Volume 8, No. 3, March – April 2017

ISSN No. 0976-5697

International Journal of Advanced Research in Computer Science

RESEARCH PAPER

Available Online at www.ijarcs.info

Sketch Based Image Retrieval with Cosine Similarity

Y. Jhansi
Research Scholar,
ANU College of Engineering
Acharaya Nagarjuna University, India

Dr. E. Sreenivasa Reddy
Professor &Principal
ANU College of Engineering
Acharaya Nagarjuna University, India

***Abstract*:** In the present paper cosine similarity is used for similarity measurement in sketch to image retrieval with focus on image matching. Similarity measurement is very challenging task in sketches due to its variable dimensions in the edges. Median filters are used to minimize the noises occur in the images due to unknown reasons. After pre-processing of the images in the benchmark database, the features are extracted using Histogram of oriented gradients (HOG). Canny edge detection algorithm is discussed to detect the image edges in several directions. Later the sketch to image matching is calculated by similarity measurement algorithm cosine similarity. A Benchmark sketch image database is considered to prove the proposed algorithms is efficient and works effectively when compared with many existing techniques. It is very important to measure accurate similarity measurement in high dimensional space for information retrieving. This paper presents efficient cosine similarity measurement algorithm. In the present exertion extensive experimental work is carried out inorder to prove the state of art in association with Sketch Based Image Retrieval (SBIR).

***Keywords*:** Image retrieval, Cosine similarity, HOG, Canny edge detection, Median filter

I. INTRODUCTION

The query based visual mining is one of the dominant eras in the recent image retrieval context. There are many methods has been involved effectively in image retrieval to produce accurate results. Content Based Image Retrieval (CBIR) called as Query by Image Content is one of the applications of image retrieval. CBIR has many advantages over traditional image retrieval approaches. In traditional approaches the image is retrieved with metadata. The metadata includes some keywords regarding images. But in this approaches human intervention highly demanded to match the similar image from the database. It is also possible to miss the images that cause less accuracy. It takes lot time and difficult to identify similar image if it retrieving from large database. On the other hand, in the CBIR the search analyses the content rather than metadata. The content include colour, shapes, textures, or any other parameter related to the image [2][3][4]. These features are either used in combination which allow easier interface with the input given for information retrieval. In CBIR as per the application, different algorithms applied for image retrieval. But, the resultant image may not have common query elements provided [5][14]. In the initial days the CBIR used few contents and partially or entirely depends on user feedback. Later iterative techniques are applied to make more intelligent than the previous technique [6].

In the present paper multidimensional spaces are considered for distance estimation, which a prime limitation in CBIR. The Sketch based modelling is used in 3D graphics of visual mining. In sketch base 3D model the user draws 2D image with graphics tools and this will be automatically converted into 3D image after pre-processing. In traditional 3D modelling and other 3D modelling, it is very difficult and lengthy for the users to create 3D images. In recent days, many researchers are pugnaciously working to address efficient algorithms in association with Sketch Based Image Retrieval (SBIR). In the sketch based image retrieval methodology the hand drawn sketches being a natural way of representing and exploring a synthetic query. Most of the existing techniques are failed in obtaining the inherent properties of an image. The inherited complexity in images can be easily subjugated in case of sketch based image retrieval [1]. In 3D model the user need not to have any idea on original 3D image and 3D programs. Sketch is a collection of approximate bit vectors of similar search. Hence in SBIR, sketch vector with compact bit features are used instead of original feature vector [19].

The SBIR has many applications and has demand in many sectors. Actually the SBIR came in to the image retrieval profile from the blaze of touch screen devices. The SBIR also extends its proliferation into many sectors like defence, social security, crime and medical investigations, detection of copy right infringement, face and finger print identification and etc. In security the hand drawn images are need to be analysed and appropriate matching must be carried out to identify the burglar. The hand drawn images are to be analysed carefully when compared with digital images[8]. In hand drawn images two similar images are diversified in their dimensions like translation, scaling and rotation. Many existing techniques are effectively involved in identifying the similar image by retrieving from large database with minimum time complexity. The measure of image retrieval is demarcated with precision and recall.

