Plant Species Identification Using Image Processing

Mr.S.Gnana Saravanan[1], Evanjalin Kirthika. Fazil. A, Hezia Golda. N, Jeevitha. R [2]

[1] Assistant Professor (SG), Dept. of ECE, Sri Shakthi Institute of Engg and Tech, Coimbatore
[2]Final Year ECE Students, Sri Shakthi Institute of Engg and Tech, Coimbatore.
(E-mail: gnanasaravanan@siet.ac.in)

Abstract**—** nowadays image processing plays a major role in identification of all aspects. Here image processing technique is used for identification of plant species. Leaves play a major role for the classification of plants. There are different varieties of tress grown in this world. They are basically identified by the leaves. The samples are taken from various plants, textures, shape. The image is captured and is fed to the computer. Now the captured image of sample is processed by converting from RGB to grey value. By applying advanced techniques of image processing and utilizing the capabilities of the recent image storage facilities and use of computer techniques for analyzing the shape,texture,color,aspect,ratio,vien structure, entropy and so on. Now the processed leaf samples are compared with the reference leaves to identify the plant species. This technique is used to give machine vision to robots in future and identify plant species with medicinal values. The detection is made using different types of process in the MATLAB software.

Keywords**—** *image processing technique ; various plants, textures, shape ; RGB ; shape,texture,color,aspect,ratio,vien structure, entropy.*

I. INTRODUCTION

Plants are important part of our ecosystem. There is huge number of plants existing in nature. Many of them are at the risk of extinction. So it is very necessary to catalogue their identity, features and useful properties. Now a day, whole world is facing various problems like global warming, biodiversity loss, effects of fast urban development, and various environmental damages. Hence there is an need to apply image processing techniques to obtain the botanical knowledge like plant taxonomy, various features of plant and make this information accessible and useful to different kinds of people like researchers, farmers, botanists, and students. Hence plant identification is the first and important task. There are many plant organs like leaves, flowers, fruits, seeds which can be used for plant identification. In this paper leaves are selected to obtain the features of plant. Because leaf can be easily obtained and scanned and also it consists of more excluding information which is useful for plant classification. These leaf images are sent to computer and then by using image processing tools, leaf features are extracted to identify the plant. Several methods have been introduced for plant identification. Here leaf contour i.e Border Tracing algorithm is used. The program interfaced to Arduino through the UART. At last the result of the species name is obtained.

A. Plant Species Identification

Plants play a critical role in preserving the delicate balance of the environment. In this plant species is identified using image processing and certain steps are followed

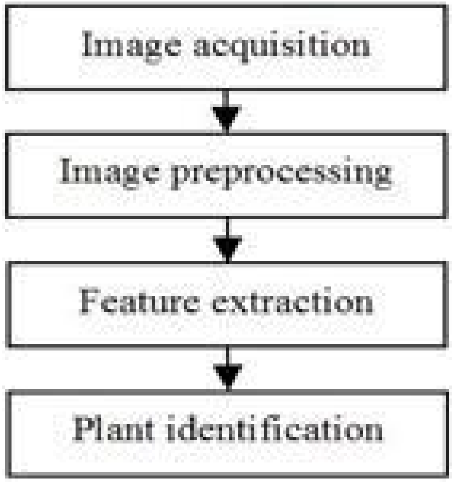


Fig 1; preprocessing stage

B. Leaf Image Acquisition

The leaf images can be acquired using a scanner or digital camera. There is no restriction for image format and resolution is resized. However, a descent sized image with reasonable resolution is enough for the proposed method. The image can be an RGB image and can be converted to gray scale image. However, the image background needs be clean preferably white or any single colored with reasonable contrast with the leaf color and the leafstalk should be removed prior to image acquisition. For our experiment, we collected the dataset of plant using our digital camera, RGB leaf images in JPEG format is collected and converted from 3D to 2D. All images were acquired using digital camera and have white background with no leafstalk. There are total 5 plant species; each plant species has around 10 to 30 sample leaves.

