Classification of Gerbera Type Flowers Based in Decision Tree Rules

Jhon Edwin Vera Vera
Vice-Rector of Research
Universidad ECCI
Bogotá, Colombia
jverav@ecci.edu.co

Sergio Mora Martinez
Vice-Rector of Research
Universidad ECCI
Bogotá, Colombia
smoram@ecci.edu.co

Andrés Torres Pérez
Dept. Electronic Engineering
Universidad ECCI
Bogotá, Colombia
atorresp@ecci.edu.co

Jonathan Avendano
Dept. Electronic Engineering
Universidad ECCI
Bogotá, Colombia
javendano@ecci.edu.co

Abstract*—In this work, a 4 Gerbera flowers subtypes are classified according to their color components in petals and in the center. We proposed first the preprocessing of each image acquired in real scenario, the preprocessing is realized in two steps, the first is to crop each flower to create a database with the tag of each subtype, the second consist in removing the background using color space transformation to Hue and filter the image according to some specific values. With all the images a decision tree is created with the 70% of the images in each category, the first rule consists of identify the highest value in the histogram, if this value corresponds to green range or orange range the classification is done, nevertheless if the value corresponds to pink, the enclosing circles are estimated, then the color into the smallest circles defined the classification. The evaluation is performed using the last 30% of the images in each category, the classification will be positive if the correct subtype is predicted, if not the classification will be negative, nevertheless the classifier could predict no class based on any existence of the flower, or different flower, in this case if non-class is predicted over a different flower the classification will be positive, if not the classification will be negative. The results obtained in each category are showed in the next list, first is the name of each subtype, followed by the percentage of positive classifications. Gerbera Renato (95%), Gerbera Marinilla (99%), Gerbera Chiper (85%), Gerbera Rio Negro (78%), None/Different (74%).

***Keywords*—Machine Vision, Flower, Classification, Decision Tree.**

I. Introduction

In the artificial vision field, the agriculture has been widely studied with the objective of automating different processes such as the crop estimate by a count of fruits; estimate the capacity of a plant to receive sunlight, measuring water requirements, excess, or deficiency of nutrients from the plant, a measurement of its foliage; and indicate the stress on the plant through measuring the stem and branches [1]. For its part, the classification of flowers is a common application in the field of machine vision and image processing. Through the development of these applications, we can make various processes of agriculture focused on flowers such as the automatic harvest and crop estimate and greenhouse production. Different works have been presented with the aim of classifying flowers; however, this is a relevant topic because there is still no universal classification algorithm because of many existing classes and varieties of flowers.

In [2], a work for the classification of flowers was presented. First, a background subtraction was applied and a feature extraction was performed based on the SURF (Speeded Up Robust Features) descriptor's method where they used the shape, color, and texture of the flowers. Finally, a multiclass SVM (Support Vector Machine) was trained with 88% accuracy on a dataset of 13 flower types. Further, in [3] the automatic classification of 8 types of flowers was presented. Feature extraction was based on SIFT (Scale Invariant Feature Transform) descriptors and segmentation using SFTA (Segmentation-based Fractal Texture Analysis). Again, in the first instance, a background removal process was performed. The extracted features were used to train two types of classifiers, SVM, and RF (Random Forest). The authors showed that flowers have different forms reach 100% precision using different features and classifiers. Finally, in [4] a system to classify orchid using a limited set of training data was presented. LBP features and AdaBoost learning were used. The classification was solved by training a binary tree of linear SVM classifiers on a limited set of training data. The authors found that even using a set of training data limited, both for the detection model, and training SVM classifiers, the result was quite impressive compared to the current research in object detection, where thousands of training samples are used to reach robust classifiers. The reason is based on the authors used controlled lighting conditions and the known configuration was used (camera position, it is possible to background, flowers of orchids).

In contrast to previous work, in [5] a model to classify flowers using neural networks was proposed. Two types of features were used to train the neural network, color, and texture. For the color features the authors used the normalized histogram in the HSV color space, and for texture used gray-level co-occurrence matrix. In this work, 19 varieties of flowers were classified, obtaining accuracies near 69%. In [6], a system for the classification of 10 types of flowers is presented. The approach presented in this paper uses a K-

978-1-7281-1491-0/19/$31.00 ©2019 IEEE

nearest neighbor classifier. The features used Hu's seven-moment algorithm for edges detection and spaces RGB and HSV color, obtaining accuracies of 80% for the classification task.

