Athens Journal of Sciences 2023, 10: 1-10
https://doi.org/10.30958/ajs.X-Y-Z

Identification of Greenish Soybean Seeds through Image Processing, under Different Types of Lighting

By Alcebíades Fogaça de Souza Sobrinho[*]*, Roberto Alves Braga Junior*[±]*, Edvaldo Aparecido Amaral da Silva*[°] *& José Luís Contado*[•]

A high occurrence of greenish soybean seeds in crops is an issue, these types of seeds have low physiological quality, which can generate seedlings with anomalies, if destined for processing in industries, the presence of chlorophyll is undesirable, requiring additional processes for its removal. This work aimed to evaluate the differentiation of mature and greenish soybean seeds, illuminated with red laser, green laser, red LED, and fluorescent lamp, using image processing. Images of mature and greenish soybean seeds were captured at a resolution of 340x480 pixels, illuminated with red laser, green laser, red LED, and fluorescent lamp. Subsequently, the averages of the gray levels of each image were obtained in the red, green, blue channels and in images converted to grayscale 8-bit. The data were submitted to tests of variance after gray level for image classification. And a validation presented results of 97% of hits for red laser, 94% for fluorescent light and 93.5% for red LED, all in red channel.

Keywords: *seed classification, agricultural automation, soybean seed quality, computer vision*

Introduction

Soybean is one of the most important crops in the world, its grains are used both for protein food and for the production of vegetable oil (Hartman et al. 2020). The total cultivated area in the world during the 2018/2019 harvest was approximately 120 million hectares, with a total grain production of around 335 million tons (USDA 2020). And for soybean production, climate instability is an increasingly recurrent problem. Crop yields in general, including soybeans, are expected to decline due to climate stress factors that include drastic temperature fluctuations, drought, flooding and high salinity of water and soil (Li et al. 2017). High temperatures and water stress can lead to the occurrence of "green seeds" in soybean, which is characterized by the retention of chlorophyll in mature seeds (Teixeira et al. 2016). At seed maturation, the occurrence of abiotic and biotic stresses, including nematode parasitism, can result in

[*]Professor, Federal University of Lavras, Brazil.
[±]Professor, Federal University of Lavras, Brazil.
[°]Associate Professor, State University Júlio de Mesquita Filho, Brazil.
[•]Associate Professor, Federal University of Lavras, Brazil.

premature death of soybean plants, resulting in the production of greenish seeds (França-Neto et al. 2012). In Brazil, where more than 60% of soy production is concentrated in tropical regions, the occurrence of high levels of green seeds is becoming an increasingly frequent problem for producers (Cicero et al. 2009). In the province of Ontario, Canada, due to the scarcity of rain, coupled with the extremely hot weather, resulted in samples collected containing 15% greenish soybeans in most fields of the 2002 crop, with samples containing up to 70% soybeans greenish (Lackey 2002). Lots with a high percentage of greenish soybean seeds or grains are classified as lower quality products, and their selling price is reduced. In the United States, classification grades for yellow soybeans are divided into 1, 2, 3, and 4, and the maximum established number of beans of other colors in samples for each classification grade is 1, 2, 5, and 10 percent respectively (USSEC 2006). Residual chlorophyll compromise the quality of seed material in terms of viability, nutritional value and shelf life, and pose a serious challenge for farmers (Smolikova et al. 2017). In rapeseed (*Brassica napus* L. var. Napus), carrot (*Daucus carota* L.), corn (*Zea mays* L.) and soybean (Glycine max L.), accelerated aging reduced germination to 61% in the seed sample with low chlorophyll content and to 30% in the sample of seeds with high chlorophyll content (Smolikova et al. 2011). The increased incidence of green seeds in soybean seed lots negatively affects their quality (Pádua et al. 2007). This reduction in quality occurs because the higher deterioration rate of green seeds results in abnormal seedlings, with non-uniform size and low root development, compared to seeds in which the chlorophyll degradation process has been completed (Forti et al. 2015). Drying soybeans at an air temperature of 25 C allowed the degradation of chlorophyll in greenish grains harvested at stage R6 and later, causing green pigments to not be detected (Sinnecker et al. 2005). Drying soybeans at low temperatures may be a solution to eliminate the greenish grains in the post-harvest stage, however for the use of soybeans as seeds, this process is ineffective. Soybean seeds which are harvested immature and dried with ambient or heated air, with temperature variations between 25 C to 29 C, exhibited low germination, less than 32% (Samarah et al. 2009). Once that green soy is unfeasible for use as seeds, and that conventional processes carried out post-harvest are not enough to solve this problem, it's clear that the development of specific methods to separate the greenish grains from others is required.

