Object Annotation using Fuzzy Object Shape Retrieval

Rotney Roy Meckamalil
P.G. Student
Department of Computer Science & Engineering
Federal Institute of Science & Technology

Abstract

In the world of object recognition, much effort has to be put to predict the particular object correctly by using the classifiers. There are only fewer amounts of trained data. Thereby the recognition part will be not accurately predicted. In the proposed system, a new technique has been introduced to recognize and annotate the object based on the Fuzzy-Object-Shape. The database contains different of set of images (apple, orange, ball etc.), each one have plenty of images. A good annotation system is by obtaining the correctly predicted the class labels when the particular image is being selected. The shapes like circle, ellipse, square, rhombus, rectangle, cone, and cylinder are being used to extract the features from the entire images in the database. The process is implemented in a MATLAB 2014.
Keywords: Object annotation, fuzzy membership function, object recognition

---

I. INTRODUCTION

Image processing is a method to convert an image into digital form and perform some operations on it, in order to get an enhanced image or to extract some useful information from it. The input is an image, a series of images, or a video, such as a photograph or video frame; the output of image processing may be either an image or a set of characteristics or parameters related to the image. Images are manually made from physical models of objects, environments, and lighting, instead of being acquired using cameras from natural scenes. There are several steps in image retrieval: Image Acquisition, Image Enhancement, Image Restoration, Color Image Processing, Wavelets and Multiresolution Processing, Compression, Morphological Processing, Segmentation, Representation and Description, Object recognition.

Object recognition is the final step in the image processing and it is an important task in image processing and computer vision. In today's world, the technology has been developed, so by increasing the images provide a door open to the development of robust and efficient object recognition techniques. With more and more information appear on the Internet and other digital multimedia as well as human beings eager for the fastest retrieval. Object recognition is a task of finding and identifying objects in an image or video sequence. Object annotation comes as the subcategory of object recognition.

The existing systems are text-based, but images are recognized based on the labels. The solution historically has been to develop text-based ontologies and classification schemes for image description. Text-based indexing has much strength including the ability to represent both general and specific notation of an object at varying levels of complexity. Assignment of terms to describe images is not solved entirely by the use of controlled vocabularies or classification schemes however. First, one could attempt to predict annotations of entire images using all information present which is annotation task. Next, one might attempt to associate particular words with particular image visual tokens which is correspondence task. Last but not least one might attempt to predict keywords for image object which is recognition task [1].Automatic Image Tagging is such a technique which associates an appropriate keyword from a given set of words or phrases based on the relevance to the content of the image [2].

The Problems with text-based access to images have prompted increasing interest in the development of image-based solutions. This is most often referred to as content-based Object Annotation Using Fuzzy-Object-Shape image retrieval. Content-based image retrieval relies on the characterization of primitive features such as color, shape, and texture that can be automatically extracted from the images themselves. Queries to CBIR systems are most often expressed as visual exemplars of the type of image or image attribute being sought.

Due to the advances in digital photography, storage capacity and networks speed, storing large amounts of high quality images has been made possible. Digital images are used in a wide range of applications such as medical, virtual museums, military and security purposes, and personal photo albums. However, users have difficulties in organizing and searching with large numbers of images in databases, as the current commercial database systems are designed for text data.

Object recognition is more complex processes that include visual inference that could constitute a drawback. Another relevant drawback is that building object recognition datasets is costly and thus the number of images that are collected is only limited. To overcome the disadvantage of having few training images per class and, which exploit the so called wisdom of crowd to populate object recognition datasets. Web based image annotation has become one of the emergent technique for the fastest recognition process. This provide a way of building large annotated datasets by relying on the collaborative effort of a large population of users. The outcome of CBIR, consists of millions of tagged images, but usually of these only few are accessible, and they are not well organized.

II. RELATED WORKS

Image recognition and annotation is a very fast growing research area in the last few years. Data is dynamically organized into groups according to relevance feedback from users. Image features that represent each of these regions are computed for indexing and searching. While semantic level annotation would be more desirable for users, given the current state of technology in image understanding, this is still very difficult to achieve.. Image recognition and annotation is a very fast growing research area in the last few years. Data is dynamically organized into groups according to relevance feedback from users. While semantic level annotation would be more desirable for users, given the current state of technology in image understanding, this is still very difficult to achieve. This is especially true when one has to deal with a heterogeneous and unpredictable image collection such as from the WWW.

