Chapter 3

Image Processing: An Overview

3.1 Introduction

Digital image processing deals with manipulation and analysis of images by using computer algorithm, so as to improve pictorial information for better understanding and clarity. This area is characterized by the need for extensive experimental work to establish the viability of proposed solutions to a given problem [42]. Image processing involves the manipulation of images to extract information to emphasize or de-emphasize certain aspects of the information, contained in the image or perform image analysis to extract hidden information.

Recent advances in Precision Farming have resulted in significant improvements in agriculture by increasing crop production, with good quality and low operating cost. In this context, image processing is used for a number of application areas that include soil testing, good quality seed selection, identification of nutrient deficiencies, monitoring and controlling of stress, weeds and diseases, assessment of crop status and yield.

3.2 Basic Image Processing and Analysis System

The Computer Vision System aims at recognizing objects of interest from given images and helps in developing the machine, that can perform visual function parallel to human vision. Computer Vision System consists of filtering, coding, enhancement, restoration, feature extraction, analysis and recognition of objects from image [43]. Processing of an image comprises of improvement in its appearance and effective representation of input image suitable for required application.

A digital image is simply a matrix where each number represents the brightness at
regularly spaced points or very small regions in the image. Mathematically, an image may be defined as a two dimensional function, $f(x, y)$, where $x, y$ are spatial (plane) coordinates and the amplitude of $f$ at any pair of coordinates $(x, y)$ is called the intensity or grey level of the image at that point.

Figure 3.1 shows the generic block diagram of Image analysis system. It includes image acquisition, preprocessing of input image, segmentation, feature extraction and presentation or classification.

Image acquisition involves capturing the images in the suitable form. Preprocessing improves the quality of the data by reducing artifacts. Segmentation groups pixel into regions, thereby defining the boundaries of the region of interest. Feature extraction and selection provides the measurement vectors. Feature extraction is followed by presentation or classification and is performed by estimating different features of the segmented region.

3.3 Image Acquisition

Many digital images are captured using visible light as the energy source; this has the advantage of being safe, cheap and easily processed with suitable hardware. Two very popular methods of producing a digital image are with Digital camera or Flat-bed Scanner. Generally, the image acquisition stage involves preprocessing, such as scaling.

3.4 Preprocessing

The preprocessing of image aims at selectively removing the redundancy present in captured images without affecting the details that play a key role in the overall process. That involves the following basic steps:
3.4.1 Image Re-sizing

Re-sizing of an image is performed by the process of the interpolation. It is a process which re-samples the image to determine values between defined pixels [44]. Thus, resized image contains more or less pixels than that of original image. The intensity values of additional pixels are obtained through interpolation if the resolution of the image is increased.

3.4.2 Filtering

Uncertainties are introduced into the image such as random image noise, partial volume effects and intensity non uniformity artifact (INU), due to the movement of the camera. This results in smooth and slowly varying change in image pixel values and lead to information loss, SNR gain and degradation of edge and finer details of image.

Spatial filters are used for noise reduction. These filters may be linear or non-linear filters. The linear filter uses spatial filtering masks and they are of two types:

Low Pass Filter

This filter opposes the edge and sharp details in an image and passes slowly varying characteristics. They are implemented by replacing the value of every pixel in an image by the average value of the gray levels in the neighborhood, defined by the filter mask. Spatial neighborhood averaging is used as a linear operation on the input image given by equation,

\[ g(x, y) = \frac{1}{M} \sum_{s} f(m, n) \]  

Where, \( S \) is the \( M \) - pixel neighborhood of points, surrounding the point \((x, y)\). However, implementing of linear low pass filter the edges are not maintained. It means noise reduction is accomplished by blurring which results in loss of fine details. The cost incurred in reducing noise and gaining contrast resolution is accompanied by loss of spatial resolution. Problem of blurring edges during smoothing can partly be overcome by ordered statistic filter or weighted averaging method [42, 43, 44].

In weighted average method, pixels are multiplied by different coefficients, giving more importance (weight) to some pixels over other pixels i.e. the pixel at the center of the mask is multiplied by a higher value than the other. The other pixels are inversely weighted as a function of their distance from the center of the mask. The smooth gray
level \( \tilde{g}(r,c) \) is obtained by a linear combination of \( g(m,n) \) over a neighborhood where weights are dependent on their relative position within the neighborhood. Thus, the basic strategy behind weighing the centre point at highest and then reducing the value of the coefficients as a function of increasing distance from the origin is an attempt to reduce blurring in smoothing process. Low pass filter mask is shown in Figure 3.2.

