Feature Analysis of Sketch Based Image Retrieval System

Kathy Khaing, Sai Maung Maung Zaw, Nyein Aye

University of Computer Studies, Mandalay,
University of Computer Studies, Mandalay,
University of Computer Studies (HpaAn)

DOI: 10.29322/IJSRP.8.5.2018.p7763
http://dx.doi.org/10.29322/IJSRP.8.5.2018.p7763

Abstract- Although many researches are increased in Sketch Based Image Retrieval (SBIR) field, it is still difficult to bridge the gap between image and sketch matching problem. Feature extraction is the critical role for SBIR to get efficient matching. In this paper, the proposed feature descriptor called Edge Orientation Histogram (EOH) for sketch based image retrieval (SBIR) system is presented. The features of database images and query sketch are extracted by EOH descriptor. And then cosine similarity measure is applied for matching features. The retrieved images similar with query sketch are displayed by rank order. Mean Average Precision (MAP) is measured as evaluation criteria. The Flickr15K benchmark dataset is used to evaluate the performance of this system.

Index Terms- EOH, IR, MAP, SBIR

I. Introduction

Due to the progress in digital imaging technology, image retrieval has become a very active research area in computer science. Moreover, searching desired image by describing hand-drawn sketch is popular, because of the emerging of touch screen technology. An image retrieval system basically consists of three components: preprocessing, feature extraction, and similarity measurement. The results of the search are usually shown as a list which is sorted according to a score given by a selected similarity measure.[1]

• Preprocessing: This component uses image processing techniques in order to get simpler representation of the input image that could be analyzed in an easier way by the subsequent steps. Some techniques applied here are noise elimination, edge detection, image enhancement, thresholding and morphological operations among others.

• Feature Extraction: This component is very critical in order to select appropriate features for the image retrieval task. These features are typically low-level features related to color, texture and/or shape. The set of features must possess a good or an optimal discriminatory power in the particular application.

• Similarity Measurement: The similarity measure is used to compute a distance value between the user query and a candidate match from the database. A good similarity measure is characterized by showing a low distance value when the underlying compared images are very similar and vice-versa, it must show a high distance value when the underlying images are very different.

Feature extraction is the main part in SBIR system. It extracts the visual information from the image and save them as the feature vectors in database. The feature extraction finds the image description in the form of feature values called feature vector for each pixel. These feature vectors are used to compare the query with the other images and retrieval. Features extracted from the whole image are called global features. Local features are extracted from an object or a segment of an image. Global features cannot provide enough information to estimate the similarity between images. Therefore, local feature descriptor is proposed in this system.

II. Related Works

This section describes the related literature and the concept of an image retrieval system. An image retrieval system is a computer system for browsing, searching and retrieving images from large databases of digital images.

In 2014, Prajakta A. More et al., presented an efficient Content Based Image Retrieval (CBIR) system using Sobel's edge detection algorithm [7]. Content Based Image Retrieval (CBIR) is a process to retrieve a stored image from database by supplying an image as query instead of text. This can be done by proper feature extraction and querying process. A universal content based image retrieval system uses color, texture and shape based feature extraction techniques for better matched images from the database.

Gang. H et al [9], proposed a novel image feature descriptor, namely texture structure histogram (TSH) for content-based image retrieval. This method using the color and edge orientation information to describe the image texture structure information. Considering the HSV color space conforms to humans' visual perception mechanism, the feature extraction is conducted in the HSV color space. The authors also proposed a novel non-equal interval quantization scheme that according to the different attribute and distribution of the *H*, *S* and *V* channels. Building texture structure map is also a key step, which provides a good feature representation for the texture information of image. TSH integrates the advantages of both color and texture features, and

shows a good retrieval performance. Their experimental results demonstrated that the proposed algorithm have a better performance for image retrieval.

