APPLICATIONS OF SPATIAL FEATURES IN CBIR : A SURVEY

R. Priyatharshini[1] Surendernath S.P [2] Dr.S.Chitrakala [3]

[1]Easwari Engineering College, [2][3]Anna University, Chennai
{priya.sneham,spsuren3}@gmail.com,{chitras@cs.annauniv.edu}

ABSTRACT

With advances in the computer technology and the World Wide Web there has been an explosion in the amount and complexity of multimedia data that are generated, stored, transmitted, analyzed, and accessed. In order to extract useful information from this huge amount of data, many content based image retrieval (CBIR) systems have been developed in the last decade. A typical CBIR system captures image features that represent image properties such as color, texture, or shape of objects in the query image and try to retrieve images from the database with similar features. Retrieval efficiency and accuracy are the important issues in designing Content Based Image Retrieval System. The Shape and Spatial features are quiet easy and simple to derive and effective. Researchers are moving towards finding spatial features and the scope of implementing these features in to the image retrieval framework for reducing the semantic gap. This Survey paper focuses on the detailed review of different methods and their evaluation techniques used in the recent works based on spatial features in CBIR systems. Finally, several recommendations for future research directions have been suggested based on the recent technologies.

KEYWORDS

Content Based Image Retrieval, Semantic Gap, Semantic Image Retrieval, Image Retrieval based on spatial features.

1. INTRODUCTION

The dramatic growth of digital media at home, in enterprises, and on the web, has over the last decade spawned great interests in developing methods for effective indexing and searching of desired visual contents to open the value of these contents. CBIR is the sub problem of content based retrieval (CBR) it is necessary to develop power tools for retrieving images from the web where the number and size of digital image is growing fast. In addition, CBIR is also the key technology for improving the interface between user and computer. CBIR is a technique used for retrieving similar images from image database. It uses various features of image to matching the query image to database images. Content Based Image Retrieval (CBIR) is a technique which uses visual contents, normally called as features, such as shape, color, texture, edge, etc. to search images from large scale image databases according to users' requests in the form of a query image. Content based retrieval of visual data requires a paradigm that differs significantly from

both traditional databases and text based image understanding systems. The challenge in CBIR is to develop the methods that will increase the retrieval accuracy and reduce the retrieval time. Content based image retrieval system consists of following modules:

1. **Feature Extraction:** In this module the features of interest are calculated for image database. This step is done in offline manner.
2. **Feature extraction of query image:** This module calculates the feature of the query image. Query image can be a part of image database or it may not be a part of image database.
3. **Similarity measure:** This module compares the feature database of the existing images with the query image on basis of the similarity measure of the interest.
4. **Retrieval and Result:** This module will display the matching images to the user.

In order to improve the effectiveness of CBIR systems we need to discover semantically meaningful patterns. Semantic concept detection is a research topic of great interest as it provides semantic filters to help analysis and search of multimedia data. It is essentially a classification task that determines whether an image or a video shot is relevant to a given semantic concept. The semantic concepts cover a wide range of topics such as those related to objects, indoor/outdoor scenes, events, etc. Automatically detecting these concepts is challenging especially in the presence of within-class variation, occlusion, background clutter, pose and lighting changes in images and video shots. Global features are known to be limited in face of these difficulties, which stimulated the development of local invariant features (key points) in recent years. Key points are salient patches that contain rich local information about an image. The most popular key point based representation is bag-of-visual-words (BoW). In BoW, a visual vocabulary is generated through grouping similar key points into a large number of clusters and treating each cluster as a visual word. By mapping the key points of an image back into visual words of the vocabulary, we can represent the image as a histogram of visual words and use it as the feature for classification.

1.1 Feature Extraction

Mapping the image pixels into the feature space is known as feature extraction. Extracted features are used to represent images for searching, indexing and browsing images in an image database. Feature extraction is a means of extracting compact but semantically valuable information from images. This information is used as a signature for the image. Similar images should have similar signatures. Feature extraction of the image in the database is typically conducted off-line so computation complexity is not a significant issue. This section introduces spatial features: texture, shape, and color, size and Edge density which are used most often to extract the features of an image. We have surveyed some already existing systems to understand the various feature extraction techniques.

In boundary detection method, the binary image is scanned until a boundary is found. A color averaging technique is used for feature extraction in the spatial domain to reduce the size of the feature vector. In Gradient and 12 Directional Feature Extraction methods first they compute the gradient of the handwritten character by applying Sobel's Mask. After computing the gradient of each pixel of the character, they map these gradient values onto 12 direction values with angle span of 30 degree between any two adjacent direction values .Once they get the direction features, they feed them into a neural network system. Principal Component Analysis method for feature extraction is used for performing dimensionality reduction. Slope Magnitude technique is

used for Extraction of Shape Features. Fast Fourier Transform, Walsh Transform, DCT and DST transform is also used for feature Extraction to design the sectors required for the search and retrieval of images from a database.

