

A Histogram-based Classification of Image Database Using Scale Invariant Features

Shashwati Mishra

Ph. D Research Scholar, Utkal University, Vani Vihar, Bhubaneswar, Odisha, India
Email: shashwati.mishra@gmail.com

Mrutyunjaya Panda

Reader, Department of Computer Science and Applications, Utkal University, Vani Vihar, Bhubaneswar, Odisha, India
Email: mrutyunjayapanda@yahoo.in

Abstract—Development of advanced technology has increased the size of data and has also created different categories of data. Classifying these different categories of data is the need of the era. We have proposed a method of classifying the image database containing four categories of images like human face, airplane, cup and butterfly. Our approach involves steps like feature extraction, bag of feature creation, histogram representation and classification using decision tree. For feature extraction SIFT (Scale Invariant Feature Transform) algorithm is used since it is invariant to rotation, change of scale, illumination etc. After extracting the features the bag of features concept is used to group the features using k-means clustering algorithm. Then a histogram is plotted for each image in the image database which represents the distributions of data in different clusters. In the final step the most robust, simple and flexible decision tree algorithm is applied on the table created from the histogram plots to obtain the classification result. The experimental observations and the calculated accuracy proves that this method of classification works well for classifying an image dataset having different categories of images.

Index Terms—Classification, bag of features, decision tree, histogram, feature extraction

I. INTRODUCTION

In the era of science and technology, development of advance technologies has made our work easier, but on the other hand it has also increased the amount and categories of data or information stored. Besides numerical and categorical data the advanced technology has created different categories of data to be stored like digitized image data, audio, video etc. [1]. In this paper our focus is in classification of image data with the help of image processing and data mining techniques. Different data mining techniques are used to handle the huge volume of data stored in databases. Similarly the advancement in image acquisition approaches, image processing approaches and the way image is stored has led to the development of a new field called image mining. The images can be analyzed to extract embedded

knowledge, image data relationship hidden in an image [2]. Image plays a very important role in different sectors. The increase in different categories of images like business images, medical images, satellite images has made the field of image mining popular [3]. Information in the form of an image can be easily understood and obtained by nontechnical people and they can also retrieve their requirements easily.

A technique of handling the information mined from image databases, finding the relationship between image and data and extracting the hidden information from the images is called image mining [4]. Image mining involves extraction of knowledge from an image, extraction of an object from an image, content-based image retrieval etc. Content based image classification and retrieval is the popular research area in the recent years. For an image color, shape, texture are the features used to describe an image [5, 6, 7]. Then the images whose features match with that of the feature in the query are extracted from the database. The main aim of image mining is to process the image sets and convert it into such a form that it is possible to apply the data mining algorithms on it [2]. Processing the low level pixel information to recognize the high level image objects and relationship among different objects is the most important challenge in image mining [2]. Like classification of datasets, image sets can also be classified into different categories. Image classification is a complex process which is affected by many factors. This process involves several steps like pre-processing and post-processing steps, selection of a suitable classification system, extraction of features from the images etc. [8]. Extraction of image content information and their use for different purposes is one of the popular and successful methods. An image contains redundant and more memory consuming raw data. The aim of extracting image content information is to extract certain data points from this huge volume of data called features which make the succeeding steps simple. A good feature must be perceptually meaningful, identifiable on different images, invariant to rotation and scale change, insensitive to noise, robust to change in illumination etc. [9].

K. Lu and D. Yang [10] proposed a decision tree based image mining technique. They extracted pixel-wise image

features and created a database like table from it. They applied the data mining algorithm on this table. Their experimental observations prove that their model is very effective and efficient for image processing and image mining.

Batik images (a traditional fabric of Indonesia) were classified by extracting features using SIFT (Scale Invariant Feature Transform), constructing bag of features and using SVM (Support Vector Machine) for classification [11].

Classification of human faces and nonfaces was done using machine learning techniques. The classification accuracy was compared for classification methods like Bayes, Function, Lazy, Meta, Misc, Rules and Tree [12].

X. Bai and V. Cherkassky [13] proposed a method for gender classification of human faces which uses inference through contradictions. R. Zhang and Z. Zhang [14] developed a concept vector model based image classification method. An approach of classifying remote sensing images using decision tree was proposed by L. Jiang et al. [15] and A. P. Pooja et al. [16].

D. Wilking and T. Röfer [17] presented the use of decision trees for recognizing a real time object. K. Stapor et al. [18] developed a method of automatic segmentation and classification of eye images using multilayer perceptron as the classification procedure.

M. P. Patil and S. R. Kolhe [19] proposed an automatic image annotation technique using decision trees and rough sets. Our aim is to develop a method which can accurately classify the images present in the image database.

This paper is organized as follows. Section 2 describes the steps like feature extraction using SIFT (scale Invariant Feature Transform), bag of features creation, histogram representation, classification using decision tree of our proposed method. Section 3 explains the environmental set up used for applying the proposed methodology. Section 4 contains the experimental observations along with the explanations and performance evaluation. The concluding remarks and future work are discussed in section 5.

II. PROPOSED METHODOLOGY

In this paper, our approach is to classify the images present in an image database into different categories. Starting from an initial set of different categories of images, feature extraction process derives informative and non-redundant values known as features. This process is also known as feature selection. These features make subsequent steps simple and help in better human interpretations. Feature extraction plays an important role in image processing, machine learning and pattern recognition. Different types of feature selection algorithms are there depending on the types of features to be extracted. In an image features may be low level features or features that are used to detect curvature, motion, shape of objects etc. SIFT (Scale Invariant Feature Transform) is one of the popular methods for low level feature selection.

