International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878, Volume-8 Issue-5, January 2020

Sketch Based Image Retrieval in Large Databases using Edge Features

G.G Rajput, Prashantha

***Abstract:** Sketch-based image retrieval (SBIR) presents better flexibility in expressing the query as sketch for retrieval of images as opposed to text based retrieval. Using a sketch, it is easier to express the orientation and pose of the objects for image retrieval from the database. We propose an efficient approach for SBIR from large databases based on hand awn rough sketch. In the proposed method, images are synthesized to yield a binary sketch that is processed in similar way to user drawn sketch. Edge features are extracted by overlaying the sketch with non-overlapping and overlapping grids, respectively. The most similar images to the query are then retrieved from the database using weighted based similarity approach. Experiments are performed on flickr15k dataset yielding excellent retrieval performance in comparison to the methods available in the literature.*

***Keywords:** contour image, edge features, grid, sketch, large databases.*

I. INTRODUCTION

With growing usage of internet, there is rapid increase in multimedia contents which includes images, videos, and music. Efficient systems are required to retrieve the images, commonly held in multimedia databases. The SBIR system, based on user drawn sketch, finds corresponding similar images from the data set. Compared to text based retrieval, sketch based image retrieval offers greater flexibility in expressing query as free hand sketch, since it is easier to represent the orientation and pose of the object of interest in retrieval process. Text based search requires user to provide a word as a query based on which the system retrieves the images tagged with similar word. However, this approach assumes that all the stored images are tagged appropriately and the system requires tremendous pre-computation. Note that such an approach does not consider shape information in the image and human visual perception.

On the other hand, in sketch based image retrieval, searching for image(s) requires user to draw a free-hand sketch, visualizing the object (or viewing the object, if available), as a query to the retrieval system. However, comparison of sketch lines to natural images is a difficult task in SBIR system. Instead, descriptors are extracted from images and sketch and using similarity measures [1, 2, 3, 4].

II. REVIEW OF LITERATURE

The early approach addressed towards SBIR is image search based on the pictorial description [5]. In SBIR one of the chilling tasks is, directly comparing the sketch image to all database images. Instead, descriptors are used for fast access. The descriptors can be represented as points in a high-dimensional space and close match is found by searching nearest neighbors in this space. Many methods have been proposed in literature to extract descriptors, global or local [6]. The difficult task in SBIR system is matching the rough sketch of lines to natural image. Searching an image based on edge matching from 205 colored oil-painting dataset against the user query is presented in [7]. Searching and retrieval is performed by matching sketch image edges against the database image edges. Images are subdivided into 8×8 local blocks. After the normalization process, for each local block, the best local correlation is computed by searching in a small window of local blocks. In the work presented by Rui Hu et al. [8], multi scale region segmentation with a bag of visual words framework is used to retrieve images, given sketch as query. Then, contours of the region map are extracted that contains various levels of details. Multi scale HOG descriptor built on GF-HOG to capture the local structure of the image. Using k-means clustering, Bag of Visual Words (BoVW) codebook is formed. Finally, images are retrieved based on object of sketch images.

In its early stages, SBIR system involved sketch with blobs of color or predefined texture [16, 17, 21]. Later systems investigated shape descriptors [18] and spectral descriptors such as wavelets [24]. In the work presented by Eitz et al. [9], image is sliced using local grid and descriptors (color and texture) computed from each cell are concatenated resulting in global image feature. Grid based approaches are also found in other methods, for example in the EHD method, viz., Eitz et al. [10]. Szantoet al.[11] proposed user interface system for retrieval, using descriptors, namely, Histogram of oriented Gradients(HOG), Edge Histogram Descriptor(EHD), Scale Invariant feature Transform (SIFT). Minkowski distance has been used in matching process. The retrieval system has been implemented on small dataset.

