Image Analysis of Foliar Greenness for Quantifying Relative Plant Health

Anton Louise P. de Ocampo[a], **Joseph B. Albo**[b], and **Kerish J. de Ocampo**[c]
Batangas State University, Batangas City, Philippines
Corresponding Author Email: anton.deocampo@gmail.com, yosef06_16@yahoo.com, kereishdeocampo@yahoo.com

ABSTRACT

Foliar color has been used by farmers and specialists as visual health indicators for plants. However, assessment of health thru naked eye is subjective; hence, Munsell Plant Tissue Color Chart, Globe Plant Color Chart, and Leaf Color Chart have been developed. More accurate measurements of plant physiological processes such as leaf development, leaf senescence, light absorption, and foliar nutrient status, require the use of Chlorophyll meters, spectrophotometers, reflectometers and spectroradiometers. The aim of this work is to develop an experimental set-up to validate the use of digital images of foliar parts to quantify relative plant health of tomato plants. With this study, quantifiable measure can be achieved to rate the relative health of any plant by taking images of its leaves. To get reliable results, the model characterization used features extracted from captured images of sample leaves taken from 40 plants. Twenty plants are identified by a plant specialist as healthy and the other 20 plants as unhealthy. Bartlett's and F-Tests were used to develop a statistical model which defines the group of healthy plants from the unhealthy ones. Randomly selected plants are then subjected to the model for classification and relative valuation of plant health.

Results show that the model successfully classified healthy plants against unhealthy plants using image analysis of foliar greenness. Twenty healthy and 20 unhealthy plants were selected by the plant specialist and then subjected to the model for classification. Thus, measures of dispersions between the pixels of foliar image as well as that of between leaves in three sections of a mother plant can be used to identify the relative health of plants.

KEYWORDS: Leaf Color, Image Analysis, Statistical Model, Relative Plant Health, Bartlett's Test, F-Test

1. INTRODUCTION

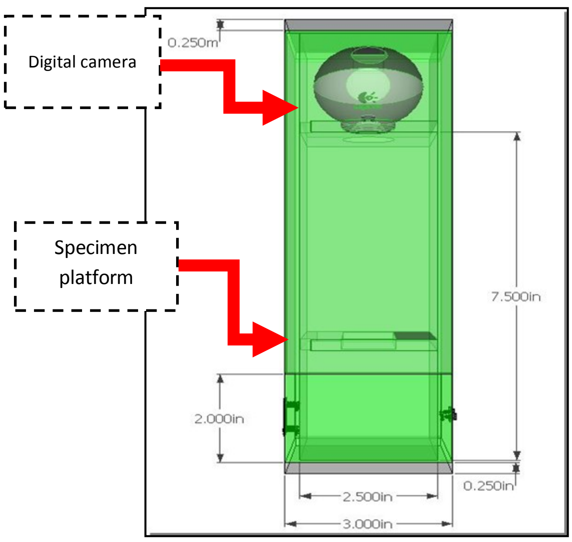
Greenness of leaf has proven to be relevant to many requirements of plant monitoring to assess the health condition of plants. Visual inspection, chemical analysis (leaf and petiole analysis), reflection or transmission of light, and image processing techniques are the major classifications of methods used for the purpose of evaluation of plants' health conditions. Among these methods, human visual inspections with or without using Leaf Color Chart (LCC) is the most common practice used by farmers in assessing whether the plants are well-fertilized or not during the season. This method, however, introduce much subjectivity in the results of observations [1]. Use of digital imaging can provide a more objective yet still practical and convenient approach in plant monitoring and plant health assessment.

The aim of this work is to develop an experimental set-up to validate the use of digital images of foliar parts to quantify relative plant health of tomato plants.

2. EXPERIMENTAL SET-UP

2.1 Hardware Set-up

For the acquisition of digital images of the leaf specimen, a 3.0-Megapixel camera is installed in a light-insulated photo box (as shown in figure 1). Captured images are directly saved in a notebook PC for image processing.



