A Literature Review on Image Processing and Classification Techniques for Agriculture Produce and Modeling of Quality Assessment system for Soybean industry Sample

Mr. Sachin Sonawane1*, Dr. Mohan Awasthy2, Dr. Nitin Choubey3

1Faculty at NMIMS, Mukesh Patel School of Technology Management and Engineering, Shirpur.
2Faculty at G. H. Raisoni Institute of Engineering and Technology, Nagpur.
3Faculty at NMIMS, Mukesh Patel School of Technology Management and Engineering, Shirpur.

*Corresponding Author: Mr. Sachin Sonawane, Faculty at NMIMS, Mukesh Patel School of Technology Management and Engineering, Shirpur.

Abstract: Soybean, the most popular golden bean of America, is widely known for its fat free food products. Richness in Protein makes it one of the best suggested meals which can be consumed in the form of pulses, oil, food for animals etc. The quality of such products is mainly dependent on the quality of Soybean procured as fresh farm produce. Governments and regional authorities have already defined the standards for quality assessment and grading of Soybean, which are meant to be followed in the commercial market while trading. Presently, visual inspection is the preferred way to conduct physical quality assessment of Soybean and it is performed by an expert person, at the time of procurement of Soybean based on the standards recommended by the buying authority. Physical parameters of Soybean kernel like color, growth corresponding to size, damage, fungi/ disease as well as mixing of other material/ objects in soybean sample influence the grade of that sample during quality assessment. Dependency of this process on human expert usually harms the accuracy in assessment. Therefore, an automated machine is desired to be a suitable solution to address this issue and can benefit in terms of increased accuracy, reliability, and reduced response time. Worldwide researchers are working on designing such automated systems for different type of fruits, grains etc. However, in case of Soybean, up till now we are successful in cleaning, sorting, color detection, and also, through various image processing techniques and algorithms researchers could detect the anomalies present in grain sample. But an automated system for the quality assessment and grading of Soybean according to an International Standard is yet to be implemented. In this paper, we propose a two-stage model for the quality assessment of Soybean; first stage focuses on image processing techniques like Image acquisition, preprocessing and feature extraction Likewise the second stage works on classification of Soybean kernels and sample grading with the help of a machine learning technique on the basis of an International Standard.

Keywords: Quality Assessment, Visual Inspection, International Standard, Automated System.

1. INTRODUCTION

Soybean is the most admired and favored food in the world. It is readily available in all the countries. Because of richness in protein, it is recommended and consumed in the daily food. Variety of value-added edible products can be prepared from Soybean including Soya-Milk, Oil, Pulses etc. for human as well as animals. The countries having major share in Soybean production are United States, Brazil, Argentina, China and India [1]. In international food market, the quality of Soybean is a major concerning order to assure the production of best tertiary product by food industry. To ease the trading process the countries like US, Canada have recommended certain criteria and standards to buyers for grading Soybean quality. Thus, the measurement of Soybean quality is essential as well as an equally important requirement in today’s scenario to protect the buyers from substandard produce [2]. Accordingly, if the Soybean quality assessment is done manually then the chance of error in measurement is significant and in case it is done with the help of a perfectly trained automatic machine then the chance of error would be minimum. On these aspects, this research work is directed towards the modeling of an automated system for the quality measurements of Soybean Sample [3].

In this research work, authors propose a novel two stage model to addresses the issues related to quality assessment of Soybean. A two-stage model comprised of; first stage to detect the physical
parameters like dimensions, area and color of Soybean by using an image processing technique and
the second stage to classify and grade Soybean sample with the help of a machine learning technique.

2. RELATED WORK

2.1. Image Enhancement Techniques and Feature Extraction

Image processing techniques involve all the techniques that are used to enhance graphic appearance of
given image or extract specific features from an image for better interpretation by human or
automated system. It is hard to develop the generalized theory of image enhancement, as the standard
for quality of an image can’t be generalized.

