




# Classification of *Cattleya Trianae* and Its Varieties by Using Colorimetry

Manuel G. Forero<sup>1</sup> (✉) , Carlos E. Beltrán<sup>2</sup>, Armando Troncoso<sup>1</sup>,  
and Christian González-Santos<sup>1</sup>

<sup>1</sup> Semillero Lún, Facultad de Ingeniería, Universidad de Ibagué, Ibagué, Colombia  
{manuel.forero, christian.gonzalez}@unibague.edu.co,  
armandotroncosolugo@gmail.com

<sup>2</sup> Semillero Lún, Facultad de Ciencias Naturales y Matemáticas, Universidad de Ibagué, Ibagué,  
Colombia  
carlos.beltran@unibague.edu.co

**Abstract.** Orchids in general, like *Cattleya trianae*, have been characterized mostly by taxonomic and visual studies. However, colour is not used for classification. Here, a new method for identifying and classifying orchids of different varieties of *Cattleya trianae* is introduced. This method is not subjective and uses the colour information obtained from the central axis of the *Cattleya trianae* lip. To this end, a new acquisition protocol was established, which uses a new device for image acquisition of the labellum's central axis from the hippocampus to the epicentre. The colour patterns found between samples of the same variety were adjusted and it was verified by using correlation they can be employed to identify each variety. Finally, a support vector machine was used to classify and identify four *Cattleya trianae* varieties, finding that a linear kernel was enough to classify them with an accuracy of 100% .

**Keywords:** Orchids · *Cattleya trianae* · Plant classification · Orchid varieties

## 1 Introduction

The family of orchids, to which the *Cattleya trianae* belongs, shown in Fig. 1 (a), is the most abundant in the world. The number of species is close to 30,000 [1, 2]. In fact, it is the family of plants that exhibits the most advanced characteristics from an evolutionary point of view, which is why it is in the process of diversification, a circumstance that is reflected in the abundance and diversity of species. Orchids are more numerous in the tropics than in other regions of the world [3]. In general terms, the orchid flora of tropical America outnumbers species in other continents [4]. Northwest of South America has the highest number of orchid species. In Colombia, orchids are found in practically all natural ecosystems and occupy the first place as the family of plants with the highest number of species threatened with extinction [5]. There is a great risk to orchids, due to the high levels of endemism in Colombia and the high conversion rates of their ecosystems, to uses such as agriculture and livestock [6].



**Fig. 1.** (a) *Cattleya trianae* orchid. (b) Parts: sepal (S), petal (P) and labellum (L).

*Cattleya trianae* is endemic of Colombia and its distribution is restricted to the upper basin of the Magdalena River, especially in the departments of Tolima, Huila, part of Caqueta and to a lesser extent in Cundinamarca, at altitudes between 600 and 1940 m. In the visual characterisation, *Cattleya trianae* is usually described by its flower, composed of lavender colour petals, twice as large as the sepals as shown in Fig. 1 (b). There are several colour variations among the species. Varieties include the standard or type, semi-alba, alba, concolor variety 1, concolor variety 2, amesiana, coerulea, bull blood or rubra and splash; being different between them by the colour distribution and intensity on petals and, in particular, on the lip.

In the last years, many species of orchids have entered into critical levels due to the difficulty in designing conservation strategies [4]. This is because of their complex life cycles and that they cannot take place without appropriate population studies [5, 8]. Most orchids studies have been limited to population analysis or to visual characterisation of terrestrial species or epiphytic habit and litofic. These studies are difficult and complex, on one hand, due to the variability of the different properties of the flower within the same variety, like shape, texture, pattern and colour and, on the other hand, flowers of different species and varieties are very similar. The characterisation and taxonomic classification of orchids is often done visually by their pattern and morphology, but also based on other types of characteristics such as colour. Therefore, identification requires good experience and it is done mainly by employing subjective methods based on visual and taxonomic properties [3, 7]. Accordingly, it is necessary to develop an objective and reproducible method, i.e., one that will produce the same result independent of the user.

The first classifications of the orchid date back to the 18th century when the Swedish botanist Carlo Linnaeus wrote about it in his compendium “Species plantarum” [7]. The popularity of the flower continued to grow, being even the object of studies of specialists like Charles Darwin, who was interested by its process of reproduction. From the nineteenth century, the work of the experts focused on finding and cataloging orchid species, emphasizing their characteristics and differences. At the beginning, these studies were visually done and nowadays, genetic analysis is also employed for the differentiation of genotypes.