In [3] B.C Russell proposes a database called Label Me database and an online Object Annotation Using Fuzzy-Object-Shape line annotation tool that allows the sharing of images and annotations. Web-based image annotation tool that was used to label the identity of objects and where they occur in images. Collecting a large number of high quality annotations, spanning many different object categories, for a large set of images, many of which are high resolution. Recover a depth ordering of objects in a scene, and increase the number of labels using minimal user supervision and images from the web. For each object present in an image, the labels should provide information about the objects identity, shape, location, and possibly other attributes such as pose. Users often provide textual tags to provide a caption of depicted objects in an image. In [4] Qingyong Li proposes Proposes a fuzzy approach to describe and to extract the fuzzy aesthetic semantic feature of art images. Apply neural network approach to model the process of human aesthetic perception and to extract the fuzzy aesthetic semantic feature vector.

In [5] Patrizio Frosini proposes a method by applying morphological operations and by Geometrical and topological properties are considered. Adding salt and pepper noise. Adopted euclidean distance. The advantages of the method is that Reduces time complexity and better retrieval result. The limitation of this method is that presence of Salt and pepper noise.In [6] G. Castellano, et.al proposes a fuzzy shape annotation approach for semi-automatic image annotation is presented. A fuzzy clustering process guided by partial supervision is applied to shapes represented by Fourier descriptors in order to derive a set of shape prototypes representative of a number of semantic categories. Next, prototypes are manually annotated by attaching textual labels related to semantic categories. Based on the labeled proto-types, a new shape is automatically labeled by associating a fuzzy set that provides membership degrees of the shape to all semantic categories.

In [7] Makoto Hasegawa et.al proposes Amplitude-only log Radon transform for geometric invariant shape descriptor. Combining the Radon transform, the amplitude extraction, and the log-mapping and invariant to shape rotation, scaling, and translation. The advantages of the above method are that the descriptor shows good performances with occlusion and additive noise. Estimation errors remain weak. Invariant to shape rotation, scaling, and translation. The limitation of the method is that Descriptor is performed on the spatial domain only.

In [8] Zhenbao Liu proposes Locality-constrained sparse patch coding for 3D shape retrieval. Geometric attributes including local, global, and topological features are considered. Then Locality constrained sparse coding is adopted. A number of patch words generated from a large set of 3D shapes. The advantages of this method are that reduces the number of point descriptors and robust retrieval against topology variation the limitationis that difficult to obtain visually suitable patches for all 3D models.

In [9] Qingyong Li et.al proposes a fuzzy approach to describe and to extract the fuzzy aesthetic semantic feature of art images. Aiming to deal with the subjectivity and vagueness of human aesthetic perception, utilize the linguistic variable to describe the image aesthetic semantics, so it becomes possible to depict images in linguistic expression such as `very action'. Applying neural network approach to model the process of human aesthetic perception and to extract the fuzzy aesthetic semantic feature vector

In [10] Dong Seon Cheng et.al proposes semantic knowledge coming from Wordnet, coupled with Google Ngram. There are Classes of different object recognition benchmarks (PASCAL VOC 2012, Caltech-256, ImageNet)and collects novel training images from the dataset. Web search by adding to the standard keyword denoting the class of objects to be detected some other related terms, Search more expressive. Terms to be added not to be arbitrarily chosen, but rather selected with a criterion that has to be explicit and meaningful. The advantage is that Datasets generalize well and are stable and not require manual intervention.The limitation is that Terms to be added not to be arbitrarily chosen and Selected with a criterion that has to be explicit and meaningful.

