An Algorithm for Plant Diseases Detection Based on Color Features

MOSBAH EL SGHAIR
John Naisbitt University
Graduate School of Comp. Sci.
Bulevar umetnosti 29, Belgrade
SERBIA
musbah.bellid@gmail.com

RAKA JOVANOVIC
Hamad bin Khalifa University
Qatar Env. and Energy Research Inst.
PO Box 5825, Doha
QATAR
rjovanovic@hbku.edu.qa

MILAN TUBA
John Naisbitt University
Graduate School of Comp. Sci.
Bulevar umetnosti 29, Belgrade
SERBIA
tuba@ieee.org

Abstract: Digital images enabled great improvements in many different areas and biology and agriculture are some of them. Plant diseases detection and classification is very important task and further more automatic detection of plant diseases is an important analysis topic because it could help in observation of giant fields of crops, and therefore mechanically find the symptoms of diseases as shortly as they appear on the plant leaves. In this paper we proposed an algorithm for automatic detection of plant leaf diseases. The proposed method consists of several steps including median filter, thresholding, and uses of different color models for segmentation. Our proposed method provides quicker and more accurate detection and classification compared to other state-of-art.

Key–Words: CIELAB, HSI, YCbCr, plant leaf disease detection, image thresholding, Kapur's method

1 Introduction

Digital images became part of everyday life. Digital image processing and analysis represents a rather popular scientific research topic. It was used in many different fields such as astronomy [1], biology [2], medicine [3], [4], etc. Increasing use of digital images resulted with numerous algorithms for image processing that are used in various applications such as multilevel thresholding [5], [6], JPEG compression [7], handwritten digit recognition [8], leaf recognition [9], lip detection [10], etc.

One of the areas where digital image analysis brought a great benefits is agriculture. Plant diseases have become an important issue because they cause important reduction in each quality and amount of agricultural products. Large amount of money is spent every year all around the world for efficient plant disease detection while the financial lost in case of late disease detection is 3, 4 times larger. One of the methods of plant disease detection and identification is an eye observation by the specialists is that. However, this needs continuous observation of specialists which could be time and financially expensive in giant farms. Further, in some developing countries, farmers could have to be compelled to go long distances to contact specialists which also is not practical and cheap. Automatic detection of plant diseases is an important analysis topic because it could significantly help in observation giant fields, and enable automatic detection the symptoms of diseases as soon as they appear on the plant leaves. Therefore, searching for quick, automatic, more cost-effective and correct technique to find disease is of the great significance. Machine learning primarily detection and later recognition of plant diseases will offer clues to spot and treat the diseases in its early stages. Also, as it was explained earlier visually distinguishing plant diseases can be inefficient and troublesome since it needs the experience of trained plant scientist. Some researchers have used image processing techniques for quick and correct detection of plant diseases. The accuracy of result depends on technique used for disease spot detection. The biggest obstacle in disease spot detection is noise, that is introduced by camera flash, modification in illumination, droning background and presence of vein within the plant leaf. Thus a technique that wipes out the noise and provides higher disease spot segmentation is required.

Veins color is the same as plant leaf color solely intensity differs. On the other hand, disease spot color is completely different from plant leaf color. Thus if image is first transformed to some color model that has intensity component threshold can be applied on it and this is one of the main techniques that are applied in plant diseases detection applications.

In this paper adjusted median filter is proposed for noise reduction. One of the well known threshold technique, Kapur's method is employed to convert filtered image into binary image Mistreatment on top of techniques disease spot is detected in totally different plants, during which largely veins are parallel and fewer visible. The proposed algorithm was tested by using different color models such as RGB, HSI and CIALAB.

The rest of the paper is organized in four sections. Our proposed algorithm is defined and described in Section 2 Experimental results are showed in Section 3. Conclusion of the paper is given in the last section.