Bag of features (BoF) model is the widely used model for image classification. A bag of visual words represents a vector of occurrence counts of a vocabulary of local image features. This model considers image features as words. To apply Bag of features model, first the features should be extracted from the images in the database. SIFT algorithm detects image features and represents each local patch as 128-dimensional numerical vectors. These vectors are known as feature descriptors which act as the input to the bag of features creation step. The aim is to create a visual vocabulary by clustering the features extracted from the images. For grouping the features according to their similarity we have applied K-means clustering algorithm. K-means is one of the simple unsupervised clustering method to generate clusters. We have also applied this algorithm to all the vectors generated in the first step. The centers of clusters formed are also called as the codewords. Similarly, the number of clusters represent the codebook size. Our second step is to create a dictionary of visual words.

The bag of features creation step maps each feature in the image to a codeword or cluster centers. Now the image can be represented as a histogram of codewords or clusters formed. This histogram of codewords or clusters formed represents the probability distribution of the codewords or features in different clusters. For this the frequency of occurrence of codewords in different range of values or the number of features present in different clusters are calculated. Our third step is to plot histogram for each image.

Finally the decision tree algorithm is applied which classifies the image present in the image set into different categories. The diagrammatic representation of the proposed approach is shown in fig. 1.

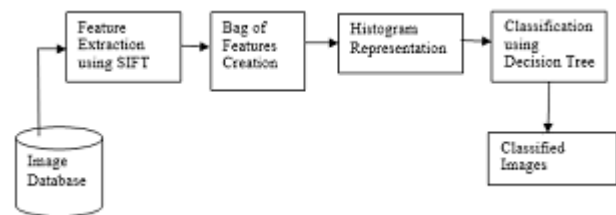


Fig.1. Proposed methodology.

A. Feature Extraction using SIFT (Scale Invariant Feature Transform)

For any object interesting point on it can be extracted, which provides a feature description of the object. These feature descriptions can be used to locate the object in an image containing other objects. These features may be affected by several factors like changes in image scale, noise, illumination etc. SIFT (Scale Invariant Feature Transform) image features provide such a set of features, which are not affected by these factors.

The feature descriptions extracted from a training image can be used to identify the object when attempting to locate the object in the test image containing many other objects. To perform reliable recognition it is required the features extracted from training image be

detectable even changes in image scale, noise and illumination. Such points usually lie on high contrast regions of the image, such as object edges. Another important characteristic of these features is that the relative position in the original scene should not change from one image to another [20]. The feature points extracted using SIFT algorithm are invariant to image scale change and rotation, robust to 3D view point, affine distortion, presence of noise, change in illumination.

In computer vision SIFT (Scale Invariant Feature Transform) algorithm is used to extract local features from an image. This algorithm was published by David Lowe from UBC (University of British Columbia). Applications of this algorithm include object recognition, localization and mapping of robot, Panorama stitching, modeling recognition and tracking of 3D scene. For extracting these features the SIFT algorithm uses a four stage filtering approach [20].

1. Scale-Space Extrema Detection:-

In this step, those locations and scales are identified that can be detected from different views of the same object. This can be achieved by using a “scale space” function, which is based on Gaussian function. The scale space is calculated by (1) for an image $I(x, y)$.

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \quad (1)$$

Where, * is the convolution operator.

$G(x, y, \sigma)$ is a variable-scale Gaussian and can be computed using (2).

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2} \quad (2)$$

Then stable key-point locations in the scale-space are detected. Out of various techniques difference of Gaussians is one technique, which helps to find out scale-space extrema, $D(x, y, \sigma)$. The scale space extrema is calculated by finding the difference between two images, one image with scale k times the other image as in (3).

$$D(x, y, \sigma) = L(x, y, k\sigma) - L(x, y, \sigma) \quad (3)$$

To detect the local maxima and minima of $D(x, y, \sigma)$ each point is compared with its eight neighbors at the same scale, and its nine neighbors up and down one scale. To be an extrema this value should be either a maximum or minimum among all these points [20].

2. Keypoint Localization:-

The aim of this stage is to check the list of keypoints extracted for finding low contrast or poorly localized points on an edge. Then such points are eliminated to reduce the size of extracted keypoints. For achieving this goal Laplacian value for each keypoint (which are found in the previous stage) is calculated. So, using (4) the location of extremum, z , can be computed.

$$z = -\frac{\partial^2 D^{-1}}{\partial x^2} \frac{\partial D}{\partial x} \quad (4)$$

The function value z is compared with a threshold value. The point is excluded if this value is below the threshold value. Thus extrema with low contrast can be removed. In case of extrema based on poor localization, there may be a large principal curvature or a small curvature. Large principal curvature is found across the edge and curvature is small in the perpendicular direction in the difference of Gaussian function. The keypoint is rejected if this difference is below the ratio of largest to smallest eigenvector, from the 2x2 Hessian matrix at the location and scale of the keypoint [20].

3. Orientation Assignment:-

Considering the local image properties this step assigns a consistent orientation to the extracted keypoints. Then the keypoint descriptors can be represented relative to this orientation and invariance to rotation can be achieved. Initially a Gaussian smoothed image is created using the keypoints scale. Assuming L as the Gaussian smoothed image, gradient magnitude m and orientation θ are computed using (5) and (6) respectively.

$$m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2} \quad (5)$$

$$\theta(x, y) = \tan^{-1}\left(\frac{L(x, y+1) - L(x, y-1)}{L(x+1, y) - L(x-1, y)}\right) \quad (6)$$

Gradient orientations of sample points are represented in the form of an orientation histogram. The highest peak and the peaks within 80% height of the highest peak of this orientation histogram are extracted. These peaks can be used to create a keypoint with the calculated orientation. Some keypoints may have multiple orientations also. To interpolate the position of peaks, a parabola is fitted to the three histogram values closest to each peak [20].

