CHAPTER 2

REVIEW OF LITERATURE

This research work, as mentioned in Chapter 1 (Introduction) is concentrating on developing techniques that automate the process of mango fruit quality assessment using skin characteristics as external quality indicator. This chapter begins with a description of history of mango along with its development details, followed by methods used during fruit quality assessment.

2.1. FRUIT QUALITY ASSESSMENT

Automatic fruit quality detection is an important aspect of fruit commercialization and grade classification. Food industries now place more emphasis on the usage of methods that can be effectively used to monitor the fruit quality. As a consequence, several researches have been conducted to develop full and semi-automated techniques that has the potential to inspect the quality of fruits and ensure safety of the fruits consumed by the public (Weyrich *et al.*, 2012; Rahman *et al.*, 2014). In almost all of these researches, the fruit appearance is an important index of detecting fruit quality (He, 2012; Schaare and Fraser, 2000). This importance has led to remarkable mechanization and automation in fruit packaging houses, where fruit quality detection and grading is considered as a highly dynamic area. Automation is highly preferred in both food technology and agricultural due to various reasons that include the following points.

(i) Increasing consumer demand of quality produce (Aleixos *et al.*, 2002).

(ii) Machines are more consistent than humans (Deck *et al.*, 1995; Njoroge *et al.*, 2002)

(iii) Scarceness of labor in developed countries (Walsh, 2005)

(iv) Opportunity to reduce labor costs (Bato *et al.*, 2000)

The past few years have envisaged several authors pointing out the major advances made in this field. Studman (2001) reviewed the operations in the postharvest industry, where computers and electronic technologies have had a huge impact. Brosnan and Sun (2004) compared different computer vision systems for horticultural produce blemish and disease detection. Garcia-Ramos *et al.* (2005) reviewed state-of-the-art non-destructive fruit firmness sensors. Butz *et al.* (2005) and Nicolai *et al.* (2006) compared different technologies to characterize the internal quality of fruits and vegetables. Janardhana *et al.* (2013) presented and analyzed various computer-aided methods that can be used for fruit inspection and quality assurance. Zhanga *et al.* (2014) provided an review on the principles, developments and applications of computer vision for external quality inspection of fruits and vegetables. In the same year, Swarnalakshmi and Kanchanadevi (2014) discussed the various methods available for grading fruits for quality inspection. Vimaladevi and Vijayarekha (2014) presented the various publications that used machine vision applications to locate fruits, detect defects. The review also included methods to remove noise from fruit images.

Fruit diseases were identified by Pydipati *et al.* (2006) applied color co-occurrence method (CCM) with statistical classification algorithms. Fernandez *et al.* (2002) introduced an automated visual inspection system to find the quality of preserved fruits. The features were extracted from the images and analyzed by on-line capabilities. The speed of the system was high.

Sirisathitkul *et al.* (2006) developed a technique to find the maturity of oranges by color grading based on image processing technique. The system consists of a CCD camera to capture the images and a classifier to make decision. Lua *et al.* (2010) proposed a method to detect dry matter of kiwifruit using near infrared spectroscopy. Synergy interval partial least square was used to select the wavelengths and four data sets were obtained to create model.

Fathi *et al.* (2011) developed a system to control osmotic dehydration and color changes of kiwi fruit using image processing techniques and artificial neural network. A multilayer network was formed to estimate the color changes, solid gain and water loss and found that the best results was provided using this system.

Iqbal *et al.* (2013) explained design and implementation of machine vision method used for automatic grading and sorting of apples. Two grading systems with same structure and for orientation of fruit (horizontal system and vertical system) were proposed. Machine vision system was also used to assist robots for fruit picking.

Fruit defects were identified by Pandiayammal et al. (2015) used unsupervised clustering algorithms to segment the regions. The color features were extracted from the images and the simulation results were high.

Although it is established that the internal quality assessment has important roles in the fruit postharvest stage, the external appearance of fresh produce also continue to be an important factor (Moreda *et al.*, 2009). The practice of using the fruit surface as an external quality indicator is common. The techniques used for this purpose should be non-destructive (Marcus *et al.*, 2012) and should use electronic devices and automated processes using advanced software.

2.1.1. Electronic Devices used for Fruit Quality Assessment

Over the years, many electronic devices have been developed, most of them trying to mimic human sensory characteristics, in order to automatically measure quality and quality-related attributes and are often oriented toward real-time and nondestructive testing. Electronics has been widely adopted in fruit grading machines since the 1980s, initially to support the use of load cells, for more accurate and faster weighing than the mechanical predecessors (Walsh, 2005). Widespread use of photocells and microprocessors enabled the development of color-sorting systems in the 1980s. This was followed by machine vision–based systems that used charge-coupled device (CCD) cameras and PC frame grabbers for sorting produce based on shape and color. Implementation of morphometric algorithms allowed estimation of fruit volume to be incorporated.

Although machines were initially marketed with vision or weighing capabilities, the combination of the two capabilities proved to be useful for the assessment of produce density. In the mid-1990s, automated blemish detection systems, employing more sophisticated algorithms (e.g., neural networks, fuzzy logics, etc.) provided new utilities to machines based on image processing. At the same time, the fact that computer images are 2D projections of a 3D scene, made researchers envisage different methods to obtain information on the whole surface of the produce, that is, imaging the fruits while rotating on rollers or using mirrors (Gutierrez *et al.*, 2012). Furthermore, use of wavelengths outside the visible spectral range, or combinations of visible and invisible information (ultraviolet, near-infrared) from images is becoming commercially available (Rizzi *et al.*, 2007).

A wide range of noninvasive imaging technologies (X-ray, ultrasonics, magnetic resonance, etc.) that are currently used in medicine are also slowly becoming available in the agri-food sector also (Süsstrunk and Fredembach, 2010). Widespread adoption of soft X-ray inspection systems by the security industry is driving prices down for this technology. This technology is also migrating into the food sorting industry (http://www.eufic.org/article/en/artid/X-rays-in-food-inspection). Spectrophotometry is also capable of extracting information related to the chemical composition of fruit and vegetables, thus providing data to estimate the quality of these products (Yamakawa *et al.*, 2012). Other technologies, such as electrical impedance (Rehman *et al.*, 2011), chlorophyll fluorescence (Chiu and Hsu, 2014), gas sensors (Tan *et al.*, 2005) and biosensors (Rana *et al.*, 2010) have also been developed in order to generate more complete and accurate assessment of quality.

The spectroscopic and imaging techniques are unique disease monitoring approaches that have been used to detect diseases and stress due to various factors, in plants and trees. Current research activities are towards the development of such technologies to create a practical tool for a large-scale real-time disease monitoring under field conditions. Various spectroscopic and imaging techniques have been studied for the detection of symptomatic and asymptomatic plant and fruit diseases (Dubey and Jalal, 2012c). Some the methods are:

(i) Fluorescence imaging (Bravo *et al.*, 2004)
(ii) Multispectral or hyperspectral imaging (Moshou *et al.*, 2006)
(iii) Infrared spectroscopy (Spinelli *et al.*, 2006)
(iv) Visible/multiband spectroscopy (Yang *et al.*, 2007, Chen *et al.* (2008)
(v) Nuclear Magnetic Resonance (NMR) spectroscopy (Choi *et al.*, 2004).

Hahn (2009) reviewed multiple methods (sensors and algorithms) for pathogen detection, with special emphasis on postharvest diseases. Several techniques for detecting plant diseases that affect growth of fruits and vegetables was reviewed by Sankarana *et al.* (2010). The review included techniques such as Molecular techniques, Spectroscopic techniques (Fluorescence spectroscopy and Visible and infrared spectroscopy) and Imaging techniques (Fluorescence imaging and Hyper-spectral imaging).