In [11] Heng Li et.al proposes Fuzzy-Object-Shape to capture the shape of the object of interest along with the degree of impreciseness in the boundary information. The Fuzzy- Object-Shape information is extracted from each object in an image. This information provides a measure of closeness of the object of interest with well-known shapes. For each object, the fuzzy membership values are calculated and considered as feature vector. A similarity measure is proposed for measuring the degree of closeness of objects present in both query and database images. The advantages of this method are that Images are segmented into non-overlapping clusters and treated as objects. The boundary of the object is ill-defined and there is impreciseness and vagueness present in the object information. The limitation is that there is no 3D fuzzy object extracted and no feature will be combined with the text based retrieval system for supporting a combined retrieval system.

III. PROPOSED METHODOLOGY

The objective of the proposed method is to identify and annotate with the labels/semantic of the selected image based the features (using shape) extracted and saved as feature vector. The dataset contains different images like Apple, Aeroplane, Ball, Orange etc. Fig 3.1 shows the schematic diagram of the working of our proposed medthod. Content-based image retrieval relies on the characterization of primitive features such as color, shape, and texture that can be automatically extracted from the images themselves. Queries to CBIR systems are most often expressed as visual exemplars of the type of image or image attribute being sought

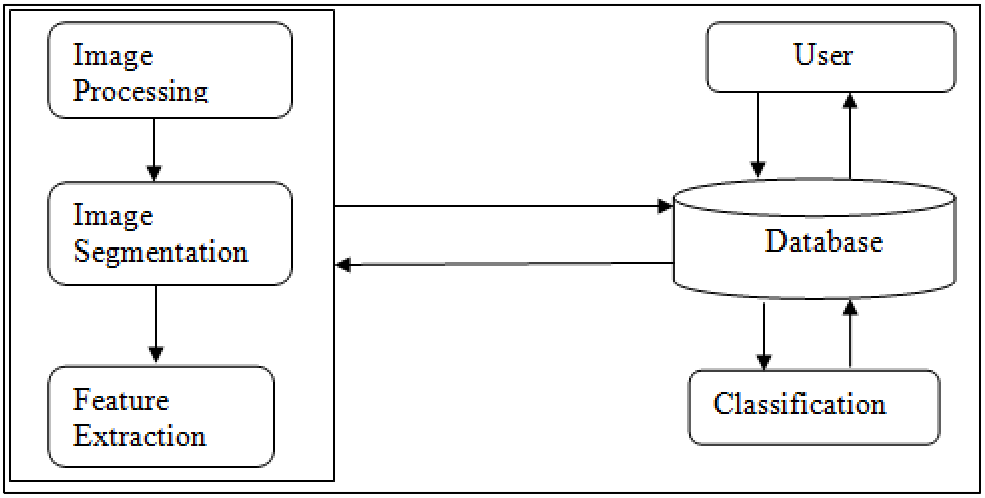

Fig. 1: Automatic Image Annotation

In the proposed system, the fuzzy based membership shape feature is calculated. The images in the database may be of different shape. So these objects should be segmented from the image using the shape feature. There are 7 shapes considered here. They are ellipse, circle, cone, cylinder, cone, rectangle, sphere.

In this paper, Mr and mR are the minimum and maximum length that bisects the diagonal in each of the primitive shapes. Mr and mR are considered as the maximum radius (Mr) and minimum radius (mR) of the object. Also, these values are second maximum (omitting the diagonal) and minimum distance between the center to edge of the object respectively. The minimum and maximum radius is measured from the center of the object.The center of the object is the centroid and from which the distance to the boundary is measured with certain angular displacement. Among the measured sample, minimum and maximum value is considered as minimum and maximum radius.

If Mr/mR->1, the object is a perfect circle shown in fig.2 (a)

Unlike circle,the ellipse has different lengths of diameter based on the orientation. The major axis is considered as Mr and the minor axis is considered as mR and two diagonal points in an ellipse are denoted as d1 and d2.The angle between Mr and mR is denoted as a1 and the angle between d1 and d2 is denoted as a2. While Mr is not equal to mR, d1=d2 and ∠a1 = ∠a2 ¼=90^0 , the object of interest is considered as a perfect ellipse(one) is depicted in fig.2(b) and for other values of d1,d2,a1and a2,the membership function value is below one.