The hypothesis of this work was that an accessible and robust optical device could address the separation of greenish from mature soybean seeds. Therefore, we aimed to test a feasible and a reliable setup with multiple artificial light sources, image capture and processing to provide the best achievement in separating greenish from mature soybean seeds.

The paper is structured in an Introduction of the theme addressing the problem of greenish soybean, and the importance of the classification. The Review of Literature pointed out the techniques developed so far to separate seeds using many methodologies. In the Material and Methods section, we detailed the methodology adopted to test the best configuration using artificial light sources and camera to acquire and process the images. The section also presented the image processing and statistical analysis of the data, finishing with a validation step. The Results and Discussion presented the best performances of the light sources adopted and the result of the validation test, finishing the paper with the Conclusion.

Literature Review

The advancement of technology has provided useful improvements in our daily activities, of which agriculture is no exception, and has paved the way for new farming techniques all over the world (Abdulhamid et al. 2019). And the search for the implementation of innovative methods that automate traditional methods in agriculture is a constant effort, and one that precedes for a long time. Electric color sorters were used to sort beans in Michigan as early as the 1930s, the first machines used were slower than manual sorting, and their use was justified based on accuracy (Boyd Junior et al. 2021). And the increase in the processing capacity of microcomputers, made it possible to develop computer programs for seed evaluation. The Groundeye® software was able to progress the growth of coleoptiles and roots of popcorn hybrids, in a saline potential of -0.9 Mpa, and the evaluations through images performed with the software were efficient in the evaluation of the vigor of these seeds (Catão et al. 2020). The greenish soybean contains chlorophyll, which was not degraded in the maturation process, this characteristic can serve as a parameter for the segregation of these undesirable grains. Using the chlorophyll fluorescence technique, samples of soybean seeds of the cultivar TMG 113 RR were separated, sub-samples containing high and low chlorophyll fluorescence, and verified that the time for the seeds to reach 50% of maximum germination was shorter for seeds with low chlorophyll fluorescence, compared to ones that showed high chlorophyll fluorescence (Cicero et al. 2009). With the practicality and flexibility that computer resources provide, its application in agriculture contributes to the development of new techniques and equipment with the most diverse purposes. A seed sorter based on color was patented: the seeds were illuminated with strobe lights, which under a digital camera captured the images and sent them to a computer (Sheldon and Affleck 1990). Using computer vision in a prototype, it successfully identified impurity fractions in grains with an accuracy of 96% for split grains, 75% for contaminated grains, and 98% for defective grains, stems, and pods (Momin et al. 2017). The basis for the success of any machine vision application is high image quality, and normally, the contrast needed comes from the fashion the object is illuminated, and it is estimated that 90% of the applications can be solved just by selecting the lighting correctly (Blackman 2017).

Materials and Methods

The soybean seeds used in the experiment were from the MG/BR 46-Conquista cultivar, produced in the 2014/2015 harvest at Fazenda Experimental Lageado, belonging to the Faculty of Agronomic Sciences - Unesp, located in the municipality of Botucatu, São Paulo, Brazil (22 °50'32" S; 48°25'29" W; 750 m altitude). Seeds were classified as mature and greenish according to the chlorophyll content determined by the High-Performance Liquid Chromatography method (Teixeira et al. 2016).

