KNN and Steerable Pyramid Based Content Based Image Retrieval Mechanism

Pritpal Kaur[1], Sukhvir Kaur[2]
[1]*Research Scholar, Department of Computer Engineering, CTIEMT, Jalandhar*
[2]*Assistant Professor, Department of Computer Engineering, CTIEMT, Jalandhar*

Abstract— Due to exponential increase of the size of the so-called multimedia files in recent years because of the substantial increase of affordable memory storage on one hand and the wide spread of the World Wide Web (www) on the other hand, the need for efficient tool to retrieve images from large dataset becomes crucial. This motivates the extensive research into image retrieval systems. In this paper, we present a hybrid approach of image classification using KNN and feature extraction using steerable pyramid based image retrieval system that uses color, contours and texture as visual features to describe the content of an image. We have initially used k-nearest neighbor image classification mechanism to fetch the appropriate images from the database image set using the query image. Further we have applied steerable pyramid to extract features from query image and candidate images retrieved from the knn. From the experimental results, it is evident that our system performs significantly better and faster compared with other existing systems.

Keywords— *Content based image retrieval (CBIR), color, shape, texture features, DWT, KNN, relevance feedback.*

I. INTRODUCTION

Content-based image retrieval (CBIR) has become an important research area in computer vision as digital image collections are rapidly being created and made available to multitudes of users through the World Wide Web. There are collections of images from art museums, medical institutes, and environmental agencies, to name a few. In the commercial sector, companies have been formed that are making large collections of photographic images of real-world scenes available to users who want them for illustrations in books, articles, advertisements, and other media meant for the public at large. Content-Based Image Retrieval (CBIR) systems are search engines for image databases, which index images according to their content. A typical task solved by CBIR systems is that a user submits a query image or series of images and the system is required to retrieve images from the database as similar as possible. Another task is a support for browsing through large image databases, where the images are supposed to be grouped or organized in accordance with similar properties. Although the image retrieval has been an active research area for many years this difficult problem is still far from being solved. There are two main reasons, the first is so called semantic gap, which is the difference between information that can be extracted from the visual data and the interpretation that the same data have for a user in a given situation. The other reason is called sensory gap, which is the difference between a real object and its computational representation derived from sensors, which measurements are significantly influenced by the acquisition conditions. In a typical CBIR system, image low level features like color, texture, shape and spatial locations are represented in the form of a multidimensional feature vector. The feature vectors of images in the database form a feature database. The retrieval process is initiated when a user queries the system using an example image or sketch of the object. The query image is converted into the internal representation of feature vector using the same feature extraction routine that was used for building the feature database. The similarity measure is employed to calculate the distance between the feature vectors of query image and those of the target images in the feature database. Finally, the retrieval is performed using an indexing scheme which facilitates the efficient searching of the image database. Recently, user's relevance feedback is also incorporated to further improve the retrieval process in order to produce perceptually and semantically more meaningful retrieval results. Content-based image retrieval research has produced a number of search engines. The commercial image providers, for the most part, are not using these techniques. The main reason is that most CBIR systems require an example image and then retrieve similar images from their databases. Real users do not have example images; they start with an idea, not an image. Some CBIR systems allows users to draw the sketch of the images wanted. Such systems require the users to have their objectives in mind first and therefore can only be applied in some specific domains, like trademark matching, and painting purchasing. Earlier CBIR systems rely on global image features, such as color histogram and texture statistics. Global features cannot capture object properties, so local features are favored for object class recognition. For the same reason, higher-level image features are preferred to lower-level ones. Similar image elements, like pixels, patches, and lines can be grouped together to form higher-level units, which are more likely to correspond to objects or object parts. Different types of features can be combined to improve the feature discriminability. For example, using color and texture to identify trees is more reliable than using color or texture alone. The context information is also helpful for detecting objects. A boat candidate region more likely corresponds to a boat if it is inside a blue region. While improving the ability of our system by designing higher-level image features and combining individual ones, we should be prepared to apply more and more features since a limited number of features cannot satisfying the requirement of recognizing many different objects in ordinary photographic images in the form of a multidimensional feature vector. The feature vectors of images in the database form a feature database. The retrieval process is initiated when a user query the system using an example image or sketch of the object. The query image is converted into the internal representation of feature vector using the same feature extraction routine that was used for building the feature database. The similarity measure is employed to calculate the distance between the feature vectors of query image and those of the target images in the feature database. Finally, the retrieval is performed using an indexing scheme which facilitates the efficient searching of the image database. The idea is that all features will be regions, each with its own set of attributes, but with a common representation. This uniform representation enables our system to handle multiple different feature types and to be extendable to new features at any time.