**High Pass Filter**

This filter oppose the slowly varying characteristics like background and passes edges and other sharp details. The filter mask includes negative coefficient at the center and positive coefficients in the outer periphery. The sum of the entire coefficient in this mask is zero. When the mask passes over a constant or slowly varying region, the output is zero or very small. Thus filter eliminates the zero frequency term. The result of this filter will be a somewhat edge enhanced image over a dark background.

Nonlinear spatial filters are Median filter, Max filter and Min filter, also operates on the neighborhoods pixels directly with coefficient. These filters compute median gray level value for noise reduction and used to find out the brightest and dimmest points in an image. High pass filter mask is shown in Figure 3.3.
3.5 Segmentation

Segmentation of an image is a process of dividing the image into homogenous, self consistent regions corresponding to different objects in the image. It separates image into meaningful regions [45]. Image can be segmented using basic properties of features of image like intensity, edge or texture.

3.5.1 Types of Segmentation

Depending on type of input image samples, segmentation can be classified as Gray scale single image segmentation and Histogram based segmentation.

Gray Scale Single Image Segmentation

In gray scale segmentation, a single image is used for feature extraction and segmentation. The gray scale single image segmentation methods can be subdivided as:

1. **Edge based segmentation methods**: Edge detection schemes suffer from incorrect detection of edges due to noise, over and under segmentation and variability in threshold selection in the edge image.

2. **Region growing segmentation methods**: The segmentations require an operator to select seeds and thresholds. Pixels around the seeds are examined and included in the region if they are within the thresholds. Results obtained with seed growing are generally dependent on the operator settings.

3. **Threshold based segmentation methods**: The most intuitive approach to segmentation is global threshold, which has been performed on various types of images. One common difficulty with this approach is to determine the value of the thresholds. Gray scale segmentation methods are generally limited to relatively simple structures.

Histogram Based Segmentation

Histogram based segmentation method uses the histogram to select the gray level for grouping pixels into regions [44, 46]. In an image, there are two entities: the background and the object, which can be differentiated by their gray level. The histogram of an image represents the gray levels of all the pixels of an image. Thresholding takes any pixel whose gray values are above threshold and sets all others to zero. The histogram peaks and the valleys between them are the keys for segmenting the image.
There are five segmentation techniques used for segmenting an image based on histogram and are described as below:

1. **Manual segmentation**: In this method, the image and its histogram are inspected manually and trial and error method is adapted to obtain better results. Even though manual thresholding gives better results, due to its trial and error methodology, it is time consuming and difficult to apply to large data set.

2. **Histogram peak technique**: This method finds the two peaks in the histogram corresponding to the background and object of the image. It sets the threshold halfway between the two peaks. By reviewing the histogram precisely, the threshold is selected for optimum results of segmentation.

3. **Histogram valley technique**: This technique uses the peaks of the histogram, but concentrates on the valleys between them. Instead of setting the midpoint arbitrarily halfway between the two peaks, the valley technique searches peaks to find the lowest valley.

4. **Adaptive histogram technique**: In this method, the peaks of the histogram are used in the first pass and adapt it to the objects found in the image in second pass. In the first pass, the adaptive technique calculates the histogram for the entire image. This smoothes the histogram and uses the peak technique to find the high and low threshold values. In the second pass, the technique works on each row and column area of the image individually. In each area, it segments using the high and low values found during the first pass. Then it computes the mean value for all the pixels segmented into background and object. It uses these means as new peaks and estimates new high and low threshold values for that rows and columns area now, it segments that area again using the new value.

5. **Triangle thresholding technique**: This method is useful in the case where the object and background occupy different ranges of gray level and object pixels produce a weak peak in the histogram. In this method, the threshold value is the maximum distance measured from the line constructed between maximum and minimum point of the brightness peak of the histogram and to brightness points of the histogram.

### 3.5.2 Edge Detection

In the area of image processing and computer vision, edges or contours of images provide valuable information towards human image understanding. Edges are often used
in image analysis for finding region boundaries [43]. Edges are the representation of discontinuities of image features.

