**Chapter II**

**REVIEW OF RELATED LITERATURE**

**Rice Economy**

Rice is one of the most cultivated source of carbohydrate and calorie requirements. In 2014, the global production of rice reached 490 million metric ton with 402 million of it being used as food and the remaining as feed and other purposes (FAO, 2014). In the Philippines, it is estimated that the national production reached 7.6 million metric ton in first half of 2015 with the 23.59% of the total production coming from the Central Luzon region. The average yield for the same region was 5.64 metric tons per hectare. It is also forecasted that there will be a total of 12.27% increase in national production in the first half of 2017. (Philippine Statistics Authority, 2016). Rice farming is a big contributor to the development and progress of the Philippine economy.

**Philippine Grains Standardization Program**

The rice market in the Philippines is one of its biggest. The reason for this is that Filipinos are one of the nationalities with rice as its staple food. Though rice is relatively bland in taste, its consumers could differentiate quality among varieties of rice. Quality can be assessed through physical, chemical, and market preferences. The chemical methods of assessing the quality of rice employ the analysis of the percentage composition and moisture content. Though this method is highly selective and accurate, its cost is not feasible for frequent evaluation. Sensory evaluation through tasting is highly subjective to the tester’s ability to differentiate and ‘tasting’ skills. The physical method of evaluation the quality of rice is deemed to balance the trade-off between economic feasibility and precision.

The Philippine Grains Standardization Program of 2002 is a government program spearheaded by the National Food Authority to integrate recommended industrial and commercial assessment that will provide inclusive growth, uniformity, compliance, and food quality and safety standards for the labelling and quality assessment of corn and rice grains produced in the Philippines. The National Grains Standard provides the standard specifications on the quality assessment, labelling, and recommended packaging for corn and rice products. The significance of providing quality assessment specifications is mainly to classify rice products so that the appropriate prices are set fairly and justifiably based on the superiority of the products. The NGS provides grading criteria to classify the rice product into Premium or any from Grade 1 to Grade 5. The specifications for milled rice grading are provided in Table 2.1.

**Table 2.1** The National Grains Standard

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **PARAMETER** | **GRADE** | | | | | |
| **PREMIUM** | **GRADE 1** | **GRADE 2** | **GRADE 3** | **GRADE 4** | **GRADE 5** |
| Grain Size | Very Long, Long, Medium, Short | | | | | |
| Degree of Milling | Over milled, Well milled | Well milled | Regular milled | | | |
| **GRADE FACTORS**  **(% by weight)** | **GRADE** | | | | | |
| **PREMIUM** | **GRADE 1** | **GRADE 2** | **GRADE 3** | **GRADE 4** | **GRADE 5** |
| Brokens, max. (total including brewers) | 5.00 | 10.00 | 15.00 | 25.00 | 35.00 | 45.00 |
| Brewers, max. | 0.10 | 0.20 | 0.40 | 0.60 | 1.00 | 2.00 |
| **Defectives:** | | | | | | |
| Damaged kernel, max. | 0.50 | 0.70 | 1.00 | 1.50 | 2.00 | 3.00 |
| Discolored kernel, max. | 0.50 | 0.70 | 1.00 | 3.00 | 5.00 | 8.00 |
| Chalky kernel, max. | 4.00 | 5.00 | 7.00 | 7.00 | 10.00 | 15.00 |
| Immature kernel, max. | 0.20 | 0.30 | 0.50 | 2.00 | 2.00 | 2.00 |
| Contrasting type, max. | 3.00 | 5.00 | 10.00 | - | - | - |
| Red kernel, max. | 1.00 | 2.00 | 4.00 | 5.00 | 5.00 | 7.00 |
| Foreign matters, max. | 0.025 | 0.10 | 0.15 | 0.17 | 0.20 | 0.25 |
| Paddy, max. (no. per 1000 grams) | 10.00 | 15.09 | 20.00 | 25.00 | 25.00 | 25.00 |
| Moisture content | 14.00 | | | | | |
| Milling degree | OMR, WMR | WMR | RMR, WMR(Super),  UMR(Ordinary) | | | |

**Definitions and classification of the characteristics**

The National Grains Standard defined the factors and parameters of the specifications. The following definitions are directly referenced from the NGS.