2 Our proposed algorithm

In this paper we propose a method for plant diseases detection based on the leaf images. The proposed method contains three main steps. The first one is image transformation into different color models, the second is noise reduction and the last one is image segmentation where disease spots are detected.

It was assumed that images of the plant leafs are available. Images can be obtained by putting cameras in the fields or satellite images can be used. All the images used in this paper are in JPEG format. These images are in RGB color model. The proposed method will be tested when the image is transformed into YC_bC_r, HIS and CIELAB color models. The color converted images are gone through median filter to get rid of spare spots and reduce the noise. In the last step Kapur's method is applied on different color components depending on the used color model. A element of CIELAB color area was used, H element of HSI color area and C_r element of YC_bC_r color area are employed to discover the disease spot. Diseases detection is obtained by all the three strategies and compared to induce the most effective methodology for disease spot detection.

In plants, leaf vein is totally different in intensity and disease spot is different in color compared to plant leaf. Therefore if Kapur's method is applied on grayscale image, vein will be detected in binary image with the disease spot. However the region of interest is simply disease spots, not vein. For minimize the effect of presence of vein, RGB color model is not suitable for segmentation. Thresholding method are often applied on color element to discover disease spot accurately.

The first color model that will be used is YC_bC_r. This color model is wide employed in digital video. In YC_bC_r color model, Y indicates brightness level element and C_b, C_r indicate color elements. C_b represents distinction between the blue elements while C_r is that the distinction between the red elements. This color model was often used in literature for different applications such as face detection [11], blood cell segmentation [12], etc. Using following formulas RGB image can be transformed into YC_bC_r color model.

$$\begin{aligned} Y &= 0.299 * R + 0.587 * G + 0.114 * B \\ C_b &= -0.168 * R - 0.331 * G + 0.500 * B \\ C_r &= 0.500 * R - 0.418 * G - 0.081 * B \end{aligned} \qquad (1)$$

Color model HSI is device dependent color model and primarily based upon human color perception. In this color model H indicates hue, that describes a pure color and is mostly associated with the wavelength of the light. Component S indicates saturation, that measures the colorfulness in HSI color model while I indicates intensity, that shows the amplitude of the light. The first step in conversion is to represent the RGB components in the range $[0, 1]$ by dividing each pixel value for each component by 255. Image further from RGB color model can be transformed into HSI model by the following equations:

$$H = \begin{cases} \theta & \text{if } B \leq G, \\ 360 - \theta & \text{if } B > G, \end{cases}$$

where θ is:

$$\theta = cos^{-1}\{\frac{\frac{1}{2}[(R-G)+(R-B)]}{[(R-G)^2+(R-B)(G-B)]^{1/2}}\}$$

$$S = 1 - \frac{3}{R+G+B}[\min(R, G, B)] \qquad (2)$$

$$I = \frac{1}{3}(R+G+B) \qquad (3)$$

CIELAB system is device independent color model that is outlined by the CIE to classify color consistent with the human vision. Within the conversion method of an image from RGB color element to CIELAB color element, first RGB image is transform into CIEXYZ by the following equation:

$$\begin{aligned} X &= 0.4124 * R + 0.3576 * G + 0.1805 * B \\ Y &= -0.2126 * R + 0.7152 * G + 0.7220 * B \\ Z &= 0.0193 * R + 0.1192 * G + 0.9505 * B \end{aligned} \quad (4)$$

Brightness and color data of research lab color model is independent of every different. In CIELAB color model, L describes color brightness, A describes the color starting from green to red while B describes the color starting from blue to yellow. Conversion formula for research laboratory color model is defined by the following equation:

$$\begin{aligned} L &= 116 * f(\frac{Y}{Y_n}) - 16 \\ A &= 500 * (f(\frac{X}{X_n}) - f(\frac{Y}{Y_n})) \\ B &= 200 * (f(\frac{Y}{Y_n}) - f(\frac{Z}{Z_n})) \end{aligned} \quad (5)$$

where function f is defined as:

$$f(t) = \begin{cases} \sqrt[3]{t} & \text{if } t > \beta^3, \\ \frac{t}{3\beta^2} + \frac{4}{29} & \text{otherwise} \end{cases}$$

where $\beta = \frac{6}{29}$. Constants X_n, Y_n and Z_n are equal to 95.047, 100.000 and 108.883, respectably.

