

Image Classification based on Subset Feature set and Optimized by Local Hill climbing Method

Preeti Choudhary
Student
SOIT, RGPV,
Bhopal

Nishchol Mishra
Assistant Professor,
SOIT, RGPV,
Bhopal

ABSTRACT

Image classification is a very challenging and important problem in the image management and retrieval system. The traditional methods are not effective to the image classification due to the high dimensionality of the image feature space. This paper proposes a method of image classification over a given data set using subset feature set and morphological profile. On the basis of subset feature set the image data set are classified. The input is the image and the result is the class of images related to that image. Using this technique, the performance is found to be 84%, which is quite acceptable.

Keywords

Discrete cosine transform, Sub-part features, Morphology, Region growing etc.

1. INTRODUCTION

A large collection of digital images are being created in many areas of today's world. These collections include digitizing existing collection of analogue diagrams, drawing, photographs, paintings etc. Many images are available in the computer from photo collections, web pages, video databases and pictures got by digital camera etc. This vast collection of images has created a great requirement to develop image management and retrieval systems that helps the user in storing, indexing, browsing and retrieving images from the vast image database. The present content-based image retrieval system can't fulfill the user's information needs. When user want to retrieve the image from the image database, the user queries are typically based on semantics (e.g., show me a plantation image) and not on image low-level features (e.g., show me a green image). So retrieval will not be very efficient and effective, if the semantic concepts aren't identified in the image database. Image classifications enhance the performance of a content based image retrieval system. After grouping the image database into meaningful class, the image retrieval will be more efficient and effective. Our purpose is to show how image classification improves the performance of the CBIR system.

In this paper we propose a novel method for image classification over a given dataset. This classification method is based on the sub-part features set (like structural information or shape information). In this method firstly we calculate the sub-part feature set for the dataset after applying the DCT, morphological constraint set and region behind classification process. After that we find out the classified data by class labeling. Then classifier mode (subset feature nearer)generate the classified set on the basis of feature set and classified data. At last we apply optimization for checking

that the feature set to classified data and give the final classification result.

The rest of the paper is organized as follows; Section 2 reviews the related work. Section 3 describes the image features which are used in this method. The proposed algorithm is described in section 5. In section 6 we discuss the result. Finally we conclude in section 7.

2. RELATED WORK

There have been many image classification methods which classify the dataset on the basis of various other image features and also used different classifier for image classification. The image classification is important for the management and retrieval system. Some special image classification methods have been used in many applications such as Urban area image classification [1], Hierarchical image classification [2], Landscape image classification [3], Medical image classification ,Multimedia and satellite image database classification. The urban area image classification with parts-based feature and morphology profile was proposed in [1]. In this method the parts description features were extracted and joint with the morphology base spatial features. Then the joint features give to the SVM classifier. The advantages of this method are the representation of physical composition of the urban area. In [2] the entire dataset classified hierarchically in different classes using several discriminative features. The image dataset is classified with hierarchical semantics into day, night and sunrise/sunset images; close-up and non close-up; indoor and outdoor; city and landscape classes.

Several discriminative features are selected after comparing multiple low level features, such as color and texture features. The classification is performed through the combination of the selected feature by using SVM classifier. This classification scheme achieves a high accuracy. In [3] an efficient approach for landscape image classification based on color and shape descriptor was proposed.

3. PROPOSED ALGORITHM

In this work first we generate the sub-part feature set for image dataset then on the basis of these feature set and label data the classified set are generated. The flow chart shows the steps of proposed algorithm