A ground-based real-time remote sensing system for detecting diseases under field conditions is developed by Moshou, (2005), which considers the early stage of disease development. Hyper-spectral reflection images of infected and healthy samples with an imaging spectrograph under ambient lighting conditions and field circumstances were used during defect and disease detection. The results showed that it was possible to detect presence of diseases and defects through the comparison of the 550 and 690 nm fluorescence images. On the other hand color based features were used by Singh *et al.* (2012) and Gupta *et al.* (2013).

2.1.2. Computer Vision-Based Software Systems

Eventhough several advanced technologies have been manufactured for defect detection using electronic devices for fruit quality analysis and assessment, recent years are assigning equal important to automatic computer vision-based technology as they have more potential during quality assessment and grading. As examples, a color vision system for peach grading (Miller and Delwiche, 1989), computer vision based date fruit grading system (Janobi, 1998), machine vision for color inspection of potatoes and apples (Tao, 1995) and sorting of bell peppers using machine vision (Shearer and Payne, 1990).

Some machine vision systems are also designed specifically for factory automation tasks such as intelligent system for packing 2-D irregular shapes (Bouganis and Shanahan, 2007), versatile online visual inspections (Garcia *et al.*, 2006; Garcia and Villalobos, 2009), automated planning and optimization of lumber production using machine vision and computer tomography (Bhandarkar *et al.*, 2008), camera image contrast enhancement for surveillance and inspection tasks (Kwok *et al.*, 2009), patterned texture material inspection (Ngan and Pang, 2009) and vision based closed-loop online process control in manufacturing applications (Cheng and Jafari, 2008). The reason behind such popularity is that the tasks of fruit quality inspecting and assessment presents numerous challenges as aspects such as appearance cannot be defined clearly and the degree of irregularity that is acceptable is not uniform.

Most of the techniques in literature places more emphasis on finding methods that increase speed of quality assessment and reduce human errors. These techniques aim to find defects present in fresh market fruits such as frost damage, flesh damage, hail damage, scald, russet, rot and bruise. All these techniques have steps that can be generally grouped as, preprocessing, segmentation, feature extraction, defect detection and classification.

2.2. PREPROCESSING

Preprocessing fruit images is a crucial initial step performed before image analysis. Many preprocessing methods are available in the literature, which use some algorithm to improve the quality of the fruit image through noise removal (Thangam *et al.*, 2009; Zelelew, 2008). This step is vital during defect detection and fruit grading processes. Fruit images are often affected by impulse noise, which degrade image quality and obscure information required for accurate defect detection.

Fruit images are often degraded by the presence of impulse noise, which is an area of research work that has attracted many researchers (Mélange *et al.*, 2011; Mohammad *et al.*, 2011; Hao *et al.*, 2012). This section presents a brief discussion on the various techniques available to remove impulse noise from digital images.

Several methods have been proposed to solve this problem and they include adaptive filter (Manikandan *et al.*, 2004; Kalavathy and Suresh, 2011), multistate median filter (Chen and Wu, 2001), weighted median filter (Yang *et al.*, 1995) and switching median filters (Ping *et al.*, 2007). Vector directional filters uses directional image vectors during denoising (Lukac, 2004).

Variations to vector directional filters are the weighted vector direction filter which implement a tracking algorithm to identify the varying signal and noise statistics. Peer Group Filters (PGF) that uses statistical properties of accumulated distances for vector median filtering has also been proposed (Smolka, 2008). This algorithm switches between vector median and the original central pixel.

Hsu *et al.* (1993) presented an adaptive separable median filter as a post filter for removing the blocking effects which generally originate from lower-bit-rate image transmission. This filter adaptively transforms from a traditional median filter to a low-pass filter progressively when the filter is close to the position of blocking effects. Simulation results demonstrated that this filter is extremely useful not only in removing the blocking effects, but also in maintaining the main advantages of the median filter such as edge preservation and noise reduction.

Sun *et al.* (1994) used a switching scheme for median filtering to eliminate impulse noise in images. The switching procedure is based on the local measurements of impulses which are called desired detection. A median and weighted median based desire detector was developed and the significant properties were discussed. Investigation exhibited that the filtering arrangement capitulates an output image which is significantly improved than those of median, weighted median and most favorable stack filters.

Chen and Wu (2001a) proposed the use of median based switching schemes, called multi-state median (MSM) filter. By using uncomplicated thresholding logic, the output of the MSM filter was adaptively switched surrounded by those of a grouping of center weighted median (CWM) filters that have different center weights.

In the same year, Chen and Wu (2001b) also devised a novel adaptive operator that was based on the differences between the current pixel and the outputs of center-weighted median (CWM) filters with varied center weights to remove impulse noise. Extensive simulations show that the proposed scheme consistently works well in suppressing both types of impulse noise (salt and pepper and random) with different noise ratios. This work of Chen and Wu (2001b) improved the work of center weighted median filter by including more threshold values. Similarly, Zhang and Karim (2002) used a Laplacian edge detector and the detected edges were preserved during noise removal.

Wang and Zhang (1999) proposed algorithms that performed noise removal in two steps. The first step detected the noisy pixels, while the second step replaced these corrupted pixels using median filters. Uncorrupted or clean pixels were left untouched. This method was efficient in removing noise in detail region, but ignored the noise present in the edges. This produced a smearing effect. This was solved by Nikolova (2004) by using modified switching median filter to incorporate a weighted median filter using two thresholds during the detection and replacement of noisy pixels.

Dong *et al.* (2007) proposed a new impulse detector based on the differences between the current pixel and its neighbors aligned with four main directions for impulse noise detection. The result was then combined with a new directional weighted median (DWM) filter for noise removal.

Kang and Wang (2009) developed a denoising algorithm that used a rank-order-based switching median filter to solve the problems posed by threshold selection in the conventional switching median filter.

Xu *et al.* (2009) applied adaptive fuzzy switching filter that adopted a fuzzy logic approach for the enhancement of images corrupted by impulse noise. In order to achieve optimal detail preservation, the algorithm used maximum- minimum exclusive median method to estimate and handle the corrupted pixels.

Esakkirajan *et al.* (2011) designed a modified decision based unsymmetrical trimmed median filter algorithm for the restoration of gray scale and color images that are highly corrupted by salt and pepper noise. This algorithm replaced the noisy pixel using a trimmed median.

A video directional weighted median filter with suitable colour correction to remove random valued impulse noise from colour video sequences was proposed by Bhupender *et al.* (2013). The switching median technique was utilized to protect noise free pixels from filtering so as to avoid blurring of frames. The threshold used for identifying isolated noisy pixels was made adaptive based on the local statistics of the current pixel component. The restoration of noisy pixels was done by utilizing the brightness and colour information obtained from directional weighted median filter and vector directional filter, respectively. Local statistics for video filtering involves a three dimensional sliding window for estimation of characteristics of current pixel. This three dimensional sliding window provided spatial as well as temporal information about the neighbourhood for independent restoration of frame under consideration but decision for filtering was made only on the basis of information obtained from current frame only because of the fact that the consecutive frames may also have special characteristics. Only noise free pixels of three dimensional sliding window, are used for restoration of frame under consideration.

Gupta *et al.* (2013) introduced a new method for the enhancement of gray scale images, where images were corrupted by fixed valued impulse noise (salt and pepper noise). This method gave better output for low-density impulse noise and worked better when compared to other famous filters like Standard Median Filter (SMF), Decision Based Median Filter (DBMF) and Modified Decision Based Median Filter (MDBMF) and so on. This method improved the Image Enhancement factor (IEF), Peak Signal to Noise Ratio (PSNR), visual perception and also reduced blurring artifacts. The proposed algorithm replaced the noisy pixel by trimmed mean value. When previous pixel values, 0's and 255's are present in the particular window and all the pixel values are 0's and 255's then the noisy pixels are replaced by mean value. Different gray-scale images are tested via proposed method. The experimental result shows better Peak Signal to Noise Ratio (PSNR) value, Image Enhancement Factor (IEF) and with better visual and human perception.

