Chapter 3

Shape Grading

3.1 Introduction

Shape is an important aspect that is considered by a consumer while purchasing any fruit [Leemans and Destain, 2004*a*] and so it is of importance while assessing fruit quality. Description of fruit shape has been often necessary in agricultural research for many purposes such as cultivar description [Beyer et al., 2002], evaluation of consumer preferences, investigating heritability of fruit shape traits [Currie et al., 2000], stress distribution analysis in the fruit skin and determining misshapen fruit in a cultivar [Rashidi and Gholami, 2008]. Abnormality in shape ultimately leads to rejection of a foreign consignment. Thus, a cost effective and objective computer vision system is badly needed to segregate misshapen Mango fruits.

In recent years, for content based image retrieval, computer vision techniques have been developed by researchers using shape descriptors. The shape features can be extracted from digital images to characterize the shape of fruits in order to discriminate the different shapes during processing [Du and Sun, 2004*b*, Leemans and Destain, 2004*a*]. Various features for shape description and measurement have been studied in the literature. Size-dependent features (including compactness, convexity, roundness, length, width, and length/width ratio), boundary encoding, invariant moments, and Fourier descriptors are the most popular shape features applied in the quality inspection of food industry [Zhang, Huang, Li, Zhao, Fan, Wu and Liu, 2014].

Size-dependent shape measurements use single or combination of two or more size parameters to form a shape description expression [Zheng et al., 2006] which can reflect the overall shape of the objects. Due to its easy realization and less time consumption, the size-dependent shape features are still widely used for the overall shape sorting of apples and orange([Kondo et al., 2000]). However, most of the time, these size-dependent features are not sufficient to determine irregular shapes of fruits and in that case, boundary encoding, invariant moments, and Fourier descriptors are more powerful size-independent features for shape descriptions and measurements. Boundary encoding represents a chain code vector which records the sequence of coordinates of boundary pixels to describe its shape. Invariant moments can describe the shape by their magnitudes which are translation, rotation, and scale invariant. Generally, in order to get a more accurate shape classification result, in most applications, more than one shape features are used to sort the fruits.

Imou et al. [2006] developed a three dimensional shape measurement system for strawberry fruit in which the 3D shape was reconstructed from nine side views by the volume intersection method, and was compared with data measured by a laser scanner. The results showed that the 3D shape of strawberries could be measured by the proposed method. Kiwifruit shape sorting using geometrical features was evaluated by [Rashidi and Gholami, 2008] in which significant difference in shape parameters such as aspect ratio and ellipsoid ratio were detected.

Therefore, the present study was designed to develop a rapid image processing algorithm that permits reproducible, quantitative description of Mango shape. The aim of this study is to evaluate the irregularity of the Mango fruit shape using Fourier based shape descriptors and Neural Networks, thus aiming at more reliable, accurate and more sophisticated automated classification system for Mango fruit. Mango cultivatrs produced in the Konkan area (District: Sindhudurg) of India are considered for experimentation. As shape deformity is rarely observed on Guava fruits; they are not considered for shape sorting.

3.2 Shape Signatures

Shape signature represents the shape by one dimensional function which is derived from shape boundaries and provides information about the shape's features. Despite the fact that shape signatures can be used to describe a shape alone, they are often used as a preprocessing tool for other algorithms that extract features. Several shape signatures have been used to derive shape descriptors such as complex coordinates, centroid function, chord length signatures, curvature signature, etc. However, shape descriptors derived from different signatures can have significantly different effect on the result of shape retrieval [Zhang and Lu, 2003].

A contour C can be denoted as an ordered sequence of N coordinate points,

C = $\lambda_t = (x(t), y(t), t = 0, 1, , N-1)$, where C is closed. i.e. λ_i+N = λ_i.

Two shape signatures derived from a contour C are discussed here, namely complex coordinates and centroid distance signature.