Manuscript published on January 30, 2020.
* Correspondence Author
G.G Rajput*, department of Computer Science Akkamahadevi Women's University, Vijayapura 586108, Karnataka, India. Email: ggrajput@yahoo.co.in
Prashantha, department of Computer Science Rani channamma University Belagavi 591156, Karnataka India. Email: prashant.koryal@gmail.com

Jain et al.[12] proposed color and shape information based image retrieval. Edge histogram and color histogram are explored to capture color and shape of object. Recent work [13] presented an estimation of image by salient region matching. Mean shift clustering is used to generate visual word groups. Position descriptors are generated for each visual words and indexing has been build for group the visual words. Results are represented with indexing and without indexing. SBIR systems also support criminal investigation and identification of unsubstantial images [14, 15].

Hu et al. [19] presented a novel approach; the local structure is captured using a depiction invariant descriptor from transformed canny edge maps. Multi scale HOG descriptor is used to mitigate important information within a BoVW representation. Multiscale HoG descriptors are computed over gradient field interpolated from orientation of strong canny edges (GF-HOG). Eitz et al. [20], in later approach, computed HoG descriptor over canny edges. The shape, texture and color features are extracted from the photo-realistic images and indexing is used for fast retrieval [18].

III. DATABASE DESCRIPTION

Flickr15k dataset [23] is used for performing experiments. The dataset contains 60 classes of natural images and 10 sketch classes. Each sketch class contain 33 sketch images, the total number of sketch images are 10×33=330, and 60 classes of natural image contain 14660 natural images. The images include objects like horse, flower, moon, and koeln_dom, etc. Each class contains 33 sketch images. Fig1. Fig.2 presents natural images and sketch images.

Fig.1. Natural images from Flickr15k dataset

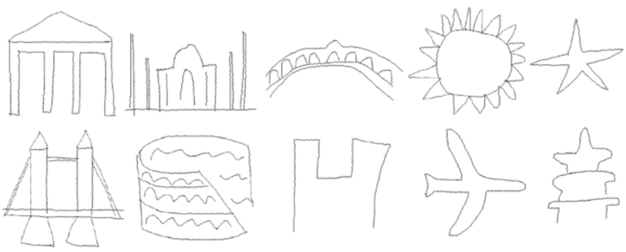

Fig.2. Sketch images from flickr15K dataset

IV. METHODOLOGY

The proposed method accepts free-hand sketch as query describing the shape of the object and retrieves images that contain similar shapes. Edge based features are computed to represent the sketch and images contained in the database. The block diagram of the proposed method is shown in Fig. 3. Global contour is extracted from sketch/natural image and edge features are computed using non-overlapping and overlapping grid approach. Employing weighted similarity approach, relevant images are retrieved from the database. Following sections describe the method in detail.

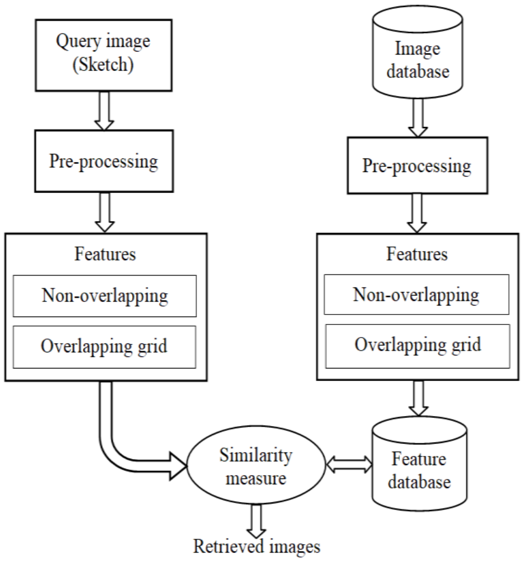

Fig.3. Block diagram of the propose method

A. Pre-processing:

A critical task in SBIR system is to represent a natural image as sketch in pre-processing stage. We extract global information from the natural images. Otsu's method [25] is used to extract the contour of the natural images. Empirically, we set the threshold as 40% to eliminate weak contours. Examples of natural images with corresponding contour images are presents in fig.4.