Edge is the boundary between two regions with relatively distinct gray level properties. In a continuous image, a sharp intensity transition between neighboring pixels is considered as an edge. Edge corresponds to fast change in gray level and thus, considered as high frequency information. Therefore, the process of separation of high frequency information is edge detection.

There are many methods available for edge determination. These are either gradient based or zero crossing based. In gradient based edge detection, a gradient of consecutive pixels is taken in x and y directions. While in derivative based edge detection, the algorithm takes the first or second derivative on each pixel in the images.

In case of first derivative at the edge of the image, there is a rapid change of intensity. While in case of second derivative there is zero pixel value, termed zero crossing. The zero crossing property of the second derivative is quite useful for locating the centers of thick edges. Taking a derivative on each and every pixel of the image consumes a lot of computer resources and hence an operation called kernel operation is carried out. A kernel is a small matrix sliding over the image matrix-containing coefficients, which are multiplied, to the corresponding image matrix elements and their sum is put at the target pixel.

The first three of the edge detection methods, Sobel, Perwitts and Roberts, use the approximation of the first derivative of the intensity function, where it returns the edges at those points, where the gradient of the images intensity is at maximum. The only difference between them is the approximation of the first derivative.

The Robert’s method in edge detecting is one of the earliest techniques. Robert’s operator is as shown in Figure 3.4. The computational speed for this detector is fast and produces finer lines than other edge detection methods. However, the performance
of this detector is quite poor with significant noise.

In Sobel edge detection technique, formulas are applied on each pixel in the image and two kernels are obtained. Sobel operators is as shown in Figure 3.5. The resultant matrix is then obtained by taking the square root of the sum on the squares of those two kernels.

The Prewitt method is simpler to implement computationally than the Sobel operator method, but tends to produce noisier results. Prewitt operators is as shown in Figure 3.6. The Sobel detector provides a smoothing effect and it reduces the noise content in the image. The Perwitt and Sobel operator are among the most used devices in practice for computing digital gradients.

In Laplacian of Gaussian (LoG) method, a Gaussian filter is being used to filter out the image. The purpose of the Gaussian filter is to smooth the image before the Laplacian operator provides the image with zero crossing, in order to establish the location of the edges. The degree of blurring the image is controlled by the standard deviation.
The Zero-Cross detector works similar to that of the LoG detector, except that it allows the user to specify the filter used for filtering the image. Zero crossing method is preferable because of its noise reduction capabilities and potential for rugged performance.

Canny’s edge detection method has good detection and localization property with minimal response, where the edge in the image is only marked once and image noise does not create false edge. This detector is considerably strong due to its ability to detect weak edges. Canny’s method determines the edges by locating the local maxima and minima of the gradient of the intensity function. To begin with, the raw images are convolved with a Gaussian filter which results in a slightly blurred image. The local gradient and the edge direction can be determined by taking the first derivative of the intensity function. The edges are traced through the image using two threshold values, high and low. The high threshold is applied first where it marks out the edges that are fairly ‘obvious’ and ‘strong’. The ‘weak’ edges are included if they are connected to the strong edges.

3.5.3 Edge Detection Using Morphological Operation

In context to digital image processing, Morphology means study of topology or structure of objects from their image. In images, morphological operations are based on relations of two sets. One is an image and the second is a small probe, called a structuring element. The structuring element systematically traverses the image and its relation to the image in each position is stored in the output image [42].

Dilation and Erosion are fundamental morphological operations. Dilation expands an object to the closest pixels of the neighborhood. Normally, dilation is used to fill small holes and narrow gulfs in objects. If the original size needs to be preserved, then dilation is combined with erosion and is explained in next subsection. Erosion shrinks the object. Erosion of an image is the operation of assigning to each pixel the minimum value found over a neighborhood of the corresponding pixel whereas dilation is the operation of assigning to each pixel the maximum value of the neighborhood. The structuring element is a function of two variables that specifies the desired local gray level property [47]. The value of the structuring element is added or subtracted for calculating maximum or minimum levels in the neighborhood.
3.6 Feature Extraction

Feature extraction is a low level image processing operation which is usually performed as the first operation on an image. A feature can be defined as the “interest” part of an image. The desirable property for a feature detector is repeatability; i.e. whether or not the same feature will be detected in different images of the same scene. Step edges, lines and junctions usually convey the most relevant information of an image; hence it is important to detect them in a reliable way.