The automatic classification of plants based on the shape, texture and colour of leaves and flowers is an active field in computer vision. These methods usually compare a new specimen against a catalog to identify the plant. Several methods have been developed to identify plants of different species based on flower images. Nilsback and Zisserman

worked in the HSV colour space and employed gradient orientation histograms and the SIFT shape descriptors to get colour, texture and shape characteristics of flowers. A multikernel framework with a SVM was employed for classification [9]. Guru employed cooccurrence matrices and Gabor filters to characterise texture, and classification was done using a k-nearest neighbours method [10]. Mabrook et al. employed SURF and Lab to get shape and colour descriptors to identify flowers of different species and also a SVM for classification [11]. Kumar et al. used RGB colour indexes, as well as leaf shape characteristics to identify plants [12]. Qi et al. employed colour SIFT descriptors and linear SVM for classification [13]. Zawbaa et al. used SIFT and SFTA to describe the characteristics of the flower and SVM and random forest to classify different kinds of flowers [14]. Khan et al. developed new colour descriptors and used them to classify flowers of the Flower-102 base, employing a non-linear SVM [15]. Guo et al. developed a deep sparse coding framework for the visual characterization of flowers of different species in the Oxford-102 flowers database [16]. They employ local contracted sparse coding to extract intermediate local features and local orientation histograms. Hong et al. classifies flowers based on colour and shape descriptors, and employs k-means and history matching to classify [17]. Among the works of identification in orchids, we can mention the works of Arwatchananukul et al. in *paphiopedillum*, who use texture and colour [18]. Sani et al. compared different classification methods Naive Bayes, k-nearest binary tree and sequential minimal optimisation employing colour and texture characteristics through the use of HSV moments, HSV histograms and segmentation based fractal textures SFTA features. They classified two species of *dendrobium* flowers using directly the images of orchids through neural networks [19]. Puttemans et al. classified *phalaenopsis* transforming the images into space Lab and using linear SVM [20].

The above methods do not take into account the colour distribution of the flower, which can be used to recognise *Cattleya trianae* varieties and do not enable the colour of the labellum of one variety to be related to the colour of another. Until now the studies of characterisation of the varieties of *Cattleya trianae* are made only of visual form. In the literature review, no reference was found where these techniques were used for *Cattleya trianae* characterisation, i.e., a study that uses objective techniques, independent of the user and therefore reproducible, since the previous studies depend on the visual concept of each specialist. Therefore, the proposed study, specifically developed for biologists, based on colorimetric analysis from photographs of the lip under controlled lighting conditions is a necessary and novel step for the characterization of the *Cattleya trianae* lip that can be extended in the future to the study of other varieties of flowers. We proposed a new method to study the flowering of *Cattleya trianae* s.p and its varieties, using techniques of image processing and pattern recognition for the characterization of their labellum, from images acquired with a digital camera. The middle axis of each lip is proposed as a kind of fingerprint to identify the flowers, see Fig. 2.



**Fig. 2.** Orchid's labellum or lip.

## 2 Materials

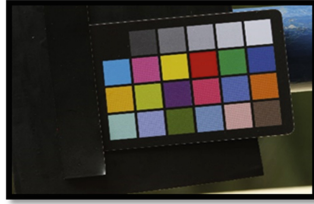
To develop the method, pictures from 99 flowers of four *Cattleya trianae* varieties were taken in situ. Flowers were dissected without damaging their original colour, following the method of classical taxonomic and phylogenetic identification based on visual traits. Images were acquired in RAW format, since it does not introduce losses by compression and stores all the information collected by the camera sensor. Images were taken with an 18 megapixel Canon 7D camera and employing a Canon EF 100 mm f/2.8 USM macro lens, used in biological research to document studies in plant species allowing to appreciate details that often go unnoticed and allowing images to be acquired at close range with low colour distortion and focus on the entire image Fig. 3. To establish an acquisition protocol, 99 flowers of orchids of the *Cattleya trianae* type, concolor var. 1, concolor var. 2 and alba varieties were photographed. Photographs were acquired using the minimum available sensitivity (ISO 100) to minimize noise amplification. The aperture was adjusted so that the entire region of interest was focused, being generally f/11 or less and the shutter speed was set so that the orchid appeared well lit avoiding saturation of the histogram and using the maximum dynamic range available. The speed was generally set to 1/8 s, whereby a trigger cable was used to avoid movement of the camera during acquisition. Various aperture and velocity values were tested, in addition to controlled ambient lighting techniques, until good quality images were obtained. It was found that the best lighting was given with natural light, without direct exposure of the sun, using a parasol. 58 flowers were necessary to adjust the protocol and ensure that in all cases they acquire good quality images that could be used for the study.