Recently, Devi and Pritamdas (2014) used a switching median filter incorporated with an iterative clustering based noise detection algorithm to efficiently restore images degraded with impulse noise. This filter engaged a noise detection algorithm for image reinstatement to detect noisy pixels by classifying the pixels into three pixel clusters iteratively. If the pixel in deliberation falls beneath the middle cluster in the last iteration, the pixel is confidential as clean otherwise it is detected as corrupted. Corrupted pixels use median filtering to convert themselves to uncorrupted form while uncorrupted pixels remain untouched.

In the same year, Dhas and Suresh (2014) also modified switching median filter to remove impulse noise. This technique first detected impulse noise, which was then, restored using the median of neighboring noise-free pixels. This method used a statistics-based detection window that measures noisy pixels using four directions. The window size was determined adaptively and the results proved that the proposed method was more efficient than the conventional switching median filter.

A similar approach was also used by Bhadouria *et al.* (2014) where the proposed denoising algorithm performed filtering procedure only on the degraded pixels, while keeping the uncorrupted pixels intact. However, this method used a coupled window scheme for the removal of high density noise. A sliding window of increasing dimension centered at any pixel was used during noise removal, where the noisy pixel was replaced consecutively by the median value of the window. However, if the entire pixels in the window are noisy, then the dimension of sliding window is increased in order to obtain the noise-free pixels for median calculation.

Mukhopadhyay *et al.* (2014) proposed a fuzzy-based switching median filtering technique that aims at noise detection and removal of impulse noises. Two types of noise models are used to obtain the noisy images. In this two-step process, the noise-free pixels remained unchanged. The proposed detection algorithm uses 5×5 window, based on all neighboring pixels on the center of the window of a noisy pixel. Two weighted median filter was developed, which was applied selectively to the noisy pixel based on the characteristics of the neighboring pixels within the window. Instead of a single threshold, two threshold values were used in the proposed fuzzy membership function to partition the noise level and accordingly, a filtering method is applied to restore the corrupted pixel.

A novel partition-based fuzzy median filter for noise removal from corrupted digital images was proposed by Lin *et al.* (2014). The proposed filter was developed using the weighted sum of the current pixel value and the output of the median filter, where the weight is set by using fuzzy rules concerning the state of the input signal sequence to indicate to what extent the pixel is considered to be noise. Based on the adaptive resonance theory, a neural network model was implemented with a new weight function where the neural network model is employed to partition the observation vector. In this framework, each observation vector was mapped to one of the M blocks that form the observation vector space. The least mean square (LMS) algorithm was applied to obtain the optimal weight for each block.

Recently, Pradeep kumar *et al*. (2015) applied fuzzy switching median filter, which was an extension to the classical switching median filter by employing fuzzy inference mechanism. Thus the experimental results showed better PSNR value compare to others.

2.3. SEGMENTATION

Image segmentation is the first step in image analysis and pattern recognition and is performed before feature extraction and disease detection. It is a critical and essential component of image analysis and pattern recognition system. It is one of the most difficult and challenging task in image processing and determines the quality of the final result of analysis. Image segmentation is the process of dividing an image into different regions such that each region is, but the union of any two adjacent regions is not, homogeneous.

A correct segmentation may be, in some cases, very difficult, especially in poor quality image. Segmentation is considered as a critical step in automatic defect detection systems and there is a large literature on segmentation dating back over 30 years (Jain and Dubes, 1988). Some examples of modern techniques are presented by Liu and Nixon (2007), Peters and Kerdels (2007), Karvelis *et al.* (2008), Zhang *et al.* (2010a), Menze *et al*., (2014) and Chen *et al.* (2015). A detailed review on the various segmentation techniques available is presented in Pham *et al.* (2005).

In general, image segmentation techniques can be classified into five categories (Raut *et al.*, 2009). They are, Threshold-based techniques, Histogram-based techniques, Edge detection techniques, Region-based techniques and Watershed Transformation techniques. The application of cluster based techniques (Pauwels and Frederix, 1999), neural network-based segmentation techniques (Pereira *et al.*, 2010; Du, 2010), compression based techniques (Biswas, 2003) and Graph partitioning methods (Sumengen and Manjunath, 2006) have also been used for the purpose of segmentation. Another popular method used is the curve evolution methods (Farzinfar *et al.*, 2010).

As this research work uses a clustering-based approach to segment fruit images, it is described in detail in this section. Clustering can be formally considered as a particular kind of NP-hard grouping problem (Hruschka *et al.*, 2009). This assumption has stimulated much research and use of efficient approximation algorithms.

The clustering-based segmentation algorithms are categorized as hard-clustering methods and fuzzy-clustering methods. The conventional hard clustering method restricts each point of the data set to exclusively just one cluster. As a consequence, with this approach the segmentation results are often very crisp, i.e., each pixel of the image belongs to exactly just one class. An example of hard-clustering is the use of K-Means algorithm.

However, in many real situations, for images, issues such as limited spatial resolution, poor contrast, overlapping intensities, noise and intensity inhomogeneities variation make this hard (crisp) segmentation a difficult task. This was solved using fuzzy set theory (Zadeh, 1965), which produced the idea of partial membership of belonging described by a membership function; fuzzy clustering as a soft segmentation method has been widely studied and successfully applied in image segmentation (Kwon *et al.*, 2003; Balafar, 2014; Feng *et al.*, 2010).

Fuzzy clustering algorithms are broadly classified into two groups (Hoppner *et al.*, 1999): i) Classical and ii) Shape-based. There exist many classical fuzzy clustering algorithms in the literature, among them the most popular and widely used algorithms are given below.

i) Fuzzy C Means (FCM) (Bezdek, 1981)

ii) Suppressed Fuzzy C Means (SFCM) (Fan *et al.*, 2003)

iii) Possibilistic C Means (PCM) (Krishnapuram and Keller, 1993)

iv) Gustafson-Kessel (GK) (Gustafson and Kessel, 1979)

Fuzzy C-Means (FCM) algorithm is the most popular method used in image segmentation because it has robust characteristics for ambiguity and can retain more information than hard segmentation methods (Bezdek *et al.*, 1993). However, it has a serious limitation: it does not incorporate any information about spatial context, which cause it to be sensitive to noise and imaging artifacts. Moreover, the time complexity of the algorithm is high. To compensate for these drawbacks of FCM, several authors have proposed methods to enhance the operation of the conventional FCM algorithm (Chatzis *et al.*, 2008; Li *et al.*, 2011) and some of these are discussed in this section.

Mazouzi and Batouche (2008) proposed new approach for improving range image segmentation, based on fuzzy regularization of the detected edges In the initial stage, a degraded version of the segmentation is generated by a new region growing-based algorithm. In the next step, the resulting segmentation is extracted by a robust fuzzy classification of the pixels on the resulting edges which correspond to borders of the extracted regions. Pixels on the boundary between two adjacent regions are labelled taking into account the two regions as fuzzy sets in the fuzzy classification stage, using an enhanced version of the Fuzzy C-Mean algorithm. The process is continued for all region boundaries in the image till convergence is reached and the output is taken as the final segmented result.

Arifin and Asano (2006) proposed a new approach of image thresholding by using cluster organization from the histogram of an image. A new similarity measure which is proposed is based on inter-class variance of the clusters to be merged and the intra-class variance of the new merged cluster.

Sudhavani and Sathyaprasad (2009) have proposed the application of modified Fuzzy C means algorithm for lip segmentation problem. The modified fuzzy C-means algorithm is able to take the initial membership function from the spatially connected neighboring pixels. Successful segmentation of lip images was possible with their proposed approach. Comparative study of the proposed modified fuzzy C-means is done with the traditional fuzzy C-means algorithm.