*Grain Size*

The grain size of a particular sample is the average of the individual sizes of the grain’s measured major axis length. With the specification of the National Grains Standards, only the major axis length of the grain is measured with disregard to the minor axis length. The size classifications are defined in Table 2.2.

**Table 2.2** The National Grain Standards Grain Size Classification

|  |  |
| --- | --- |
| **Grain Size** | **Description** |
| Very Long | Rice with 80% or more of whole milled rice kernels having a length of 7.5mm and above. |
| Long | Rice with 80% or more of whole milled rice kernels having a length of 6.4 to 7.4mm. |
| Medium | Rice with 80% or more of while milled rice kernels having a length of 5.5 to 6.3mm. |
| Short | Rice with 80% or more of the whole milled rice having a length of less than 5.5mm. |

*Degree of Milling*

The rice seed is coated with plant material called bran. The degree of milling is defined as the extent of how much bran layers and germ have been removed in the milled rice. The classifications of the degree of milling are defined in Table 2.3.

**Table 2.3** The National Grains Standard Degree of Milling Classification

|  |  |
| --- | --- |
| **Degree of Milling** | **Description** |
| Regular milled | Rice kernel from which the hull, the germ, the outer bran layers and the greater part of the inner bran layers have been removed but parts of the lengthwise streaks of the bran layers shall be within the range of 20-40% of the kernels. |
| Well milled | Rice kernels from which the hull, the germ, the outer bran layers and the greater part of the inner bran layers have been removed, but parts of the lengthwise streaks of the layers shall be less than 20% of the kernels. |
| Over milled | Rice kernel from which the hull, the germ and the bran layers have been completely removed. |

*Broken Kernels*

The broken kernels are described as the pieces of kernels smaller than 75% of the average length of the unbroken kernel.

*Brewers*

The brewers are grain samples that can pass through sieves having round perforations of 1.4 mm in diameter.

*Damaged Kernels*

The damaged kernels are those that are sprouted or distinctly damaged by insects, water, fungi, and/or any other means.

*Discolored Kernel*

The discolored kernels are kernels that have changed their original color as a result of heating and other means. They are also known as ‘yellow kernels’ or ‘fermented kernels’.

*Chalky Kernel*

The chalky kernels are those, whole or broken, one-half or more of which is white like the color of white chalk and is brittle upon removal of the hull for palay.

*Immature Kernel*

The immature kernels are those, whole or broken, which are light green and chalky with soft texture.

*Contrasting Type*

Palay/rice kernels of different varieties other than the variety designated, wherein the size, shape, and color differ distinctly from the characteristics of kernels of the variety designated.

*Red Kernel*

The red kernels are those that have red bran covering, wholly or partly.

*Foreign Matter*

Organic and inorganic components other than whole or broken rice kernels (e.g. foreign seeds, husks, bran, sand, dust, and other crop seeds).

*Paddy*

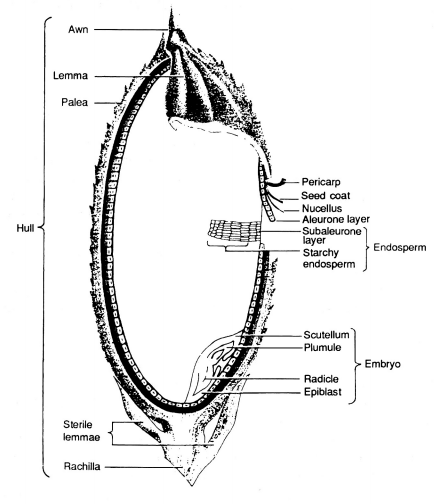
Paddy is the cut part of the rice plant other than the seeds.

*Moisture Content*

The moisture content is the water content of palay, milled rice and corn, expressed in percent (%) as received.

