A Survey Paper On Content Based Image Retrieval

[1]Roshni Shinde, [2]Buddhapriya Shinde, [3]Mansi Sonawane, [4]Dipali Telangi, [5]Jaya Dewan

Information Technology Department,
PCCOE, Nigdi, Pune, India.

Abstract

Content based image retrieval is the task of retrieving the images from the large collection of database on the basis of their own visual content.In CBIR, each image that is stored in the database has its features extracted and compared to the features of the query image. It involves two steps: Feature Extraction is the process is extracting image features to a distinguishable extent.Matching is the second step involves matching these features to yield a result that is visually similar. To increase efficiency, accuracy and speed of the image retrieval from the large dataset, the second approach known as Content-Based Image Retrieval (CBIR) techniques are used. In a CBIR system, images are automatically indexed by visualizing their valued features such as color, texture, and shape. The performance of Content-Based Image Retrieval (CBIR) system depends on efficient feature extraction and accurate retrieval of similar images.

Keywords—*Bag of visual words(BoVW), Gist, Histogram of orientation gradient(HoG), Scale invariant feature transform (SIFT), Harris corner point detection, Content based image retrieval(CBIR)*

I.]INTRODUCTION

CBIR is Content Based Image Retrieval. In CBIR we retrieve the images based on local features and global features such as colour, texture and shape. Reasons for its development are that in many large image databases, traditional methods of image indexing have proven to be insufficient, laborious, and extremely time consuming.

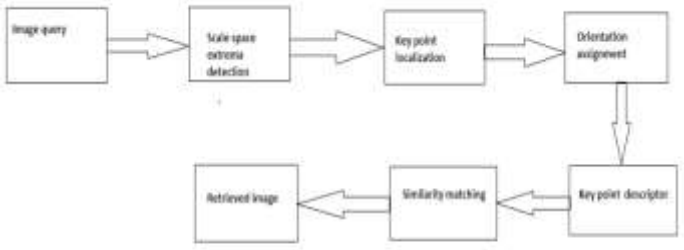
These method of image indexing ranging from storing an image in the database and associating it with a keyword or number, to associating it with a categorized description, have become obsolete. In the content based image retrieval, each image that is stored in the database has its features extracted and compared to the features of the query image CBIR works as follows:

1.] Query Image
2.] Feature Extraction
3.] Similarity Matching
4.] Retrieved Similar Image

1. SIFT:

Sift algorithm used in content based image retrieval .
Scale Invariant Feature Transform (SIFT) is an algorithm used to detect and describe the local features of the image . SIFT was proposed by David Lowe .
Image content is transformed into local feature coordinates that are invariant to translation, rotation, scale. The main SIFT algorithm consists of four major stages :

Scale-space extrema detection
Keypoint localization
Orientation assignment
Key point descriptor



Block diagram of SIFT steps

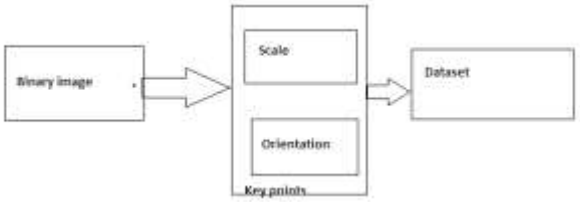
[1.1]SIFT algorithm (Descriptor) are used here for Binary and gray scale images. Binary images are those which has only two colours Black and white. This binary image used in CBIR system for retrie 92% found in this paper. It is calculated by using Recall rate and Precision rate. But SIFT has high computational complexity, speed is less and much more expensive. Therefore SURF (speeded up robust feature) is used .SURF works faster than SIFT and also have less computational complexity. The drawback of SIFT and SURF is both are patented. So it is not possible to use them. Not possible for low power devices. SURF has better time cost than SIFT and also RGB (noise) removable as compared to SIFT.

Algorithms Used	SIFT	k-means, Clustering, SIFT
Accuracy	94%	81.3%
Advantages	Easy manipulations. Cheap to store Faster to process.	Efficient Exact
Disadvantages	Not effective to low power devices	

PCA-SIFT is different algorithm from simple SIFT ,it does not use histogram for normalizing the patch. It uses principal component analysis.

[1.2] Hybrid method is used , where local global features extraction is done here. This descriptor is named as Histogram of local and global features (HLG-SIFT).For extraction of global feature new descriptor Upper-Lower of local binary pattern (UL-LBP). This descriptor works for the individual colour channel red , green and blue. After the combined result of both descriptor local feature by HLG-SIFT and global feature by UL-LBP it is quantized by Bag of visual words (BoW) similarity matching and image retrieval is done by large dataset present. The comparison between research papers is given below.