Sivanand (2013) proposed novel method for performing image segmentation. An adaptive local threshold algorithm was proposed, to overcome the problem of conventional threshold segmentation method. This method works well with images having non-uniform illumination. The problem of Fuzzy C-Means clustering method can be avoided by using kernel based clustering technique. Kernel Fuzzy C-Means clustering method can be used for images having unequal sized clusters.

Yu and Wang (2007) proposed a two-dimensional fuzzy C Means (2DFCM) algorithm for the molecular image segmentation. The first stage consists of the noise suppression by utilizing a method combining a Gaussian noise filter and anisotropic diffusion techniques. The second stage is the texture energy characterization using a Gabor wavelet method. The third stage is introducing spatial constraints provided by the denoising data and the textural information into the two-dimensional fuzzy clustering. The incorporation of intensity and textural information allows the 2DFCM algorithm to produce satisfactory segmentation results for images corrupted by noise (outliers) and intensity variations.

Logeswari and Karnan (2010) proposed an approach which describes segmentation method consisting of two phases. In the first phase, the MRI brain image is acquired from patients' database, after that film artifact and noise are removed. And finally Hierarchical Self Organizing Map (HSOM) is applied for image segmentation.

Roy and Bandyopadhyay (2012) proposed an interactive segmentation method that enables users to perform segmentation in a quick and efficient manner. In addition to area of the region and edge information the proposed method uses a type of prior information also its symmetry analysis. Padole and Chaudhari (2012) proposed an efficient method that combined two standard algorithms, mean shift and normalized cut is performed to detect defects in medical images.

Ahmed *et al.* (2002) introduced a regularization term into the standard FCM to impose the neighborhood effect. Later, Zhang *et al.* (2003) incorporated this regularization term into a kernel-based fuzzy clustering algorithm. Similarly, Li *et al.* (2003) incorporated this regularization term into the adaptive FCM (AFCM) algorithm (Pham and Prince, 1999) to overcome the noise sensitivity of AFCM algorithm. Although the latter two methods are claimed to be more robust to noise, they show considerable computational complexity.

Shankar and Pal (1994) presented a progressive sub sampling method called fast fuzzy c-means (FFCM) (Uma Shankar & Pal, 1994). FFCM generates a sequence of extended partitions of the entire dataset by applying the original FCM to a nested sequence of increasing size subsamples. It terminates when the difference between successive extended partitions is below a threshold.

Speed is always a concern during clustering using FCM, which is directly proportional to the size of the image. Pal and Bezdek (2002) developed the extensible FFCM (eFFCM) clustering algorithm for the segmentation of very large digital images. Hathaway and Bezdek (2006) discuss an extension of the eFFCM method, geFFCM, to non-image object data.

O'zdemir and Akarun (2002) proposed a Partition Index Maximization (PIM) algorithm by adding the partition coefficient (PC) (Bezdek, 1974) into the FCM objective function and successfully applied PIM for a color quantization of images. This modification can form a cluster core for each cluster and data points inside the cluster core will have membership values of zero or one. Since the volumes of each cluster core in PIM are equal, Wu *et al.* (2005) proposed a fuzzy compactness and separation (FCS) algorithm. The volumes of each cluster core generated by FCS are different. Moreover, Yang *et al.* (2008) proposed the alpha-cut implemented fuzzy clustering algorithm (FCMa) whose cluster cores are generated by the alpha-cut concept.

Several authors have attempted to improve FCM by modifying the objective function of the algorithm. Tolias and Panas (1998) proposed a spatial constraint rule-based system that spatially smoothes membership function to enhance the results of FCM. Similarly, Pham (2002) produced a modified objective function by incorporating a spatial penalty. Admed *et al.* (2002) also attempted to modify the objective function to improve FCM algorithm. In this method the objective function that allowed the pixel to be labeled by the influence of its neighbourhood labels.

Zhang and Chen (2004) proposed spatially constrained kernelized FCM, which replaced similarity measurement of objective function by a kernel-induced distance measure and incorporated a different penalty term that used spatial neighbourhood information. Later, Shen *et al.* (2005) proposed an improved FCM by adding neighbourhood attraction to the distance of a pixel from a class center using intensity and distance attractions. In both cases, the weighted average of neighbouring pixels membership was taken into consideration.

Acton and Mukherjee (2000) incorporated multiscale information to enforce spatial constraints. Dave and Krishnapuram (1997) proposed the idea of a noise cluster to deal with noisy data using the technique, known as Noise Clustering. This technique separated the noise into one cluster but experimental results showed that this technique is more suitable to remove noises than image segmentation.

Zhang *et al.* (2009) developed a robust clustering technique by deriving a novel objective function for FCM. Kang *et al.* (2009) proposed another technique which modified FCM objective function by incorporating the spatial neighborhood information into the standard FCM algorithm.

Yang *et al.* (2005) proposed a novel penelized FCM (PFCM) for image segmentation where penalty term acts as a regulizer in the algorithm which is inspired by neighborhood maximization (NEM) algorithm and is modified in order to satisfy the criterion of FCM algorithm. Shen *et al.* (2005) presented an algorithm called Improved FCM which changed the distance function used in FCM, i.e. the distance between pixel intensity and cluster intensities and applied neural network optimization technique to adjust parameters in the modified distance function.

Chuang *et al.* (2006) presented a FCM algorithm that incorporated spatial information into the membership function of clustering. The membership weighting of each cluster is altered after the cluster distribution in the neighborhood is considered. Rhee and Hwang (2001) proposed Type 2 fuzzy clustering. Type 2 fuzzy set is the fuzziness in a fuzzy set. In this algorithm, the membership value of each pattern in the image is extended as type 2 fuzzy memberships by assigning membership grades (triangular membership function) to Type 1 fuzzy membership.

Van Huy pham (2015) proposed a k-means clustering algorithm to segment the fruit regions based on the Euclidean distance. The experiment showed good results in terms of human observation and processing time.

2.4. FEATURE EXTRACTION

An image feature is a descriptor of an image, which can avoid redundant data and reduce the effects of noise and variance. In computer imaging, feature detection, extraction and selection are vital for image analysis and interpretation. Feature detection is a low-level image processing application. For an image, the feature is the “interest” part in image. In the pattern recognition literature the name feature is often used to denote a descriptor.

Within the field of feature extraction there are many areas of specialization. The techniques for these different specializations can vary substantially. Research into automated feature extraction from digital images dates back to the seventies and since that time, technology has improved and commercial access to imagery has continued to expand. Feature extraction and selection are based on the mathematical selection, computation and manipulation of image features with high efficiency, robustness and invariance (Lichun *et al.*, 2009).

Interest in automatic feature extraction has increased significantly since the advent of digital imagery and the possibilities associated with electronic processing. Focused conferences provide an overview of many of the techniques available (http://www.imagefeatures.org/events, Baltsavias, *et al.*, 2001; Gruen, *et al.*, 1997; 1995). In addition, several commercially available photogrammetric workstation systems cited in the review by Plugers (1999) now incorporate some automated feature extraction capability. Other companies such as Definiens (2003) and Visual Learning Systems (VLSI, 2003) are developing software specifically targeted at feature extraction.

Repeatability is the desirable property of a feature detector (Padmavathi, 2012). Many approaches, such as principal component analysis, minimum noise fraction transform, discriminant analysis, decision boundary feature extraction, non-parametric weighted feature extraction, wavelet transform and spectral mixture analysis (Lu and Weng, 2007) may be used for feature extraction, in order to reduce the data redundancy. The common aim of all these techniques is to detect and extract features hidden inside an image so as to improve the classification performance.

Feature extraction and selection are considered as the important tasks that allow the determination of the most relevant features for fault detection in fruit images. Extracting and selecting suitable features are critical steps for successfully implementing a defect classification system. In general, this research work considers color and texture features of mango fruit during defect detection, which are described in the following sections.

