Chapter 1

Prologue

1.1 Preamble

In modern times, people have sought ways to cultivate, buy, wear, or otherwise be around flowers and blooming partly because of their agreeable appearance and smell. Around the world, people use flowers for a wide range of events and functions that cumulatively encompass their lifetime. People therefore grow flowers around their homes, dedicate some parts of their living space to flower gardens, pick wildflowers, or buy flowers from florists who depend on an entire network of commercial growers and shippers to support their trade. Though, flowers provide less food than any other major parts of plants (seeds, fruits, roots, stem and leaves), they provide several important foods and spices. Hence, flowers are used in a day to day life for various events which necessitates an automatic guide for classification.

Nevertheless, floriculture has become one of the important commercial trades in agriculture owing to steady increase in demand of flowers. Hence commercial floriculture has emerged as a hi-tech activity, which takes place under controlled climatic conditions inside a greenhouse. Floriculture in India is being viewed as a high growth Industry. Commercial floriculture is becoming important from the export angle. The liberalization of industrial and trade policies paved a way for development of export-oriented production of flowers. It has been found that commercial floriculture has higher potential per unit area than most of the field crops and is therefore a lucrative business. Floriculture industry comprises of flower trade, nursery and potted plants, seed and bulb production, micro propagation and extraction of essential oil from flowers. All these essentially involve flower classification.
The taxonomy originally contained approximately 8000 plants, but has since been extended to encompass more than 2,50,000 flower species around the world. For flowers, classification may require dissection. However, even when an image is sufficient, classifying a flower may still need a guidebook. This is particularly frustrating now because with advances in digital and mobile technology it is easy to take pictures of flowers, but it is still difficult to find out what they are. Once we know the name of a flower, we can find more information about the flower on the web, but the link between obtaining an image of a flower and acquiring its name is missing. Therefore, it does require an automatic guide that classifies an image of a flower based on its visual content. Further, flower recognition is used for searching patent flower images to know if a flower image applied for a patent is already present in the patent image database or not (Das et. al., 1999). Since these activities are being done manually and they are mainly labor dependent, automation is necessary.

In this thesis, we develop and evaluate a flower classification and retrieval scheme. Novel algorithmic models are designed to segment, track, detect and classify different types of flowers from flower videos. Suitable feature extraction schemes and classifiers are developed to classify flowers into various classes. A retrieval model is also designed to retrieve similar flowers for a given query by sketch of a flower. To index a large dataset of flowers effectively, a multi-dimensional data structure is explored. Due to the non-availability of a large benchmark dataset of flowers, we have created our own dataset of flowers which includes both images and videos. Experimental results on these relatively large datasets are presented to bring out the superiorities of the proposed method over the other existing contemporary models.

In the following sections, we present an overview of classification and retrieval of flowers, major tasks involved in classification and retrieval of flowers. Applications of proposed system is listed. A survey on existing related works on flowers classification is covered. Subsequently, objectives for the current work are presented. Thereafter, the measures which are generally used for evaluation of segmentation and classification models are discussed. Finally motivation for the current work and organization of the entire thesis are presented.
1.2 Classification and Retrieval of Flowers: An Overview

The general architecture of a typical flower classification and retrieval system is shown in figure 1.1.

1.2.1 Acquisition

Flowers used for imaging are randomly selected from a real environment. Flower images are captured using a digital camera at different ecological conditions, such as sunny, cloudy and rainy with different complex background and deformation etc.

1.2.2 Segmentation

Flowers in images are often surrounded by greenery in the background. Hence, the background regions in images of two different flowers can be very similar. In order to avoid matching the green background region, rather than the desired foreground region, the image is segmented. We can find several segmentation techniques such as Pixel labeling method (Das et. al., 1999), Contour-based methods (Kass et. al., 1987, Chan and Vese, 2001, Mortenson et. al., 1995, Saitoh et. al., 2004), Graph-based methods (Boykov and Jolly, 2001; Rother et. al., 2004; Kumar et. al., 2005; Nilsback and Zisserman, 2007) and Region based methods (Ning et al., 2010, Calderero and Marques, 2008 & 2010).