2.1 Image Smoothing

The next step in our proposed algorithm is image smoothing. During image assortment, some noise is also introduced due to camera flash. This noise might have an effect on the detection of disease. To remove unneeded spots, image smoothing technique is required. In this paper adjusted median filter is employed for this purpose.

Median filter can be a higher order statistics filter. Median filter is nonlinear in nature and replaces the value of the central pixel by the median of the gray levels within the image area surrounded by the filter.

So in order to perform median filtering, first, window is rapt and the pixels enclosed by the window area are sorted. When median is computed, it is assigned to the center pixel. The number of the pixels in $K \times K$ window is odd. In this paper size of the window was empirically determined to be 5×5. Smaller window size was not able to remove some larger noise spots, while the larger window size cause damaging clear parts of the image.

2.2 Disease Detection by Kapur's Method

After image smoothing, a method to detect and isolate the disease spot is required. It is necessary to find a threshold value that will differentiate the disease spots from plant leaf. If the histogram has sharp and deep depression between two peaks, bottom of the depression may be chosen as threshold. However the problem is when depression does not exists of it is flat. In such case this method can not be accustomed to separate objects from background. One of the most used method for thresholding is Kapur's method that is based on the entropy. This method maximize the amount of information between the two parts of a intensity histogram that are separated by concrete threshold value or better to say maximize the entropy measure of the part of the histogram in order to each part has a more centralized distribution.

In this paper, Kapur's method was used for diseases spot detection. Threshold value was searched for different components. In case of RGB image, threshold value was searched for gray level image obtained by averaging all pixels components. If HSI model was used, threshold value for H component was searched while in the case of the YC_bC_r model, Kapur's method was applied to C_r component. When CIELAB color model is used, Kapur's method was used to find optimal threshold value for component A.

3 Experimental results

In this paper, first was experimented with images in RGB model. Disease spots are detected by applying Kapur's threshold method on gray scale image that is result of averaging pixels components. In Fig. 1 are shown experimental results for one example. It is obvious that plant diseases were not recognized successfully. As it was assumed, RGB color model is not suitable for detection disease spots in plant leafs.

To improve detection further experiments were done. In the second technique RGB image is initial transformed into YC_bC_r color model by using color transform formula previously described. Then median filter is employed for image smoothing. Disease spots are detected by applying Kapur's threshold on C_r component of filtered YC_bC_r color image. Experimental results are shown in Fig. 2. Diseases were detected more successful in this case comparing to the detection in RGB model. When YC_bC_r model was used detections of the spots was larger then actual diseases. This means that some parts of the leafs without diseases were recognized as infected. False disease

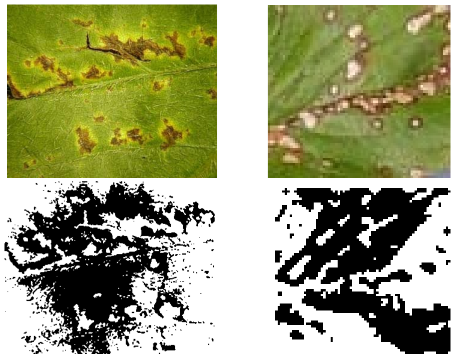

Figure 1: Plant diseases detection in RGB image

detection is not good characteristic of the application.

