Mean and Range Color Features Based Identification of Common Indian Leafy Vegetables

Ajit Danti[1], Manohar Madgi[2] and Basavaraj S. Anami[3]

[1]*J. N. N. College of Engineering, Shimoga – 577 204, Karnataka, India*
[2,3]*K. L. E. Institute of Technology, Hubli – 580 030, Karnataka, India*
ajitdanti@yahoo.com, manohar.madgi@gmail.com, anami_basu@hotmail.com

Abstract

The computer vision techniques are required for the identification of leafy vegetables for the development of veggie vision applications. This paper presents the mean and range color features based identification of leafy vegetables. Initially, a total of 18 RGB and HSI color features are chosen. A reverse engineering process is adopted for reduction of features. Finally 12 mean and range features of RGB and HSI color features are selected based on the performance. A BPNN based classifier is used for identification of vegetables. The identification rate is in the range 92-100% for ten types of vegetables. The work finds applications in automatic vending, packing and grading of vegetables, food preparation and the like.

Keywords: *Vegetable Identification, Color features, Neural Networks, Agricultural/horticultural produce*

1. Introduction

In the recent years, a considerable amount of developments have been taken place in the area of robotics. Robots have found a widespread applications in the aerospace, automotive, consumer goods, food, metal fabrication, medical, pharmaceutical, solar panel, agricultural and many other industries. They carry pre-programmed tasks very efficiently with an incorporated computer vision systems. Computer vision is concerned with building artificial systems that obtain information from images. Computer vision systems have many applications such as detecting events, performing agricultural tasks, organizing information, modeling objects or environments interacting with human, controlling of processes and the like. Computer vision is considered a complement to the biological vision. There are certain computer vision systems developed for agricultural/horticulture applications, namely, weed identification, leaf disease identification, classification of grains and recognition of food in food industry, medical plant recognition, classification of crops and weeds etc. In all these systems, the images are subjected to image processing techniques to extract useful features that are necessary for further recognition and classification.

It is evident that majority of activities in the real world with the development of robotics will be carried out automatically. One such application is envisaged, wherein the automatic identification of images of leafy vegetables. Many types of leafy vegetables are found in supermarket that are grown in India are shown in Figure 1. The details in terms of English name, botanical name and Indian common names are listed in Table 1. To know the state-of-the-art applications in the area of agriculture in general and vegetable identification in particular, a literature survey is carried out. Following is the gist of the survey.

Figure 1. Samples of Common Indian Leafy Vegetables

Table 1. Details About the Common Indian Leafy Vegetables

English Name	Botanical Name	Common name in India		
		Hindi	Kannada	Telugu
Green Amaranth	Amaranthus viridis	Chaurai	Dhantina Soppu / Rajgiri	Thota aaku
Coriander leaves	Coriandrum sativum	Dhania patta	Kotambari soppu	Kothimeera aaku
Curry leaves	Murraya koenigii	Karipatta	Karibevu soppu	Karivepaku
Dill	Anethum graveolens L.	Suwa	Sabsige soppu	Soya aaku
Fenugreek	Trigonella foenumgraecum	Methi	Menthya Soppu	Menthi aaku
Sorrel leaves	Rumex acetosa	Meshta / Pitwaa	Pundi soppu	Gongura
Mint leaves	Mentha spicacata	Pudina patta	Pudina	Pudina aaku
Onion leaves	Allium cepa	Pyaaz	Eerulli soppu	Ullipaaya aaku
Red Amaranth	Amaranthus cruentus	Lal Chaurai	Kempu Rajgiri	Koyya Thota aaku
Spinach	Spinacea oleracea	Paalak	Palak Soppu	Palak aaku

[1] have developed an automatic fruits and vegetables classification methodology from images. A unified approach that combines many features and classifiers which requires less training, and is more adequate to establish image categories automatically using histograms, colors and shape descriptors with an unsupervised learning method. The results show that, the solution is able to reduce the classification, error is up to 15% points with respect to the baseline.

[11] have developed an approach to classify potato chips using pattern recognition from color images. Ten chips were prepared for each of the following 6 conditions: two pretreatments (balanced and unbalanced) at three temperatures (120 ℃, 150 ℃, and 180 ℃).