2.1.1. Image Smoothing

Image smoothing operation is generally performed on noisy images. Noise in an image is the
undesired variation in the pixel values introduced while image acquisition, transformation and
communication process. The image pixels, affected by noise, look noticeably different from their
neighborhood pixels. Noise can’t be removed completely but a smoothing process can reduce it
significantly. The basic method of noise reduction is to consider few nearest neighborhood pixels of
each pixel in the image analyze them and apply the suitable noise filter. Often averaging or median
filter is applied for image smoothing. In averaging filter, every pixel component is substituted by the
average value of its neighborhood pixels. Likewise, in case of median filtering, every pixel
component is substituted by the median value of its neighborhood pixels. There is another class of
image smoothing filters known as spatial filter and its basic principle is low pass filtering. Spatial
filters conduct a discrete convolution of the input image with designed mask [4]. literature describe
that the Gaussian smoothing is very effective for removing Gaussian noise. The degree of smoothing
is controlled by adjusting the factor $\sigma$. However, Gaussian filtering blurs images and reduces fine
details while removing noise. Hence, on the original images where fine details and high frequency
information can’t be loosed, the Gaussian filtering is a least recommended technique. Instead of this,
the best edge preserving image smoothing technique can be applied [5]. Graham Treece (2016) [6],
investigated on image smoothing techniques and proposed a new filter; The Bitonic Filter for linear
filtering within a morphological framework intended for preserving edges. Author found this filter to
be better in edge and detail preserving properties than a median, similar to a Gaussian in noise
reduction capability, and applicable to many signal and noise types.

2.1.2. Image Sharpening

The contrast between adjacent colors is considered as Sharpness. If there is a quick changeover from
black to white then it appears sharp and on the contrary if there is a gradual change in color from
black to gray and then gray to white then it appears blurry. Sharpening image means increasing the
contrast along the edges where different colors meet. Image sharpening improves the high frequency
component in an image. High frequency components, like edges of objects in an image are desired to
be sharp to enhance the visual appearance of an image. Image sharpening adds high pass filtered
values of the original image to the original image resulting in improved edges, fine details in an image
and amplification of high frequency noise component [7]. Image enhancement techniques are widely
applied in printing and photographic industries. Chitra P et al (2017) [8] worked on Sharpening of
Edges in Radiographic Images Using Edge Preserving Filter. Authors have used two edge preserving
filters namely modified Local Edge Preserving Filter and Adaptive Weighted Mean Filter and claimed
to have obtained better results. Kuo-Shiuan Peng et al (2013) [9], proposed an edge directed unsharp
masking sharpening method aiming for effective super resolution. Authors have specifically
mentioned the advantages and disadvantages of two image sharpening techniques namely
conventional unsharp masking sharpening and edge directed super resolution methods and tried to
combine these techniques for achieving better performance.

2.1.3. Shadow Removal

Many a times, in the imaging of Soybean seed samples, shadows of the seeds can be seen, due to
insufficient light. Such shadows are not useful and distort the visibility, while counting the number of
seeds present in an image as well as detecting the edges or boundaries of seeds in an image. Hence, in
all such cases, the shadow effect is treated as a noise and the removal of shadow noise becomes essential. The treatment of shadow effect means nothing but the brightening or restoration of luminance of the shadowed areas in an image [10].