2.2 Methodology of sampling

According to Mr. Ruben O. Recede, a plant specialist from the Department of Agriculture-STIARC, the following sampling procedures must be considered to obtain the Standard Values for Healthy Tomato Leaves:

1. Random samples must be 2 meters away from the border line of the field.
2. Collecting samples must be done with clean and dry hands. Samples should be placed in a tissue and stored in a cool place (refrigerator) immediately after picking to prevent leaves from deteriorating. Avoid samples being contaminated unless samples are kept in a cool place.
3. Three parts of tomato plant (Top, Middle and Bottom) must be considered because the nutrient distribution in plants vary with its height from the ground to top as shown in Figure 2a.
4. Twenty plants with 3 different parts each (Top, Middle and Bottom) are enough for sampling.
5. Choose the mid-shoot leaf in each part of the tomato plants as shown in Figure 2b.

(a) (b)

Figure 2a: Tomato plant sections; Top, Middle and Bottom sections. 2b:Mid-shoot leaf part to be selected

2.3 Training and Execution process

Training and learning process includes selecting 20 plants that are classified by a plant specialist as healthy based on morphological and physiological aspect of the plant. For every plant, 3 leaves are taken: one from each section of the plant (Top, middle, bottom). then each leaf is subjected to feature extraction by image processing, which is used to define a statistical model for healthy plants. Execution process starts by randomly sampling 20 plants. Again, top, middle, and bottom leaves were taken and their features extracted. These features are then subjected to the developed model to quantify the relative plant health of each plant. Figure 3 summarizes the training and execution process.

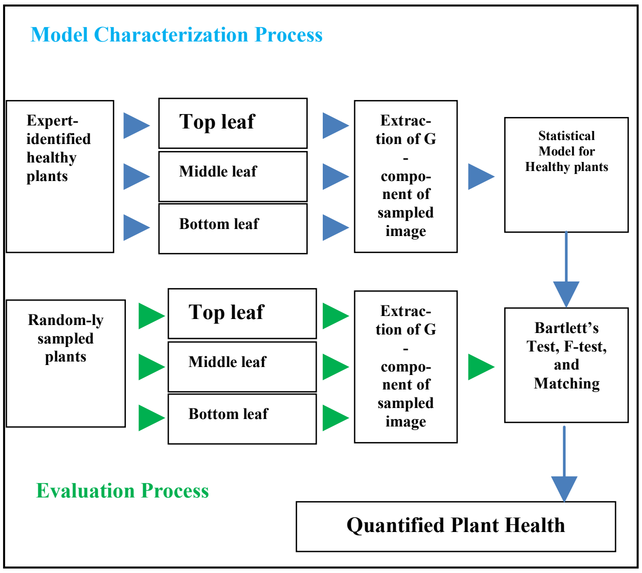


Figure 3 Model Training and Execution Process

Captured images of the leaf specimens are subjected to enhancement prior to RGB decomposition where green component of the image are extracted for statistical analysis.

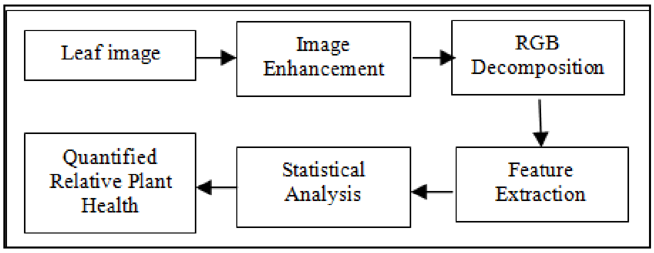


Figure 4.Process flow for analysis of foliar color

An example of captured image is shown in Figure 5 (a). The image is then planted at maximum green area possible as shown in figure 5(b).