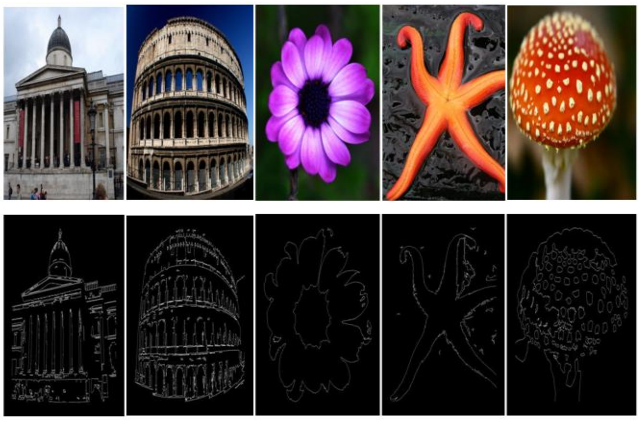

Fig.4. Global contour: The top row is original images and second row is contour images.

B. Non-overlapping grid approach

The global contour image resulting from the natural images in pre-processing stage are resized to 260×260 pixels, the size being determined empirically. The image is then divided into 100 sub images by placing a non-overlapping grid of size 10x10 (Fig. 5). Each sub image includes 26×26 pixels representing edge information of the object.

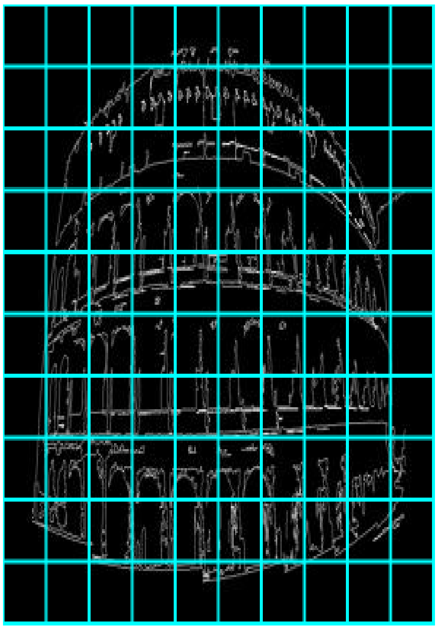

Fig.5.Example of non-Overlapping grid placing over the contour image.

The edge pixels for each sub image are computed and noted as edge features presents (Fig.6). feature vector is created by consolidating features extracted from all sub images .Similar procedure is followed for sketch image, given an input query, by first resizing the sketch image to 260×260 pixels. Euclidean distance has been used to measure the distance between query feature and database feature. Based on the minimum distance, images in the order of ascending distance values are then retrieved from the database.

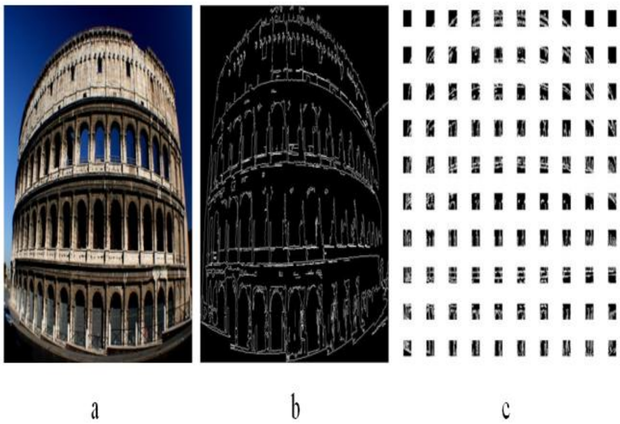

Fig.6. Example of non-overlapping grid based feature extraction (a) natural image (b) contour image (c) feature extraction

The retrieval results are presented in four steps; top 5, top10, top 15, and top 20. Out of top 5 resulting images, 5 images matched, out of 10 images 8 images are matched and 2 images are mismatched, out of 15 resulting images 13 are matched and 2 images are mismatched, and 17 images are matched and 3 images are mismatched out of 20 resulting images. Fig.7. shows results of non-overlapping approach, first column represents query image, second column represents top retrievals (5, 10, 15, and 20) and third column represents output results as mentioned in column two.

