CHAPTER 4

Results and Discussion

Section 1

Edge detection via $3 \cdot x (x = 1, 2, 3)$ spatial convolution windows

Section 2

Evaluation of image enhancing techniques for remotely sensed vegetation image interpretation

Section 3

Thresholding techniques for segmentation of remotely sensed vegetation data
SECTION 1

Edge detection via $3 \times x (x = 1, 2, 3)$

spatial convolution windows
4.1.1 INTRODUCTION

We propose a detailed analysis of spatial domain filtering via $3X (X = 1, 2, 3)$ convolution windows using Mather's image processing system (MIPS). A total image-derivative filter, augmenting to the category of second order derivative filters is proposed and evaluated for Landsat (TM, MSS) image data. A generalization of first order and second order derivative filters is suggested for the ease of implementation. A brief description of digital image environment is also given for novice readers. Digital image processing (DIP) is emerging as a leading branch of computer science which is reverberating with almost all the state of the art disciplines of science and technology, like remote sensing, digital pattern recognition, medical image segmentation and delineation, forensic impression identification, gas bubble size distribution estimation for various small-scale air-sea interaction processes, in situ microscopic analysis of cells in bioreactors, fluorescent measurement of air-water gas transfer, thermographic analysis in botany, X-ray astronomy and planetology, image depth measurements, and in dynamic calibration of robots etc. The general goal of DIP for scientific and technical applications is to use radiation emitted by objects. An imaging system collects the radiation to form an image. Then, image-processing techniques are used to perform an area-extended measurement of the object feature of interest. For scientific and technical applications, area-extended measurements constitute a significant advantage over point measurements [42], since also the spatial and not only the temporal structure of the signals can be acquired and analyzed. The goal of 2-D DIP can be divided into three principal categories: determine the geometry, photometry, and spatial structure of objects. In the simplest case, the measuring task requires only measurement of position, size, and form of the objects.
4.1.2 A Digital Image

“A photograph or an image is equivalent to thousand word” (a Chin’s quotation). But at the advent of digital images we can consider an image to be composed of million of words, which we call bites or gray-levels or in latest terminology, pixel (or pel/picture, elements). With the aid of existing very fast and high memory computing facilities, we, now are independent to assess, arrange, manipulate, and play with these wonderful words to write our own stories (or even novels.)

The term image or scene stands for a landscape or a view as seen by a spectator. The scene, indoor or outdoor, we see around us, in general, are three-dimensional. However, there are scenes, which can be considered as two dimensional, for example, flat terrain seen by satellite, X-ray graphs, microscopic biological images and documents printed on paper. An image may be defined in a number of ways.

(i) Photographic definition

According to Oxford dictionary, the term image is defined as the optical appearance of something produced in a mirror or through a lens. Thus, a two dimensional representation of three-dimensional world is known as an image or a photograph.

(ii) Physical definition

In physical terminology, an image is defined as a two dimensional distribution of energy or intensity. It is called an analog image if the intensity distribution is continuous and digital or discrete if intensity distribution is discrete.
(iii) Mathematical Definition

In mathematical point of view on image may be defined as a two-dimensional function \( f(y, x, t) \), where \( x, y \) are position coordinates and \( t \) is time. \( f(x, y, t) \) is a continuous function, and in the case of digital image it is a discrete function. Since the image intensity can never be negative, neither it is infinite so the image function is a bounded function.

\[ 0 \leq f(x, y, t) \leq \alpha \]

If the intensity is constant at every position with respect to time, then \( t \) does not come into play, the function is written as \( f(x, y) \) otherwise.

(iv) Digital Definition

According to digital terminology an image is defined as a two dimensional array of numbers. Each number represents a particular gray level or a pixel. In this way a digital image can be represented in form of a matrix of intensity function.

<table>
<thead>
<tr>
<th>( f(0,0) )</th>
<th>( f(0,1) )</th>
<th>( f(0,2) )</th>
<th>( f(0,3) )</th>
<th>( f(0,4) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( f(1,0) )</td>
<td>( f(1,1) )</td>
<td>( f(1,2) )</td>
<td>( f(1,3) )</td>
<td>( f(1,4) )</td>
</tr>
<tr>
<td>( f(2,0) )</td>
<td>( f(2,1) )</td>
<td>( f(2,2) )</td>
<td>( f(2,3) )</td>
<td>( f(2,4) )</td>
</tr>
<tr>
<td>( f(3,0) )</td>
<td>( f(3,1) )</td>
<td>( f(3,2) )</td>
<td>( f(3,3) )</td>
<td>( f(3,4) )</td>
</tr>
</tbody>
</table>

Each function \( f(x, y) \) has a certain weight of binary digits. If the weight of binary digit is \( 'K' \) then total number of gray levels in the image are given as \( L = 2^k \). The image represented by this method is said to \( L \) – bit data. In particular case if \( K = 1 \) then \( L = 2 \), the image is called a binary image.
(v) Proposed definition

Thinking towards some irrelevant way, we can say that, “Every object in this world is an image; we can’t see any object, what we see is only its image formed at retina.”

“Thus, a digital image is a two or three dimensional space where the intensity function obeys some discrete mathematical laws.”

\[ I_{2d} = \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} f(x,y,t) \]

\[ I_{3d} = \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} \sum_{z=0}^{L-1} f(x,y,z,t) \]

4.1.2.1 Why a Digital Image

The philosophy of DIP is deeply rooted in the digitalization (Quantization and Sampling). The processing is meaningless for an analog image unless it is not digitalized. It may be cleared by following examples;

(i) Every construction starts with bricks. Smaller is the size of the brick; more is the freedom to have variety of design.

(ii) Living world is made of cells (digits)

(iii) Matter as well as radiation is not continuous, i.e. made of discrete entities.

(iv) A continuous chain does not exist, if so, it will have no flexibility.

Here, it is the term flexibility in processing or arrangement, which is the hero of the plot. Once, any system is broken or available in form of small units then we have a great
degree of freedom to play with according to our own desirer. The smallest unit of a
building is brick, of the living world is cell, of the matter/radiation is atom/photon, of the
chain is its links, and so the same way when we are talking about an image, picture or
photo, the smallest unit or building block is a pixel. Once we break an image into pixels,
then it is very easy to play with them, to arrange them and to manipulate them as per our
requirement. The reason is that in a discrete system, there is more degree of freedom in
comparison to a continuous system that is why most of the systems in nature are discrete,
if not, they are made to be.

4.1.3 FILTERING

Any one can say that a device, which stops or checks some particular elements and
lets others to pass or vice versa, from a variety of individuals, is a filter and it is true. An
image-processing filter does not perform any extraordinary job apart from one stated
above. A digital image is composed of variety of gray levels. For a single band 8-bit
image, there are 256 gray levels, ranging from minimum 0 to maximum 255. For a three
channel 8-bit image, there would be 256×256×256 gray levels. The visual appearance of
an image depends upon the degree of variation in the gray levels. If degree of variation is
small, the image appears smooth and pleasant. It is true up to a certain limit, if crossed;
the image informations become blurred and meaningless. On the other hand, if there are
incomparable variations or large fluctuations in the gray levels, the image appears sharp,
which is also bounded to a certain limit after which image loses its meaning. The slowly
varying smoothening gray levels are called, as low frequency components while highly
varying sharpening gray levels are known as high frequency components. Image filtering
is the process of masking low or high frequency components to extract some features of
interest at the expense of some information loss.
The tools and algorithms performing the operation are called low or high pass filters. There are two categories of image domain; spatial domain and frequency domain, and so are the corresponding filters. *The present study is strictly carried out in spatial domain.* The image field in which all the operations are performed directly on the pixels is known as spatial domain. Filtering in spatial domain is a neighborhood operation where the influence of some neighboring pixels is considered on the central pixel. This is achieved by taking a specifically weighted two-dimensional array of numbers, known as sub-image, mask, kernel, template, or window. The process of moving a kernel through the image is called as convolution, where the value of central pixel is modified by the influence of the content of kernel, according to some specific law. A 3×3 and a 5×5 kernel masks are shown below.