3.6.1 Colour Image Processing

It is important to know that there are many factors that may affect the colour from the nature to the computer screen. Camera and computer equipment can also change the appearance of colour on the computer screen from that in reality. To keep source of the error to a minimum, it is important to ensure consistency of equipment and processing so that two images that are to be compared have been processed in the same way.

Colour Spaces

To be able to describe colours, a reference frame, referred to as a colour space, is needed. The main idea is that every colour can be described by three vectors, one for every dimension in the colour space. Sometimes a fourth is added for convenience. The most frequently used colour spaces are as follows:

1. **RGB (Red, Blue, Green)**: The RGB space, which is made up of red, green and blue intensity components, is the most common within digital imagery. Any colour can be described by giving its red, green and blue values. The origin is black and the maximum intensity is white as shown in Figure 3.7.

![RGB colour space model](image-url)

Figure 3.7: RGB colour space model
2. **CMY and CMYK**: The CMYK (cyan, magenta and yellow) model is an exact inverse of the RGB space, which makes the origin white and not black as in the RGB model, so it is often used in the printing industry, where image start with white paper. An extension of the model is to add a black component, K. The idea of this comes from the fact that any CMY colour has an underlying gray component that consists of equal amount of cyan, yellow and magenta. This gray level can be obtained by the black colour instead, which is cheaper for printing.

3. **HSV and HSI**: The HSV (hue, saturation and value), HSI (hue, saturation and intensity) models are very similar. HSV colour space model is shown in Figure 3.8. The Hue defines the colour according to wavelength, while Saturation is the amount of colour. An object that has a deep, rich tone has a high saturation and one that looks washed-out has a low saturation. The last component differs between the models but they all describe the amount of light in the colour.

The main disadvantage of the RGB model and also CMY model is that humans do not see colour as a mix of three primary colours. Rather, our vision differs between hues with high or low saturation and intensity, which makes the HSV colour model closer to the human colour perception than the RGB model. In the HSV space, it is easy to change the colour of an object in an image and still retain variations in intensity and saturation such as shadows and highlights. It is simply achieved by changing the hue component, which would be impossible in the RGB space. The feature implies that the effects of shading and lighting can be reduced. A shaded area within an object can be detected as belonging to the object, even though it is darker. The reason for this is that the hue component does not change very much shading and lighting much.
4. **YUV and YIQ:** The YUV and YIQ models based on luminance and chrominance and are used in the PAL and NTSC broadcasting system respectively. The reason for their suitability in broadcasting system is their compatibility with hardware systems. The Y component, the luminance, corresponds to the brightness and is the only component that a black and white television receiver uses. The U and V (I and Q) components are used to describe the colour. The disadvantage of these models is that since only two colour components are used, not all colours that a computer screen is capable of displaying can be produced [42].

### 3.7 Representation or Classification

Representation almost always follows the output of feature extraction stage, which usually is raw pixel data, constituting either the boundary of a region or all the points in the region itself. In either case, converting the data to a form suitable for computer processing is necessary. The first decision that must be made is whether the data should be represented as boundary or as complete region. Boundary representation is appropriate when the focus is on external shape characteristics, such as corners and inflections. Regional representation is appropriate when the focus is on internal properties, such as texture or skeletal shape [45].

Choosing a representation is only a part of the solution for transforming raw data into a form suitable for subsequent computer processing. A method must also be specified for describing the data so that features of interest are highlighted.

Classification, also called feature selection, deals with extracting attribute that result in some quantitative information of interest or are basic for differentiating one class of object from the other [42].

### 3.8 Concluding Remarks

Image analysis system explained in the chapter comprises of preprocessing, segmentation, feature extraction and colour image processing. Preprocessing improves the quality of the data by reducing artifacts. Segmentation groups the pixels into regions, and hence defines the boundaries of the regions of the interest. Segmentation is followed by labeling of the regions. Feature extraction and selection provides the measurement vectors using which the image segmentation is performed.
The preprocessing of the images is performed by image resizing and filtering. Smoothing operation is used primarily to diminish the effects of noise due to motion and shift of the camera. The most common approach in detecting meaningful discontinuities in gray level is edge detection. Derivative operator or Gradient operator detects edge pixels by taking derivative followed by thresholding; two dimensional derivative operators calculate derivatives by edge masks.

Colour images can be processed by using different colour space. Most commonly used colour models are RGB, HSV and HSI.

Next chapter presents the algorithm based on image processing theory for seed selection of sugarcane.