In traditional distance measurement scheme like L1 distance [15] and cosine similarity [16], each bit in feature vector of sketches is determined by the position of feature vector with reference to random hyperplane. The distance of such sketches is measured by hamming code. Hamming code implementation is very simple and the filtering process is faster than scanning of feature vectors. But it is a crude method for approximation of distance measurement. The accuracy is also very low in compact sketches. There are many distance metric algorithms effectively involved to improve the accuracy in distance measurement.

In distance measurement two sketches are compared to estimate the distances. This phenomenon is called symmetric. If the distance estimation performed in two different spaces then it is called asymmetric estimators. In the present paper a novel symmetric estimator is proposed to compare the queried feature vector and a sketch of a data object.

Hierarchical k-medoids: In this method the databased images are pre-processed to frame the chains. These chains are represented by connections of segments which contains object of boundaries. This is implemented in the following two methods.

a) Extracting long chains in segment network
b) Extracting boundaries of segment

The chains similarities are computed by fast Dynamic Programming based partial matched algorithm. This method produces Hierarchical k-medoids based indexing structures. These structures help to produce all chains of segments of all database images. Finally geometric verification is done to verify multiple chain matches inorder to improve the results [7].

Mahalanobis distance metric: It is used in cluster base image retrieval method. This algorithm minimizes the square distances between similarly labelled images. For differently labelled inputs a lower bound is assigned on sum of distances [9].

Neighbourhood Component Analysis: It is a distance metric learning algorithm which is developed to overcome the snags in kNN classification technique. The algorithm is Mahalanobil distance learning algorithms implemented on a trained data set by minimize the leave-on-out cross validation error of kNN classier. The distance to probability is converted with the help of softmax activation function. Hence the gradient computation is expensive because of softmax activation function [10].

Weinberger method: This algorithm is also used to improve the kNN classification. This algorithm is projected to learn a matrix design to further enhance the performance of kNN methodology. The object function is used to minimize the distance between neighbouring targets and correspondingly realizes that the distance is one unit closure between target neighbours and other classes. It may not be applicable for achieving good accuracy results in labelled image retrieving [11].

Davis method: With a prior knowledge of Mahalanobis distance and wide range of constraints a theoretical approach is proposed to learn Mahalanobis distance. This method regularizes the matrix design for further enhancement of distance metric measurement. The closeness between two Gaussian distributions of two neighbours are measured in Kullback-Leibler divergence between two Gaussian distributions [12][13].

Cosine Similarity: Cosine similarity is the angle measurement between two images on the vector space. The metric measurement is a measurement of orientation, and not magnitude. The cosine similarity is measured with dot product represented in equation 1.

$$\cos\theta = \frac{\vec{a}.\vec{b}}{\|\vec{a}\|.\|\vec{b}\|} \qquad \text{----------- (1)}$$

Where a and b are vector attributes,

The cosine similarity measurement delivers the metric about, how similar the two images are equal in angle instead of scaling. If the angle is ‘1’ it represents the images are in same direction otherwise images are in different directions.

II. PRE-PROCESSING

The feature vectors are pre-processed, so that the data objects analogous to query can be quickly find out. The compact sketches are scanned based on the query to generate small sets of vectors very quickly. They are ranked afterwards to obtain the accurate search results. This is often called filtering. The key challenges in SBIR are to obtain accurate distance estimation and accuracy in sketch size. The sketches are having smaller magnitudes than their feature vectors, which makes easy and enhances the search speed and optimizes the space significantly [15][17][18]. In pre-processing step the noise is removed to improve the signal in the later process. The following discussion supports the reason why Median filter is mostly used in image processing.

Median Filtering

In the present work median filter is used to reduce the noise in the image. Median filter is a nonlinear digital filtering technique, often used in reducing some kind of noise occurs on signal or image. After certain stage of processing, some kinds of edges are left out on the images while removing noises. The median filter works better for small to moderate Gaussian errors than Gaussian blur while removing noise. It shows effective results in speckle noise and salt and pepper noise [21].

Canny Edge Detection: canny edge detector widely used while wide ranges of edges are to be detected in images. Canny edge detection is used to extract the use full structural information from the objects and reduces the data size. The canny edge detection algorithm follows in 5 steps [22].

