International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:10, No:8, 2016

# Modern Detection and Description Methods for Natural Plants Recognition

Masoud Fathi Kazerouni, Jens Schlemper, Klaus-Dieter Kuhnert

***Abstract*—**Green planet is one of the Earth's names which is known as a terrestrial planet and also can be named the fifth largest planet of the solar system as another scientific interpretation. Plants do not have a constant and steady distribution all around the world, and even plant species' variations are not the same in one specific region. Presence of plants is not only limited to one field like botany; they exist in different fields such as literature and mythology and they hold useful and inestimable historical records. No one can imagine the world without oxygen which is produced mostly by plants. Their influences become more manifest since no other live species can exist on earth without plants as they form the basic food staples too. Regulation of water cycle and oxygen production are the other roles of plants. The roles affect environment and climate. Plants are the main components of agricultural activities. Many countries benefit from these activities. Therefore, plants have impacts on political and economic situations and future of countries. Due to importance of plants and their roles, study of plants is essential in various fields. Consideration of their different applications leads to focus on details of them too. Automatic recognition of plants is a novel field to contribute other researches and future of studies. Moreover, plants can survive their life in different places and regions by means of adaptations. Therefore, adaptations are their special factors to help them in hard life situations. Weather condition is one of the parameters which affect plants life and their existence in one area. Recognition of plants in different weather conditions is a new window of research in the field. Only natural images are usable to consider weather conditions as new factors. Thus, it will be a generalized and useful system. In order to have a general system, distance from the camera to plants is considered as another factor. The other considered factor is change of light intensity in environment as it changes during the day. Adding these factors leads to a huge challenge to invent an accurate and secure system. Development of an efficient plant recognition system is essential and effective. One important component of plant is leaf which can be used to implement automatic systems for plant recognition without any human interface and interaction. Due to the nature of used images, characteristic investigation of plants is done. Leaves of plants are the first characteristics to select as trusty parts. Four different plant species are specified for the goal to classify them with an accurate system. The current paper is devoted to principal directions of the proposed methods and implemented system, image dataset, and results. The procedure of algorithm and classification is explained in details. First steps, feature detection and description of visual information, are outperformed by using Scale invariant feature transform (SIFT), HARRIS-SIFT, and FAST-SIFT methods. The accuracy of the implemented methods is computed. In addition to comparison, robustness and efficiency of results in different conditions are investigated and explained.

Masoud Fathi Kazerouni is with the Institute of Real-time Learning Systems, University of Siegen, D-57076 Siegen, Germany (corresponding author to provide phone: +49-(0)271 740 2559; e-mail: masoud.fathi@uni-siegen.de).

Jens Schlemper and Prof. Dr. –Ing. Klaus-Dieter Kuhnert are with the Institute of Real-time Learning Systems, University of Siegen, D-57076 Siegen, Germany (e-mail: schlemper@fb12.uni-siegen.de, kuhnert@fb12.uni-siegen.de).

***Keywords*—**SIFT combination, feature extraction, feature detection, natural images, natural plant recognition, HARRIS-SIFT, FAST-SIFT.

## I. Introduction

ALTHOUGH that earth is approximately 4.5 billion years old, the history of life began about 0.7 billion later, and both mentioned years are just two estimated values. It is worth mentioning that life has always developed on earth gradually, although this pneumonia can be considered as a constant occurrence. In order to look back through the past years, fossil records prove the mentioned fact. Plants have played a great and irrefutable role in change of the earth's climate after Ordovician extinction which happened more than 425 million years ago. Discovery of earth history is possible by studying microscopic ancient plants. In addition to their impacts on earth history, they have had different influences on human life too. Plants contributed to development of human civilization as they appeared close to rivers where they are origins of modern life. They influence on the climate and its variations, and their impacts are undeniable seriously. Due to scientific findings, perspiration and breath of plants lead to cool the atmosphere. They consume carbon dioxide in the process of photosynthesis and lead to reduce the amount of carbon dioxide. This reduction has an indirect cooling effect. Furthermore, climate change alters the life cycles of plants. It is an interesting point and approves the relationship between them. Additionally, plant species traits are the attributes that most directly affect ecosystem processes. They contribute to healthiness of ecosystem. Plants produce all food of living organisms, even their own food in order to survive and grow.

Plants have been used by human for centuries to soothe and improve discomfort and various health problems. Recently people prefer natural treatments for their ailments. Therefore, medicinal role of plants has been increased. Rich resources of ingredients can be utilized in medicine and industry which can be considered as a fundamental role. These types of resources are medicinal plants which are usable in drug development and synthesis. Horizon of plant genetics is brightening in agricultural applications. Advancements of plant genetics have led to new research areas, fields of studies, and technologies which are needed to overcome new challenges of human life. In order to secure global food, plant genetics are key components which can be applied in agriculture too.

According to new applications of plants in different fields, importance of them has been highlighted more than before. Computer vision is field related to new technologies. Many

new applications are connected to advances in this field. Automatic plants recognition is an exigency in modern world. Importance of an automatic plant recognition system increases, while recognition of plants is a big challenge, even for experts and botanists. They cannot recognize a plant with just either a glance or a blink. Therefore, they use some books and references that contain plants information to identify the species. It is a common way for identification and recognition of plants. It is a time-consuming method too. Time is an important factor in research and industry. Having an automatic recognition system helps scientists, researchers, managers, and engineers in labs, offices, and factories. Efficient plant recognition system contributes involved people in plant projects doing their tasks without presence of botanists.

Some components of plants are common and can be taken in consideration for plant species recognition's purposes. Finding a useful characteristic of plant is an important step to develop a recognition system. Stability, consistency, and independency have been considered as three important factors of selected part. Due to nature of leaves and their seasonal characteristics, they have been selected and used as basis for identifying plants. Leaves characteristics like size and color might vary in a single plant, but their typical shapes are the same as others.

To develop an accurate and general system, additional considerations and factors are added and should be solved carefully. Plant recognition systems used today work under constrained conditions. Generalization can be helpful and useful in various applications of system. To generalize the system, recognition system should be robust in different natural environments. Light intensity changes every hour. This parameter affects performance of a plant recognition system. Variation of light intensity plays a significant role for improving and enhancing existing recognition systems. Therefore, this parameter is taken into consideration to generalize plant recognition system. Another considered parameter is distance from camera to plants. In agricultural application, this parameter is useful. Some new features will be added to the system. The system will be more applicable and efficient. There is no consideration in camera selection. Therefore, the type camera does not affect the proposed algorithm. Due to importance of generalization, it contributes to have general system. The used dataset has some unique characteristics. They will be explained in detail. The mentioned parameters and considerations have been used to put the research in right direction.

One essential requirement of a computer vision system is to acquire camera images as an important process. The system should act similar to a human brain to recognize different plant species, and images are like the images which are seen by human eye. The basic idea of plant recognition system is originated from shape of leaf. Shape and structure of leaves are usually used to invent a plant recognition system. Due to distance changes, extracting useful features and information are critical points and beginning steps. It is essential to analyze the constituent shapes. Using an appropriate technique aims at processing information correctly and efficiently. Collection of information is another necessity. Many methodologies have been proposed and used to analyze plant leaves in automatic procedures. Most of them use shape recognition methods to find out contour of leaves' shapes and use the found contours for recognition task. In [1], geometrical parameters are applied such as area, maximum length, maximum width, and perimeter. These methods are not effective to have a general recognition system. When the distance between camera and plants changes, there are not the same numbers of leaves in each image. Therefore, similar methodologies are useless. Color features have been reported in [2] in addition to geometrical features. A large number of methodologies have been introduced in [3]-[8] for shape representation. Curvature scale space (CSS) method has been proposed in [9] and K-nearest neighbor (KNN) method has been applied to classify chrysanthemum leaves. Region-based shape recognition techniques have been used in [10] for leaf image classification. Artificial images have been utilized in mentioned literatures.

Nowadays, researchers and scientists tend to use modern approaches such as SIFT [11], [12] and SURF [13] to get features of images. Stable local feature detection and representation is a fundamental characteristic of these methods. SIFT algorithm is resistant to common image deformations. SIFT algorithm includes a feature detector and a feature descriptor which would be used separately due to considered application. The detector detects and extracts a large collection of keypoints from one image. In fact, keypoints are oriented disks attached to blob-like structures of the image. It is an invariant method to image translation, scaling, and rotation. Therefore, it is an accurate feature detector and descriptor. In [14], SIFT algorithm is the basis of used methods to recognize plants automatically. In [15], SIFT features are extracted to classify flowers. SIFT method can be applied in different fields such as object recognition, video tracking, and even robotic mapping.

There are other modern methods to do feature detection. FAST [16] algorithm is a machine learning approach to detect corners. From real time application point of view, it is a fast algorithm. It provides a lot of features in a short time. This approach has an important disadvantage. In fact, it is not robust to high levels of noise. In [14], [17], FAST approach is a part of combinational methods for automatic plant recognition.

Another interesting approach for feature detection is HARRIS [18] algorithm. It was proposed by Chris Harris and Mike Stephens. Since corners show a variation in the gradient in the image, this variation can be utilized to do detection procedure perfectly. HARRIS method was used as a component of used combined methods in [14], [17] to do automatic plant classification.

In most natural language processing approaches, documents or sentences are represented by a sparse bag of visual words (BoW) [19] representation. In this model, one text can be shown as the bag of its word and order of words is ignored by the model. This technology has become major method of image classification and object categorization in computer

vision recently. In [14], [17], BoW method is applied to implement an automatic plant recognition system. This approach has been useful, and results show its efficiency.

An initial step is to choose a useful dataset. There are several plant datasets such as Flavia dataset [20], ImageCLEF dataset, Leafsnap dataset, and Intelengine dataset. Each dataset contains different plant species. Flavia dataset was used in [14], [17]. A suitable dataset is made and prepared according to final targets. This dataset fulfills demands of generalized system. The created dataset includes 1000 natural images of four different plant species. These plant species are common plants of Siegerland, a region in Germany. The types of the plants species are Hydrangea, Amelanchier Canadensis, Acer Pseudoplatanus, and Cornus.

Computer vision, known as a field with many discoveries and influences on technology, has the goal to make computers more efficient to perceive, process, and understand visual data such as images and videos. In comparison to the last decades, applications of computer vision techniques have been spread in various fields. Invention of a general system to recognize and identify plant species is one necessity. This system should be automatic.

This paper reports a new general system to recognize natural plant species. Modern methodologies have been used to implement the automatic system. In addition to find accuracy of implemented algorithms, some experimental calculations have been done.