C. Leaf Contour

The contour of the leaf two types of features is extracted. Geometric and moment based features. The geometric features are features like rectangularity. Circularity and area of convex hull the moment features are features like moment in variants. First a digital camera is used to take images of the whole plant species under study. The images are taken carefully so that the majority of the leaves are parallel to the camera lens. Individual leaves are then manually segmented. Experiments are presented to verify the effectiveness of the approach. In this project we present an experiment on different leaf species, the results show that the algorithm is a promising 2D contour identification.

D. Matlab

MATLAB is a fourth-generation programming language and numerical analysis environment. MATLAB, an abbreviation for „matrix laboratory", is a platform for solving mathematical and scientific problems. Cleve Moler, a professor of Computer Science at the University of New Mexico, created MATLAB in the 1970s to help his students. MATLAB's commercial potential was identified by visiting engineer Jack little in 1983.It is a proprietary programming language developed by Math Works, allowing matrix manipulations, functions and data plotting, algorithm implementation, user interface creation and interfacing with programs written in programming languages like C, C++, Java and so on.

In MATLAB, the IPT is a collection of functions that extends the capability of the MATLAB numeric computing environment. It provides a comprehensive set of reference-standard algorithms and workflow applications for image processing, analysis, visualization and algorithm development. It can be used to perform image segmentation, image enhancement, noise reduction, geometric transformations, image registration and 3D image processing operations. Many of the IPT functions support C/C++ code generation for desktop prototyping and embedded vision system deployment.

E. Arduino

The Arduino project started at the Interaction Design Institute Ivrea(IDII) in Ivrea. In 2003 Hernando Barragan created the development platform Wiring as a Master"s thesis project at IDII, under the supervision of Massimo Banzi and Casey Reas, who are known for work on the Processing language. The project goal was to create simple, low cost tools for creating digital projects by non-engineers.

The Wiring platform consisted of a printed circuit board(PCB) with an A Tmegal 168 microcontroller, an IDE based on processing and library functions to easily program the microcontroller. In 2003, Massimo Banzi, with David Mellis, another IDII student, and David Cuartielles, added support for the cheaper ATmega8 microcontroller to Wiring. But instead of continuing the work on Wiring, they forked the project and renamed it Arduino.

The project"s products are distributed as open-source hardware and software, which are licensed under the GNU Lesser General Public License(LGPL) or the GNU General Public License(GPL), permitting the manufacture of Arduino boards and software distribution by anyone. Arduino boards are available commercially in preassembled form, or as do Arduino board designs use a variety of microprocessors and controllers. The boards are equipped with sets of digital and analog input/output(I/O) pins that may interfaced to various expansion boards and other circuits. The boards feature serial communications interfaces, including Universal Serial Bus(USB) on some models, which are also used for loading programs from personal computers. The microcontrollers are typically programmed using a dialect of features from the programming languages C and C++. In addition to using traditional compiler tool chains, the Arduino project provides an integrated development environment (IDE) based on the processing language project. Most Arduino boards consists of an Atmel 8-bit AVR microcontroller(ATmega8, ATmega328, ATmega1280, ATmega2560) with varying amounts of flash memory, pins and features. The boards use single or double-row pins or female headers that facilitate connections for programming and incorporation into other circuits.