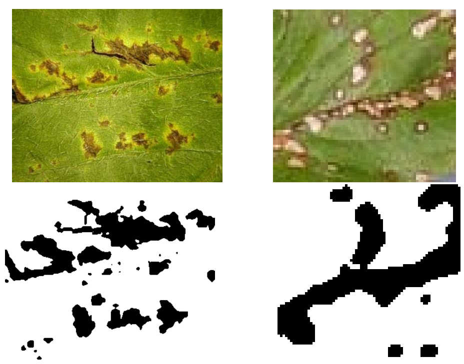

Figure 2: Plant diseases detection using YC_bC_r color model

In next experiment RGB image was transformed into HSI color model and disease spots were detected like before by applying Kapur's threshold but on H component of filtered HSI color area. Detections of plant diseases by our proposed method by HSI color model are shown in Fig. 3. As it can be seen, diseases were mostly discovered correctly. Comparing to the previous examples this recognition was the most precise. Some not infected parts are still marked as a disease, which is visible in the second test example. The first example was almost completely correctly marked.

The last color model that need to be tested is CIALAB. Once more the proposed algorithm was used but with images in CIELAB color model. Disease spots are segmental by applying Kapur's threshold on component A of filtered laboratory color area and the results are presented in Fig. 4. It can be noticed that the diseases were not completely recognized. This leads to the conclusion that this color model is not appropriate for early disease detection which is one of the goals of the proposed algorithm.

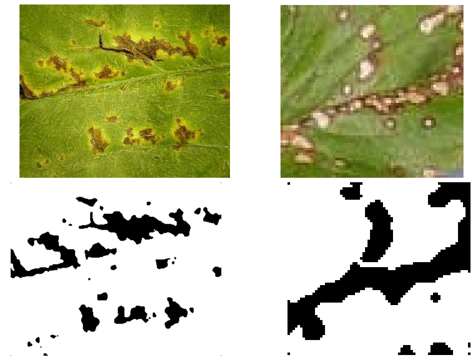

Figure 3: Plant diseases detection using HSI color model

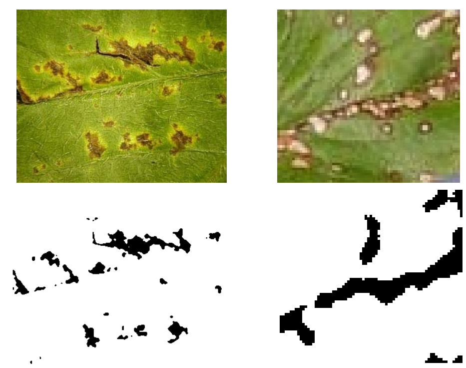

Figure 4: Plant diseases detection using $CIELAB$ color model

Based on the experimental results some conclusions can be made. We tested the proposed algorithm and used different color models. Results can be summarized as the follows:

- Using threshold on RGB image disease spot is almost unusable.
- Using threshold on A element of CIELAB color model and C_r element of YC_bC_r color model, disease spots are detected however not correctly. Therefore, results are not satisfying.
- Results show that exploitation threshold on H component of HSI color model disease spots are detected accurately and results better comparing to the other tested in this paper.

Figure 5: Diseases detection with different color models on images with the background

Next, experiments included two additional images of the plant but with the background. In these images noise was introduced by adding the background that can disturb diseases detection. Experimental results for disease spot detection of some leafs suffering from different diseases are shown in Fig. 5. In the first row are the original images. In the second row are shown results of our proposed method by using RGB color model, in the third YC_bC_r model, in the forth row are shown detection when HSI color model was used and at the end, in the fifth row are results when CIELAB color model.

Based on this results, it can be concluded that exploitation threshold on RGB image, disease spots are not detected and disturbance due to vein is present. Using threshold on C_r element of YCbCr color model, some disease spots area unit detected effectively, but again in the first two examples more parts of the leaf were recognized as disease than they really are. Using threshold on H element of HSI color model diseases were quite correctly recognized. Comparing to the previous color models, it is far the best one. Finally, by using threshold on A element of CIELAB color model disease spots may be detected accurately altogether. Obtained results are comparable to the one obtained with HSI color model.

