RGBEXCEL: AN RGB IMAGE DATA EXTRACTOR AND EXPORTER FOR EXCEL PROCESSING

Peter A. Larbi[1,2]

[1]College of Agriculture and Technology, Arkansas State University, Jonesboro, Arkansas, USA
[2]University of Arkansas, Division of Agriculture, Fayetteville, AR

ABSTRACT

The objective of this paper was to develop a means of rapidly obtaining RGB image data, as part of an effort to develop a low-cost method of image processing and analysis based on Microsoft Excel. A simple standalone GUI (graphical user interface) software application called RGBExcel was developed to extract RGB image data from any colour image files of any format. For a given image file, the output from the software is an Excel file with the data from the R (red), G (green), and B (blue) bands of the image contained in different sheets. The raw data and any enhancements can be visualized by using the surface chart type in combination with other features. Since Excel can plot a maximum dimension of 255 by 255 pixels, larger images are downscaled to have a maximum dimension of 255 pixels. Results from testing the application are discussed in the paper.

KEYWORDS

RGB image, Image processing, Microsoft Excel, Software Development, MATLAB

1. INTRODUCTION

It is not surprising that a Google search for image processing or analysis using Microsoft Excel leads to almost no useful finds. The simple reason is that Excel is a spreadsheet program which is not immediately thought of as an option for image processing. As a spreadsheet program, Excel is used for a wide spectrum of applications from performing simple calculations to solving complex problems by both professionals and non-professionals around the world. Some examples include word processing, pilot training, data analysis, leads tracking, software creation, or even just using it for fun [1]. Because it serves as general purpose software, it offers great flexibility and may win a deeper affection by its dedicated users as an image processor.

Image processing is utilized in agriculture, medicine, engineering, and a wide range of other fields to help in feature identification or classification, comparison, and diagnoses of different conditions [2, 3, 4]. While the raw images may not be diagnostic in themselves, enhancements can be obtained using different image processing operations. The enhanced images may either remove unwanted components or enhance certain features to reveal information that are less apparent in the raw images. [2] used a method comprising image segmentation and feature

extraction and selection to successfully classify different animals. [4] recently conducted a review of different imaging techniques applied in plant phenotyping revealing quite varied applications. Whereas several image processing concepts are well established, this area is still evolving and new methods and tools are still being explored. [3] proposed, in their publication, a novel method of hybrid segmentation for magnetic resonance imaging (MRI) and applied it to pathological brain MRI.

An image is basically a two dimensional array of square pixels arranged in rows and columns, each pixel representing intensity of light sensed by the camera's detector [5]. Pixel values in a (8-bit) greyscale image ranges from 0 to 255. A colour image obtained with a digital camera is made up of three such images from the red (R), green (G), and blue (B) spectral bands of the electromagnetic spectrum. In this regard, therefore, a colour image can be considered as a multispectral image. While Excel can handle the arrays of data extracted from an image, the primary hurdle to overcome is getting the data into Excel. [6] developed a simple approach to clustering in Excel based on k-means algorithm. They were able to acquire RGB data from images using an add-in called "loadImageArray". Unfortunately, the webpage which they referred to as the source of this add-in was no longer available when this manuscript was written. [7] used a Visual Basic for Applications (VBA) macro program to visualize terrain image data and concluded that its function and quality is similar to geographic information system (GIS). However, a macro may not work properly when using it on versions other than the Excel version used in creating it.

For image processing, one can find a range of dedicated programs most of which are commercially available. Such programs can perform specialized operations more quickly, effectively, and accurately [1]. However, they also offer less flexibility compared to Excel. For instance, Excel spreadsheets hold over a million rows of data and is very useful for automating number crunching [1]. The ability to see the data itself, not just the visual representation, means that the user can also perform several custom operations on the data. In this regard, however, the MATLAB program offers even more flexibility than Excel. Nonetheless, MATLAB uses significant hard drive space, requires some level of programming experience, is not inexpensive, and may require an annual license renewal (which limits availability).

RGB images can be decomposed into their respective R, G, and B components in MATLAB. The data can then be written into an Excel file. A graphical user interface (GUI) with appropriate components provides an avenue for automating the extraction and export of the data. The MATLAB Compiler™ that comes with MATLAB R2014b enables the sharing of programs developed in MATLAB as standalone applications or shared libraries [8]. Deployment of these applications and libraries to non-MATLAB users in packaged or web downloadable forms is allowed royalty-free. Therefore, MATLAB lends itself as an appropriate platform for making RGB data available in Excel.