1. Apply Gaussian filter to remove the noise. This step will smooth the image by reducing the noise which causes a great effect on edge detect results. Here it is very important to consider the size the Gaussian kernel which may affect the performance of detection.
2. Measure the intensity gradients of the image. Generally an image may contain errors in different directions. Hence the Canny Edge detection algorithm is used four filters to reduce the noise at four different directions. The edge direction and gradient are determined from the edge detection operator [23].
3. Apply non-maximum suppression to gradient magnitude for thin edges. The algorithm is applied every pixel of gradient pixel. In this step, compare the strength of current pixel with the strength of positive and negative gradient pixels. If the strength of current pixel is greater than other pixels in the same direction, then preserve it otherwise suppress it.
4. After step 3 the image contains more clear edges. Measure possible edges occur in the smooth image by applying double threshold. In this step the preserve edge pixels filtered with high gradient value and out edge pixels with weak gradient value.

If the gradient value of edge pixel is greater than high threshold value it is preserved as strong edge pixel. If the gradient value is lower than higher threshold value and greater than lower threshold value it is treated as lower edge pixel. If the gradient value is lower than lower threshold value it will be suppressed [24].

5. Finally detect prompt edges by neglecting poor edges not connected to strong edges.

The performance of the canny detection algorithm can be improved by replacing Gaussian filter with adaptive filter to increase the accuracy.

III. FEATURE EXTRACTION

The image features are extracted after completion of noise reduction. The abstracted features are used to retrieve the image from query image. The feature vectors formed from these features are used for further processing. In further processing the file names are processed to differentiate the images and subsequently the global features are processed such as Edge Histogram descriptor, Histogram of oriented gradients, SIFT and SURF features. In the present paper, Histogram of Oriented Gradient (HOG) is discussed for feature extraction. HOG is a feature descriptor used for object detection in image processing [25][26][27].

IV. COSINE SKETCH

In many applications the Euclidean distance measuring technique is used for matching the user perception image. In Euclidean distance measurement there is a chance of uncertainty in similarity measurement in the process of image retrieval. The uncertainty occurs due to the mismatch between the computed features by algorithm and user professed image content. In the present work the similarity measurement is processed with cosine similarity measurement method [28][29]. The Cosine similarity is used in many applications like in distance measurement of document retrieval. In the present work cosine similarity is applied for sketch based data retrieval. For any vector a and b $\in R^n$, the cosine similarity is defined as shown in equation 2,

$$d_{\cos}(a,b) = \frac{a.b}{\|a\|.\|b\|} \qquad \text{------------- (2)}$$

The similarity measurement is always in the range between -1 to +1. This makes the objective function more simple and effective. The cosine similarity, CS(i,j) in the transformed subspace derived as shown in equation 3 [20],

$$CS(i,j,A) = \frac{(A_i)^T (A_j)^T}{\|A_i\|.\|A_j\|} = \frac{i^T A^T A_j}{\sqrt{i^T A^T A_i}\sqrt{j^T A^T A_j}}$$

-------------------- (3)

By substituting CS(i,j,A) with $d_{\cos}(a,b)$ where i and j are input vectors.

The sum of positive (Pos) and negative (Neg) sample index sets will give the total number of samples. When similarities are measured the linear transformation minimizes cross validation error. This is done by objective function f(A) which is shown in equation 4.

$$f(A) = \sum_{k\in Pos} CS(i_k, j_k, A)) - \alpha \sum_{k\in Neg} CS(i_k, j_k, A)) - \beta \|A - A_0\|^2$$

---------------- (4)

The first two factors in the objective functions combinely used to increase the positive and Negative samples margin to a large extent. This will help to minimize the training error. At the same time the third factor is used to regularise the matric *A.*

The asymmetric sketch distance is derived as shown in equation 5 [19].

$$d_\sigma^*(i,j) = \frac{1}{m}\sum_{i=0}^{m} |i_k . j| \times [\sigma_k(i) \oplus \sigma_k(j)]$$

$$= \frac{1}{m}\sum_{k=0}^{m} -sign(\rho_k . i) \times ((\rho_k . j) \times [\sigma_k(i) \oplus \sigma_k(j)]$$

-------------------- (5)