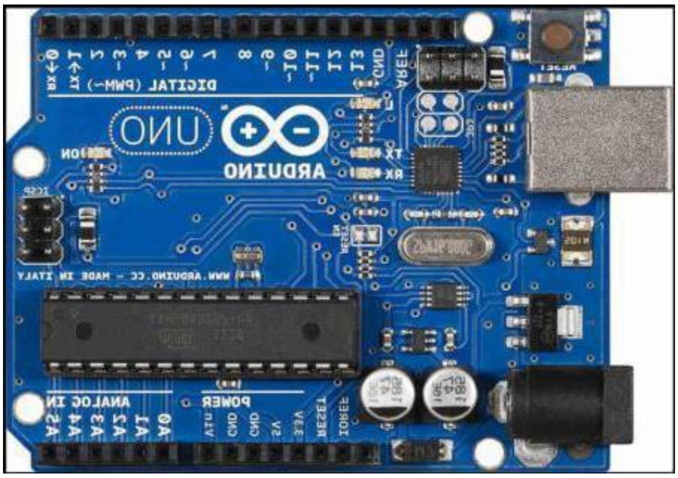


Fig 2; Arduino Board

II. EXISTING SYSTEM

The contour of a leaf first is extracted then processed and a feature 'Vector is formulated. The feature 'Vectors are stored in a database to be used for identifying plants in the worksite. An image of the leaf is first feed to the computer, the feature 'Vector is then extracted from the leaf and finally the species is identified by matching the leaf feature vector with the feature vectors stored in the database. The main aim of this paper is to identify shape descript-ors which are used in a computer system to match or identify similar leaves [1]. Computer-Aided Plant Species Identification acts significantly on plant digital museum system and systematic botany, which is the groundwork for research and development of plant. This paper presents a new method for plant species identification using leaf image. [2]. The main idea of FAST is based on SUSAN corner detector. Binary Robust Independent Elementary Features are introduces by M. Calonder et al in 2010. This is a multi-purpose point-feature descriptor which can be combined with desired detectors. EnginTola et al in 2010 introduced a local image descriptor, DAISY, which is very efficient in compact computation. Binary Robust Invariant Scale Key point (BRISK) detector is a method for recognition, description, key point adaptation established in 2011 by S. Leutenegger et al. This method is a combination of DAISY and BRIEF descriptors which has the advantage of fast convergence and good numeric stability associated with occupying the least amount of computer recording memory. [3].

Plant identification is needed for weed detection, herbicide application or other efficient chemical spot spraying operations. The key to successful detection and identification of plants as species types is the segmentation of plants from background pixel regions. In particular, it would be beneficial to segment individual leaves from tops of canopies as well. The segmentation process yields an edge or binary image which contains shape feature information. Results indicate that red-green-blue (RGB)formats might provide the best segmentation criteria, based on models of human color perception. The binary image can be also used as a template to investigate textural features of the plant pixel region, using gray image co-occurrence matrices. Texture features considers leaf venation, colors, or additional canopy structure that might be used to identify various type of grasses or broadleaf plants.This is concentrated on just processing a few images for features, our work has concentrated on finding new methods for segmenting plant pixels from background pixels under various environmental and lighting conditions. [4]. There is an increasing interest in automating the process of species identification. The availability and ubiquity of relevant technologies, such as, digital cameras and mobile devices, the remote access to databases, new techniques in image processing and pattern recognition let the idea of automated species identification become reality. This paper is the first systematic literature review with the aim of a thorough analysis and comparison of primary studies on computer vision approaches for plant species identification. After a careful analysis of these studies, we describe the applied methods categorized according to the studied plant organ, and the studied features, i.e., shape, texture, color, margin, and vein structure. Furthermore, we compare methods based on classification accuracy achieved on publicly available datasets. [5].

III. PROPOSED SYSTEM

Decomposing an image into local regions of interest or features is a widely applied technique in Computer Vision used to alleviate complexity while exploiting local appearance properties. Image representation, object recognition and matching, 3D scene reconstruction and motion tracking all rely on the presence of stable, representative features in the image. In Plant identification using neural network has, however despite the clear advantage in speed, the latter approach suffers in terms of reliability and robustness as it has minimal tolerance to image distortions and transformations, in particular to in-plane rotation and scale change. The inherent difficulty in extracting suitable features from an image lies in balancing two competing goals: high quality description and low computational requirements. This is where this work aims to set a new milestone with the BRISK methodology. Perhaps the most relevant work tackling this problem is SURF which has been demonstrated to achieve robustness and speed, only, as evident in our results; BRISK achieves comparable quality of matching at much less computation time. This disadvantage can be resolved by using BRISK method for generating key points from an image, BRISK composes of two sub-methods.