II. RELATED WORK

James Z. Wang et al. (2001) presented simplicity (Semantics sensitive Integrated Matching for Picture Libraries), an image retrieval system, which uses semantics classification methods, a wavelet-based approach for feature extraction, and integrated region matching based upon image segmentation. As in other region based retrieval systems, an image is represented by a set of regions, roughly corresponding to objects, which are characterized by color, texture, shape, and location. Yixin Chen et al. (2002) proposes a fuzzy logic approach, UFM (unified feature matching), for region-based image retrieval. In their retrieval system, an image is represented by a set of segmented regions, each of which is characterized by a fuzzy feature (fuzzy set) reflecting color, texture, and shape properties. Yixin Chen et al. (2005) introduces a new technique, cluster-based retrieval of images by unsupervised learning (CLUE), for improving user interaction with image retrieval systems by fully exploiting the similarity information. R. Fergus et al. (2005) developed a new model, TSI-pLSA, which extends pLSA (as applied to visual words) to include spatial information in a translation and scale invariant manner. Savvas A. Chatzichristofis et al. (2008) deals with a new low level feature that is extracted from the images and can be used for indexing and retrieval. This feature is called "Color and Edge Directivity Descriptor" and incorporates color and texture in formation in a histogram. CEDD size is limited to 54 bytes per image, rendering this descriptor suitable for use in large image databases. Chuen-Horng Lin et al. (2008) proposes three feature vectors for image retrieval. In addition, a feature selection technique is also brought forward to select optimal features to not only maximize the detection rate but also simplify the computation of image retrieval. The first and second image features are based on color and texture features, respectively called color co-occurrence matrix (CCM) and difference between pixels of scan pattern (DBPSP) in this research work. Michal Perdoch et al. (2009) proposes a novel method for learning discretized local geometry representation based on minimization of average reprojection error in the space of ellipses. Herve Jegou et al. (2010) addresses the problem of image search on a very large scale, where three constraints have to be considered jointly the accuracy of the search, its efficiency, and the memory usage of the representation. Amandeep Khokher et al. (2012) has stated that expanding image collections on the Internet have attracted significant research efforts in providing tools for effective retrieval and management of visual data. The need to find a desired image from a large collection is shared by many professional groups, including journalists, design engineers and art historians. Difficulties faced by text-based image retrieval brought the researchers to develop new solutions to represent and index visual information. Yanzhi Chen et al. (2012) proposed a discriminative criterion for improving result quality. This criterion lends itself to the addition of extra query data, and they showed that multiple query images can be comb index to produce enhanced results. Experiments compare the performance of the method to state-of-the-art in object retrieval, and show how performance is lifted by the inclusion of further query images.Lakhdar et al. (2016) presents an efficient region based image retrieval method, which uses multi-features color, texture and edge descriptors. In contrast to recent image retrieval methods, which use discrete wavelet transform (DWT), they proposed using shape adaptive discrete wavelet transform (SA-DWT). The advantage of this method is that the number of coefficients after transformation is identical to the number of pixels in the original region. Abbas H et al. (2016) studied the influence on performance of reducing the colors number contained in images. Accomplishing this task poses an extra overhead on the system, which requires more computation time, but, on the other hand, can accelerate the comparison process. Aman Saini et al. (2016) evaluates the performance of CBIR system with parameters precision, recall, and NMRR and retrieval time by using feature extraction techniques based upon colour - histogram, texture - GLCM and region –boundary descriptors of an image. NMRR and retrieval time of different query images have been evaluated.

