Content Based Image Retrieval Systems using SIFT: A Survey

Kimaya S. Meshram[1],Ajay M. Agarkar[2]
[1]*M.E. Student, Digital Electronics, SSGMCE, Shegaon, Maharashtra, India*
[2]*Professsor, Department of Electronics and Telecommunication, SSGMCE, Shegaon, Maharashtra, India*

Abstract

This paper provides the novel information about Image Retrieval and Content based image retrieval (CBIR) system since it is now a big need of society. CBIR is a method for finding similar images from large image databases. Different approaches are used for Content based image retrieval, out of which Scale invariant feature transform is very popular. In this paper, we discuss the Industrial product development using SIFT Algorithm and results obtained in several experiments proposed to evaluate the application of Scale Invariant Feature Transform (SIFT) in content based Image Retrieval (CBIR) systems.

Keywords— *Content Based Image Retrieval (CBIR), Scale Invariant Feature Transform (SIFT)*

I. INTRODUCTION

What is an image? An Image is defined as the Picture which represents something that has been created or copied and can be stored in electronic form. Image size is measured in terms of pixel value. Resolution of an image defines the Quality of an image.

Images are everywhere. The current rapid improvement in the digital storage media, image capturing devices like scanners, web cameras, digital cameras and rapid development in internet provide a huge collection of images. This leads to the retrieval of these images for visual information efficiently and effectively in different fields of life like medical, medicine, art, architecture, education, crime preventions. To achieve this purpose many image retrieval systems have been developed.

Since 1970's the research is active in the field of Image Retrieval, which concerned with the study of searching and browsing digital images from database collection. Different Approaches are used for the image retrieval. Free browsing, keyword based image retrieval or text based image retrieval; these different approaches are previously used in image retrieval. But these methods having some drawbacks. Because of its limitations and scalability problem, the new image retrieval method is emerged named as Content based Image retrieval (CBIR). Though there is a progress in image retrieval, the performance is still not satisfied, since the presence of semantic gaps. Semantic gap is defined as the gap between low level feature and High level feature. Many techniques are used to capture scene semantics in image processing and computer vision. Out of which Scale Invariant feature transform (SIFT) is used in many applications.

Prof. David Lowe proposed SIFT Algorithm in ICCV 1999.Which is refined in 2004 in IJCV, which is cited more than 31,000 times till now. SIFT is used in various applications [1].

II. RELATED WORK

Extraction of visual content of query image and database image is obtained by using CBIR with 128 dimensional SIFT feature vector. For indexing and matching SIFT features, KDtree with the Best Bin First (BBF), an Approximate Nearest Neighbors (ANN) search algorithm is used. For retrieving the top similar images, a modified voting scheme called Nearest Neighbor Distance Ratio Scoring (NNDRS) is used. To re-check the results and for removing the false matches, Random Sample Consensus (RANSAC) is adopted as a geometry verification Method. The famous image database ZuBud is used to obtain a high recall and high precision in the context of CBIR [3]. This can be used in Application of Image SIFT Features to the Context of CBIR. The SIFT feature is invariant to rotation, image scaling and transformation and partially invariant to illumination changes and affine transformation. Based on ROI i.e. Region of Interest, Content based image retrieval algorithm is implemented by using scale invariant featuretransform. This gives the improved stability and precision in image retrieval [4]. Indicating that SIFT algorithm can be implemented effectively for image retrieval based on Region of Interest.

Visual attention analysis is used for ranking all keypoints in an image with its saliency weight, and only the most distinctive keypoints will be reserved. So that, the background features are reduced evidently, and the accuracy of retrieved images can be improved at the same time. Effective alternative of standard SIFT is AF-SIFT [5]. SIFT Algorithm is used for the video tracking, where problem in tracking object and scene background change during tracking is possible [6].Application areas where SIFT is used for Video tracking which provide the good efficiency in the experimental results. For improved output AF-SIFT can be the

good alternative. Sometimes the Content based image retrieval may fail on capturing some local features which are representing the details and duplicate degree of difference of scenes. So, Scale invariant feature transform is used, since it becomes effective when applied to retrieve images [8], [20]. Since some previously used techniques of image retrieval may fail on capturing some local features representing the details and nuances of scenes.