The flowers depending on the variety can be similar between different types of flowers, as for *gerberas*, where in some varieties only change the color and shape of the center of the flower, while the color and shape of the petals are similar. In [7] a work with the aim of generating an algorithm to train expert systems that classify types of similar flowers, using a lower number of training images is presented. The training data were generated artificially using low-resolution images and multiple mapping techniques.

The methods for classifying first perform a background removal, then a feature extraction is performed and finally a supervised classifier is trained. Further, we can evidence that most proposed approaches used color, shape, and texture features, and petal count etc. However, we can evidence that by using color information, the accuracy is high only if the colors of the flowers differ. For its part by using shape information and counting petals present problems related to the perspective from which the images were acquired. This work proposes the development of a system that classifies 4 types of gerbera using information of the petals and the center of the flower. Data were acquired in the field, under different lighting conditions, and from different viewpoints. The automatic classification of flowers using machine vision is a good alternative in agriculture to optimize the production processes and supervision of crops in a greenhouse.

This paper is organized as follows: Section 2 presents the Pre-processing data step, Section 3 presents the algorithm proposed, and in Section 4 show the results and discussion of the proposed method. Finally, Section 6 presents the key conclusions.

II. Pre-Processing Data

The database consists of 149 images taken with the same camera in a local garden center, due to the roof of this place the light condition will be considered uniforms for all taken images. The flower images have been cropped and manually classified from the 149 raw images, an example of each class is shown in the Fig. 1. This image shows in each row the name of the category and 3 examples of the flower cropped from the original raw image.

Table 1 – Range of colors in Hue channel for each RGB color

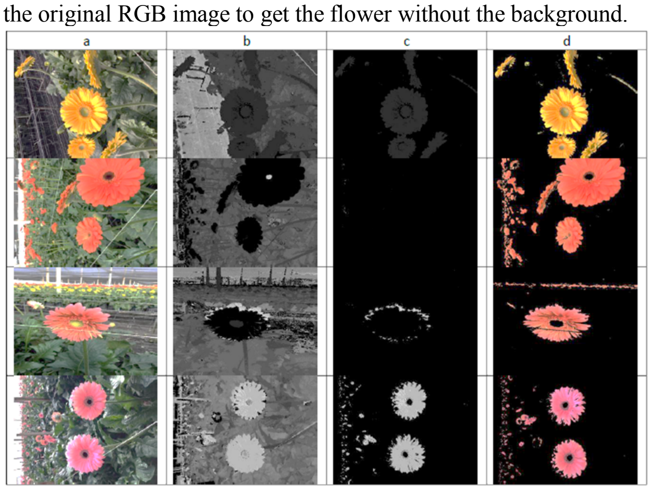
Color	Hue	
	Min.	Max.
Orange	0	0.122
Yellow	0.123	0.211
Green	0.212	0.416
Blue	0.417	0.722
Violet	0.723	0.888
Red	0.889	1.0

Fig. 1. Examples of each category in our database

III. Algorithm

The algorithm for classification starts with the background subtraction function, this particular function takes the image and transform the color space from RGB to HSV, this technique has been widely used to filter specific colors in image processing [8][9]. The color space HSV (Hue, Saturation, and Value) uses a nonlinear color space transformation from RGB, the Hue (H) is a value where the color is defined in several ranges, and these values are presented in the Table 1. The filter removes all green values from the Hue Image, this image is used as a mask to filter the color RGB image, all the process for each category is shown in the Table 2. The background subtraction allows to identify the flower petal color, this value is the first feature to classify the image. The Table 2 shows the final color image filtered, as the color feature comes from the Hue image, the filtered image is only demonstrative to illustrate background subtraction.

Table 2– Steps in the background subtraction in the petal color. In the image a) the original RGB image is shown, then it is translated to HSV color-space, in b) the Hue channel is presented. The green filter is applied and the result is the image c), finally this mask is used in the original RGB image to get the flower without the background.



The classifier takes the remain histogram in Hue after green filter is applied, as the Hue value is a single value, the image has one channel, and the histogram has one dimension. The histogram for each class of the image is shown in the Fig. 2, the x-axis corresponds to the value of Hue from 1 to 360 (scaled from the 0 to 1 in the Table 1), the y-axis corresponds to the number of pixels in the image that has the corresponding value. The value 0 has always the bigger number of pixels due to the background subtraction, this value is irrelevant so it is removed to avoid any bias that could damage the classification process.

