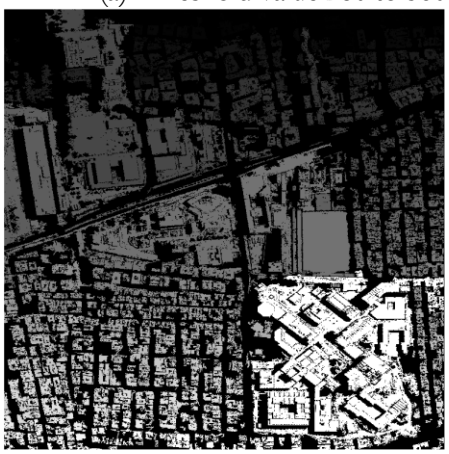
(c) Threshold value 1100 to 1600

Figure 5. 31 Thresholding output of Subset-5 image

In histogram of Subset-6 image (Figure 5.1(f)) two peak and one transition values are observed, accordingly the subset image is segmented based on three threshold range values. These are threshold value from 280 to 560 which extract small shadows of buildings, from 560 to 880 extracts mainly the main road and small buildings and the last threshold value from 880 to 1570 identifies mainly large footprint of buildings and bare lands in the subset image. Figure 5.25 presents the output of the thresholding.



(a) Threshold value 280 to 560 (b) Threshold value 560 to 880



(c) Threshold value 880 to 1570

Figure 5. 32 Thresholding output of Subset-6 image

Object identification from the decomposed image is mainly depending on the size of the object that is related to the pulse size obtained from the DPT. It is also depending on the composition or the amount of information contained in the image. The histogram of Subset-1 image (Figure 5.1(a)) shows that at a minimum three different classes object is present in the image. Based on visual interpretation, small and medium buildings and the bare land (open space) are visible in the scene. Since, small and medium sized buildings exist, they were identified by small sizes of pulses of the decomposed image as presented in Figure 5.21. When the extracted buildings compared with the result from the thresholding method (Figure 5.27), buildings are better extracted using the DPT from LULU operator. The reason behind is thresholding method only use the brightness value as the only criteria to segment the image in over a certain ranges of values.

The visual interpretation of the Subset-2 image (Figure 5.1(b)) exhibits that buildings, roads and open spaces are exists in the image. Also, the histogram of the subset image (Figure 5.1(b)) shows nearly four peaks values and one of the peak is dominant in the image. The structures of the buildings also complex and their size is varies from small to medium. The objects in this subset are better extracted by the DPT

algorithm than the threshold method since the thresholding method (Figure 5.28) extract only some of the area of interest like the road from the subset while the DPT algorithm extract all areas of interest in the medium size pulses from the decomposed image.

Objects extracted from the decomposed Subset-3 image ranges from small size building to the large objects (Figure 5.1(c)).The histogram of this subset image shows that at least four distinct classes are present in the image. The thresholding method (Figure 5.29) mainly extracts the road and some of the buildings. Whereas, the DPT-algorithm is extracts all areas of interest distinctly from the decomposed image based on multi-resolution analysis.

Objects such as the Great Giza Pyramid from Subset-4 image, complex building structures from Subset-5 and distinct large and medium sized buildings from Subset-6 are better extracted from the corresponding decomposed image based on the multi-resolution analysis application of the LULU operators in image processing than the thresholding method. The histogram (Figure 5.2(d)) of the Subset-4 image shows two peak values which districts the Pyramids from the settlement areas, while the thresholding method did not identify the Pyramids from the surrounding objects (Figure 5.30). In the case of Subset-5 image histogram (Figure 5.2(e)), three peaks exist and one of them is dominating in the image. The corresponding thresholding method result (Figure 5.31) shows that only some of the buildings are extracted from the image. Subset-6 image histogram (Figure 5.2(f)) has two distinct peak points, since the scene consist mainly the buildings and the roads. The thresholding method results in Figure 5.32identifies the collection of buildings rather than individual building. The multi-resolution analysis allows extracting of object of interest at a particular level of scale from the decomposed image.

5.2. Evaluation

Quantitative assessment of the segmentation output to the reference data objects are selected such as buildings and roads. For image segmentation accuracy assessment three criteria were chosen. These are calculating the number of segmented object inside the reference objects, the overlap percentage of a segmented object which is calculated based on the area of the largest extracted object from the segmentation that fit best to the reference objects and the Area fit Index. The results of the three criteria presented in Table 5.4.