Result shows that disease spots might be detected accurately using HSI color model and CIELAB showed also promising results.

4 Conclusion

In this paper a method based on different color models and Kapur's thresholding for plant diseases detection was proposed. Four different color models were tested and compared: RGB, YCbCr, HSI and CIELAB color model. The best results were obtained when HSI color model was used. Component H was used for image segmentation where diseases were separated from the leaf. Median filter was applied to color transformed image. At the end, disease spots area are determined by applying Kapur's threshold on different color components. Experimental result shows that noise that is introduced due to background, vein and camera flash makes the least problem for HSI color model. Following this technique totally different disease spots are detected accurately and results do not seem to be laid low with background, sort of leaf, type of disease spot and camera. In the further work, disease may be classified by calculative dimensions of disease spot.

Acknowledgments: This research is supported by Ministry of Education, Science and Technological Development of Republic of Serbia, Grant No. III-44006.

References:

[1] C. Tsiotsios and M. Petrou, "On the choice of the parameters for anisotropic diffusion in image

processing," *Pattern recognition*, vol. 46, no. 5, pp. 1369–1381, 2013.

[2] G. Stevenson, A. Welsh, L. Impey, J. A. Noble, and S. Collins, "An image processing technique for the visualization and quantification of blood flow entering the placenta using 3d power doppler ultrasound (pd-us)," *Placenta*, vol. 45, p. 85, 2016.

[3] M. Nikolic, E. Tuba, and M. Tuba, "Edge detection in medical ultrasound images using adjusted Canny edge detection algorithm," in *24th Telecommunications Forum (TELFOR)*. IEEE, 2016, pp. 691–694.

[4] A. Stojak, E. Tuba, and M. Tuba, "Framework for abnormality detection in magnetic resonance brain images," in *24th Telecommunications Forum (TELFOR)*. IEEE, 2016, pp. 687–690.

[5] A. Alihodzic and M. Tuba, "Bat algorithm (BA) for image thresholding," *Recent Researches in Telecommunications, Informatics, Electronics and Signal Processing*, pp. 364–369, 2013.

[6] I. Brajevic and M. Tuba, "Cuckoo search and firefly algorithm applied to multilevel image thresholding," in *LNCS, Cuckoo Search and Firefly Algorithm: Theory and Applications*, X.-S. Yang, Ed. Springer International Publishing, 2014, vol. 516, pp. 115–139.

[7] M. Tuba and N. Bacanin, "JPEG quantization tables selection by the firefly algorithm," in *International Conference on Multimedia Computing and Systems (ICMCS)*. IEEE, 2014, pp. 153–158.

[8] E. Tuba and N. Bacanin, "An algorithm for handwritten digit recognition using projection histograms and SVM classifier," in *23rd Telecommunications Forum Telfor (TELFOR)*. IEEE, 2015, pp. 464–467.

[9] M. Lukic, E. Tuba, and M. Tuba, "Leaf recognition algorithm using support vector machine with Hu moments and local binary patterns," in *15th International Symposium onApplied Machine Intelligence and Informatics (SAMI)*. IEEE, 2017, pp. 000 485–000 490.

[10] A. Arsic, M. Jordanski, and M. Tuba, "Improved lip detection algorithm based on region segmentation and edge detection," in *23rd Telecommunications Forum (TELFOR)*. IEEE, 2015, pp. 472–475.

[11] R. C. Hrosik, M. Tuba, and M. Vukovic, "Face detection algorithm based on skin detection and invariant moments," in *Recent Advances in Knowledge Engineering and Systems Science*, 2013, vol. 10, pp. 110–115.

[12] N. Abbas, D. Mohamad *et al.*, "Microscopic RGB color images enhancement for blood cells segmentation in YCbCr color space for k-means clustering," *Journal of Theoretical and Applied Information Technology*, vol. 55, no. 1, pp. 117–125, 2013.