3.2.1 Complex Coordinate

It is simply a complex number generated from boundary points.

$$R(t) = [x(t) - x_c] + i[y(t) - y_c] \tag{3.1}$$

Where (x_c, y_c) is the centroid of the shape and can be given as an average of boundary points.

$$x_c = \frac{1}{N}\sum_{t=0}^{N-1} x(t) \quad and \quad y_c = \frac{1}{N}\sum_{t=0}^{N-1} y(t) \tag{3.2}$$

3.2.2 Centroid distance signature

The centroid distance function is expressed by the distance of the boundary points to the centroid (x_c, y_c) of the shape.

$$r(t) = \sqrt{[x(t) - x_c]^2 + [y(t) - y_c]^2} \tag{3.3}$$

Fig. 3.1 shows centroid distance signatures of the four Mangoes. The position of the centroid and the center of gravity is fixed in relation to any shape. Since this shape signature is only dependent on the location of the centroid and the points on the boundary, it is invariant to the translation of the shape.

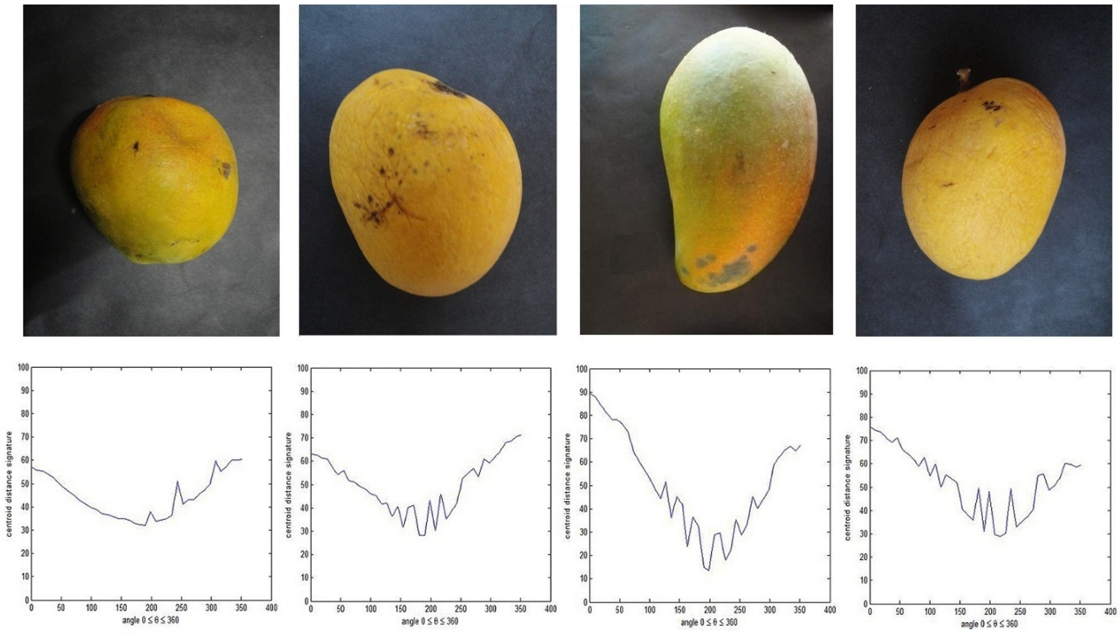

FIGURE 3.1: Mango shapes (at top) and their respective centroid distance signature(at bottom)

3.3 Shape Descriptors

Fourier descriptor (FD) is one of the widely used shape descriptors. In general, the FD is obtained by applying a Fourier transform on a shape signature. A shape signature is any 1-D function representing 2-D areas or boundaries which uniquely describes shape. The performance of FD method is affected by the shape signature as different shape signatures give different FD. Many shape signatures, such as complex coordinates, centroid distance, tangent angle, curvature, cumulative angle and so on, have been proposed for deriving FD. Zhang and Lu [2002] compared six different FDs which are derived from different shape signatures. They concluded that FD derived from centroid distance function and proposed area function are the most suitable for shape representation.

3.3.1 Generic Fourier descriptor (GFD)

All pixels within the shape region are considered in the region based shape descriptor. The most commonly used region based descriptors are Zernike moments, Legendre moments, Generic Fourier Descriptors, and Wavelet Descriptors. A region based Fourier descriptor is called a generic FD as it can be applied to any application of shape representation and classification.

Shape analysis using Fourier Transform is backed by well developed and well-understood Fourier Theory. However, it is not desirable to acquire shape features using FT directly, because the acquired features are not rotation invariant and not compact. Therefore, a modified polar Fourier Transform (MPFT) is proposed by treating the polar image in polar space as a normal two-dimensional rectangular polar image [Zhang and Lu, 2002].