To assess the accuracy of the extracted object, first the objects were overlaid with reference polygons. This is to decide whether the extracted object from the decomposed image is located in the reference object or not based on their geometric center. This method is only applied to the extracted building. Then evaluation of the segmented output for accuracy assessment is performed on the subset of Subset-6, Subset-2 and Subset-1 image. Figure 5.33 shows the segmented object from the LULU DPT and the corresponding reference data digitized from the GeoEye-1 image.

(a) Subset-6 image (b) Subset-2 image (c) Subset-1 image

Figure 5. 33 Segmented output of Subset-6, Subset-2 image and Subset-1 image

The reference data for subset of Subset-6 has two objects, the building and the road. The open space covers large area of the subset than the other objects.

Reference	Area (m^2)	Number of pixels	Number of objects
Building	29012.27	115681	41
Road	973.70	3882	1
Open space	45960.53	183254	1
Total	75946.5	302817	43

Table 5. 2 Referenced data for evaluation of Subset-6 image

The number of buildings extracted from the subset image is 39 in number and 2 buildings are missed during identification when it is compared with the reference data. The area covered by the extracted object and the number of pixel are presented in the following Table 5.3.

Segmented	Area (m^2)	Number of Pixels	Number of objects	Missed object	Overestimated	Total
Building	26110.8	104110	39	2	3	44
Road	769.223	3067	1	0	0	1
Open space	49066.47	195640	1	0	0	1
Total	75946.5	302817	41	2	4	46

Table 5. 3 Segmented output from Subset-6 image

For the extracted building from the subset Subset-1 image the overlap percentage is 90% to the reference data. The road is overlapped at 79 %. The correctness of the output is calculated from the percentage of number of objects to the Total number of objects, missed and overestimated objects. The building achieved 88.63%, whereas the road achieved 100% since the reference and the extracted road has 1 element. The AFI of the building is 0.10 and the overlap percentage is less than hundred percent which indicates that the reference object is over-segmented.

	Correctness %	Overlap %	AFI
Building	88.63%	90.00%	0.10
Road	100.00%	79.00%	0.21
Open space	100.00%	106.76%	-0.07

Table 5. 4 Outputs of evaluation for Subset-6 image

Using boundary matching, the segmented object boundary of Subset-6 image was fitted to reference object boundary to calculate the $D(b)$. The lowest fit boundary $D(b)$ = 1.5 is obtained since, the objects were over segmented. Therefore, the corrected value $D(B)_{corr}$ applied and obtained $D(B)_{corr}$=0.90 and $D(b)$= 0.7 at this point the number of boundary pixels is proportional to the segmented image and the referenced data. If a perfect fit exist the value of $D(b)$ is equals 0.0.

The output segmented objects from Subset-2 image also evaluated for accuracy assessment by comparing from the reference data. The reference data for Subset-2 has two objects, the building and the road. The open space covers almost 85% of the subset total area. The following table presents the area covered by the reference object, number of pixel contained and the count of those object.

Reference	Area (m^2)	Number of pixels	Number of objects
Building	65814.38	262419	77
Road	27293.79	108827	1
Open space	471191.50	1878754	1
Total	564300.00	2250000	79

Table 5. 5 Referenced data for evaluation of Subset-2 image

The number of buildings extracted from the subset image is 75 in number and 2 buildings are missed during identification when it is compared with the reference data. The area covered by the extracted object and the number of pixel are presented in the following Table 5.6.

Segmented	Area (m^2)	Number of Pixels	Number of objects	Missed object	Overestimated	Total
Building	58574.80	233552	75	2	6	83
Road	22107.97	88150	1	0	0	1
Open space	483617.24	1928298	1	0	0	1
Total	564300.00	2250000	77	2	6	85

Table 5. 6 Segmented output from Subset-2 image

For the extracted building from the subset Subset-1 image the overlap percentage is 89% to the reference data. The road is overlapped at 81 %. The correctness of the output estimated 90.36% to the building, whereas the road achieved 100%. The AFI of the building is 0.11 and the overlap percentage is less than hundred percent which indicates that the reference object is over-segmented.

	Correctness %	Overlap %	AFI
Building	90.36%	89.00%	0.11
Road	100.00%	81.00%	0.19
Open space	100.00%	102.64%	-0.03

Table 5. 7 Outputs of evaluation for Subset-2 image

The output segmented objects from Subset-1 image also evaluated for accuracy assessment by comparing from the reference data. The reference data for Subset-1 has one object, the building. The open space covers almost 90% of the subset total area. The following table presents the area covered by the reference object, number of pixel contained and the count of those object.

Reference	Area (m^2)	Number of pixels	Number of objects
Building	53994.10	215293	147
Open space	510304.63	2164025	1
Total	564300.00	2250000	148

Table 5. 8 Referenced data for evaluation of Subset-1 image

The number of buildings extracted from the subset image is 141 in number and 6 buildings are missed during identification when it is compared with the reference data. The area covered by the extracted object and the number of pixel are presented in the following Table 5.9.