1.2.3 Feature Extraction

This stage concerns with description of flowers in terms of features. The description of flowers involves extraction of features. Different features are chosen to describe different properties of a flower. Some flowers are with very distinctive shapes, some are with very distinctive colors, some are with very characteristic texture patterns, and some are characterized by a combination of these properties.

1.2.3.1 Color Features

Flowers exist in a wide variety of colors, but many have a distinctive color. The color of a flower can help to narrow down the possible species, but it doesn’t enable us to determine the exact species of the flower. To handle this problem the color feature is
Figure 1.1: Block diagram of a flower classification and retrieval system

1.2.3.2 Texture Features


1.2.3.3 Shape Features

The shapes of individual petals, their configuration, and the overall shape of the flower can all be used to distinguish flowers. The difficulty of describing a shape is increased due to natural deformations of a flower. Nilsback and Zisserman (2006, 2008) describe the shape features using SIFT descriptors. Tsong and Yun (2003) studied the first 36 Zernike moments and found the dependence relations between them. A different matching scheme called SURF (Speeded Up Robust Features) was presented by Bay et al., (2006). Given a set of points (usually edges) in an image, the shape-context descriptor (Belongie and Malik 2002) can be used for describing the

1.2.4 Classification

The problem of flower classification is basically a task of partitioning the feature space into regions each representing a category. It is to establish a correspondence between a flower and a category to which it belongs. Ideally, one would like to arrange this portioning so that none of the decisions is ever wrong. When this cannot be done, one would like to minimize the probability of errors, or if some errors are more costly than the others, the average cost of errors. In this case, the problem of classification becomes a statistical decision theory, a subject that has many applications to pattern classification (Schalkoff, 1992). Nilsback and Zisserman (2006 & 2008) used nearest neighbor classifier and support vector machine to classify the flowers. Varma and Ray (2007) used multiple kernel classifiers to classify flowers.

1.2.5 Retrieval

Given a query flower, the preprocessing steps are carried out and it is represented using the same representation scheme adopted in the training stage. Based on the representation, a matching scheme is used to compare the given flower image against the flower images stored in the knowledge base (Das et al., 1999; An-xiang et al., 2004). After comparing the query flower image with the flowers in the database, the flowers are ranked and retrieved based on their proximity values. The leaf image retrieval system is developed using nearest neighbor search algorithm (Nam et al.,
2005), centroid-contour distance (Wang. et. al., 2000) and shape features (Park et. al., 2008).

1.3 Application

Classification and retrieval of flowers find their applications in the following areas.

- **Floriculture**: Flower industry comprises of flower trade, nursery and potted plants, seed and bulb production, micro propagation and extraction of essential oil from flowers.

- **Medical**: Flowers are alternative medical treatments for a variety of conditions. It is useful in manufacturing medicines for the diseases like cancer and muscle pain. Further it is also useful in manufacturing several beauty creams.

- **Patent Image database**: It is useful in verifying if the flower image applied for patent is already present in the patent image database or not.

- **Industrial**: The flowers are useful in manufacturing of several tea flavors and also for making up the various sweets.

1.4 Challenges

In this section, we list out the challenges involved in each stage of classification and retrieval of flowers.

- **Light variations**: It can be due to different intensities or colors of the light falling from different angles. Ambient light also changes with the time of a day. These changes in illumination can cause shadowing effects and changes in intensity. In addition, flowers are often transparent to some degree.

- **Viewpoint variations**: It may result in significant changes in flower's appearance. It can lead to changes in the perceived size, shape, pose and rotation of the flower.
• **Occlusion:** It occurs when a part of a flower is hidden by either itself or by another object. If characteristic parts, for example texture patterns, are occluded then it might be difficult to classify a flower.

• **Clutter:** In flower images the backgrounds are often similar, hence the background clutter can mislead a classifier.