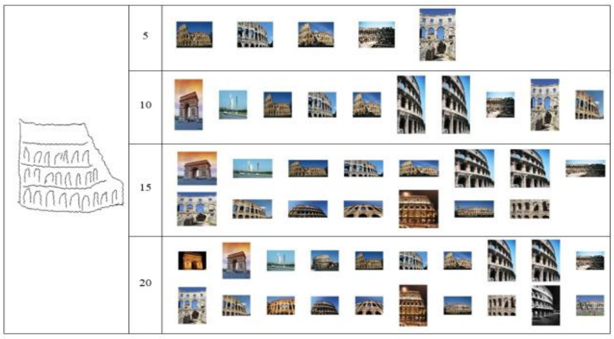

Fig.7. Retrieval results presents in top 5,10,15,20 by propose non-overlapping grid approach

C. Overlapping grid approach

The global contour image is resized to 410×410 pixels and is divided into 100 sub images of size 10 x 10 with 20% overlapping of left side and upper side portions of each block. The contour image overlaid by the overlapping grid presents in fig.8.

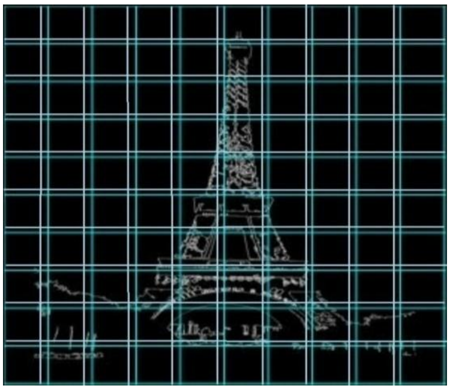

Fig.8.Example of overlapping grid placing over the contour image

The process involved in feature extraction of natural images is also applicable to query sketch, where the sketch is resized to 410×410 pixel. Feature extraction by overlapping grid shown in fig.9.The Euclidean distance is used to measure the distance between sketch feature and database feature. Based on the minimum Euclidean distance measure, images in the order of ascending distance values are then retrieved from the database.

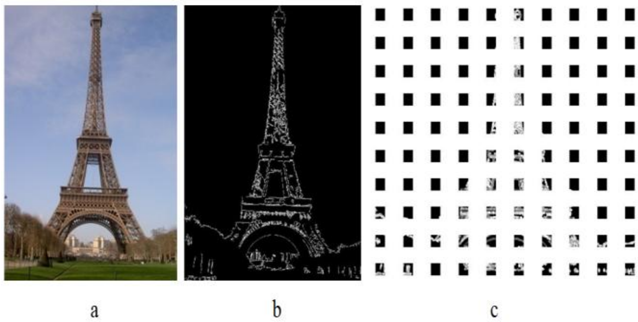

Fig 9: Feature extraction using overlapping grid approach. (a) Original image (b) global contour (c) feature extraction

The retrieval results are presented in four steps top 5,10,15,20. 5 images are matched in top 5 images, 8 images are matched in top 10, 13 images are matched in top 15, and 18 images are matched in top 20 shown in fig.10.

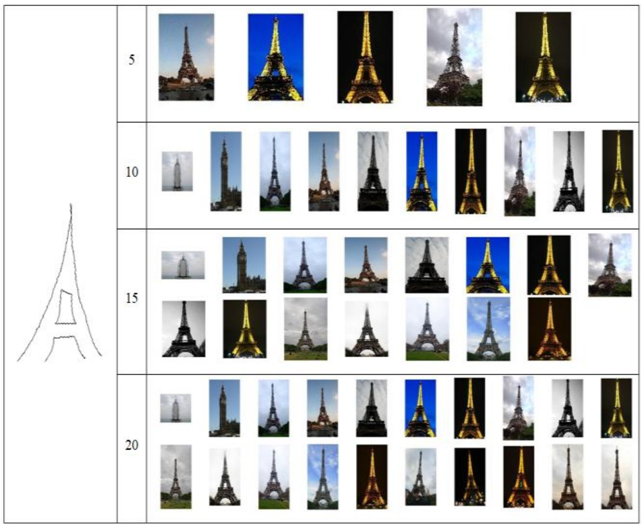

Fig.10. Retrieval results presents in top 5,10,15,20 by propose overlapping grid approach