The remainder of this paper is organized in five sections. In Section II, a general overview is provided. Section III introduces the forming steps of the algorithm of system. As a result, the experiments and results are presented in Section IV. Section V discusses conclusion, some remaining problems, and future work.

## II. General Review

Plants are enormous resources of energy and also play an important role in environment obviously. Every plant has its own properties. Some of them are even unique. In a world with plenty species of plants, collection of their information is useful in different areas definitely. Data collection becomes more important, when there are some extinct plants. A reliable, accurate, and general system is helpful to identify plants without needs of botanists or experts. The system should be compatible with different conditions, thus generalization of it is essential, useful, and undeniable. Design and implementation of such a system can be done step by step. Some data are highly very complex, variable, and redundant. Images, videos, and real world date are only some of them. Each image can be considered as a structure. This structure is consisted of adjacent pixels. Naturally a good modeling of them leads to better understanding of the structure and hidden information in image. A decisive part of classification tasks is always right, correct, and efficient information extraction. Determination of this automatic task is very important to achieve desired goals. In general, it is essential to extract as much information as possible form image. In real world applications, the choice of specific techniques or algorithms depends on the goals of each individual project basically.

Correct information extraction contributes to facilitate the other steps. In fact, the first operation on image is a low-level operation and it is called feature detection. It is an essential step to fulfill projects and researches of some fields like pattern recognition and image processing. It has various usages and applications in image analysis and computer vision applications. Afterwards, feature extraction will be performed. Both operations provide attributes. Interest points can be corners, blobs, or edges in an image. Efficiency of a detection method depends on how much rich information it gives about the nature of image. Feature detection's concept refers to techniques that help to compute abstractions of image's information and investigate and make local decisions at every image point and pixel whether there is a feature of a given type at that point or not. There are different feature detection techniques; for instance, HARRIS and FAST are two common techniques to do features detection. Kazerouni et al. [14], [17] used combination of different modern techniques to create an automatic plant recognition system and achieved high accurate results. In addition, feature extraction plays an important and effective role, since this step affects the reduction of dimensionality. A reduced set of features is expected by this step. Also, relevant features should be extracted to form feature vectors. To distinguish and recognize one specific object from others, a good feature set should be extracted and discriminating information should be included. It must be as robust as possible, because robustness is a factor which aims to prevent generating different feature codes for the objects in the same class. Extraction of features is not sufficient to get the information from image. This step has observable impacts on efficiency of a recognition system. Thus, selection of an appropriate method is very important and should be done according to raw images. SIFT algorithm is a modern technique to do feature extraction. It is the most widely used local feature-based algorithm according to its unique characteristics and properties. Typically SIFT descriptors can be visualized as boxes with many arrows. SIFT method provides a set of features of an object that are not affected by scaling and rotation. In SIFT method, a keypoint descriptor can be represented consistent with the orientation, it assures to have a method which is invariance to image rotation. In [14], SIFT has been used to distinguish plant species.

A set of descriptors are needed for further process. A current efficient technique is BoW to represent the needed set of descriptors in a compact organization. Fundamentally, it has the role of translating a very large set of descriptors into a single sparse vector. Additionally, this allows use of machine learning algorithms that by default assume that the input space is vectorial. In [14], [17], this algorithm is applied to recognize plant species. There are different learning techniques, either supervised or unsupervised. Support vector machines (SVMs) [21]-[23] and Bayesian classifiers [21], [24], [22] are the most common and popular techniques between supervised learning methods.

Training and testing sections are two fundamental steps for solving a classification problem. Decision theory approaches

are behind classification solutions. There are different classification methods for training. Supervised and unsupervised learning methods can be performed to solve machine learning problems. Meanwhile, statistical learning theory is the origin of SVM, which is actually a supervised classification method and often yields nice classification results from complex and noisy data. In order to achieve high accuracies, SVM was selected in [14], [17] as a fundamental component of implemented algorithms. This SVM was based on radial basis function (RBF).

Four different plant species are included in this dataset, and the number of classes is 4. All experiments are performed over the dataset to observe the results. At the end, a test dataset, comprised of 336 images, is used for test of the implemented systems.

## III. Plant Recognition System Approach

The proposed system undergoes four stages, image preprocessing, feature detection and extraction, modeling and training, SVM classification. These stages are debriefed in the four subsections, *A*, *B*, *C*, and *D*, respectively. The simplified block diagram is shown in Fig. 1. It summarizes the overall process of proposed system.

### *A. Image Preprocessing*

The images which have been used as inputs have very complicated nature. Each image contains a scene of each plant with different objects such as leaves, stems, backgrounds, etc. All inputs are RGB images. RGB to grayscale conversion is performed to initialize the approach. This phase is converting an input image from one space to another one, RGB to grayscale. This conversion is done by (1).

$$Y = 0.299\,R + 0.587\,G + 0.114\,B. \tag{1}$$

where *R*, *G*, and *B* correspond to color of each pixel.

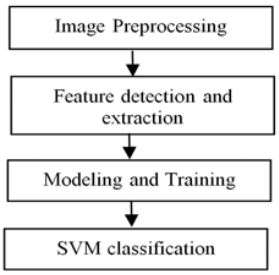


Fig. 1 Overall process of proposed plant recognition system

### *B. Detection and Description of Features*

Second stage of the system is keypoint detection and description which are so important parts. This stage has to have a powerful stability for different condition of various scenes, different light intensities and illuminations, scaling, geometric, and shift transformations. To obtain relevant information from image data, this stage is essential and inevitable. There are some modern methods like SIFT, FAST, and HARRIS to detect keypoint. The mentioned methods can be combined [14], [17], therefore it would be possible to combine advantages of different methods and use the advantages simultaneously.

Detection of a set of keypoints is performed for each used method. Corners are very important local features in an image as these regions always vary largely in intensity in all directions. Corner detection leads to extract accurate information from image. Another advantage is reduction of calculations. Two common methods are HARRIS and SUSAN (Smallest Univalue Segment Assimilating Nucleus) [25]. Basically, Harris corner detection method is selected as one of the used method. It is superior to SUSAN corner detection method on the whole.

HARRIS method is one of the performed detection methods. Gradient of each pixel is calculated in HARRIS corner detection method, and absolute gradient values are computed in two directions. Then, a comparison will be done. If the values are both great, then the pixel is considered as a corner. In (2), HARRIS detection method is defined.

$$\begin{aligned} &R = \det(M) - k\,tr^2(M), \\ &M(x,y) = \begin{pmatrix} I_u^2(x,y) & I_{uv}(x,y) \\ I_{uv}(x,y) & I_v^2(x,y) \end{pmatrix} \\ &I_u^2(x,y) = X^2 \otimes h(x,y), \\ &I_v^2(x,y) = Y^2 \otimes h(x,y), \\ &I_{uv}(x,y) = XY \otimes h(x,y), \\ &h(x,y) = \frac{1}{\Pi} e^{-\frac{x^2+y^2}{2}} \end{aligned} \tag{2}$$

where $I_u(x,y)$ and $I_v(x,y)$ have been considered as partial derivatives of the gray values in both directions at point $(x,y)$, direction *u* and direction *v*, respectively, and the second-order mixed partial derivative is shown by $I_{uv}(x,y)$ [26]. In (2), *k* is an empirical value and $h(x,y)$ is a Gaussian function. Convolution of the gray values and difference operators in direction u and v can be carried out for the first-order directional differentials, *X* and *Y*. In order to reduce the effect of noise, Gaussian function is applied since first-order directional differentials are sensitive to noise. If *R* exceeds certain threshold, then take the point as a corner. This method is not usable in re-scaled applications. Although it is a rotation invariant method, but change of point of view, light intensity, and illumination can affect the method's efficiency.

It is a required to have high repeatable local information content. To achieve this purpose, another well known method, FAST, is utilized. The reason behind this method is to develop an algorithm for use in real time applications. In addition to the method's fastness, it has adequate repeatability and efficiency. FAST method compares pixels only on a circle of fixed radius (16 pixels) around the corner candidate point. Every circle's pixel will be labeled from 1 to 16 clockwise. In this method, pixels are investigated and checked whether they can be desired and interest points. *P* is a pixel with *Ip*, which is assumed to be the intensity of the pixel. A threshold intensity value, 20% of the current pixel, is set. Here, it is called *Threshold*. Then, a circle of 16 pixels around *P* is considered for further procedure. This pixel will be a corner if

there exist n contiguous pixels in the surrounding circle which are lighter than *Ip+Threshold* or darker than *Ip-Threshold*. *n* is defined as 12 in original method. If the value is not defined less than 12, this method does not reject many candidates. A machine learning approach is applied to solve this weakness and failure. To speed up the FAST method, this procedure can test only four pixels at 1, 9, 5, and 13. First of all, 1 and 9 are tested if they are too brighter or darker. If so, then we check 5 and 13. If none of them are the case, then this pixel is not a corner. This procedure will be continued for all pixels of image. Although this method is faster than other corner detection methods, it is not robust to high levels of noise and there is dependency on the mentioned threshold.

The next used method is SIFT, a method to detect keypoints and extract descriptors. This is a popular and modern method in computer vision application and proposed by Lowe in [12]. Its detector extracts a collection of keypoints from an input image. Building Gaussian scale space, keypoint detection and localization, orientation assignment, and keypoint descriptor are the method's steps [27].

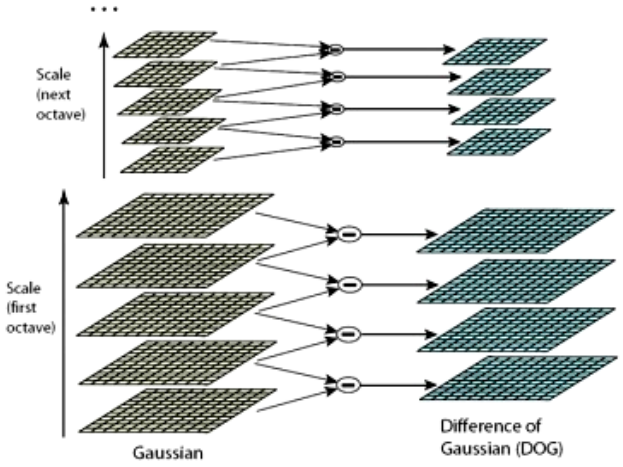


Fig. 2 Octave of scale space and procedure

To detect corners, it is impossible to apply the same windows for detection of keypoints with different scales. And that is why Difference of Gaussians (DoG) is utilized as it is an approximation of Laplacian of Gaussian (LoG). It can be considered as one type of scale-space filtering. The difference of Gaussian blurring of an image in various octaves is computed to obtain DoG. For each octave of scale space, the input image is convolved with Gaussians to create a set of scale space images. It is done repeatedly. In next step, adjacent Gaussian images are subtracted to build the DoG images. After that, the Gaussian image is down-sampled by a factor of 2, and the procedure repeated. The convolved images are grouped by octave, and we obtain a fixed number of DoG per octave. The next step of SIFT method is keypoint detection and localization. Obviously, local maxima or minima of the DoG images are keypoint candidates. In this step, comparison of pixels in DoG images to neighbors is performed. 26 neighbors in 3x3 regions in the current and adjacent scales of each pixel are taken into consideration and compared to the intended pixel. In Fig. 2, the described procedure is shown.