![3x3 kernel](image1)

![5x5 Kernel](image2)

In the normal filtering process, the central pixel is replaced by the algebraic weighted average of all the content of the mask. It \( w_i \) is the weight of sub-image pixel and \( w_k \) is the corresponding weight on the kernel then the normal filter operation is defined as:

\[
I_0 \approx \frac{1}{n^2} \sum (w_i \times w_k)
\]
Where \(I_{f_0}\) is the output-filtered image. This is conventional neighborhood operation. We can define and evaluate other neighborhood operations as given below:

(i) \[I_{f_0} \approx \frac{1}{n^2} \sum (w_i + w_k)\]

(ii) \[I_{f_0} \approx \frac{1}{n^2} \sum (w_i - w_k)\]

(iii) \[I_{f_0} \approx \frac{1}{n^2} \sum \frac{w_k}{w_i}\]

(iv) \[I_{f_0} \approx \frac{1}{n^2} \sum \left[\frac{(w_i - w_k)}{(w_i + w_k)}\right]\]

(v) \[I_{f_0} \approx \frac{1}{n^2} \sum \left\{\frac{(w_i + w_k)}{(w_i - w_k)}\right\}\]

For all such operations, the pixels lying on the border remain unaffected. It can be proved that for \(N \times M\) image, convolved with an \(n \times n\) kernel masks, the number of border pixels that remain unaffected is \((n-1) [M + N - (n-1)]\). Thus a \(3 \times X (X = 1, 2, 3)\) kernel is the most suitable configuration for convolution operation.

4.1.4 EDGES AND DISCONTINUITIES

Edges and discontinuities are the building block elements of the information content of an image data. Contour of images or objects, or in other words ‘edges’ in the paradigm of DIP and computer vision provide valuable informations towards human image understanding [43]. Probably the most important image-processing step in human picture recognition system consists of edge detection process. Naturally, edge detection has
become a serious challenge to the community of image processors and analysts and since the last two decades, in particular, numerous publications have been detailing methodologies and algorithms for edge detection. How do edges come in an image? Probably the answer to this question provides the early important clue for locating edges in an image. The variations of image features, usually brightness give rise to edges [44]. More objectively, the edges are the representations of the discontinuities of image intensity function. Therefore, edge detection algorithm is essentially a process of detection of these discontinuities in an image data. Since abrupt change in brightness value indicates edge, its detection in binary, or segmented images is quite straightforward. However, the process of edge location is quite complex in the case of gray level or intensity images. The transition in intensity in gray-scale images is relatively smooth in nature rather than abrupt as in the case of segmented or binary images. Application of derivative operators on intensity images produces another image, usually called ‘gradient image’ as it reveals the rate of intensity variation [45].

In grayscale images, the edges are local features that within a neighborhood separate two regions in each of which the gray level is more or less uniform with different values on the two sides of the edge [46]. So an ideal edge has step like cross-section. The process of edge detection is broadly classified into two categories; one is derivative approach in which edge-pixels are detected by taking derivative followed by thresholding (e.g. Roberts operator and 4-neighbor operator). They occasionally incorporate noise-cleaning scheme (e.g. Prewitt operator and Sobel operator). Two-dimensional derivatives are computed by means of what we call edge mask. Second is pattern fitting approach, in which a series of edge approximating functions in the form of edge templates over a small neighborhood are analyzed. Parameters along with their properties corresponding to the
best fitting function are determined. Based on these informations the presence or absence of edge is decided. In the present study, we have employed the former approach in various modified forms. First order and second order derivatives are used. The beauty of these operators is that they can be easily procured in discrete formulation and their implementation via a $3 \times X$ ($X = 1, 2, 3$) kernel windows is facilitated in all image processing software of any level. Higher order derivatives may also be considered but their implementation via a $3 \times X$ ($X = 1, 2, 3$) convolution windows is very complex. Further, if they are managed to operate, anyway, the results are not satisfactory. The derivatives, or in discrete domain difference operators, which yield high values at places where graylevel changes abruptly, are used to find the gradient of an image [47]. The derivatives of a digital function are defined in terms of differences. There are various ways to define these differences. However, it is required that any definition we use for a first order derivative: (1) Must be zero in the flat segments; (2) Must be non-zero at the onset of a graylevel step or ramp; and (3) Must be non-zero along ramp [48]. Same is true for second order derivatives. Since we are dealing with digital quantities whose values are finite, the maximum possible graylevel change is also finite, and the shortest distance over which that change can occur is the distance between adjacent pixels (Gonzales and Wood 2002).

4.1.5 FIRST ORDER DERIVATIVE AND CORRESPONDING $3 \times 3$ KERNEL MASKS

From elementary calculus, we know that first order derivative of a continuous function $f(x)$ respect to $x$ is defined as: -
\[
\frac{df(x)}{dx} = \lim_{h \to 0} \frac{f(x + h) - f(x)}{h}
\]

Where \( h \) is the minimum value that variable \( x \) can have. In case of a discrete function, the variable \( x \) has quantized values, so that minimum value in this case can not be zero.

Further, as the quantized function can be efficiently represented by index (1, 2, 3, ..., \( n \)), the minimum value of \( 'x' \) (i.e. \( h \)) can be taken as equal to unity. Thus, putting \( h = 1 \) in the preceding expression we get a transformed first order derivative as:

\[
\frac{df(x)}{dx} = f(x + 1) - f(x)
\]

Clearly, it is merely a difference between the values of \( f(x) \) at \( (x + 1) \) and \( x \). This transformation provides an ease in computation of derivative via a digital interface. Since the operation of \((d/dx)\) on function \( f(x) \) merely yields a difference, the operator can be addressed as difference operator and we will denote it by \( D \), i.e. \( D [f(x)] = f(x + 1) - f(x) \).

### 4.1.5.1 Difference operators for an Image Function

For a two-dimensional image function \( f(x, y) \), the difference operators may be implemented in a number of ways as given below;

(i) \( D [f(x, y)] = -f(x+1, y) + f(x, y) \)

(ii) \( D [f(x, y)] = -f(x + 1, y) + f(x, y) \)

(iii) \( D [f(x, y)] = -f(x - 1, y) - f(x, y) \)

(iv) \( D [f(x, y)] = f(x + 1, y) - f(x - 1, y) \)
(v) \[ D[f(x, y)] = f(x, y + 1) - f(x, y) \]

(vi) \[ D[f(x, y)] = f(x, y - 1) - f(x, y) \]

(vii) \[ D[f(x, y)] = f(x, y - 1) - f(x, y + 1) \]

(viii) \[ D[f(x, y)] = f(x, y + 1) - f(x, y - 1) \]

These operators can be implemented by a 3x3 spatial convolution window. The corresponding kernel masks will have only two weights +1 and -1. The 3x3 kernel masks corresponding to above expressions are show below.