1.2 Feature Selection

It is generally believed that a better image recognition effect can be achieved with more features descriptors used, but this is not absolutely true. Not all features are helpful for image recognition. Irrelevant features are actually interfering signals and cause drops in the recognition rate, especially if the effect of the irrelevant features exceeds that of the effective ones. Therefore, these features must be dropped. The goal of feature selection technique is to select the best features from the original ones. It can not only achieve maximum recognition rate but can also simplify the calculation of the image retrieval process. Working with extremely high dimensional data, as given in case of a spectral data set, may result in high computational costs. Furthermore we have to face the so-called "curse of dimensionality", as the number of samples in relation to the number of features is very low. Therefore, researchers perform dimensionality reduction or feature selection where either a subset or a transform of the original features is chosen for the clustering process.

1.3 Similarity Measurements

The CBIR process is used for general applications like recognizing patterns and matching fingerprints or other biometric data. Basically we need to compare two images and check whether they are similar or they match or not. To compute this, it is required to have certain techniques by which one can statistically evaluate if two Images are similar or not. Once we extract a good set of features, we compare the extracted feature for similarity. It is believed that if good sets of features are extracted, the similarity between 2 images is given by how close the extracted features are of the two images. There are various kinds of similarity measurement techniques: the Euclidean Distance, Mean Square Error and Sum of Absolute Differences are common for CBIR applications.

2. RELATED WORKS

Semantic concept detection aims to annotate images or video shots with respect to a semantic concept. In existing Works, this task is often conducted in a diverse setting where the emphasis usually includes feature selection, multimodality fusion, and machine learning on huge multimedia data sets. Here we focus our review on feature-level analysis which is related to our latter experimental comparison. In [1] Edge Histogram Descriptor (EHD) captures the spatial distribution of edges like the color layout descriptor does by capturing colors. It is useful in image matching when the texture is not homogeneous. A novel image feature representation method, namely color difference histograms (CDH) has been proposed in [2] for image retrieval. This method is entirely different from the existing histograms; most of the existing histogram techniques merely count the number of frequency of pixels.

Fig.2. An example of image retrieval using the CDH algorithm on the Corel 10Kdataset.Thequery is an image of fruit, and all the returned images are correctly retrieved and ranked within the top12images.(The top left image is the query image, and the similar images returned include the query image itself) as in [2].

In [3] three approaches to utilizing object-level spatial contextual information for semantic image analysis are presented and comparatively evaluated. Contextual information is in the form of fuzzy directional relations between image regions. All techniques, namely a Genetic Algorithm (GA), a Binary Integer Programming (BIP) and an Energy-Based Model (EBM), are applied in order to estimate an optimal semantic image interpretation, after an initial set of region classification results is computed using solely visual features.
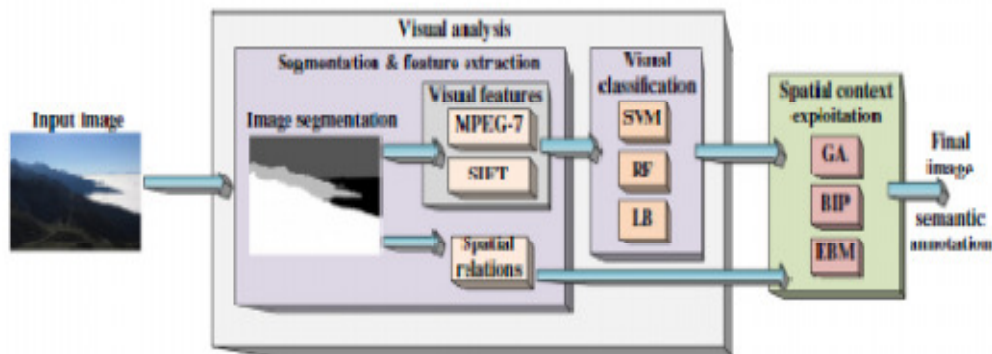

Fig.3. Developed Evaluation Framework as in [3]

In [4] an effective descriptor based on Angular Radial Transform (ART) and Polar Hough Transform (PHT), which is capable of fulfilling the above requirements. ART is used as a region based shape descriptor, which represents the global aspects of an image is proposed. PHT is used as a local shape descriptor for detecting linear edges in an edge image. The detected linear edges represent the association among adjacent edge points. The perpendicular distance of each linear edge to the centroid of the edge image is computed to build histograms, which exhibit rotation, scale and translation invariant properties.