Local maximum or minimum pixels are potential keypoints finally. Now, keypoint candidates should be filtered for getting more accurate results. The solution is to use Taylor series expansion of scale space. It leads to have more accurate location of extrema. The intensity at this extremum is compared to a contrast threshold (0.03). If this value is less than the threshold, it will be rejected automatically for next process. It is essential to remove edges as DoG has higher response for them. To remove the edges, a similar method to Harris method is used. A 2x2 Hessian matrix (H) is applied for finding curvature. An edge threshold is defined as a ratio. This threshold equals 10 in [27]. If it is greater than the threshold, the keypoint is removed. Therefore, strong keypoints are obtained.

In next step of the method orientation of keypoint is determined by means of computing gradient histogram in the neighborhood of the keypoint. This step contributes to achieve invariance to image rotation. A 36-bin orientation histogram is produced. It is weighted by the gradient magnitude and a Gaussian window with a σ that is 1.5 times the scale of the keypoint. Highest peak of the histogram and other peaks with 80% of highest peak are taken to compute the orientation. The outcome will be keypoints with same location and scale, but different directions.

The final step of SIFT method is to compute keypoint descriptors. For each keypoint, a neighborhood is considered and divided into 16 sub-blocks and this neighborhood is 16x16 one. The size of each sub-block is also important and it is defined to be 4*4. An 8-bin orientation histogram is built for each sub-block. Therefore, 4*4*8 is the size vector which is equal to 128. The obtained vector represents keypoint descriptor. One normalization is performed to enhance invariance to changes in illumination [12].

SIFT method is used as a component of combination methods to extract detected keypoints of HARRIS and FAST methods.

### *C. Modeling and Training*

As it was described, BoW is one of the new and most popular representation methods for object categorization and classification to achieve promising results. Concepts of visual words and BoW are applied in this stage. Certainly, visual words have revolutionized the field of computer vision. Pixels of an image can be represented as words in a text. It contributes to index for local image keypoints. Images of dataset have very large variations and it is one of the main problems. For instance, various clutters in background and foreground can be observed in images. Scale, distance, and illumination have been varied in different images. It can solve the problem of having an understandable image representation for machines.

The idea of BoW is to use compact descriptors from local features, have a finite number of clusters, and create a visual vocabulary. Splitting an image into small image patches steers to have a set of words. By using learning algorithms, it would

be possible to group them. Each group can be considered as one word. The next step will be mapping each part of image to one of the obtained visual words.

In the other words, clustering many descriptors is needed to create a visual vocabulary. To do clustering step, K-means method [28] is applied as an efficient method. This method is an unsupervised learning algorithm, which can be applied and utilized in most of clustering problems. The method's procedure follows a simple way. The main idea is to use k centers for each cluster. Each cluster's center will be utilized as a visual word. The goal of this procedure is to assign a cluster to each obtained feature descriptors. This method finds the positions $y_i$, i= 1, 2, 3, …, k, of the clusters that minimize the square of the distance from the obtained feature descriptors to the cluster.

$$\arg\min_{c} \sum_{i=1}^{k} \sum_{x \in c_i} d(x, \mu_i)^2 = \arg\min_{c} \sum_{i=1}^{k} \sum_{x \in c_i} \| x - \mu_i \|_2^2 \quad (3)$$

K-means clustering utilizes the Euclidean distance. One important point is to decide the number of clusters. Correct selection of it is very important, otherwise the whole procedure and system will be invalid. One empirical solution is to find the best number of clusters by doing the method with various numbers. Then, it would be possible to check the whole performance of the system with different numbers of clusters. It is essential to emphasize that this K-means clustering is utilized for quantization of the feature space. Quantized feature space suggests a vector representation of images that indicates the frequency of the visual words, which can be utilized in conjunction with some vector-based kernels or similarity measures for matching or categorization of image content. This method is fast, robust, and understandable simply. Due to change of distance from plants to the camera, a vocabulary is created for each distance. It is a critical and vital point in designing of the system.

In the other words, the image's features can be denoted into words by specifying which visual word are actually nearest in the feature space based on the Euclidean distance between the cluster centers and the input descriptor. In addition to, each extracted region from image has to be assigned to the corresponding visual word in test phase. This model will be used for new images in a certain procedure. Firstly, keypoints of the new image will be detected. Secondly, descriptors will be extracted from them. Thirdly, nearest neighbor in vocabulary will be computed for each descriptor. The last step is to build a histogram. In the histogram, $i^{th}$ value will be the frequency of the $i^{th}$ vocabulary word. The histograms will be fed to a classifier to predict labels and classes for images. Classifiers need fixed dimension feature vectors.

### *D. SVM Classification*

This step includes training the dataset. SVM is one powerful tool in computer vision problems. The main concept of SVM is to rely on preprocessing the data to represent patterns in a higher dimension than the original feature space. For example, separation of data from two categories can be carried out by a hyper-plane when an appropriate mapping is applied. Also, the hyper-plane has the largest distance to the nearest training data point of any class (it is then called functional margin). The original SVM method is extended to regression, classification, and clustering problems. SVM method has some special benefits. As it was mentioned before, this method is so effective in high dimensional spaces. This method does not lose its efficiency even where number of dimensions is greater than the number of samples. It uses an iterative training algorithm to construct an optimal hyper-plane and minimizing error. In other words, given labeled training data (*supervised learning*), the algorithm outputs an optimal hyper-plane which categorizes new examples. In [29]-[33], mathematical aspects of SVM have been explained. SVM tries to minimize an error function which is helpful for dividing SVMs. SVM models have been divided into four different types due to the form of the error function.

Basically, the first step is to train the SVM on a set of images and construct the training matrix. This step is followed by putting histogram responses for each class. Then, labels must be set up for each training image. For example, if there are two classes, leaf and non-leaf based on images, it would be necessary to specify which row in the training matrix corresponds to a leaf and a non-leaf. In this example, if the $2^{nd}$ element of the label matrix is -1, it proves that the $2^{nd}$ row of the training matrix falls into the non-leaf class. Therefore, a 1D label matrix is defined and each element of this matrix corresponds to one row in 2D matrix. The next step is consideration and setting up of SVM parameters. Accurate adjustment of parameters has to be done to use advantages of SVMs. Nu-Support vector classification is the selected and used method. It has the benefit of utilizing a parameter nu in order to control the number of support vectors. Each SVM model has some parameters like C, P, and Nu. Nu parameter is an important parameter for optimizing the problem. It should be in the interval (0, 1]. This parameter is investigated in the next section. Also, the Nu parameter represents the lower and upper bounds on the number of examples that are support vectors and that lie on the wrong side of the hyper-plane, respectively. In the other words, a SVM is useful because it is able to locate a separating hyper-plane in the feature space and classify points in the space without representing the space explicitly by utilizing and applying kernel function. It plays the role of the dot product in the feature space. This technique prevents from the computational burden of explicitly representing the feature vectors which can be considered as a helpful factor. As kernel function determines the feature space in which classification should be performed, it is very important to select a suitable kernel function. It is worthy to mention that an SVM's operation can be correct completely however the designer does not know how it is really working and doing its task. Thus, it does not depend on knowledge of the designer.

The kernel's type is RBF, which is a good choice in most cases. The origin of the reason is their localized and finite responses across the whole range of the real x-axis.

$$K(x_i, x_j) = e^{-\gamma \|X_i - X_j\|^2}, \quad \gamma > 0. \qquad (4)$$

In general, Gamma is an adjustable parameter of RBF kernel function. Additionally, this parameter defines how far the influence of a single training example reaches, with low values meaning 'far' and high values meaning 'close'.

Doing the test procedure is the last step of the whole system. Prediction of labels is performed for the test dataset. Thus, the whole system can be investigated and evaluated to know its final accuracy and practical efficiency. In the next section, some experiments and computations are done.

## IV. Experiments

Experiments have been conducted, and the results have been investigated in current section. The dataset is divided into two subsets, named training dataset and test dataset. The training dataset contains 664 images, while the test dataset has 336 images. Table I shows number of images for each distance and subset. The prepared dataset can be considered as a dataset of natural images. There are large variations in appearance, such as scale, illumination, pose, and background clutter in these natural images. The images have been taken from different distances, angles, and views. Light intensities and illumination, weather conditions, and time of taking the images are changed during preparation of the dataset.

TABLE I
Number of Images in Different Distances

| Dataset | 25 cm | 50 cm | 75 cm | 100 cm | 150 cm | 200 cm |
|---|---|---|---|---|---|---|
| Number of images for Training | 160 | 160 | 160 | 160 | 12 | 12 |
| Number of images for Test | 80 | 80 | 80 | 80 | 8 | 8 |

Intel® Core™ i7- 4790K, CPU @ 4.00 GHz, and installed memory (RAM) 16.0 GB are specifications of the used machine. Different aspects of the system are taken into consideration and the results are illustrated in detail completely.

Accuracy of classification is calculated in different distances. In distance 25 cm, the highest accuracy belongs to the system implemented by SIFT method. The percentage of accuracy is 92.50%. In this distance, HARRIS-SIFT method has 88.75% as percentage of accuracy. The lowest accuracy has been gained by means of FAST-SIFT method. These results have been shown in Table II.

The lowest percentage of accuracy in distance 50 cm is 92.50% and shown in Table III and this accuracy belongs to implemented system by FAST-SIFT method. The other systems have the same accuracies which are equal to 96.25%.

When the distance increases from 50 cm to 75 cm, performance of the FAST-SIFT method is the same as HARRIS-SIFT method. The accuracy of both systems equals 92.50%. When SIFT method is applied in this distance, the percentage of accuracy is 93.75%, with 75 correct and 5 wrong predictions. Table IV shows accuracy of classification in 75 cm.