Where $\sigma_k(i) = \begin{cases} 0 \;\; if\, \rho_k . i < 0 \\ 1 \;\; if\, \rho_k . j \geq 0 \end{cases} \qquad \forall_k = 1, 2, ..., m$ ----

-------------- (6)

V. RESULTS AND DISCUSSIONS

Performance metric: The performance is measured in precision and recall. The precision is also represented as positive predicted value. This value signifies the fraction of retrieved value that is relevant to the input information. The recall is also called to sensitivity. The general meaning of sensitivity is the fraction of output responded with respect to a small change in the input. In the present context the sensitivity defined as the fraction of relevant indices retrieved. The precision and recall can be expression in equation format shown in equation 7 and 8,

$$precision = \frac{\text{Number of retrieved images relevant to the query image}}{\text{Total number of images retrieved}}$$

---------(7)

$$recall = \frac{\text{Number of retrieved images relevant to the query image}}{\text{Total number of relevant images in the database}}$$

---------(8)

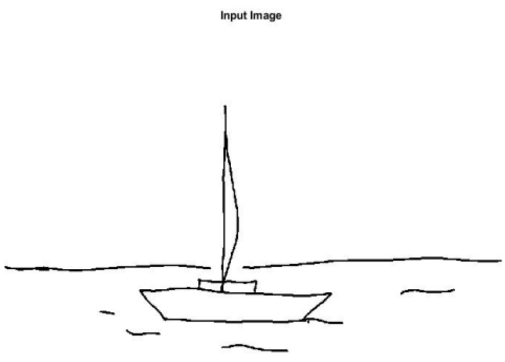

Figure 1: (a) Input Sketch

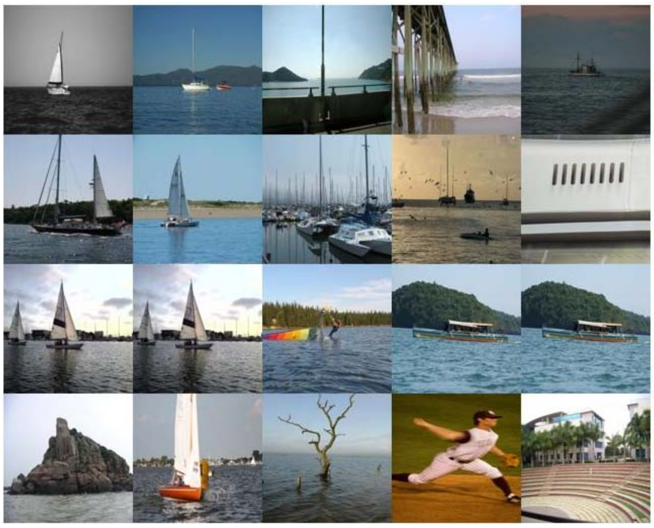

Figure 1:(b) Output images

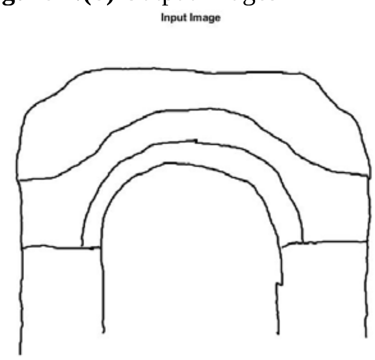

Figure 2: (a) Input Sketch

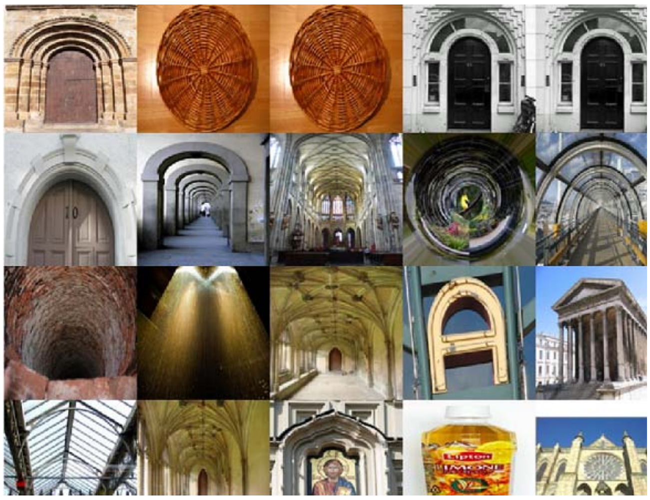

Figure 2: (b) Output images

The precision and recall is evaluated and compared with other methods and they are tabulated in table 1. The evaluated logical values of retrieval measurement are given in the table 1. In the proposed cosine similarity measurement method the precision value at particular retrieval count, is improved when compared with Euclidean distance measurement method. At the medium retrieval count it is observed good accuracy in cosine similarity measurement method.