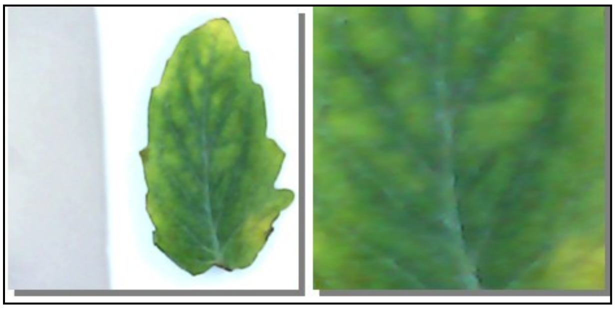


Figure 5. (a)captured image of the leaf specimen; (b)plantped image focusing on the center of the leaf.

Green component of the image is then extracted. Green values are mapped together with the other samples from the same plant and/or same group of tomato plants. The mapped data are then fed to a Simulink model for statistical analysis.

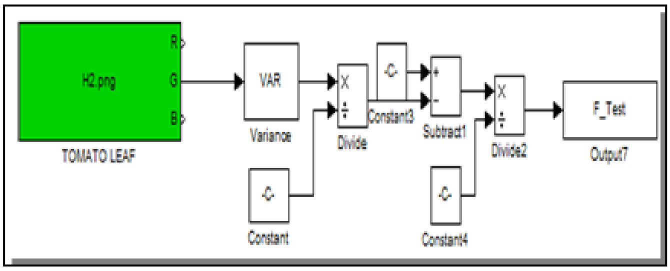


Figure 6 . Simulink model for data extraction and analysis

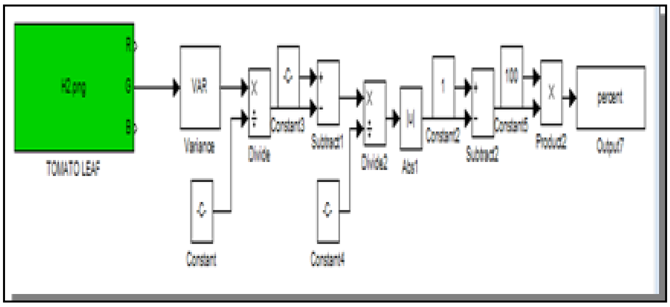


Figure 7 . Simulink model for feature matching and evaluation

III. RESULTS AND DISCUSSION

3.1 Data Gathering

In this experiment, 60 leaves from 20 expert-identified healthy plants were used to develop the statistical model for the healthy plants. Three leaves, one per section of the plant (top, middle, bottom), were sampled. Each leaf was subjected to feature extraction to obtain its foliar greenness.

Table 1. Variance of means of the greenness level of all Top, Middle and Bottom leaves

Leaf	Variance of Means	Description
Bottom S^2_m	0.007383963	Variance of Means from the Data of Bottom Leaf
Middle S^2_m	0.00522995	Variance of Means from the Data of Middle Leaf
Top S^2_m	0.00289721	Variance of Means from the Data of Top Leaf

Table 2. Averages of variances of the greenness level of all Top, Middle and Bottom leaves

Part of Leaf	Average of Variances	Description
Bottom A_v	0.007383963	Average of Variances from the Data of Bottom Leaf
Middle A_v	0.00522995	Average of Variances from the Data of Middle Leaf
Top A_v	0.00289721	Average of variances from the Data of Top Leaf
I A_v	0.002060119	Average of variances from the Data of Bottom, Middle and Top Leaf

3.2 Bartlett's Test

This method was used to determine the equality of variances of two groups (Healthy group and Randomly-sampled group). It is assumed that if the variances of the groups are the same, then the randomly selected specimen belongs to the same group of healthy plants. If not, then the hypothesis that the sample came from a healthy plant cannot be accepted.