A new image indexing and retrieval system for CBIR is proposed in [5]. The characteristics of image are computed using color histogram and spatial orientation tree (SOT). The SOT defines the spatial parent–child relationship among wavelet coefficients in multi-resolution wavelet sub-

bands. First the image is divided into sub-blocks and then constructed the SOT for each low pass wavelet coefficient is considered as a vector point of that particular image. Similarly the color histogram features are collected from the each sub-block. The vector points of each image are indexed using vocabulary tree. Vector points of each image are indexed using vocabulary tree. A specific CBIR system for hyper spectral images exploiting its rich spectral information is proposed on [7]. The CBIR image features are the end member signatures obtained from the image data by end member induction algorithms (EIAs). End members correspond to the elementary materials in the scene, so that the pixel spectra can be decomposed into a linear combination of end member signatures.
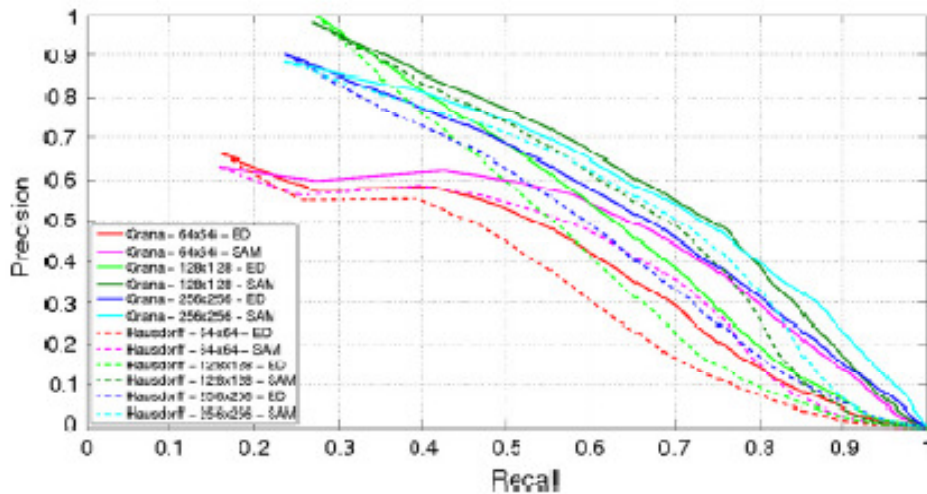

Fig.4.Precision and recall results for the10-Edataset using the EIHA end member induction algorithm for the Granadistance and the Hausdorff distances in [6]

3. OVERVIEW OF PROPOSED METHOD

The spatial locations are the key points in an image carry important information for classifying the image. For example, an image showing a beach scene typically consists of sky-like key points on the top and sands-like key points at the bottom. To integrate the spatial information, we follow to partition an image into equalized rectangular regions, compute the visual-word feature from each region, and concatenate the features of these regions into a single feature vector. There can be many ways of partitioning, e.g., 3x3 means cutting an image into 9 regions. This region-based representation has its downside in terms of cost and generalizability. First, if we divide each image into M x N regions, and compute a K-dimensional feature on each region, the concatenated feature vector is of K x M x N dimension, which can be prohibitively expensive to deal with. Besides, encoding spatial information can make the representation less generalizable. Suppose an image class is defined by the presence of a certain object, say, airplane, which may appear anywhere in an image. Using region-based representation can cause a feature mismatch if the objects in the training images are in different regions from those in the testing images. Many objects may cross region boundaries. These considerations prefer relatively coarse partitions of image regions to fine-grained partitions.