1.1 Objectives

The primary goal of this paper is to describe a software tool allowing RGB image data to be processed and analysed in Microsoft Excel. The specific objectives of the study were: 1) to develop a MATLAB application that will extract from any image the R, G, and B channel data and export the data to an Excel file; and 2) to deploy the application as a standalone compiled software application, thus allowing it to run on computers without a MATLAB license.

2. MATERIALS AND METHODS

2.1. Software Design, Development and Deployment

The concept behind the software application was to extract the dataset from a spatial or other image comprising R, G, and B data and export to an Excel file. The basic design of the application, henceforth referred to as RGBExcel, is shown in Figure 1. Three buttons – namely, Browse Button, Load Image Button, and Export Data Button – are required which must be used sequentially. This design was implemented in MATLAB R2014b with a graphical user interface and deployed as a standalone application for testing. The extract/export process was timed and logged at the end of each run in MATLAB.
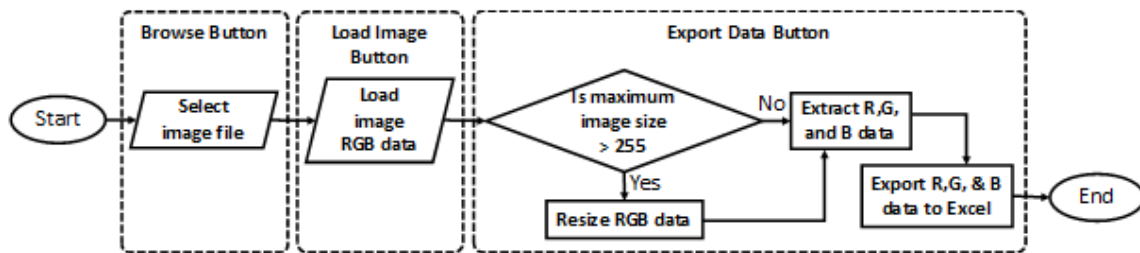

Figure 1. Conceptual design of the RGBExcel program.

2.2. Software Testing

To test the RGBExcel program, it was ran on a computer having an Intel® Core™ i7-4810MQ processor with a clock speed of 2.8 GHz speed. The application was used to extract and export RGB data from several image files of various file formats. An experiment was conducted to determine the effect of the original image size (i.e. original pixel count) on the runtime. To do this, six copies of an image (Figure 2) with a square dimension were created and used to test the application. In order to eliminate any influences from programs running on the background, five replications of tests were carried out for each image size. Summary information on the images used in this test are provided in Table 1.

Figure 2. Image of soybean plants in the field used in testing the RGBExcel program.

Table 1. Summary information on images used to test for the effect of image size on runtime spent in extract/export process of the RGB program.

Image ID	Image Size	Pixel Count	File Size
1	100 x 100	10,000	7.64 kB
2	255 x 255	65,025	38.5 kB
3	500 x 500	250,000	128 kB
4	1000 x 1000	1,000,000	397 kB
5	2400 x 2400	5,760,000	1.32 MB
6	4800 x 4800	23,040,000	3.27 MB

2.3. Data Visualization in Microsoft Excel

Data was extracted from the Figure 2 image (4800 x 4800 pixels) and exported to Excel to demonstrate Excel's potential. Visualization of such a spatial image in Excel is achieved by using a surface chart. The surface chart has four options which include 3-D surface, wireframe 3-D surface, contour, and wireframe contour, the latter two being 2-D charts. These charts have a maximum number of 255 data series per chart. The third dimension (z-data) in the 3-D charts are the pixel values in the 2-D charts which appear as different colours (0 - 255) corresponding to their signal intensities. Thus, the maximum dimension of the surface chart is 255 x 255 x 256. In order to briefly demonstrate the potential of processing the extracted data in Excel, two image processing operations involving the combination of the data from multiple bands were implemented. The first operation was the implementation of a simple ratio of G image on R image. The second was the implementation of segmentation to remove the background (soil). Since the actual methodologies involved in these operations are outside the scope of this paper, only the outputs are shown in the Results and Discussion section below.

3. RESULTS AND DISCUSSION

3.1. Running RGBExcel

The GUI of the RGBExcel program is shown in Figure 3. When the program starts, both the Load Image Data Button (henceforth called Load Button) and the Export RGB Image Data to Microsoft® Excel Button (henceforth called Export Button) are inactive (Figure 3a). The Browse button, when clicked, first checks the existence of the reserved folder ('C:/RGBEXCEL Files') for files created by the program and creates it if it does not exist. If the folder already exists and contains files from a previous extract/export process, the files are deleted. After these steps, a dialog box opens for the user to select the target image file of JPG, JPEG, JPE, JFIF, GIF, TIF, BMP, or PNG format. Once this is done, a copy of the target image file is placed in the program reserved folder and the Load Button becomes active (Figure 3b).