If we plot a polar image into cartesian space, it is the normal rectangular image. Therefore, if we apply 2-D Fourier Transform on this rectangular image, the polar FT has the similar form to the normal 2-D discrete FT in cartesian space.

For a given shape image *F(x, y)*, the modified polar FT is obtained as,

$$pf(\rho,\phi) = \sum_r \sum_i F(r,\theta_i)\exp[j2\pi(\frac{r}{R}\rho + \frac{2\pi i}{T}\phi)] \tag{3.4}$$

Where $0 \leq$ r (t) = $\sqrt{[x(t)-x_c]^2 + [y(t)-y_c]^2}$ and θ_i = i($\frac{2\pi}{T}$)
where ($0 \leq i \leq T$); $0 \leq \rho \leq$R; $0 \leq \phi <$ T and (x_c, y_c) is the center of mass of the shape; *R* and *T* are the radial and angular resolutions.

The physical meanings of ρ and ϕ are clear; they are the ρ^{th} radial frequency and the ϕ^{th} angular frequency selected to describe shapes. The determination of the number of ρ and ϕ for shape description is physically achievable, because shape features are normally captured by the few lower frequencies.

The acquired Fourier coefficients are translation invariant. Rotation and scaling invariance are achieved by the following normalization.
GFD = $\frac{|pf(0,0)|}{area}, \frac{|pf(0,1)|}{|pf(0,0)|}, \ldots.. \frac{|pf(0,n)|}{|pf(0,0)|}, \cdots \frac{|pf(m,0)|}{|pf(0,0)|} \cdots \frac{|pf(m,n)|}{|pf(0,0)|}$

Where *area* is the area of the bounding circle in which the polar image resides; m is the maximum number of the radial frequencies selected and n is the maximum number of angular frequencies selected. Finally m and n can be adjusted to achieve hierarchical coarse to fine representation requirement.

The advantage of MPFT over FT is that the acquired spectrum is rotation invariant and more concentrated to the origin. MPFT is also more advantageous than conventional polar FT, because both the radial features and the circular features captured by the coefficients are physically meaningful.

3.3.2 Contour Based Fourier Descriptors (FD)

In most of the applications, region of shape is not important; only outline or boundary is enough to describe that shape; such kind of shape descriptors are called as contour based shape descriptor. Simple one dimension Fourier transform is used by researchers to represent the shape.

Fourier Descriptor is obtained by taking the Fourier transform of shape signature. Discrete Fourier Transform (DFT) and Inverse Fourier Transform (IDFT) is given by *f(k)* and *r(t)* respectively.

$$f(k) = \frac{1}{N}\sum_{t=0}^{N-1} r(t)\exp\frac{-j2\pi kt}{N} \tag{3.5}$$

Where *K* = - *N/2,......, N/2 -1*

$$r(t) = \frac{1}{N}\sum_{k=0}^{N-1} f(k)\exp\frac{j2\pi kt}{N} \tag{3.6}$$

Where *K* = - *N/2,......, N/2 -1*

Contour of shape in spectral domain is represented by Fourier coefficients *f(k)*. For scale invariance, all coefficients are normalized by dividing them by the magnitude of the first coefficient. As the phase of coefficients gives rotation information; so to achieve rotation invariance, phase information is ignored. Zhang and Lu [2001] concluded that FD outperforms any other contour based shape descriptor; but for disjoint shapes, contour based descriptors are not suitable.

3.3.3 Wavelet Based Fourier descriptor (WFD)

Wavelet Descriptor is a set of wavelet transformed coefficients applied to a contour of shape. Wavelet transform analyzes image information at different resolutions by dilating the scale of the wavelet function and constructs a time scale representation of the signal. It thus relates local properties of the signal to the evolution of wavelet transform coefficients when the scale varies. It outperforms Fourier transform as it captures both frequency and location information at a time. Wavelet coefficients are generated from mother wavelet function as shown by equation 3.7.

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \int Z(t)\psi(\frac{t-b}{a})dt \tag{3.7}$$

where b is the shifting parameter, a is the scaling parameter, and $\sqrt{a}$ for energy normalization at different scales. For any specific scale a, wavelet coefficients can be obtained.