Recently, content based image retrieval is becoming a way of fast retrieval. Users are not satisfied with the usual information retrieval techniques. Retrieving the relevant information from keywords is very major task and this technique has been focused by researchers in many ways. Sketch is more natural and informative, that breaking down the language barrier and one of the strong media for communication in old era as well as in current scenarios. Hence currently sketch based image retrieval (SBIR) is also in focus for research purposes [10].

In 2015, in order to increase the accuracy of image retrieval, an intent-based image retrieval system (IBIR) based on interactive genetic algorithm (IGA) is proposed [12] by Kumar K.R and Sindhu S.N. In the query stage of their work, the feature descriptors of a query image were extracted and then used to evaluate the similarity between the query image and those images in the database. In the evolution stage, the most relevant images were retrieved by using the IGA. IGA is employed to help the users identify the images that are most satisfied to the users' need.

In 2017, Tolias, G. and Chum proposed a novel concept of asymmetric feature maps (AFM), which allows evaluating multiple kernels between a query and database entries without increasing the memory requirements [13]. They also derive a short vector image representation that supports efficient scale and translation invariant sketch based image retrieval system. The authors states that the performance is further boosted by image-based average query expansion combined with AFM for object outline localization.

Bozas, K. and Izquierdo proposed a large scale sketch based image retrieval using patch hashing [11]. Sketch recognition and designing of descriptor are included in their system. The system in this paper is closely related to that paper. The differences are feature descriptor and similarity measure. The result of our system is compared with theirs. In this work, the feature descriptor is proposed and cosine similarity measure is used for feature matching.

This is the extended work of our prior work in [15]. The performance of this work is outperforming the prior work and the experiments also emphasize on the feature analysis of our proposed descriptor.

III. Proposed SBIR System

The aim of this paper is to develop a sketch based image retrieval system, which can retrieves images quickly, easily and efficiently. Therefore, the SBIR system is presented by proposing the feature descriptor.

The proposed feature descriptor called Edge Orientation Histogram (EOH) is a local approach that is based on computing a histogram of edge orientation in each block and matching them for the whole image. The features of database images and query sketch are extracted by EOH descriptor. And then cosine similarity method is applied for matching features. The system architecture of this approach is described in figure 1.

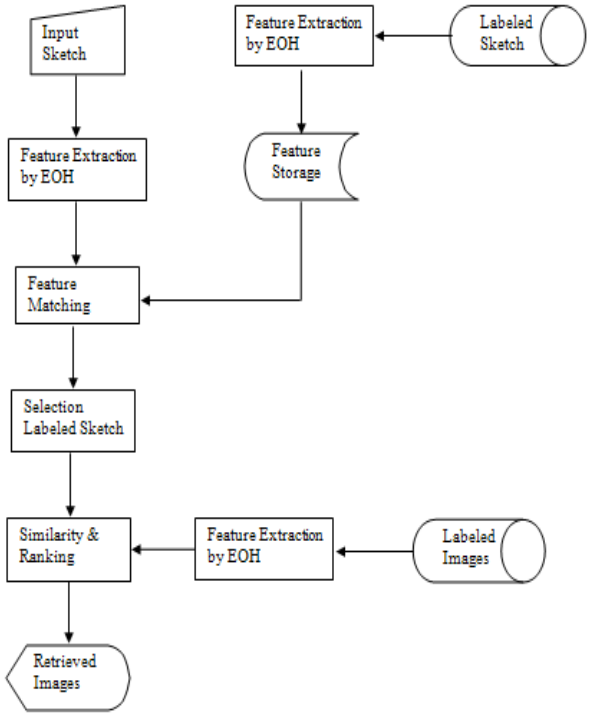

Fig. 1System Architecture of EOH Approach

In this work, sketches are simple curve drawings and define the boundaries of an object or a scene. There are 10 folders of 33 sketches in dataset. The sketches of 5 folders are trained and that of the other 5 folders are tested. Therefore, the features of 165 sketches are pre-extracted by the proposed descriptor and stored in the file. On the other hand, the all (14660) images in the dataset are labelled with categories and stored. Then, the features of all labelled images are also extracted by the proposed descriptor. When the query sketch is inputted, the features of the sketch are extracted by the proposed descriptor (EOH). The next step is matching the features of both. Cosine similarity is used for matching. Finally, the retrieving images are displayed by rank.