It is known that if Mr = mR and d1 = d2, the shape of the object of interest is a perfect square and for other combination of values, the output value is less than one. The logical representation of square with minimum and maximum radius is depicted in Fig. 1(c). A four sided flat shape, where lines with all interior angles are right angles. The locations of the angles on the coordinate plane are determined by four vertices opposite angles which are parallel and congruent. Also,the diagonal bisects each other and they are congruent. Considering Fig. 2(a) and (c) for circle and square, both are perfect in their shape.In this case,the value of Mr and mR are the same for both the shape.If there is a distortion in the shape, the values of Mr and mR changes and the value of both is different. Thus, additional parameters are required to differentiate a perfect circle from a perfect square and as a result the d1 and d2 are introduced. For a perfect circle the value of Mr=mR=d1=d2 and for square Mr=mR and both d1 and d2 is different. This is the reason d1 and d2 are used to distinguish a perfect circle and square.

The logical representation of rectangle with minimum and maximum radius is depicted in Fig. 2(d). It is noticed from the above figure that mR~Mr, d1~d2 and angle between mR and Mr, which a1 is 90^0. However, a1 should not be equal to the angle between diagonals (d1 and d2), i.e. a1 not equal to a2.

A rhombus is a four sided shape with equal length on all sides. Also, opposite sides are parallel as well as opposite angles are equal. The diagonals d1 and d2 meet in the middle at right angles. The logical representation of a rhombus with minimum and maximum radius is shown in Fig. 2(e).

A cone is a shape that has a circle at the bottom and sides that narrow to a point. The pointy end of the cone is called vertex shown in fig.2(f).

A solid object with two identical flat ends that are circular or elliptical and one curved side is called as Cylinder. The view of the cylinder is presented in Fig. 2(g).