4. EXPERIMENTAL RESULTS AND DISCUSSION

Combining the important sub features of color, texture and shape improves the efficiency of image retrieval .PCA is used for feature selection which results in reduced set of features which is the set of selected features. These reduced features can be used for image retrieval and give the better performance[1].Experimental results demonstrate that CDH is much more efficient than the existing image feature descriptors that were originally developed for content-based image retrieval, such as MPEG-7 edge histogram descriptors, color auto correlograms and multi-texton histograms. It has a strong discriminative power using the color, texture and shape features while accounting for spatial layout [2]. Spatial context is efficient in improving the initial region-concept association results exhibiting an overall increase of up to 9.25% in the current evaluation framework by exploiting spatial context in semantic image analysis. The highest on average performance is achieved when complex partial constraints are acquired and their weight against the visual and co-occurrence information is efficiently adjusted [3].The combination of both global (ART based) and local (PHT based) features yields improved retrieval accuracy, which is analyzed over various sorts of image databases. An extensive set of experiments witness the superiority of the proposed hybrid system over other major contour based, region based and hybrid approaches in [4].The importance of spatial information can be seen by comparing the classification performance between the plain visual-word features (BoW) and the region-based ones. four ways of partitioning images, including 1x1(whole image), 2x2 (4 regions), 3x3 (9 regions), and 4x4 (16 regions) is used. Figure 6 shows the performance using different spatial partitions on various key point representation techniques such as vocabulary sizes, and weighting schemes. We see that the 2x2 partition substantially improves the classification performance. As the partition changes from 2x2 to 4 x 4, the MinfAP drops for most of the vocabulary sizes. When investigating the per-concept performance, we find that spatial information is more useful for classifying scenes than for classifying objects, since the former usually occupy a whole key frame, while the latter can appear anywhere in a key frame. For large scale semantic detection in diversified data set, using 2x2 partitions might be enough. The spatial information always helps as we have investigated various representation choices in BoW feature for semantic concept detection.
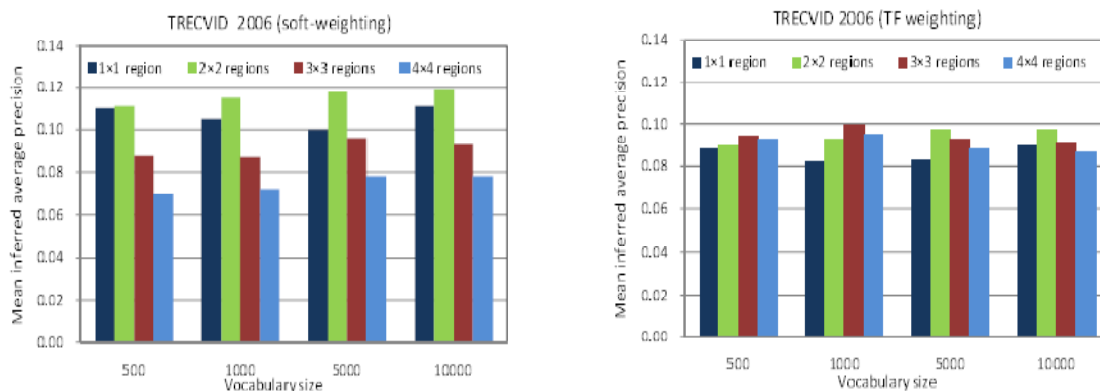

Fig.5.Concept detection performance on TRECVID 2006 using region-based features computed from different spatial partitions. Due to the feature mismatch problem caused by spatial partition, relatively coarse region partition such as 2x2 is preferred.

5. APPLICATIONS

The CBIR technology has been used in several applications such as fingerprint identification, biodiversity information systems, digital libraries, crime prevention, medicine, historical research. Various applications of CBIR and the approaches used are shown in the Table 1. The selection of the CBIR process depends on many constraints and has to be chosen to best suit the final goal. Constant efforts are being made to reach a 100% accuracy rate.

Table 1.Applications of CBIR

Application	Feature Extraction method	Feature Vector Generated	Similarity Measuring Technique	Performanc e Measuring Technique	No of Images in databas e	No of querie d images	Accuracy Rate
Recognition of English Handwritten Characters	Boundary Detection Feature Extraction	Fourier Descriptors	Multilayer Perception Neural Network	--	250	250	94%
Image Retrieval From Database	Color Averaging Techniques	RCM, CDM, FDM	Euclidean Distance	Precision and Recall	1000	55	--
Handwritten Character Recognition	Gradient Feature Extraction	12 directional Values	Back Propagatio n Neural Network	--	250	250	97%
Feature Vector Dimension Reduction	Principal Component Analysis	Eigen Faces	Euclidean Distance	Precision and Recall	1080	75	Complexity Reduced by 99.9%
Image Retrieval from Database [4]	Slope Magnitude on Gradient	Shape Features	Mean Square Error	Precision and Recall	1000	55	
Iris Recognition[19]	Histogram of Gray Images	Correlation metric at different Resolutions	Deviation and Threshold	False Accept Rate, False Reject Rate	26400	2400	EER=12.48 %
Digital Image Search and Retrieval	Fast Fourier Transform Sectorizatio n	Eight Component s	Euclidean Distance	Precision and Recall	170	7	Good Accuracy Rate.

6. INFERENCES MADE

- EDH technique can be extended to spectral and frequency domain to extract the features.
- In addition to spatial context exploitation, additional contextual information sources and their combination with spatial context can be explored.
- Due to data variations due to the image capture condition study of appropriate image normalization procedures
- Developing relevance feedback CBIR approaches for hyper spectral images.
- By mining semantically meaningful patterns we can improve the image search and retrieval process.