SIFT is used for retrieving the Medical images with the help of bag-of-words framework [9]. The development in digital image equipment's providing the data produced in hospitals and medical centers is increasing exponentially leads to the image retrieval in Medical area. The craze and need of Internet surfing is increasing day to day, this leads to the web image retrieval. With the help of Content based image retrieval by Scale invariant feature transform Web images can be retrieve [10]. Which proceed towards the web image search engines by using Scale invariant feature transform algorithm. Keypoints features are stored as a XML format which becomes helpful for improving the searching performance. Related technology and Hibernate framework is used for establishing a link with the database of all of the information of image and completed the development of persistent object. At last, through the HTML the results displayed to the user [11], [14]. By combining the distance measure of global image and sub-image the image retrieval is possible using Scale Invariant Feature Transform Algorithm [13]. Along with these feature if more features such as color feature is involved it can give the better results in the future.

As the Scale invariant feature transform is invariant to illumination, partial image retrieval is done providing image retrieval which is robust to apparent changes occurred in the color [15]. Because of semantic gaps the performance of image retrieval is not good, for that integration of Scale invariant feature transform and Local binary pattern is used based on the bag-of-feature model. Along with this model, when weighted K-means algorithm is used then the performance obtained is superior than other existing algorithms [17]. Image retrieval using Scale invariant feature transform is quietly costly. Therefore, binary quantization of SIFT feature vectors with the help of medians of vector components is proposed. binary quantized features provide significant computational and storage gains [18] For image retrieval, SIFT-based elastic sparse coding is developed based on bag-of-features, which integrates the bag-of-feature model and sparse representation based on Elastic Net regression model. It uses the Scale Invariant Feature Transforms feature descriptors to construct the codebook, and simultaneously it provides the Elastic sparse coefficient vector, which considers some existed correlation between the local feature and multiple visual words, as well as eliminates the redundant information among sparse-coefficient vectors, At last, the vector sum of the sparse-coefficient vectors from an image is the representation of the image. With the help of Coil20 dataset, the effectiveness of the model is demonstrated [21].Ultimately, Better Performance is important in the retrieval of images. Therefore, using various models to have novel experimental results.

Retrieving images at different viewpoints, using combined feature descriptors. Scale Invariant Feature Transform (SIFT) and Deformation and view point Invariant Color Histogram (ICH) in HSV color space are used to extract the content of images with these two descriptors. The properties of SIFT shows that it is the most reliable descriptor for rotation, translation and partially to illumination and affine or 3D projection invariant image matching. Since, it is designed for gray images. SIFT features are combined with ICH in HSV color space for Image Retrieval to increase the deformation and viewpoint invariance capability and thus to improve image recognition. By using this method the efficiency is improved to 10% than the other methods [22]. By using the different image databases like MPEG-7, COIL-20 and ZuBuD, for binary and grayscale level images, the content based image retrieval using scale invariant feature transform is done. The images having corners and edges provides more keypoints of SIFT features which make the retrieval results better than image that has lesser corners [23]. Corners refers to the place or angle where two sides or two edges meets means when it is used with SIFT computation of feature vector at each point, large no. of vectors are produced for the more keypoints.

An embedded real-time scene recognition system is proposed. The computing speed is High, because of three Descriptors used. (1) Scale invariant feature transform, (2) 18-dimensional new descriptor, and (3) five-branch-tree data structure. Firstly SIFT is adapted and implemented in FPGA, then second descriptor is used and finally Third descriptor is used which is designed for fast searching and matching of descriptor. The random access to SDRAM is the limitation on speed in the system. In this case, one keypoint can be matched to multiple keypoints and thus multiple images with overlapping content can be recognized from the library simultaneously, this is an Advantage of system. So, in this system, it enables to remember a sequence of scenes along the path for navigation when equipped on a mobile robot [24]. Canny edge detection (CED) along with the Scale invariant feature transform is used to retrieve the similar matching images. To easily identify and trace the specified image from large the Database images reduced SIFT with Canny Edge Detection (CED) algorithm is used [25]. The retrieval speed is

improved and retrieves accurate resultant outcome with the help of Reduced Scale invariant feature transform (RSIFT). To extract the features of the images and to improve the performance of the system, the CIBR systems using the SIFT algorithm. The SIFT led to the excellent performance in image matching and to retrieve desired images from the database [26].CBIR can be possible by different transforms also [27].These are the various ways for Content Based Image Retrieval by using SIFT and Other Algorithms.