According to the Köppen classification, the climate of the region is characterized as humid hot temperate (mesothermal), CFA type, with an average annual rainfall of 1,501.4 mm and an average annual temperature of 20.3º C.

Images of mature and greenish soybean seeds were captured at a resolution of 340x480 pixels, illuminated with red laser, green laser, red LED, and fluorescent lamp. Subsequently, two groups of images were formed for each type of lighting, one group with 200 images of mature seeds, and the other with 200 images of greenish seeds.

Image Acquisition and Lighting Adjustment

The intensity of the laser and LED lighting was controlled employing a neutral filter so that there was no saturation of the pixels by excess light. The luminosity of the fluorescent lamp was 245 lumens, measured with a digital lux meter. Images were captured using a Dino-Lite model AM211 digital camera, connected to a computer via a USB port. The seed was placed on a matte black background.

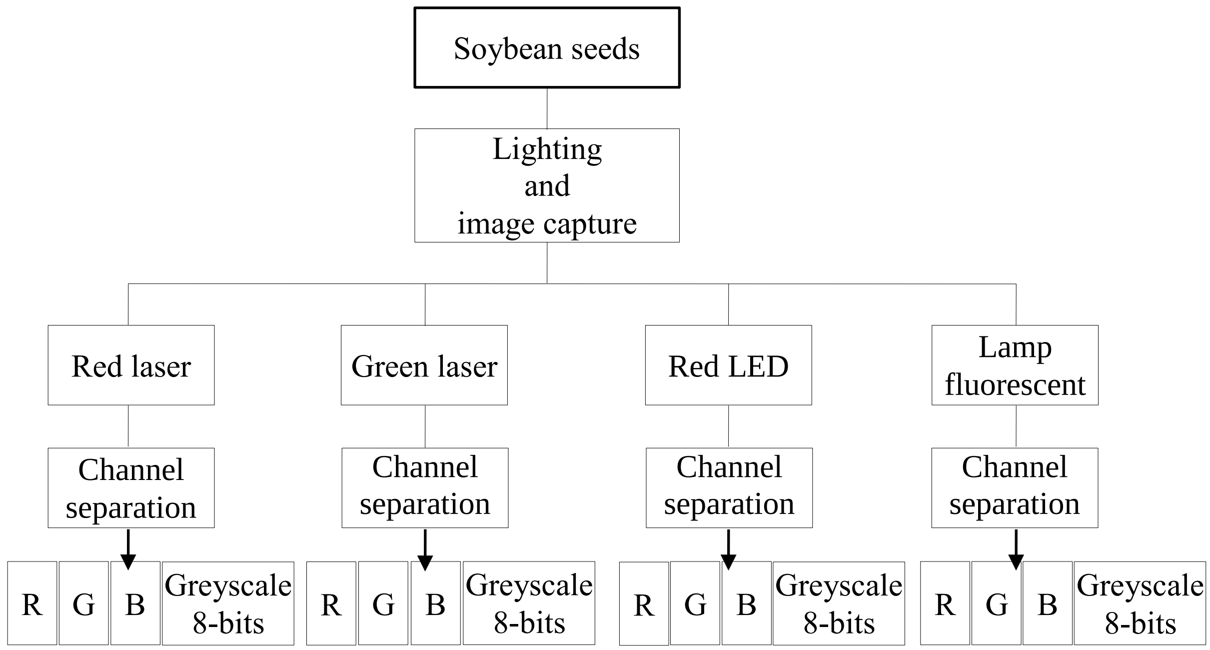
The control of the laser and LED light was done by monitoring the histogram of the image of a yellow-colored paper, which was used as a default background for the adjustment of the illumination intensity, with the averages of the gray levels shown by the histogram, stabilized at approximately 160 with the maximum value not reaching 255, indicating the absence of pixels saturated by excess lighting.