III. KNN

An instance based learning method called the K-Nearest Neighbor or K-NN algorithm has been used in many applications in areas such as data mining, statistical pattern recognition, image processing. Successful applications include recognition of handwriting, satellite image and EKG pattern. In data mining, we often need to compare samples to see how similar they are to each other. For samples whose features have continuous values, it is customary to consider samples to be similar to each other if the distances between them are small. Other than the most popular choice of Euclidean distance, there are of course many other ways to define distance. The k-means clustering algorithm attempts to split a given anonymous data set (a set containing no information as to class identity) into a fixed number (k) of clusters. Initially k number of so called centroids are chosen. A centroid is a data point (imaginary or real) at the center of a cluster. In context each centroid is an existing data point in the given input data set, picked at random, such that all centroids are unique (that is, for all centroids ci and cj, ci ≠ cj). These centroids are used to train a KNN classifier. The resulting classifier is used to classify (using k = 1) the data and thereby produce an initial randomized set of clusters. Each centroid is thereafter set to the arithmetic mean of the cluster it defines. The process of classification and centroid adjustment is repeated until the values of the centroids stabilize. The final centroids will be used to produce the final classification/clustering of the input data, effectively turning the set of initially anonymous data points into a set of data points, each with a class identity. CI is a categorical variable, and there is a scalar function, f, which assigns a class, $y = f(x)$ to every such vectors. We do not know anything about f (otherwise there is no need for data mining) except that we assume that it is smooth in some sense. We suppose that a set of T such vectors are given together with their corresponding classes.

IV. STEERABLE PYRAMID

The Steerable Pyramid is a linear multi-scale, multi-orientation image decomposition that provides a useful front-end for image-processing and computer vision applications. We developed this representation in 1990, in order to overcome the limitations of orthogonal separable wavelet decompositions that were then becoming popular for image processing (specifically, those representations are heavily aliased, and do not represent oblique orientations well). Once the orthogonality constraint is dropped, it makes sense to completely reconsider the filter design problem (as opposed to just re-using orthogonal wavelet filters in a redundant representation, as is done in cycle-spinning or undecimated wavelet transforms!).

The basic functions of the steerable pyramid are Kth-order directional derivative operators (for any choice of K), that come in different sizes and K+1 orientations. As directional derivatives, they span a rotation-invariant subspace, and they are designed and sampled such that the whole transform forms a tight frame. An example decomposition of an image of a white disk on a black background is shown to the right. This particular steerable pyramid contains 4 orientation sub bands, at 2 scales. The smallest sub band is the residual low pass information. The residual high pass sub band is not shown. The block diagram for the decomposition (both analysis and synthesis) is shown to the right. Initially, the image is separated into low and high pass sub bands, using filters L0 and H0.

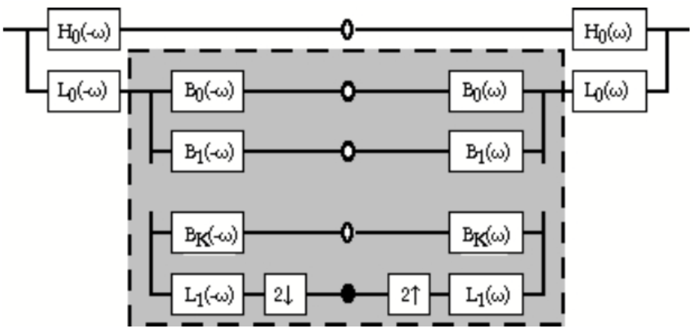

Figure 1. Decomposition using Steerable Pyramids

The low pass sub band is then divided into a set of oriented band pass sub bands and a lower-pass sub band. This lower pass sub band is subsampled by a factor of 2 in the X and Y directions. The recursive (pyramid) construction of a pyramid is achieved by inserting a copy of the shaded portion of the diagram at the location of the solid circle (i.e., the low pass branch). The right side of the diagram is the synthesis part. The synthesized image is reconstructed by up sampling the lower low-pass sub band by the factor of 2 and adding up with the set of band-pass sub bands and the high-pass sub band.

V. RESEARCH METHODOLOGY

- First of all we have multiple query images in the query folder and rest of the images are database images
- Extract the low level features like color, shape and texture of all database as well as query images.
- Match all the database images with the query images by using KNN. Then the result will comes as classified images and unclassified images. Classified images are stored in labeled folder and unclassified images are stored in unlabeled folder.
- Apply Steerable Pyramid on classified images
- Then Match the classified images with query images and display output.