<table>
<thead>
<tr>
<th>0 0 0</th>
<th>0 0 0</th>
<th>0 0 0</th>
<th>0 0 0</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 -1 +1</td>
<td>+1 -1 0</td>
<td>+1 0 -1</td>
<td>-1 0 +1</td>
</tr>
<tr>
<td>0 0 0</td>
<td>0 0 0</td>
<td>0 0 0</td>
<td>0 0 0</td>
</tr>
</tbody>
</table>

Some other important independent forms of the above kernel masks may also be augmented as:

<table>
<thead>
<tr>
<th>+1 0 0</th>
<th>-1 0 0</th>
<th>0 0 0</th>
<th>0 0 0</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 -1 0</td>
<td>0 +1 0</td>
<td>0 -1 0</td>
<td>0 +1 0</td>
</tr>
<tr>
<td>0 0 0</td>
<td>0 0 0</td>
<td>0 0 0</td>
<td>0 0 0</td>
</tr>
</tbody>
</table>

The above four kernels and their remaining four counterparts incorporate diagonal difference. Actually, the above one-dimensional simple kernels are the building blocks of more complex convolution operators which we will work out in further discussion.
4.1.5.1.1 Operation of D operator

The operation \( D \) operator is very simple and straightforward. It operates via a well-established convolution theorem [49]:

\[
f(x, y) * h(x, y) = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f(m,n)h(x - m, y - n)
\]

Where \( f(x, y) \) is the image intensity function, \( h(x, y) \) is the mask-image or kernel in more familiar terminology, and \( * \) is the sign of convolution operator. The kernel mask sits on the image-array and replaces the central pixel by the weighted average of all the pixels in the mask (including central pixel). To have a well-defined centre, the kernel dimension must necessarily be odd. In case of a \( D \) operator, it multiplies the two adjacent pixels by its positive and negative unit weights and sums up the result. In the present study the weight of divisor \( (MN) \) is taken as unity to maintain the intensity (in general, it is taken as the product of kernel dimensions). The result is nothing but the difference of adjacent pixel gray levels. The operation is performed throughout the image and we get a new set of DN values which when displayed yield what we call a difference image. The beauty of the \( D \) operator is that it completely vanishes out the image regions which are digitally uniform, and any variation, how much subtle it may be, is reported. Another advantage of using \( D \) operator is that it is computationally very fast as only a single difference is required to calculate, followed by a division with unity, which is discarded, from routines.

4.1.5.1.2 Implementation of D Operator

What the output of a \( D \) operator, is a network of directional edges with a dark background. We can enhance the directional discontinuities of an image data by
superimposing the $D$ image on the original image. From expression (1) of the section 4.1.5.1, adding $f(x, y)$ on both sides we get

$$D [f(x, y)] + f(x, y) = 2f(x, y) - f(x + 1, y)$$

$$\Rightarrow \quad [D + 1] f(x, y) = 2f(x, y) - f(x + 1, y)$$

$$\Rightarrow \quad \text{The operator} \quad [D + 1] \approx +2, -1$$

The corresponding 3×3 kernel mask will be (in fact it is a 3×1 kernel)

$$[D + 1] \approx \begin{bmatrix} 0 & 0 & 0 \\ -1 & +2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Implementing the kernel $k$ ties we can get a more generalized form of the $[D + 1]$ kernel masks as

$$k.D \ [(f(x, y)] + f(x, y) = (k + 1)f(x, y) - kf(x + 1, y)$$

$$\Rightarrow \quad \text{The operator} \quad [k.D + 1] \approx [k + 1, -k]$$

Thus the generalized kernel mask corresponding to above operator will be

$$[k.D + 1] \approx \begin{bmatrix} 0 & 0 & 0 \\ -k & k + 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

The implementation of the above formulations is performed on standard FCCs (RGB, 432) of ‘Little Colorado River’ and ‘Morro Bay, California’ image data sets. There results of implementation are illustrated in Figs. (4.1.1, 4.1.2, 4.1.3, 4.1.4). All the images were first automatically stretched to highlight the features of the raw data. We can clearly notice the quality enhancement by comparing the results with the original data. Further,
increasing the value of $k$ (2, 3, 4, ...) the information content enhances up to a certain limit (for a particular image data), beyond which garbage output results.

4.1.6 SECOND ORDER DERIVATIVE AND CORRESPONDING 3×3 KERNEL MASKS

Edge detection by second order derivative operator corresponds to the detection of zero-crossing [50]. The most widely used second derivative operator is the Laplacian ($\nabla^2$) operator. In one of its useful variations, ($\nabla^2$) operator is preceded by the noise smoothing operations commonly known as Laplacian of Gaussian (LOG) or Marr-Hildreth (Marr and Hildreth, 1980) operator. A second order derivative is much more aggressive than a first order derivative in highlighting the sharp changes. Thus, we can expect a second order derivative to enhance fine details (including noise) much more than a first order derivative.

4.1.6.1 Laplacian Operator ($\nabla^2$)

For a two-dimensional function $f(x, y)$ the continuous form of Laplacian is given as

$$\nabla^2 = \left[ \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right]$$

The discrete neighborhood operation of this operator on an image function $f(x, y)$ yields

$$\nabla^2 f(x, y) = f(x + 1, y) + f(x - 1, y) + f(x, y + 1) + f(x, y - 1) - 4f(x, y)$$

This implies that kernel weights associated with Laplacian operator (coefficients of various terms in the above expression) are
\[ \nabla^2 \approx +1, +1, -4, +1, +1 \]

And so the corresponding simplest 3x3 kernel mask (commonly known as Laplacian high pass filter) will be as

\[
\begin{bmatrix}
0 & 1 & 0 \\
1 & -4 & 1 \\
0 & 1 & 0 \\
\end{bmatrix}
\quad \text{or} \quad
\begin{bmatrix}
0 & -1 & 0 \\
-1 & 4 & -1 \\
0 & -1 & 0 \\
\end{bmatrix}
\]

We can observe that \( f(x, y) \) is a two-dimensional function, so the variations in the direction of \( xy \) must also be taken into account. Considering this factor, we should incorporate additional terms \( \partial^2/\partial x \partial y \) in the expansion of discrete Laplacian i.e.

\[
\nabla^2 f(x, y) = f(x + 1, y) + f(x + 1, y + 1) + f(x - 1, y - 1) + f(x + 1, y - 1) \\
\quad + f(x - 1, y + 1) + f(x - 1, y) + f(x, y + 1) + f(x, y - 1) - 8f(x, y)
\]

Associated kernel weight

\[ \nabla^2 \approx +1, +1, +1, +1, -8, +1, +1, +1, +1 \]

And corresponding 3x3 kernel mask will be as:

\[
\begin{bmatrix}
-1 & -1 & -1 \\
-1 & +8 & -1 \\
-1 & -1 & -1 \\
\end{bmatrix}
\quad \text{or} \quad
\begin{bmatrix}
+1 & +1 & +1 \\
+1 & -8 & +1 \\
+1 & +1 & +1 \\
\end{bmatrix}
\]

**4.1.6.1.1 Operation of \( \nabla^2 \) Operator**

We can readily notice from the geometry of the \( \nabla^2 \) operator that it is nothing but a little complicated version of previously established \( D \) operator, where the difference is taken in all directions (including the diagonal elements). Of course, the diagonal elements should be less than unity (in fact \( 1/\sqrt{2} \) because they fall at relatively larger distance from the central pixel, but for the sake of computational efficiency, they are incorporated as
being unity [51]. The operation of $\nabla^2$ operator is just akin to that of previously described $D$ operator, with an additional advantage of all-direction edge detection. As a result of operation we get a new image (of the size of original) called as Laplacian-image representing all types of discontinuities in the form of a bright network of linear features in a dark background.

4.1.6.1.2 Implementation of $\nabla^2$ Operator

The output of $\nabla^2$ operator is a network of bright edges in a dark background. This network can be superimposed (added or substracted) on the original image to get edge-enhanced version, which is far most superior in quality in comparison to the original image (with a disadvantage that edges are one pixel displaced from their original position).

As we have

$$\nabla^2 f(x, y) = f(x + 1, y) + f(x - 1, y) + f(x, y + 1) + f(x, y - 1) - 4f(x, y)$$

Subtracting the Laplacian from original image we get

$$f(x, y) - \nabla^2 f(x, y) = 5f(x, y) - f(x + 1, y) - f(x - 1, y) - f(x, y + 1) - f(x, y - 1)$$

The operator

$$[1 - \nabla^2] \approx -1, -1, +5, -1, -1.$$ 

And corresponding $3 \times 3$ spatial convolution mask will be

$$[1 - \nabla^2] \approx \begin{pmatrix}
0 & -1 & 0 \\
-1 & +5 & -1 \\
0 & -1 & 0
\end{pmatrix}$$
The above operator can be envisaged in a more general form by subtracting the Laplacian \( k \) times from the original image. In this situation;

\[
[1 - k \nabla^2] f(x, y) = (1 + 4k) f(x, y) - k f(x + 1, y) - k f(x - 1, y) - k f(x, y + 1) - k f(x, y - 1)
\]

The operator

\[
[1 - k \nabla^2] \approx -k, -k, (1 + 4k), -k, -k
\]

And the corresponding generalized 3×3 Laplacian kernel mask will be.

\[
\begin{bmatrix}
0 & -k & 0 \\
-k & 1+4k & -k \\
0 & -k & 0
\end{bmatrix}
\]

To augment, considering the influence of diagonal elements, we get another form of generalized Laplacian operator as:

\[
\begin{bmatrix}
-k & -k & -k \\
-k & 1+8k & -k \\
-k & -k & -k
\end{bmatrix}
\]

In the application of above generalized cases \( k \geq 1 \), the divisor weight must be properly chosen to maintain the intensity.