2.4.1. Color Features

Color is one of the most important features of images. Color features are defined subject to a particular color space or model. For the colour feature extraction the first step is segmentation. For the segmentation the choice of the colour space is also vital. A number of color spaces have been used in literature, such as RGB, LUV, HSV and HMMD (Stanchev *et al.*, 2003). Various popular colour spaces have been described in Table 2.2 (Cheng *et al.*, 2001; Gonzalo, 2005).

A number of important color feature extraction methods have been proposed in the literatures, including color histogram (Jain and Vailaya, 1996), Color Moments (CM) (Flickner *et al.*, 1995), Color Coherence Vector (CCV) (Pass and Zabith, 1996) and color correlogram (Huang *et al.*, 1997). Table 2.3 provides a summary of different color methods excerpted from the literature (Zhang *et al.*, 2012), including their strengths and weaknesses.

Apart from these methods like the following were also used for color feature extraction from fruits. More details of these features are provided by (Zhou *et al.*, 2012).

- Dominant Color Method (DCD)
  - Strawberry Fruit (Liming and Yanchao, 2010)
- Intensity Distribution Method
  - Dates (Ohali, 2011)

TABLE 2.2

COLOR SPACES USED FOR FEATURE EXTRACTION

Color Space	**Merits**	**Demerits**
RGB	Convenient for display	Due to its high correlation not good for color image processing
YUV	Less computation time and solves the problem of RGB's correlation	Correlation eventhough lower than RGB still exists
HIS	Based on human color perception Hue is invariant of different types of highlights, shadows and shading	At low saturation, it is numerically unstable because of its non-linear transformation
CIE Spaces (L*u*v* or L*a*b*)	Color and intensity information can be controlled independently Efficient in measuring small color differences	Singularity problem as other non-linear transformations

TABLE 2.3

COMPARISON OF COLOR EXTRACTION METHODS

Method	Merits	Demerits
Histogram	Simple to compute, intuitive	High dimension, No spatial information, sensitive to noise
CM	Compact and Robust	Cannot describe all color and no spatial information
CCV	Spatial Information	High dimension and high computation cost
Correlogram	Spatial Information	Very high computation cost, sensitive to noise, rotation and scale

- Mean of Color Images
    - Mango Fruit (Razak *et al.*, 2012)
    - Oil Palm Fruit (May and Amaran, 2011)
    - Palm Oil Fresh Fruit Bunches (Alfatni *et al.*, 2008)
    - Apple Fruit (Leemans *et al.*, 2002; Leemans and Destain, 2004, Blasco *et al.*, 2003)
    - Fruit Classification (Seng and Mirisaee, 2009)
- Nine Color Characteristics Data
    - Apple Fruit (Nakano,1997)
    - Apple Fruit and Grapes Fruit (Xiaobo *et al.*, 2007)
- HIS color model was used by
    - Lemon (Khojastehnazhand *et al.*, 2010)
    - Apple Fruit (Rao and Renganathan, 2002)
    - Citrus Fruit (Blasco *et al.*, 2007),
    - Starfruit (Abdullah *et al.*, 2006)
- The simple RGB color space was used by
    - Palm Oil Fresh Fruit Bunches (Alfatni *et al.*, 2008),
    - Apple Fruit (Qiabao *et al.*, 2009).
- Direct color mapping method
    - Dates and Tomatoes (Lee *et al.*, 2010)

According to the analysis performed by Vyas *et al.* (2013), most of the above mentioned techniques have the following disadvantages and the usage of simple RGB model is more advantageous due to the following.

- Computationally complex
- Works well with high Contrast images
- Not accurate as other methods
- Requires color transformation as a mandatory step
- Sensitive to lighting and other environmental conditions

Thus the above mentioned demerits are further confirmed by Carrillo and Peñaloza, (2009) (Coffee-Excelso Beans quality) and Jamil *et al.* (2009) (Oil Fresh Fruit Bunches). Both RGB and HSI were also used in online detection of fruit defects (Qiabao *et al.*, 2009).

2.4.2. Texture Features

Texture is a very useful characterization for a wide range of images. It is generally believed that human visual systems use texture for recognition and interpretation. In general, color is usually a pixel property while texture can only be measured from a group of pixels (Haralick, 1979). Based on the domain from which the texture feature is extracted, they can be broadly classified into spatial texture feature extraction methods (STFEM) and spectral texture feature extraction methods (SpTFEM). For STFEM, texture features are extracted by computing the pixel statistics or finding the local pixel structures in original image domain, whereas SpTFEM transforms an image into frequency domain and then calculates feature from the transformed image. Both have advantages and disadvantages as listed in Table 2.4.

TABLE 2.4

TYPES OF TEXTURE FEATURES

Texture Method	Merits	Demerits
Spatial Texture	Meaningful, easy to understand and can be extracted from any shape without losing information	Sensitive to noise and distortion
Spectral Texture	Robust, but are less complex	Need square image regions with sufficient size

Approaches to texture analysis are usually categorized into four groups of methods, namely, structural-based, statistical-based, model-based and transform-based (Materka and Strzeleck, 1998). Structural approaches represent texture by well-defined primitives (micro-texture) and a hierarchy of spatial arrangements (macro-texture) of those primitives.

In contrast to structural methods, statistical approaches do not attempt to understand explicitly the hierarchical structure of the texture. Instead, they represent the texture indirectly by the non-deterministic properties that govern the distributions and relationships between the grey levels of an image. Methods based on second-order statistics (i.e. statistics given by pairs of pixels) have been shown to achieve higher discrimination rates than the power spectrum (transform-based) and structural methods.

Model based texture analysis, using fractal and stochastic models, attempt to interpret an image texture by use of, respectively, generative image model and stochastic model. The parameters of the model are estimated and then used for image analysis. In practice, the computational complexity arising in the estimation of stochastic model parameters is the primary problem. The fractal model has been shown to be useful for modelling some natural textures. It can be used also for texture analysis and discrimination. However, it lacks orientation selectivity and is not suitable for describing local image structures.

Transform methods of texture analysis, such as Fourier, Gabor and wavelet transforms represent an image in a space whose co-ordinate system has an interpretation that is closely related to the characteristics of a texture (such as frequency or size). Methods based on the Fourier transform perform poorly in practice, due to its lack of spatial localization. Gabor filters provide means for better spatial localization; however, their usefulness is limited in practice because there is usually no single filter resolution at which one can localize a spatial structure in natural textures.

Texture feature extraction techniques include methods like First-order histogram based features (Pratt, 2007; Papoulis, 2002; Levine, 1985), Co-occurrence matrix based features (Julesz, 1975; Haralick, 1979), Multiscale features (Cohen, 1989; Teuner, 1995; Dunn, 1995) and Gray-Tone Difference Matric (GTDM) (Amadasun, 1989). Out of these, co-occurrence matrix-based features are more frequently used during texture feature extraction.

Several researches have used texture features for fruit inspection and analysis. A detailed review of several of texture extraction techniques was given by Reed and Dubuf (1993) and Zhang *et al.* (2009). The following lists some of the texture features used by fruit inspection and classification systems.