[1.2] Hybrid method is used , where local and global features extraction is done here. This descriptor is named as Histogram of local and global features (HLG-SIFT).For extraction of global feature new descriptor Upper-Lower of local binary pattern(UL-LBP).This descriptor works for the individual colour channel red , green and blue. After the combined result of both descriptor local feature by HLG-SIFT and global feature by UL-LBP it is quantized by Bag of visual words (BoW) similarity matching and image retrieval is done by large dataset present.



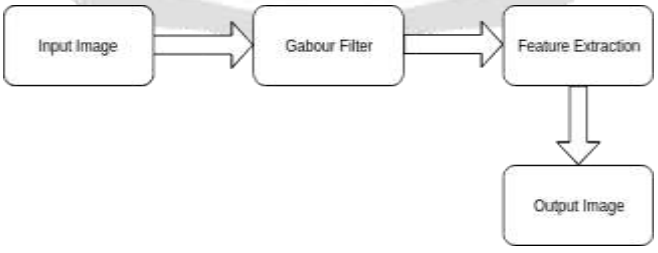
Block diagram of SIFT in binary image

2]Gist global method

Humans can recognize the gist of a novel image in a single glance, independent of its complexity.How is this remarkable feat accomplished? On the basis of behavioral and computational evidence, describes a formal approach to the representation and the mechanism of scene gist understanding,based on scene-centered, rather than object-centered primitives. We show that the structure of a scene image can be estimated by the mean of global image features, providing a statistical summary of the spatial layout properties (Spatial Envelope representation) of the scene. Global features are based on configurations of spatial scales and are estimated without invoking segmentation or grouping operations.The scene-centered approach is not an alternative to local image analysis but would serve as a feed-forward and parallel pathway of visual processing, able to quickly constrain local feature analysis and enhance object recognition in cluttered natural scenes. Gist is the global feature extraction technique, in which the whole image is considered for feature extraction.
That means on whole image the features are extracted like color,shape,texture. It is mainly used for content based image retrieval,where the input image is given to the system and on that query image gist performs the feature extraction so that to find or retrieve similar images from the large databGist is mainly used for scene recognition, Regarding the global descriptor Gist, it has been observed that , a higher capability and possibly low cost, which indicates a good potential for image matching from huge database.

Oliva and Torralba introduced gist mainly to recognize the spatial envelope of scene(its volume, shape, level of naturalness, level of roughness, level of openness, etc),so that to match the image with other images in huge database. The gist features are also no changes to scale because the majority of the scenes (background) are not moving and the system is trained at all viewing distances. The gist features achieve a solid illumination no change when trained with different lighting conditions. Last, the combined-sites experiment shows that the number of differentiable scenes can be quite high.

Different Approaches of Algorithms:



Gabor Filter is the technique used for filtering the image, in order to do feature extraction on whole image. After filtering the image, divide it into 4x4 grid cells and then combine its response and then extract the features.

Gist indexing structure known as GISTIS, the first filtering step of our image search system, relies on an efficient gist indexing structure known as GISTIS. It is derived from which employs an inverted file .In this structure, a quantized index identifies the descriptor lists that are likely to contain the nearest descriptors.The two stage re-ranking, indexing structure proposed above dramatically reduces the number of images assumed to be relevant. For a billion image dataset, it would typically return one million images. Having filtered the majority of the images using GISTIS,

we can now apply more precise image comparison methods. Here, we propose two strategies that are used either independently or jointly to refine the ranking provided by our efficient structure.

3] Harris corner point detector:

A corner is nothing but the intersection of two edges.The corner is important for detecting the any object. Corner detection is a popular research area in image processing and therefore many corner detectors have been presented. Some of them are widely used in industries.Gaussian filtering is used to blur images and remove noise .

Following are the steps used to detect corner:

• Calculate the local Harris matrix A

• Compute the response function R(A)

• Choose the images for corners detection by selecting thresholds on the response function R(A)

• Apply non-maximal suppression

• Filter the image with a Gaussian filter.

• Estimate intensity gradient in two coordinate directions

Harries is a local technique to extract the feature from the large image. The corner is important for detecting the any object .Harris corner detector is based on the auto correlation function of the signal. The basic idea of this detector is we find whether point shows significant change in all direction or not. If yes then point is marked as a corner point . If both of the Eigen values of the second moment matrix are large and nearly equal than that point are considered as the corner point. The Harris corner detector is invariant to translation, rotation and illumination change This detector is most repetitive and most informative. The points with the higher curvature along with the corner points. Here we find the the second moment matrix which requires finding the gradients of an image which is sensitive to noise and computationally expensive.