The implementation of the above formulation is carried out on standard false color composites of ‘Tanzanian Coast’ [RGB-754] and ‘Paris’ [RGB-432] image data sets. The result of implementation of the discussed kernel masks are illustrated in the Figs. (4.1.5, 4.1.6). All the images were automatically stretched prior to processing to highlight the features in the raw image data.
4.1.6.2 Total Differential Operator $T$

As an augment to the study of image derivative operators, we have established and tested the performance of the Total differential operator ($T$), and, fortunately, it paid well.

We know from the theory of elementary calculus that total differential of a function $f(x, y)$ in continuous formulation is given as:

$$d[f(x,y)] = \frac{\partial f(x,y)}{\partial x} dx + \frac{\partial f(x,y)}{\partial y} dy$$

For a discrete function, we can have

$$\frac{\partial f(x,y)}{\partial x} dx = \frac{f(x+1,y) - f(x,y)}{\delta x} dx$$

$$\frac{\partial f(x,y)}{\partial y} dy = \frac{f(x,y+1) - f(x,y)}{\delta y} dy$$

In the case of a discrete and quantized function, the minimum difference can be taken as equal to unity.

$$\delta x = dx \text{ and } \delta y = dy$$

In this way, can rewrite above expression as

$$d[f(x,y)] = f(x+1,y) - 2f(x,y) + f(x,y+1)$$

As the operator is 'total differential' we will denote it by symbol $T$

$$T[f(x,y)] = f(x+1,y) - 2f(x,y) + f(x,y+1)$$
Operator

\[
T \approx +1, -2, +1
\]

\[
\begin{array}{ccc}
0 & +1 & 0 \\
+1 & -2 & 0 \\
0 & 0 & 0 \\
\end{array}
\]

T \equiv \begin{array}{ccc}
0 & +1 & 0 \\
+1 & -2 & 0 \\
0 & 0 & 0 \\
\end{array}

Other canonical forms may also be incorporated which are equally valuable, as

\[
\begin{array}{ccc}
0 & 0 & 0 \\
+1 & -2 & 0 \\
0 & +1 & 0 \\
\end{array}
\]

\[
\begin{array}{ccc}
0 & 0 & 0 \\
0 & -2 & +1 \\
0 & +1 & 0 \\
\end{array}
\]

\[
\begin{array}{ccc}
0 & +1 & 0 \\
0 & -2 & +1 \\
0 & 0 & 0 \\
\end{array}
\]

4.1.6.2.1 Operation of T operator

The mechanism of operation is just akin to that described for previous derivative operators. It replaces the central pixel by the weighted average of two orthogonal pixels, as a consequence of which we get orthogonal set of edges. In fact, the edges which are orthogonal in the reference frame of the kernel are much more highlighted. The performance is superior in the case of images having a number of orthogonal edges (like images of crop lands and aerial images of planned cities.)

4.1.6.2.2 Implementation of T Operator

The \(T\) operator can be implemented via same mechanism as that for \(D\) and \(V^2\) operators.

i.e. \(T [f(x, y)] - f(x, y) = f(x + 1, y) - 3f(x, y) + f(x, y+1).\)

\([T - 1] f(x, y) = 1.f(x + 1, y) - 3.f(x, y) +1.f(x, y +1)\)

Operator

\([T-1] \approx +1, -3, +1\)
And so, the corresponding 3×3 kernel mask will be as

$$[T - 1] \approx \begin{bmatrix} 0 & 0 & 0 \\ +1 & -3 & 0 \\ 0 & +1 & 0 \end{bmatrix}$$

Other canonical forms of $[T-1]$ operator are equally valuable, which are

$$\begin{bmatrix} 0 & +1 & 0 \\ +1 & -3 & 0 \\ 0 & +1 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & -3 & +1 \\ 0 & +1 & 0 \end{bmatrix}$$

More generalized form may be achieved as per the reference of previous operators by implementing it $k$ times as

$$[k.T - 1] f(x, y) = +k.f(x +1, y) - (1 + 2k)f(x, y) + k.f(x, y +1)$$

The operator

$$[k.T - 1] \approx +k, -(1+2k), +k$$

And the corresponding 3×3 Kernel mask is

$$[k.T - 1] \approx \begin{bmatrix} 0 & +k & 0 \\ +k & -(1+2k) & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Other canonical forms of $[k.T - 1]$ may also be obtained as

$$\begin{bmatrix} 0 & 0 & 0 \\ +k & -(1+2k) & 0 \\ 0 & +k & 0 \end{bmatrix}$$

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & -(1+2k) & +k \\ 0 & +k & 0 \end{bmatrix}$$

$$\begin{bmatrix} 0 & +k & 0 \\ 0 & -(1+2k) & +k \\ 0 & 0 & 0 \end{bmatrix}$$

We have implemented the above kernels for various values of $k$ weights on Littleport, Cambridgeshire image data. The results of implementation of the above worked out kernel masks for varying values of $k$ weights are illustrated in the Fig. (4.1.7). The main characteristic, to worth notice, of a total differential operator ($T$) is that from $k$
= 2 onwards it give, like a 3D impression of a 2D image and so the features seem to be much more delineated.

4.1.7 EPILOGUE

As to sum up, we have initiated the discussion with the concept of digital image understanding and tried to address the complex aspects in an interesting and vantageous way. The study is carried out using Mather's Image Processing System (MIPS) with due acknowledgement of Paul. M. Mather.
For $k \cdot D$ operator

(a) Little Colorado River image for $k = 1$
(b) Little Colorado River image for $k = 2$
(c) Original image [RGB-432]

\[
\begin{array}{ccc}
0 & 0 & 0 \\
-k & k & 0 \\
0 & 0 & 0 \\
\end{array}
\]

Fig. 4.1.1
For $kD$ operator

(a) Morrobay image for $k = 1$
(b) Morrobay image for $k = 2$
(c) Original image [RGB-432]

Fig. 4.1.2
For $[k, D+1]$ operator

(a) Little Colorado River image for $k = 1$
(b) Little Colorado River image for $k = 2$
(c) Original image [RGB-432]

Fig. 4.1.3
For $|k.D+1|$ operator

(a) Morrobay image for $k = 1$
(b) Morrobay image for $k = 2$
(c) Original image [RGB-432]

Fig. 4.1.4
For $[1 - k \nabla^2]$ operator

(a) Tanzanian Coast image for $k = 1$
(b) Tanzanian Coast image for $k = 2$
(c) Original image [RGB-754]

Fig. 4.1.5
For $[1 - k \nabla^2]$ operator

(a) Paris image for $k = 1$
(b) Paris image for $k = 2$
(c) Original image [RGB-432]

Fig. 4.1.6
For \([kT-1]\) operator

(a) Littleport image for \(k = 2\)
(b) Littleport image for \(k = 3\)
(c) Original image [RGB-432]

Fig. 4.1.7
SECTION 2
Evaluation of image enhancing techniques for remotely sensed vegetation image interpretation
4.2.1 INTRODUCTION

For visual interpretation of remotely sensed data; histogram equalization (HE), principal component analysis (PCA), and intensity hue saturation (IHS) are commonly utilized. Incorporating another simple method, namely, manual histogram stretching (MHS), as to augment, we propose a comparative quest for these recipes in special case of vegetation data discrimination. It was found that the proposed method (MHS), however, slightly tedious, serves much better in the discrimination of minutes forest classes. In the construction of FCCs for the Landsat (TM) data under study, for MHS, the DN range for band -4, -3, and -2 were selected such as to encompass maximum vegetation content, for these bands for the given data. In the realm of MHS the FCCs so produced appear in quite different themes, but minute forest-class discrimination is promising.