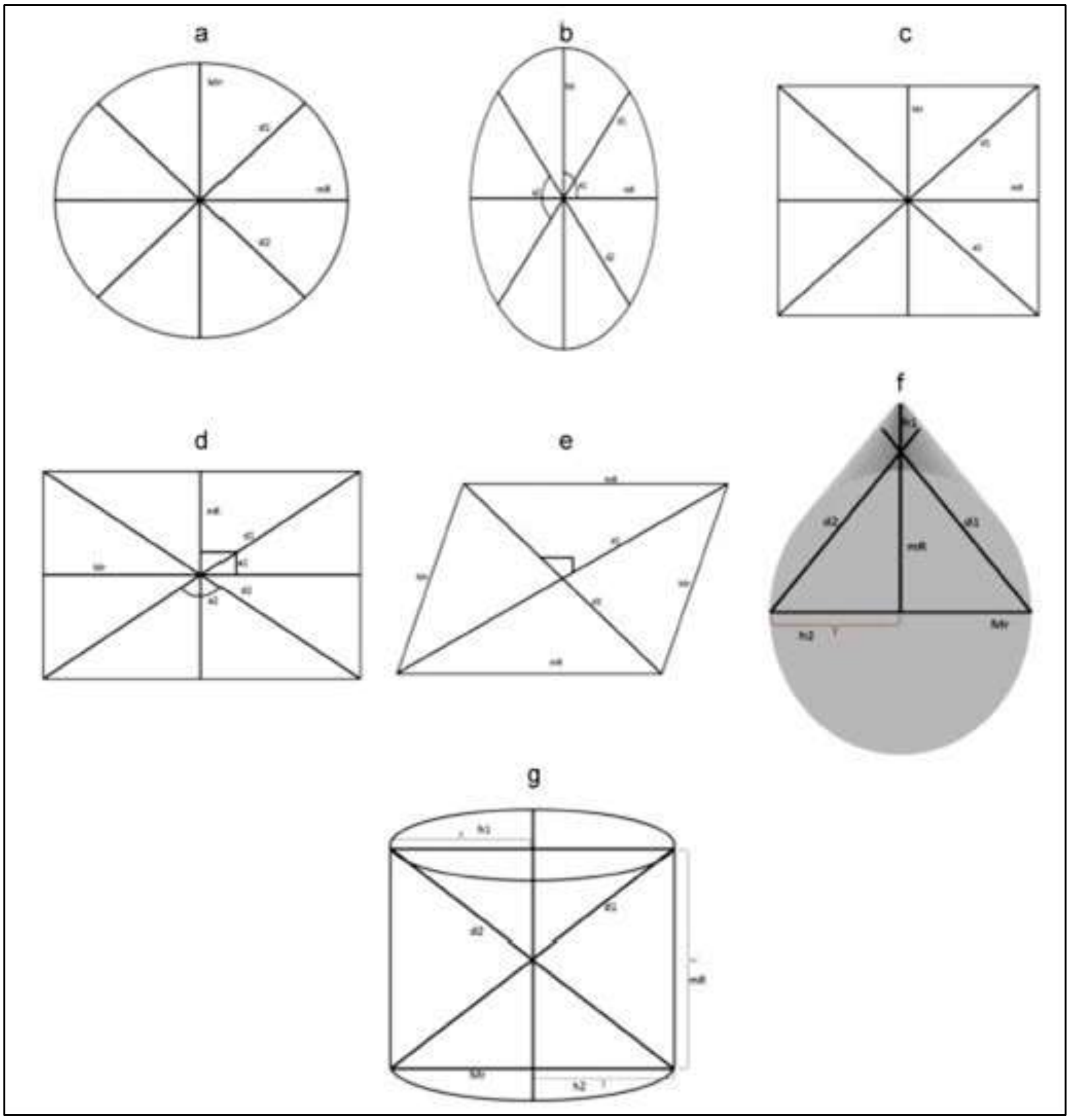

Fig. 2: The logical representation of various shapes using minimum and maximum radius: (a) logical representation of circle; (b) logical representation of ellipse; (c) logical representation of square; (d) logical representation of rectangle; (e) logical representation of rhombus; (f) logical representation of cone; and (g) logical representation of cylinder.

A. Pixel grouping and fuzzy-object level similarity measure:

In this paper, the procedure for obtaining the objects in images is depicted in Fig. 3. The pixel grouping method used in this work is very traditional. However, it is noticed that the performance of the method used here is encouraging. A similar traditional method, say, connected component detection is based on graph theory principle and the connected components are uniquely labeled for a given heuristic. Once the first pixel of a connected component is found, all the connected pixels of the connected component are labelled and then ext pixel is considered. Data structures such as linked list and a queue is required for processing the labelled pixels. In this work, we use canny edge detection along with morphological operations such as Dilation and Close with appropriate parameter using MATLAB tool. It is well-known that the canny edge detection algorithm is optimal edge detector and the error rate by canny operator is very low. It is important that edges occurring in images should not be missed and that response of non-edges should be very low with localization of edge points.