Wavelet Fourier descriptors [Yadav et al., 2007] are obtained by applying a Fourier transform on wavelet coefficients $\psi_{a,b}$ and is defined as (3.8),

$$f^a(k) = \frac{1}{N} \sum_{b=0}^{N-1} \psi_{a,b}(t) \exp \frac{-j2\pi b}{N} \tag{3.8}$$

Thus, advantage of multi scale representation and Fourier shape descriptors are combined in this description. WFDs are invariant to translation because of the centroid function and to make it rotation invariant, phase information is ignored. For scale invariance, this descriptor has been normalized.

Shape contour is represented in the frequency domain using WFD. Wavelet gives two kinds of coefficients namely approximation and detailed. Detailed coefficients give finer details about object shape. By applying Fourier Transform, N number of coefficients can be obtained. Depending upon interest or application, these numbers of coefficients can be limited.

3.4 Material and methods

The samples of Mangoes with various shapes and sizes used in this study were collected from the Regional Fruit Research Center, Vengurla (Dist: Sindhudurg, Maharashtra), Shailesh Nursery, Malkapur (Dist: Kolhapur) and Mango orchards of Devgad and Ratnagiri (Fig. 3.2). A database of 300 images, consisting of Alphonso (150), Payari (50), Totapuri (40), Langada (30), Dashhari (30) was created; out of which 190 graded as well formed and 110 graded as deformed Mangoes.

3.4.1 Image Acquisition and Pre-processing

The fruits were photographed using a Sony DSC-W290 camera with a 28-140 mm lens at 35 mm. The camera resolution was 12.1 Mega pixels. The D65 daylight illumination was used while acquiring images. The captured scene from the experimental set-up had the fruit image including the background. To facilitate the segmentation process, black paper was used in the background.

All the algorithms for image pre-processing and segmentation, contour extraction and shape descriptors and classification are written in MATLAB v7.11.0-R (2010b)) (Math Works, Inc., USA).

FIGURE 3.2: Mango database of different varieties (From left to right: Dashhari, Alphonso, Totapuri, Payari, Langada)

By means of a selective threshold value, the background was totally removed and some preprocessing image operations such as *'imfill'* and *'erode'* were applied. The captured image from experimental set-up is in RGB color space which was converted to grayscale and resized to

512×512 pixels to simplify computations. Contour (outline of fruit) was extracted from this resized image as shown in Fig. 3.3 which was analysed later using various shape descriptors. Finally, these shape features were used to classify deformed Mangoes from well_formed ones.

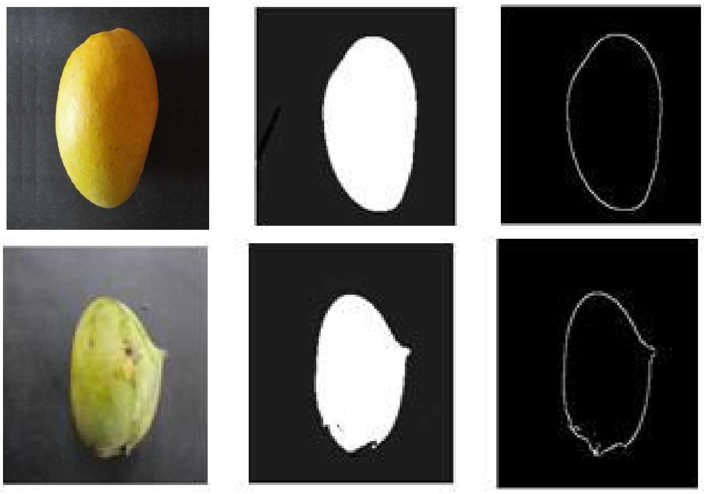

FIGURE 3.3: Contour Extraction of well formed (top row) and deformed shaped (bottom row) Mango fruit

3.4.2 Feature Extraction

Shape features were calculated using three different shape descriptors namely Fourier Descriptor (FD), Generic Fourier Descriptors (GFD) and Wavelet Fourier Descriptor (WFD). Centroid distance was used as a shape signature by all three shape descriptors. Initially, according to human experts, deformed images were sorted out from the database. A total of 20 boundary points from the contour image were considered to describe the fruit shape. All these descriptors were used as input in the ANN features which were input to the Artificial Neural Network (ANN) for grading purposes. Invariance of these descriptors regarding rotation, scale and translation was also tested by giving images of the same fruit with various scales and rotation.