Bartlett's test uses a statistic **B** which describes the departure of the sample from the normality which, in this study, is the Healthy group. **B** is expressed as follows

$$B = \frac{\left[(s_1^2)^{n_1-1}(s_2^2)^{n_2-1}\ldots\ldots\ldots(s_k^2)^{n_k-1}\right]^{\frac{1}{(N-k)}}}{s_p^2}$$

$$S_p^2 = \frac{1}{N-k}\sum_{i=1}^{k}(n_i - 1)$$

Where:

Total Variances of Mean of Leaf Sample

S^2_k =variance of Means of designated Leaf

$$s_p^2 = \frac{1}{60-3}\sum_{i=1}^{3}(20-1)(.007383963+.00522995+.00289721)$$

N = $s_p^2 = 0.005170374$ Total Number of Leaf

$$B = \frac{\left[(.007383963)^{19}(.00522995)^{19}(.00289721)^{19}\right]^{\frac{1}{(60-3)}}}{0.005170374}$$

$$B = 2.8653E-132 < \text{the critical values of } .8964$$

n = Leaf Sample

3.3 Quantifying Relative Plant Health

Each image sample contains approximately 3 Megapixels (RGB) where two-thirds of which is the leaf sample as shown in the figure. Central tendency and dispersion of the green values are computed to represent the leaf sample.

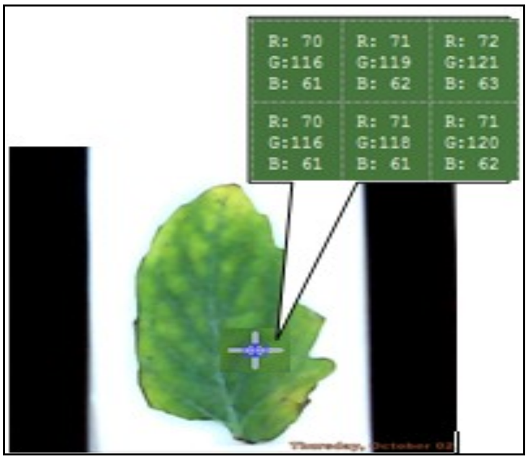


Figure 8 .RGB components of 6 pixels sample from a leaf

Consistency of greenness among the pixels of the planted leaf image is used to relatively quantify the health of the leaf which implies the health of its mother plant. To do this, F-test is conducted to test the homogeneity of the pixels of the leaf images from top, middle and bottom sections of the plant.

$$F = \frac{Average\ of Variances from Healthy Leaf}{Variance of New Sample Leaf}$$

The F statistic is tested against a 95% confidence interval for F- distribution. If F-value is within the range of 0.5319 to 1.8800 of the F-distribution table, the mother plant from which the sample leaf was taken from is considered healthy.

The value of the F-statistic is also used to quantify the status of the relative health of the mother plant.

$$\%\ relative\ health = \left(1 - \frac{\left|\left(X - \frac{1.88+.5319}{2}\right)\right|}{(1.88 - .5319)}\right) X\ 100\%$$

3.4 Testing and Evaluation

To test the model, 40 leaf samples from 40 randomly selected tomato plants were subjected to classification by a plant specialist. Twenty plants were identified as healthy and the other 20 were classified unhealthy. Then the same leaf samples were tested using the model developed. The table below shows the F-value of the sampled plants, the % relative health, the model's and the plant expert's classifications.