Segmented	Area (m^2)	Number of Pixels	Number of objects	Missed object	Overestimated	Total
Building	48540.69	193543	141	6	10	157
Open space	515759.31	2056457	1	0	0	1
Total	564300.00	2250000	142	6	10	158

Table 5. 9 Segmented output from Subset-1 image

For the extracted building from the subset Subset-1 image the overlap percentage is 89.90% to the reference data. The correctness of the output estimated 89.81% to the building. The AFI of the building is 0.10 and the overlap percentage is less than hundred percent which indicates that the reference object is over-segmented.

	Correctness %	Overlap %	AFI
Building	89.81%	89.90%	0.10
Open space	100.00%	101.07%	-0.01

Table 5. 10 Outputs of evaluation for Subset-1 image

6. DISCUSSION

This study shows an attempt to study the application of LULU operators for extraction of information from remotely sensed image. The method used to identify objects of interest based on these operators in one and two dimensional analysis is novel for analysis of remotely sensed images.

One of the advantages of remotely sensed image is it provides information that is confined in the multi-spectral bands of the image. Identification of man-made features from the remotely sensed VHR image is advantageous in application of urban studies. The LULU-DPT one dimension algorithm as well as The Discrete Pulse Transform in 2D algorithm is limited to a single band implementation, then, to compromise the limitation the first principal component analysis is applied to the input multi-spectral image. The PCA is used to extract information from the multispectral image by reducing the number of bands and retaining the most of the variability of the bands.

LULU operators for image segmentation and object detection studied based on the merits of the operator such as theoretically determined non-linearity property and multi-resolution analysis which is supported by shape preservation, consistent separation, and hierarchal decomposition properties. Since the operator basically developed for one dimensional analysis and later on extended to multi-dimension analysis of signals, so it worth to explore to the opportunities in one and two dimensions for identification of objects from the remote sensed image in this study. The other advantage of these operators is it provides discrete pulses. The recursive application of the operator decomposes input signals into discrete pulses, which have a constant value in connected set and zero elsewhere. These discrete pulses are used for identification of object during the process of reconstruction of the object of interest. Depending on the size and variability of input image, the decompose image, for example, 500 rows by 500 columns image may have from 100, 000 to 200,000 pulses. One real problem is to relate an object of interest (e.g. a building) and pulses produce by DPT, since the object can be related to many pulses smaller than the object and also are a part of large objects, and not all pulses are equally important during identification of the objects of interest. To overcome this problem, in this study, the scale space analysis is applied to automatically identify the appropriate pulses for identification of object interest from the generated pulses.

The discrete pulses from the multi-resolution analysis of LULU operators preserve trend in the ranges of scale level. This advantage of trend preservation properties of the operator is mainly used in the scale-space analysis to identify objects of interest, since pixels in the lower scale become elements of the pulses in the higher scale. The principle of scale-space in this study is based on a given point from the image by linking the growth region of these points to each pulses of the decomposed image, calculating the mean, standard deviation, the change in mean and standard deviation over the scale and determine the breaking points based on threshold values at which object of interest change its state in the decomposed image. The other advantage of scale-space analysis is it excludes large size pulses and lower size pulses since the threshold values are stated in the algorithm. At higher level of scale, pulses become larger in size and include different objects in the image, also at lower level of scale; pulses become smaller in size, which might not be important in identification of objects of interest.

Identification of object of interest from a VHR remotely sensed image using pulse generated by LULU operators is one of the possible ways of extraction of information. It creates the opportunities to explore its applicability in the field of image processing. Compared to the thresholding segmentation method, complex urban features such as buildings and apartment blocks are easily identified by pulses. While, the continuity between successive scale levels was not observed during identification of object that consist heterogeneous pixels, since the heterogeneity of the area limits the connectivity of pixel at higher scale level and the output is not satisfactory at this point. Another unsatisfactory result is obtained when the object of interest pixel brightness value is similar with the surrounding pixel brightness value, the connectivity of pixel in the area increased and the object of interest merged with other objects and represented by a single pulse.