• **Object deformations:** The size of the flower may change or a part may go missing or flower petals may fall.

• **Intra-class vs inter-class variations:** Intra-class variations are variations between flowers within a category and inter-class variations are variations between flowers across different categories. What distinguishes one flower from another can sometimes be the color, sometimes the shape, and sometimes patterns on the petals. The difficulty lies in finding suitable features to represent color, shape and texture patterns and also for the classifier having the capacity to learn features to use for better discrimination.

### 1.5 Related Works

In this section, we provide a brief review of existing related works on classification and retrieval of flowers.

Das et al., (1999) proposed an indexing method to index patent images using the domain knowledge. This method has illustrated a solution to the problem of indexing images of flowers for searching a flower patent database by color. The flower was segmented using an iterative segmentation algorithm with the domain knowledge driven feedback. The domain knowledge eliminates most of the frequently occurring elements of the background in flower images by deleting pixels which do not represent colors of flowers. The color of the flower is defined by the color names present in the flower region and their relative proportions. The database can be queried by example and by color names. In their work, the image color is mapped to names using ISCC-NBS color system and X Window system. Each flower image is discretized in HSV color space and each point on the discretized HSV space is
mapped to a color name in ISCC-NBS and X Window system. These names are used to index the flowers which can be used during retrieval of images. Retrieval method discusses how a database of flower patent images may be queried using both an example flower image as well as using the names of colors. Flower images are submitted as a part of the process of applying for flower patents from the U.S. Patent and Trademark Office. A person, who would like to check whether a new flower submitted for patenting is unique, can provide an example image from the patent application to retrieve similar flowers that already exist in the database. On the other hand, a person looking for flowers to cultivate may only be able to specify the flower type and a color name while querying the database.

Saitoh et al., (2000) describes an automatic recognition system for wild flowers by considering both flowers and their leaves. The objective is to extract both flowers and leaves from each image using a clustering method, followed by recognition using a piecewise linear discriminant function on 10 features of a flower and 11 features of a leaf. The experiment is performed on 20 sets of 34 species of wild flowers.

Saitoh et al., (2004) describes an automatic method for recognizing a blooming flower based on a photograph taken with a digital camera in natural scene. They have also proposed a method for extracting flower regions. It is based on “Intelligent Scissors” (Mortensen and Barrett 1995), which finds the path between two points that minimizes a cost function dependant on image gradients. The method works under the assumption that the flower is in focus that too in the centre of the photograph and that the background is out of focus. Under this assumption the cost between any two points on the flower is smaller than the cost between a point in the background and a point in the foreground. By fixing up the midpoint of the image as a part of the flower, this can be used as a starting point for finding the flower region. This method requires no prior color information.

Yoshioka et al., (2005) in their work performed a quantitative evaluation of petal colors using principal component analysis (PCA). Before scanning, the flowers are separated into five petals with a cutter and obtained petal images in a digital image scanner. The images were saved in RGB color format with 256 grey levels. They set a
region of interest in each petal as a region that represented the petal color pattern and defined the maximum square on each petal as the region of interest (ROI). Then converted each ROI image to a 10 × 10 pixel mosaic image and defined a total of 300 variables per mosaic image. Finally, they summarized the information on the 300 variables by PCA, and redrew the mosaic images to corresponding principal component score to determine the effect of each principal component on color pattern.

Nilsson and Zisserman (2006) designed a flower classification system by extracting visual vocabularies representing color, shape and texture features of flower images. First the flower was separated from background using contrast dependent prior MRF cost function (Boykov and Jolly, 2001), then optimized using graph cuts. To extract color vocabulary the flower image is mapped onto HSV color space, the HSV values for each pixel in the training images are clustered using k-means clustering to get color vocabulary. The SIFT descriptors are used to represent the shape features and responses of MR8 filter bank at different orientations are used as texture features. Nilsson and Zisserman considered a dataset of 17 species each containing 80 images and achieved an accuracy of 81.3% for combination of all the three features.