TABLE II
Accuracy of Classification (Distance 25 cm- K = 1000)

| Used Method | Correct Predictions | Wrong Predictions | Percentage of Accuracy |
|---|---|---|---|
| SIFT | 74 | 6 | 92.50 |
| FAST-SIFT | 69 | 11 | 86.25 |
| HARRIS-SIFT | 71 | 9 | 88.75 |

TABLE III
Accuracy of Classification (Distance 50 cm- K = 1000)

| Used Method | Correct Predictions | Wrong Predictions | Percentage of Accuracy |
|---|---|---|---|
| SIFT | 77 | 3 | 96.25 |
| FAST-SIFT | 74 | 6 | 92.50 |
| HARRIS-SIFT | 77 | 3 | 96.25 |

TABLE IV
Accuracy of Classification (Distance 75 cm- K = 1000)

| Used Method | Correct Predictions | Wrong Predictions | Percentage of Accuracy |
|---|---|---|---|
| SIFT | 75 | 5 | 93.75 |
| FAST-SIFT | 74 | 6 | 92.50 |
| HARRIS-SIFT | 74 | 6 | 92.50 |

Consideration of system's performance is carried out, while the distance increases again. The last calculation of accuracy belongs to three distances, which considered as one set in calculations and shown in Table V. SIFT method has a good performance, and its percentage of accuracy is equal to 96.87%. The other methods, FAST-SIFT and HARRIS-SIFT, have the percentage of accuracy 95.83%.

TABLE V
Accuracy of Classification (Distance 100 cm, 150 cm, and 200 cm- K = 1000)

| Used Method | Correct Predictions | Wrong Predictions | Percentage of Accuracy |
|---|---|---|---|
| SIFT | 93 | 3 | 96.87 |
| FAST-SIFT | 92 | 4 | 95.83 |
| HARRIS-SIFT | 92 | 4 | 95.83 |

Table VI, the next table, shows accuracy of classification for each method in general. All distances have been taken in consideration and total accuracy of each method has been obtained.

TABLE VI
Accuracy of Classification (K = 1000)

| Used Method | Correct Predictions | Wrong Predictions | Percentage of Accuracy |
|---|---|---|---|
| SIFT | 319 | 17 | 94.9404 |
| FAST-SIFT | 309 | 27 | 91.9642 |
| HARRIS-SIFT | 314 | 22 | 93.4523 |

Confusion matrix is a visional tool to evaluate performance of a proposed model or system in classification and prediction tasks. In classification tasks, it shows predictive capability. In fact, it is a square matrix with *n* dimensions, where *n* is the

number of target classes and it equal to 4 in this specific case. For three used methods, confusion matrixes are created for each distance and shown in Tables VII-XVIII. Investigation of confusion matrix helps to realize the roles of elements in each row and column, where columns show the predicted class or label and the rows correspond to the true and actual class or label [34].

TABLE VII
CONFUSION MATRIX (DISTANCE 25 CM- K = 1000)

| SIFT Method | Hydrangea | Amelanchier Canadensis | Acer Pseudoplatanus | Cornus |
|---|---|---|---|---|
| Hydrangea | 18 | 0 | 0 | 2 |
| Amelanchier Canadensis | 1 | 17 | 0 | 2 |
| Acer Pseudoplatanus | 1 | 0 | 19 | 0 |
| Cornus | 0 | 0 | 0 | 20 |

TABLE VIII
CONFUSION MATRIX (DISTANCE 50 CM- K = 1000)

| SIFT Method | Hydrangea | Amelanchier Canadensis | Acer Pseudoplatanus | Cornus |
|---|---|---|---|---|
| Hydrangea | 19 | 1 | 0 | 0 |
| Amelanchier Canadensis | 0 | 19 | 0 | 1 |
| Acer Pseudoplatanus | 1 | 0 | 19 | 0 |
| Cornus | 0 | 0 | 0 | 20 |

TABLE IX
CONFUSION MATRIX (DISTANCE 75 CM- K = 1000)

| SIFT Method | Hydrangea | Amelanchier Canadensis | Acer Pseudoplatanus | Cornus |
|---|---|---|---|---|
| Hydrangea | 18 | 0 | 0 | 2 |
| Amelanchier Canadensis | 0 | 19 | 0 | 1 |
| Acer Pseudoplatanus | 0 | 0 | 20 | 0 |
| Cornus | 2 | 0 | 0 | 18 |

TABLE X
CONFUSION MATRIX (DISTANCE 100 CM, 150 CM, 200 CM - K = 1000)

| SIFT Method | Hydrangea | Amelanchier Canadensis | Acer Pseudoplatanus | Cornus |
|---|---|---|---|---|
| Hydrangea | 22 | 1 | 0 | 1 |
| Amelanchier Canadensis | 0 | 24 | 0 | 0 |
| Acer Pseudoplatanus | 0 | 0 | 23 | 1 |
| Cornus | 0 | 0 | 0 | 24 |

TABLE XI
CONFUSION MATRIX (DISTANCE 25 CM - K = 1000)

| FAST-SIFT Method | Hydrangea | Amelanchier Canadensis | Acer Pseudoplatanus | Cornus |
|---|---|---|---|---|
| Hydrangea | 16 | 0 | 0 | 4 |
| Amelanchier Canadensis | 1 | 16 | 0 | 3 |
| Acer Pseudoplatanus | 0 | 0 | 20 | 0 |
| Cornus | 0 | 3 | 0 | 17 |

TABLE XII
CONFUSION MATRIX (DISTANCE 50 CM - K = 1000)

| FAST-SIFT Method | Hydrangea | Amelanchier Canadensis | Acer Pseudoplatanus | Cornus |
|---|---|---|---|---|
| Hydrangea | 18 | 1 | 0 | 1 |
| Amelanchier Canadensis | 0 | 17 | 0 | 3 |
| Acer Pseudoplatanus | 0 | 1 | 19 | 0 |
| Cornus | 0 | 0 | 0 | 20 |

TABLE XIII
CONFUSION MATRIX (DISTANCE 75 CM - K = 1000)

| FAST-SIFT Method | Hydrangea | Amelanchier Canadensis | Acer Pseudoplatanus | Cornus |
|---|---|---|---|---|
| Hydrangea | 19 | 0 | 0 | 1 |
| Amelanchier Canadensis | 0 | 19 | 0 | 1 |
| Acer Pseudoplatanus | 1 | 0 | 19 | 0 |
| Cornus | 2 | 1 | 0 | 17 |

TABLE XIV
CONFUSION MATRIX (DISTANCE 100 CM, 150 CM, 200 CM - K = 1000)

| FAST-SIFT Method | Hydrangea | Amelanchier Canadensis | Acer Pseudoplatanus | Cornus |
|---|---|---|---|---|
| Hydrangea | 21 | 2 | 0 | 1 |
| Amelanchier Canadensis | 0 | 24 | 0 | 0 |
| Acer Pseudoplatanus | 0 | 0 | 24 | 0 |
| Cornus | 1 | 0 | 0 | 23 |

TABLE XV
CONFUSION MATRIX (DISTANCE 25 CM - K = 1000)

| HARRIS-SIFT Method | Hydrangea | Amelanchier Canadensis | Acer Pseudoplatanus | Cornus |
|---|---|---|---|---|
| Hydrangea | 16 | 0 | 0 | 4 |
| Amelanchier Canadensis | 1 | 15 | 0 | 4 |
| Acer Pseudoplatanus | 0 | 0 | 20 | 0 |
| Cornus | 0 | 0 | 0 | 20 |

TABLE XVI
CONFUSION MATRIX (DISTANCE 50 CM - K = 1000)

| HARRIS-SIFT Method | Hydrangea | Amelanchier Canadensis | Acer Pseudoplatanus | Cornus |
|---|---|---|---|---|
| Hydrangea | 20 | 0 | 0 | 0 |
| Amelanchier Canadensis | 1 | 18 | 0 | 1 |
| Acer Pseudoplatanus | 1 | 0 | 19 | 0 |
| Cornus | 0 | 0 | 0 | 20 |

TABLE XVII
CONFUSION MATRIX (DISTANCE 75 CM - K = 1000)

| HARRIS-SIFT Method | Hydrangea | Amelanchier Canadensis | Acer Pseudoplatanus | Cornus |
|---|---|---|---|---|
| Hydrangea | 18 | 0 | 0 | 2 |
| Amelanchier Canadensis | 0 | 19 | 0 | 1 |
| Acer Pseudoplatanus | 1 | 0 | 19 | 0 |
| Cornus | 2 | 0 | 0 | 18 |

TABLE XVIII
CONFUSION MATRIX (DISTANCE 100 CM, 150 CM, 200 CM - K = 1000)

| HARRIS-SIFT Method | Hydrangea | Amelanchier Canadensis | Acer Pseudoplatanus | Cornus |
|---|---|---|---|---|
| Hydrangea | 20 | 3 | 0 | 1 |
| Amelanchier Canadensis | 0 | 24 | 0 | 0 |
| Acer Pseudoplatanus | 0 | 0 | 24 | 0 |
| Cornus | 0 | 0 | 0 | 24 |

Apart from providing information of the system, two common matrices are derived by using confusion matrix: (1) precision and (2) recall. In order to compute precision, it is needed to divide the number of correctly classified positive examples by the number of examples labeled by the system as positive. In addition to precision, the next parameter to investigate is recall which is computed by dividing by the number of correctly classified positive examples and the number of positive examples in the used data.

$$precision_i = \frac{M_{ii}}{\sum_j M_{ji}}. \quad (5)$$

$$recall_i = \frac{M_{ii}}{\sum_j M_{ij}}. \quad (6)$$

For each distance, precision and recall results of all methods have been computed. When the distance is 25 cm and the SIFT method is applied, variation of precision and recall results is less than other methods, FAST-SIFT and HARRIS-SIFT. In Figs. 3 and 4, it is completely clear. FAST-SIFT method has the lowest precision result. It was predictable as it has the lowest accuracy among three methods in this distance.

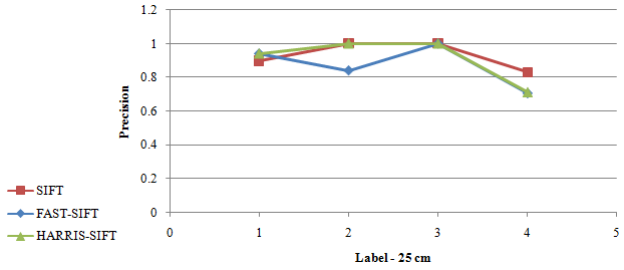


Fig. 3 Precision measurement for SIFT, FAST-SIFT, and HARRIS-SIFT methods (distance 25 cm)

Comparison of results can be done by investigation of area under the curves. If the area under one curve is high, it shows both high recall and high precision, and high precision and high recall relate to a low false positive rate and a low false negative rate, respectively [35].