The very basic step in process of removal of shadow effect is to detect the shadow and all the methods of shadow detection can be broadly categorized into two groups; one is model based and the other is feature based. Out of these two, the feature-based methods are used most widely. In shadow detection and removal which is necessarily based on extracted features, by using a threshold value, the shadow region present in an image can be detected. Till date researchers have proposed various feature-based methods, which mainly have the basis of shadow detection index or features obtained from near infrared channels or multispectral view etc [11]. Y. I. Shedlovksa et.al (2016) [12], published their research paper which provides a simple and enhanced technique to detect as well as remove shadow in satellite images which can be useful for the same operation on Soybean seed sample images. In this method, the input image is transformed from RGB (Red, Green, Blue) to HSV (Hue, Saturation, Value) color scheme; the normalized saturation-value difference index is calculated to determine the shadowed region in an image as it can indicate higher values for the image components in shadowed region as compared to image components in non-shadowed regions. Manoj K Sabnis et al (2013) [13], published their research work on detection of shadows in static images by new thresholding method. The shadow of an object of interest having shape identical to that of the object is a major hurdle in tracking and identifying the object. As RGB can’t perform well for outdoor images, the input image is converted to HSV color scheme. By using threshold values, binary image is obtained from HSV image, which acts as a mask for shadow detection and can be applied on input color image. Arti Tiwari (2016) [14], conducted a survey on detection and removal shadows present in images as well as sequential video frames. Authors analyzed 10 methods on image processing and 10 methods on video processing. However, the issues related with all these techniques have not been stated specifically with respect to any of the techniques. Common statements on all those techniques have been passed in both the aspects like failure and success of a technique.

2.1.4. Color, Geometry and Shape

Color, geometry and shape are the physical parameters of Soybean mainly comprised of length of major axis, length of minor axis, perimeter, area, and roundness of Soybean Kernel. Abubakar Yamin et al (2017) [15], worked on the color images for detection and classification of blood group. Authors applied RGB to HSV conversion, morphological operations along with other image processing techniques like histogram derivative, segmentation then binary image is used to measure the density of pixels in each object and based on observations, were able to classify the blood sample according to blood group. Choubeila Maaoei et.al (2005) [16], attempted to design a robust color object detection and recognition system. Their work is divided into two stages, first stage concerned to the color segmentation and then image conversion into binary for the detection of objects and background. The Connected Component Labelling algorithm is used to distinguish among various objects present in an image. Second stage is concerned to the recognition of extracted objects and utilizes the shape descriptor based on Zernike moments for this purpose. The support vector machine classifies objects. M.J.C.S. Reis et al (2012) [17] conducted research on automated color based grape bunch detection system for outdoor natural location. In the development phase, authors determined the possible and valid RGB values of pixels of Red grapes as well as white grapes independently, during application phase, the designed system searched for the pixel values falling within the bounds of those central values respectively. Zalhan, Mohd Zin et al (2016) [18], could successfully design a simple vision-based egg grade classifier. Authors could accurately measure diameters of eggs of various sizes with the help of a Coordinate Measuring Machine (CMM), which is widely used in industry to measure the physical geometry of an object. To validate the same, authors used image processing technique in which the RGB image of an egg is converted into grayscale image and then by using thresholding technique, it is converted to binary image. In binary image, sum of all the pixels correspond to area of an egg, the radius of an egg can be easily calculated and used to classify eggs.
Surya Prabha et al (2013) [19], conducted a research for the banana fruit maturity assessment through image processing and the analysis of variance between all maturity stages based on features, length of major axis, length of minor axis, perimeter, area. The obtained results indicate that the mean color intensity and area features are key parameters in predicting the banana maturity. To solve the purpose authors converted color image into a binary image and then used it as a mask on the original color image for calculating color intensity and area of banana.