4. Keypoint Descriptor:-

To create keypoint descriptors local gradient data is also used. First to line up with the orientation of the keypoint, the gradient information is rotated. Then weighted by a Gaussian with a variance of $1.5 * \text{keypoint scale}$. A set of histograms over a window centered on the keypoint is created from this data.

Typically keypoint descriptors use a set of 16 histograms which are aligned in a 4×4 grid and each having 8 orientation bins. One orientation for each of the main compass directions and one orientation for each of the mid points of the main compass directions. Finally it creates a feature vector having 128 dimensions which are nothing but the SIFT (Scale Invariant Feature Transform) keys [20].

Advantages of using SIFT (Scale Invariant Feature Transform) algorithm are:-

1. Locality:-Features are local, so robust to occlusion and clutter.
2. Distinctiveness:-Individual features can be matched to large database of objects.
3. Quantity:-Many features can be generated for even small objects.
4. Efficiency:-Close to real-time performance.

B. Bag of Features Creation

In this step the concept of bag of features is used for generating the codebook. Aim is to represent an image as orderless collection local features which are extracted in the previous step using SIFT (scale Invariant Feature Transform) algorithm. The name bag of features is used in case of images just like the bag of words representation is used for the text documents. The concept of bag of features for image representation can be explained in two ways. One approach is by comparing with the concept of bag of words representation and another is from a codebook perspective. Generally bag of words is used for textual documents and in this technique a document is converted to a normalized histogram considering the number of occurrences of words in the text document. All the words of the document are counted which may use a single term to represent the synonyms and may exclude non-informative words like 'the'. The text document is represented by a sparse vector. Each element of this vector is a term in the dictionary. The number of times that term appears in the document divide by the total number of dictionary words in the document represents the value of the element. Bag of words representation of a document is the term vector. In this term vector there is no ordering of words which is there in the original document. The concept of bag of features is similar to that of bag of words. The difference is only that here we consider the features extracted from the images. These features are grouped according to their similarity. Since no information is available regarding the similarity between the extracted features, generally unsupervised technique like clustering is used to group these features and create a discrete vocabulary from a large number of local features extracted from the images. These clustered features represent the visual vocabulary which acts as the dictionary and represent local areas of the image. So finally we can able to group features in such a way that each feature is assigned to its nearest cluster center. Each cluster represents a visual word. The quantized features extracted from the image are represented in the form of a normalized histogram which acts as the term vector [21].

In the second approach initially features are extracted from the images and then a visual codebook is generated using the concept of vector quantization [21]. A quantization technique that allows the modelling of probability density functions by the distribution of prototype vectors is nothing but the vector quantization technique. In this technique a large set of points (or vectors) is divided into groups having approximately equal number of points closest to them. The centroid of a group represents that group as in case of a clustering algorithm. So finally we will get a visual codebook

where the image features are assigned to the nearest code in the codebook. Now the image is in the form of a set of codes and the histogram is plotted which represents these set of codes present in the codebook. The concept of normalized histogram of visual words and normalized histogram of codes is exactly same. We can conclude that a term vector in case of bag of features is a way of representing an image compactly discarding the large scale spatial information, the positions, change in orientations and scales of the extracted features.

For extracting a bag of visual features k-means algorithm is used on the extracted feature vectors in our proposed approach. In this paper dictionary size is assumed as 20. This means after applying k-means algorithm 20 mutually exclusive clusters can be obtained, where each cluster center represents a visual word. Then the extracted feature vectors of a particular image are assigned to the nearest cluster. The number of feature vectors assigned to a particular cluster represents the length of a histogram for a particular bin.

K-means:-

K-means is a partitioning base clustering method. A partitioning based method partitions a dataset D of n objects into K partitions ($k \leq n$). Each partition represents a cluster. K-means algorithm partitions the set of input objects into k clusters in such a way that the intracluster similarity is high but the intercluster similarity is low. The mean value of the objects in a cluster is considered as the cluster similarity. This mean value represents the cluster's centroid or center of gravity [22].

Algorithm:-

This algorithm takes as input the objects present in the dataset and the number of clusters to be created. Let d is a dataset having n number of objects and the number of clusters to be created is k. Finally this algorithm creates k number of clusters from the dataset d.

Steps:

1. Randomly choose k objects from the given dataset d. These k objects act as the initial cluster centers.
2. Repeat the following steps until there is no change
 - a. Find the mean value of the objects in the cluster.
 - b. Compare the value of the object with the calculated mean value, accordingly assign the object to the most similar cluster.
 - c. Recalculate the cluster means [22]

C. Histogram Representation

Distribution of numerical data is represented graphically with the help of a histogram. It was first introduced by Karl Pearson. Histogram representation involves two steps.

1. The first step is to divide the range of values into a series of intervals. Each interval is a bin which is consecutive, non-overlapping, adjacent and of equal size.

2. In the second step frequency of each interval is calculated which is nothing but the number of values in each interval.

We have created 20 clusters so our histogram plot shows 20 bins. This process is repeated for each image in the database.

D. Decision Tree

After finding the histogram for each image decision tree algorithm is applied to classify the images present in the database into different categories. For finding the best splitting criterion information gain is considered as the attribute selection measure.

The algorithm used to generate a decision tree from the given class-labeled training datasets is known as the decision tree induction. A decision tree is nothing but a tree like structure where the topmost node is the root node, each internal node denotes a test on an attribute, each branch of the tree represents the result of the test and each leaf node contains the class label. From these decision trees, classification rules can be easily generated. This classifier is very popular because it can handle high dimensional data, easy to understand, no domain knowledge is required, simple, fast and accuracy is also good.