In this study, different segmentation accuracy assessment methods were applied and different accuracy values were obtained depending on the composition of the subset image. The AFI values for building identification were 0.10 and 0.11. Thus, the largest segmented objects was less than the reference object that means the reference objects are larger than the segmented object, and the maximum overlap was 90%. For the extracted roads, the AFI values were 0.19 and 0.21. The object boundary fit between the segmented object and the reference is 0.7 after the correction factor 0.90 obtained and applied; at this point the number of boundary pixels is proportional to the segmented image and the referenced data. The reference objects are over-segmented since the largest segmented object was less than the reference object. The maximum overlap between the extracted object and the reference was 81%.The accuracy for the building object as well as the roads was high since the resolution of the GeoEye-1 image has fine resolution in multi-spectral bands for detection of individual buildings and roads. On the other hand, in respect to spectral responses of each pixel, identifying the edge of building and the road is extremely difficult since the brightness value of the pixel at the edge of the building and the road shows high variability.

7. CONCLUSION AND RECOMENDATIONS

7.1. Conclusion

The main objective of this research was to explore LULU operators and associated DPT (LULU-DPT) for object identification from remotely sensed images. To achieve the objectives of the study, the corresponding research questions were formulated and have been addressed during the study.

To develop a segmentation method based on the LULU operators for remotely sensed image, first the essential properties of the operator such as shape preservation, consistent separation, hierarchal decomposition and total variation were studied. These properties of the operator reviewed from previous researches and analysed their importance in image segmentation and object detection.

To apply the LULU-DPT segmentation method to VHR multi-spectral remotely sensed image, the image was processed to obtain the first PCA. Since bands are correlated each other PCA is applied to reduce the number of bands comprising a dataset and also to identify hidden patterns in the data.

To fulfil objectives of the study, such as multi-resolution analysis of LULU operators, the application of DPT decomposes the image into lower components. To identify object from the decomposed image, the study was implemented in one and two dimensions analysis. LULU operators inherently has MRA in which images are decomposed into $M \times N$ scale level and each level may have different objects of the image in different size and shape except for zero valued DPTs' levels. i.e., there are some DPT levels that have zero values since the local maximum–minimum or the local minimum–maximum does not exist in those particular scale levels. Image is reconstructed by either by full reconstruction from all non-zero DPT levels or partial reconstruction of objects of interest from certain level of non-zero DPT levels.

For objectives of scale space application automatic identification of object of interest, first by providing all points from the image the algorithm generates the breaking point when the object show drastic change in size and shape based on the analysis of the change in mean, standard deviation and scale level by introducing threshold values. Therefore, scale (scale of DPT) selection for identification of object of interest mainly based on the selection of those break points in MRA of decomposed image over the scale function. The first derivative of the function is used to identify when the size of object is drastically changed in scale space.

Different accuracy assessment methods used to evaluate the output of the segmentation from the LULU operators. This methods are area fit index, percent overlap and correctness percentage. In addition the results of LULU operators are compared with thresholding segmentation methods. Generally the output results are high value based on the segmentation accuracy assessment criteria such as index values and percentages.

7.2. Recommendations

In this study a method to identify objects is successfully implemented based on LULU operators to serve image segmentation on a VHR remotely sensed image. In addition, further analysis on this method may contribute to enrichment of the research in image segmentation method in image analysis field of study. So, the recommendations from this study that needed more exploration are:

- Extend the method to work with multi-bands analysis.
- Apply the method for identification of objects in regular pattern natural features.
- Explore the potential of the method in three dimensions images analysis.

LIST OF REFERENCES

Acharya, T., and Ray, A. (2005). *Image Processing: Principles and Applications*. Hoboken, New Jersey: John Wiley & Sons.

Anguelov, B. (2008). *Discrete Pulse Transform of images: Algorithm and applications.* Paper presented at the ICPR 2008: 19th International Conference on Pattern Recognition, 8-11 December, Tampa, Florida, USA.

Anguelov, R. (2006). LULU Operators and Locally δ Monotone Approximations *Constructive Theory of Functions* (pp. 22-34). Sofia: Marin Drinov Academic Publishing House.

Anguelov, R., and Fabris-Rotelli, I. (2008). Discrete Pulse Transform of Images In A. Elmoataz, O. Lezoray, F. Nouboud & D. Mammass (Eds.), *Image and Signal Processing* (Vol. 5099, pp. 1-9): Springer Berlin / Heidelberg.

Anguelov, R., and Fabris-Rotelli, I. (2010). LULU operators and discrete pulse transform for multidimensional arrays. *IEEE Transactions on Image Processing, 19*(11), 3012-3023. doi: http://dx.doi.org/10.1109/tip.2010.2050639

Anguelov, R., and Rohwer, C. (2009). LULU operators for functions of continous argument. *Quaestiones Mathematicae, 32*(2), 187-202. doi: http://dx.doi.org/10.2989/qm.2009.32.2.2.795

Bhanu, B., Sungkee, L., and Das, S. (1995). Adaptive image segmentation using genetic and hybrid search methods. *IEEE Transactions on Aerospace and Electronic Systems, 31*(4), 1268-1291. doi: http://dx.doi.org/10.1109/7.464350

Bong, C.-W., and Rajeswari, M. (2011). Multi-objective nature-inspired clustering and classification techniques for image segmentation. *Applied Soft Computing, 11*(4), 3271-3282. doi: http://dx.doi.org/10.1016/j.asoc.2011.01.014

Congalton, G. (1991). A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment, 37*(1), 35-46.