2.2. Utility of Neural Network for Classification

Since last few years, researchers are investigating on the utility of neural network while developing an automated machine vision for the quality assessment and grading of agricultural produce. To inculcate the abilities of the neural network and humans is in a single machine, which can perform in a best possible way for a particular application is the main aim of this study. So far, researchers are successful in designing and implementing various types of neural networks to classify input entities by achieving an accuracy of 80–100 % and significant improvement in minimizing response time. Somya Sharma et al (2014) [20], proposed a neural network designed for detection of agricultural activities. Activities like bed-making, plantation, and harvesting, moving and steady position are detected by using mobile accelerometer sensors and fed to the forward propagation and backward propagation neural network. Authors mention that the input data is filtered using low pass filter and input features are fed. However, authors did not describe those 27 input features based on which the improved accuracy with respect to extended number of hidden layers has been claimed. Qing Yao et al (2009) [21], attempted to apply Support Vector Machine (SVM) for rice diseases detection on the basis of geometry and color features. With the help of image segmentation, technique disease spots of rice were detected and based on extracted features SVM is trained to classify diseases into rice blast, rice sheath blight and rice bacterial leaf blight with an accuracy of 97.2%. Panitrat Yimyam et al (2012) [22], proposed an approach to generate computer vision systems for specific tasks from generic components for grading of agriculture produce by machine learning using genetic programming. This approach worked on issues like surface inspection & maturity assessment of Mango, discrimination of apple variety, classification of barley and wheat, and grading of purple gluey rice. It’s claimed that the neural network classifies with more accuracy than SVM and genetic programming when the input features like shape, color and texture are used. But the detailed design of neural network is not provided. Nur Badariah Ahmad Mustafa et al (2011) [23], proposed a novel approach to develop an intelligent, fast and effective fruit sorting system by utilizing digital image processing techniques as well as probabilistic neural network. Five fruit samples i.e. banana, apple, mango, oranges and carrot were analyzed based on seventeen morphological and color features. Authors mentioned that shapes and sizes can significantly differentiate among different agricultural produce and color helps in identifying the object. Authors could perform sorting accurately because of combination of shape and color features as well as with increased speed of operation due to easy interpretation of output of probabilistic neural network. Ksh. Robert Singh et al (2016) [24], described a technique for classification of 4 types of rice grain using back propagation neural network (BPNN) and wavelet decomposition. System is trained to perform classification based on 27 textures, 24 wavelet, 18 color features, 45 combined features of color and texture. Input signal included 3 different data set of images captured under dissimilar environmental conditions. Authors found that BPNN successfully classifies input rice varieties even for low quality input images. In addition, the classification on the basis of wavelet features outperforms as compared to classification using other feature slower solution images. The authors could further investigate that proposed BPNN achieves better accuracy than support vector machine and naive Bays, k-nearest neighbor classifier for given input images. But the details about design of BPNN have not been published by authors.

3. MATERIAL AND PROPOSED METHODOLOGY

3.1. Soybean Industry Sample

Soybean industry sample is the collection of few samples of fresh Soybean approximately of weight 1000 – 1050 grams (FGIS) taken from the big lot of Soybean which farmers bring to the industry for commercial sale. This sample represents whole lot, which usually includes foreign material, splits, damaged kernels, discolored kernels etc. Therefore, it becomes necessary to perform physical quality assessment and grading of Soybean before it is processed. If a lot carries infested Soybean kernels...
then it is rejected because the disease may spread out during storage and major lot of Soybean get infected from it. It also degrades the quality of tertiary product of Soybean. In proposed two – stage model, a rough sample of Soybean will be abstracted and weighed to prepare the Soybean Industry sample [25].

![Flowchart Diagram](image.png)

**Fig1. Work flow Diagram of Proposed Two – Stage Model for Soybean Quality Assessment and Grading**

### 3.2. Digital Imaging of Soybean Industry Sample

The Soybean industry sample will be exposed to high resolution Digital Camera in order to obtain high quality image. It is essential to keep sufficient illumination for capturing actual shade, color of Soybean kernels and to avoid appearance of shadow. The controlling of illumination can be done by considering suitable light source and light filter if required. Tuning of appropriate distance between digital camera and sample after spreading sample elements on black background will help in capturing high quality image [26].

### 3.3. Pre – Processing on Sample Image

The captured digital image can be accessed on Desktop Computer and with the help of Mat lab software, pre – processing operations such as image smoothing and sharpening of edges can be performed on it. As morphological information is carried by the edges of objects present in sample image, the application an efficient edge preserving image smoothing algorithm can provide best results. Subsequently, sharpening of preserved edges increases the accuracy of edge detection algorithm. As per the previous Canny edge is the best edge detection algorithm for grain images.