Delineation and identification of minor and dispersed forest classes is of particular interest in forest and vegetation RS. In the present study, a comparative approach of some well-known enhancing tools is centered from a variety of widely available techniques. From a host of enhancing techniques, we have concentrated attention towards four given below: -

(i) Principal component Analysis (PCA).
(ii) Intensity Hue Saturation (IHS).
(iii) Histogram Equation (HE).
(iv) Manual Histogram stretching (MHS).
Out of which the MHS is the *proposed method*, which is evaluated and compared with remaining three methods. The evaluation is based on the visual appearance of the output images so a human observer, here, is the ultimate judge.

4.2.2 PCA

PC analysis among band-4, -3, and -2 of Mississippi river image data (provided by NASA and accompanied with the MIPS package, with due acknowledgement of Paul M. Mather) is performed. This was followed by making a FCC by these three principal components through assigning RGB colour scheme. The final results were obtained on the basis of covariance and correlation matrix, of which the former was found to be much suitable.

The multispectral or multidimensional nature of RS image data can be accommodated by constructing a vector space with as many axes or dimensions as there are spectral components associated with each pixel [52]. In the case of Landsat (TM) data it will have seven dimensions while for SPOT-HRV data it will be three-dimensional. For hyperspectral data there may be several hundred axes. A particular pixel in an image is plotted as a point in such a space with coordinates that correspond to the brightness values of the pixels in the appropriate spectral components [53]. The position of pixel points in multispectral space can be described by vectors, whose components are individual spectral responses in each band. Strictly, these are the vectors drawn from the origin do the pixel point. For a multispectral space with a large number of pixels, with each pixel described by its appropriate vector $x$, the mean position of the pixels in the space is defined by the expected value of pixel vector $x$ as $m = \frac{1}{k} \sum_{x=1}^{k} x$, where
$x_i$ are the individual pixel vectors of total number $k$; $\delta$ is the expectation operator [54]. While the mean vector is useful to define the average or expected position of the pixels in the multispectral vector space, it is of value to have available a means by which their scatter or spread is described. This is served by covariance matrix, which is defined as
$$\Sigma_x = \delta[(x - m)(x - m)^\prime].$$
An unbiased estimate of covariance matrix is given as
$$\Sigma_x = \frac{1}{k-1} \sum_{k=1}^{k} (x_i - m)(x_i - m)^\prime.$$ The covariance matrix is one of the most important mathematical concepts in the analysis of multispectral RS data [55]. If there is a correlation between the responses in a pair of spectral bands the corresponding off-diagonal elements in the covariance matrix will be large by comparison to the diagonal terms. On the other hand if there is a little correlation, the off-diagonal terms will be close to zero [56]. This behaviour can also be described in terms of the correlation matrix whose elements are related to those of the covariance matrix by $Q_{ij} = v_{ii} / \sqrt{v_{ii}v_{jj}}$ where $Q_{ij}$ is an element of the correlation matrix, $v_{ii}$ are the elements of the covariance matrix, $v_{ii}$ and $v_{jj}$ are the variances of the $i$" and $j$" bands of data. The $Q_{ij}$ describes the correlation between band $i$ and band $j$. Using PCA in practice the user is not involved in this level of detail rather only three steps are necessary, presuming software exists for implementing each of those steps., (i) Assembling the covariance matrix of the image to be transformed (ii) Determination of the eigenvalues and eigenvectors of the covariance matrix (iii) Finally to form the components using the eigenvectors of the covariance matrix as the weighting coefficients [57].

In constructing the colour display of remotely sensed data only three dimensions of information can be mapped to the three colour primaries of the display system. In the case of Landsat (MSS) data, usually bands -4, -5 and -7 are chosen for this. band-6. A less
ad hoc means for colour assignment rest upon performing a PC transform and assigning the first three components to the red, green and blue colour primaries.

### 4.2.2.1 Interpretation of PCA image

The original image data and its PC transform are shown in Fig. (4.2.1) and Fig. (4.2.2). In the PC transform of original image we found that:

(i) The dense forest areas in the original image with poor leaf health, appear in a slightly varying dark-red tones in which the minute vegetation classes are not available to notice. Some rare patches of healthy vegetation and cultivated fields are of course visible in scattered form.

(ii) Healthy vegetation, on the other hand, shows some remarkable tonal variations, but in places where these variations strongly exist. Small vegetation classes are still very hard to detect.

(iii) Sandy area in the middle of the river exhibits some yellowish tinge, thereby, indicating the existence of any type of sparse vegetation (like rare grass).

### 4.2.3 IHS

Intensity, Hue, and Saturation are image characteristics, which are most commonly linked to human perceptions [58]. In other words, these are the parameters in terms of which a common man identifies and describes an image. Hue is the colour attribute that describes the purity of a colour, Saturation gives a measure of the degree to which a pure colour is diluted with white light, and Brightness is a subjective descriptor that is practically colour sensation. We do know that the intensity is the most useful descriptor that is practically impossible to measure [59]. It embodies the achromatic notion of
intensity and is one of the key factors in describing colour sensation. We do know that the intensity is the most useful descriptor of monochromatic images. That is the quality, which is most easily measurable and interpretable.

On the basis of RGB colour Model, the RGB → IHS and converse transform can be described and implemented in a computer system as

\[
H = \begin{cases} 
\theta & \text{if } B \leq G \\
360 - \theta & \text{if } B > G \end{cases} \text{ with } \theta = \cos^{-1} \left( \frac{1}{2} \left[ (R - G) + (R - B) \right] \left[ (R - G)^2 + (R - B)(G - B) \right] \right),
\]

\[
S = [1 - 3 \times \text{Min}(R, G, B)/(R + G + B)], \text{ and } (R + G + B)/3 \text{, and conversely } B = I[(1 - S)],
\]

\[
R = I \times [1 + S \times \cos H/\cos (60 - H)], \; G = [1 - (R + B)] [60]. \text{ These are the formulations in terms of which the concept is incorporated in MIPS.}
\]

The intermediate stage between two transforms is the stage in which desired and suitable operations may be performed. In the present study, the simplest operation, which was performed, is that the 'intensity' and 'saturation' were linearly stretched.

The IHS is useful in two ways: first, as a method of image enhancement and secondly, as a means of combining coregistered images from different sources. In the former case, the RGB representation is first converted to IHS, as described by Foley et al (1990), Hearn and Baker (1994), and Shih (1995). Next the Intensity (I) and saturation (S) components are stretched independently and the IHS representation is converted back to
RGB for display purpose. It does not make sense to stretch the hue ($H$) component, as this would upset the colour balance of the image (Gillespie et al., 1986, provide this in detail).

IHS Transformation has been found to be particularly useful in geological applications (Jutz and Chorowicz, 1993, and Nalbant and Alptekin, 1995). Further details of IHS transform are given in Blom and Daily (1982), Foley et al. (1990), Green (1983), Hearn and Baker (1994), Phol and Van Genderen (1998), and Mulder (1980). Terhalle and Bodechtel (1986) illustrate the use of IHS transformation in the mapping of arid geomorphic features, while Gillespie et al. (1986) discuss the role of IHS in the enhancement of highly correlated images. Massonat (1993) gives details of an interesting use of IHS in which the amplitude, coherence and phase components of an interferometric image are allocated to hue, saturation, and intensity respectively which highlights the details of coherent and incoherent patterns.

4.2.3.1 INTERPRETATION OF IHS IMAGE

The original image data after IHS transformation (in fact, back to RGB) along with histograms and cross-section is illustrated in Fig. (4.2.3). We can extract following informations.

1. The healthy vegetation areas on the left of the river and also on the right but in scattered form, appear in bright red tone with a poor textural variation.