In Figs. 5-7, precision and recall measurements of all methods have been shown separately. The relationship between recall and precision can be observed for each method in separate figures.

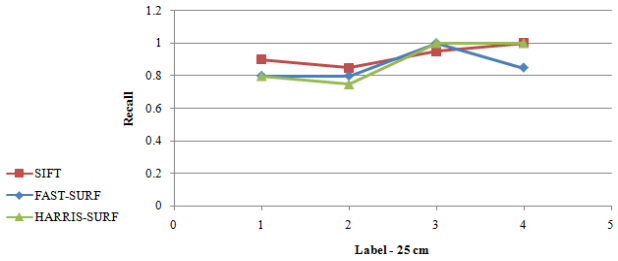


Fig. 4 Recall measurement for SIFT, FAST-SIFT, and HARRIS-SIFT methods (distance 25 cm)

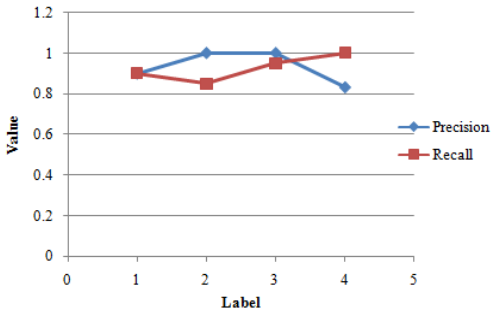


Fig. 5 Precision and recall values for SIFT measurement (distance 25 cm)

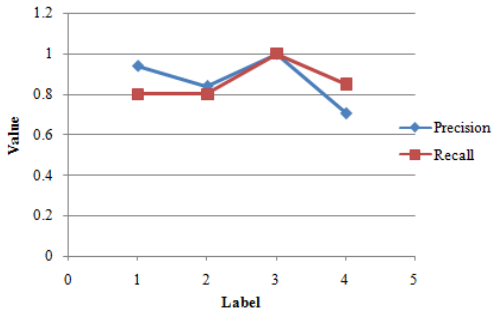


Fig. 6 Precision and recall values for FAST-SIFT measurement (distance 25 cm)

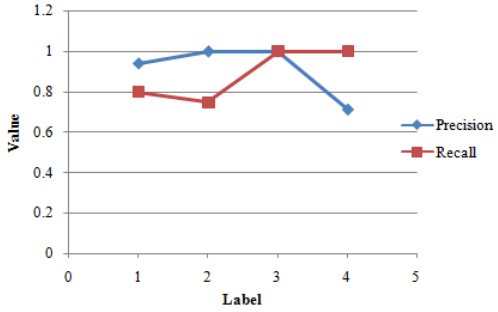


Fig. 7 Precision and recall values for HARRIS-SIFT measurement (distance 25 cm)

When the distance increases from 25 cm to 50 cm, recall and precision results are changed. The SIFT method has the lowest range of variations between the applied methods for the system. In this special case, obtained results of HARRIS-SIFT method are comparable to the results of SIFT method, but the

FAST-SIFT method does not have a good performance in this distance too. In Figs. 8 and 9, the results of this part have been shown.

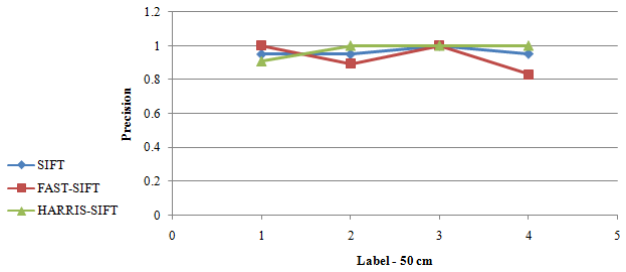


Fig. 8 Precision measurement for SIFT, FAST-SIFT, and HARRIS-SIFT methods (distance 50 cm)

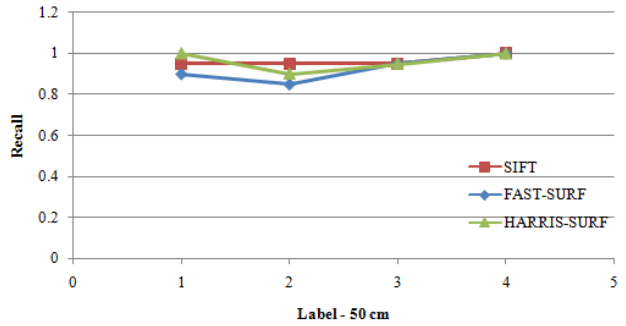


Fig. 9 Recall measurement for SIFT, FAST-SIFT, and HARRIS-SIFT methods (distance 50 cm)

For distance 50 cm, measurements of precision and recall for each method are shown in the Figs. 10-12.

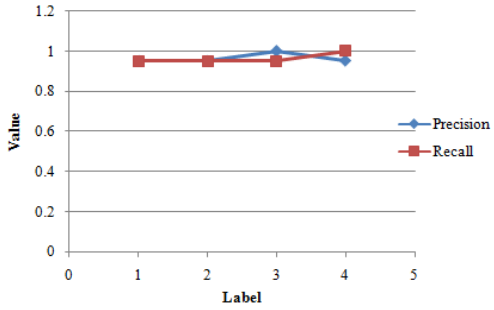


Fig. 10 Precision and recall values for SIFT measurement (distance 50 cm)

The next distance is 75 cm. Figs. 13 and 14 show the results of 75 cm distance. In this distance, the results of FAST-SIFT are comparable to the results of HARRIS-SIFT method. In addition to, accuracy results of them are the same. SIFT method has the best results in this distance too. When precision results are investigated, FAST-SIFT method has lower range of change than HARRIS-SIFT method, however the difference value is roughly 0.006.

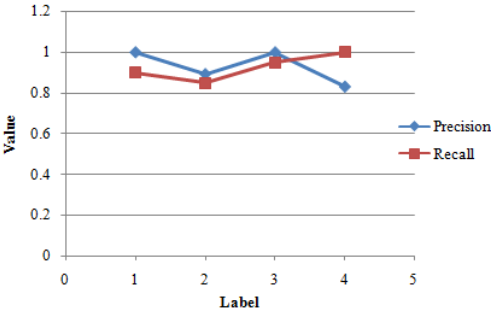


Fig. 11 Precision and recall values for FAST-SIFT measurement (distance 50 cm)

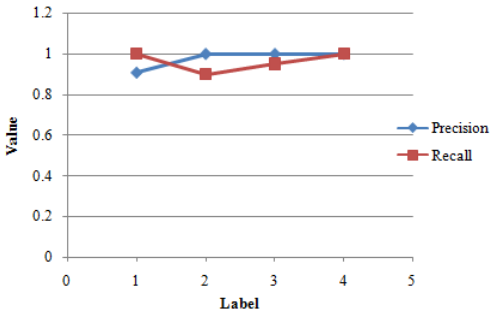


Fig. 12 Precision and recall values for HARRIS-SIFT measurement (distance 50 cm)

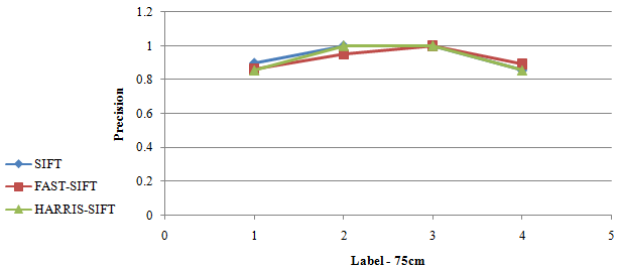


Fig. 13 Precision measurement for SIFT, FAST-SIFT, and HARRIS-SIFT methods (distance 75 cm)

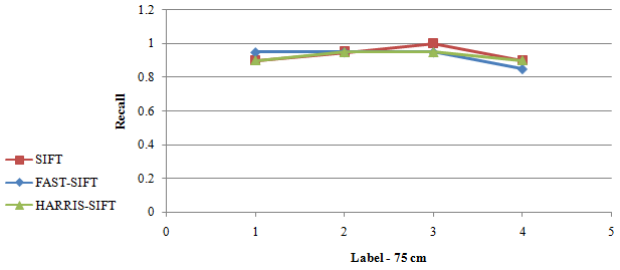


Fig. 14 Recall measurement for SIFT, FAST-SIFT, and HARRIS-SIFT methods (distance 75 cm)

Comparison of methods is possible by consideration of the measurements separately. Therefore, Figs. 15-17 have been used to find out a better understanding of the measurements. In addition to, each figure belongs to one method.

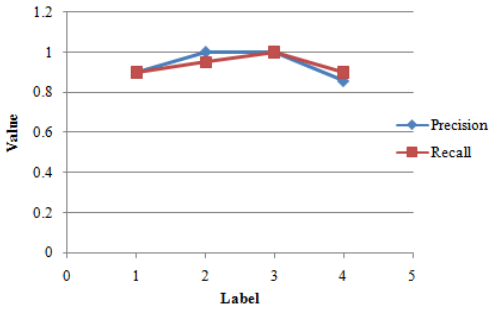


Fig. 15 Precision and recall values for SIFT measurement (distance 75 cm)

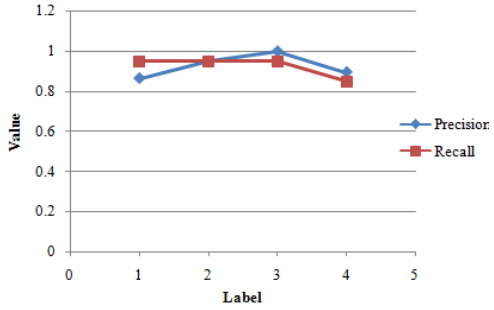


Fig. 16 Precision and recall values for FAST-SIFT measurement (distance 75 cm)

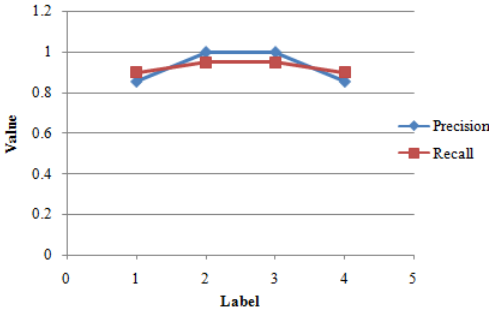


Fig. 17 Precision and recall values for HARRIS-SIFT measurement (distance 75 cm)

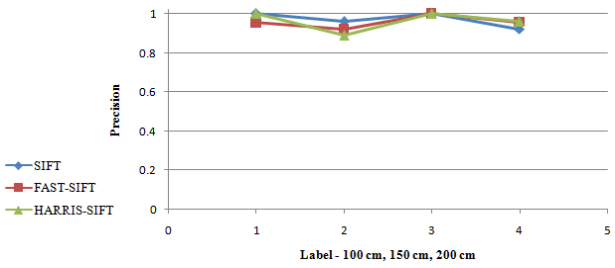


Fig. 18 Precision measurement for SIFT, FAST-SIFT, and HARRIS-SIFT methods (distance 100 cm, 150 cm, and 200 cm)

Three distances have constructed the last considered distance. This group consists of 100 cm, 150 cm, and 200 cm distances. The lowest variations belong to SIFT method in this case the same as other distances. Variations of FAST-SIFT method are lower than HARRIS-SIFT method. The results are shown in Figs. 20-22.