### 3.4. Feature Extraction

In many literatures, it is described that feature extraction is one of the ways to reduce dimensionality. In image processing features are extracted so that the part of an image to be processed is represented...
A Literature Review on Image Processing and Classification Techniques for Agriculture Produce and Modeling of Quality Assessment system for Soybean industry Sample

This approach is useful to quickly complete tasks when image sizes are large. In proposed method, two types of features namely morphological and color are considered for Soybean kernel classification.

3.4.1. Morphological Features

These are the features associate with form, shape, or structure of object in an image. When captured input image is converted into a binary form, the total number of pixels covered by individual Soybean kernel roughly represent its shape which mostly appears to be elliptical. From this recognized geometrical shape of a Soybean kernel, the morphological features contributing effectively in classification of Soybean kernels can be computed as [27] [28]:

![Fig2. Ellipse](attachment:image)

If ‘a’ and ‘b’ are measured from the center then

**Length of Major Axis**: Longest line passing through a kernel represents its major axis length.

Length of Major Axis = 2a

**Length of Minor Axis**: The longest line passing through a kernel and which is perpendicular to the major axis length, represents its minor axis length.

Length of Minor Axis = 2b

**Area**: By using below mentioned formula, area of a Soybean kernel can be calculated.

Area = $\pi ab$

**Perimeter**: The continuous line forming the boundary of a closed geometrical shape is called its perimeter.

$$h = \frac{(a-b)^2}{(a+b)^2}$$

**Roundness**: Area and Perimeter of an object are used to find its roundness.

Roundness = $\frac{4\pi \times Area}{Perimeter^2}$

3.4.2. Color Features

The color features of sample image can be extracted from its histogram in designated color scheme. The captured image follows RGB color scheme. Hence, it’s all three planes i.e. Red, Green and Blue will be analyzed individually in order to compute the parameters like mean, standard deviation. So that the value for most dominant color in each image can be obtained [23] [27].

**Mean** = $\frac{1}{n} \sum_{i=1}^{n} P_i$

**Standard Deviation** = $\sqrt{\frac{1}{n} \sum_{i=1}^{n} (P_i - Mean)^2}$

Where, $n$ = Number of Pixels, $P_i$ = Value of $i^{th}$ Pixel

To validate the results obtained with RGB image, it can be transformed into the HSI(Hue, Saturation, Intensity) image, which is more perceptual color model. In HSI, Hue indicates colors, Saturation shows shades of gray in the color space and Intensity is the brightness of color. Once again, the similar analysis is to be performed on three images of hue, saturation and intensity.
3.5. Rescaling of Extracted Features (Normalization)

In case, the range of values of features vary widely, machine learning algorithms might not function properly then feature scaling method can be used to normalize the range of independent features of an image. Following equation is used while normalizing data features [23];

\[
\text{Normalized value} = \frac{x - \min(x)}{\max(x) - \min(x)}
\]  

(9)

**Classification of Soybean Kernels**

The morphological and color features extracted in previous stage act as an input for the classifier. There are two important aspects of classification in proposed method; one is the maximum accuracy in classification and second is minimum response time. In many literatures, published so far, it can be checked that the SVM (Support Vector Machine) is one of the best classifiers, which effectively deals with the problem of over fitting and can produce accurate results. Also, the less complexity in designing of SVM contributes in minimizing its response time. Another classifier which can quickly respond for the given input is Probabilistic Neural Network (PNN). PNN generates probabilistic output which can be easily interpreted, and hence the training time required is also short. Being a supervised neural network, PNN can act as an alternative to other supervised neural networks like back-propagation neural network [23] [28].

3.6. Grading of Soybean Sample

Based on the results obtained from classification of Soybean kernels, the percentage of other material mixed in Soybean sample can be calculated and it can be graded by referring the standards recommended by regional authorities. Following table represents one of such standards defined by FGIS for grading of Soybean in US market.