A. Preprocessing

The input sketch and all of database images are resized to 240x320 pixels. Then, the edges of resized images are detected by Canny. The edge images are divided into various blocks: 4x4, 6x6, 8x8, 9x9, 10x10, 12x12, 14x14, 16x16, 17x17, 20x20 and 21x21 respectively. Finally, the features are extracted by EOH descriptor for each block.

B. Proposed Edge Orientation Histogram (EOH) Descriptor

EOH descriptor considers the orientation only, magnitude of edge is neglected. It contains:

1. Masking
2. Edge detection by Canny

3. Cropping into 3*3 kernel

Five types of edge are defined for masking such as horizontal mask, vertical mask, 45˚ diagonal mask, 135˚ diagonal mask and non-directional mask respectively.

Fig.2. 5 types of Edge

The Canny edge detector uses a filter based on the first derivative of a Gaussian, because it is susceptible to noise present on raw unprocessed image. The raw image is convolved with a Gaussian filter [8]. The result is a slightly blurred version of the original which is not affected by a single noisy pixel to any significant degree.

Gaussian operator is:

$$g(x, y, \sigma) = e^{\frac{-(x^2+y^2)}{2\sigma^2}}$$

In this system, 5×5-size Gaussian mask with σ = 1.4 is used. Threshold value of Canny is [0.25, 0.6] for weak and strong edge respectively.

The final stage of EOH descriptor is cropping the edge image into 3*3 kernel size and the features are extracted by mapping various bin histograms. The best result is got by 8 bin histogram. EOH descriptor is used in three phase: for training sketch and images, for feature extracting of query sketch and that of database images.

C. *Similarity Measure*

The extracted features are matched by measuring cosine similarity. The cosine similarity between two vectors is:

$$cos\theta = \frac{\vec{a}.\vec{b}}{\| \vec{a} \| \| \vec{b} \|}$$

In this system,

cosine = AdotB/(A * B)

where, A = sqrt(sum(skfeature)^2)
B = sqrt(sum(imfeature)^2)

IV. Experimental Evaluation

To evaluate the performance of the proposed SBIR System, the Mean Average Precision (MAP) rate is calculated. The Precision provides information related to effectiveness of the system. Consequently, the mean average precision (MAP) is the mean of the average precision scores over all queries:

$$MAP = \frac{1}{Q}\sum_{q \in Q}^{NG\,(q)} \mathrm{AP}(q)$$

Where, Q= no: of queries images displayed with similar shape accuracy of the system. To calculate the MAP, average precision is firstly calculated. Average Precision is averaging the precision values from the rank positions where a relevant image was retrieved. As a global estimate of performance using a single value, it is standard to use the average precision (AP). The average precision for a single query q is the mean over the precision scores at each relevant item:

$$AP\,(q) = \frac{1}{NG\,(q)} \sum_{k=1}^{NG\,(q)} p_q(R_k)$$

where R_k is the recall after the k^{th} relevant image was retrieved. In this work, 5 categories of sketches are defined depend on their MAP value. Sketch number 1, 3, 12 and 20 are included in the category C5 and their MAP value is 1. The MAP value of category C1 is 0.8 and consists of sketch number 5, 6, 8, 14, 15, 16, 21 and 23. Sketch number 2, 4, and 24 are comprised in the category C3 and their MAP value is 0.6. The MAP value of C2 category is 0.4 and sketch number 10, 17, 18, 19, 27 and 29 are included. The category C4 sketches are 9, 11, 22, 25, 28, 31 and 33 and their MAP value is 0.2. The sketch categories and their related MAP values are illustrated in the figure3.