2. Sparse, unhealthy dense forest areas on the right of the river appear in an average dark-gray tone with minute red tinges at some places. But tonal and textural variations are still satisfactory.
(3) The sandy areas in the middle of the river, here, also appear white without any red tinct, denying, therefore, the presence of any kind of sparse vegetation or grass.

(4) The IHS image incorporates a remarkable contrast between healthy and stressed vegetation, the former appearing in bright red while the later in a dark-gray tone, thus, HIS transform is a good indicator of biomass vigor.

4.2.4 HE

The histogram of a digital image with gray-levels in the range \([0, L-1]\) is a discrete function \(h(r_k) = n_k\), where \(r_k\) is the \(k^{th}\) gray level and \(n_k\) is the number of pixels in the image having gray-level \(r_k\). It is common practice to normalize the histogram by dividing each of its values by the total number of pixels in the image. Thus a normalized histogram is given by \(P(r_k) = n_k/n\), for \(k = 0, 1, \ldots, L-1\). Loosely speaking, \(P(r_k)\) gives an estimate of the probability of occurrence of gray-level \(r_k\) [61]. The shape of gray-level histogram gives an idea about the overall appearance of an image. For example, an image with a positively-skewed gray-level histogram looks brighter than an image with negatively-skewed gray level histogram. An image may be enhanced by mapping the input gray-levels in such a way that the gray-level histogram of the processed image attains a desired shape. HE is the simplest kind of these techniques. Theoretically, an equalized histogram has the shape of the uniform histogram, in which, in principle, each bar has the same height. Such a histogram has associated with it an image that utilizes the available brightness levels equally and thus should give a display in which there is good representation of detail at all brightness values. In practice a perfectly uniform histogram
can not be achieved for digital image data, a quasi-uniform on the average can, however be realized [62]. The development of the method is as follows.

Consider for a moment a continuous function, and let the variable \( r \) represents the gray-levels of the image to be enhanced. We assume that \( r \) has been normalized to the interval \([0, 1]\), with \( r = 0 \) representing black and \( r = 1 \) representing white. Later, we consider a discrete formulation and allow pixel values to be in the interval \([0, L - 1]\). For any \( r \) satisfying the aforesaid conditions, we focus attention on transformation of the form \( s = T (r) \) for \( 0 \leq r \leq 1 \), that produces a level \( s \) for every pixel value \( r \) in the original image. The transformation function is such that; (i) \( T (r) \) is single valued and monotonically increasing in the interval \( 0 \leq r \leq 1 \) and, (ii) \( 0 < T (r) < 1 \) for \( 0 < r < 1 \). The requirement in (i) that \( T (r) \) be single valued is needed to guarantee that the inverse transformation will exist, and the monotonicity condition preserves the increase order from black to white in the output image. Condition (ii) guarantees that the output gray-levels will be in the same range as the input levels. The gray-levels in an image may be viewed as random variables in the interval \([0, 1]\). One of the most important and fundamental descriptor of a random variable is its probability density function (PDF). If \( P_r (r) \) and \( P_s (s) \) denote the PDFs of random variables \( r \) and \( s \), then from the elementary probability theory \( P_s (s) = P_r (r) \frac{dr}{ds} \). Thus the PDF of transformed variable \( s \) is determined by the gray-level PDF of input image and by chosen transformation function [63]. A transformation function of particular importance in DIP has the form \( s = T (r) = \int p(\omega) d\omega \), where \( \omega \) is a dummy variable of integration. The right side of this equation is recognized as the cumulative distribution function of random variable \( r \). Following the Leibnitz's rule that the derivative of a definite integral with respect to its
upper limit is simply the integrand evaluated at that limit i.e. \( ds/dr = P_r (r) \Rightarrow P_x (s) = 1 \) for \( 0 \leq s \leq 1 \).

Because \( P_x (s) \) is a PDF, it follows it must be zero outside the interval \([0, 1]\) in this case because its integral overall values of \( s \) must equal to 1. This form of \( P_x (s) \) as given above is recognized as a uniform PDF. For discrete values, probabilities and summations are considered instead of PDF and integral, and so the discrete version of transformation function is given as

\[
s_k = \sum_{i=0}^{k} P_r (r_i) = \frac{n_i}{n} \text{ for } k = 0, 1, \ldots, L - 1.
\]

Thus a processed image is obtained by mapping each pixel with gray-level \( r_k \) in the input image into a corresponding pixel with level \( s_k \) in the output image. This transformation is what we call histogram equalization or histogram linearization [64]. The method developed here has the additional advantage that it is fully ‘automatic’, that is, given an image, the process of HE consists simply the implementation of the above equation, which is based on information that can be extracted directly from the given image, without the need for further parameter specification.

### 4.2.4.1 Interpretation of HE image

The original image after histogram equalization, along with histogram and cross section is shown in Fig. (4.2.4). We can readily perceive the following informations:

(i) The overall brightness of the original image is very much improved after HE as expected due to approximately uniform distribution of gray-levels for all classes.
(ii) The tonal variations in the dense forest area on the right of the river are clearly visible, thereby, improving the degree of discrimination. These variations are due to variety of tree species or vegetation health.

(iii) The healthy forest areas on the left of the river appear in slightly pink colour, but variations are observable in terms of texture.

(iv) The small sandy portions in the middle of the river appear milky-white, in contrast to the original image where they are slightly pinkish, that is, the information of sparse vegetation in this area is lost.

(v) Tonal variations in the river water are much prominent in comparison to original image, which is due to turbidity or shallowness the river water.

**4.2.5 MANUAL HISTOGRAM STRETCHING (MHS)**

The MHS (Proposed method) provides more control over gray-level stretching for various bands of a given multispectral image data set. It is basically a hit and trial process in which each band is stretched on the basis of a predefined observation. The gray level ranges of the data for various bands and the requirement of interpreter are the two key factor which support a good prior setting of limits (gray-level ranges for RGB in the present work) from non-trivially large number of combinations (\(^nP_3\), \(n = \) number of spectral bands). The present study is carried out on image data sets of Mississippi river (a seven band Landsat –TM image data), Tanzanian coast (a three band Landsat – MSS image data), and Rio de Janeiro (a seven band Landsat-TM image data). The spectral bands for each image data that were considered are NIR, Red and Green. For Landsat–TM, these bands are designated as band-4, band-3 and
band-2, while for Landsat-MSS, these are designated as band-7, band-5 and band-4 respectively. The main results for all of the three image data are in the form of standard FCCs, remaining other combinations (permutation FCCs) are also constructed and illustrated, as they are also extremely valuable.

4.2.5.1 Mississippi river image data

The original histogram extremum for the Mississippi river image data for NIR, Red and Green bands are as follows:

<table>
<thead>
<tr>
<th></th>
<th>NIR</th>
<th>Red</th>
<th>Green</th>
</tr>
</thead>
<tbody>
<tr>
<td>Max</td>
<td>79</td>
<td>80</td>
<td>169</td>
</tr>
<tr>
<td>Min</td>
<td>4</td>
<td>13</td>
<td>4</td>
</tr>
</tbody>
</table>

As we are basically interested in the vegetation content, it is logically justice to extract the vegetation information associated with each of the three bands. This is done by investigating the DN values of vegetation patches in the image for each band. For each band, the vegetation ranges between some upper and lower limits. These limiting values are decided manually by analyzing the brightest and darkest vegetation pixels (and so is the name MHS). Though the method is somewhat painstaking, yet it offers much control to the human observer. For Mississippi river image data, the vegetation limiting values for NIR, Red and Green band are worked out as:

<table>
<thead>
<tr>
<th></th>
<th>NIR</th>
<th>Red</th>
<th>Green</th>
</tr>
</thead>
<tbody>
<tr>
<td>Max</td>
<td>48</td>
<td>24</td>
<td>21</td>
</tr>
<tr>
<td>Min</td>
<td>17</td>
<td>19</td>
<td>17</td>
</tr>
</tbody>
</table>

For NIR band, as expected, the range is maximum because of the fact that in NIR band vegetation has highest reflectance. RGB-432 [17-48, 19-24, 17-21] is shown in
Fig. (4.2.5) as a standard MHS result to compare it with other three methods. Remaining other permutation FCCs are as follows:

<table>
<thead>
<tr>
<th>RGB Combination</th>
<th>Bands 1 Band 2 Band 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>RGB - 423</td>
<td>17-48, 17-21, 19-24</td>
</tr>
<tr>
<td>RGB - 342</td>
<td>19-24, 17-48, 17-21</td>
</tr>
<tr>
<td>RGB - 324</td>
<td>19-24, 17-21, 17-48</td>
</tr>
<tr>
<td>RGB - 243</td>
<td>17-21, 17-48, 19-24</td>
</tr>
<tr>
<td>RGB - 234</td>
<td>17-21, 19-24, 17-48</td>
</tr>
</tbody>
</table>

These permutation FCCs are shown in Fig. (4.2.7). These combinations represent vegetation in a variety of themes and provide enormous case for interpretation of vegetation classes.