**Morphological Indicators**

The NGS is based on the physical characteristics and morphology average quantification of the rice grains. The characteristic set of each sample could indicate and be classified into grades.



**Figure 2.1** Longitudinal section of a rice grain (FAO, 2014)

The rice grain is the seed of a rice plant that is sexually mature. The color of a rice grain begins from being light green to progressively yellow to golden. This change in morphology could indicate the ripeness of the grain for harvest. Milled rice is rice grain where its bran is removed to an extent. The NGS defined the extent classification of milling. In Figure 2.1, the hull is the outer layer of the grain which defines the color of the grain. As the hull is removed, the bran will be left behind. The bran is a layer in the grain structure which lines the starchy endosperm of the grain.

**Grain size, broken and brewer grains**

The white part of a hulled rice grain is the starchy endosperm cells. These cells are composed of starch structures that provide sustenance for the embryo of the rice grain. The grain size is determined by the major axis length of the endosperm. Although the NGS does not provide the length to width ratio specification, it classifies rice grain size based on the major axis length. This means that the width or length of the minor axis could be disregarded for classification parameter. The grain size is significant to assess the volumetric property of the rice product.

Broken rice grains degrade the quality of the overall product. The breaking of the rice grains is caused by milling procedures or the general property and chemical quality of the rice grains. Broke rice grains are originally whole rice grains but have lengths that are less than 75% of the average length. Market preference indicates that the lesser the amount of broken grains per whole grains means the quality of the grain products is higher (Dalen, 2004). Brewers are similar to the broken grains. However, brewers are rounded. The brewer grains also decrease the perceived quality of the product. The major axis length of the rice grain is used to indicate if the sample grain is a broken, brewer, or a regular desired rice grain.

Image processing dimensional measurements are often used in the grading. In a study made in 2004, the dimensional length and width of the rice grains are used to determine if it is broken (Dalen, 2004). Broken kernels have length less than 75% of the average length. The author used image processing tools to segment, binarize, and measure the dimensions.



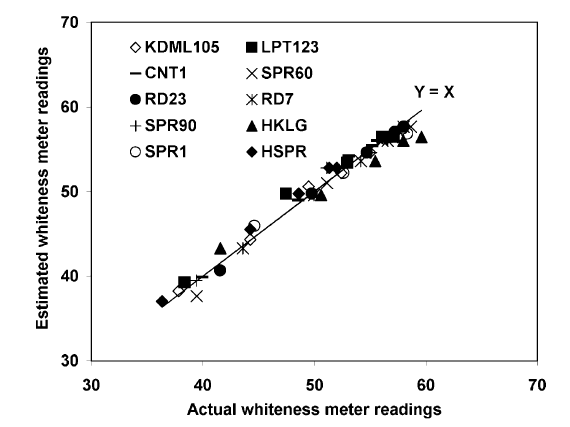
**Figure 2.2** Individual connected component analysis of grain threshold image with count labels (Bambole, et al., 2015)

The classification of grain size depends on the average length of the sample. The individual size must be compared to the mean length to provide a statistically significant classification of the rice. Using connected-component algorithm, threshold rice grains can be analyzed individually by labelling them. In this way, the total count of the rice grain sample can also be obtained (Bambole, et al., 2015).

**Degree of milling**

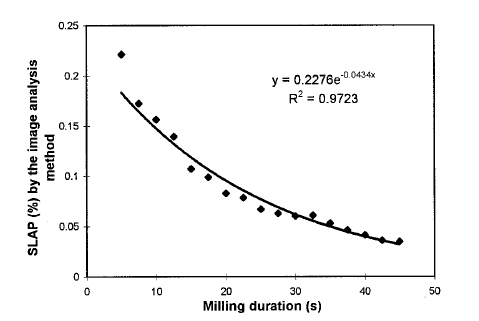
Milling is the process of removing the husk, bran, and the embryo of the rice grain. The main purpose of milling to a degree is to manage the starch to protein content of the resulting grain. As mentioned, the starch resides in the endosperm. However, proteins and lipids are found in the bran which is removed in the milling process. High degree of milling means the starch to protein content ratio is higher (Paiva, et al., 2014).