Pre-Processing of Images

Samples were taken from the region of interest for analysis and their images captured using the method previously described. Images of the seeds illuminated with fluorescent lamp were extracted in the resolution of 225x225 pixels, and images of the seeds illuminated with laser and LED were extracted in the resolution of 200x150 pixels.

Subsequently, the samples were duplicated into two groups, from one group the red, green, and blue (RGB) channels were separated, and the other group was converted to grayscale 8-bit. This procedure was done with the images of mature and greenish soybean seeds, in each type of lighting, organization chart 1 exemplifies the process of pre-processing the images. The total number of images used for analysis in each type of lighting was 1600, totaling, 6400 processed images.

Organization Chart 1. *Stages of Image Pre-Processing*



Data Acquisition and Analysis

The analyzed data is the average of the gray levels of the image samples, calculated by equation 1.

$$\mu i = \frac{\sum_{k=1}^{n}(\text{ncki})}{n} \tag{1}$$

Where ***µi*** is the average value of the gray levels of an image; ***nc*** is the pixel gray level value; and ***n*** is the total number of pixels in the image. The gray level averages were obtained using the ImageJ software. To verify the existence of statistical difference between the means, the data was submitted to the variance test at 5% probability.

Classification of Images – Biasing and Validation

The creation of a reference value of median for the gray scale images to separate the seeds was done in a group of 200 images of mature seeds and 200 images of greenish seeds. The median values of gray occurrences of 100 images, randomly selected from the two groups, were obtained and a umbral values – the mean of both medians – were set as references for classification. The classification of mature and greenish seeds was conducted, and the level of hits measured. Since all the seeds were tagged individually, with the primary classification in the two groups already known.

A procedure with one million interactions was carried out with the median and percentage of hits for mature classification identified. The means of medians values and hits were set as global reference to the validation step.

The classification references were set to the four types of illumination as well as to the channels of RGB and global gray.

The validation step used the 400 images – 200 from mature and 200 from greenish mixed – with the median values with the highest percentage of hits to mature seeds for all the treatments.

Results and Discussion

References for Classification

In the analysis of variance at 5% of probability, in all types of lighting, the differences between the average values of the gray levels of the images concerning the mature and greenish seeds, in the RGB channels and the images in grayscale 8-bits, were statistically significant with the P-value being 0.000 in all cases (Table 1). The next step was then to evaluate the best approach using the lighting and the channels of RGB or global gray.

Table 1. *Variance Analysis*

Variables	*CV (%)	General average	P value
	Fluorescent lamp		
Red	4.32	174.89	0.000
Green	2.82	124.43	0.000
Blue	5.25	98.66	0.000
Grayscale 8-bits	4.16	132.65	0.000
	Green laser		
Red	8.73	9.42	0.000
Green	6.72	87.51	0.000
Blue	9.35	38.49	0.000
Grayscale 8-bits	6.63	45.14	0.000
	Red laser		
Red	7.77	136.62	0.000
Green	19.78	6.41	0.000
Blue	5.55	13.95	0.000
Grayscale 8-bits	6.03	52.33	0.000
	Red LED		
Red	6	140.21	0.000
Green	25.21	4.56	0.000
Blue	10.60	8.12	0.000
Grayscale 8-bits	4.24	50.98	0.000

*Coefficient of variation

In Table 2 it is shown the average gray levels of the images. The channel that conveyed the most considerable difference between the averages of gray levels was red, in the lighting with red laser, red LED and fluorescent lamp respectively. In lighting with green laser, the channel that revealed the most significant difference between the averages of gray levels was green; however, this difference was smaller than the differences found in the red channel, of the other types of lighting evaluated. Evidencing that the images formed in the red channel, except when illuminated with green laser, were the ones that provided the best differentiation between the images of mature and greenish soybean seeds, and the efficiency in the differentiation is increased using the red laser.