More than 1500 features are extracted from each of the 60 images. Finally, 11 features are selected according to their classification attributes. Seven different classification cases are analyzed using the selected features. Although samples are highly heterogeneous, using a simple classifier and a small number of features, it is possible to obtain a good performance value in all cases: classification of the six classes is in the confidence interval between 78% and 89% with a probability of 0.95.

[3, 5] have proposed a methodology for identification and classification of boiled food grains based on the level of boiling using two color models HSV and L*a*b*. These color models provide a good texture definition of an image. The classification is performed at two levels; Level-1 determines the type of grain image and Level-2 estimates the amount of boiling. They have reported average accuracies of 80% and 96% for Level-1 and 70% and 90% for Level-2 classifications for HSV and L*a*b* respectively.

[10] have presented a machine vision algorithm to characterize target features in color space from a set of training data so that the color classification can be performed accurately and efficiently. They have evaluated the algorithm in the context of food processing applications and examine the effects of the color characterization on computational efficiency.

[2] have designed an algorithm for identification and classification of medicinal plants based on edge and color descriptors. They have considered the edge and color descriptors that have two dimensions. The classification is done using SVM based classifier. The classification accuracies for color and edge features are 74% and 80% respectively. [5] color features to classify the grapes.

[9] have presented fruit detection using improved multiple features based algorithm. It helps to locate the fruit on the tree. The algorithm is designed for calculating different weights for features like intensity, color, orientation and edge of the input test image. The weights of different features represent the approximate locations of the fruit within an image. The detection efficiency is achieved up to 90% for different fruit image on tree and captured at different positions.

[12] have proposed a method for fruit recognition using color and texture features. An efficient fusion of color and texture features for fruit recognition is proposed. The recognition is done by the minimum distance classifier. They have reported a recognition rate of 45% for color features, 70% for texture features and 86% for combined features.

[4] have presented a method for automatic produce ID system, intended to ease the vegetable checkout process. A variety of features namely color, texture and shape are used. Each texture measure is computed using the segmented image from the green channel output only. Two texture measures are developed. The first measure convolves the image with two crossed bar masks, with the bars parallel to the image x and y directions, second measure is a “centre-surround" operator, a kind of first-order statistic. It consists of computing and histogramming the deviation of the image intensity of a pixel from the average of its neighbours in a moderate-sized block centered on that pixel. Depending on the certainty of the classification, the final decision is made either by the system or by a human from a number of choices selected by the system. Over 95% of the time, the correct produce classification is achieved.

[6] have developed an automated classification system of pizza sauce spread using color vision and support vector machines (SVM). To characterize pizza sauce spread with low dimensional color features, a sequence of image processing algorithms are developed. The image is transformed from red, green, and blue (RGB) color space to hue, saturation, and value (HSV) color space. A vector quantifier is designed to quantify the HS (hue and saturation) space to 256-dimension, and the quantified color features of pizza sauce spread are represented by color histogram. Finally, principal component analysis (PCA) is applied to

reduce the 256-dimensional vectors to 30-dimensional vectors. With the 30-dimensional vectors as the input, SVM classifiers are used for classification of pizza sauce spread.

[14] have described image analysis algorithm based on color and morphological features for identifying different varieties of rice seeds. Seven color features and fourteen morphological features are used for the discrimination of analysis. A two-layer neural network model is developed for classification. [7] have developed an Automated Visual Inspection System (AVIS) for quality control of preserved Orange segments .The method has adopted a three layer neural network perception topology.

[13] have developed an image-based food classification system to grade six types of fresh produce using average CIE L*a*b* and Bayesian interface technique to classify food ripeness.

From the literature survey, it is observed that not much work on identification of leafy vegetables is reported and hence it is the motivation for the present work. In this work, we have considered leafy vegetables which are commonly grown and used in India. They are classified based on mean and range RGB and HSI color features using ANN classifier.

The paper is organized into five sections. Section 2 contains the proposed methodology and image acquisition. Section 3 describes feature extraction and reduction. Results and discussions are presented in Section 4. Section 5 concludes the work.

2. Proposed Methodology

The proposed methodology comprises of four stages, namely image acquisition, feature extraction, feature reduction and classification. The block diagram depicting these stages is given in Figure 2.