**Table1. Official U.S. Grades for Soybeans (FGIS, 1994)**

<table>
<thead>
<tr>
<th>Grading Factors</th>
<th>Grades U. S. Nos.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td><strong>Maximum percent limits of</strong></td>
<td></td>
</tr>
<tr>
<td>Damaged Kernels:</td>
<td></td>
</tr>
<tr>
<td>Heat (part of total)</td>
<td>0.2</td>
</tr>
<tr>
<td>Total</td>
<td>2.0</td>
</tr>
<tr>
<td>Foreign material</td>
<td>1.0</td>
</tr>
<tr>
<td>Splits</td>
<td>10.0</td>
</tr>
<tr>
<td>Soybeans of other colors(^I)</td>
<td>1.0</td>
</tr>
<tr>
<td><strong>Maximum count limits of</strong></td>
<td></td>
</tr>
<tr>
<td>Other Materials:</td>
<td></td>
</tr>
<tr>
<td>Animal filth</td>
<td>9</td>
</tr>
<tr>
<td>Castor beans</td>
<td>1</td>
</tr>
<tr>
<td>Crotalaria seeds</td>
<td>2</td>
</tr>
<tr>
<td>Glass</td>
<td>0</td>
</tr>
<tr>
<td>Stones(^2)</td>
<td>3</td>
</tr>
<tr>
<td>Unknown Foreign Substance</td>
<td>3</td>
</tr>
<tr>
<td><strong>Total(^3)</strong></td>
<td>10</td>
</tr>
</tbody>
</table>

U.S. sample grade are soybeans that:
(a) Do not meet the requirements for U.S. Nos. 1, 2, 3, or 4; or
(b) Have a musty, sour or commercially objectionable foreign odor (except garlic odor); or
(c) Are heating or otherwise of distinctly low quality.

\(^I\) Disregard for Mixed soybeans.
\(^2\) In addition to the maximum count limit, stones must not exceed 0.1 percent of the sample weight.
\(^3\) Includes any combination of animal filth, castor beans, crotalaria seeds, glass, stones and unknown foreign substances. The weight of stones is not applicable for total other material.

4. DISCUSSION AND CONCLUSION

Soybean is a globally consumed food and it is available in the form of variety of tertiary products. Food industries manufacturing or processing Soybean are more concerned about the procurement of clean, uncontaminated and good quality Soybean. To fulfill this requirement, in commercial market, while farmers bring fresh farm produced Soybean for trading, it is graded on the basis of certain
standards defined for physical quality assessment of Soybean by regional authority or government. Few of such standards can be the US standard, and Canadian standard. However, the grading processes inferior in terms of accuracy and reliability because even today, it is conducted manually. The quality grade assigned to a Soybean lot is highly dependent on the expertise of a person assessing it and by following his/her remarks; the selling rate for whole Soybean lot is decided. In this scenario, it is highly possible that a particular lot is not fixed to its true selling rate. So, the desired solution to address this long prevailing issue is to use an automated machine for quality assessment and grading of Soybean which entirely follow the government standards. As pointed in this literature review, presently we are able to only sort, count Soybean kernels as well as detect their color and geometry. Recently, the researchers have directed their focus towards the development of image processing techniques and algorithms to detect the anomalies present in grain sample. But an automated system for the quality assessment and grading of Soybean, just like many other grains and fruits is yet to be implemented. In this paper, we propose a methodical two-stage model for the designing of an automated system working on image processing techniques and machine learning which is enabled to perform effectively by its methodical approach. This model will be a great facility to food industry and farmers.

REFERENCES

A Literature Review on Image Processing and Classification Techniques for Agriculture Produce and Modeling of Quality Assessment system for Soybean industry Sample


[25] Thomas J. Brumm, Ph.D., Iowa State University, Chaper 2, Quality and Grading Factors of IP Soybeans.


Copyright: © 2018 Authors. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.