- GLCM Features
    - Mango Fruit (Pujari *et al.*, 2013)
    - Tomato Fruit (Yamamoto *et al.*, 2014)
    - Maize (Patil and Kumar., 2012)
    - Arable (Moshou *et al.*, 2011)
    - Tobacco (Guru *et al.*, 2011; Lu *et al.*, 2014)
    - Grapes (Zsófi *et al.*, 2014)
    - Pomogranate (Khoshroo *et al.*, 2009)
- Co-Occurrence Method
    - Citrus Fruit (Bandi *et al.*, 2013)
    - Soyabean (Cui *et al.*, 2010)

    - Apple (Chassangne-Berces *et al.*, 2010)
    - Grapes (Segadea *et al.*, 2011, 2013)
- Wavelets and Curvelets
    - Orange Fruit (Jafari *et al.*, 2012)
    - Guava Fruit and Citrus Fruit (Khoje *et al.*, 2013)
- Local Binary Pattern Based Texture Feature
    - Apple Fruit (Dubey and Jalal, 2012a, 2012b)

Advanced techniques in agricultural application are also using a combination of color and texture features during fruit quality and grading tasks. Examples include Semary *et al.* (2015) for tomato fruit grading, Payne *et al.* (2014) for mango fruit and Arivazhagan *et al.* (2010) for fruit recognition. Pujari *et al.* (2013c) on the other hand used the fusion of RGB color features and GLCM texture features to classify fruits as normal and defective.

2.5. FRUIT DEFECT DETECTION AND CLASSIFICATION

Fruit disease identification is generally viewed as an instance of image classification. Recent years have encountered lot of activities in the area of using image classification for defect detection (Qiabao *et al.*, 2009; Katyal and Srivastava, 2012; Zheng and Lu, 2012). Classification methods attempt to partition pixels into different classes using different methods and it is being used by many different researches. Kotsiantis (2007) and Kotsiantis *et al.* (2006) provide detailed discussion on the various algorithms available for classification. Wei *et al.*(2005) on the other hand, focused on providing a review of machine-learning methods and analyzed several state-of-the-art machine-learning methods for automated classification and disease detection in food products.

While classification is a well-studied problem, its usage on defect detection has been found only in the last few decades (Catlett, 1991; Chan and Stolfo (1993a, 1993b). Several classification models have been proposed over the years, e.g. neural networks (Lippmann, 1987), statistical models like linear/quadratic discriminants (James, 1985), decision trees (Breiman *et al.*, 2007) and genetic models (Goldberg, 2006).

Cubero *et al.* (2011) presented some modern developments in the field of the inspection of the interior and outside quality of fruits and vegetables. Similarly, Gomes *et al.* (2012) also presented an appraisal of the major publications in the last ten years with admiration to innovative technologies and to the spacious application of systems of visual inspection in the sectors of exactitude farming and in the food manufacturing.

According to Kleynen *et al.* (2005), Bayesian classifier and Support Vector Machine (SVM) classifier are the most frequently used classifiers for defect detection and classification problems. Leemans *et al.* (1998, 1999) developed a defect detection model where pixels are compared with a pre-calculated model and classified as defective or healthy fruit.

Dubey and Jalal, (2012d, 2013), Dubey (2013) and Dubey *et al.* (2013), proposed a framework for fruit recognition and defect detection. The study considered 15 types of fruits and vegetables for recognition. The system performed defect detection using three steps, namely, region of interest segmentation, feature extraction and classification. The system used improved sum and difference histogram as texture feature along with color features to train multi-class Support Vector Machine (SVM). The segmentation or region of interest extraction was performed using simple threshold approach (Li, Wang and Gu, 2002; Mehl *et al.*, 2002).

Lu *et al.* (2011) proposed a Principal Component-Support Vector Machine (PC-SVM) and Support Vector Machine (SVM) models collective with fractal examination were residential and contrast with categorization models based on RGB concentration values for sorting and classification of fruits. Zheng *et al.* (2010) used Chlorophyll Fluorescence-Support Vector Machine (C-SVM) for fruit categorization.

Sivamoorthi and sujatha (2015) proposed a novel approach for detection and classification of fruit based on local binary patterns. The k-means clustering was applied to segment the regions and SVM classifier was used to find the type of detects in the cluster.

Defect detection and classification method used depends on the fruit being examined and this section presents some works related to ten types of fruits, namely, apple, dates, blueberries, grapes, peach, pomegranate, watermelon, banana, orange and mango (Kodagali, and Balaji, 2012; Gill *et al.*, 2014).

2.5.1. Apple

Woodford *et al.* (1999) have detailed the image processing and neural network classification methods like neural network classifier using wavelets applied to the task of identifying the pest that caused the damage to apple fruits and leaves in orchards. The system obtained high classification rate on a standard neural network without any special alteration to the learning algorithm.

Rao *et al.* (2002) used geometrical models to design and develop six methods for determining size apples. The models used were circle method, parabola method, ellipse method, principal axis method, radius and area signature method and coefficient of variation method.

Kavdir and Guyer (2003) applied Fuzzy logic (FL) as a decision making support to grade apples. Quality features such as the color, size and defects of apples were measured through different equipment. In the same year, Woodford *et al.* (2003) offered a solution to the problem of training and testing a neuro-fuzzy system for the purpose of image recognition when there are a limited number of images. Features of interest were segmented from each image and then used to train a neural-fuzzy system. This increased the number of data examples used to train the system.

Leemans and Destian (2004) investigated the defect segmentation of "Golden Delicious" apples using machine vision. To segment the defects, each pixel of an apple image was compared with a global model of healthy fruits by making use of the Mahalanobis distances. The proposed algorithm was found to be effective in detecting various defects such as bruises, russet, scab, fungi or wounds.

Unay and Gosselin (2005) introduced a computer vision based system to automatically sort apple fruits. Artificial neural network was used to segment the defected regions on fruit by pixel-wise processing. Linear Discriminant, nearest neighbor, fuzzy nearest neighbor, Adaboost and support vector machines classifiers were tested for fruit grading, where the last two were found to produce best recognition results.

Unay *et al.* (2006) proposed a method for Apple Defect Detection and Quality Classification with MLP-Neural Networks. Here, the initial analysis of a quality classification system for "Jonagold" and "Golden Delicious" apples were shown. Later, Color, texture and wavelet features are extracted from the apple images. Principal components analysis was applied on the extracted features and some preliminary performance tests were done with single and multi layer perceptrons.

Bennedsen *et al.* (2007) used artificial neural networks and principal components to detect surface defects on apples in near infrared images. Neural networks were trained and tested on sets of principal components, derived from columns of pixels from images of apples acquired at two wavelengths (740 nm and 950 nm). In an iterative process, different ways of preprocessing images prior to training the networks were attempted. Best results were obtained by removing the background and applying a Wiener filter to the images.

Bin *et al.* (2007) introduced an automated computer recognition system for inspection of apple quality based on Gabor feature-based kernel principal component analysis (PCA) method. First, Gabor wavelet decomposition was employed to extract appropriate Gabor features. Then, the kernel PCA method with polynomial kernels was applied in the Gabor feature space to handle non-linear separable features. Mirzaei and Saraee (2007), in the same year, proposed an algorithm to detect the apple skin bruises by applying a threshold and a few morphological operations such as opening and closing.

Sudhakara *et al.* (2009) adopted HSI model for sorting and grading of fruits by colour and developed a prototype for on-line sorting of Apples based on colour, size and shape. In the same year, Stajnko *et al.* (2009) proposed a fruit detection algorithm to estimate the diameter and number of apple fruits on a tree.

Later, Lak *et al.* (2010) developed an edge detection-based and color-shape based algorithm to segment images of red apples obtained under natural lighting, which were later used to detect defects.

Unay and Debeir (2011) presents an application work for grading of apple fruits by machine vision. The segmentation of defects was performed by minimal confusion with stem/calyx areas on multispectral images, from which the statistical, textural and geometric features were extracted. Using these features, statistical and syntactical classifiers were trained for two- and multi-category grading of the fruits.

Dubey *et al.* (2013) proposed a novel defect segmentation method of fruits based on color features with K-means clustering unsupervised algorithm. Color images of fruits were used for defect segmentation, carried out into two stages. At first, the pixels were clustered based on their color and spatial features, where the clustering process was accomplished. Then the clustered blocks were merged to a specific number of regions. This two step procedure was useful in increasing the computational efficiency by avoiding feature extraction for every pixel in the fruit image.