The Load Button implements the MATLAB 'imread()' function which loads an m x n x 3 array data of the image copy (where m and n are the spatial dimensions and the '3' consists of the R, G, and B layers). The data is saved and the Export Button is activated (Figure 3c). The Export Button first checks the maximum dimension of the RGB data. If the maximum dimension is greater than 255, the MATLAB 'imresize()' function is implemented to scale it down before

separately extracting 2-D arrays of the R, G, and B data. A new Excel file is created with a reserved name ("imagedata.xlsx") and, using the MATLAB 'xlswrite()' function, the data are written into three sheets named respectively as R, G, and B.
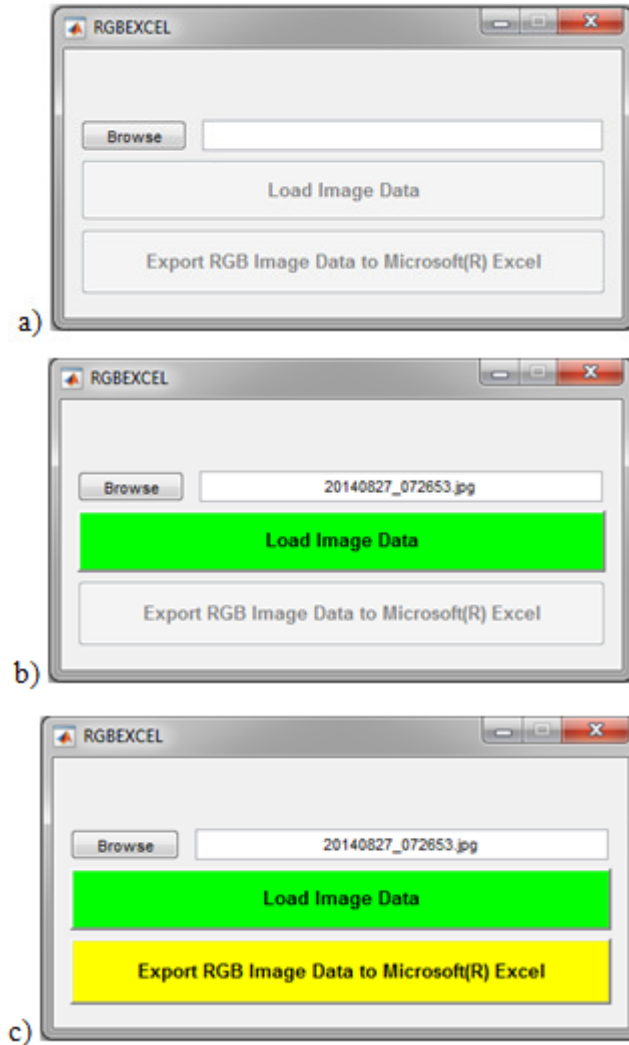

Figure 3. GUI of the RGBExcel program.