[3.1] In this paper used an object recognition and identification system using the Harris Corner Detection method.

Following steps are used to detect the object using Harris corner detection:
1. Select colour image to which image we have to search. The image colours are changed to grayscale.
2. Gaussian Filter: Gaussian filtering is used to blur image and remove noise and detail. Log (Laplacian of Gaussian) to localize interest points in scale. Gaussian Filter is used for noise reduction and blurs the image by applying the Gaussian Filter convolution with the original images. Algorithms for noise reduction are derived from the calculation of the mask size.
3. Gradient Calculation Divide the Gradient Calculation into two parts: x, partial derivatives P[i, j] and Q [i, j] was built from the smooth image Magnitude and direction of gradient was built from x, y partial derivatives which can be calculated by rectangular to polar conversion.

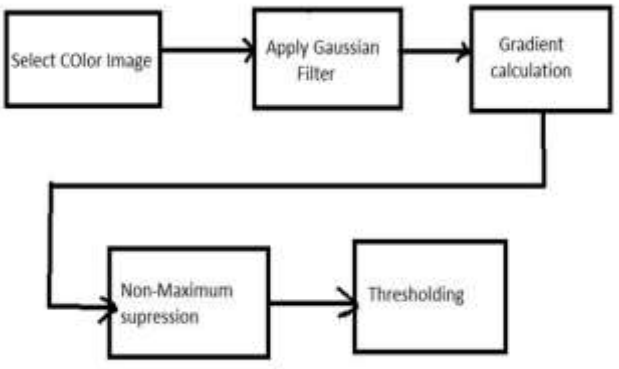


Fig Object detection using Harris corner detection technique

4.Non Maxima Suppression The pixels of edge detection from canny method are the highest values of gray scale level and the same direction of gradient. Such a method has a thin edge only 1 pixel. The Image pixels after making Non Maxima Suppression are zero at all points except the local maxima points which are still the old value
5.Thresholding This paper used the concept of threshold in the final process of the edge detection for reduced noise or the patterned surface of the object in the image.

[3.2] In this paper shows that Harris corner detection method haves the best result as compared to previous method . The experimental results show that this method can reduce the time of detection, and also detect the corner points which are distribution proportion, avoid the clustering phenomenon.

1.For each pixel (x, y) in the image calculate the autocorrelation matrix
2. For each pixel of image has Gaussian filtering, get new matrix M , and discrete two-dimensional zero-mean Gaussian function.
3. Choose the local maximum point. Harris method considers that the feature points are the pixel value which corresponding with the local maximum interest point.

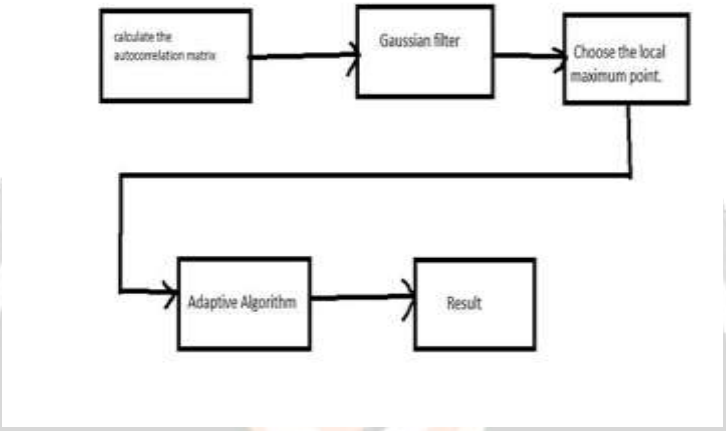


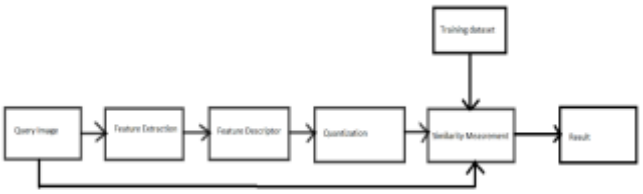
Fig Harris corner detection

4. Set the threshold T, detect corner points. Adaptive Algorithm of Harris Detection: In this process we adopt neighboring point eliminate method to avoid setting threshold T, ensure the corner will uniform distribution and avoid corner clustering. We can select one stencil (e.g. 3×3) to process for image. If there is not only one corner point under this stencil, we retain the corner which has the largest value of R.