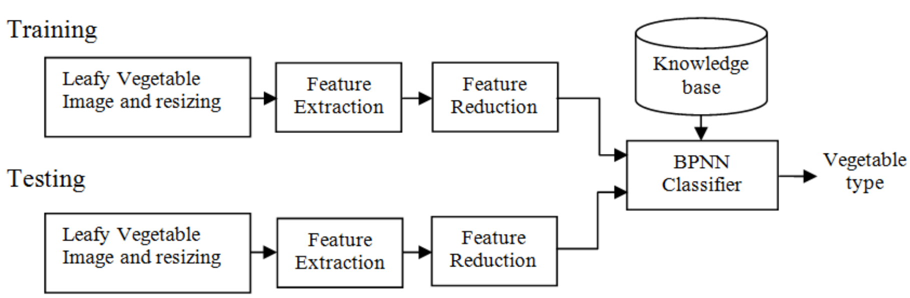

Figure 2. Block Diagram of Proposed Methodology

2.1 Image Acquisition and Resizing

The ten different types of fresh leafy vegetable bunches are selected from the market. These vegetables are placed against a plain white background vegetable type at a time and images are taken. The images are acquired in a day light using a digital camera placed at a distance of 40 centimeters. A Sony digital camera with an image resolution of 10 mega pixels is used to capture the images. The images obtained are of size 3264 x 2448 pixels. The images are cropped and resized to 150 x 150 pixels for computational reasons.

3. Feature Extraction and Reduction

The leafy vegetables are identified by the human being mainly by color, texture and shape of the leaves. Color is the most important feature for identification of vegetables. However, color is more or less green across most of the vegetables, but the shades of green color vary from vegetables to vegetables. The vegetables are also identified by the shape of the leaves, when a single leaf is taken. The variation in leaves sizes give rise to variation in texture when leafy vegetable bunch is used. The texture is another feature one can use to identify vegetables. In this work, we have considered mean and range color features for identification of vegetables.

3.1 Color Feature Extraction

The values of R, G and B color components from vegetable images are normalized so as to have the values in the range [0, 1]. The HSI color features are obtained from the RGB components.

The leafy vegetables images have distribution of color and change in color over the image. These discriminating features are represented by mean, variance and range values for the given image and are listed in Table 2. These 18 color features are used to characterize an image. All these 18 features obtained by the equations (Gonzalez, 2008) for leafy vegetables are shown in Figure 3, 4, 5 and 6.

3.2 Feature Reduction

From the graph it is observed that, for all types of vegetables the variance feature values are very low (< 0.1). Such features do not contribute to the identification process. Hence, the variance feature is ignored. The 18 features are reduced to 12 features and termed as reduced color features set which are listed in Table 3. For the purpose of feature reduction, a threshold based approach is used. A threshold of 0.5 for the first phase and 0.1 for the second phase are used. The threshold values are obtained by the experimentation. Each feature is compared with first threshold value 0.5. If the value is larger than the threshold then it is considered as a good feature. The feature that fails the first threshold value, it is compared with the second threshold as not to miss selection. If the feature is greater than the threshold, then it is labeled as a good feature.

Table 2. 18-Color Features

Sl. No.	Feature	Sl. No.	Feature
1	Red mean	10	Hue mean
2	Red variance	11	Hue variance
3	Red range	12	Hue range
4	Green mean	13	Saturation mean
5	Green variance	14	Saturation variance
6	Green range	15	Saturation range
7	Blue mean	16	Intensity mean
8	Blue variance	17	Intensity variance
9	Blue range	18	Intensity range

Table 3. Reduced Color Features Set

Sl. No.	Feature	Sl. No.	Feature
1	Red mean	7	Hue mean
2	Red range	8	Hue range
3	Green mean	9	Saturation mean
4	Green range	10	Saturation range
5	Blue mean	11	Intensity mean
6	Blue range	12	Intensity range