Table 2. *Overall Average of the Gray Levels of the Images*

Seeds	8-bit gray scale	Red	Green	Blue
				Fluorescent lamp
Mature	137.07	183.23	127.1	100.88
Greenish	128.23	166.54	121.77	96.45
				Green Laser
Mature	47.25	9.13	91.68	40.94
Greenish	43.02	9.7	83.33	36.02
				Red Laser
Mature	56.92	151.51	4.77	14.49
Greenish	47.73	121.74	8.06	13.39
				Red LED
Mature	53.18	148.84	3.41	7.25
Greenish	48.77	131.57	5.72	9.00

Classification of Images – Biasing and Validation

Table 3 contains the data obtained in the validation of seeds classification. The red laser, red LED and fluorescent lamp, in the red RGB channel, achieved an efficiency rate above 90%.

When we consider the parameters False positive and False negative, which indicates the proportion of images of greenish seeds classified as mature, and images of mature seeds classified as greenish, respectively, we can conclude that the red laser illumination in the red RGB channel was the one that showed the highest efficiency, when compared to the other types of lighting and channels.

The errors can be attributed to the classification process proposed and also by the primary classification of the seeds in the field. In the field, the greenish seeds came from a parcel identified as greenish, but that can have some mature within. The high level of hits observed here leads us to conclude that the proposed method can be applied efficiently.

Gomes-Junior et al. (2016) evaluated a classification system by color of Swingle citrumelo to identify at which stage of maturation the harvest should be carried out, to obtain seeds with greater physiological potential. And they found that the images in the red channel obtained greater differentiation between the gray levels of the images, of totally green, yellow-green, and yellow fruits. The results obtained with Swingle citrumelo corroborate those of soybean seeds, since during their maturation stages, the color of the seeds also varies from green to yellow, as in the fruits of Swingle citrumelo, consequently in both cases, the red channel of the images showed the greatest differentiation.

Table 3. *Validation of Seed Classification – Test of Efficiency in Mature Seeds*

	Hits (%)	False POSITIVE (%)	False NEGATIVE (%)
Red laser			
Red channel	97	6.5	3
Green channel	88	1.5	12
Blue channel	82	33	18
Grayscale 8-bits	96.5	7.5	3.5
Green laser			
Red	53.5	68	46.5
Green	85.5	28	14.5
Blue	82	37	18
Grayscale 8-bits	83.5	33	16.5
Red LED			
Red	93.5	16.5	6.5
Green	78.5	7	21.5
Blue	76.5	7.5	23.5
Grayscale 8-bits	93	17	7
Fluorescent lamp			
Red	94	14.5	6
Green	82.5	37.5	17.5
Blue	75	39.5	25
Grayscale 8-bits	87	27.5	13

Conclusion

This work was able to successfully propose a methodology to classify soya bean seeds in mature and greenish. The optical method proposed identified the best option of lighting and channel of RGB to classify the seeds efficiently. The red channel of RGB images with red laser lighting presented an efficiency of 97% of hits. Other options as gray level outcome or the use of red and fluorescent light also presented high values of hits. The achievement offers to users robust and accessible alternatives to classify seeds non-destructively.

Acknowledgments

This study was partially financed by São Paulo Research Foundation (FAPESP; grant number 2017/50211-9) and productivity fellowship to the E.A.A.S. (code 311526/2021-7) and has been supported by the following Brazilian research agencies: FAPEMIGPPM 163-17, CAPES, and CNPq.

References

Abdulhamid UF, Daniel S, Babawuro U (2019) Classification of Soya beans based image processing techniques and artificial neural network. *Journal of Advances in Mathematics and Computer Science* 26(6): 1–9.

Blackman G (2017) Contrast is king. *Imaging and Machine Vision Europe* 8: 24–27. Available at: https://www.imveurope.com/feature/contrast-king.

Boyd Junior AH (2021) Potential applications of electric color sorting techniques in seed technology. *Seed Technology Papers* 128. Available at: https://scholarsjunction.msstate.edu/seedtechpapers/128POTENTIAL.