A researcher in machine learning J. Ross Quinlan, expanded the work on concept learning systems described by E. B. Hunt, J. Marin and P.T. Stone and developed a decision tree algorithm known as ID3(Iterative Dichotomiser) during the late 1970s and early 1980s. Later Quinlan developed a successor of ID3 called C4.5 [22].

The decision tree algorithm uses an attribute selection measure for finding the splitting criterion that best separates a given data set into individual classes. Since these measures determine the way the tuples at a given node are to be split, so these are also known as splitting rules. Information gain as the attribute selection measure is used in the ID3 algorithm.

Change in the information entropy H from a prior state to a state that takes some information is nothing but the expected information gain [22]. Considering $H(X)$ as the entropy and $H(X|a)$ as the conditional entropy, information gain can be represented mathematically as in (7).

$$IG(X, a) = H(X) - H(X|a) \quad (7)$$

Entropy:-

Assume X is a discrete random variable having values $\{x_1, x_2, \dots, x_n\}$. Suppose the probability mass

function is $P(X)$, E is the expected value operator and I is the information content of X . Then entropy can be defined as in (8).

$$H(X) = E[I(X)] = E[-\ln(P(X))] \quad (8)$$

Equation (8) can be rewritten as in (9).

$$H(X) = \sum_{i=1}^n P(x_i) I(x_i) = -\sum_{i=1}^n P(x_i) \log_b P(x_i) \quad (9)$$

Where, b is the base of the logarithm used.

Conditional Entropy:-

Let X and Y are two events taking values x_i and y_j respectively. Then conditional entropy of these two events is computed using (10).

$$H(X|Y) = \sum_{i,j} P(x_i, y_j) \log \frac{P(y_j)}{P(x_i, y_j)} \quad (10)$$

Where, $P(x_i, y_j)$ is the probability that $X = x_i$ and $Y = y_j$

Decision Tree Algorithm:-

This algorithm creates a decision tree from a given dataset D . Input to this algorithm are the dataset D , `list_of_attributes`, `select_attribute` method (finds the splitting criteria which partition the dataset in best way) and `partitioning_attribute`.

Steps:

1. Create the root node R
2. Check the class of the tuples in D . If all are of same class then make R as leaf node, label R with class label information and return R .
3. If `list_of_attributes` is empty then make R as a leaf node, label R with the majority class information and return R .
4. Use `select_attribute(D, list_of_attributes)` to find the best `partitioning_criteria`.
5. Label R with the `partitioning_criteria`
6. If `partitioning_attribute` has discrete values and multiway splits are allowed then
7. `list_of_attributes = list_of_attributes - partitioning_attribute`
8. Repeat the following steps for each output o of the `partitioning_criteria`
 - a. let D_o be the set of data tuples in D satisfying the outcome o
 - b. if D_o is empty then attach a leaf labeled with the majority class in D to node R
 - c. else attach the node returned by the decision tree algorithm to node R
9. return node R [22].

III. EXPERIMENTAL SETUP

For the implementation of the proposed approach, Matlab 2013a, which runs on a computer having 64 bit windows 7 operating system, core i5 processor with 2.30 GHZ speed, 4 GB RAM is used. Four different categories (five from each category) of images such as human face,

airplane, cup and butterfly are considered to create an image database of size 20. Human face is assumed as class 1, airplane as class 2, cup as class 3 and butterfly as class 4. The aim is to classify these different categories of objects using the proposed methodology. As explained earlier initially key points or feature points are extracted using SIFT algorithm on a gray scale image. In case of color image, the image is first converted to gray scale image before applying SIFT algorithm. This process is repeated for all the images in the image database and the feature points are extracted from each image separately. In the next step separate groups of features are created using bag of features concept and then the histograms are plotted. One histogram represents the distribution of data points for a particular image. In the final step, decision tree algorithm gives the final classified result using the table created from the histogram plots.

IV. EXPERIMENTAL OBSERVATIONS AND DISCUSSIONS

The experimental results obtained using the proposed methodology are discussed in this section. Fig. 2 shows five different human faces, its gray scale conversions, features extracted from each image and the histogram plots for it. The number features extracted is different for different images. Features extracted from first image is 1740 (shown in 1(c)), second image is 1411 (shown in 2(c)), third image is 1663 (shown in 3(c)), fourth image is 1673 (shown in 4(c)), fifth image is 1523 (shown in 5(c)).

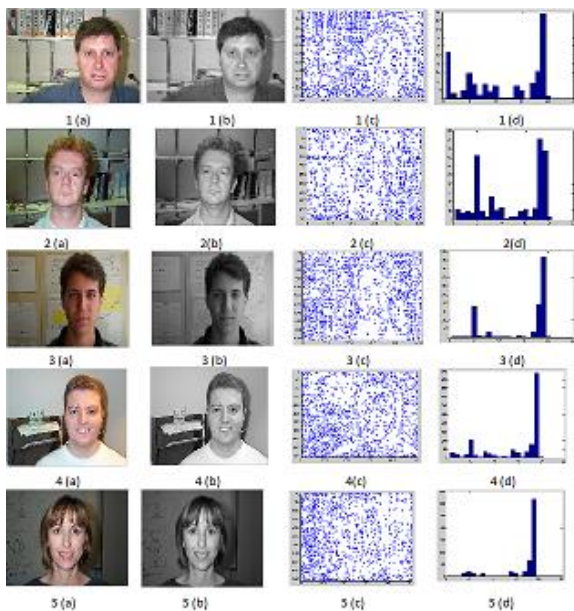


Fig.2. Original human face image[1(a), 2(a), 3(a), 4(a), 5(a)], Conversion to gray scale[1(b), 2(b), 3(b), 4(b), 5(b)], Extracted feature points[1(c), 2(c), 3(c), 4(c), 5(c)], Histogram plot[1(d), 2(d), 3(d), 4(d), 5(d)] .