VI. EXPERIMENTAL SETUP

WANG Database contains 1000 images which can be classified into different domains namely Buses, Dinosaurs, Flowers, Building, Elephants, Mountains, Food, African people, Beaches and Horses with JPEG format which used in a general purpose for experimentation. These images are stored with size 256×256 and each image is represented with RGB color space.

Figure 2. WANG dataset images

VII. OPENCV

The open source computer vision library, OpenCV, began as a research project at Intel in 1998. It has been available since 2000 under the BSD open source license. OpenCV is aimed at providing the tools needed to solve computer-vision problems. It contains a mix of low-level image-processing functions and high-level algorithms such as face detection, pedestrian detection, feature matching, and tracking.

VIII. EXPERIMENTAL RESULTS

Multiple number of experiments have been conducted on different categories of images from the WANG dataset. The different categories like buses, buildings, flowers, elephants, mountains, dinosaurs, beaches, human and horses have been used for testing and analysis. A retrieved image is considered to be correct if and only if it is in the same category as the query. The experiments are carried out in a personal computer with Intel Core i5 processor with 8GB RAM. The program is developed using OpenCV libraries and Visual Studio IDE.

Table 1. Results evaluated after testing

Category Name	Relevant After Feedback	Irrelevant After Feedback	Total Retrieved Images After KNN
Human	18	18	9
Beech	29	26	14
Building	25	22	10

IX. PERFORMANCE EVALUATION

The performance of a retrieval system is evaluated based on several criteria. The precision of the retrieval is defined as the fraction of the retrieved images that are indeed relevant for the query. A good retrieval system should have high values for precision. Different set of images are taken in each experiment. Table 1 illustrates the results retrieved by running the multiple number of experiments. All the categories have been taken for testing purposes.

Table 2. Precision Value

Total Retrieved Images After Steerable Pyramid	Precision
9	0.7
14	0.7
10	0.9

In the table 2, we have calculated the precision for the same set of images that are used in table 1. The overall execution time is computed in milliseconds and has been optimized.

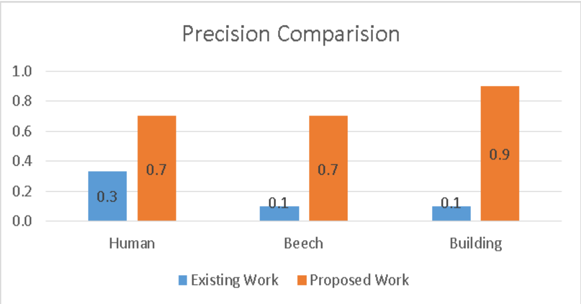

Figure 3. Precision Comparison

Figure 3 depicts the precision comparison of the existing work and the proposed work. There is lot of improvement in the precision in the present work. Improvement in precision means CBIR system is retrieving more number of relevant images.

X. CONCLUSION

Visual feature such as color, texture and contour are extracted using KNN and steerable pyramid. Features are extracted on both whole image level and database image level to better capture salient object descriptions. To negotiate the gap between low-level visual features and high-level concepts, median vector mechanism is applied and integrated with these content-based retrieval techniques in a vector space model. Experiments show that combining the color, texture and feature vector achieves the best performance in the comparison of various approaches. Finally, since it is obvious that neither single color feature nor textual features are sufficient to capture the overall contents of visual data, we propose a seamless integration of all the feature vectors such as color, texture and contour, taking advantage of using our vector space model. The combined feature vector, on which latent semantic indexing will be performed afterwards, is normalized and weighted. Preliminary results reveal that it is a very promising approach to further bridging the semantic gap and achieving better retrieval performance. After evaluating the results, we have reached up to the solution that we have been able to improve the CBIR mechanism using the proposed mechanism in this work. We will further test and benchmark this integrated image retrieval framework over various large image databases, along with tuning the relevance feedback to achieve optimal performance with highly reduced dimensionality.