3.2. RGBExcel Test Results

The results indicate that image dimension correlates well with file size in a power relationship (Figure 4). Therefore, file size will have a similar relationship with runtime as image dimension. According to Figure 5, the pixel count in the images influenced the runtime spent in extracting and exporting RGB data into Excel. However, there were very little differences in runtime between the images with dimensions of 255 x 255, 500 x 500, 1000 x 1000, and 2400 x 2400 pixels. There was an increase in runtime of about 100% from the image with 2400 x 2400 pixels to the image with 4800 x 4800 pixels. Apart, from the image with 4800 x 4800 pixels, all the other images tested resulted in a runtime of less than 1 s.
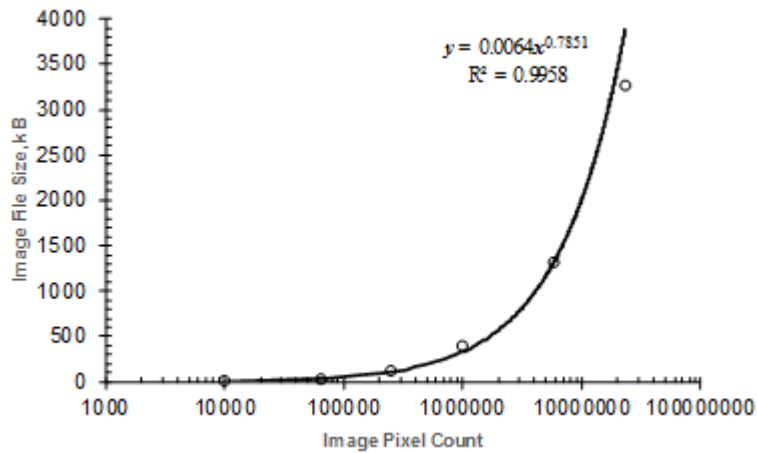
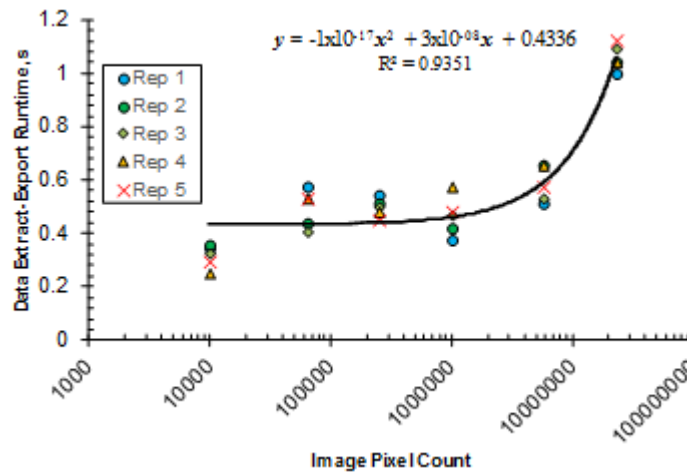Figure 4. Relationship between image dimension and image file size

Figure 5. Runtime used in extracting-exporting RGB data to Microsoft Excel

The RGB data of the test image with 4800 x 4800 pixels were extracted and exported to Excel using the RGBExcel program. The data was plotted using the contour chart type and applying the appropriate colour scheme. The results are shown in Figure 6 with legends showing the range of pixel values corresponding to various colours. The data from different bands can be combined to generate enhanced data which can also be visualized in a similar way as depicted here, using different colour schemes. Figure 7 shows an enhanced image obtained by computing a simple ratio of the green image over the red image. This enhancement eliminates the differences in illumination created by the shadows in the original image. Figure 8 is a further enhancement of Figure 7 with the background (soil) removed.
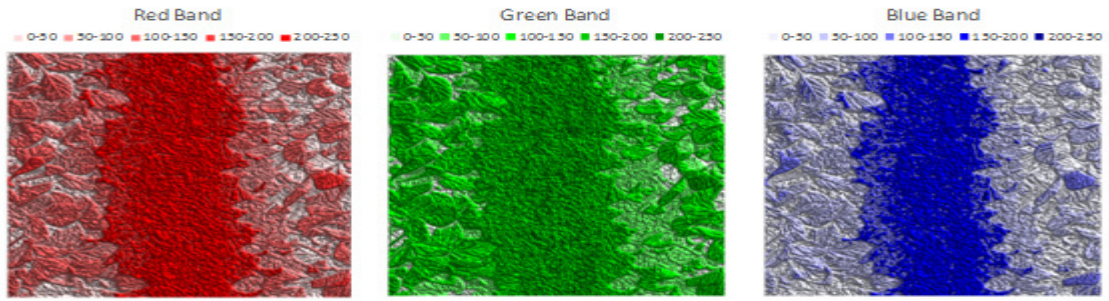
Figure 6. R, G, and B images of Figure 2 visualized in Excel after exporting the data with the RGBExcel program
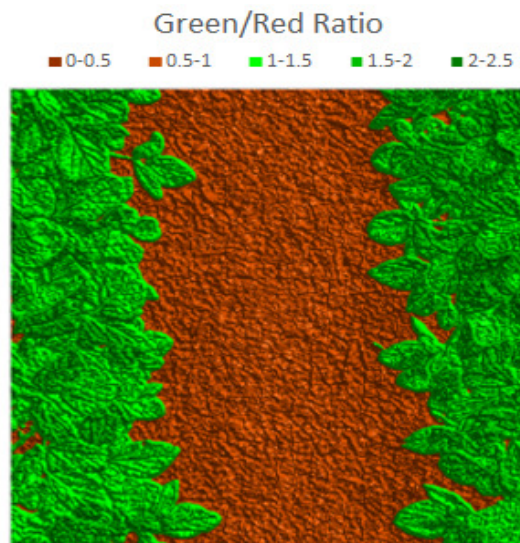