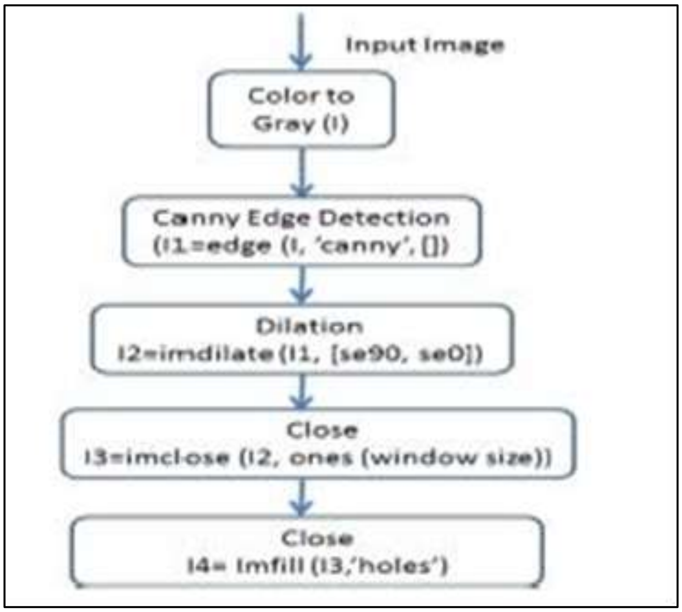

Fig. 2: The sequence of processes followed for obtaining object from images

The Canny operatoral ways tries to minimize the distance between the estimated edge pixels and the actual edge with one response to a single edge. However, it is very difficult to completely eliminate the possibility of multiple responses to an edge. This issue has been handled by the canny edge detector by smoothing the images to eliminate the noise and find the image gradient to highlight regions with high spatial derivatives. The algorithm then tracks along these regions and suppresses any pixel that is not at the maximum and the gradient array is now further reduced by hysteresis. The hysteresis isused to track along the remaining pixels that have not been suppressed and use two thresholds for edge detection. Because of the above mentioned procedural advantages and known as an optimal edge detector, we use canny edge detection algorithm. This has saved us time in implementing the feature extraction procedure and the performance is also encouraging. The input image is converted as gray image over which the canny edge detection algorithm is applied. The canny filter significantly reduces the amount of data in an image, while preserving the structural properties to be used for further processing.

We use MATLAB tool for carrying out both canny edge detection and morphological operations. We have chosen an edge function in MATLAB, which automatically choose low and high values of the threshold. This value is relative to the highest value of the gradient magnitude of the images to that the manual intervention for each and every image is avoided. The default sigma of the functionis sqrt(2). The output of the Canny shows lines of high contrast in the image.These lines do not adequately delineate the outline of the object of interest. While comparing with the original image, gaps in the lines surrounding the object in the gradient mask are predominant. These line argaps can be handled by dilating the linear structuring elements, which are created with the strel function. In this work,we hav eused se90=strel('line', 3, 90) and se0=strel('line', 3, 0). Here, se90 and se0 refers to the structuring element used in MATLAB. As a result, the binary gradient mask is dilated using the vertical structuring element followed by the horizontal structuring element. Finally, the close and fill functions are used for filling the hole to obtain a clear object.

IV. EXPERIMENTAL RESULTS

The method was implemented in a MATLAB 2014 prototype and tested with an image collected from the database.The feature extraction and the classification were performed on a desktop PC with the following characteristics:Intel Core i3 CPU, 3.4 GHz, 4 GB RAM.

The database was built with the images from the LabelMe dataset. During the feature extraction, the shapes like circle, ellipse, rectangle, square, rhombus,cone and cylinder shapes are extracted from the images to segment an image.The features are stored in the form of a feature vector shown in fig.4.