Now, we will see the Details of SIFT Algorithm.

A. *SIFT Algorithm*

Following are the major stages of computation in SIFT Algorithm used to generate the set of image features:

1) *Scale-Space Extrema Detection:*

All scales and image locations are searched in the first stage of computation and it seeks to identify the potential interest point by difference-of-Gaussian (DoG) function it is implemented efficiently.

The scale space of an image $I(x,y)$ is computed by the convolution of an image with variable-scale Gaussian.

$$L(x,y,\sigma\ ) = G(x,y,\sigma\ ) * I(x,y) \quad (1)$$

Where, $*$ is the convolution operation in$x\ and\ y$, $I(x,y)$ is an input image, $G(x,y,\sigma\ )$ is variable-scale Gaussian and $L(x,y,\sigma\ )$ is scale space of an image.

$$G(x,y,\sigma) = \frac{1}{2\pi\sigma^2} e^{\frac{-(x^2+y^2)}{2\sigma^2}} \quad (2)$$

Difference of Gaussian is an algorithm for detecting edges in an image. Which help to identify potential interest points that are invariant to scale and orientation. DoG i.e. Difference of Gaussian function $D(x,y,\sigma)$ can be computed from the difference of two nearby scales separated by a constant multiplicative factor k.

$$D(x,y,\sigma) = \big(G(x,y,k\sigma) - G(x,y,\sigma)\big) * I(x,y) \quad (3)$$

$$D(x,y,k\sigma) = L(x,y,k\sigma) - L(x,y,\sigma) \quad (4)$$

2) *Keypoint Localization*

In the second stage, a detailed model is used to determine location and scale it at each candidate location. Based on the measure of keypoints stability they are selected, keypoints resistant to image distortion are stable keypoints.

For detecting local maxima and minima of DoG each point is compared with its eight neighbors on same scale and nine neighbors in the scale above and below as shown in fig 1.

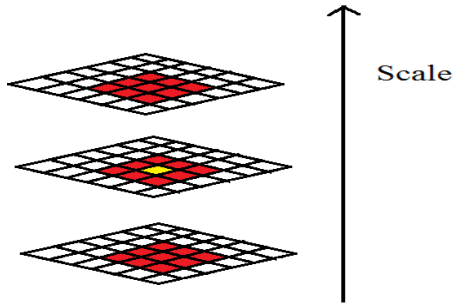

Fig 1.Keypoint Localization at Different Scales.

3) *Orientation Assignment*

In the third stage, SIFT computes the direction of each image gradient. Based on local image gradient directions, one or more orientations are assigned to each keypoint location; its feature vector can be represented to this orientation so that rotation invariance can be achieved.

Keypoint orientation can be calculated by orientation of histogram of local gradients from the closest smoothed image $L(x,y,\sigma\ )$. The gradient magnitude $m(x,y)$and orientation $\theta(x,y)$can be computed using pixel differences for each image sample $L(x,y)$.

$$\frac{\partial L}{\partial_x}(x,y) = L(x+1,y) - L(x-1,y) \quad (5)$$

$$\frac{\partial L}{\partial_y}(x,y) = L(x,y+1) - L(x,y-1) \quad (6)$$

$$m(x,y) = \sqrt{\frac{\partial L}{\partial_x}(x,y)^2 + \frac{\partial L}{\partial_y}(x,y)^2} \quad (7)$$

$$\theta(x,y) = \tan^{-1}\left(\frac{\frac{\partial L}{\partial_y}(x,y)}{\frac{\partial L}{\partial_x}(x,y)}\right) \quad (8)$$

*4) **Keypoint Descriptor***

In the fourth stage of computation, the local image gradients are measured at the selected scale in the region around each keypoint. These are transformed into a representation which allows for significant levels of local shape distortion and change in illumination.