Figure 7. A simple ratio of G and R images in Figure 6 obtained in Excel. This enhanced image eliminates the differences in illumination created by the shadows in the original image.
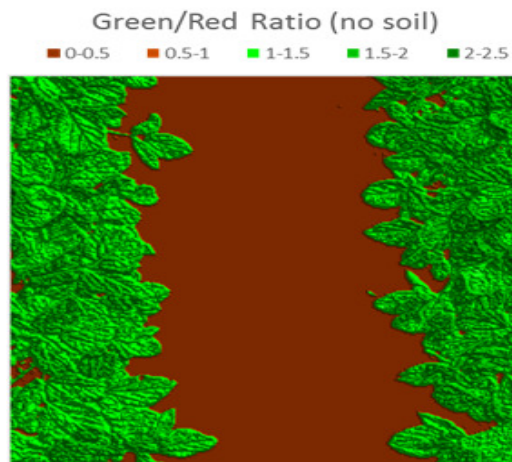

Figure 8. A simple ratio of G and R images in Figure 6 with the background (soil) removed using image segmentation operation.

4. CONCLUSIONS

In this paper, the potential of Microsoft Excel for processing and analysing RGB colour images has been demonstrated. This potential was proved through the development of the RGBExcel software application that extracts and exports RGB data from any image file of JPG, JPEG, JPE, JFIF, GIF, TIF, BMP, or PNG format to Excel. The runtime required for this extract/export operation was shown to be under 1 s for images less than 4800 x 4800 pixels. While this rapid runtime speed is positive, there is still the need to run the program once for every image to be processed in Excel. This is a limitation especially when needing to process a large number of images. Once in Excel, however, the raw data and any enhancements can be visualized by using the surface chart type in combination with other features. This visualisation of the data in Excel has also been demonstrated in the paper.

5. FUTURE WORK

This work is part of an effort to develop a low-cost method of image processing and analysis based on Microsoft Excel. Future direction, therefore, will include improving the RGBExcel software application as well as developing step-by-step methods for accomplishing different image processing and analysis operations.

ACKNOWLEDGEMENTS

REFERENCES

[1]   Bradley, Helen, (2011) "Use Microsoft Excel for (Nearly) Everything", PCWorld Digital Magazine: http://www.pcworld.com/article/220782/use_microsoft_excel_for_everything.html.

[2]   Kumar, Y.H. & Divya, C.D., (2014) "Feature Selection Approach in Animal Classification", Signal & Image Processing: An International Journal, Vol. 5, No. 4, pp55-66.

[3]   Abderrezak, M.Z., Chibane, M.B. &   Mansour, K., (2014) "A New Hybrid Method for the Segmentation of the Brain MRIs", Signal & Image Processing: An International Journal, Vol. 5, No. 4, pp77-84.

[4]   Li, L. Zhang, Q., & Huang D., (2014) "A Review of Imaging Techniques for Plant Phenotyping", Sensors, Vol. 2014, No. 14, pp20078-20111.

[5]   Liew, Soo Chin (2001) "Principles of Remote Sensing", Space View of Asia CD-ROM Tutorial, 2nd Edition, Centre for Remote Imaging, Sensing and Processing, National University of Singapore. Available at: [http://www.crisp.nus.edu.sg/~research/tutorial/image.htm]

[6]   Aravind, H., Rajgopal, C., & Soman, K.P., (2010) "A Simple Approach to Clustering in Excel", International Journal of Computer Applications, Vol. 11, No. 7, pp19-25.

[7]   Lin, Chih-Chung & Lin, Yen-Ling (2009) "Apply Excel VBA to Terrain Visualization", The 22th
        IPPR Conference on Computer Vision, Graphics and Image Processing, pp1707-1711.

[8]   Mathworks (2014) "MATLAB® Compiler™ User's Guide", 22 pp. Available at:
        [http://www.mathworks.com/help/releases/R2014a/pdf_doc/compiler/compiler.pdf].

AUTHOR

Dr. Peter Ako Larbi joined the College of Agriculture and Technology, Arkansas State University, in August 2014 as an assistant professor of Agricultural Systems Technology, with a joint appointment with the Division of Agriculture, University of Arkansas. Prior to this, he had about three-and-a-half years of postdoctoral research experiences from University of Florida's Citrus Research and Education Center in Lake Alfred, FL and Washington State University's Center for Precision and Automated Agricultural Systems in Prosser, WA. He has a bachelor's and a master's degrees in Agricultural Engineering from the Kwame Nkrumah University of Science and Technology in Ghana and a Ph.D. in Agricultural and Biological Engineering from the University of Florida. His teaching interests include Modern Agricultural Systems, Remote Sensing, Modern Irrigation Systems, and Precision Application Technology. His research interests include precision agriculture technologies, remote sensing, and agricultural machinery automation.