Extracted features points from five different airplane images are shown in Fig. 3. For the first image number is 1033 (shown in 1(c)), for the second image it is 1346 (shown in 2(c)). Similarly for third, fourth and fifth images number of features extracted are 1054 (shown in

3(c)), 710 (shown in 4(c)) and 1358 (shown in 5(c)) respectively.

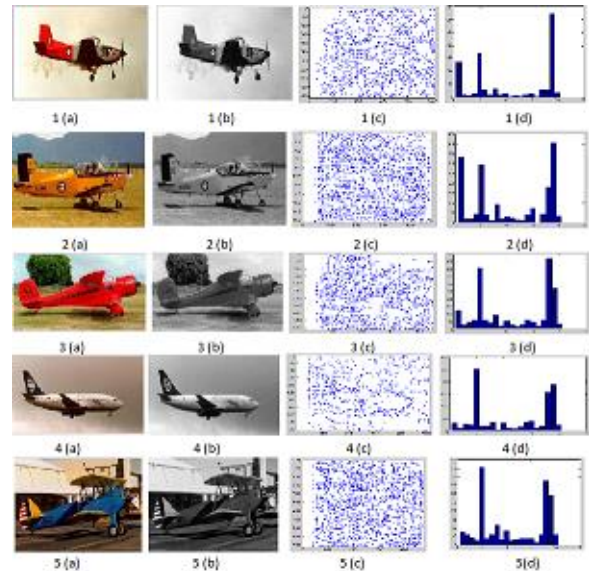


Fig.3. Original airplane image[1(a), 2(a), 3(a), 4(a), 5(a)], Conversion to gray scale[1(b), 2(b), 3(b), 4(b), 5(b)], Extracted feature points[1(c), 2(c), 3(c), 4(c), 5(c)], Histogram plot[1(d), 2(d), 3(d), 4(d), 5(d)] .

Features extracted from first, second, third, fourth and fifth cup images of Fig. 4 are 960 (shown in 1(c)), 1381 (shown in 2(c)), 726 (shown in 3(c)), 1453 (shown in 4(c)), 1234 (shown in 5(c)) respectively.

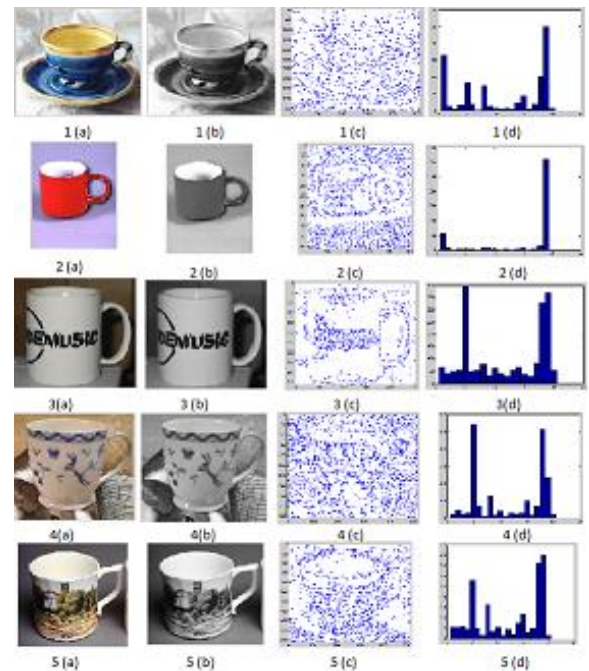


Fig.4. Original cup image[1(a), 2(a), 3(a), 4(a), 5(a)], Conversion to gray scale[1(b), 2(b), 3(b), 4(b), 5(b)], Extracted feature points[1(c), 2(c), 3(c), 4(c), 5(c)], Histogram plot[1(d), 2(d), 3(d), 4(d), 5(d)] .

Like other three categories in the butterfly category also features are extracted from the five different images. Number of features extracted are 2018, 1571, 2052, 1027, 1114 for first, second, third, fourth and

fifth image of butterfly category respectively. The results obtained from butterfly images are shown in Fig. 5.

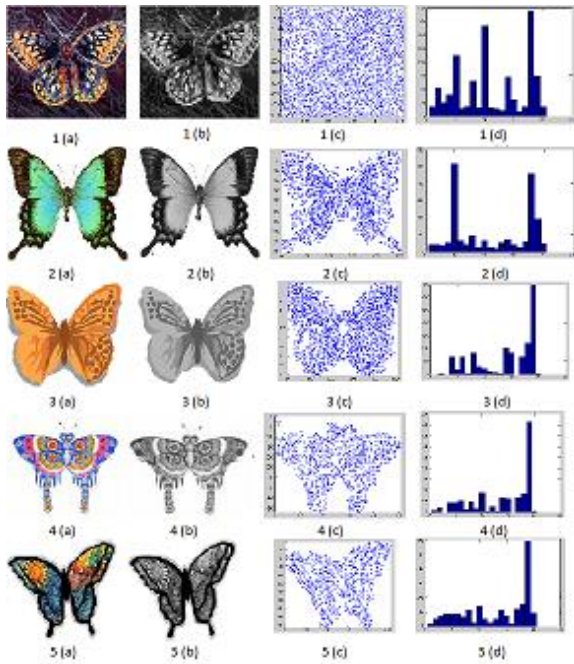


Fig.5. Original butterfly image[1(a), 2(a), 3(a), 4(a), 5(a)], Conversion to gray scale[1(b), 2(b), 3(b), 4(b), 5(b)], Extracted feature points[1(c), 2(c), 3(c), 4(c), 5(c)], Histogram plot[1(d), 2(d), 3(d), 4(d), 5(d)]).