D. Weighted Based

Weighted combination of the features computed as described in previous two sections is used for image retrieval [26]. The feature vectors computed using non-overlapping and overlapping grid over the contour image, defined in section 4.1 and 4.2, are weighted appropriately, weight is being assigned to both feature vector in the range of [0,1].

The feature set defined as S_n^m where the number of dataset images defined by n and m denotes in case of different methods. Non-overlapping grid-based features are defined as S_n^1 and overlapping grid features defines as S_n^2. In other end we extracted the sketch image features and it is defined as S^m, Let hyp1(n) and hyp2(n) denotes the similarity of non-overlapping grid based S_n^1 and S^1 and the similarity of the overlapping grid S_n^2 and S^1. Now we can measure the similarity of sketch image and database image by using Non-overlapping grid based and overlapping grid-based method as follows:

$$S(n) = w \times S1(n) + (1-w) \times S2(n), \qquad n-1,2,.......,N \qquad 1$$

The similarity of features is depending on the weight, weight assigned to the feature vector is denoted by *w* and number of dataset images denoted by N. In our experiment weight is set to w=0.7. Finally, similarity score is sorted in ascending order to retrieve the top n-ranked images from the database. In top 5 resulting images all are matched, in top ten, 9 images are matched and one is mismatched; in top fifteen resulting images, 14 are matched and 1 mismatched, and 19 images are matched in top 20 images. The results are presents in Figure 11.

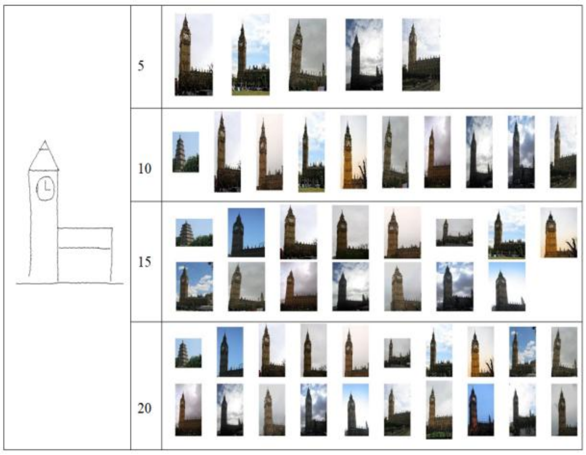

Fig.11. Retrieval results presents in top 5,10,15,20 by propose weighted based approach

V. RESULTS OF PROPOSE METHODS

The experimental results of proposed methods are based on minimum distance measured by using Euclidean distance. The retrieval results are analyzed as among top 5, top10, top 15, and top 20, respectively.

Non-overlapping grid, overlapping grid and weighted based approaches, respectively, yielded encouraging results and are presented in fig7, fig, 10 and fig11, respectively.

Table 1 presents Mean Average Precision (MAP) [27] on top 20 retrieval.

Table.1.Result of propose methods in MAP

Methods/MAP	MAP
Non-overlapping grid	0.130
Overlapping grid	0.145
Weighted based	0.153

VI. EXPERIMENTAL EVALUATION

The proposed method is evaluated on flickr15k dataset [23]. We have chosen 15k natural images and 330 sketch queries. Retrieval results are presented in section 4 where it is observed that weighted approach yielded better response. Mean average precision (MAP) is used to measure the retrieval performance of propose method.

$$MAP = \frac{1}{Q}\sum_{Q=1}^{Q} AP(q), \frac{1}{S}\sum_{i=1}^{N} p(k) * rel(k) \qquad 2$$

Where Q denotes number of Sketch's; S is number of similar images in retrieved images; N denotes total number of retrieved images; p(k) represents top-k precision score.