**Fig. 3.** Image acquisition method.

### 3 Methods

To assure that colours in the different images were the same and coincided with the real colour, they were calibrated by using the colour chart presented in Fig. 4 and following the protocol established by the chart manufacturer [7].



**Fig. 4.** Chart used to adjust the colour.

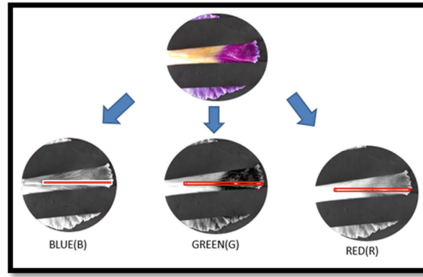
Images were obtained during bloom and most of them served to develop the acquisition protocol. To fix the labellum, the device shown in Fig. 5 was created, consisting of three brass sheets, a square one used as a base where the labellum is placed and two smaller rectangular sheets to fix the labellum. They were fastened to the square base by using two magnets located under the base. The central axis of the labellum, from the epichile to hipochilo, was chosen as fingerprint because this section of the flower contains all the range of colour and change of intensity present in *Cattleya trianae*, given that the labellum is used by the plant to attract pollinators.



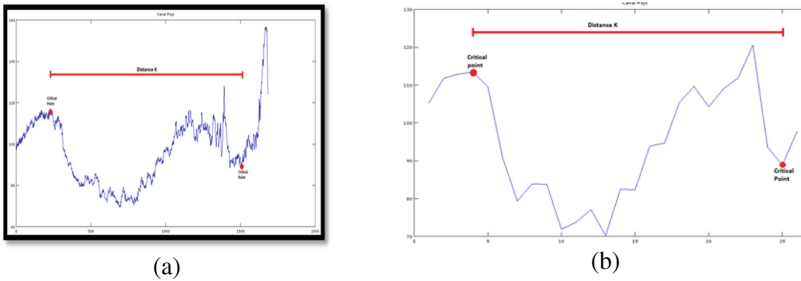
**Fig. 5.** Base developed to fix the labellum. (a) Top and bottom view. (b) Side view.

The colour profiles of the central area of each labellum were obtained, as shown in Fig. 6. To reduce noise, a fringe of the central axis was used and the average of the colour values in each column of the profile was calculated.

Given that flower size is variable and the distance from the *Cattleya trianae* to the camera also changes from one acquisition to another, profiles were normalised by using a method similar to the one employed in electroencephalography (EEG), used to adjust EEG signals to obtain evoked potential responses, as shown in Fig. 7.a. To this aim, the colour channel profile that presented maximum amplitude variation was chosen. Then, all signals were aligned in such a way that the first and last local minimum or maximum local of the profile coincide. The profiles of the other two channels were also aligned by using the same calibration.