Nilsson and Zisserman (2007) proposed a two step model to segment flowers in color images. Where first step is to separate foreground and background, and the second step is to extract the petal structure of the flower. Segmentation starts using general foreground and background color distributions. These distributions are learnt by labeling pixels in a few sample images of each class in the dataset as foreground (i.e. part of the flower), or background (i.e. part of the greenery), and then averaging the distributions across all classes. Given these general foreground and background distributions, a binary segmentation is obtained using the contrast dependent prior Markova random field cost function, optimized with graph cuts. A generic flower shape model is then fitted to this initial segmentation in order to detect petals. The model selects petals which have a loose geometric consistency using an affine invariant Hough like procedure. The image regions for the petals deemed to be consistent are used to obtain a new image specific color foreground model. This image specific foreground model replaces the general foreground model, and the
MRF segmentation is repeated. In the cases where the initial segmentation was not perfect, the use of the image specific foreground often harvests more of the flower. For experimentation they considered a dataset of 17 species each containing 80 images.

Varma and Ray (2007) proposed a method for learning the trade-off between invariance and discriminative power for a given classification task. They learn the optimal, domain-specific kernel as a combination of base kernels corresponding to base features which achieve different levels of trade-off such as rotation invariance, scale invariance, affine invariance, etc. The knowledge of the trade-off can directly lead to improved classification and also it can be used to perform analogous reasoning where images are retrieved on the basis of learnt invariant properties rather than just image content. Classification is carried out on the basis of vocabularies of visual words of shape, color and texture descriptors. The background in each image is removed using graph cut segmentation approach. Shape distances between two images are calculated as the $\chi^2$ statistic between the normalized frequency histograms of densely sampled, vector quantized SIFT descriptors of the two images. Similarly, color distances are computed over vocabularies of HSV descriptors and texture over MR8 filter responses. An entire set of weights is learnt, spanning the full range from shape to color. The texture features are probably ignored because they are very strongly correlated with shape features.

To study the effect of classification accuracy on a large data set, Nilsback and Zisserman (2008) have considered a dataset of 103 classes each containing 40 to 250 samples per class. The low level features such as color, histogram of gradient (HOG) orientation and SIFT features on both foreground region and its boundary are used. The HOG is applied here over the entire flower region and it captures more global spatial distribution of the flower, such as the overall arrangement of petals. At the end they combine all the features using a multiple kernel frame work with a SVM classifier. On this large dataset they have achieved an accuracy of 72.8% using multiple kernel classifiers.
From the above literature survey it is understood that, there are quite a good number of attempts towards development of flower classification systems, but all of them have been demonstrated only on small datasets. In addition, those approaches use simple and conventional descriptors for preserving shape, texture and color features. All aforementioned methods use the entire flower region for classification. However, it is known fact that whorl of flowers of different species are different in appearance. Whorl region of a flower contains more discriminating features when compared to entire flower region, and is helpful to classify the flowers. As flower images can be captured always, a sketch is a rapidly executed through freehand drawing. It might record something that the human being sees, it might record or develop an idea for later usage or it might be used as a quick way of graphically demonstrating flower images. It is an excellent way to quickly explore concepts. Due to advance mobile technology, it is easy to take videos of flowers, but it is still difficult to find out what they are, but in literature, we do not find any work on flower videos. For the purpose of classification a simple classifier such as nearest neighbor classifier, support vector machine have been used. Furthermore only a little work on indexing and retrieval of flowers has been reported. Further, there is no benchmark datasets of flowers which is reasonably large size.

1.6 Objectives

With the above backdrop, in this research work, we are motivated to design algorithmic models for the following objectives.

- Classification of flowers using fusion strategies of different features and classifiers.
- Developing a classification system based on only whorl part of a flower instead of considering entire flower.
- To support classification task on a large database it is also proposed to design an efficient multi level indexing scheme.
• Detection of a query sketch in flower videos and developing sketch based flower retrieval system.

• Creating an automatic guide that labels flowers in videos.