Cetişli and Büyükçingir (2013) proposed a new prediction model for the early warning of apple scab based on artificial intelligence and time series prediction. The infection period of apple scab was evaluated as the time series prediction model and the important hours of duration were determined with the feature selection methods, such as Pearson's correlation coefficients (PCC), Fisher's linear discriminant analysis (FLDA) and an adaptive neuro-fuzzy classifier with linguistic hedges (ANFC_LH). The experimental dataset with selected features was classified by ANFC_LH and predicted by an adaptive neural network (ANN) model.

2.5.2. Dates

Ali *et al.* (2003) developed a neural network to classify seven major varieties of date fruit: Berhi, Khlass, Nubot Saif, Saqei, Sefri, Serri and Sukkari using models incorporating selected physical and color features of each variety.

Ohali (2011) had implemented a prototypical computer vision based date grading and sorting system. A set of external quality features such as flabbiness, size, shape, intensity and defects are defined and extracted. Based on the extracted features the system classified dates into three quality categories (grades 1, 2 and 3: defined by experts) using back propagation neural network classifier.

Khalid and Tamer (2012) classified dates based on attributes extracted from a computer vision system (CVS). Two models of neural networks have been applied as classifiers: Multi-Layer Perceptron (MLP) with BackPropagation and Radial Basis Function (RBF) networks.

A customized feature extraction followed by classification was used by Haidar *et al.* (2012) to perform automatic detection and classification of date fruits. The parameters of the classifier was tuned in a generalized manner and features were extracted based on their color, size, shape and texture. Three different classifiers, namely, K Nearest Neighbour, Linear Discrimination Analysis and Back Propagation Neural Network (BPNN) were used for classification. Among the three classifiers, BPNN classifier produced best results.

2.5.3. Blueberries

Defects in frozen blueberries were also investigated by Sugiyama *et al.* (2010). Several methods including NIR (Near InfraRed) spectral imaging, discriminant analysis, wavelets and machine learning algorithms were used during defect detection. Ahlawat *et al.* (2011) used hyperspectral remote sensing as a tool to detect leaf rust at an early stage in blueberries. The tool measured reflectance in the wavelength range from 350 to 2500nm using a handheld hyperspectral spectro-radiometer and observed differences in spectral reflectance at a number of wavelengths in the visible, NIR and SWIR (Short Wave InfraRed) regions. The NIR region showed significant difference between the three varieties of blueberry. The results indicated the possibility to detect differences in healthy and leaf rust infected blueberry plants at an early stage.

Leiva-Valenzuela and Aguilera (2013) proposed a simple and non expensive computer vision method to remove blueberry unities with fungal damage. It automatically segregated unities with fungal decay, shriveling and mechanical damage from health unities and successfully classified images with fungal decay and global damage (fungal decay, shriveling or mechanic damage).

2.5.4. Grapes and Citrus

Janik *et al.* (2007) conducted a study to compare the performance of Partial Least Squares (PLS) regression analysis and ANN for the prediction of total anthocyanin concentration in red-grape homogenates from their visible-near-infrared (Vis-NIR) spectra. The proposed method combined the advantages of the data reduction capabilities of PLS regression with the non-linear modeling capabilities of ANN. ANN with PLS scores required fewer inputs and was less prone to over-fitting than using PCA scores.

Kim *et al.* (2009) designed technologies of color imaging and texture feature analysis that is used for classifying citrus peel diseases under the controlled laboratory lighting conditions. A total of 39 image texture features were determined from the transformed hue (H), saturation (S) and intensity (I) region-of-interest images using the color co-occurrence method for each fruit sample. The model using 14 selected HSI texture features achieved the best classification accuracy, which suggested that it would be best to use a reduced hue, saturation and intensity texture feature set to differentiate citrus peel diseases.

Kumar *et al.* (2012) proposed a method for detecting infected citrus using airborne hyperspectral and multispectral imaging. This work used a hyperspectral imaging software (ENVI, ITT VIS) during image analysis. The infected areas were identified using image-derived spectral library, mixture tuned matched filtering (MTMF), spectral angle mapping (SAM), and linear spectral unmixing. The experimental results showed that the accuracy of the MTMF method was greater than the other conventional methods. A similar work was previously done by Polek *et al.* (2007).

2.5.5. Peach

Esehaghbeygi *et al.* (2010) introduced a machine vision system for evaluation and classification of peach. Physical features such as size and colour were measured to categorize peaches into three quality classes of red-yellow, yellow-red and yellow. The HSI model was used to extract color features and boundary values were selected to extract size features.

Alipasandi *et al.* (2013) introduced a machine vision and Neural Network system to classify three varieties of peach namely Anjiri peach cultivar, Shalil Nectarine cultivar and Elberta peach cultivar. Image acquisition and processing toolboxes of MATLAB software were used to visualize, acquire and process the images directly from the computer. Some qualitative information was extracted and fed as inputs to classify the objects into different categories.

2.5.6. Pomegranate

Blascoa *et al.* (2009) studied the agricultural developments of Instituto Valenciano de Investigaciones Agrarias (IVIA) during the past 15 years. The institute has developed computer vision systems for the automatic, on-line inspection of fresh and processed fruits. One such system was a machine for the automatic inspection of pomegranate arils for fresh consumption. This machine inspects, classifies and separates the arils in four categories, removing those that do not fulfill the minimal specifications. Multivariate analysis models were used to classify the arils.

Drying of pomegranate arils was predicted by Motevali *et al.* (2010) by creating ANN and mathematical models. Ten semi-theoretical and empirical models were fitted to the experimental data to evaluate and select the best model for thin-layer drying of pomegranate arils. Experiments were conducted at six temperature levels of 45, 50, 55, 60, 65 and 70°C and three levels of air velocity (0.5, 1 and 1.5 m/s). Regression analysis of mathematical models showed that the Midili model fitted best to the measured data. However, regarding R2 and MSE criteria, ANN modelling yielded a better prediction of pomegranate arils moisture ratio during drying of arils compared to all the mathematical models studied.

Later, Blasco *et al.* (2011) developed machine for automatic inspection of pomegranate arils for fresh consumption wherein this machine individualizes, inspects, classifies and separates the arils in four categories. Multivariate analysis models were used to classify the arils.

2.5.7. Watermelon

Hassan *et al.* (2007) have classified the long type watermelon depending on the fruit shapes. Physical characteristics of watermelon such as mass, volume, dimensions, density, spherical coefficient and geometric mean diameter were measured. Relations and correlations coefficient were obtained between characteristics for normal and non-standard fruit shape. It was found that weight of normal watermelon could be determined by image analysis.

2.5.8. Banana

An electronic nose based system, which employed an array of inexpensive commercial tin-oxide odour sensors, was developed by Llobet *et al.* (1999) to analyze the state of ripeness of bananas. Readings were taken from the headspace of three sets of bananas during ripening over a period of 8-14 days. A principal components analysis (PCA) and investigatory techniques were used to define seven distinct regions in multi-sensor space according to the state of ripeness of the bananas, predicted from a classification of banana-skin colors. Then three supervised classifiers, namely Fuzzy ARTMAP, LVQ and MLP were used to classify the samples into the observed seven states of ripeness. It was found that the Fuzzy ARTMAP and LVQ classifiers outperformed the MLP classifier.

Wang *et al.* (2009) proposed a non-destructive measuring and evaluating method for fruits based on color identification. The color images of fruits were taken and RGB histograms were calculated and used as quality parameters to feed as input to BP neural network classifier with three layers. For verifying the proposed method, the cultivar used was bananas.

2.5.9. Orange

Simoes *et al.* (2001) investigated on the applicability of color classification using an artificial neural network in the fruit-sorting domain. Further this approach was used for the segmentation of colored images represented by the RGB color system. Jointly with color analysis, shape analysis was done to generate a robust and real time system. The model was tested for orange classification according to a Brazilian standard and was able to provide fruit classification under less restricted visual conditions.