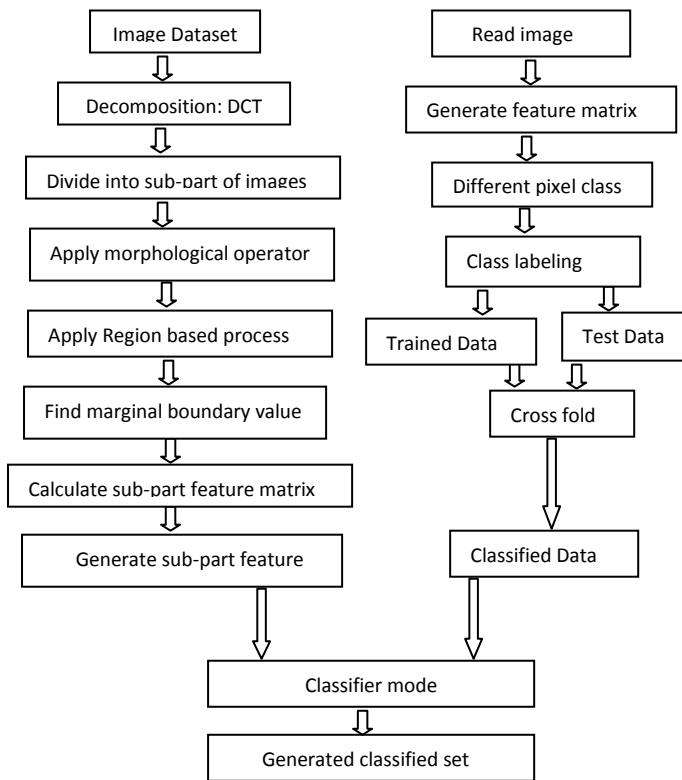


Figure1: Flow graph for image classification

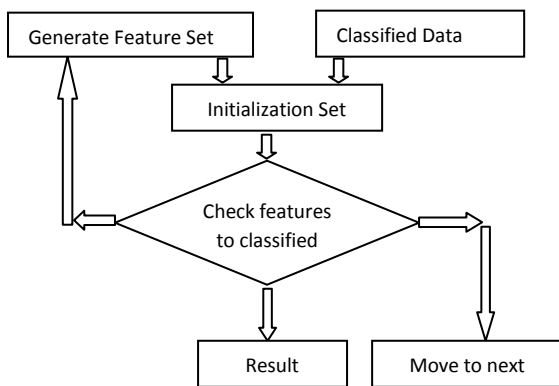


Figure2: Flow graph for optimization

4. PROPOSED ALGORITHM

Step 1

1.1 Firstly load image from the database. The image dataset consist of 100 images of 10 categories, each category containing 10 images. The categories are lemon, tomato, car, frog, juice cup, duck, flower, horse etc. The images are collected from the Internet for making the dataset.

1.2 Now we apply the discrete cosine transform (DCT) for block decomposition of images of given dataset. The discrete cosine transform (DCT) is a technique for converting a signal into elementary frequency components. It is widely used in image compression. We used the Two-Dimensional discrete Cosine Transform for decomposing the image into blocks or converting the given image into its elementary frequency component (Blocks). The one-dimensional DCT is useful in processing one-dimensional signals such as speech

waveforms. For analysis of two-dimensional (2D) signals such as images, we need a 2D version of the DCT.

1.3 In this step we apply the morphological constraint set on sub-part of image. The morphology gives the structural information about each block of input image and dataset images. The structural information describes the shape of block or sub part i.e. the block is rectangular or circular etc. Granulometries are popular and the powerful tools derived from the mathematical morphology (MM) theory. They are classically used for the analysis of the size distribution of particles in an image. The two fundamental operators in mathematical morphology are erosion and dilation. They are applied to an image with a set of known shapes, called the structuring elements (SEs).

1.4 Now we apply the region based process on the sub-parts of images after applying the morphology, so that the remaining blocks that were uncovered are included for feature extraction. Region based Segmentation is the art of partitioning an image into different regions where each one has some degree of uniformity in its feature space. It uses intrinsic image features, such as pixel intensity, color components and texture. It is an efficient alternative to partitioning schemes for image processing. Its main goal is to separate image into disjoint, nonoverlapping regions where each one has a uniformity in its predefined feature space.

1.5 Then we calculate the sub-part feature matrix.

1.6 In this step we find the marginal boundary value between two blocks.

1.7 In this step we generate the subset feature set like a_1, a_2, \dots, a_n .

Step 2

2.1 Firstly read the image for class labeling.

2.2 Then generate the feature matrix for image.

2.3 Calculate the different pixel classes.

2.4 On the basis of different pixel class the class labeling are being done.

2.5 In this step the trained data and test data are cross fold for validation of classified data.

Step 3

3.1 The generated subset feature set in step 1 given to the classifier mode.

3.2 The feature subset nearer classifier mode generate the classified set based on the sub-part feature set and label data. The most straightforward classifier in the machine learning techniques is the Nearest Neighbor Classifier – classification is achieved by identifying the nearest neighbors to a query example and using those neighbors to determine the class of the query. The Nearest Neighbor classification focusing on mechanisms for assessing similarity (distance), computational issues in identifying nearest neighbors and mechanisms for reducing the dimension of the data.

Step 4

In this step we apply optimization by using the local hill climbing method. This method checks features subset to classified data, If feature set and classified data match then give the final result and move to the next. Local hill climbing is a mathematical optimization technique which belongs to the family of local search. It is an iterative algorithm that starts with an arbitrary solution to a problem, then attempts to find a better solution by incrementally changing a single element of the solution. If the change produces a better solution, an

incremental change is made to the new solution, repeating until no further improvements can be found.