3.4.3 Radial Basis Function Neural network

Radial basis function appeared as a variant of artificial neural networks in the late 80s. RBFs are rooted in three layered architecture (input, hidden and output layer) where each hidden layer implements a radial activated function as network neurons. The output unit is a weighted sum of hidden unit outputs as shown in Figure 3.4. RBF neural network converts non-linear input into linear output.

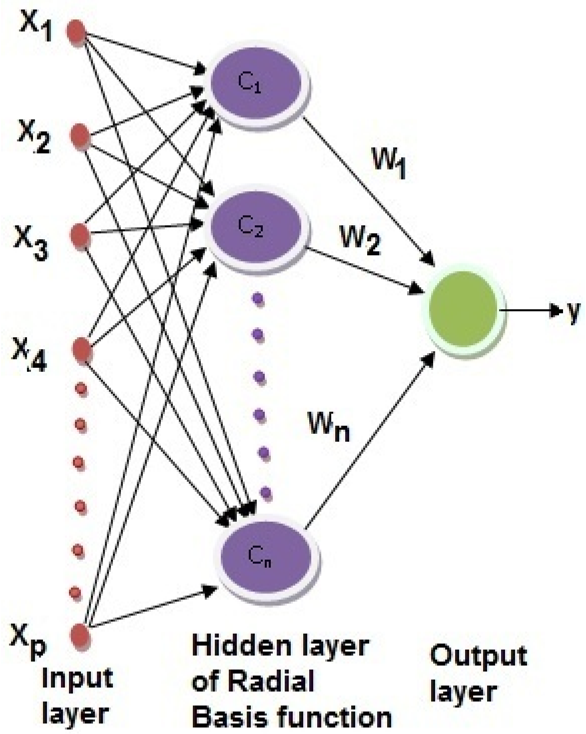

FIGURE 3.4: Radial Basis Function Neural network

Radial basis function has been employed for functional approximation in pattern recognition where inputs are feature vectors corresponding to each class. Various functions are used to activate the radial basis function which computes the Euclidean distance between the input feature vector and the center of that unit. In pattern classification, Gaussian function is preferred and thus output is limited to the interval (0,1) by a sigmoid function.

Modelling of a given mapping function is done through training of the neural network. Thus, weights and topology to model such a mapping has to be found. RBF Neural Networks are fast learners as they establish the local mapping. The radial basis neural network was designed by keeping the number of input nodes in the input layer equal to the input feature vectors. Length of shape descriptor was limited to 20 coefficients which was input to the neural network. The number of output nodes is equal to number of the classes in which data to be classified. Thus, two output nodes were selected to identify well-formed and deformed shapes. There were two tiers in the hidden network wherein the first tier, the number of neurons were equal to number of input features required to train network and second tier was consists of 2 neurons. The network parameters were adjusted, so that the network could be trained by approximating the underlying functional relationship between the training set (30% randomly selected data) and the target shape classes. To avoid over training, the network performance under training, was tested on a randomly selected unseen test dataset (20% validation set) to check how well the network models the input-target relationship. The remaining data (50%) was used as the testing set to measure the classification accuracy of the neural network on unseen data.

3.5 Results and Discussions

This section evaluates the proposed shape grading methods using a Mango database. The effectiveness of extracted shape descriptors was tested using two classifiers namely ANN and SVM. Experiments were carried out to find the following results:

1. Computation of ANN parameters
2. Selection of mother wavelet family for WFD descriptor.
3. Performance comparison of ANN and SVM classifier.

3.5.1 Artificial Neural Network Design

Initially, trial and error tests were carried out to determine the number of hidden neurons in order to build the neural classifier. From these trial tests, it was observed that increasing neurons in the hidden layer gave an improvement in classification accuracy; but at the same time it leads the network to ”memorize” these classification patterns instead of ”learning” it. This leads to decrease in accuracy for the test data set. In turn, after a typical number of hidden layer neurons, though we increase their number, there is only a minor improvement in classifiers accuracy. For a small number of hidden layer neurons, the network gives high training errors due to under fitting. On the other hand, a large number of hidden layer neurons produce high generalization error due to over fitting. Thus a compromise between these two is the optimization of neurons in a network which can be selected. The classifier architecture used in this study is shown in Fig. 3.5.