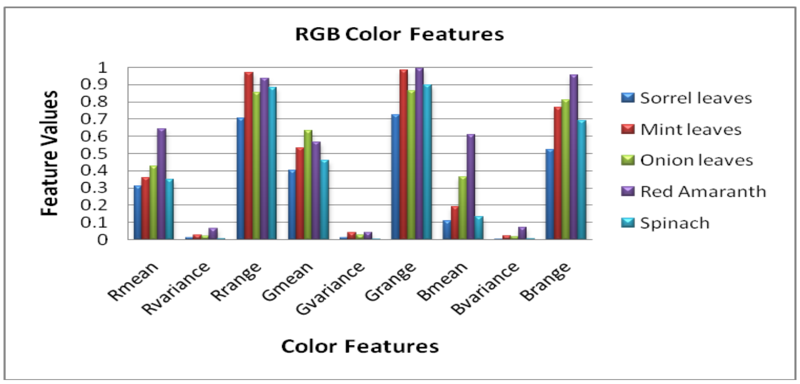

Figure 3. RGB Color Features Histogram of Green Amaranth, Coriander Leaves, Curry Leaves, dill and Fenugreek

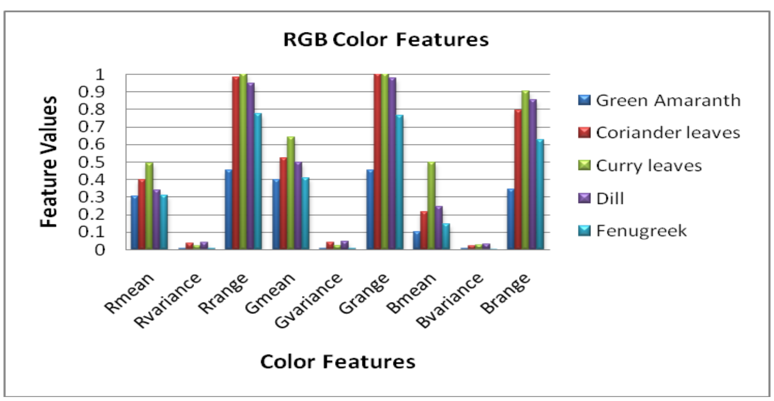

Figure 4. RGB Color Features Histogram of Sorrel leaves, Mint leaves, Onion leaves, Red Amaranth and Spinach

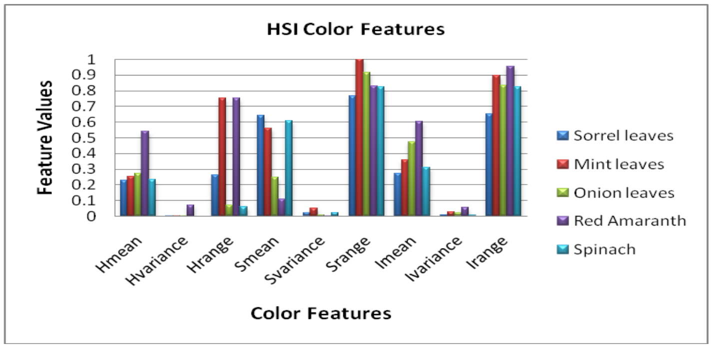

Figure 5. HSI Color Features Histogram of Green Amaranth, Coriander leaves, Curry leaves, dill and Fenugreek

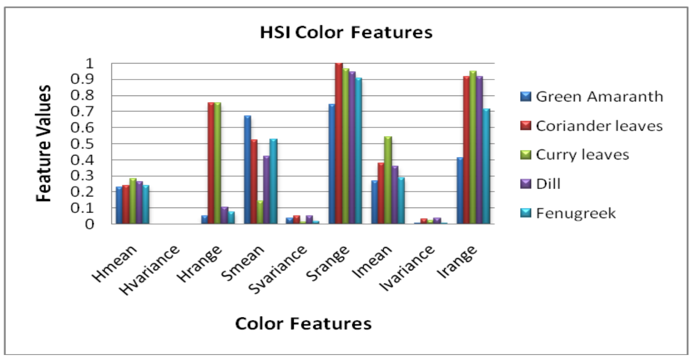

Figure 6. HSI Color Features Histogram of Sorrel leaves, Mint leaves, Onion leaves, Red Amaranth and Spinach

In this work, the mean and range features are adopted, since these satisfied the feature reduction strategy. The procedure to obtain these features is given in Algorithm 1.