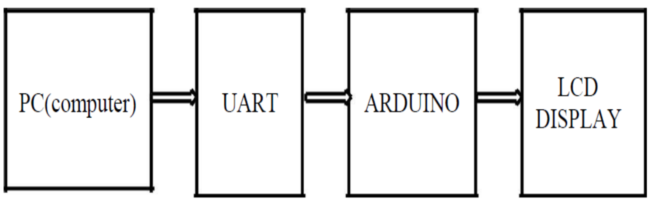


Fig 3 ; Block Diagram

Edges of an image are considered a type of crucial information that can be extracted by applying detectors with different methodology. This research paper presents a brief study of the fundamental concepts of the edge detection operation, theories behind different edge detectors, and some simple self-written Mat lab edge detection functions with the simulation results.

Edge detection is a type of image segmentation techniques which determines the presence of an edge or line in an image and outlines them in an appropriate way. The main purpose of edge detection is to simplify the image data in order to minimize the amount of data to be processed. Generally, an edge is defined as the boundary pixels that connect two separate regions with changing image amplitude attributes such as different constant luminance and tristimulus values in an image.The detection operation begins with the examination of the local discontinuity at each pixel element in an image. Amplitude, orientation, and location of a particular subarea in the image that is of interest are essentially important characteristics of possible edges. Based on these characteristics, the detector has to decide whether each of the examined pixels is an edge or not. Also, several Matlab functions that underlie the principle of

first and second order derivative edge detection techniques are written. The result of the simulations were analyzed and compared to the theoretical result of the edge detectors.

CANNY'S EDGE DETECTION

The Canny edge detector is an edge detection operator that uses a multi-stage algorithm to detect a wide range of edges in images Canny's aim was to discover the optimal edge detection algorithm. To satisfy these requirements Canny used the calculus of variations a technique which finds the function which optimizes a given functional.

The optimal function in Canny's detector is described by the sum of four exponential terms, but it can be approximated by the first derivative of a Gaussian. The algorithm runs in 5 separate steps:

1. Smoothing: Blurring of the image to remove noise.
2. Finding gradients: The edges should be marked where the gradients of the image has large magnitudes.
3. Non-maximum suppression: Only local maxima should be marked as edges.
4. Double thresholding: Potential edges are determined by thresholding.
5. Edge tracking by hysteresis: Final edges are determined by suppressing all edges that are not connected to a very certain (strong) edge.

FAST & FREAK Algorithms in Selected Object Tracking

Object tracking is defined as a method applied in tracking and recognizing the state of a moving object selected by the user or found based on a specific feature in different image frames. For this purpose different features of the selected object like: corners, colour, geometric shape, dimensions etc. are extracted accordingly, and in the next frame, the new location of the object is recognized which makes the tracking of the moving path possible. Several methods are proposed among which extracting and using the key points of the image is the one mostly applied. Here, the FAST is applied and the binary descriptors and indexes of each key point found thereof are introduced by FREAK, inspired by human eye, applied in comparing and recognizing the objects in each frame. Based on this the accuracy and speed of the recognition increase with less memory space needed for implementation. The performance evaluation of this newly proposed method is made through the data set introduced by Mikalajczyk and Schmid. The obtained results indicate a 99% precision on images not subjected to transformation.

Fast Algorithm For Key Point Extracting

The main idea of FAST is based on SUSAN Corner detector , where the corner of a circular area is used for determining lighter and darker neighboring pixels. However, in FAST, not the whole area, but just the pixels on discrete circle of descriptor segment, is evaluated. The criteria for a pixel to be a corner, based on the Accelerated Segment Test (AST), where the circle must have at least 6 connected pixels lighter or darker than the threshold value of the center pixel. The values of the other 16-S pixels are not important. Hence, S value defines the most determined corner angle. Keeping the value of S, as big as possible, while it keeps suppressing the edges (where, S = 8), the repeatability of the corner detector increases. FAST with segment size of 9 (FAST-9) is usually the preferred version. The prevailing question is which one of the first, second, third and the fourth pixels is compared first.