A study indicated that the degree of milling and the whiteness of the rice grains are related with each other (Yadav, et al., 2001). In this study, using image processing techniques, the head rice yield and the degree of milling are estimated. The head rice yield is the ratio of the weight of the milled rice grains to the total amount of unhusked rice grains. This ratio aims to provide a general quantification of the extent of the removal of the husks and other internal coverings. The whiteness of the rice is said to be proportional to the lipid concentration in the rice grain. High lipid concentration means the rice grain is whiter. As the degree of milling increases, the amount of whiteness in the grain intensifies. The whiteness is defined by the overall white level of the sample and not by region only. Therefore, a gray level distribution mean was used to indicate the overall whiteness. Increasing degree of milling corresponds to increase in mean gray level.



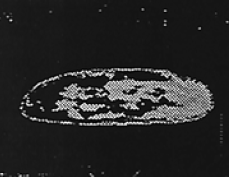
**Figure 2.3** Comparison between whiteness estimation using mean gray levels and the actual whiteness (Yadav, et al., 2001)

In another study, the surface lipid concentration on the grain was used to estimate the degree of milling and found that higher degree of milling results in lower surface lipid concentrations (Liu, et al., 1998). A vision system is used to acquire images of the both sides of the rice grain. Applying thresholding, a threshold value of 10 is used to create a binary image. The total area of the grain is obtained based on the pixel binary values. The same method was used to get the total area of the surface lipid. A threshold value of 10<T<255 was used to normalize the pixels within the range. The total area of the surface lipid is obtained based on the pixel binary values. The T value is adjusted based on the amount of feature extraction for the surface lipid concentration. The surface lipid area percentage is the ratio between the total area surface lipid and the total area of the grain.



**Figure 2.4** The relationship between the surface lipid area percentage versus the milling duration with T = 130 (Liu, et al., 1998)

By measuring the relative whiteness of the rice grain and the region percentage of the color imperfections caused by varying surface lipid, the degree of milling can be estimated.



**Figure 2.5** Detecting bran and other internal coverings using thresholding (Liu, et al., 1998)

Furthermore, low to medium degree of milling retains some bran coverings on the rice grain. Through image processing, the bran coverings’ area can be compared to the whole grain area. Comparing the ratio of bran to whole area to some calibration values, the degree of milling can be estimated. Ultimately, this is similar to the method mentioned previously by measuring the amounts of imperfection on the surface of the grain.

**Damaged kernel**

External factors like the temperature, humidity, presence of pests, fungi, and rots causes the damage in the kernels of the rice. In excess humidity, unhusked rice grains could begin to sprout. This is visible even when the rice grains are milled. Also, fungal contamination like molds contribute to the decreasing quality of the rice grains. Damaged kernels can be distinguished by the visible and unnatural looking spots in the rice grain or the presence of sprout like structures.



**Figure 2.6** Collection of damaged kernels (Buhler, 2017)

Damaged kernels show unnatural discolorations. In a study made for classifying defects in rice grains, gray scale images were analyzed for features of discoloration (Chandra, et al., 2014). Using statistical parameters, the texture properties of the rice grain were analyzed. The gray level co-occurrence was analyzed to provide an insight to the texture of the rice grains. The damage ratio is described as the ratio of damage texture to the healthy texture. As with the detection of milling degree, the surface imperfections can be used to estimate the damage in the kernels by getting the ratio of the classified damaged region and the total area of the region.

**Discolored kernel**

Discolored kernels are those that have gone through the process of ‘fermentation’. Rice products that have undergone this process indicates that the products are stored for a long time. Unnecessary high storage temperature also causes discolored kernels. Fermentation in the rice can be indicated by the collective yellowness of the grains.