• It is to create a reasonably large database of flowers consisting of varieties of flowers, flower videos and sketches of flowers.

1.7 Performance Measures

In this section, we present different measures that are generally used to evaluate the performance of segmentation and classification models.

1.7.1 Evaluation of Segmentation Approaches

Evaluation results vary significantly between different evaluators, because each evaluator may have distinct standards for measuring the quality of the segmentation.

Rand Index

Consider two images, say ground truth and segmented respectively: S1 and S2 of N points \( X = \{x_1, x_2, x_3, \ldots, x_N\} \); that assigned labels \( \{l_i\} \) and \( \{l'_i\} \) respectively to point \( x_i \). The Rand Index can be computed as the ratio of the number of pairs of vertices having the compatible label relationship in S1 and S2. It can be defined as:

\[
R(S_1, S_2) = \frac{1}{2N} \sum_{i \neq j} \left[ I(l_i = l_j \land l'_i = l'_j) + I(l_i \neq l_j \land l'_i \neq l'_j) \right] 
\]  

(1.1)

Where, I is the identity function, and the denominator is the number of possible unique pairs among N data points. This gives a measure of similarity ranging from 0 to 1.
**Variation of Information**

It measures the sum of information loss and information gain between the two clustering, and thus it roughly measures the extent to which one clustering can explain the other. For segmentations, it can be interpreted as the average conditional entropy of one segmentation given the other.

\[
VI(S_{test}, S_K) = H(S_{test} | S_K) + H(S_K | S_{test})
\]

(1.2)

The first term in the above equation measures the amount of information about \( S_{test} \) that we lose, while the second term measures the amount of information about \( S_K \) that we have to gain, when going from segmentation \( S_{test} \) to ground truth \( S_K \). Where, \( H(\cdot | \cdot) \) is the conditional entropy.

**Global Consistency Error**

Measures the extent to which the regions in one segmentation are subsets of the regions in second segmentation (i.e. the refinement). Let \( R(S, p_i) \) be the set of pixels in segmentation \( S \) that contains pixel \( p_i \), then the local refinement error is defined as:

\[
E(S_1, S_2, p_i) = \frac{|R(S_1, p_i) \cap R(S_2, p_i)|}{|R(S_1, p_i)|}
\]

(1.3)

This error is not symmetric \((i.e., E(S_1, S_2, p_i) \neq E(S_2, S_1, p_i))\) w.r.t. the compared segmentations, and takes the value of zero when \( S_1 \) is a refinement of \( S_2 \) at pixel \( p_i \).

Global Consistency Error is then defined as:

\[
GCE(S_1, S_2) = \frac{1}{n} \min \left\{ \sum_i E(S_1, S_2, p_i), \sum_i E(S_2, S_1, p_i) \right\}
\]

(1.4)

where, \( n \) is the number of pixels.

**Boundary Displacement Error (BDE)**

The BDE is a boundary based metric to evaluate the segmentation quality. It defines the error of one boundary pixel as the distance between the pixel and its closest pixel in the other boundary image. Let \( B1, B2 \) represent respectively the boundaries of segmentation and Ground truth. The BDE can be computed using the minimum
absolute difference from arbitrary point \( x \) in \( B_1 \) to all the boundary points in \( B_2 \). A near-zero mean and small standard deviation of BDEs computed for all the points in \( B_1 \) indicate the quality of the image segmentation.

i.e., \( BDE = \min(|x - y_i|) \), where \( x \in B_1 \) and \( y_i \in B_2, i = 1, 2, ..., n \), \( n \) is the number of boundary points in \( B_2 \).

### 1.7.2 Evaluation of Classification System

To evaluate the correctness of flower detection algorithms in videos, one should look into confusion matrix. A confusion matrix is a matrix plot of predicted versus actual classes of the samples.

Usually in detection of flowers, the system may identify flowers which may not be correct flower. Therefore, to measure the performance of a system, two other different statistics known as precision and recall are employed.