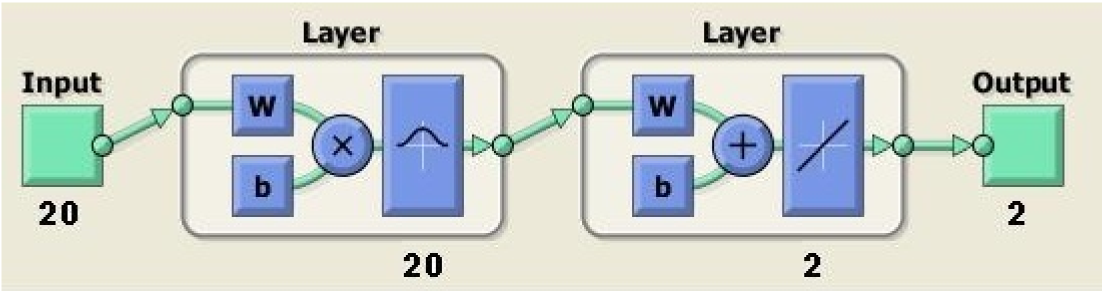

FIGURE 3.5: Radial Basis Function Neural network

The selected classification architecture is composed of 20 features as input and a hidden layer composed of 20 neurons, learning rate of 0.1 and 2 output neurons corresponding to well formed and deformed fruit shapes.

The network was trained for various numbers of epochs as well by changing the length of input vector (wavelet coefficients) for WFD and the observations are summarized in Table 3.1. Normally, when both validation and test datasets produce good and consistent classification results, it can be assumed that the network generalizes well on unseen data. It has been observed from table 3.1 that there is no significant improvement in the classification accuracy for twenty or higher number of WFD coefficients. However, there is a drastic reduction in the number of epochs required to train the ANN on the penalty of increased WFD coefficients.

3.5.2 Selection of the mother wavelet family for WFD descriptor

The Mango fruit shape grading is performed with designed ANN using WFD shape descriptor and experimentation is carried out to select efficient mother wavelet family.

Initially, the image was converted to grayscale using MATLAB command *'rgb2gray'* and was resized to 512×512 as discrete wavelet transforms demands an image of size 2^{k}(dyadic) for efficient analysis. Thus, each selected decomposition level for wavelet family yields horizontal,

TABLE 3.1: Impact of number of WFD coefficients on Training performance of Neural Network and Classification accuracy.

Number of WFD Coefficients	**Number of epochs required to train Radial Basis Neural Network.**	**Classification Accuracy(%)**
5	37	77.97
10	37	83.05
15	37	86.44
20	37	89.93
25	37	89.93
30	37	89.93
35	31	89.93
40	18	89.93
45	18	89.93
50	18	89.93

vertical and diagonal detail coefficients and approximated coefficients. Six common wavelet types Daubechies (db), Biorthogonal (bior), Reverse Biorthogonal (rbior), Coiflet (Coif), Discrete Meyer (Dmey) and Symlet (sym) were explored for their ability to describe the Mango shape. In this study, the feature vector for each mother wavelet is derived using the approximated value of decomposed wavelet coefficients. The proposed wavelet texture feature vector was derived using the MATLAB command *'wavedec'* and *'appcoef'*. For 512×512 image size, the image was decomposed at the first and second level by a wavelet transform and shape classification results using ANN are summarized in Table 3.2.

TABLE 3.2: Classification accuracy of various wavelet families for Mango shape sorting

Wavelet Families	**Classification Accuracy at scale 1 (%)**	**Classification Accuracy at scale 2(%)**
Debauches	85.45	81.82
Coiflet	79.09	81.82
Symlet	81.82	83.63
Discrete Meyer	89.93	81.82
Bi-orthogonal	85.45	80.91
Reverse Orthogonal	85.45	80.91

The analysis of Mango shape sorting using various wavelet mother families shows the following facts:

- Discrete Meyer function and Debauches show a high classification efficiency of 89.93% and 85.45% for discriminating deformed Mangoes respectively.