Conradie, W., de Wet, T., and Jankowitz, M. (2006). Exact and asymptotic distributions of LULU smoothers. *Journal of Computational and Applied Mathematics, 186*(1), 253-267. doi: http://dx.doi.org/10.1016/j.cam.2005.03.073

Delves, L. M., Wilkinson, R., Oliver, C. J., and White, R. G. (1992). Comparing the performance of SAR image segmentation algorithms. *International Journal of Remote Sensing, 13*(11), 2121-2149. doi: http://dx.doi.org/10.1080/01431169208904257

du Toit, J. (2007). *The Discrete Pulse Transform and Applications.* Msc Thesis, University of Stellenbosch, South Africa.

Fabris-Rotelli, I. (2009). *LULU operators on multidimensional arrays and applications.* MSc. Thesis, University of Pretoria, Pretoria, South Africa.

Fabris-Rotelliy, I., and van der Walt, S. (2009). *The Discrete Pulse Transform in Two Dimensions.* Paper presented at the proceedings of the 20th Annual Symposium of the Pattern Recognition Association of South Africa (PRASA), November 2009, Stellenbosch, South Africa.

Fu, S., and Mui, K. (1981). A survey on image segmentation. *Pattern Recognition, 13*(1), 3-16. doi: http://dx.doi.org/10.1016/0031-3203(81)90028-5

Gertel, J., and Samir, S. (2002). *Cairo: urban agriculture and visions for a "modern" city in Growing Cities, Growing Food: Urban Agriculture on the Policy Agenda: A Reader on Urban Agriculture, edited by Nico Bakker, Marielle Dubbeling, Sabine Guendel, Ulrich Sabel-Koschella, & Henk de Zeeuw.* Available at RUAF Foundation, http://ruaf.org/sites/default/files/Cairo.PDF.

Glasbey, C., and Horgan, G. (1995). Image Analysis for the Biological Sciences: Wiley.

Haar, A. (1910). Zur Theorie der orthogonalen Funktionensysteme. *Mathematische Annalen, 69*(3), 331-371. doi: 10.1007/bf01456326

Howeidy, A., Shehayeb, D., Göll, E., Abdel Halim, K., Séjourné, M., Gado, M., . . . Cobbett, W. (2009). *Cairo's Informal Areas Between Urban Challenges and Hidden Potentials: Facts. Voices. Visions.* Egypt: German Technical Cooperation.

Jankowitz, M. (2007). *Some Statistical Aspects of LULU Smoothers.* Dissertation PhD, Stellenbosch University, South Africa.

Kao, O. (2001). Modification of the LULU operators for preservation of critical image details. In H. R. Arabnia (Ed.), *Proceedings of the International Conference on Imaging Science, Systems and Technology, Vols I and Ii* (pp. 280-286). Athens: C S R E a Press.