5. RESULTS AND DISCUSSION

To test and validate subset features effectiveness for image classification, the experiments are conducted on image dataset. The image dataset consist of 100 images of 10 categories, each category containing 10 images. The categories are lemon, tomato, car, frog, juice cup, duck, flower, horse etc. The platform used is MATLAB 7.8.0. The experiment shows that this technique enhances the performance of image classification. In figure 3 the input image is loaded for image classification. Figure 4 and figure 5 give two class of images related to that input image. The table 1 shows the comparison of result of classification of image by how many image similar or not similar.



Figure3: The input image

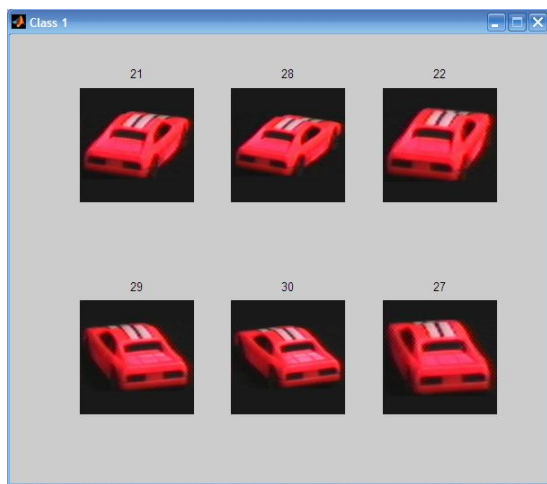


Figure4: classification result of class 1

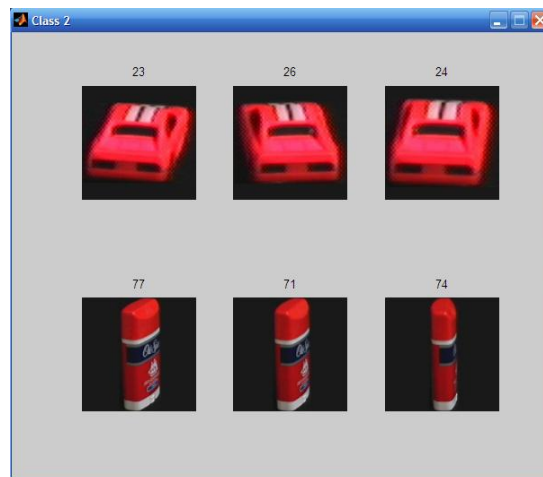


Figure5: classification result of class 2

Table 1: Classification Performance

Image	Similar	Non similar	Classification accuracy in %
Car	9	3	75
Frog	10	2	83.3
Duck	10	2	83.3
Juice cup	9	3	84
Tomato	10	2	83.3
Lemon	9	3	75

6. CONCLUSION AND FUTURE WORK

In image retrieval and management system the image classification is very challenging task. In this paper we proposed an approach which achieves acceptable accuracy in image classification. The subset features of images used for classification. The experimental results show that this approach is very effective in image retrieval and image management system. For future work we can apply this method on the urban area image classification by using SVM.

7. REFERENCES

- [1] HUANG Yuancheng, ZHANG Liangpei, LI Pingxiang, ZHONG , “High-resolution Hyper-spectral Image Classification with Parts-based Feature and Morphology Profile in Urban Area”, Springer volume 13, issue 2, june2010.
- [2] Yanni Wang, Bao-Gang Hu, “Hierarchical Image Classification Using Support Vector Machines”, ACCV2002: The 5th Asian Conference on Computer Vision, 23--25 January 2002.
- [3] A. Vailaya, A.K. Jain, and H.J. Zhang, “On Image Classification: City Image vs. Landscapes,” Pattern Recognition, Vol. 31, No. 12, pp.1921-1936, 1998.

- [4] M. Szummer and R.W. Picard, “Indoor-Outdoor Image Classification,” IEEE Intl Workshop on Content-based Access of Image and Video Databases, Jan 1998.
- [5] Andrew B. Watson, “Image Compression Using the Discrete Cosine Transform”, *Mathematica Journal*, 4(1), 1994, p. 81-88.
- [6] Saurabh Agrawal, Nishchal K Verma, Prateek Tamrakar, Pradip Sircar, “Content Based Color Image Classification using SVM”, 2011 Eighth International Conference on Information Technology.
- [7] Luc Vincent, “Morphological grayscale reconstruction in Image analysis: applications and efficient algorithms”, *IEEE Transactions on Image Processing*, 2(2): 176-201, 1993.
- [8] Philippe H. Gosselin, Matthieu Cord “A Comparison of Active Classification Methods for Content Based Image Retrieval” *ACM*, 2006.
- [9] Selim Aksoy, Robert M. Haralick “A Classification Framework for Content-Based Image Retrieval” *IEEE publication*, December 2002.
- [10] M. Dash, H. Liu, “Feature Selection for Classification”, *Intelligent Data Analysis* 1 (1997)131-15