5.In order to verify the validity of the algorithm, we compare the original algorithm and the adaptive algorithm.

4.] Bag of visual words:

The bag-of-words (BoW) model is also called as Bag of features. BoW model in CBIR is mainly designed for the local descriptors of images which describe regions around the key points detected in the images. Different from global features describing a picture in a holistic way. And local descriptor like SIFT is demonstrated to be a good way to represent the characteristics of these patches. But once we extract such local descriptors for each image, the total number of them might be huge. And searching nearest neighbours for each local descriptor in the query image becomes too time consuming. Therefore, BoW is proposed as a way to solve this problem by quantizing descriptors into "visual words", which decreases the descriptors' amount dramatically.



Initially, Bovw was used for text but now it can used for images.
There are three steps involved in Bovw
A.] Extract the interest point and descriptor from the query image.
B.] Quantize the interest point and descriptor to form a codebook.
C.] Calculate the occurences of each visual words in the image to generate a histogram.

Parameter	Paper1	Paper2
Author	Naimeh Sadat Mansoori, Mansour Nejati, Parvin Razzaghi and Shadrokh Samavi	Ding Yuxin1, Zhao Bin1 , You Qingzhen1 , Chai Guangren1,2
Publication	IEEE 2013	IEEE 2012
Technique	SIFT descriptor, hue descriptor, k-means algorithm	SIFT descriptor, k-means algorithm, DoG detector,Mean average precision
Advantage	Color sensitive	Decrease number of wrong matches, collect the information lost in the quantization
Disadvantage	Information lost in the quantization	

5.1]SIFT for Feature Extraction:
The scale-invariant feature transform (SIFT), it is an algorithm in computer vision which is used to detect and describe local features in the images.
SIFT isn't just scale invariant. You can change the following, and still get good results:

- Scale
- Rotation
- Illumination
- Viewpoint

5.2]Histogram:
A histogram gives an precise representation of the distribution of numerical data. It is an approximation of the probability distribution of a continuous variable (quantitative variable).

5.3] K-means clustering:
K-means is one of the easy and simple unsupervised learning algorithms that solve the clustering problem. The procedure have a simple and easy way to classify a given data set through a certain number of clusters

IV. CONCLUSION

Content based image retrieval is mainly used for retrieving the similar images from the large dataset efficiently.After a survey the previous CBIR works, the paper explored the low-level features of color and texture extraction for CBIR. After comparing the two color histogram features as well as comparing color and texture features, the paper implemented a CBIR system using color and texture fused features. Similar images can be retrieved quickly and accurately by inputting an image. More low-level features such as shape and spatial location features etc. will be fused to make the system more robust in the future.

V. REFERENCE

[1.1] Content-Based Image Retrieval using SIFT for Binary and Greyscale Images. 2013 IEEE International Conference on Signal and Image Processing Applications (ICSIPA)

[1.2] Content Based Image Retrieval using Local and Global features descriptor. 2nd International Conference on Advanced Technologies for Signals and Image Processing - ATSIP'2016 March 21-24, 2016, Monastir, Tunisia

[2.1]Content-Based Image Retrieval using Local Features Descriptors and Bag-of-Visual Words, Mohammed Alkhawlani Ibb University Ibb, Yemen

[2.2]OBJECT RETRIVAL BASED ON VISUAL WORD PAIRS Ding Yuxin1, Zhao Bin1, You Qingzhen1, Chai Guangren1,2 1 Harbin Institute of Technology Shenzhen Graduate School, China 2 State Key Laboratory of Computer Architecture, Institute of Computing Technology Chinese Academy of Sciences, China

[3.1] An Object Recognition and Identification System Using the Harris Corner Detection Method T. Kitti, T. Jaruwan, and T.Chaiyapon

[3.2] An Adaptive Algorithm for Harris Corner Detection Zhiyong Ye Yijian Pei Jihong Shi School of Information School of Information School of Information Yunnan University Yunnan University Yunnan University Kunming, China Kunming, China Kunming, China

[4.1] A. Oliva and A. Torralba, "Modeling the Shape of the Scene: Holistic Representation of the Spatial Envelope",IEEE publication.

[4.2] Rapid Biologically-Inspired Scene Classification Using Features Shared with Visual Attention Christian Siagian, Student Member, and Laurent Itti, Member, IEEE.

[4.3] Evaluation of GIST descriptors for web-scale image search,IEEE publication.