Freak Algorithm

Many of the sampling grids have the ability to compare pairs of pixel intensities. BRIEF and ORB use random pairs. BRISK uses a circular structure, in a sense that points are equally located on concentric circles, like DAISY, use retinal sampling grids which are circular, with the difference of having higher point density around the center. The density of the points decreases exponentially. Each sampling point needs to be smoothened, for having less sensitivity toward noise. BRIEF and ORB use the same kernel for all points in the patch. In order to match the retina model, different kernel sizes are used for each sample point, as in BRISK. The difference here is in exponential change in size and the overlapping receptive fields.

BRISK: Binary Robust Invariant Scalable Key points

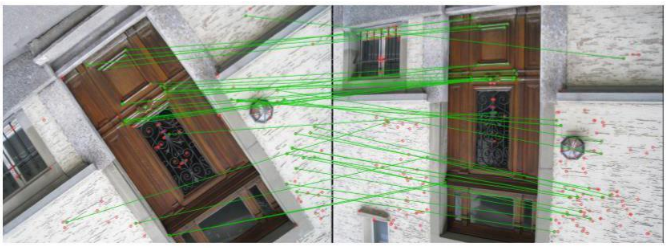


Fig 4 ; Example of BRISK

The inherent difficulty in extracting suitable features from an image lies in balancing two competing goals: high- quality description and low computational requirements. This is where this work aims to set a new milestone with the BRISK methodology. Perhaps the most relevant work tackling this problem is SURF which has been demonstrated to achieve robustness and speed, only, as evident in our results, BRISK achieves comparable quality of matching at much less computation time. In a nutshell, this paper proposes a novel method for generating key points from an image, structured as follows:

Scale-space key point detection: Points of interest are identified across both the image and scale dimensions using a saliency criterion. In order to boost efficiency of computation, key points are detected in octave layers of the image pyramid as well as in layers in-between. The location and the scale of each key point are obtained in the continuous domain via quadratic function fitting.

Key point description: A sampling pattern consisting of points lying on appropriately scaled concentric circles is applied at the neighborhood of each key point to retrieve gray values: processing local intensity gradients, the feature characteristic direction is determined. Finally, the oriented BRISK sampling pattern is used to obtain pair-wise brightness comparison results which are assembled into the binary BRISK descriptor. Once generated, the BRISK key points can be matched very efficiently thanks to the binary nature of the descriptor. With a strong focus on efficiency of computation, BRISK also exploits the speed savings offered in the SSE instruction set widely supported on today's architectures. The key stages in BRISK, namely feature detection, descriptor composition and key- point matching to the level of detail that the motivated reader can understand and reproduce. It is important to note that the modularity of the method allows the use of the BRISK detector in combination with any other key point descriptor and vice versa, optimizing for the desired performance and the task at hand. Their AGAST is essentially an extension for accelerated performance of the now popular FAST, proven to be a very efficient basis for feature extraction. With the aim of achieving invariance to scale which is crucial for high-quality key points.

Brisk Detector Repeatability

The detector repeatability is calculated as the ratio between the corresponding key points and the minimum total number of key points visible in both images. The correspondences are identified by looking at the overlap area of the key point region in one image (i.e. the extracted circle) and the projection of the key point region from the other image (i.e. ellipse-like): if the region of intersection is larger than 50% of the union of the two regions, it is considered a correspondence. Note that this method is largely dependent on the assignment of the key-point circle radius, i.e. the constant factor between scale and radius. We choose this such that the average radii obtained with the BRISK detector approximately match the average radii obtained with the SURF and SIFT detectors.