The proposed approaches are compared with existing approaches present in literature. Table.2 presents the comparison of the propose method with existing methods HOG [23], GF-HOG [23], SSIM [28] Structure Tensor [10], and Shape Context [29].

Table.2. Comparison of proposed methods with other methods based on MAP(top 20 retrieval)

Methods	MAP
Structure Tensor	0.079
Shape Context	0.081
HOG	0.111
GF-HOG	0.122
SSIM	0.095
Non-overlapping grid (proposed)	**0.130**
Overlapping grid (proposed)	**0.145**
Weighted based (proposed)	**0.153**

The comparison reveals that proposed methods yielded better MAP values compared to MAP values of other methods on flickr15k database and that performance of weighted base approach is better with MAP value of 0.153.

VII. CONCLUSION

An efficient and efficient approach for SBIR is presented in this paper. Edge features are computed using non-overlapping grid and overlapping grid approach. Euclidean distance measure is adopted for image retrieval. Experiments are preformed on flickr15k dataset. The proposed approaches successfully retrieve images related to input query from the large database. The method retrieved the most similar images from chaotic dataset as compared to the results of other methods available in the literature.

REFERENCES

1. Arnold W.M. Smeulders, Marcel Worring, Simone Santini, Amarnath Gupta, Ramesh Jain. "Content-Based Image Retrieval at the End of the Early Years", IEEE transactions on pattern analysis and machine intelligence, 2000 vol. 22, pp.1349-1380.
2. Yong Rui and Thomas S. Huang. "Image Retrieval: Current Techniques, Promising Directions, and Open Issues", Journal of Visual Communication and Image Representation, 1999, pp.39–62.
3. R. Datta, D. Joshi, J. Li, and J. Z. Wang. "Image retrieval: ideas, influences, and trends of the new age", ACM Computing Surveys, 2008,40(2):1–60.
4. H. H. S. Ip, A.K.Y. Cheng, W.Y.F. Wong, and Feng, "Affine-invariant sketch-based retrieval of images," in Proc. IEEE Int. Conf.Comput. Graphics, 2001, pp. 55–61.
5. CHANG N., FU K.: Query-by-pictorial-example. In the IEEE Computer Society's Third International Computer Software and Applications Conference, Proceedings. COMPSAC,1979, pp. 325–330.
6. HIRATA K., KATO T.: Query by visual example-content based image retrieval. In Proceedings of the 3rd International Conference on Extending Database Technology: Advances in Database Technology, Springer-Verlag London, UK, 1992 pp. 56–71.
7. Hirata K., Kato T. "Query by visual example-content based image retrieval", In Proceedings of the 3rd International Conference on Extending Database Technology: Advances in Database Technology, Springer-Verlag London, UK, 1992, pp. 56–71.
8. Rui Hu, Tinghuai Wang and John Collomosse. "A bag of regions approach to sketch based image retrieval", Proceedings from 18th IEEE international conference on image processing, 2011 pp.3661-3664.
9. M. Eitz, K. Hildebrand, T. Boubekeur, and M. Alexa. "A descriptor for large scale image retrieval based on sketched feature lines", Sketch Based Interfaces and Modeling, 2009 pp. 29-36.
10. M. Eitz, K. Hildebrand, T. Boubekeur, and M. Alexa. "An evaluation of descriptors for large-scale image retrieval from sketched feature lines", Computers and Graphics, 2010 vol. 34, pp. 482–498.
11. B. Szanto, P. Pozsegovics, Z. V amossy, Sz. Sergyan. "Sketch4Match – Content-based Image Retrieval System Using Sketches", IEEE Xplore, 2011, pp.183-188.