This approach has been named as the Scale Invariant Feature Transform (SIFT) because it transforms image data into scale-invariant coordinates relative to local features. Fig 2 shows the computation of the keypoint descriptors. A 2x2 descriptor array is computed from an 8x8 set of samples is as shown in figure 2, whereas in some experiments 4x4 descriptors are computed from a 16x16 sample array [1].

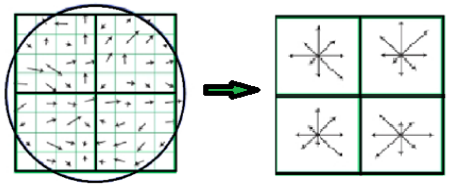

Fig 2 Shows the Computation of the Keypoint Descriptors.

III. INDUSTRIAL PRODUCT DEVELOPEMENT USING SIFT ALGORITHM

Industries are developing very faster and it requires the lot of human labor. But using human labor everywhere in industry is not possible. Since, it may cause damage as Industrial Environment is very Messy and needs Very efficient management.

Scale invariant feature transform Algorithm can be used in various Industries like Automatic object searching system based on real time scale invariant Algorithm [16]. With the help of texture information, wood image retrieval is done using Scale invariant feature transform. No color information is used in this case [12]. Wood image retrieval improves the quality in timber production, furniture manufacturing and also in other fields due to the advent and increasing popularity of hand-held devices.For Blind people, intelligent navigation system was proposed based on image matching with the help of computer vision using SIFT Algorithm [19].

Quality, Time and cost are the important factors in any industrial product, for that purpose Automated Visual Inspection (AVI) system was developed using SIFT Algorithm for small to medium size industries [7]. For extracting a set of characteristic features SIFT algorithm is used in visual pattern recognition technology which is a basic primitive in the Robotics [2].

IV. PROPOSED FRAMEWORK

For content based image retrieval system, when Query image is given to retrieve the same kind of image from the database. Different databases are available, here for Experimental Results Wang Database is used. The SIFT algorithm is used for feature extraction providing single feature vector for input image and number of feature vector for database images. Then classifier is used at the feature vector of database providing two types of images.

(1) Relevant images
(2) Non-Relevant images

After that Matching function is used for relevant images and Query image. However, Non-Relevant images are discarded. Various kind of matching functions are available like Euclidean distance function,Mahalanobis distance function, Earth Mover's distance function depending on their efficiency. After that we will get the Array of an images then image sorting is carried out to retrieve the relevant images.

V. FLOWCHART

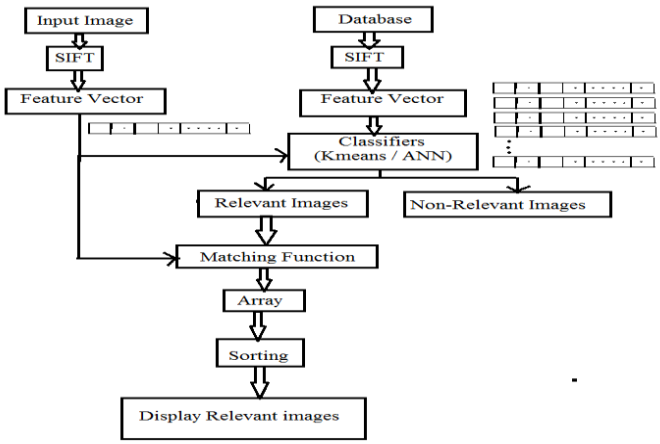

Fig. 3 Flowchart of the Content Based Image Retrieval System using SIFT Algorithm

VI. EXPERIMENTAL RESULTS

Experimental results shows the Input Image, Output Image, Time required for Gaussian scale space construction, Time for Differential scale space construction, Time for finding keypoints, Total number ofkeypoints extracted, Time for calculating descriptor. Here the Value of y and z is fixed. Value of y = 256 256 and Value of z = 128 128 So that Output image obtained will provide the output of same size of images. Experimental Results for the SIFT keypoints calculations are given below in Table I.