XI. REFERENCES

[1] M. Fakheri, T. Sedghi, M. G. Shayesteh1 and M. C. Amirani, "Framework for image retrieval using machine learning and statistical similarity matching techniques," IET Image Process, pp. 1-11, 2013.
[2] P. MANIPOONCHELVI and K. MUNEESWARAN, "Multi region based image retrieval system," Indian Academy of Sciences, pp. 333-344, 2014.
[3] H. J´egou, M. Douze, C. Schmid and P. P´erez, "Aggregating local descriptors into a compact image representation," IEEE, pp. 3304-3311, 2010.
[4] Y. Chen, J. Z. Wang and R. Krovetz, "CLUE: Cluster-Based Retrieval of Images by Unsupervised Learning," IEEE, pp. 1187-1201, 2005.
[5] R. Fergus, L. Fei-Fei, P. Perona and A. Zisserman, "Learning Object Categories from Google's Image Search," IEEE, 2005.
[6] Y. Chen, X. Li, A. Dick and A. v. d. Hengel, "Boosting Object Retrieval with Group Queries," IEEE, pp. 765-768, 2012.
[7] R. Arandjelovi´c and A. Zisserman, "Three things everyone should know to improve object retrieval," IEEE, pp. 2911-2918, 2012.
[8] M. Perd'och, Chum and J. Matas, IEEE, pp. 9-16, 2009.
[9] S. A. Chatzichristofis and Y. S. Boutalis, "CEDD: Color and Edge Directivity Descriptor: A Compact Descriptor for Image Indexing and Retrieval," Springer-Verlag Berlin Heidelberg, pp. 313-322, 2008.
[10] Y. Chen and J. Z. Wang, "A Region-Based Fuzzy Feature Matching Approach to Content-Based Image Retrieval," IEEE, pp. 1252-1267, 2002.
[11] C.-H. Lin, R.-T. Chen and Y.-K. Chan, "A smart content-based image retrieval system based on color and texture feature," ELSEVIER, p. 658–665, 2009.
[12] Z. Wang, J. Li and G. Wiederhold, "SIMPLIcity: Semantics-Sensitive Integrated Matching for Picture LIbraries," IEEE, pp. 947-963, 2001.
[13] S. Gandhani, R. Bhujade and A. Sinhal, "AN IMPROVED AND EFFICIENT IMPLEMENTATION OF CBIR SYSTEM BASED ON COMBINED FEATURES," IET, pp. 353-359.
[14] S. M. H. Khan, A. Hussain and I. F. T. Alshaikhl, "Comparative study on Content Based Image Retrieval (CBIR)," IEEE, pp. 61-66, 2013.
[15] Sreedevi S. and Shinto Sebastian, "Fast Image Retrieval with Feature Levels," IEEE, 2013.

[16] S. Ezekiel, Mark G. Alford, David Ferris and Eric Jones,, "Multi-Scale Decomposition Tool for Content Based Image Retrieval," IEEE, 2013.
[17] K. Juneja, A. Verma , S. Goel and S. Goel , "A Survey on Recent Image Indexing and Retrieval Techniques for Low-level Feature Extraction in CBIR systems," IEEE, pp. 67-72, 2015.
[18] K. BELATTAR and S. MOSTEFAI, "CBIR using Relevance Feedback: Comparative Analysis and Major Challenges," IEEE, pp. 317-325, 2013.
[19] D. Jeyabharathi and A. Suruliandi, "Performance Analysis of Feature Extraction and Classification Techniques in CBIR," IEEE, pp. 1211-1214, 2013.
[20] H. Xie, Y. Ji and Y. Lu, "An Analogy-Relevance Feedback CBIR Method Using Multiple Features," IEEE, pp. 83-86, 2013.
[21] B. Kaur and S. Jindal, "An implementation of Feature Extraction over medical images on OPEN CV Environment".
[22] S. Kumar, S. Jain and T. Zaveri, "ARALLEL APPROACH TO EXPEDITE MORPHOLOGICAL FEATURE EXTRACTION OF REMOTE SENSING IMAGES FOR CBIR SYSTEM," IEEE, pp. 2471-2474, 2014.
[23] K. BELATTAR and S. MOSTEFAI, "CBIR with RF: which Technique for which Image," IEEE, 2013.
[24] G. Raghuwanshi and V. Tyagi, "Texture image retrieval using adaptive tetrolet transforms," ELSEVIER, pp. 1-8, 2015.