In Fig. 2, Fig. 3, Fig. 4 and Fig. 5 the x-axis of the histogram plot of an image represents number of bins. Y-axis represents number of data items in each bin. By observing the histogram plots of different categories of images it can be concluded that it is very difficult to classify images from their histogram plots. Therefore a classification algorithm is needed. As class label information is available a supervised classification algorithm can be applied. In this paper, the decision tree classification algorithm is applied on the data set (shown in table 1) generated from these histogram plots.

Applying decision tree algorithm on the database table created from histogram plots a decision tree as in Fig. 6 can be obtained.

In Fig. 6 class 1 represents human face, class 2 represents airplane, class 3 represents cup and class 4 represents butterfly images. 4, 20, 15 etc. indicate the bin numbers.

So $4 < 53.50$ means values in bin 4 is less than 53.50.

Thus from Fig. 6 it can be concluded that, if $\text{bin4} \geq 53.50$ and $\text{bin1} < 82.50$ then $\text{class}=4$

Class 4 means this is the image of a butterfly. Similarly rules for class 1, class 2 and class 3 category of images can also be generated.

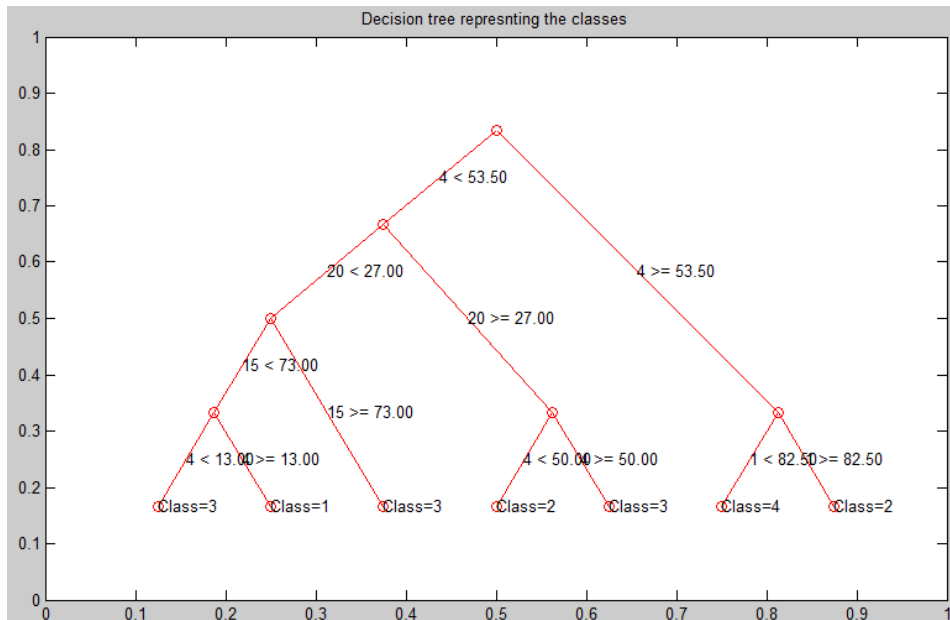


Fig.6. Decision tree representation of the database table.