IV. EXPERIMENTAL RESULTS

MATLABversion2015 software version is used for this project Digital Image Processing technique. The MATLAB coding for Plant Species Identification is given as below

Fig 5; BRISK matching point

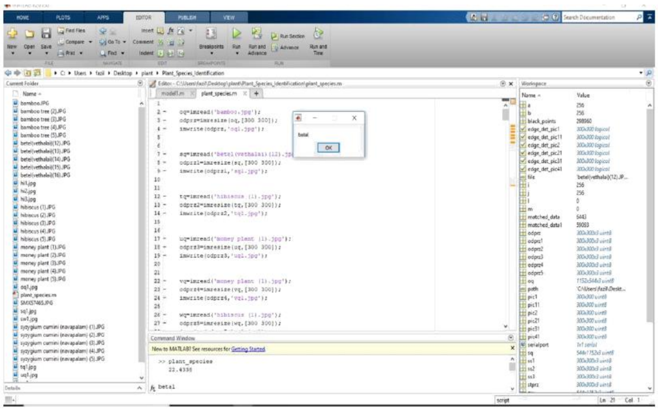


Fig 6; simulated output plant species name

Fig 7; Arduino Kit

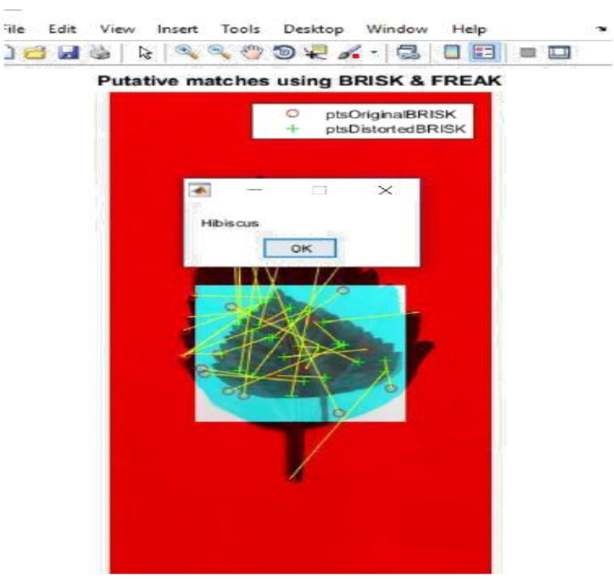


Fig 8; Name of the matched plant species

Figure 9; Hibiscus leaf

Fig 10; cuminis leaf

Fig 11; Betal leaf

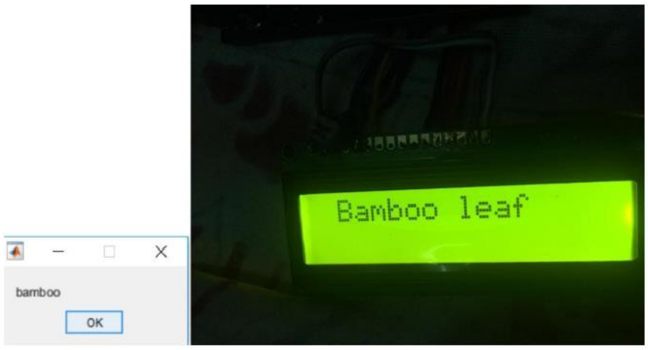


Fig 12; Bamboo leaf

Fig 13; No matching images

V. CONCLUSION AND FUTURE SCOPE

The Image processing is a vast field with numerous applications, In this project the identification of plant spaces is done by digital Image processing. The main aim is to identify the contour of the leaf using edge detection which is done in MATLAB. A basic image of a leaf is resized and a 3D image to 2D image. The algorithm like edge detection, Fast and Freat, Brisk are used. Brisk relies on an easily configurable circular sampling pattern. Fast is used to detect the lighter and darker neighboring pixels. Edge detector used to perform certain methods of evaluation, provides as a good understanding and a possible ways of finding out the effect of each developed detection model. The main aim of this method is that contour features are mainly used in the matching process; the method is scale invariant and computationally efficient. Robotics play a major role in day to day life. AI's have become a part in all fields. This species Identification of plants will be helpful in future for agriculturists these process can be embedded in robots so that the robot can pick the fruits and vegetables from the plants making it easier for workers. Leaves are most common part of the plant they can be used for species identification from this the medicinal uses, disease detection can be done Millions of plants are endangered so these will provide the existence of those plants and their uses. Robots can also provide fertilizers as required.