**Fig. 6.** Acquisition of RGB profiles



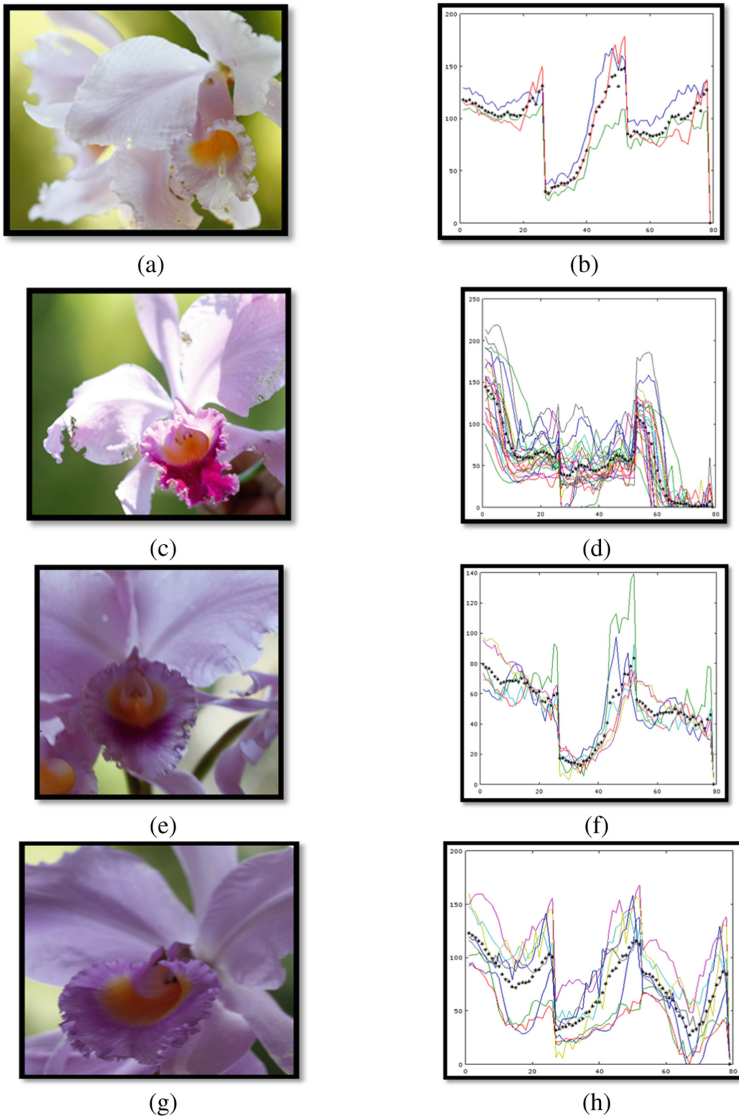
**Fig. 7.** Normalization and sampling of channel profiles. (a) Channel chosen. (b) Channel sampled.

Although the images of the central axis were taken by using the same protocol, some profiles were longer; thereby, to assure that profiles corresponded to the same equivalent part of the central axis, the distance between the first and the minima or maxima, used as reference points, was calculated and 20 equally spaced samples were taken. Three additional samples were taken before the first reference point and three others after the last one, making a total of 26 equally spaced samples per channel profile, and a total of 78 samples used as descriptors for each flower, significantly reducing data processing and noise, as shown in Fig. 7.b.

## 4 Results

From the 99 images acquired, only 41 were usable, given that most were used to establish the acquisition protocol. Figure 8 shows the orchids varieties used in this study. To verify if the profiles could be used as a kind of fingerprint to identify *Cattleya trianae* varieties, the correlation between the samples of each variety were calculated. The correlation between the profiles of the different varieties of *Cattleya trianae* are displayed in Table 1. As it can be seen, a high correlation appears between several profiles from the different varieties and the highest correlation between samples of the same variety.

To classify and identify each variety of *Cattleya trianae*, some popular classification algorithms were studied. Among them, it was decided to use support vector machines (SVM) for this work, which has important advantages over neural networks (ANN), another very popular method. The advantages of SVM are that the solution converges towards a global minimum instead of local minima and models depend on a few parameters, so modelling is easier.



**Fig. 8.** Orchid varieties. (a) Alba. (c) Standard or type. (e) Concolor var. 1. (g) Concolor var. 2. (b), (d), (f) and (h) show the profiles and the average of the three channels altogether, one behind the other for the four varieties. It can be observed that there is a characteristic pattern for each variety, where the pattern dotted with black asterisks represents the average value for that variety. This average pattern was employed to find the correlations.

To train the SVM, the 26 samples obtained by each channel were used as characteristic descriptors. In this way, a descriptive vector of 78 characteristics was formed, to which a label was added according to the variety to which the sample belongs (groups 1, 2, 3 and 4). The SVM was implemented using a linear kernel, given the low number

**Table 1.** Average correlation of the three colour profiles, between samples of the same variety and against the other varieties.

Orchid	Type	Alba	Concolor var. 1	Concolor var. 2
Type	0.901	0.1224	0.606	0.5757
Alba	0.1261	0.9515	0.7006	0.6824
Concolor v. 1	0.5414	0.6426	0.8342	0.7643
Concolor v. 2	0.586	0.602	0.82	0.871

**Table 2.** Vector classification results of the vector machine.

Orchid	Number of samples	True positives	True negatives	False positives	False negatives
Type	24	100%	0%	0%	0%
Alba	3	100%	0%	0%	0%
Concolor 1	8	100%	0%	0%	0%
Concolor 2	6	100%	0%	0%	0%

of samples and the data set was subdivided into three randomly distributed groups, 60% for training, another 20% for model fit and the remaining 20% for validation, using cross-correlation with which a 100% successful classification result was obtained.