Table 1. Database table created from histogram plots

| Class Labels | Bin1 | Bin2 | Bin3 | Bin4 | Bin5 | Bin6 | Bin7 | Bin8 | Bin9 | Bin10 | Bin11 | Bin12 | Bin13 | Bin14 | Bin15 | Bin16 | Bin17 | Bin18 | Bin19 | Bin20 |
|--------------|------|------|------|------|------|------|------|------|------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| 1 | 268 | 30 | 9 | 51 | 147 | 84 | 23 | 84 | 41 | 73 | 6 | 10 | 4 | 72 | 52 | 0 | 89 | 157 | 492 | 19 |
| 2 | 269 | 19 | 13 | 25 | 341 | 51 | 18 | 56 | 21 | 29 | 8 | 16 | 8 | 24 | 26 | 6 | 54 | 50 | 643 | 34 |
| 2 | 330 | 18 | 17 | 42 | 296 | 39 | 9 | 93 | 19 | 29 | 24 | 12 | 9 | 37 | 73 | 3 | 44 | 180 | 404 | 33 |
| 3 | 334 | 23 | 5 | 32 | 170 | 42 | 2 | 141 | 28 | 13 | 5 | 14 | 6 | 47 | 82 | 4 | 39 | 209 | 501 | 14 |
| 3 | 228 | 17 | 2 | 10 | 12 | 9 | 0 | 28 | 22 | 9 | 1 | 16 | 0 | 13 | 26 | 2 | 16 | 56 | 1244 | 0 |
| 4 | 37 | 106 | 52 | 80 | 228 | 28 | 34 | 131 | 33 | 333 | 32 | 27 | 20 | 145 | 61 | 19 | 34 | 391 | 148 | 36 |
| 4 | 39 | 39 | 36 | 55 | 519 | 61 | 44 | 92 | 23 | 66 | 25 | 20 | 50 | 75 | 45 | 23 | 59 | 462 | 194 | 48 |
| 3 | 67 | 43 | 49 | 52 | 394 | 48 | 56 | 80 | 33 | 60 | 44 | 27 | 52 | 67 | 36 | 31 | 83 | 329 | 371 | 53 |
| 2 | 126 | 29 | 44 | 58 | 408 | 60 | 41 | 92 | 20 | 54 | 27 | 16 | 29 | 59 | 51 | 23 | 56 | 474 | 272 | 36 |
| 2 | 65 | 28 | 56 | 48 | 508 | 39 | 39 | 77 | 14 | 62 | 26 | 17 | 31 | 69 | 81 | 27 | 38 | 320 | 380 | 50 |
| 1 | 59 | 38 | 48 | 41 | 364 | 49 | 20 | 134 | 50 | 68 | 10 | 20 | 21 | 49 | 59 | 18 | 63 | 454 | 390 | 20 |
| 2 | 80 | 60 | 40 | 29 | 458 | 39 | 56 | 118 | 34 | 82 | 28 | 34 | 37 | 38 | 36 | 17 | 56 | 380 | 290 | 63 |
| 3 | 18 | 46 | 24 | 30 | 543 | 72 | 11 | 126 | 17 | 42 | 10 | 33 | 19 | 31 | 104 | 13 | 69 | 513 | 233 | 21 |
| 3 | 54 | 59 | 62 | 45 | 284 | 52 | 14 | 164 | 33 | 59 | 15 | 45 | 22 | 75 | 120 | 31 | 53 | 366 | 402 | 20 |
| 1 | 7 | 17 | 8 | 16 | 357 | 23 | 7 | 65 | 11 | 16 | 7 | 12 | 1 | 20 | 9 | 3 | 62 | 388 | 938 | 8 |
| 4 | 0 | 1 | 0 | 140 | 38 | 133 | 3 | 165 | 58 | 42 | 17 | 7 | 1 | 201 | 162 | 0 | 132 | 247 | 693 | 1 |
| 1 | 66 | 44 | 6 | 45 | 214 | 39 | 16 | 74 | 59 | 36 | 15 | 29 | 10 | 93 | 64 | 13 | 76 | 162 | 977 | 3 |
| 1 | 6 | 2 | 0 | 42 | 83 | 47 | 6 | 42 | 12 | 16 | 5 | 16 | 5 | 86 | 44 | 0 | 84 | 260 | 1280 | 5 |
| 4 | 9 | 41 | 0 | 77 | 81 | 104 | 19 | 94 | 28 | 173 | 6 | 43 | 4 | 130 | 117 | 0 | 126 | 163 | 825 | 1 |
| 4 | 21 | 53 | 71 | 90 | 86 | 89 | 63 | 111 | 29 | 140 | 48 | 27 | 48 | 112 | 70 | 20 | 121 | 153 | 594 | 95 |

Table 2. Confusion matrix

| Total Number of Instances=20 | Types of Image Objects | Predicted: NO | | | | Predicted: YES | | | |
|------------------------------|------------------------|---------------|----------|-----|-----------|----------------|----------|-----|-----------|
| | | Human Face | Airplane | Cup | Butterfly | Human Face | Airplane | Cup | Butterfly |
| Actual: NO | Human Face | 0 | | | | 0 | | | |
| | Airplane | | 0 | | | | 0 | | |
| | Cup | | | 0 | | | | 0 | |
| | Butterfly | | | | 0 | | | | 0 |
| Actual: YES | Human Face | 0 | | | | 5 | | | |
| | Airplane | | 0 | | | | 5 | | |
| | Cup | | | 0 | | | | 5 | |
| | Butterfly | | | | 0 | | | | 5 |

Confusion matrix:-

Confusion matrix is typically used in supervised learning and is also known as error matrix. In unsupervised learning it is called as matching matrix. The main purpose of using this matrix is to find out the performance of a classifier. According to the definition of confusion matrix it is a tabular representation used to describe the performance of a classifier or classification model on a set of test data for which the true values are known. Table 2 shows the confusion matrix representation of the results obtained by our proposed method. From these values accuracy can also be calculated using (11).

$$Accuracy(ACC) = \frac{\sum TP + \sum TN}{\sum TotalPopulation} \quad (11)$$

Where,

TP= True Positive

TN= True Negative

So, accuracy = (20+0)/20=1

Therefore the accuracy is 100%. Hence it can be concluded that our approach works well for classifying the image databases.

Computational Cost:-

Time taken for getting the final classified result in Matlab 2013a for classifying twenty (20) images of four (4) categories is 14.365 seconds. This time will increase with increase in the number of images and types or classes of images in the image database.

V. CONCLUSION AND FUTURE WORK

Our proposed approach classify the image database considering the features extracted from the images. Due to the characteristics of SIFT (Scale Invariant Feature Transform) algorithm, the features extracted remain invariant to scale change, rotation, illumination etc. After constructing the histogram plots from the bag of features the decision tree algorithm is applied. The flexibility and robustness of this classification algorithm can be observed from the experimental results. The performance evaluation proves that this method gives better accuracy than other methods of image database classification. Our future work is to reduce the time taken for classifying the images in the database and develop a classification method to classify the image dataset containing a large number of images. Future research scope includes development of a technique which can extract a specific category of object according to the user requirement from an image containing different categories of objects.