Separate representations of recall and precision measurements are done in Figs. 20-22 for all used methods.

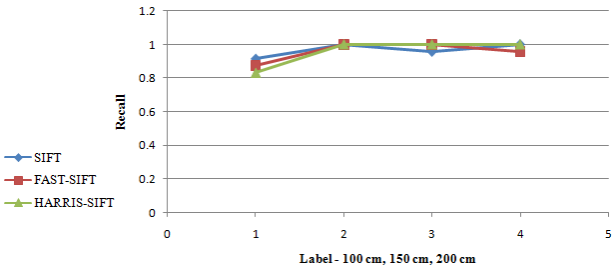


Fig. 19 Recall measurement for SIFT, FAST-SIFT, and HARRIS-SIFT methods (distance 100 cm, 150 cm, and 200 cm)

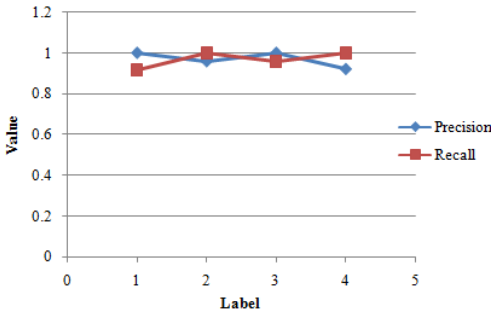


Fig. 20 Precision and recall values for SIFT measurement (distance 100 cm, 150 cm, and 200 cm)

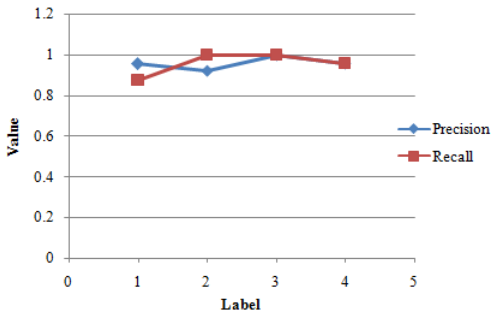


Fig. 21 Precision and recall values for FAST-SIFT measurement (distance 100 cm, 150 cm, and 200 cm)

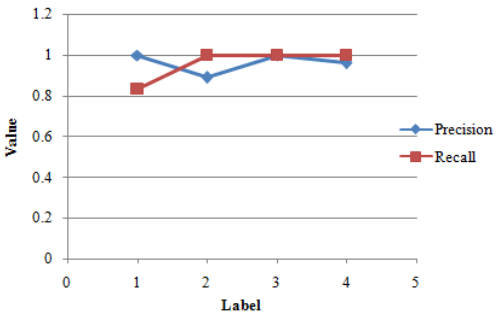


Fig. 22 Precision and recall values for HARRIS-SIFT measurement (distance 100 cm, 150 cm, and 200 cm)

As a conclusion of the precision and recall measurements, the sequence of best results and performances is SIFT, HARRIS-SIFT, and FAST-SIFT methods. The areas under curves are evidences of this sequence.

Each SVM has different parameters which can affect the performance of the system. NU parameter is a new regularization parameter after new reformulation of SVM and applies a slightly different penalty. Change of NU parameter is performed to investigate its effects over the system. Another investigated parameter is Gamma, which will be discussed in this section too.

Type of kernel is one important factor in designing of a system by using SVMs. Linear kernel, polynomial kernel, and RBF kernel are just some kinds of existed kernel. Kernel, one similarity function, is provided to a machine learning algorithm. In the other words, a kernel function can be specified to compute similarity of images and do the needed computations quicker. Different kernel types have been applied to implement the system and compare their results. In Table XIX, the accuracies of the different implemented system are shown in 25 cm, when the used method is SIFT.

TABLE XIX
ACCURACY OF SYSTEM FOR DIFFERENT KERNEL TYPES (DISTANCE 25 CM, K=1000, SIFT METHOD)

| Kernel Type | Correct Predictions | Wrong Predictions | Error | Percentage of Accuracy |
|---|---|---|---|---|
| Linear | 37 | 43 | 0.5375 | 46.2500 |
| Polynomial | 20 | 60 | 0.750 | 25.0000 |
| RBF | 75 | 5 | 0.0625 | 93.7500 |

TABLE XX
VARIATION OF NU PARAMETER (DISTANCE 25 CM, K=1000)

| SIFT Method | Correct Predictions | Wrong Predictions | Error | Percentage of Accuracy |
|---|---|---|---|---|
| NU = 0.1 | 75 | 5 | 0.0625 | 93.75 |
| NU = 0.2 | 75 | 5 | 0.0625 | 93.75 |
| NU = 0.5 | 69 | 11 | 0.1375 | 86.25 |

TABLE XXI
VARIATION OF NU PARAMETER (DISTANCE 50 CM, K=1000)

| SIFT Method | Correct Predictions | Wrong Predictions | Error | Percentage of Accuracy |
|---|---|---|---|---|
| NU = 0.1 | 79 | 1 | 0.0125 | 98.75 |
| NU = 0.2 | 79 | 1 | 0.0125 | 98.75 |
| NU = 0.5 | 76 | 4 | 0.0500 | 95.00 |

TABLE XXII
VARIATION OF NU PARAMETER (DISTANCE 75 CM, K=1000)

| SIFT Method | Correct Predictions | Wrong Predictions | Error | Percentage of Accuracy |
|---|---|---|---|---|
| NU = 0.1 | 77 | 3 | 0.0375 | 96.25 |
| NU = 0.2 | 76 | 4 | 0.0500 | 95.00 |
| NU = 0.5 | 74 | 6 | 0.0750 | 92.50 |

NU parameter has two main characteristics. The first characteristic relates to its value. The value is always bounded between 0 and 1. There is a range for this parameter instead of a single value. The second characteristic relates to its direct interpretation. Meanwhile, it has much more meaningful interpretation because it represents upper and lower bounds on the fraction of training samples and the samples that are respectively errors and support vectors.

In Tables XX-XXXI, the effects of nu changes can be observed. To investigate the parameter's influence, another important parameter, gamma parameter is fixed to 1.0. The type of SVM is nu-support vector classification and kernel type is RBF.

TABLE XXIII
VARIATION OF NU PARAMETER (DISTANCE 100 CM, 150 CM, AND 200 CM, K=1000)

| SIFT Method | Correct Predictions | Wrong Predictions | Error | Percentage of Accuracy |
|---|---|---|---|---|
| NU = 0.1 | 94 | 2 | 0.0208330 | 97.9167 |
| NU = 0.2 | 94 | 2 | 0.0208330 | 97.9167 |
| NU = 0.5 | 92 | 4 | 0.0416667 | 95.8333 |

TABLE XXIV
VARIATION OF NU PARAMETER (DISTANCE 25 CM, K=1000)

| FAST-SIFT Method | Correct Predictions | Wrong Predictions | Error | Percentage of Accuracy |
|---|---|---|---|---|
| NU = 0.1 | 71 | 9 | 0.1125 | 88.75 |
| NU = 0.2 | 70 | 10 | 0.1250 | 87.50 |
| NU = 0.5 | 64 | 16 | 0.2000 | 80.00 |

TABLE XXV
VARIATION OF NU PARAMETER (DISTANCE 50 CM, K=1000)

| FAST-SIFT Method | Correct Predictions | Wrong Predictions | Error | Percentage of Accuracy |
|---|---|---|---|---|
| NU = 0.1 | 78 | 2 | 0.0250 | 97.50 |
| NU = 0.2 | 76 | 4 | 0.0500 | 95.00 |
| NU = 0.5 | 73 | 7 | 0.0875 | 91.25 |

TABLE XXVI
VARIATION OF NU PARAMETER (DISTANCE 75 CM, K=1000)

| FAST-SIFT Method | Correct Predictions | Wrong Predictions | Error | Percentage of Accuracy |
|---|---|---|---|---|
| NU = 0.1 | 75 | 5 | 0.0625 | 93.75 |
| NU = 0.2 | 74 | 6 | 0.0750 | 92.50 |
| NU = 0.5 | 72 | 8 | 0.1000 | 90.00 |

TABLE XXVII
VARIATION OF NU PARAMETER (DISTANCE 100 CM, 150 CM, AND 200 CM, K=1000)

| FAST-SIFT Method | Correct Predictions | Wrong Predictions | Error | Percentage of Accuracy |
|---|---|---|---|---|
| NU = 0.1 | 93 | 3 | 0.0312500 | 96.8750 |
| NU = 0.2 | 92 | 4 | 0.0416667 | 95.8333 |
| NU = 0.5 | 89 | 7 | 0.0729167 | 92.7083 |

TABLE XXVIII
VARIATION OF NU PARAMETER (DISTANCE 25 CM, K=1000)

| HARRIS-SIFT Method | Correct Predictions | Wrong Predictions | Error | Percentage of Accuracy |
|---|---|---|---|---|
| NU = 0.1 | 71 | 9 | 0.1125 | 88.75 |
| NU = 0.2 | 70 | 10 | 0.1250 | 87.50 |
| NU = 0.5 | 66 | 14 | 0.1750 | 82.50 |

TABLE XXIX
VARIATION OF NU PARAMETER (DISTANCE 50 CM, K=1000)

| HARRIS-SIFT Method | Correct Predictions | Wrong Predictions | Error | Percentage of Accuracy |
|---|---|---|---|---|
| NU = 0.1 | 79 | 1 | 0.0125 | 98.75 |
| NU = 0.2 | 78 | 2 | 0.0250 | 97.50 |
| NU = 0.5 | 72 | 8 | 0.1000 | 90.00 |

When the distance is 25 cm, changes of NU parameter have lower effects on SIFT method, shown in Table XX, in comparison to FAST-SIFT, shown in Table XXIV, and HARRIS-SIFT, shown in Table XXVIII, methods. In this case, performance of the system by means of SIFT method is also better. The range of NU variation in FAST-SIFT is higher than other two methods. For distance 50 cm, the system is more robust to changes of NU parameter when the SIFT method is used. In comparison to using FAST-SIFT method, performance of the system is higher when the system is implemented by HARRIS-SIFT. Additionally, increase of NU parameter has less effects on SIFT and FAST-SIFT methods. The obtained values can be observed in Tables XXI, XXV, and XXIX.