Boonmung *et al.* (2006) proposed an approach for quality inspection of orange fruit using color and texture features. Input image was segmented using histogram based thresholding technique to identify whether the portion of image was defective or non-defective. For defected samples, texture features were extracted from 3D co-occurrence distribution. Sum of squared distance (SSD) was calculated between texture features of training and test fruits. Out of a database of 150 samples, 80 images were fed for training and remaining was used for testing. Experimental results showed that the proposed method improved classification rate when compared with conventional classifiers.

Mercol *et al.* (2008) presented an automatic orange classification system that used visual inspection to extract features from images. Several data mining algorithms were used to classify the fruits in one of the three pre-established categories. These were five decision trees (J48, Classification and Regression Tree (CART), Best First Tree, Logistic Model Tree (LMT) and Random Forest), three artificial neural networks (Multilayer Perceptron with Backpropagation, Radial Basis Function Network (RBF Network), Support Vector Machine) and a classification rule. The results were encouraging due to good accuracy achieved and low computational costs.

Rasekhi *et al.* (2011) developed a comprehensive algorithm to combine image processing and neural network techniques for sorting orange fruits into size groups (Small, Medium and Large). RGB color features were extracted and fed to a back propagation network model with a number of training functions including variable learning rate back propagation (MLP-GDM), resilient back propagation (MLP-RP) and scaled conjugate gradient (MLP-SCG) were used for ANN modeling. The results showed that the multi layer perceptron with RP and SCG transfer functions had the least error.

Bama (2011) have proposed an approach for quality inspection of orange fruits using combined color texture features. In proposed method, input image is segmented using histogram based thresholding technique to identify whether the portion of image is defective or non defective. For defected samples, texture features were extracted from 3D co-occurrence distribution. Sum of Squared Distance (SSD) is calculated between texture features of training and test fruits. Non defective fruits were considered as extra class whereas class I and class II were moderately and highly defective fruits respectively.

Balogun *et al.* (2013) developed a non-destructive method to predict the status of orange fruits, based on internal quality. Histogram analysis and the features extracted from Magnetic Resonance Imaging (MRI) were applied as input to train artificial neural networks in order to predict the orange fruit status. Different structures of multi-layer perceptron neural networks with back-propagation learning algorithms were developed using MATLAB. Levenberg-Marquardt algorithm (trainlm) gave the best performance fitness as compared to backpropagation algorithm used with least Mean Square Error (MSE). The findings proved that ANN and MRI had the capability of predicting the internal content and detecting the defect based on water proton content.

Zaragoza (2013) evaluated an industrial image analysis system equipped in mobile platform device, in comparison to two other devices; a characterized computer vision system and a spectrophotometer in the analysis of color on food. The device was employed in the field while the fruit was being harvested. The results proved that the mobile platform device predicted the color index of citrus with a good reliability and is effective for classification of the fruit according to its color.

2.5.10. Mango

Salim *et al.* (2005) presented the study of using an artificial olfactory system as a non-destructive instrument to measure fruit ripeness. The cultivar chosen for this study was Harumanis mango. The system comprised of an array of semiconductor gas sensors as well as data acquisition and analysis components. It used readings taken from Harumanis mangoes of different ripeness over a period of time. Each stage of ripeness of the mangoes left a different pattern or fingerprint onto the sensors array. Artificial Neural Network (ANN) was used to classify mango into different stages of ripeness.

Hasnah *et al.* (2005) proposed automatic nondestructive classification system of Harum Manis mango. The defect detection and quality classification was performed using infested Harum Manis mango fruits by integrating X-ray imaging techniques and Artificial Immune Systems (AIS). The classification was performed by applying the AIS self and non-self recognition process unto the Harum Manis mango X-ray images.

Yimyam *et al.* (2005) described image processing techniques that can detect, segment and analyze the mangoes physical properties such as size, shape, surface area and color from images. Sixty mango sample images taken by a digital camera were analyzed and segmented based on hue model. The results show the technique to be a good alternative for grading mango as compared to manual one.

Size is one of the major parameters that the consumer identifies to be related to the quality of mango. Teoh and Syaifudin (2007) presented a model to measure the weight of Chokanan mangoes using image processing techniques. Number of pixels of mango region in the captured image was counted by the PCI software and a relationship between mango pixels and mango weights was found using statistical method of regression. The findings suggest that the image processing and analysis techniques are practical, feasible and effective in estimating weight of Chokanan mangoes.

Slaughter (2009) surveyed the literature on nondestructive methods for objective assessment of maturity in mango. Future research topics were proposed, with a focus on the mango cultivars to address needs for the development of nondestructive methods for objective assessment of maturity in mango.

Zakaria *et al.* (2012) introduced the classification of mangoes maturity and ripeness levels using fusion of the data of an electronic nose and an acoustic sensor. Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) were able to discriminate the mango harvested at week 7 and week 8. However the best performance was shown by hybrid LDA-Competitive Learning Neural Network classifier.

In the same year, Razak *et al.* (2012) implemented methodologies and algorithms that utilize digital fuzzy image processing, content predicated analysis and statistical analysis to determine the grade of local mango production in Perlis. The algorithms used knowledge from fuzzy image clustering and experimental results proved that the performance of the proposed algorithm is better than their conventional counterparts.

Zhang and Wu (2012) introduced a least-squares support vector machine (LS-SVM) classifier to detect the degree of browning on mango fruits as a function of fractal dimension (FD) and L*a*b* values. Results showed that correct classification rates of 85.19% and 88.89% were achieved for the LS-SVM models based on different values of FD and L*a*b*, respectively.

Khoje and Bodhe (2013) performed case study of mango fruit from Maharashtra, India for size grading purpose. Mango images were captured using CCD camera and size analysis was carried out using the MATLAB package. Various size estimation metrics were used as feature vectors for two classifiers namely Feed Forward Neural network (FFNN) and Support Vector Machines. Experimental results showed that Statistical method gave an average size grading efficiency of 97% irrespective of classifiers for mango size grading.

2.6. CHAPTER SUMMARY

From the literature survey, it is understood that although a variety of algorithms have been developed for the various operations of defect detection, the solutions produced focus on each operation separately and an unified framework that enhances each and every step of defect detection system is sparse, but highly desired.

Defect detection from fruit images is a complex operation for which several algorithms have been proposed, but however, although many approaches have been developed, which approach is suitable for a given application area is not fully realized. Selection of a suitable algorithm for the various steps requires consideration of many factors, such as its effect on classification accuracy, algorithm performance and computational resources.

The literature study shows that computer vision systems have been used increasingly in industry for inspection and quality assessment purposes as they can provide rapid, economic, hygienic, consistent and objective assessment. However, difficulties still exist, evident from the relatively slow commercial uptake of computer vision technology in all sectors. Even though adequately efficient and accurate algorithms have been produced, processing speeds still fail to meet modern manufacturing requirements. However, the survey reveals that still much of the work needs to be concentrated on the fruits defect detection in variety of angles of the fruit.

From the various publications, it has been deduced that soft computing models have shown a remarkable performance in fruit defect detection and classificiation. While a number of promising technologies exist, non destructive assessment of fruit classification is achieved through computer vision systems and soft computing models. In future food industry is expected to make big profits through the use of more robust soft computing models. However, the research on these areas is still an on-going process where improvements in the area of increasing accuracy, reducing error rate and increasing speed of defect detection and classification is still highly desired.

This research work was to design and develop a framework that enhances the various operations of defect detection proposes several enhancement methods that can improve the quality, segmentation and defect detection. The algorithms were designed for mango fruit and the research design and various methods used in the methodology are introduced in the next chapter, Chapter 3, Methodology.