## 5 Discussion

Table 1 presents the averaged values of the correlations between varieties of *Cattleya trianae* and the average profile of each of them. It is observed that the highest correlation is always between the average profile and the profiles of their respective varieties. Therefore, it is possible to use the profile of the lip central axis to perform the *Cattleya trianae* classification and identification. As it can be seen in Table 2, there is a very high correlation between two similar varieties (concolor var. 1 and concolor var. 2), being coincident with the colour profile of both flowers. It also can be observed in Fig. 8 (b), (d), (f) and (h), the profiles of the green channels of the alba, concolor var. 1 and concolor var. 2 are very similar. Therefore, by analysing the correlation between channels can provide information about the similarity between colour components of the varieties, making it also possible to be used to get information about the common origin of different varieties of *Cattleya trianae*.



## 6 Conclusions

A new method to identify and classify orchids of the different varieties of the *Cattleya trianae* was presented. This method is objective and reproducible, allowing to automate this procedure. It was shown the central axis of the label (from hipochilo to epichilo), can be employed as a fingerprint to identify varieties of *Cattleya trianae*. A protocol was set up to images of the labellum central axis of the *Cattleya trianae* (from the epichile to hipochilo) and its varieties. The results show the profile of this zone can be used as a fingerprint to classify and identify its varieties. The highest correlation was between similar varieties (concolor var. 1 and concolor var. 2), being coincident with the colour profile of the flowers, suggesting this new tool could be employed to study the common origin of different varieties.

**Acknowledgements.** This work was supported by project #14-314-INT Universidad de Ibagué.

## References

1. Dressler, R.: How many orchid species? In: Proceedings Second International Orchid Conservation Congress, pp. 155–158 (2005)
2. van Raveschot, V.G.: Aclimatación de plántulas de *Chloraea virescens* (Willd.) Lindl cultivadas in vitro. Universidad Austral de Chile (2009)
3. Dressler, R.: Phylogeny and Classification of Orchid Family. Dioscorides Press, Portland (1993)
4. Gentry, A.H., Dodson, C.H.: Diversity and biogeography of neotropical vascular epiphytes. Ann. Missouri Bot. Gard. **74**(2), 205–233 (1987)
5. Calderón Sáenz, E.: Libro rojo de plantas de Colombia. Orquídeas, primera parte. Instituto Alexander von Humboldt, vol. 6 (2007)
6. Dixon, K., Phillips, R.D.: The orchid conservation challenge. Lankesteriana Int. J. Orchid. **7**(1–2), 11–12 (2007)
7. Linnaeus, C.: Species Plantarum, 1753rd edn. Ray Society, London (1957)
8. Meisel, J.E., Kaufmann, R.S., Pupulin, F.: Orchids of Tropical America. Cornell University Press, Ithaca (2014)
9. Philbin, J., et al.: Lost in quantization: improving particular object retrieval in large scale image databases. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 1–8 (2008)
10. Guru, D., et al.: Texture features and KNN in classification of flower images. In: Fifth Indian International Conference on Artificial Intelligence, vol. 1 (2010)
11. Mabrouk, A.B., et al.: Image flower recognition based on a new method for color feature extraction. In: 2014 International Conference on Computer Vision Theory and Applications, vol. 2, pp. 201–206 (2014)
12. Mishra, P.K., et al.: A semi automatic plant identification based on digital leaf and flower images. In: IEEE International Conference on Advances in Engineering, Science and Management (ICAESM-2012), pp. 68–73 (2012)
13. Qi, L.S., et al.: Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. Cell **152**(5), 1173–1183 (2013)
14. Zawbaa, H., et al.: An automatic flower classification approach using machine learning algorithms (2014)

15. Khan, N., et al.: Discriminative color descriptors. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 2866–2873 (2013)
16. Guo, L., Guo, C.: A deep sparse coding method for fine-grained visual categorization. In: International Joint Conference on Neural Networks, pp. 632–639 (2016)
17. Hong, S.W., Cho, L.: Automatic recognition of flowers through color and edge based contour detection. In: International Conference on Image Processing Theory, Tools and Applications (2012)
18. Arwatchananukul, S., et al.: POC: *paphiopedilum* orchid classifier. In: IEEE 14th International Conference on Cognitive Informatics & Cognitive Computing, pp. 206–212 (2015)
19. Sani, M.M., et al.: Classification of orchid species using neural network. In: IEEE International Conference on Control System, Computing and Engineering, pp. 586–589 (2013)
20. Puttemans, S., Goedemé, T.: Visual detection and species classification of orchid flowers. In: International Conference on Machine Vision Applications, pp. 505–509 (2015)