TABLE XXX
VARIATION OF NU PARAMETER (DISTANCE 75 CM, K=1000)

| HARRIS-SIFT Method | Correct Predictions | Wrong Predictions | Error | Percentage of Accuracy |
|---|---|---|---|---|
| NU = 0.1 | 76 | 4 | 0.050 | 95.00 |
| NU = 0.2 | 74 | 6 | 0.075 | 92.50 |
| NU = 0.5 | 70 | 10 | 0.125 | 87.50 |

TABLE XXXI
VARIATION OF NU PARAMETER (DISTANCE 100 CM, 150 CM, AND 200 CM, K=1000)

| HARRIS-SIFT Method | Correct Predictions | Wrong Predictions | Error | Percentage of Accuracy |
|---|---|---|---|---|
| NU = 0.1 | 92 | 4 | 0.0416667 | 95.83330 |
| NU = 0.2 | 91 | 5 | 0.0520833 | 94.79167 |
| NU = 0.5 | 88 | 8 | 0.0833333 | 91.66667 |

In 75 cm distance, the highest accuracy belongs to the system used SIFT method. In this distance, change of NU parameter has more effects on implemented system by HARRIS-SIFT. Tables XXII, XXVI, and XXX show the obtained values. For the last group of distances, 100cm, 150 cm, and 200 cm, variation of error in implemented system by using FAST-SIFT is higher than other systems, which used SIFT and HARRIS-SIFT methods in this case, and Tables XXIII, XXVII, and XXXI show them.

There is a direct relationship between increase of NU parameter and error. As the value of NU parameter increases, error of the system increases too. In all distances, the best accuracy for the proposed system is obtained by using the SIFT method. Robustness of this method is almost higher than other used methods in different distances. In performed experiments, the highest accuracy belongs to SIFT method, where NU is 0.1 and the distance is 50 cm. The worst result has been observed in distance 25 cm for FAST-SIFT method, where the NU parameter is 0.5.

Regarding RBF kernel, there is an important parameter to control the width of the kernel and the ball-shaped curve which is called gamma parameter. Values in the range from 1 to 10 usually work well, but sometimes other values (smaller values like 0.1 or larger values like 75) give good results too. The larger value of gamma, the narrow will be the bell. Small values yield bells. In the next tables, variation of gamma parameter is performed for different methods and distances. In this experiment, NU parameter is fixed to 0.1 in all cases. The type of SVM and kernel type are the same as previous section for changing NU parameter. While the distance is 25 cm and gamma parameter increases, the SIFT method has the highest accuracy. The second best accuracy belongs to FAST-SIFT method when gamma parameter changes. It is interesting that changes of gamma parameter affect HARRIS-SIFT method more than two other methods in this distance. Related tables are Tables XXXII, XXXVI, and XL. When the distance increases to 50 cm and this parameter changes, the FAST-SIFT method has the worst result. Table XXXVII proves this fact.

In distance 75 cm, increase of gamma parameter has more effects on results of FAST-SIFT method and HARRIS-SIFT method which are shown in Tables XXXVIII and XLII. The changes of gamma parameter have less negative effects on the results of SIFT method in this distance when it is compared to changes of NU parameter in the same distance, 75 cm. Increase of gamma parameter leads to increase of error in three methods.

TABLE XXXII
VARIATION OF GAMMA PARAMETER (DISTANCE 25 CM, K=1000)

| SIFT Method | Correct Predictions | Wrong Predictions | Error | Percentage of Accuracy |
|---|---|---|---|---|
| Gamma=25 | 75 | 5 | 0.0625 | 93.750 |
| Gamma=30 | 72 | 8 | 0.1000 | 90.000 |
| Gamma=35 | 71 | 9 | 0.1125 | 88.750 |

TABLE XXXIII
VARIATION OF GAMMA PARAMETER (DISTANCE 50 CM, K=1000)

| SIFT Method | Correct Predictions | Wrong Predictions | Error | Percentage of Accuracy |
|---|---|---|---|---|
| Gamma=25 | 79 | 1 | 0.0125 | 98.7500 |
| Gamma=30 | 79 | 1 | 0.0125 | 98.7500 |
| Gamma=35 | 79 | 1 | 0.0125 | 98.7500 |

TABLE XXXIV
VARIATION OF GAMMA PARAMETER (DISTANCE 75 CM, K=1000)

| SIFT Method | Correct Predictions | Wrong Predictions | Error | Percentage of Accuracy |
|---|---|---|---|---|
| Gamma=25 | 78 | 2 | 0.0250 | 97.5000 |
| Gamma= 30 | 78 | 2 | 0.0250 | 97.5000 |
| Gamma= 35 | 77 | 3 | 0.0375 | 96.2500 |

TABLE XXXV
VARIATION OF GAMMA PARAMETER (DISTANCE 100 CM, 150 CM, AND 200 CM, K=1000)

| SIFT Method | Correct Predictions | Wrong Predictions | Error | Percentage of Accuracy |
|---|---|---|---|---|
| Gamma=25 | 94 | 2 | 0.0208 | 97.9200 |
| Gamma=30 | 94 | 2 | 0.0208 | 97.9200 |
| Gamma=35 | 94 | 2 | 0.0208 | 97.9200 |

TABLE XXXVI
VARIATION OF GAMMA PARAMETER (DISTANCE 25 CM, K=1000)

| FAST-SIFT Method | Correct Predictions | Wrong Predictions | Error | Percentage of Accuracy |
|---|---|---|---|---|
| Gamma = 25 | 73 | 7 | 0.0875 | 91.2500 |
| Gamma = 45 | 72 | 8 | 0.1000 | 90.0000 |
| Gamma = 75 | 70 | 10 | 0.1250 | 87.5000 |

In the last group of distances, FAST-SIFT method has better accuracy than HARRIS-SIFT. Additionally, best results have been obtained by using SIFT method. The results have been shown in Tables XXXV, XXXIX, and XLIII.

TABLE XXXVII
VARIATION OF GAMMA PARAMETER (DISTANCE 50 CM, K=1000)

| FAST-SIFT Method | Correct Predictions | Wrong Predictions | Error | Percentage of Accuracy |
|---|---|---|---|---|
| Gamma = 25 | 78 | 2 | 0.0250 | 97.5000 |
| Gamma = 45 | 78 | 2 | 0.0250 | 97.5000 |
| Gamma = 75 | 78 | 2 | 0.0250 | 97.5000 |

TABLE XXXVIII
VARIATION OF GAMMA PARAMETER (DISTANCE 75 CM, K=1000)

| FAST-SIFT Method | Correct Predictions | Wrong Predictions | Error | Percentage of Accuracy |
|---|---|---|---|---|
| Gamma = 25 | 75 | 5 | 0.0625 | 93.7500 |
| Gamma = 45 | 75 | 5 | 0.0625 | 93.7500 |
| Gamma = 75 | 74 | 6 | 0.0750 | 92.5000 |

TABLE XXXIX
VARIATION OF GAMMA PARAMETER (DISTANCE 100 CM, 150 CM, AND 200 CM, K=1000)

| FAST-SIFT Method | Correct Predictions | Wrong Predictions | Error | Percentage of Accuracy |
|---|---|---|---|---|
| Gamma = 25 | 93 | 3 | 0.0312 | 96.8800 |
| Gamma = 45 | 93 | 3 | 0.0312 | 96.8800 |
| Gamma = 75 | 93 | 3 | 0.0312 | 96.8800 |

TABLE XL
VARIATION OF GAMMA PARAMETER (DISTANCE 25 CM, K=1000)

| HARRIS-SIFT Method | Correct Predictions | Wrong Predictions | Error | Percentage of Accuracy |
|---|---|---|---|---|
| Gamma = 15 | 71 | 9 | 0.1125 | 88.7500 |
| Gamma = 35 | 71 | 9 | 0.1125 | 88.7500 |
| Gamma = 75 | 70 | 10 | 0.1250 | 87.5000 |

TABLE XLI
VARIATION OF GAMMA PARAMETER (DISTANCE 50 CM, K=1000)

| HARRIS-SIFT Method | Correct Predictions | Wrong Predictions | Error | Percentage of Accuracy |
|---|---|---|---|---|
| Gamma = 15 | 79 | 1 | 0.0125 | 98.7500 |
| Gamma = 35 | 79 | 1 | 0.0125 | 98.7500 |
| Gamma = 75 | 79 | 1 | 0.0125 | 98.7500 |

TABLE XLII
VARIATION OF GAMMA PARAMETER (DISTANCE 75 CM, K=1000)

| HARRIS-SIFT Method | Correct Predictions | Wrong Predictions | Error | Percentage of Accuracy |
|---|---|---|---|---|
| Gamma = 15 | 76 | 4 | 0.050 | 95.0000 |
| Gamma = 35 | 76 | 4 | 0.050 | 95.0000 |
| Gamma = 75 | 74 | 6 | 0.075 | 92.5000 |

TABLE XLIII
VARIATION OF GAMMA PARAMETER (DISTANCE 100 CM, 150 CM, AND 200 CM, K=1000)

| HARRIS-SIFT Method | Correct Predictions | Wrong Predictions | Error | Percentage of Accuracy |
|---|---|---|---|---|
| Gamma = 15 | 92 | 4 | 0.0416 | 95.8400 |
| Gamma = 35 | 91 | 5 | 0.0520 | 94.8000 |
| Gamma = 75 | 91 | 5 | 0.0520 | 94.8000 |

For test step of the implemented system, needed time is calculated. Tables XLIV-XLVI show the needed time for different used methods in different distances. The number of images is mentioned for each distance and measurement. FAST detector performs faster than other two methods. Furthermore, the needed time for description part is more than detection part for all methods. From a real time application point of view, FAST algorithm offers higher performance of corner detection in low distance, 25 cm. For distance 25 cm, the FAST-SIFT method is the fastest one for execution of system too, however the accuracy of this system is lower than two other systems. The system with SIFT method needs more time than other developed systems, but it has the best performance among three proposed systems according to obtained accuracy results.