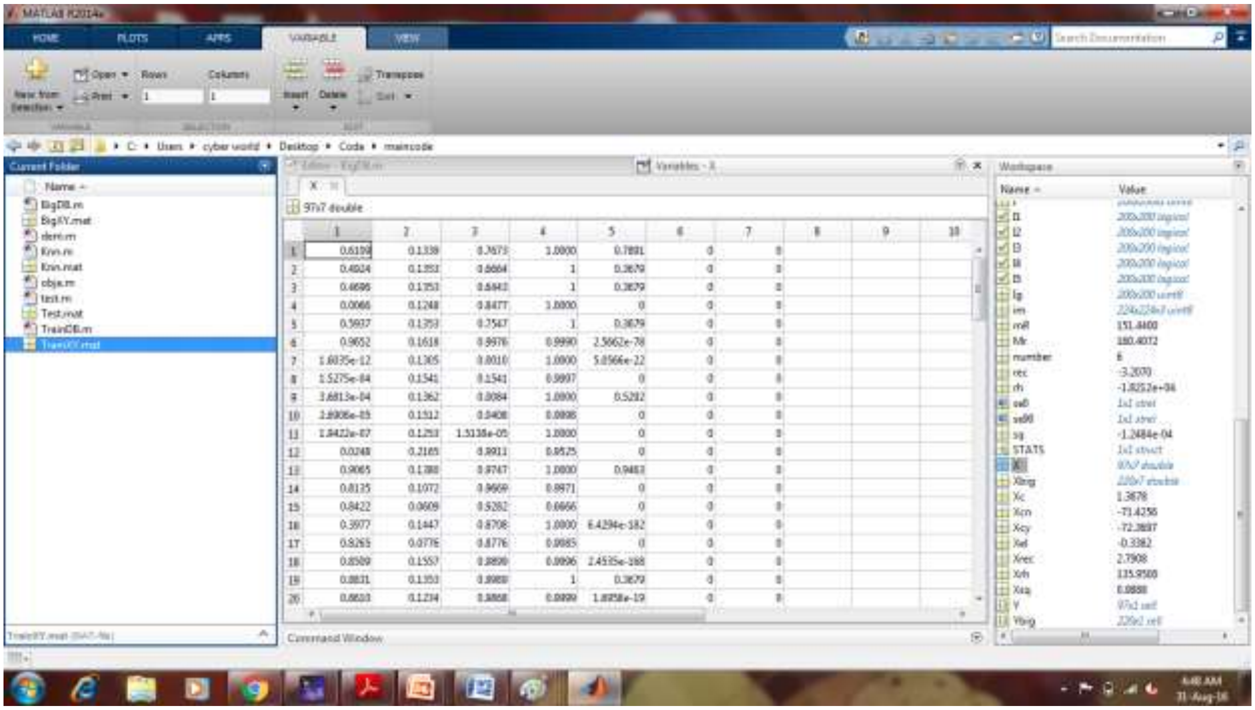

Fig. 4: shows the extracted feature vector

The proposed system uses the knn classifier to identify and annotate the object. The Knn model will predict the particular object by using the features that have been already trained earlier. After predicting the particular class by using the knn model, the final output will be of the label/text of the object have been identified successfully. In the generated knn model, it contain the features of trained database images .At the final stage, select a particular image from the database, the features of the selected object are extracted.Now the selected object's features and the features from the knn trained model are compared and predicted.The final output will be the label/text or recognize the particular object from the database and annoate it with its label nearby the object(Fig.5).