The center of the chart in Fig. 2 is void, due to 2 reasons, the first is that the values from 76 to 150 are removed in the background filter, and the second is that values from 150 to 260 correspond to blue color, which has low levels in all our images., nevertheless, this range is shown in order to define the first rule for the classifier.

The Fig. 3 shows that 2 classes are separable using a single condition from Hue value, those classes are *Gerbera Marinilla* and *Gerbera Renato*, the condition is defined as:

Rule N°1:

- If Hue value is between 0 and 25, the class is Gerbera Marinilla
- If Hue value is between 26 and 55 the class is Gerbera Renato
- If Hue value is between 325 and 360, continue to next Rule.
- Otherwise, no class is the result.

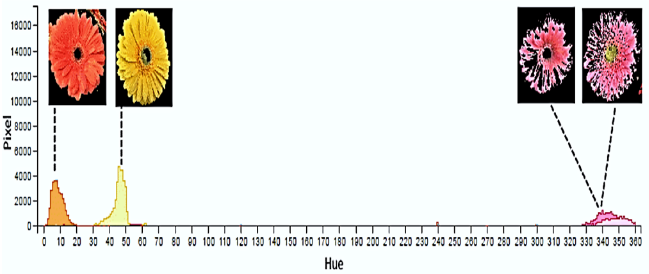
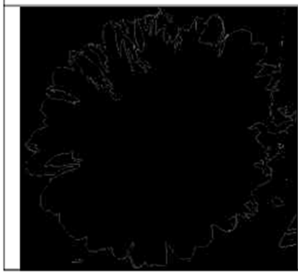


Fig. 2 – Hue histogram for each class of flower, the histogram is calculated over the extracted flower to the algorithm presented in chapter 3.

In Fig. 2 it can be seen that two types of gerberas have HUE values in their petals, therefore, another rule should be defined considering the texture of each flower and the color of its center. The center color is the naïve approximation due that color features have been working over all this project, the color center is green for *Gerbera Chiper* and Black for *Gerbera Rio Negro*, due to the background subtraction the flower have issues in the center color, then a green center could be detected as black, then from the extracted flower image an enclosure is estimated to get all colors inside the enclosure. The enclosure uses the function findContours in OpenCV 2.4, this function is implemented using the algorithm in [10], and then the contour is filled and used as mask to obtain the flower with the center as is shown in the Table 3.

Table 3 – Image cropped from the enclosure mask with the color center

Mask	Cropped
	

With the color center the histogram could have a value in the range 76 to 150, so the second rule is defined as:

Rule N°2:

- If Hue value is between 76 and 150, the class is Gerbera Chiper
- Otherwise, the class is Gerbera Rio Negro

To summarize the tree rule decision a diagram is shown in the Fig. 3. In this diagram the average of the Hue value in petals histogram is denoted as H', and the average of the second value of Hue in the histogram of the enclosure mask is denoted as C'.

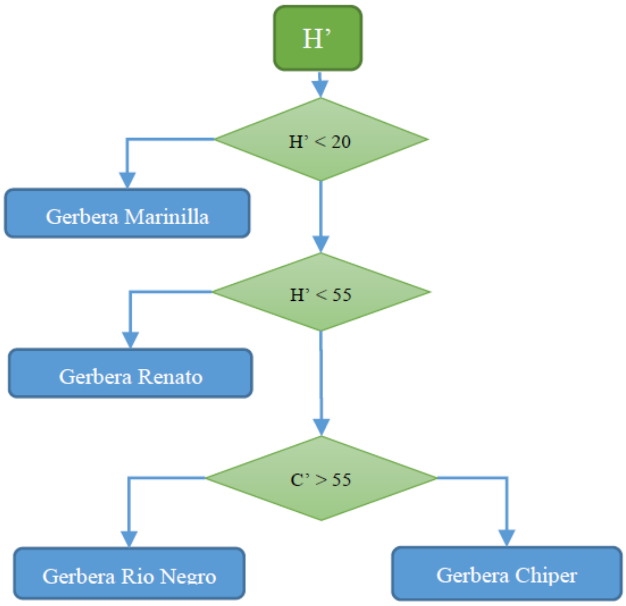


Fig. 3 – Tree rule decision diagram for flower classification.