TABLE XLIV
TOTAL NEEDED TIME FOR THE SYSTEM USED SIFT

| Distance | Operator SIFT | Rest of Test Step |
|---|---|---|
| 25 cm (80 Images) | 591.2721 | 0.008773 |
| 50 cm (80 Images) | 610.7227 | 0.008389 |
| 75 cm (80 Images) | 604.5098 | 0.008681 |
| 100, 150, 200 cm (96 Images) | 769.3853 | 0.009380 |

TABLE XLV
TOTAL NEEDED TIME FOR THE SYSTEM USED FAST-SIFT

| Distance | Operator FAST-SIFT | Rest of Test Step |
|---|---|---|
| 25 cm (80 Images) | 298.8551 | 0.007758 |
| 50 cm (80 Images) | 317.2979 | 0.007411 |
| 75 cm (80 Images) | 314.4157 | 0.007308 |
| 100, 150, 200 cm (96 Images) | 384.0809 | 0.009635 |

TABLE XLVI
TOTAL NEEDED TIME FOR THE SYSTEM USED HARRIS-SIFT

| Distance | Operator HARRIS-SIFT | Rest of Test Step |
|---|---|---|
| 25 cm (80 Images) | 337.2988 | 0.008195 |
| 50 cm (80 Images) | 343.7508 | 0.007914 |
| 75 cm (80 Images) | 360.5383 | 0.007015 |
| 100, 150, 200 cm (96 Images) | 438.4270 | 0.00938 |

(a) (b)

Fig. 23 (a) Representation of keypoints for HARRIS-SIFT method in distances 50cm and 75 cm. (b) Representation of keypoints for SIFT method in distances 50 cm and 75 cm

Computation of number of keypoints is another performed experiment. In Table XLVII, number of keypoints has been shown for two methods in different distances. Fig. 23 represents detected keypoints of HARRIS-SIFT and SIFT methods in 50 cm and 75 cm distances.

TABLE XLVII
NUMBER OF KEYPOINTS

| Number of keypoints according to method and distance | SIFT method | HARRIS-SIFT |
|---|---|---|
| 50 cm | 2385 | 314 |
| 75 cm | 1438 | 435 |
| 100 cm | 2251 | 1000 |

## V. CONCLUSION

In this paper, a new system is implemented to recognize natural plants. The used dataset contains natural images, which have been taken in different angles and views, illuminations, light intensities, weather conditions, and distances. Therefore, the implemented system is generalized to use in different situations in various environments such as windy and cloudy weather. The classification system is efficient, reliable, and accurate by using modern methods and combination of them. Some experimental tests have been performed, and the quantitative results have been compared in details. Obtained accuracy of the system is higher when SIFT method is applied in developed system. A complete discussion and demonstration of applied methods are provided. Combination of other modern methods can be used in proposed system to improve accuracy and efficiency. The next steps of the work will be use and investigation of other methods.

## ACKNOWLEDGMENT

Masoud Fathi Kazerouni has created and prepared the natural image dataset for plants. The dataset belongs to Institute of Real-time Learning Systems, University of Siegen, Germany.

## REFERENCES

[1] N. Sakai, S. Yonekawa, and A. Matsuzaki, “Two dimensional image analysis of the shape of rice and its applications to separating varieties”, Journal of Food Engineering, Vol. 27, 1996, pp. 397-407.
[2] A. J. Perez, F. Lopez, J. V. Benlloch, and S. Christensen, “Color and shape analysis techniques for weed detection in cereal fields”, Computers and Electronics in Agriculture, Vol. 25, 2000, pp. 197-212.
[3] A. Del Bimho, P. Pala, and S. Santini, “Image retrieval by elastic matching of shapes and image patterns,” in *Proc. 1996 Int. Conf. Multimedia Computing and Systems*, Hiroshima, Japan, June 1996, pp. 215–218.
[4] R. Mehrotra and J. E. Gary, “Feature-based retrieval of similar shapes,” in *Proc. 9th Int. Conf. Data Engineering*, Vienna, Austria, Apr. 1993, pp. 108–115.
[5] W. Niblack and J. Yin, “Pseudo-distance measure for 2d shapes based on turning angle,” in *Proc. IEEE Int. Conf. Image Processing*, Vol. 3, Washington, DC, USA, Oct. 1995, pp. 352–355.
[6] E. Saber and A. M. Tekalp, “Image query-by-example using region based shape matching,” *Proc. SPIE*, vol. 2666, 1996, pp. 200–211.
[7] S. Sclaroff, “Deformable prototypes for encoding shape categories in image databases,” *Pattern Recognition.*, Vol. 30, no. 4, Apr. 1997, pp. 627–641.
[8] S. Sclaroff and A. P. Pentland, “Modal matching for corresponding and recognition,” *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. 17, June 1995, pp. 545–561.
[9] S. Abbasi, F. Mokhtarian, and J. Kittler, “Reliable classification of Chrysanthemum leaves through curvature scale space”, Lecture Notes in Computer Science, Vol. 1252, 1997, pp. 284-295.
[10] C-L Lee, and S-Y Chen, “Classification of leaf images”, 16th IPPR Conference on Computer Vision, Graphics and Image Processing (CVGIP), 2003, pp. 355-362.
[11] Lowe, D. G., “Object Recognition from Local Scale-Invariant Features”, in *Proc. Of the International Conference on Computer Vision*, 1999, pp. 1150-1157.
[12] F. Estrada, A. Jepson, and D. Fleet, “Local Features Tutorial”, *http://www.cs.toronto.edu/~jepson/csc2503/tutSIFT04.pdf*, Accessed: Dec. 01, 2016.
[13] Herbert Bay, Tinne Tuytelaars and Luc Van Gool, “SURF: Speeded Up Robust Features”. Computer Vision – ECCV 2006, Lecture Notes in Computer Science, 2006. 3951: pp. 404-417.
[14] Masoud Fathi Kazerouni, Jens Schlemper, and Klaus-Dieter Kuhner, “Comparison of Modern Description Methods for the Recognition of 32 Plant Species”, Signal & Image Processing: An International Journal (SIPIJ), Vol. 6, No. 2, April 2015, pp.1-13.
[15] Hossam M. Zawbaa, Mona Abbass, Sameh H. Basha, Maryam Hazman, Abul Ella Hassenian, “An Automatic Flower Classification Approach Using Machine Learning Algorithms”, *Advances in Computing, Communications and Informatics (ICACCI, 2014 International Conference on*. IEEE, 2014, 2014, pp. 895-901.
[16] E. Rosten, T. Drummond. “Machine Learning for High-Speed Corner Detection,” in European Conference on Computer Vision, Vol. 1, May 2006, pp. 430–443.
[17] Masoud Fathi Kazerouni, Jens Schlemper and Klaus-Dieter Kuhnert; “Efficient Modern Description Methods by Using SURF Algorithm for Recognition of Plant Species”. Advances in Image and Video Processing, Vol. 3 No 2, April (2015); pp: 10-24.
[18] C. Harris, M. Stephens, “A Combined Corner and Edge Detector”, in Proceedings of the Fourth Alvey Vision Conference, 1988. pp. 147-151
[19] L. Fei-Fei, R. Fergus, and A. Torralba. “Recognizing and Learning Object Categories, CVPR 2007 short course”, 2007.
[20] S. Wu, F. Bao, E. Xu, Y.-X. Wang, Y.-F. Chang, and Q.-L. Xiang, “A leaf recognition algorithm for plant classification using probabilistic neural network,” In *2007 IEEE international symposium on signal processing and information technology*, IEEE, 2007, pp. 11 -16.
[21] G. Csurka, C. R. Dance, L. Fan, J. Willamowski, and C. Bray, “Visual categorization with bags of keypoints,” In *Workshop on statistical learning in computer vision, ECCV*, Vol. 1, no. 1-22, 2004, pp. 1-2.
[22] D. Larlus and F. Jurie. “Latent mixture vocabularies for object categorization”, In Proceedings of the British Machine Vision Conference, Vol. 3, 2006, pp. 959-968.
[23] J. Zhang, M. Marszalek, S. Lazebnik, and C. Schmid, “Local features and kernels for classification of texture and object categories: A comprehensive study”, 2006, pp. 13–13.
[24] D. Gokalp and S. Aksoy, “Scene classification using bag-of-regions representations”. In proceedings of CVPR, 2007, pp. 1–8.
[25] S. Smith, J. Brady, “SUSAN—A new approach to low level image processing,” *International Journal of Computer Vision*, Vol. 23, No.1, 1997, pp. 45-48.
[26] Suryadeep Roy, S. S Tripathy, “A Novel Approach for the Detection of Malaria Parasites and Measure its Severity Using Image Processing and Fuzzy Logic,” International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, Vol. 5, Issue 4, April 2016, pp. 3080-3088.
[27] D. G. Lowe, “Distinctive Image Features from Scale-Invariant Keypoints,” International Journal of Computer Vision, Vol. 60, No. 2, Nov. 2004, pp. 91-110.
[28] T. Leung and J. Malik, “Representing and recognizing the visual appearance of materials using three-dimensional textons”. International Journal of Computer Vision 43 (1), 2001, pp. 29–44.
[29] V. Vapnik, *The natural of statistical theory*. New York: Springer-Verlag, 1995.
[30] S. Abe, “Support vector machines for pattern classification,” Vol. 53., London: Springer-Verlag, 2005.
[31] C. Burges, “A tutorial on support vector machines for pattern recognition,” Data mining knowledge discovery 2 Vol. 2, 1998, pp 121-167.
[32] V. Kumar, M. Steinbach, and P. N. Tan, "Introduction to data mining," Pearson, Addison Wesley, London, 2006.

[33] J. Hyuk Hong, and C. Sung Bae, "A probabilistic multi-class strategy of one-vs.-rest support vector machines for cancer classification," Neuro computing Vol. 71, 2008, pp. 3275-3281.
[34] Powers, David M W (2011). "Evaluation: From Precision, Recall and F-Measure to ROC, informedness, Markedness & Correlation," (PDF). *Journal of Machine Learning Technologies* 2 (1), pp. 37–63.
[35] "Precision-recall — scikit-learn 0.18.1 documentation," 2010. (Online). Available: http://scikit-learn.org/stable/auto_examples/model_selection/plot_precision_recall.html. Accessed: Dec. 13, 2016.