Laurie, D. (2011). The Roadmaker's algorithm for the discrete pulse transform. *IEEE Transactions on Image Processing, 20*(2), 361-371. doi: http://dx.doi.org/10.1109/tip.2010.2057255
Lee, H. (2007). Wavelet Analysis for Image Processing. Retrieved from http://disp.ee.ntu.edu.tw/henry/wavelet_analysis.pdf
Lucieer, A., Kraak, M. J., and Stein, A. (2004). *Uncertainties in segmentation and their visualisation.* 113, ITC, Enschede. Retrieved from http://www.itc.nl/library/Papers_2004/phd/lucieer.pdf
Malkowsky, E., and Rohwer, C. (2004). The LULU-semigroup for envelopes of functions. *Quaestiones Mathematicae, 27*(1), 89-97. doi: http://dx.doi.org/10.2989/16073600409486086
Mallat, G. (1989). A theory for multiresolution signal decomposition: the wavelet representation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on, 11*(7), 674-693. doi: http://dx.doi.org/10.1109/34.192463
Ogden, R. T. (1997). *Essential wavelets for statistical applications and data analysis*. Boston etc.: Birkhäuser.
Pal, R., and Pal, K. (1993). A review on image segmentation techniques. *Pattern Recognition, 26*(9), 1277-1294. doi: http://dx.doi.org/10.1016/0031-3203(93)90135-j
Peijun, L., Jiancong, G., Benqin, S., and Xiaobai, X. (2011). A multilevel hierarchical image segmentation method for urban impervious surface mapping using very high resolution imagery. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 4*(1), 103-116. doi: http://dx.doi.org/10.1109/jstars.2010.2074186
Pitas, I., and Venetsanopoulos, N. (1990). *Nonlinear Digital Filters: Principles and Applications*: Springer.
Rohwer, C. (1989). Idempotent one-sided approximation of median smoothers. *Journal of Approximation Theory, 58*(2), 151-163. doi: http://dx.doi.org/10.1016/0021-9045(89)90017-8
Rohwer, C. (1999). Projections and separators. *Quaestiones Mathematicae, 22*(2), 219-230. doi: http://dx.doi.org/10.1080/16073606.1999.9632077
Rohwer, C. (2002a). Multiresolution Analysis with Pulses. In M. Buhmann & D. Mache (Eds.), *Advanced Problems in Constructive Approximation* (Vol. 142, pp. 165-186): Birkhäuser Basel.
Rohwer, C. (2002b). Variation reduction and LULU-smoothing. *Quaestiones Mathematicae, 25*(2), 163-176. doi: http://dx.doi.org/10.2989/16073600209486008
Rohwer, C. (2004). Fully trend preserving operators. *Quaestiones Mathematicae, 27* (3), 217-229. doi: http://dx.doi.org/10.2989/16073600409486096
Rohwer, C. (2005). *Nonlinear Smoothing and Multiresolution Analysis* (Vol. 150): Birkhäuser Basel.
Rohwer, C. (2007). The estimation of moments of an unknown error distribution in the discrete pulse transform. *Numerical Algorithms, 45*(1), 239-251. doi: http://dx.doi.org/10.1007/s11075-007-9116-0
Rohwer, C., and Laurie, D. (2006). The Discrete Pulse Transform. *SIAM Journal on Mathematical Analysis, 38*(3), 1012-1034. doi: http://dx.doi.org/10.1137/040620862
Rohwer, C., and Toerien, L. (1991). Locally monotone robust approximation of sequences. *Journal of Computational and Applied Mathematics, 36*(3), 399-408. doi: http://dx.doi.org/10.1016/0377-0427(91)90019-G
Rohwer, C., and Wild, M. (2002). Natural alternatives for one dimensional median filtering. *Quaestiones Mathematicae, 25*(2), 135-162. doi: http://dx.doi.org/10.2989/16073600209486007
Rohwer, C., and Wild, M. (2007). LULU Theory, Idempotent Stack Filters, and the Mathematics of Vision of Marr. In H. Peter (Ed.), *Advances in Imaging and Electron Physics* (Vol. 146, pp. 57 - 162): Elsevier.
Serra, J. (1982). *Image Analysis and Mathematical Morphology*. London: Academic Press.
Serra, J. (1988). Mathematical Morohology for Boolean Lattices. In J. Serra (Ed.), *Image Analysis and Mathematical Morphology Volume2:Theoretical Advances*. London: Academic Press.
Sezgin, M., and Sankur, B. (2004). Survey over image thresholding techniques and quantitative performance evaluation. *Journal of Electronic Imaging 13*(1), 146-165.
Stein, A., and De Beurs, K. (2005). Complexity metrics to quantify semantic accuracy in segmented Landsat images. *International Journal of Remote Sensing, 26*(14), 2937-2951. doi: http://dx.doi.org/10.1080/01431160500057749
Sweldens, W., and SchrSder, P. (2000). Building Your Own Wavelets at Home In R. Haagmans (Ed.), *Wavelets in Geosciences*. Berlin: Springer.
Tarbush, N. (2012). Cairo 2050: Urban Dream or Modernist Delusion? *Journal of International Affairs, 65*(2), 171-186.
Thakare, P. (2011). A study of image segmentation and edge detection techniques. *International Journal on Computer Science and Engineering (IJCSE), 3* (2), 899-904.