Table I

Sr. No.	Input Image	Output Image	y	z	Time for Gaussian scale space construction	Time for Differential scale space construction	Time for finding key points	Total number of key points extracted	Time for calculating descriptor
1. (40.jpg)			256 256	128 128	21.239 s	0.028 s	0.533 s	779	2.979 s
2. (100.jpg)			256 256	128 128	21.423 s	0.027 s	0.610 s	1142	3.761 s
3. (639.jpg)			256 256	128 128	21.449 s	0.027 s	0.286 s	277	0.986 s
4. (896.jpg)			256 256	128 128	21.425 s	0.027 s	0.422 s	467	1.701 s
5. (1110.jpg)			256 256	128 128	21.816 s	0.030 s	0.478 s	521	1.935 s

6. (1670.jpg)			256 256	128 128	21.318 s	0.028 s	0.492 s	923	3.199 s
7. (2334.jpg)			256 256	128 128	22.149 s	0.027 s	0.548 s	848	3.159 s
8. (3292.jpg)			256 256	128 128	24.537 s	0.029 s	0.674 s	1166	6.978 s
9. (4038.jpg)			256 256	128 128	16.533 s	0.017 s	0.320 s	1318	4.548 s
10. (4698.jpg)			256 256	128 128	24.564 s	0.037 s	0.518 s	665	3.013 s

VII.CONCLUSION

The purpose of this survey is to provide an overview of Content based image retrieval System by using Scale invariant feature transform Algorithm. The use of SIFT algorithm for various Industrial Products in small and medium level industries for better product, optimization of time, less cost and reasonably less human labor. Ultimate goal of this survey is to study SIFT algorithm for image retrieval and various other applications in the field of image processing. Experimental Results of given system shows that more key points will require the more time. And all key points are found successfully. Proposed system will provide the more efficiency compared to previously used methods for content based image retrieval. By using different algorithm along with SIFT algorithm with Good speed and less cost.

REFERENCES

[1] David G. Lowe," Distinctive Image Features from Scale-Invariant Keypoints"International Journal of Computer Vision, vol. 6, no. 2,pp. 91-110,2004.

[2] Mario E. Munich, Paolo Pirjanian, Enrico di Bernardo, Luis Goncalves, Niklas Karlsson, and David Lowe,"SIFTing through features with ViPR",Robotics & Automation Magazine,IEEE,2006.

[3] Xu Wangming, Wu Jin, Liu Xinhai, Zhu Lei, Shi Gang," Application of Image SIFT Features to The Context of CBIR", International Conference on Computer Science and Software Engineering,IEEE,2008.

[4] Zhuozheng Wang, KebinJia, Pengyu Liu,"A Novel Image Retrieval Algorithm Based on ROI by Using SIFT Feature