**Figure 2.7** Yellow discolored Basmati rice grain (Buhler, 2017)

By estimating the mean yellow of the rice grains and then compared to calibrated values, the degree of fermentation can be obtained. Also, the yellowness tends to form a gradient on the rice grain meaning there is a low tendency for a region to be sharply yellow compared to the other regions. The yellowness tends to diffuse from the starting point of the fermentation.

**Chalky kernel**

Rice grains with colors that can be compared to the color of a white chalk tends to break easily (Bambole, et al., 2015). Chalky grains, when stored or milled, turns into broken grains and powdery substrates. The degree of whiteness on a region compared to the overall whiteness of the sample is used to indicate the chalkiness of the rice grain.

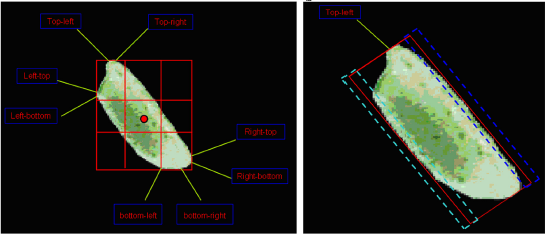
The degree of chalkiness could be processed as the measured opaqueness of the rice grains by passing light through them. However, more recent methodologies use thresholding algorithms that compare the degree of whiteness of the rice grains with the threshold value and then computing the summed ratio of the degree of whiteness (Bambole, et al., 2015). Based from the study, the chalkiness of the sample is obtained by the ratio of the number of the grains classified as chalky and the total number of the sample grains. To determine if a grain is chalky or not, a threshold was applied to the binarized image. Chalky regions are represented by black pixels. The ratio between the chalky region to the whole region represents the chalkiness of the grain. If the chalkiness is over 20%, then the grain is considered chalky.

Using thresholding, the colors can be differentiated and reduced to standard preset values. This makes the color distribution easier to analyze. The chalkiness of the rice can be measured by the ratio of the chalky white regions to the overall area of the region (Chandra, et al., 2014).



**Figure 2.8** (a) original image to apply threshold, (b) resulting threshold binary image to extract chalky regions (Chandra, et al., 2014)

In another study, a support vector machine coupled image processing method was developed to identify the chalky regions and the dimensional features of the grain. The application of linear classifiers proved to be successful with a maximum success rate of 98.5% classification. By segmenting the chalky threshold regions, the area of those regions was quantified and compared to the overall grain area region. This method is similar to the previously mentioned classification techniques (Sun, et al., 2014)



**Figure 2.9** Segmented chalky regions through thresholding (Sun, et al.,2014)

**Immature kernel**

Immature kernels are those that are harvested before the desired maturation of the seeds. Immature kernels exhibit their young age through the color of the endosperm. As mentioned, the desired maturity of the grains is indicated by the yellow to golden color of the husk. Immature kernels have greenish endosperms. The degree of green in the endosperm indicates the immaturity of the rice kernel.

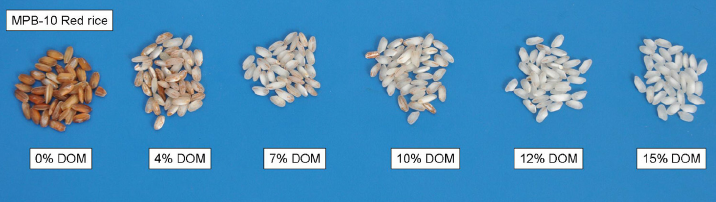


**Figure 2.10** Immature rice grains (Riceland, 2016)

Similar to fermented kernels, the diffusion of the color is gradient. In this case, the green color tends to diffuse from tips radiating towards the center of the grain. No visibly sharp regions are present.

**Red kernel**

Red kernel properties are almost similar to the degree of milling. The bran is the covering of the endosperm. Red kernels still have this covering depending on the milling. Well-milled rice grains are supposed to contain less red kernels. However, for regular milled rice grains, the red kernels are relatively abundant. In essence, red kernels could determine the degree of milling of the rice grains.