REFERENCES

- [1] Z. Aghbari and A. Makinouchi, "Semantic approach to image database classification and retrieval," *NII Journal*, No. 7, pp. 1-8, 2003.
- [2] A. Hazra, K. Bhattacharjee, A. Ghosh and T. Biswas, "Image mining techniques for efficient searching of images from internet," *International Journal of Enhanced Research in Science, Technology and Engineering*, Vol. 5, No. 4, pp. 83-87, 2016.
- [3] A. Kannan, V. Mohan and N. Anbazhagan, "Image clustering and retrieval using image mining techniques," *IEEE International Conference on Computational Intelligence and Computing Research*, pp. 1-5, 2010.
- [4] R. Sudhir, "A survey on image mining techniques: theory and applications," *Computer Engineering and Intelligent Systems*, Vol. 2, No. 6, pp. 44-53, 2011.
- [5] P. S. Sarvani and Y. M. Latha, "Image mining," *International Journal of Science and Research*, Vol.-2, No.-4, pp. 401-403, 2013.
- [6] A. A. Chaugule, S. N. Mali, "Evaluation of shape and color features for classification of four paddy varieties," *International Journal of Image, Graphics and Signal Processing*, Vol.-6, No.-12, pp. 32:38, 2014.
- [7] K. Arai, I. N. Abdullah, H. Okumura, "Image retrieval based on color, shape and texture for ornamental leaf with medicinal functionality images," *International Journal of Image, Graphics and Signal Processing*, Vol. 6, No. 7, pp.10:18, 2014.
- [8] D. Lu and Q. Weng, "A survey of image classification methods and techniques for improving classification performance," *International Journal of Remote Sensing*, Vol. 28, No. 5, pp. 823-870, 2007.
- [9] A. Aichert, "Feature extraction techniques," *CAMP Medical Seminar WS0708*, pp. 1-8, 2008.
- [10] K. Lu and D. Yang, "Image processing and image mining using decision trees," *Journal of Information Science and Engineering*, pp. 989-1003, 2009.
- [11] R. Azhar, D. Tuwohingide, D. Kamudi, Sarimuddin and N. Suciati, "Batik image classification using SIFT feature extraction, bag of features and support vector machine," *Elsevier*, pp. 24-30, 2015.
- [12] L. Hussain, W. Aziz, Z. H. Kazmi and I. A. Awan, "Classification of human faces and non faces using machine learning techniques," *International Journal of Electronics and Electrical Engineering*, Vol. 2, No. 2, pp. 116-123, 2014.
- [13] X. Bai and V. Cherkassky, "Gender classification of human faces using inference through contradictions," *International Joint Conference on Neural Networks*, IEEE, pp. 746-750, 2008.
- [14] R. Zhang and Z. Zhang, "Image database classification based on concept vector model," *International Conference on Multimedia and Expo*, IEEE, pp. 93-96, 2005.
- [15] L. Jiang, W. Wang, X. Yang, N. Xie and Y. Cheng, "Classification methods of remote sensing image based on decision tree technologies," *Computer and computing technologies in agriculture IV*, Vol. 344, pp. 353-358, 2011.
- [16] A. P. Pooja, J. Jayanth and S. Koliwad, "Classification of RS data using decision tree approach," *International Journal of Computer Applications*, Vol.-23, No.-3, pp. 7-11, 2011.
- [17] D. Wilking and T. Röfer, "Realtime object recognition using decision tree learning," *Lecture notes in Artificial Intelligence*, Springer, pp. 556-563, 2005.
- [18] K. Stapor, L. Pawlaczyk, R. Chrzątek, H. Niemann and G. Michelson, "Automatic segmentation and classification of fundus eye images for glaucoma diagnosis," *Springer*, pp.65-69, 2004.
- [19] M. P. Patil and S. R. Kolhe, "Automatic image annotation using decision trees and rough sets," *International Journal of Computer Science and Applications*, Vol.11, No. 2, pp. 38-49, 2014.
- [20] D. G. Lowe, "Distinctive image features from scale-invariant key points," *International Journal of Computer Vision*, PP. 91-110, 2004.
- [21] S. O'hara and B. A. Draper, "Introduction to the bag of features paradigm for image classification and retrieval," *Arxiv preprint arxiv:1101.3354*, pp. 1-25, 2011.
- [22] J. Han and M. Kamar, "Data mining concepts and techniques," *Elsevier*, 2006.

Authors' Profiles



Shashwati Mishra has completed her M.C.A from Ravenshaw University in 2009 and M. Tech from KIIT University in 2011. She has received Gold medal from Ravenshaw University and Vice Chancellor's Silver Medal from KIIT University for securing highest marks in M.C.A. and M. Tech respectively. She is presently pursuing Ph. D. in Utkal University and working as a lecture in Ravenshaw University. Her areas of interest include Data mining, Image mining, Artificial intelligence, Pattern recognition etc. She has more than 5 years of teaching experience.



Mrutyunjaya Panda working as Reader in the Computer Science and Applications department of Utkal University. He has 18 years of teaching and research experience. His teaching and research domains include Data mining, Soft computing, Digital signal processing, Mobile communication, Social networking, Image processing, Wireless sensor network, Text

mining etc.. He has published 70 research papers in different reputed journals and conferences. He has also written 2 text books, 3 edited books, 7 book chapters in different fields. He is a Reviewer of conferences of repute like: WICT, KES, SOCPROS, CCCS, IIMSS etc., Reviewer of Elsevier, Springer, WSEAS Transaction, Bentham Science, Wiley journals, TPC

member of SOCPAR, SOCPROS, WICT series conferences, Editorial member of Inderscience Journal, Listed in Marquis Who's who in engineering in 2013, Nominated for Best Teacher award by NFED, Hyderabad.

How to cite this paper: Shashwati Mishra, Mrutyunjaya Panda, " A Histogram-based Classification of Image Database Using Scale Invariant Features", *International Journal of Image, Graphics and Signal Processing(IJIGSP)*, Vol.9, No.6, pp.55-64, 2017.DOI: 10.5815/ijigsp.2017.06.07