Matching",International Conference on MultiMedia and Information Technology,IEEE,2008.
[5] Ke Gao, Shouxun Lin, Yongdong Zhang, Sheng Tang, Huamin Ren,"Attention Model Based SIFT Keypoints Filtration for Image Retrieval",Seventh IEEE/ACIS International Conference on Computer and Information Science,2008.
[6] Xuelong Hu, YingchengTang, Zhenghua Zhang, "Video Object Matching Based on SIFT Algorithm",IEEE Int. Conference Neural Networks & Signal Processing, Zhenjiang,China, June 8~10, 2008.
[7] MehdyBohlool, SorooshRahimi Taghanaki," Cost-efficient Automated Visual Inspection System for small manufacturing industries based on SIFT"IEEE, 2008.
[8] Jurandy Almeida, Ricardo da S. Torres and Siome Goldenstein "SIFT applied to CBIR", J. Almeida, R. S. Torres and S. Goldenstein/ Revista de Sistemas de Informacao da FSMA n. 4 pp. 41-48, (2009).
[9] ZHI Li-jia, ZHANG Shao-min, ZHAO Da-zhe, ZHAO Hong, LIN Shu-kuan, ZHAO Da-zhe, ZHAO Hong ,"Medical Image Retrieval Using SIFT Feature", 978-1-4244-4131-0/09/$25.00 ©2009 IEEE.
[10] Zhuozheng Wang, KebinJia, Pengyu Liu,"An Effective Web Content-based Image Retrieval Algorithm by Using SIFT Feature",World Congress on Software Engineering,IEEE Computer Society, 2009.
[11] Zhuozheng Wang, Yalei Mei, Fang Yan,"A New Web Image Searching Engine by Using SIFT Algorithm",International Conference on Web Information Systems and Mining,IEEE Computer Society,2009.
[12] ShaoLi HUANG, Cheng CAI*, Yang ZHANG,"Wood image retrieval using SIFT descriptor",Science Foundation from The Northwest A&F University, China and the Students Research Fund from the Northwest A&F University, China,IEEE,2009.
[13] Bin Li, Xiangwei Kong, Zhe Wang, Haiyan Fu,"SIFT-based image retrieval combining the distance measure of global image and sub-image",Fifth International Conference on Intelligent Information Hiding and Multimedia Signal Processing, IEEE Computer Society,2009.
[14] Zhuozheng Wang1, Quanying Zheng2, Jingqi Sun "An Effective Content-based Web Image Searching Engine Algorithm",Proceedings of the IEEE ICMIT, 2010.
[15] Masaki KOBAYASHI, Keisuke KAMEYAMA,"Partial Image Retrieval using SIFT based on Illumination Invariant Features",MCIT,IEEE,2010.
[16] G.Kishore Kumar, G.V. Raghavendra Prasad, G. Mamatha,"Automatic Object Searching System Based On Real Time SIFT Algorithm",ICCCCT,IEEE,2010.
[17] Xiaoli Yuan, Jing Yu, Zengchang Qin, and Tao Wan, "A SIFT-LBP IMAGE RETRIEVAL MODEL BASED ON BAG-OF-FEATURES",18th IEEE Conference on Image Processing, 2011.
[18] Kadir A. Peker, "Binary SIFT: Fast Image Retrieval Using Binary Quantized SIFT Features",CBMI,IEEE,2011.
[19] SHI Shuo, YU Ming, XUE Cui-hong, and ZHOU Ying, "The Application of SIFT Algorithm in Blind Road Environmental Image Matching",Applied Mechanics and Material,Vols 155-156, pp 1137-1141,2012.
[20] Mamta Kamath, Disha Punjabi, Tejal Sabnis, Divya Upadhyay, Seema Shrawne,"Improving Content Based Image Retrieval using Scale Invariant Feature Transform", International Journal of Engineering and Advanced Technology (IJEAT),ISSN: 2249 – 8958, Volume-1, Issue-5, June 2012.
[21] Jun Shi, Zhiguo Jiang, Hao Feng, Liguo Zhang,"SIFT-BASED ELASTIC SPARSE CODING FOR IMAGE RETRIEVAL",ICIP , IEEE,2012.
[22] P.S.Suhasini, Dr. K.Sri Rama Krishna, Dr.I.V.Murali Krishna,"Combining SIFT and Invariant Color Histogram in HSV Space for Deformation and Viewpoint Invariant Image Retrieval",International Conference on Computational Intelligence and Computing Research,IEEE,2012.
[23] Suraya Abu Bakar, Muhammad Suzuri Hitam, Wan Nural Jawahir Hj Wan Yussof,"Content-Based Image Retrieval using SIFT for Binary and Greyscale Images",International Conference on Signal and Image Processing Applications (ICSIPA), IEEE, 2013.
[24] Han Xiao, Wenhao He, Kui Yuan and Feng Wen,"Real-time Scene Recognition on Embedded System with SIFT Keypoints and a New Descriptor", International Conference on Mechatronics and Automation, IEEE, August 4-7, Takamatsu, Japan, 2013.
[25] RajannaBandaru, Dinesh Naik,"Retrieve the Similar Matching Images Using Reduced SIFT with CED Algorithm",International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT), IEEE, 2014.
[26] Dr.K.VELMURUGAN,"A Survey of Content-Based Image Retrieval Systems using Scale-Invariant Feature Transform (SIFT)", International Journal of Advanced Research in Computer Science and Software Engineering,Volume 4, Issue 1, January 2014.
[27] Nawade Rajlaxmi1, Dr.Mrs.Lokhande S.S.," Content Based Image Retrieval Using Spectral Feature Extraction Methods", SSRG International Journal of Electronics and Communication Engineering (SSRG-IJECE) volume 2 Issue 4 April 2015.