**Figure 2.11** Degree of milling comparison (Paiva, et al., 2014)

**Machine Learning Classifiers**

Machine learning is the use of computing process in order to mimic or model a behavior, classification, or identity of a certain application which can include image recognition, object detection and more. One type of machine learning encompasses supervised learning methods. In a supervised learning environment, the algorithm generates a model fed with a set of training data which are usually made up of feature vector and class pairs. The feature vector is a vector space containing the quantification of the characteristics that distinguish a class. An example of a supervised machine learning algorithm is the support vector machine algorithm (SVM).

**Support vector machine (SVM)**

A support vector machine is often used as a classifier based on supervised learning. The algorithm separates or classify data by an optimal hyperplane inside a hyperspace. The position of the data when plotted against the hyperplane determines its class. The input vectors’ dimensional space are increased non-linearly (Support Vector Networks). The SVM’s hyperplane is derived to maximize the margins on both sides of the hyperplane. For low dimensional linearly separable data with *n*-dimensions, the hyperplane could be described in a *(n-1)*-dimensional space. As illustrated in figure 2.12, the two-dimensional data set is separated with a one-dimensional hyperplane which margin sizes are optimized by the support vector machine. The margin size is the distance of the margin plane to the hyperplane. The training data used is a set described by . Where , and for all *i=*1,2…*u*. The optimal hyperplane, , is the basis of the decision where to classify a point.

Hyperplane

Margin

Figure 2.12. An SVM Classifier Plot with Optimized Hyperplane

Finding the hyperplane is an optimization problem. The support vectors are the nearest vectors to the hyperplane and in which the equation 2.1 holds true. The margin size can be described by the inequality 2.2. The optimal hyperplane is a linear combination of support vectors. The vector for the hyperplane, , can be found by equation 2.3.

**(2.1)**

**(2.2)**

**(2.3)**

**Adaptive boosting**

Weak classifiers are often evaluated as those whose accuracy is below random chance (A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting). Several algorithms are developed in order to ‘boost’ the accuracy of these classifiers. In general, this process is called boosting. The most common method of boosting is the adaptive boosting or AdaBoost . In general, the algorithm takes any classification algorithms and statistically boost the probability of the erroneous data labelled data of being able to train the model better. Each labeled data is randomly selected and given weight based on their influence on the error. A new model will be generated and the weights will be updated also based on the individual errors of each labeled data. The error of the model data is:

Where = error

= probability distribution of labelled *i*thdataat iteration *t*

= output of the classifier given

= supposed output of the classifier given

Normalizing the error,

Thus, the weights of the vectors for the next iteration can be updated by:

Since the AdaBoost algorithm generates many models of different accuracy, at the end of the algorithm, a convening step is needed to produce a single output of the model and is described by equation 2.4. The *H* is the overall hypothesis of the models generated. The ‘1’ and ‘0’ represents the classes of the weak classifier.

**(2.4)**

**Histogram of oriented gradients**

The histogram of oriented gradients (HOG) descriptor can be used as a feature set for classification. The HOG descriptor is a histogram (distribution) of the orientation of gradients (Histograms of Oriented Gradients for Human Detection). The process begins with the calculation of gradients over a small section of an image. The image will be divided into subsections and the histogram of the gradients from the subsections will be created. The block normalization process is applied to the image by sliding. A normalized vector is the result of process which is called the HOG feature vector. These feature vectors are concatenated into a single feature vector with a high dimension and can be used as input vector to a classifier.

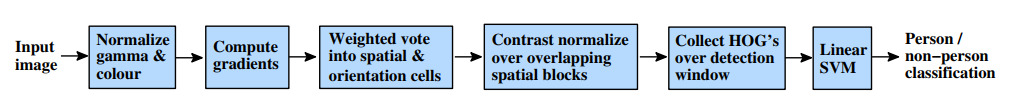


Figure 2.13. The process of using HOG features as input to a Linear SVM (Histograms of Oriented Gradients for Human Detection)