Precision measures the ratio of total number of correctly detected flowers to the total number of detected flowers by the system. Recall evaluates a fractional value of the total number of correctly detected flowers to the total number of expected flowers in a video. The precision value indicates that how the system is accurate in detecting only correct flowers, while recall value signifies that to what extent the system is capable in recalling and detecting all expected flowers.

i.e., \[ \text{Precision} = \frac{\text{Total Number of Correctly Detected Flowers}}{\text{Total Number of Detected Flowers}} \] (1.5)

i.e., \[ \text{Recall} = \frac{\text{Total Number of Correctly Detected Flowers}}{\text{Total Number of Expected Flowers}} \] (1.6)

The overall system precision and recall are computed by taking the average of the precision and recall values calculated for all \( N \) number of flower videos of a flower video database.
\[ P = \text{Avg Precision} = \frac{\sum_{i=1}^{N} P_i}{N} \] (1.7)

\[ R = \text{Avg Recall} = \frac{\sum_{i=1}^{N} R_i}{N} \] (1.8)

where, \( P_i \) and \( R_i \) are the precision and recall values of the \( i^{th} \) flower video of a database respectively.

In order to balance precision and recall, higher-order statistics such as F-measure is suggested, and it is defined as,

\[ \text{F-measure} = \frac{2 \times P \times R}{P + R} \] (1.9)

**1.8 Outline of the Thesis**

In Chapter 2, detail on the datasets which we have created during the course of this research is given along with details on other publicly available datasets.

In chapter 3, a novel method to segment a flower from its background based on whorl part of the flower is presented. A novel technique to detect the whorl region of flowers based on Gabor filter responses is proposed. Evaluation of the proposed segmentation method using region based performance measures is also presented. Experiments are conducted on our own real dataset to evaluate the performance of the proposed segmentation model.

In chapter 4, we propose a classification model using different features and classifiers. The features such as Leung- Malik filter responses, Schmid filter responses, maximum responses, Gabor responses, color texture moments, Scale invariant feature transform are extracted from the segmented flower images. Fusion of these features in different combinations is studied. Probabilistic neural network, nearest neighbor, support vector machine and their possible fusions are used for the purpose of classification. The effect of classification of flower images considering only whorl parts of flowers is also presented. Extension of the proposed classification model on
flower videos dataset is also presented. To evaluate the performance of the proposed classification model, experiments are carried out on our own real dataset and also on a benchmark flower datasets. A comparative analysis of the proposed classification model with the state-of-the-art models is also presented.

In chapter 5, we propose novel approaches to classify flower images based on skeletons. The shape context features, distance features and spatial topology are envisaged to represent the flower skeletons. The distance features and spatial topology are aggregated and represented in the form of interval-valued type data. Classification of flowers is accomplished using Nearest neighbor and Symbolic classifiers. Experiments are conducted on our own real dataset to evaluate the performance of all the three approaches. Results of the experiments are tabulated to bring out the efficiency of our proposed model.

In chapter 6, we propose a model for representation and indexing of flower images for the purpose of retrieval flowers of interest based a query sketch. We present a Kd-tree indexing mechanism for flower retrieval using shape descriptors. To corroborate the efficacy of the proposed method, an experiment has been conducted on our own data sets.

In chapter 7, we proposed a system for detecting, tracking and then segmenting of a flower in a flower video based on a query sketch of the flower. The proposed system has three stages detection, tracking and segmentation. In first stage a sketch of a flower of interest is given as an input. The edge orientation information of the given sketch is matched against that of an individual frame in search of a location of the flower of interest. In second stage, the minimum computed cost between query sketch and video frames is used for tracking the flower in entire flower videos. Finally, the tracked flower regions in all frames are segmented. Experiments are conducted on our real dataset to evaluate the performance of the proposed method. To study the effectiveness of the proposed method we have compared the obtained results with the ground truth provided by five human experts.
In chapter 8, overall summary, the major contributions of the research work presented in this thesis and avenues for further research work in the proposed directions are presented.