Algorithm 1: Feature Extraction and Reduction

Input: Original 24-bit RGB color images of different vegetable samples.
Output: Reduced features set of 12.
Start
Step 1: Accept the leafy vegetable image.
Step 2: Separate R, G and B components from the original 24-bit input color image.
Step 3: Obtain the color features Hue (H), Saturation (S) and Intensity (I) components from RGB components.
Step 4: Extract all the 18 color features. (Gonzalez, 2008)
Step 5: For feature $\leftarrow$ 1 to 18 do
if (feature value > 0.1) then It is a good feature
Add the feature to Reduced Feature set.
Stop

3.3 Classifier Parameters

A three-layer feedforward backpropagation Artificial Neural Network (BPNN) is implemented using neural network toolbox of MATLAB version R2009a. The mean square error is set to 0.001. There are 12 nodes in the input layer. The hidden layer has thirty-three nodes. The output layer has 10 nodes, corresponding to the ten different vegetable classes. The learning rate is set to 0.01.

4. Results and Discussions

The developed leafy vegetable identification method is tested on 1000 images of leafy vegetables bunches of 10 different types. 500 images with 50 images from each type are used to train the neural network classifier; the remaining 500 images are used to test the accuracy. Confusion matrix is shown in Table 4, to indicate the correct and incorrect classification of vegetables.

We have trained and tested the neural network considering all 18 color features, namely mean, variance and range of RGB and HSI color models. During experimentation, it is observed that some of the color features have very low values which does not contribute to the accuracy of the classifier. Hence, we have reduced the feature set to twelve. The identification accuracy of different types of leafy vegetables is shown in Figure 7. It is observed from the graph that, the highest identification rate observed is 100% and the lowest observed is 92%. The average identification rate is 96.4%. From the confusion matrix given in Table 4, it is observed that two samples of Green Amaranth are misclassified as Curry leaves and Dill. Two samples of the vegetable Dill are misclassified as Fenugreek and one sample as Spinach. The three samples of Fenugreek are classified as Dill and one sample as Green Amaranth. Three samples of Sorrel leaves are misclassified as Green Amaranth, Curry and Fenugreek one each. One sample of Mint leaves is misclassified as Dill. The two samples of Onion are misclassified as Curry leaves and Dill. One sample of Red Amaranth is misclassified as Green Amaranth. The two samples of Spinach are misclassified as Green Amaranth and Dill. The overall performance of the developed method is good.

Table 4. Confusion Matrix of Leafy Vegetables

I \ O	GA	Co	Cu	Di	Fe	So	Mi	On	RA	Sp
GA	48	0	1	1	0	0	0	0	0	0
Co	0	50	0	0	0	0	0	0	0	0
Cu	0	0	50	0	0	0	0	0	0	0
Di	0	0	0	47	2	0	0	0	0	1
Fe	1	0	0	3	46	0	0	0	0	0
So	1	0	1	0	1	47	0	0	0	0
Mi	0	0	0	1	0	0	49	0	0	0
On	0	0	1	1	0	0	0	48	0	0
RA	1	0	0	0	0	0	0	0	49	0
Sp	1	0	0	1	0	0	0	0	0	48

Legend: GA-Green Amaranth, Co-Coriander leaves, Cu-Curry leaves,
Di-Dill, Fe-Fenugreek, So-Sorrel leaves, Mi-Mint leaves,
On-Onion leaves, RA-Red Amaranth, Sp-Spinach

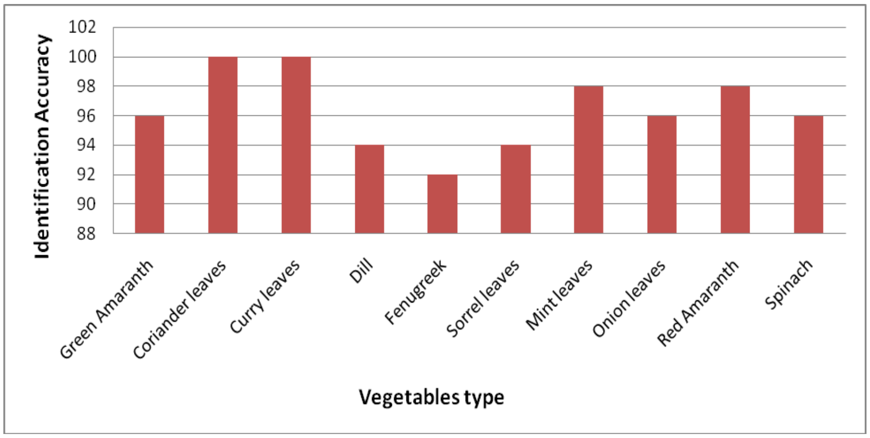

Figure 7. Identification Accuracy for Different Vegetables based on Color Features