Catão H, Rodrigues Moreira C, Caixeta L, Lopes AM, Nery-Silva FA (2020) Antioxidant activity and physiological performance of popcorn seed after saline stress and analysis of seedling images. *Ciencia e Agrotecnologia* 44: 1–9.

Cicero SM, Van der Schoor, Jalink H (2009) Use of chlrophyll fluorescence sorting to improve soybean seed quality. *Revista Brasileira de Sementes* 31(4): 145–151.

Forti VA, Cicero SM, Sliwinska E, Inomoto MM, Van der Schoor, Jalink H (2015) Meloidogyne javanica infection of soybean plants: plant response, seed quality and green seeds occurrence. *Seed Science and Technology* 43(3): 409–420.

França-Neto JB, Pádua GP, Krzyzanowski F, Carvalho MLM, Henning AA, Lorini I (2012) *Semente Esverdeada de Soja: Causas e Efeitos Sobre o Desempenho Fisiológico. – Série Sementes*. (Green soybean seed: causes and effects on physiological performance. – Seed series). Technical Report. Brazil: Embrapa.

Gomes-Junior FG, Arruda N, Marcus-Filho J (2016) *Swingle citrumelo seed vigor and storability associated with fruit maturity classes based on RGB parameters*. Scientia Agricola 0103–9016 357–363.

Hartman GL, West ED, Herman TK (2020) Crops that feed the world 2. Soybean-worldwide production, use, and constraints caused by pathogens and pests. *Food Security* 3(1): 5–17.

Lackey R (2002) *Green soybeans - Can they be fed?* Technical Report. Canada: Ministry of Agriculture, Food and Rural Affairs.

Li MW, Xin D, Gao Y, Li KP, Fan K, Munoz NB, et al. (2017) Using genomic information to improve soybean adaptability to climate change. *Journal of Experimental Botany* 68(8): 1823–1834.

Momin MA, Yamamoto K, Miyamoto M, Kondo M, Grift T (2017) Machine vision based soybean quality evaluation. *Computers and Electronics in Agriculture* 140: 452–460.

Pádua GP, França-Neto JB, Carvalho MLM, Costa O (2007) Tolerance level of green seed in soybean seed lots after storage. *Revista Brasileira de Sementes* 29(3): 128–138.

Samarah NH, Mullen RE, Goggi AS, Gaul A (2009) Effect of drying treatment and temperature on soybean seed quality during maturation. *Seed Science and Technology* 37(2): 469–473.

Sheldon A, Affleck L (1990) *Apparatus for sorting seeds according to color*. United States Patent 19, 4,946,046, p. 5.

Sinnecker P, Braga N, Macchione ELA, Lanfer-Marquez UM (2005) Mechanism of soybean (Glycine Max L. Merrill) degreening related to maturity stage and postharvest drying temperature. *Postharvest Biology and Technology* 38(3): 269–279.

Smolikova GN, Laman N, Boriskevich OV (2011) Role of chlorophylls and carotenoids in seed tolerance to abiotic stressors. *Russian Journal of Plant Physiology* 58(6): 965–973.

Smolikova GN, Dolgich E, Vikhinina M, Frolov A, Medvedev S (2017) Genetic and hormonal regulation of chlorophyll degradation during maturation of seeds with green embryos. *International Journal of Molecular Sciences* 18(9): 5–10.

Teixeira, Renake N., Ligterink W, França-Neto JB, Hilhorst HWM, Silva EAA (2016) Gene expression profiling of the green seed problem in soybean. *BMC Plant Biology* 16(1): 37.

USDA (2020) *World Agricultural production*. Foreign Agricultural Service. Available at: https://apps.fas.usda.gov/psdonline/circulars/production.pdf.

USSEC (2006) *Quality standards for U.S. soybeans and soy products*. U.S. Soy: International Buyers' Guide. Available at: http://ussec.org/wp-content/uploads/2012/08/Chap2.pdf.