Vision Research lab-UCSB. (2008, 07-August-2008). Image segmentation Retrieved 20-May, 2012, from http://vision.ece.ucsb.edu/segmentation/

Wang, J., Tang, J., Liu, J., Ren, C., Liu, X., and Feng, J. (2009). Alternative fuzzy cluster segmentation of remote sensing images based on adaptive genetic slgorithm. *Chinese Geographical Science, 19*(1), 83-88. doi: http://dx.doi.org/10.1007/s11769-009-0083-3

Zhan, Q., Molenaar, M., Tempfli, K., and Shi, W. (2005). Quality assessment for geo-spatial objects derived from remotely sensed data. *International Journal of Remote Sensing, 26*(14), 2953-2974. doi: http://dx.doi.org/10.1080/01431160500057764

APPENDIX

Appendix

Code for LULU-DPT one dimension algorithm, Using R

```
#=====================================================================
minop <-function(y)         # Minimum operator
{
      n <- length(y)
      y_ls <- c(y[1],y[-n]) # Adds the copy of the first element on the left
      return(pmin(y,y_ls)
)
}
maxop <-function(y)         # Maximum operator
{
      n <- length(y)
      yr <- y[n]
      y_rs <- c(y[-1],yr)    # Adds the copy of the last element on the right
      return(pmax(y,y_rs))
}
Ln <-function(y,n)         # Lower operator
{
      val <- y
      for(i in 1:n)
      {
            val <- minop(val)
      }
      for(i in 1:n)
      {
            val <- maxop(val)
      }
      return(val)
}
Un <-function(y,n)        # Upper operator
{
      val <- y
      for(i in 1:n)
      {
            val <- maxop(val)
      }
      for(i in 1:n)
      {
            val <- minop(val)
      }
      return(val)
}
Cn <-function(y,n)          # C 'ceiling' operator for LnUn
{
      if(n==0)return(y)
      else return(Ln(Un(Cn(y,n-1),n),n))
}
Fn <-function(y,n)           # F 'floor' operator for UnLn
```

```
{
      if(n==0)return(y)
      else return(Un(Ln(Fn(y,n-1),n),n))
}
# Input image
Root <- "D:/"
img <- read.table(paste(Root,"/",".txt",sep=""), header = FALSE)
class(img)
dim(img)
G = as.matrix(img)
class(img)

# Histogram of y as a 1D signal
windows()
h <- hist(as.vector(img), breaks=2000, xlim=c(-1, 3000),  xlab='DN
value', main='Histogram of subset Image')
x <- h$mids
y <- h$counts
n <- length(y)

# Add zero values on the both sides:
m <- length(y)
y <- c(rep(0,m),y,rep(0,m))

# Compute the complete DPT without recursion
DPT <- array(0,c(length(y),n))
Cyn <- array(0,c(length(y),n))
n.pulses <- array(0,n)
pulses <- array(0,c(n,n))
k<-1
Clow <- y
while(k<=n)
{
      Cup  <- Ln(Un(Clow,k),k)
      Cyn[,k] <- Clow
      tmp <- Clow - Cup
      DPT[,k] <- tmp
      ind <- which(abs(tmp)>0)
      if(length(ind>0))
      {
            n.pulses[k] <- length(ind)/k
            for(l in 1:n.pulses[k])
            {
                  pulses[k,l] <- ind[1]
                  ind <- ind[-(1:k)]
            }
      }
      Clow <- Cup
      k <- k+1
}
```

```
# removes zeros
y <- y[-(1:m)]
y <- y[-((m+1):(2*m))]
DPT <- DPT[-(1:m),]
DPT <- DPT[-((m+1):(2*m)),]
k<-max(n.pulses)
pulses <- pulses[,1:k]

windows()
par(mfrow=c(1,2))
plot(x,y,type="s", main="Original signal")
plot(x,rowSums(DPT),type="s",main="Reconstruction from DPT",
xlab='x', ylab='Row Sums (DPT)')

E <- array(0,n)
for(k in 1:n)E[k]<-sqrt(sum((DPT[,k])^2)/k)
ind <- which(E>0)
n1 <- length(ind)
if(ShowAll)
{
      if(n1>0) for(k in 1:n1)
      {
      windows()
      plot(DPT[,ind[k]],type="s",main=paste("level",ind[k],sep=""))
      }
}
# Selective reconstruction from DPT
Recon <- array(0,length(y))
min_level <- 1
max_level <- length(y)
Recon <- rowSums(DPT[,min_level:max_level])
windows()
par(mfrow=c(1,2))
plot(x,y,type="s", main="Original signal")
plot(x,Recon,type="s",main=paste("Reconstruction from DPT Levels
",min_level," to ",max_level,sep=""), xlab='x',
ylab='Reconstruction')

# Selective reconstruction from DPT
# Only non-empty levels
Recon <- array(0,length(y))
ind <- which(E>0)
if(length(ind)>0)
Recon <- rowSums(DPT[,ind])
windows()
par(mfrow=c(1,2))
plot(x,y,type="s", main="Original signal")
plot(x,Recon,type="s",main=paste("Reconstruction from non-empty
levels ",sep=""), xlab='x', ylab='Reconstruction')

#--- The END ---#
```