ACKNOWLEDGMENT

The authors wish to thank Head of the department ECE and management of Sri Shakthi Institute of Engineering and Technology Coimbatore for his support throughout completion of this project.

REFERENCES

[1] Hassan Hajjdiab and lman AI Maskari(2011)" Plant Species Recognition Using Leaf Contours" College of Engineering and Computer Science,Abu Dhabi University 978-1-61284-896-9/111$26.00 ©2011 IEEE

[2] Lei Zhang,JunKong,Jiayue Ren(2008)" Plant Species Identification Based on Neural Network " Network Information Center, Northeast Changchun, Jilin Province China 2Computer School, Northeast Normal University, Changchun, 978-0-7695-3304-9/08 $25.00 © 2008 IEEE

[3] Jana Wäldchen(2016) "Plant Species Identification Using Computer Vision Techniques" November 2016 / Accepted: 24 November 2016

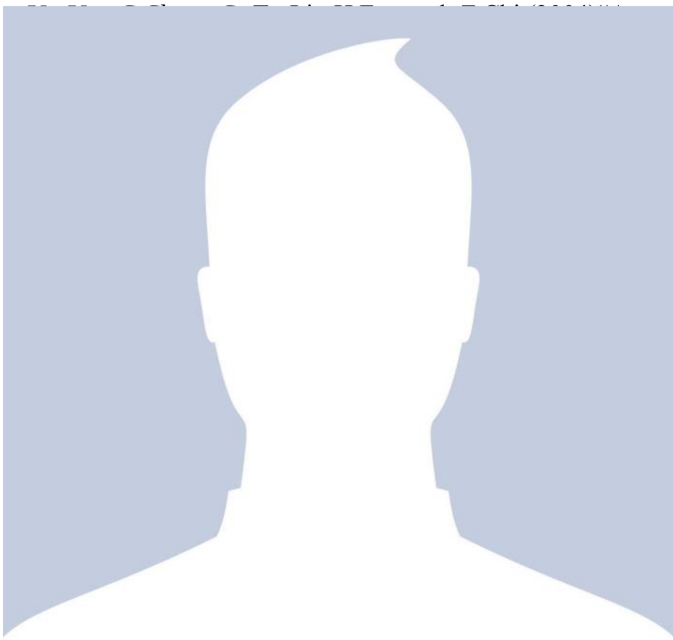
[4] George E. Meyer, Tim Hindman, and Koppolu Laksmi(2016) "Machine Vision Detection Parameters for Plant Species Identification" University ofNebraska, Lincoln NE 68583

[5] S.Khachikian,M.Emadi (2016)" Applying FAST & FREAK Algorithms in Selected Object Tracking" MS Student, Dept. of Electrical Engineering, Jahad Daneshgahi Institute, DOI:10.15662/IJAREEIE.2016.0507003

[6] D. Huang, G. C. Guo, Y. C. Lin, N. C. Shen, and Y. S. Zhang(2006), "Short response Hilbert transform for edge detection," November 2016 / Accepted: 24 November 2016 J. J. Ding, S. C. Pei, J. CVGIP, 2007.

[7] Harris and M. Stephens(2007) "A Combined Corner and Edge Detector," Proceedings S. C. Pei and J. J. Ding, "Improved Harris' algorithm for corner and edge detections," accepted by ICIP 2007.

[8]



[9]

[10]

[11]