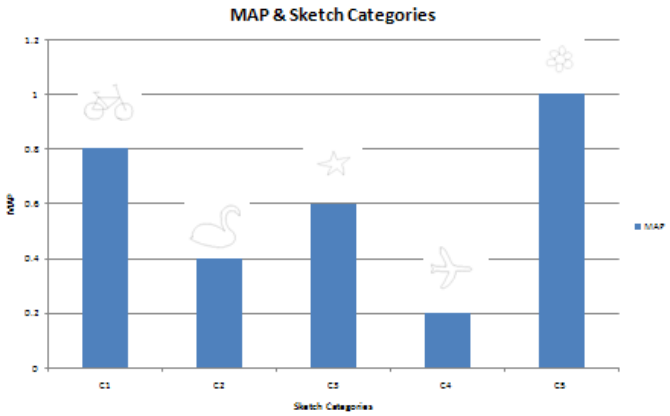

Fig 3: MAP Score with related sketch categories

A. *Dataset*

Flickr15k dataset is used to evaluate the performance of the proposed SBIR system. Flickr15k dataset consists of approximate 15k photographs and manually labeled into 33 categories based on shape, and 330 free-hand drawn sketch queries drawn by 10 non-expert sketchers. In this system, the half of sketch and database images are trained by EOH and tested by the other half.

B. *MAP Comparison over various Block sizes*

The experiments are analyzed over various block sizes and the comparison of MAP score on each block size to the PH-HOG descriptor is also illustrated in the figure 4 and table 1. The MAP value over all the sketch queries in the testing set is 0.485 and comparison with that of the state-of-the-art methods is described in table 2.

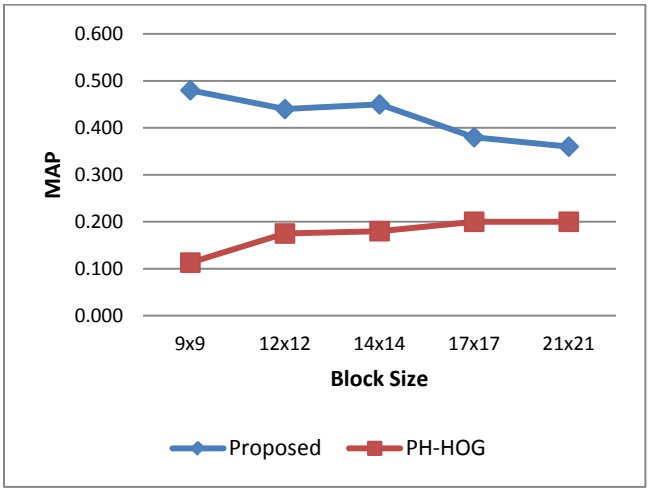

Fig 4: MAP Score Comparison over various block sizes

TABLE 1. MAP COMPARISON WITH THE PH-HOG

Block Size	**Proposed Method (EOH)**	**PH-HOG**
9x9	0.48	0.113
12x12	0.44	0.175
14x14	0.45	0.18
17x17	0.38	0.2
21x21	0.36	0.2

For each sketch, testing time is at least 5times in this work. Therefore, the MAP score is calculated over 165 sketches while the number of wrong sketches is counted over 33 sketches.

Table2. SBIR Result Comparison (MAP)

Methods	**Vocabulary Size**	**MAP**
EOH (Ours)	-	0.4848
PH-HOG	-	0.200
Siamese CNN	-	0.1954
Perceptual Edge	Non-BoW	0.1513
GF-HOG	3500	0.1222
HOG	3000	0.1093
SIFT	1000	0.0911
SSIM	500	0.0957
Shape Context	3500	0.0814
Structure Tensor	500	0.0798

V. CONCLUSION

In this paper the sketch based image retrieval system is presented which is developed to retrieve desired image efficiently. The feature descriptor called Edge Orientation Histogram (EOH) is proposed for feature extraction. The feature analysis is tested over various block sizes and histogram bins. 8x8 block size with 8 bin histogram get the best MAP result. Cosine similarity measure is used for feature matching. The half of and sketches are trained by EOH and then tested on the other half of both. The result is depend on the quality of input sketch.