12. Jain A., Vailaya A. "Image retrieval using color and shape", Pattern Recognition,1996, 1233–1244.
13. X. Qian, Y. Zhao, and J. Han, "Image Location estimation by Salient Region Matching" IEEE Trans. Image Processing, 2015, vol.24, no.6, pp.4348-4358.
14. A.K. Jain, J.E. Lee, and R. Jin, "Sketch to photo matching: a feature-based approach," Proc. SPIE, Biometric Technology for Human identification VII, vol. 7667, pp. 766702–766702, 2010.
15. G.G. Rajput, Prashantha, Geeta B, "Face Photo Recognition from Sketch Images Using HOG descriptor" 2nd International Conference on Inventive Communication and Computational Technologies (ICICCT) IEEE Xplore, 2018, ISBN:978-1-5386-1974-2
16. J. Ashley, M. Flickner, J. L. Hafner, D. Lee, W. Niblack, and D. Petkovic, ―The query by image content (QBIC) system, in SIGMOD Conference, 1995, pp.- 475
17. J.R. Smith and S.F. Chang, ―Visualseek: a fully automated content-based image query system, in ACM Multimedia, New York, NY, USA, 1996, pp. 87–98, ACM.
18. E. Sciascio, M. Mongiello, and M. Mongiello, ―Contentbased image retrieval over the web using query by sketch and relevance feedback, in In Proc. of 4th Intl. Conf. on Visual Information Systems, 1999, pp. 123–130.
19. R. Hu, M. Barnard, and J. P. Collomosse, Gradient field descriptor for sketch based retrieval and localization, in ICIP, 2010, pp. 1025–1028.
20. M. Eitz, K. Hildebrand, T. Boubekeur, and M. Alexa, Sketchbased image retrieval: Benchmark and bag-of-features descriptors, IEEE TVCG, 2010, vol. 99.
21. D. Hoiem, A. A. Efros, and M. Hebert, Geometric context from a single image, in ICCV,2005, vol. 1, pp. 654– 661, IEEE.
22. A. Chalechale, G. Naghdy, A. Mertins, "Sketch-based image matching using angular partitioning", IEEE Transactions on Systems, Man and Cybernetics, Part A 35 (1) 2005 28–41.
23. Hu R. and Collomosse J, "A performance evaluation of gradient field hog descriptor for sketch based image retrieval," Computer Vision and Image Understanding, 2013, vol. 117, no. 7, pp. 790-806.
24. Jacobs, C. E., Finkelstein, A., & Salesin, D. H. Fast multi resolution image querying, in Proc.ACM SIGGRAPH, 1995, pp.-277-286.
25. Otsu, N "A Threshold Selection Method from Gray-Level Histograms" IEEE Transactions on Systems, Man, and Cybernetics, 1979, Vol. 9, No. 1, pp. 62–66.
26. YutingZhang, XuemingQian, Xianglong Tan "Sketch-based Image Retrieval Using Contour Segments", IEEE Xplore Xiamen, China,2015.
27. Fei Huang , Cheng Jin , Yuejie Zhang , Tao Zhang "Towards sketch-based image retrieval with deep cross-modal correlation learning," Proceedings of the IEEE International Conference on Multimedia and Expo (ICME) ,2017,pp-907-912
28. Shechtman E, Irani M "Matching local self-similarities across images and videos," CVPR, 2007, pp. 1-8.
29. Mori G., Belongie S. and Malik J "Efficient shape matching using shape contexts," TPAMI, 2005, vol. 27, no. 11, pp. 1832-1837.

AUTHORS PROFILE

G. G Rajput, Professor, department of computer science Akkamahadevi Women's University, Vijayapura Karnataka, India. His research interest includes Digital Image processing, Pattern recognition, and data analytics.

Prashantha, Research scholar department of computer science Rani channamma University Belagavi, Karnataka, India. Completed M.Tech from University of Mysore Karnataka, India, in 2013 and completed M.Sc from Gulbarga University Gulbarga, Karnataka India in 2011. Research area: Digital Image Processing and Pattern recognition.