IV. Results

The algorithm exposed in the section 3 is tested against the database created from raw images in section 2. There is a class that corresponds to *none/different*, to test this category 54 images from other types of flowers are classified with the rules presented in this work. The parameter measured in this test is the correct classification of each labeled image of a single flower, the purpose of this test is to verify the accuracy of the algorithm proposed.

The results are shown in Table 4, the first 2 classes have a high accuracy due to their separability in Hue value, as it was presented in Fig. 3, the error in this flower comes in the range values 20-30, where the flower *Gerbera Marinilla* has a more yellow component and the flower *Gerbera Renato* has more orange component. The next 2 classes presented at the same value in Hue histogram for petals, then the Hue value in the center is used to determine the class obtaining a significantly high accuracy. Nevertheless, the *Gerbera Chiper* flower tends to be higher than *Gerbera Rio Negro*, this is explained for some green textures over the petals in some *Gerbera Rio Negro* flowers, and in some images the center of the flowers is invisible at all, this affected the result of this flower. Finally, the None/Different class is the lowest accuracy in the test, the result is explained due to the diversity of flowers used for the test, some have the color components in their petals that are close to the colors of the flowers classified.

Table 4 – Classification accuracy results for each flower.

Flower Name	Accuracy
Gerbera Marinilla	99%
Gerbera Renato	95%
Gerbera Chiper	85%
Gerbera Rio Negro	78%
None/different	74%

V. CONCLUSIONS

In this work we have presented an application of image classification process in the flowers using decision tree rules and Hue histogram values, the results demonstrate that the accuracy is higher for some flowers with special values, and that the other categories have values of accuracy significantly high. We conclude that for a few categories the decision tree rule is useful, the results obtained are satisfactory for the project developed. The level of precision achieved in this work is similar to those reported in [2-6]. Finally, we conclude that the use of machine learning in agricultural processes is a relevant topic to increase the productivity and the quality of the process.

As further work we propose to increase the number of categories, also use a different classification algorithm as NN or SVM. We consider that the texture in the center of the flowers is a topic that deserves to be analyzed for classification purposes.

REFERENCES

[1] D. Dey, L. Mummert and R. Sukthankar, "Classification of plant structures from uncalibrated image sequences," Applications of Computer Vision (WACV), 2012 IEEE Workshop on, pp. 329-336, 2012.

[2] A. B. Mabrouk, A. Najjar and E. Zagrouba, "Image flower recognition based on a new method for color feature extraction," Computer Vision Theory and Applications (VISAPP), 2014 International Conference on, vol. 2, pp. 201-206, 2014.

[3] H. Zawbaa, M. Abbass, S. Basha and M. Hazman , "An automatic flower classification approach using machine learning algorithms," Advances in Computing, Communications and Informatics (ICACCI, 2014 International Conference on, pp. 895-901, 2014.

[4] S. Puttemans and T. Goedemé, "Visual detection and species classification of orchid flowers," Machine Vision Applications (MVA), 2015 14th IAPR International Conference on, pp. 505-509, 2015.

[5] F. Siraj , H. M. Ekhsan and A. N. Zulkifli, "Flower image classification modeling using neural network," Computer, Control, Informatics and Its Applications (IC3INA), 2014 International Conference on, pp. 81-86, 2014.

[6] T. Tiay, P. Benyaphaichit and P. Riyamongkol, "Flower recognition system based on image processing," Student Project Conference (ICT-ISPC), 2014 Third ICT International, pp. 99-102, 2014.

[7] S. Aich and C.-W. Lee, "Flower classification: Training augmentation using manifold images," Ubiquitous Robots and Ambient Intelligence (URAI), 2015 12th International Conference on, pp. 204-205, 2015.

[8] J. E. Vera, S. F. Mora, J. A. Torres and J. Avendano, "Analysis of images SAR to flood prevention implementing fusion methods," 2016 XXI Symposium on Signal Processing, Images and Artificial Vision (STSIVA), Bucaramanga, 2016, pp. 1-5. doi: 10.1109/STSIVA.2016.7743334

[9] C. Gamit, P. Swadas and N. Prajapati, "Literature Review on Flower Classification," International Journal of Engineering Research & Technology, vol. 4, no. 2, pp. 510-512, 2015.

[10] S. Suzuki and K. be, "Topological structural analysis of digitized binary images by border following," Computer Vision, Graphics, and Image Processing, vol. 30, no. 1, pp. 32-46, 1985.