Table 3. Relative plant health

Randomly selected plants	F-value	Relative Plant Health	Model's Classification	Expert's Classification
Plant 1	2.3532	14.90%	Unhealthy	Unhealthy
Plant 2	1.0518	88.56%	Healthy	Healthy
Plant 3	1.3092	92.34%	Healthy	Healthy
Plant 4	1.2765	94.77%	Healthy	Healthy
Plant 5	1.0192	86.15%	Healthy	Healthy
Plant 6	0.8537	73.87%	Healthy	Healthy
Plant 7	0.5774	53.37%	Healthy	Healthy
Plant 8	1.2292	98.27%	Healthy	Healthy
Plant 9	0.4644	44.99%	Unhealthy	Unhealthy
Plant 10	1.6625	66.13%	Healthy	Healthy
Plant 11	0.9292	79.47%	Healthy	Healthy
Plant 12	0.6155	56.20%	Healthy	Healthy
Plant 13	1.248	96.88%	Healthy	Healthy
Plant 14	1.1411	95.19%	Healthy	Healthy
Plant 15	0.7573	66.72%	Healthy	Healthy
Plant 16	0.9385	80.16%	Healthy	Healthy
Plant 17	0.848	73.44%	Healthy	Healthy
Plant 18	0.8186	71.27%	Healthy	Healthy
Plant 19	1.0642	89.49%	Healthy	Healthy
Plant 20	0.7504	66.21%	Healthy	Healthy
Plant 21	3.2223	out-of-range	Unhealthy	Unhealthy
Plant 22	2.5068	3.51%	Unhealthy	Unhealthy
Plant 23	1.8846	49.66%	Unhealthy	Unhealthy
Plant 24	2.8138	out-of-range	Unhealthy	Unhealthy
Plant 25	2.0914	34.32%	Unhealthy	Unhealthy
Plant 26	4.997	out-of-range	Unhealthy	Unhealthy
Plant 27	2.906	out-of-range	Unhealthy	Unhealthy
Plant 28	2.3834	12.66%	Unhealthy	Unhealthy
Plant 29	2.167	28.71%	Unhealthy	Unhealthy
Plant 30	2.4156	10.27%	Unhealthy	Unhealthy
Plant 31	3.3053	out-of-range	Unhealthy	Unhealthy
Plant 32	4.459	out-of-range	Unhealthy	Unhealthy
Plant 33	2.7702	out-of-range	Unhealthy	Unhealthy
Plant 34	0.9529	81.23%	Healthy	Healthy
Plant 35	2.0172	39.83%	Unhealthy	Unhealthy
Plant 36	1.5462	74.76%	Healthy	Healthy
Plant 37	2.0512	37.30%	Unhealthy	Unhealthy
Plant 38	2.9835	out-of-range	Unhealthy	Unhealthy
Plant 39	3.4369	out-of-range	Unhealthy	Unhealthy
Plant 40	2.1078	33.10%	Unhealthy	Unhealthy

Plants with leaf F-value goes beyond the 95% confidence interval are readily classified unhealthy. In addition, plants with leaf F-value falls within the 95% C.I. are rated on how close they are to the model's definition of healthy plant. Plants having the rate of 50% and above are labeled as healthy, otherwise, unhealthy.

The model performed significantly well since out of 40 plants, it has classified exactly the same as the classification of the plant specialist

REFERENCES

[1] P. Murakami, M. Turner, A. an Den Berg, and P. Schaberg, "*An Instructional Guide For Leaf Color Analysis Using Digital Imaging Software Delaware*", USDA Forest Service. 2005

[2] A. Gupta, and K. Amandeep, *"Simulink Model Based Image Segmentation"*, Vol.2, Issue 6. International Journal of Advanced Research in Computer Science and Software Engineering. June 2012

[3] D. Kaiser, J. Lamb, and C. Rosen, *"Plant Analysis Sampling and Interpretation"*, University of Minnesota Extension: Minnesota's Agricultural Fertilizer Research and Education Center (AFREC), 2013.

[4] R. Kalaivani, S. Muruganand, S. and A. Perisami, "*Identifying the Quality of Tomatoes in Image Processing Using MATLAB*", Vol. 2 Issue 8.International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering. 2013.

[5] S. Kawashima, and M. Nakatani, *"An Algorithm for Estimating Chlorophyll Content in Leaves Using a Video Camera*", Japan: Annals of Botany Company, 1997.

[6] M. O'Neill, "*A Guide To Linear Mixed Models In An Experimental Design Context*". Statistical Advisory Training Service Pty Ltd. August 2010

[7] V. Abergos, P. Boreta, R. Comprado, S. Soltes, and A. Tatel," *Android-Based Image Processing Application for Rice Nitrogen Management"*, Ateneo de Naga University, 2012.

[8] K. Schaller, *"Leaf nutrient diagnosis – towards a more sustainable viticulture",* 2007

[9] A.Vibhute, Sk. Bodhe, & Bm. More. *"Nitrogen Estimation for Grapevine (In Veraison) Using RGB Color Image Processing"*. Department of Electronics and Telecommunication Engineering at BMIT, Solapur. India, 2013.