- Bi-orthogonal and reverse Bi-orthogonal function gives similar performance for classification.
- Thus, the discrete Meyer wavelet function shows that it is feasible to perform Mango shape sorting.

3.5.3 Performance comparison of ANN and SVM classifier

Finally, three different algorithms were developed for shape feature extraction using Fourier Descriptor (FD), Generic Fourier Descriptor (GFD) and Wavelet Fourier Descriptor (WFD) using discrete Meyer wavelet and classification results are summarized in Table 3.3. As seen

TABLE 3.3: Mango shape classification accuracy, using ANN classifier based on the FD, GFD and WFD techniques.

Shape Descriptor	**Classification accuracy(%)**
Fourier descriptor	85
Generic Fourier Descriptor	86.44
Wavelet Fourier Descriptor	89.93

in Table 3.3, FD gives lowest accuracy of 85% which is slightly improved, by GFD upto 86.44% and finally, WFD is superior among all three with an accuracy of 89.83%.

Experiments were carried out to compare the performance of ANN and SVM (polynomial kernel) classifier (in terms of classification accuracy and computation time) using WFD for Mango shape grading.

3.5.3.1 Classification accuracy

The results of the shape classification using both classifiers are presented in Table 3.4 as confusion matrices. The overall classification accuracy is given by equation 3.9,

$$Classifiaction\ Accuracy(\%) = \frac{m-n}{m} x100 \tag{3.9}$$

where m is the total number of classifies samples and n is the total number of misclassifed samples.

TABLE 3.4: Confusion matrices for Mango shape grading using WFD

Classifier Used	Actual Shape	Graded in		Total	Overall Accuracy
		Well Formed	Deformed		
ANN	Well Formed	169	21	190	89.92%
	Deformed	10	100	110	
SVM	Well Formed	165	25	190	86.46%
	Deformed	15	95	110	

The overall correct classification performance of the testing set of Mangoes was 86.36% and 89.93% using SVM and ANN respectively. Leave one out cross validation was opted by both the classifiers. It is observed that, there is no significant difference between classification accuracy for well formed and deformed Mangoes using SVM. However, it can be noted that the generalization capability of the neural network can be deemed satisfactory. This deduction is based on the experimental results of the Radial Basis function-based neural network, thus it cannot be generalized to other classification techniques.

The results here demonstrate that the neural network classifier performs well on the validation dataset. The misclassification rate is higher in the well formed test set, but the overall accuracy is 89.9% which shows that the extracted WFD feature bear the distinction between different Mango shapes and can be used for classification.

3.5.3.2 Computational time

The computational time for feature extraction (from the input image to feature vector) and classification (from feature vector to classifiers output) using ANN is given in Table 3.5. All the algorithms are implemented in MATLAB 7.11.0 Version R (2010b) on a PC with 2-GB RAM and a 2.8 GHz processor. The computational time for the proposed WFD shape descriptor is less than GFD as GFD considers all pixels within the shape to describe it; but more than FD because multi-resolution capability of wavelet lowers the speed on account of better accuracy.

3.6 Concluding Remarks

In this work, a Mango shape grading algorithm based on computer vision has been proposed. Three different shape descriptors which are derived from Fourier transform and wavelet transform has been tested using two statistical classifiers. The chosen classifiers were Artificial

TABLE 3.5: Computational time complexity of different Mango shape descriptors.

Shape Descriptor	**Features Extraction Time (s)**	**Classification Time (s)**	**Total Time (s)**
Fourier descriptor	0.070533	0.033898	0.104431
Generic Fourier Descriptor	6.450573	0.029566	6.480139
Wavelet Fourier Descriptor	0.246070	0.026594	0.272664

Neural Network and Support Vector Machine, and several kernels and parameters were evaluated. The best shape classification result was achieved using WFD with discrete Meyer mother wavelet and radial basis function ANN classifier. The results also show the importance of selecting a good mother wavelet family as well as number of WFD coefficient that is the length of feature vector while shape grading. The proposed system simplifies the first important task in fruit grading while packaging with an accuracy of 89.93% on the account of more computational complexity as compared to FD.