The RGB combinations worked out above discriminate various features just like, as they were classified on a thematic map. In this output false colour environment, the spatial features, of course, lose their standard false colour definition (i.e. the output is spectrally blind) and they appear in surprisingly different themes. For example, the sandy soils which appear bright white on standard FCC, appear on RGB-432 [17-48, 19-24, 17-21] in blood red tone, and hence are easily perceptible and delineable from their surrounding environs.

We can see a remarkable discrimination in the forest area on the right side of the river. The different tree species appear in different tones, so we can easily make a rough idea about the number of plant species present in the study area. Other combinations are equally useful and it is up to the desire and convenience of an interpreter, which combination he/she prefers.
4.2.5.2 Tanzanian Coast image data

The original histogram extremum for Tanzanian Coast data for NIR (MSS band-7), Red (MSS band-5), and Green (MSS band-4) are found as follows:

<table>
<thead>
<tr>
<th></th>
<th>NIR</th>
<th>Red</th>
<th>Green</th>
</tr>
</thead>
<tbody>
<tr>
<td>Max</td>
<td>60</td>
<td>142</td>
<td>54</td>
</tr>
<tr>
<td>Min</td>
<td>0</td>
<td>0</td>
<td>12</td>
</tr>
</tbody>
</table>

After manually examining the vegetation pixels, the limiting optimum values for vegetation for each band are worked out as:

<table>
<thead>
<tr>
<th></th>
<th>NIR</th>
<th>Red</th>
<th>Green</th>
</tr>
</thead>
<tbody>
<tr>
<td>Max</td>
<td>52</td>
<td>17</td>
<td>18</td>
</tr>
<tr>
<td>Min</td>
<td>29</td>
<td>14</td>
<td>14</td>
</tr>
</tbody>
</table>

The six permutation FCCs which can be constructed are as follows:

- RGB – 754: [29–52, 14–17, 14–18]
- RGB – 745: [29–52, 14–18, 14–17]
- RGB – 574: [14–17, 29–52, 14–18]
- RGB – 547: [14–17, 14–18, 29–52]
- RGB – 475: [14–18, 29–52, 14–17]
- RGB – 457: [14–18, 14–17, 29–52]

Original Tanzanian coast image is shown in Fig. (4.2.8). Out of the above six, the standard FCC (RGB–754) is taken into consideration. It is illustrated in Fig. (4.2.9) along with histogram and cross-section. Remaining five permutation FCCs are also equally valuable and they are separately shown in Fig. (4.2.10).

As it is evident from figures, apart from discrimination of vegetation classes, this method remarkably delineated water qualities. The turbidity, pollutants, silts and shallowness of water is highlighted in form of different tones on each of the permutation
FCC, thereby indicating the usefulness of the method for water quality mapping applications.

4.2.5.3 Rio de Janeiro image data

The original histogram extremum for Rio de Janeiro image data for NIR, Red and Green bands are found to be as:

<table>
<thead>
<tr>
<th></th>
<th>NIR</th>
<th>Red</th>
<th>Green</th>
</tr>
</thead>
<tbody>
<tr>
<td>Max</td>
<td>123</td>
<td>254</td>
<td>92</td>
</tr>
<tr>
<td>Min</td>
<td>2</td>
<td>7</td>
<td>6</td>
</tr>
</tbody>
</table>

After manually examining the vegetation pixels, the optimum limiting values of vegetation for each band are worked out as follows:

<table>
<thead>
<tr>
<th></th>
<th>NIR</th>
<th>Red</th>
<th>Green</th>
</tr>
</thead>
<tbody>
<tr>
<td>Max</td>
<td>125</td>
<td>30</td>
<td>25</td>
</tr>
<tr>
<td>Min</td>
<td>35</td>
<td>12</td>
<td>12</td>
</tr>
</tbody>
</table>

As mentioned previously, the six permutation FCCs that can be envisioned will be:

- RGB – 432 [35–125, 12–30, 12–25]
- RGB – 423 [35–125, 12–25, 12–30]
- RGB – 342 [12–30, 35–125, 12–25]
- RGB – 324 [12–30, 12–25, 35–125]
- RGB – 243 [12–25, 35–125, 12–30]
- RGB – 234 [12–25, 12–30, 35–125]

Original Rio de Janeiro image data is shown in Fig. (4.2.11). Again, out of the above six, the standard FCC (RGB – 432) is considered, and is illustrated in Fig. (4.2.12) along with histogram and cross-section. Remaining five permutation FCCs are separately shown in Fig (4.2.13). From RGB – 432 [35-125, 12-30, 12-25] composite of Rio de Janeiro image we can draw following information:
(1) The sparse vegetation, scattered inside the settlement areas is very well discriminated on this MHS thematic image, occurring in blood-red tone, while the settlement in cyan colour.

(2) The densely vegetated rocky area in the lower portion are separated from scrubbed rocks. The scrubbed rocks with much sandy outcrops appear in blood-red tones, while the densely vegetated rocky regions appears in a pinkish colour with much tonal and textual variations.

(3) The changes in water colour near the port are also of worth notice (as in the case of Tanzanian Coast image), which is due to turbidity, siltation, pollution, and shallowness of coastal water.

4.2.6 HISTOGRAMS, CROSS-SECTIONS, AND ZOOMS

For the sake of simple understanding of the mechanisms and qualities of all the four methods under study, histograms and cross-sections are plotted, and 4x zooms of certain areas of interest are provided.

4.2.6.1 Histograms

Histograms of original image and of the output images after each of the four operations are shown along with images. They depict important inferences about the distribution of pixels in various classes of gray-levels.

4.2.6.2 Cross–Sections

Cross-section is a plot between the DN value and pixel number along a line between two-fixed points on the image. For Mississippi River image we have selected two clearly visible points A, and B in the forest area, on the right hand side of the river. The cross-
section plots show the variations of pixels DN along the straight line. A comparative analysis of cross-section infers the quality of discrimination of various enhancing operations. For Tanzanian Coast and Rio de Janeiro image data the cross-sections are taken between top left and bottom right points to cover the entire image.