VI. FUTURE WORK

The success of the proposed EOH descriptor is dependent on the quality of input sketch and neglected the magnitude of edge feature. Since the both of input sketch and the database image are resized in the preprocessing stage, there is no scaling invariant. Future work should focus on the preprocessing steps of SBIR system. Before feature extraction, background and foreground of the image should be segmented. Moreover it should be extending EOH descriptor to be invariant to scaling.

REFERENCES

[1] Datta, R., Joshi, D., Li, J. and Wang, J.Z., 2008. Image retrieval: Ideas, influences, and trends of the new age. ACM Computing Surveys (Csur), 40(2), p.5.
[2] Rafael Gonzalez and Richard Woods. Digital Image Processing. Pearson Prentice Hall, New Jersey, USA, third edition, 2008.
[3] Desai, M.A.A., Shinde, M.A.S. and Malathi, M.P., Sketch Based Image Retrieval System. International Journal of Advanced Research in Computer Engineering & Technology,3(9), pp. 3016-3020.
[4] Belongie, S., Malik, J. and Puzicha, J., 2002. Shape matching and object recognition using shape contexts. IEEE transactions on pattern analysis and machine intelligence, 24(4), pp.509-522.
[5] Lim, J.J., Zitnick, C.L. and Dollár, P., 2013. Sketch tokens: A learned mid-level representation for contour and object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 3158-3165).
[6] KR, D. and Hariharan, S., 2014. Approaches and Trends in Content based Image Retrieval. Int. Conf. on Advances in Communication, Network, and Computing, CNC
[7] More, P.A. and Nalavade, K., 2014, January. Content Based Image Retrieval Using Sobel's Edge Detection Algorithm. In International Conference of Advance Research and Innovation (ICARI), ISBN (pp. 978-93).
[8] John Canny. A computational approach to edge detection. IEEE Tranactions on Pattern Analysis and Machine Intelligence, 8(6):679_698, 1986.
[9] Gang, H., Qinghe, F., Xiaoxue, Z., Jun, K. and Ming, Z., 2013, November. Content-based image retrieval using texture structure histogram. In 3rd International Conference on Multimedia Technology (ICMT-13). Atlantis Press.
[10] Aishwarya D. and Dr.B.L. Gunjal., A survey on sketch-based image retrieval technique. International Journal of Advance Research and Innovative Ideas in Education, 2(2):337 _341, 2016.

[11] K. Bozas and E. Izquierdo, "Large scale sketch based image retrieval using patch hashing," Adv. Visual Comput., vol. 7431, pp. 210–219, 2012.
[12] Kumar, K.R. and Sindhu, S.N., 2015.Image Retrieval by using digital image processing and GA. In International Journal of Engineering & Science Research 5(5).
[13] Tolias, G. and Chum, O., 2017. Asymmetric Feature Maps with Application to Sketch Based Retrieval. arXiv preprint arXiv:1704.03946.
[14] Chatzichristofis, S.A., Iakovidou, C., Boutalis, Y.S. and Angelopoulou, E., 2014. Mean Normalized Retrieval Order (MNRO): a new content-based image retrieval performance measure. Multimedia tools and applications, 70(3), pp.1767-1798.
[15] Kathy.K et al. , 2017. Sketch based Image Retrieval System based on Block Histogram Matching. 14th International conference on Modeling, Simulation and Visualization Methods (MSV '17), pp. 94-99.

AUTHORS

First Author – Kathy Khaing, Ph.D candidate ,University of Computer Studies, Mandalay, kathrine.truth@gmail.com

Correspondence Author – Sai Maung Maung Zaw, Lecturer, University of Computer Studies, Mandalay, saisaimmz@gmail.com

Correspondence Author – Nyein Aye, Professor, University of Computer Studies (HpaAn), nyeinaye@gmail.com