7. CONCLUSION AND FUTURE WORK

In this paper we reviewed about various feature extraction methods, It has been found that variation in feature extraction methodologies can ensure the better and more accurate retrieval of relevant images from the large database. We have also explored the feature set in spatial domain of CBIR system. In every kind of application we strive to achieve a 100% accurate retrieval rate. Lots of research is being currently done for the improvement in the feature extraction methods. In future the features can be extracted using spectral and frequency domain and we can also extract the image by combining all the domains. Investigation of additional contextual information sources, like scene type related information, and their combination with spatial context for achieving will also further improve the performance of Image Retrieval Systems. In order to improve the effectiveness of CBIR systems we need to discover semantically meaningful patterns. The spatial locations are the key points in an image carry important information for classifying the image. By integrating the spatial information with key point selection technique we can detect semantic concept from which semantically meaningful patterns can be discovered.

REFERENCES

[1] Minakshi Kaushik, Rahul Sharma,, Ankit Vidhyarthi ,” Analysis of Spatial Features in CBIR System“,International Journal of Computer Applications (0975 – 8887) Vol.54 No.17,pp. 11-15, September 2012 .

[2] Guang-Hai Liu , Jing-YuYang “Content based image retrieval using color difference histogram”, ELSEVIER journal on Pattern Recognition,Vo.l 46, pp. 188–198,2013.

[3] G.Th. Papadopoulos,C. Saathoff, H.J. Escalante V. Mezaris I. Kompatsiaris, M.G. Strintzis,“A comparative study of object-level spatial context techniques for semantic image analysis” Elsevier journal on Computer Vision and Image Understanding pp.1288–1307,2011.

[4] Chandan Singh, Pooja, “An effective image retrieval using the fusion of global and local transforms based features”, Optics & Laser Technology Vol. 44 pp.49–2259, 2012.

[5] M. Subrahmanyam , R.P. Maheshwaria, R. Balasubramanian ,” Expert system design using wavelet and color vocabulary trees for image retrieval”, Elsevier journal on Expert Systems with Applications Vol.39,pp. 5104–5114,2012.

[6] Henrike Stephani1, Michael Herrmann, Frank Bauer, and Bettina Heise,” Wavelet-Based Dimensionality Reduction for Hyper spectral THz Imaging ““,Terahertz Science and Technology, ISSN 1941-7411 Vol.3 No.3, September 2010 .

[7] Manuel Gran, Miguel A.Veganzones,”An end member-based distance for content based hyper spectral image retrieval”, Elsevier journal on Pattern RecognitionVol.45, pp. 3472–3489,2012.

[8] Rafiee, G., Dlay, S.S., Woo, W.L.: A Review of Content-Based Image Retrieval”, in CSNDSP pp. 775–779 IEEE 2010.

[9] Srinivasulu, S., Sakthivel, P., ,” Extracting Spatial Semantics in Association Rules for Weather Forecasting Image” Trends in Information Sciences & Computing (TISC),pp.no.54-57,Dec 2010.

[10] Pradnya Rane, Pallavi Kulkarni, Suchita Patil and B.B.Meshram, “Feature based image retrieval of images for CBIR”, IJCEM International Journal of Computational Engineering & Management, Vol. 14, October 2011

[11] C. G. M. Snoek, M. Worring, J. C. van Gemert, J.-M. Geusebroek, and A. W. M. Smeulders, “The challenge problem for automated detection of 101 semantic concepts in multimedia,” in ACM Multimedia, 2006.

[12] G. Rafiee, S.S. Dlay, and W.L.,” A Review of Content-Based Image Retrieval” CSNDSP,pp. 775-779, IEEE 2010.

[13] Hui Hui Wang, Dzulkifli Mohamad, N.A Ismail,” Image Retrieval: echniques, Challenge, and Trend“ World Academy of Science, Engineering and Technology, pp. 716-718,2009.

[14] Lokesh Setia, Machine Learning Strategies for Content Based Image Retrieval, M.Sc Dissertation, 2008. Albert-Ludwigs-Universität, Freiburg im Breisgau.

[15] Pradnya Rane ,Pallavi Kulkarni, Suchita Patil and B.B.Meshram, "Feature based image retrieval of images for CBIR", IJCEM International Journal of Computational Engineering & Management, Vol. 14, October 2011

[16] Swati V. Sakhare & Vrushali G. Nasre, "Design of Feature Extraction in Content Based Image Retrieval (CBIR) using Color and Texture", International Journal of Computer Science & Informatics, Vol.I, IssueII, 2011