5. Conclusion

In this work, we have classified the leafy vegetables into few types. The HSI color model is considered in the method. The feature set is subjected to reduction based on the thresholding values. A BPNN classifier is used for the identification and classification. The maximum identification rate of 100% is observed with Coriander and Curry leaves and a minimum of 92% is observed with Fenugreek leaves. The work carried out has relevance to real world applications in automatic identification of the different vegetable types, their grading, valuation, packing, sales in the supermarket, food preparation and the like.

References

[1] A. Rocha, D. C. Hauagge, J. Wainer and S. Goldenstein, "Automatic fruit and vegetable classification from images", In: Computers and Electronics in Agriculture, vol. 70, **(2010)**, pp. 96-104.

[2] B. S. Anami, S. S. Nandyal and A. Govardhan, "A Combined Color, Texture and Edge Features Based Approach for Identification and Classification of Indian Medicinal Plants", In: International Journal of Computer Applications (0975-8887), vol. 6, no. 12, **(2010)**, pp. 45-51.

[3] B. S. Anami and V. C. Burkpalli, "Color Based Identification and Classification of Boiled Food Grain Images", In: International Journal of Food Engineering, vol. 5, Issue 5, Article 7, **(2009)**, pp. 1-19.

[4] R. M. Bolle, J. H. Connell, N. Haas, R. Mahon and G. Taubin, "VeggieVision: A Produce Recognition System", Technical Report forthcoming, IBM **(1996)**.

[5] B. S. Anami, D. G. Savakar, A. K. Kannur and M. V. Karali, "Classification of Grapes Using Artificial Neural Network", Proceedings of the International Conference on Cognition and Recognition, **(2008)**, pp. 436-442.

[6] C. -J. Du, Da-Wen, "Pizza sauce spread classification using color vision and support vector machines", Journal of Food Engineering, vol. 66, Issue 2, **(2005)**, pp. 137-145.

[7] C. Fernandez, J. Suardiaz, C. Jimenez, P. J. Navarao, A. Toledo and A. Iborra, "Automated Visual Inspection System for the Classification of Preserved Vegetables", in Proc. of the IEEE International Symposium, vol. 1, **(2002)**, pp. 265-269.

[8] R. C. Gonzalez and R. E. Woods, "Digital Image Processing", Third Edition, Pearson Education Inc., **(2008)**.

[9] H. N. Patel, R. K. Jain, M. V. Joshi, "Fruit Detection using Improved Multiple Features based Algorithm", In: International Journal of Computer Applications (0975-8887), vol. 13, no. 2, **(2011)**, pp. 1-5.

[10] K.-M. Lee, Q. Li and W. Daley, "Effects of Classification Methods on Color-Based Feature Detection With food Processing Applications", In: IEEE Transactions on Automation Science and Engineering, vol. 4, no. 1, **(2007)**, pp. 40-51.

[11] F. Pedrerschi, D. Mery, F. Mendoza and J. M. Aguilera, "Classification of Potato Chips Using Pattern Recognition", Journal of Food Science, vol. 69, **(2004)**, pp. E264-270.

[12] S. Arivazhagan, R. N. Shebiah, S. S. Nidhyanandhan and L. Ganesan, "Fruit Recognition using Color and Texture Features", In: Journal of Emerging Trends in Computing and Information Sciences", vol. 1, no. 2, **(2010)**, pp. 90-94.
[13] S. Somatilake, A. N. Chalmers, "An Image-Based Food Classification System", In: Proceedings of Image and Vision Computing, **(2007)**, pp. 260-265, New Zealand.
[14] Z. -y. Liu, F. Cheng, Y. -b. Ying and X. -q. Rao, "Identification of rice seed varieties using neural network", Journal of Zhejiang University SCIENCE B, vol. 6, no. 11, **(2005)**, pp. 1095-1100.