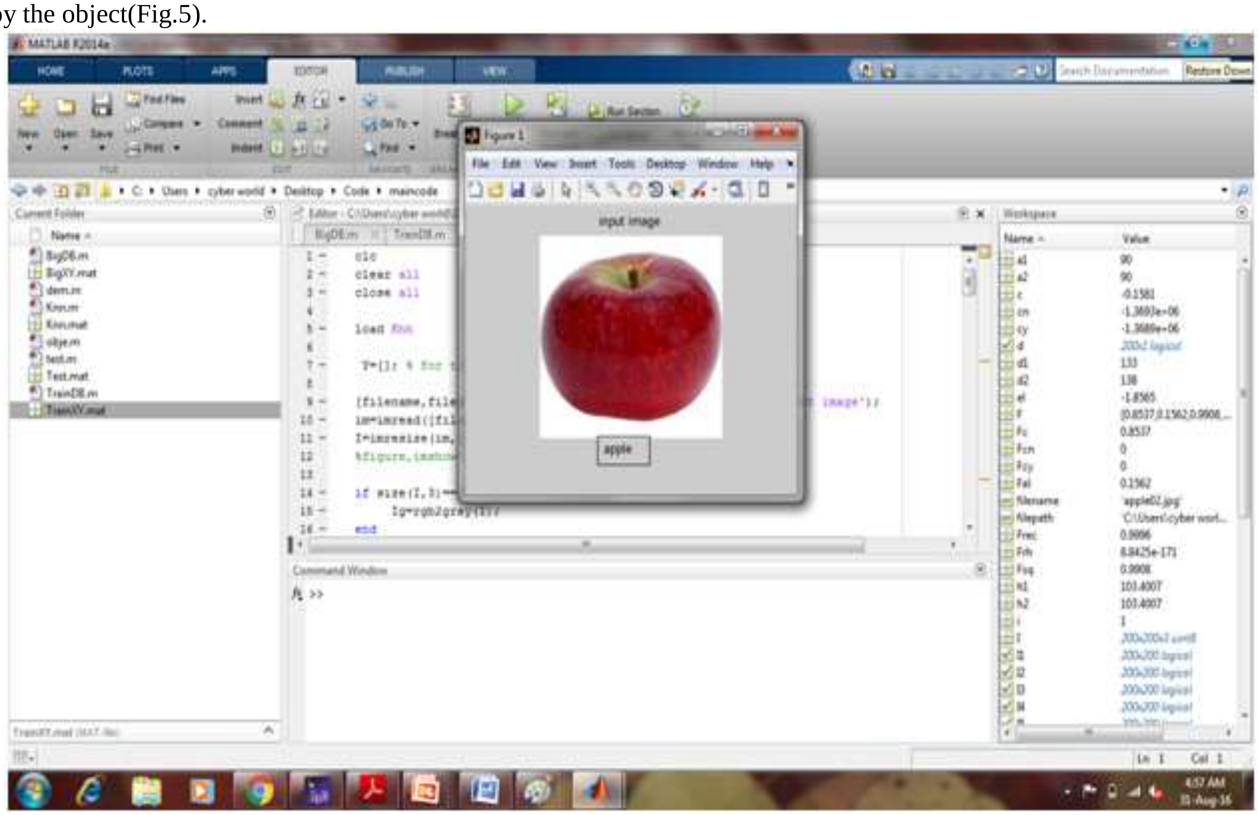

Fig. 5: Annoation of the selected object displayed

CONCLUSION

Human annotation is tedious and slow, and cannot cope with the huge volume of images generated by advanced image acquisition techniques, such as high content screening used in biological and medical research. There is a need to automate the process of annotating or indexing images in laboratories, at customs check posts, in art galleries or on the internet. In the proposed method the objects are segmented from the image using the Fuzzy membership shape function. There are 7 shapes considered in the proposed method. The proposed method can be extended for the multiple object annotation.

REFERENCES

[1] Pierre Andrews, Ilya Zaihrayeu, and Juan Pane, "A Classification of Semantic Annotation Systems", Int.J.Comput.(2011) pp.2231-2307.
[2] Pranay Tiwari, P.M.Kamde,"Automatic Image Annotation and Retrieval Using Contextual Information", International Research Journal of Engineering and Technology (IRJET) Volume: 02 Issue: 07, Oct-2015.
[3] Dong Seon Cheng,Francesco Setti, Nicola Zeni, Roberta Ferrario, Marco Cristani,"Semantically-driven automatic creation of training sets for object recognition", ELSEVIER, Computer Vision and Image Understanding 131 (2015) 5671.
[4] Bryan C. Russell, Antonio Torralba,"LabelMe: a database and web-based tool for image annotation" ELSEVIER, International Journal, vol. 28, no. 4, Apr.2006.
[5] Qingyong Li, Siwei Luoa, Zhongzhi Shi,"Fuzzy Aesthetic Semantics Description and Extraction for Art Image Retrieval", ELSEVIER(2012).
[6] Patrizio Frosini, Claudia Landi,"Persistent Betti numbers for a noise tolerant shape- based approach to image retrieval" ELSEVIER ,Pattern Recognition (2014)Vis. 77
[7] G. Castellano, Fanelli, Torsello "Shape annotation by semi-supervised fuzzy clustering", Int.J.of Computer Science and Information Security (IJCSIS), Vol. 8, No. 1, April 2014.
[8] Makoto Hasegawa, Salvatore Tabbone, "Amplitude-only log Radon transform for geometric invariant shape descriptor ", in IEEE International Conference on Computer Vision (ICCV), 2013.
[9] Zhenbao Liu, "Locality-constrained sparse patch coding for 3D shape retrieval ", ELSEVIER, Applied Soft Computing 14 (2014) 407-418.
[10] Qingyong Li, Siwei Luoa, Zhongzhi Shi, "Fuzzy Aesthetic Semantics Description and Extraction for Art Image Retrieval ", ELSEVIER, Expert Systems with Applications 41 (2014) 3276-3283.26
[11] Dong Seon Cheng,Francesco Setti, Nicola Zeni, Roberta Ferrario, Marco Cristani, "Semantically-driven automatic creation of training sets for object recognition", ELSEVIER, Computer Vision and Image Understanding 131 (2015) 5671.
[12] P.Shanmugavadivu, P.Sumathy, A.Vadivel, "FOSIR: Fuzzy-Object-Shape for Image Retrieval applications", ELSEVIER, Neurocomputing171(2016)719-735