4.2.6.3 Zooms

The 4x zooms for original and other four outputs for Mississippi river image are given; a window as top left and bottom right corners as point A and B. Zooms facilitate for good visual perception and also infer the potential of discrimination of various methods discussed (Fig. 4.2.6).
Fig. 4.2.1

(a) Original Mississippi image (RGB-432)
(b) Histogram
(c) Cross-section between points A and B
(a) Mississippi image after IHS transform
(b) Histogram
(c) Cross-section between points A and B

Fig.4.2.2
(a) Mississippi image after PCA transform
(b) Histogram
(c) Cross-section between points A and B

Fig. 4.2.3
Histogram/LUT Plots

(a) Mississippi image after HE

(b) Histogram

(c) Cross-section between points A and B

Fig. 4.2.4
Histogram/LUT Plots

(a) Mississippi image after MHS (RGB-432)
   a   (b) Histogram
   b  c   (c) Cross-section between points A and B

Fig.4.2.5
Fig. 4.2.6

(a) after PCA
(b) after IHS
(c) after HE
(d) MHS
(e) For original Points A & B as top left and bottom right corners.
Permutation FCCs of Mississippi image for MHS

(a) RGB-423  (b) RGB-342  (c) RGB-324  (d) RGB-243  (e) RGB-234  (f) Original

Fig. 4.2.7
(a) Original Tanzanian coast image (RGB-754)

(b) Histogram

(c) Cross-section between top left and bottom right corners

Fig. 4.2.8
(a) Tanzanian coast image after MHS (RGB-754)

(b) Histogram

(c) Cross-section between top left and bottom right corners

Fig.4.2.9
Permutation FCCs of Tanzanian Coast image for MHS

(a) RGB-745  (b) RGB-574  (c) RGB-547  (d) RGB-475  (e) RGB-457  (f) Original

Fig. 4.2.10
(a) Original Rio de Janeiro image (RGB-432)
(b) Histogram
(c) Cross-section between top left and bottom right corners

Fig. 4.2.11
Histogram/LUT Plots

(a) Rio de Janeiro image after MHS (RGB-432)
(b) Histogram
(c) Cross-section between top left and bottom right corners

Fig.4.2.12
Permutation FCCs of Rio de Janeiro image for MHS

(a) RCB-423  (b) RGB-342  (c) RCB-324  (d) RGB- 243  (e) RGB-234  (f) Original

Fig. 4.2.13
SECTION 3
Thresholding techniques for segmentation of remotely sensed vegetation data
4.3.1 INTRODUCTION

We propose a simple method to segment and delineate the vegetation content in a given Remotely Sensed image data utilizing the rudimentary concepts of multiple thresholding and filtering, using Mather’s Image Processing System (MIPS). MIPS package is accompanied with the book of Paul M. Mather, entitled, “Computer Processing of Remotely Sensed Images”. The procedure is worked out to minimize the problem of broken boundaries. A manual threshold selection approach is used which seems somehow tedious for novice practitioners, but after a little exercise, much satisfactory results, to serve as a layer of GIS input may be procured.

4.3.2 DATA UNDER STUDY

The digital image data under the study is of an area of Rio de Janeiro, Brazil.

4.3.3 THRESHOLDING

Thresholding is an image segmentation technique, which suppresses (or highlights) the DN values of an image beyond a certain limit. That limiting value is known as threshold. In general, the threshold can be chosen as the relation, \( t = t [r, c, p(r, c)] \), where \( p(r, c) \) is the feature value at pixel \((r, c)\). In the case of gray-level thresholding, \( p(r, c) = g(r, c) \) for all \((r, c)\) within the image domain. If \( t \) depends on the feature \( p(r, c) \) only, it is called local threshold, if \( t \) depends on the pixel position as well, it is called dynamic threshold, otherwise it is global or position independent threshold [65]. It should be noted that feature thresholding is the simplest method of image segmentation; usually it precedes the selection of appropriate feature to obtain a useful result. Because of its intuitive properties and simplicity in implementations, digital image thresholding
enjoys a central position in the applications of image segmentation [66]. Secondly, the selection of threshold is also a non-trivial task. Any inappropriate threshold would incur significant and non-acceptable error of classification. Thus the use of thresholding techniques for image segmentation needs to solve the two problems: (1) Choice of feature properties to achieve desired segmentation and, (2) Selection of optimum threshold that would incur least classification error [67].

In the present study, the gray-level thresholding (simplest thresholding) is focused on image segmentation \{i.e., \( b(r, c) = 1 \) for \( p(r, c) \leq t \), and \( = 0 \) for \( p(r, c) > t \) or vice versa\} and desired feature-boundary (vegetation) delineation. In the case of vegetation boundary delineation, NIR gray band images are used because in this band vegetation is having highest reflectance.

4.3.4 THRESHOLD SELECTION

One should require sitting with patience when going to select an optimum threshold. We have tried to use such an approach in our study. The study is performed on the Rio-image data of NIR band. The NIR band was chosen for the prime purpose of vegetation delineation and vegetation boundary detection.

The primary step of the presented procedure is the interpretation of image data for vegetation to confirm whether a given patch on the image belongs to vegetation. Once this is confirmed with confidence, the investigation of DN values starts. This is the point where the real test of nerves is evaluated. We have adopted a very careful and patient methodology to analyze the vegetation on DN value basis. The DN values lie between a lower and an upper limit. What we have to do is only to find out these bounding limits very accurately and carefully because these limit confirm the boundaries of vegetation
patches. If the DN values of vegetation lie between limits $T_1$ and $T_2$, then we cut the histogram between $T_1$ and $T_2$ and get a histogram ranging between $T_1$ and $T_2$. The corresponding image approximately contains only vegetation patches in a dark background. For Rio-image data under study, the DN values for vegetation range from 30-120. Thus cutting the histogram between $T_1 = 30$ and $T_2 = 120$ renders only vegetation information. Stretching the 30-120 histogram to full dynamic range (0-255) of the display system makes the matter more perceivable. The resultant after this operation is illustrated in Fig. 4.3.2. We can easily notice the brightest patches of vegetation, other informations are suppressed.

4.3.5 VEGETATION PATHCING

A 7x7 low pass filter is applied to the resultant of the above treatment for making the vegetation areas more uniform and smooth (Fig. 4.3.3). After this operating the DN values range 43-90. Thresholding 43 and selecting all the above DN values equal to zero, clear-cut segmentation results. We can notice very bright patches of vegetation in a completely dark background, shown in the Fig.4.3.4.
The systematic steps of the procedure

1. Interpretation of image data for vegetation

2. DN value analysis for vegetation. In the present case it lies between 30 to 120

3. Histogram equalization $30 \rightarrow 0$, $120 \rightarrow 255$.

4. Application of $7 \times 7$ low pass filter, now vegetation ranges $43-90$

5. Thresholding $43$ and selecting all the above DN value equal to zero

6. Application of $3 \times 3$ laplacian Kernel, vegetation boundaries are detected.

7. Application of $5 \times 5$ low pass followed by histogram equalization, better unbroken boundaries

8. Thresholding with zero, vegetation patches in a dark background

9. Application of $3 \times 3$ laplacian Kernel, gentle and mild boundaries
4.3.6 BOUNDARY DETECTION AND DELINEATION

The job is over with the application of a 3×3 Laplacian filter (high pass) on the product of the previous operations. These boundaries possess some discontinuities and are broken (Fig. 4.3.5). This abnormality is removed by the application of a 5×5 low pass filter followed by histogram equalization. The boundaries are now unbroken (Fig. 4.3.6). To get gentle boundaries the resultant image is thresholded with zero (Fig. 4.3.7), which is followed by a 3×3 Laplacian kernel. Very gentle boundaries are detected (Fig. 4.3.8).
(a) Standard False Color Composite (FCC) [RGB, 432] of original, Rio-image data. Red patches are vegetation.

(b) NIR (TM-band-4) image of original Rio-image data. Brightest patches are vegetation.

Fig. 4.3.1
(a) Rio-NIR (TM-band-4) image after cutting the histogram between 30-120 followed by histogram equalization [30→0, 120→255].

Fig. 4.3.2
(a) Image (4.3.2) after application of a $7 \times 7$ low pass filter. Vegetation areas are more uniform and smooth. DN value range 43-90.

Fig. 4.3.3
(a) Image (4.3.3) after thresholding 43 and selecting all above DN value equal to zero. Clear cut segmentation.

Fig. 4.3.4
a (a) Image (4.3.4) after application of $3 \times 3$ high pass filter. Clear vegetation boundaries, but slightly broken.

**Fig.4.3.5**
(a) Image (4.3.5) after application of 5×5 low pass filter, followed by histogram equalization, thick but unbroken boundaries.

Fig. 4.3.6
a  (a) Image (4.3.6) after thresholding with zero. Binary like segmentation.

Fig. 4.3.7
(a) Image (4.3.7) after application of $3 \times 3$ high pass filter. Very mild, gentle and unbroken boundaries

Fig. 4.3.8