Code for Scale space automatic object identification algorithm, Using R

```
#=============================================================
# Read DPT pulse images (geotif)
#=============================================================
setwd(Path_DPT)

Files <- list.files(".",pattern=".tif")
Nf <- length(Files)
level_arr <- array(0,Nf)
Ns <- Nf
DPT <- array(0,M*N)
DPTm <- array(0,c(M,N)
indpix <- (pt_x-1)*M + pt_y
j0 <- floor(indpix/M)+1
i0 <- indpix - (j0-1)*M
lev_contain <- array(0,0)
gamma_arr <- array(0,0)
pmean <- array(0,0)
psd   <- array(0,0)

for(n.f in 1:Nf)
{
      fstr <- Files[n.f]
      npos <- str_locate(fstr,"_t44")[1]
      level_txt <- substr(fstr,start=3,stop=npos-1)
      level <- as.numeric(level_txt)
      level_arr[n.f] <- level
      inputfile <- fstr
      Pulse <- readGDAL(inputfile)
      D <- Pulse@data$band1
      Dm <- matrix(D,nrow=M,ncol=N,byrow=FALSE)
      if(D[indpix]!=0)
      {
            lev_contain <- c(lev_contain,level)
            ind2 <- which(D!=0)
            gamma_n <- length(ind2)/level
            gamma_arr <- c(gamma_arr,gamma_n)
            if(gamma_n>1)
            {
                  jarr <- floor(ind2/M)+1
                  iarr <- ind2 - (jarr-1)*M
                  seg <- Segment(Dm,D[indpix],i0,j0)
                  ind2 <- seg$pn
                  #test
                  if(TEST==TRUE)
                  {

      plot(iarr,jarr,main=paste("level=",level,sep=""),xlim=c(1,M),yl
im=c(1,N),xlab="col",ylab="row")
                        yarr <- floor(ind2/M)+1
                        xarr <- ind2 - (yarr-1)*M
```

```
                        points(xarr,yarr,col="red")
                    }
                }
                pvec <- P@data$band1[ind2]
                msmat <- MS@data[ind2,]
                pmean <- c(pmean,mean(pvec))
                psd   <- c(psd,sd(pvec))
        }
}
Nlev <- length(lev_contain)
scalespace <- data.frame(array(0,c(Nlev,3)))
names(scalespace) <- c("scale","mean","sd")
ind <- order(lev_contain)
scalespace$scale <- lev_contain[ind]
scalespace$mean <- pmean[ind]
scalespace$sd <- psd[ind]

write.table(scalespace,file=paste(Root,"\\scalespace_x_",pt_x,"_y_",
pt_y,".txt",sep=""))
}
setwd(Start_Path)

# Compute changes
deriv_fun <- function(f){
      val <- c(f[-1],0) - c(0,f[-length(f)])
      # Cut off boundaries
      val[1] <- 0
      val[length(val)] <- 0
      0.5*val
}
scalespace$dm_ds <- deriv_fun(scalespace$mean)
scalespace$dsd_ds <- deriv_fun(scalespace$sd)
ticks <- seq(1, Nlev, length.out=5)
windows()
par(mfrow=c(2,2))
plot(1:Nlev,scalespace$mean,type="b",xlab="Scale",ylab="Mean",xaxt='
n',pch=20, col=1)
axis(side=1,ticks,labels=format(scalespace$scale[ticks]))
title("Mean")
plot(1:Nlev,scalespace$sd,type="b",xlab="Scale",ylab="SD",xaxt='n',p
ch=20, col=1)
axis(side=1,ticks,labels=format(scalespace$scale[ticks]))
title("Standard Deviation")

ind <- which((abs(scalespace$dsd_ds)>=
1.5)&(deriv_fun(scalespace$scale)>=50)& scalespace$scale >500)

plot(1:Nlev,scalespace$dsd_ds,type="b",xlab="Scale",ylab=expression(
paste(Delta," ","SD")),,xaxt='n',pch=20, col=1)
axis(side=1,ticks,labels=format(scalespace$scale[ticks]))
title("Change in the SD")
```

```
points(ind,scalespace$dsd_ds[ind], cex=2,col=1)
plot(1:Nlev,deriv_fun(scalespace$scale),type="b",xlab="Scale",ylab=e
xpression(paste(Delta," ","scale")),xaxt='n',pch=20, col=1)
axis(side=1,ticks,labels=format(scalespace$scale[ticks]))
title("Change in scale")

points(ind,deriv_fun(scalespace$scale)[ind], cex=2,col=1)
scalespace[ind,]

#--- The END ---#
```