Table 1: Comparision results of performance metric using Euclidean, Cross correlation and Cosine similarity methods

	Ours previous Method I [Euclidean distance[30]]		**Ours previous method II [Cross correlation[26]]**		**Proposed method [Cosine Similarity]**	
Retrieval Count	**Precision**	**Recall**	**Precision**	**Recall**	**Precision**	**Recall**
20	0.9402	0.65	0.89	0.63	0.951	0.51
30	0.8646	0.4215	0.851	0.324	0.832	0.435
40	0.7969	0.3790	0.764	0.25	0.784	0.415
50	0.7503	0.3790	0.6271	0.3298	0.764	0.347

In the present work the algorithm are implemented in MATLAB on the Eitz *et al.*dataset [30] mainly designed to measure retrieval performance in a large database. The images are conceded through various stages like filtering, pre-processing, feature extraction, and similarity measurement. For a given query image as shown in figure 1(a) and figure 2(a) the output images are obtained as shown in figure 1(b) and 2(b). The output images contain several similar images related to the human drawn sketch. Almost 90% images are similar to the query image given to the proposed cosine similar measurement system.

VI. CONCLUSION

Median filter is used for noise reduction and cannon edge detection is used after filtering. In this work a cosine similarity distance measurement is used for sketch retrieval from query. In the present work 90% of image retrieval and matching is obtained with cosine similarity measurement technique. Cosine similarity is very simple and in many cases it is used in content and face retrieval information. In this paper cosine similarity is proposed since it affords effective objective function.

VII. REFERENCES

[1] Dan Xu et al., Sketch-based Image Retrieval" , IEEE, ICPR'2016.

[2] Michael Lew et al., "Content based Multimedia Information Retrieval: State of Art and Challenges" ACM Transactions on Multimedia Computing, Communications, and Applications, pp 1-19, 2006.

[3] Eakins et al., "Content Base Image retrieval", University of Northubria at Newcastle, (https://en.wikipedia.org/wiki/Contentbased_image_retrieval).

[4] Rui, Yong; Huang, Thomas S.; Chang, Shih-Fu (1999). "Image Retrieval: Current Techniques, Promising Directions,

and Open Issues". Journal of Visual Communication and Image Representation. 10: 39–62.

[5] Datta, Ritendra; Dhiraj Joshi; Jia Li; James Z. Wang (2008). "Image Retrieval: Ideas, Influences, and Trends of the New Age". ACM Computing Surveys. 40 (2): 1–60.

[6] Cardoso, Douglas; et al. "Iterative Technique for Content-Based Image Retrieval using Multiple SVM Ensembles" (PDF). Federal University of Parana(Brazil). Retrieved 2014-03-11.

[7] Sarthak Parui et al., “Sketch-based Image Retrieval from Millions of Images under Rotation, translation and Scale Variations”, computer Vision and Pattern recognition, arXiv:1511.00099, Cornell university Library.

[8] Rong Zhou et al., “Sketch-based Image Retrieval on a Large Scale Database”, MM’12, ACM, October 29–November 2, 2012, Nara, Japan.

[9] Xing, E.P., Ng, A.Y., Jordan, M.I., Russell, S.: Distance metric learning, with application to clustering with side-information. In: Advances in Neural Information Processing Systems 15. Volume 15. (2003) 505512.

[10] Goldberger, J., Roweis, S., Hinton, G., Salakhutdinov, R.: Neighborhood component analysis. (In: NIPS), https://papers.nips.cc/paper/2566-neighbourhood-components-analysis.pdf.

[11] Weinberger, K., Blitzer, J., Saul, L.: Distance metric learning for large margin nearest neighbor classication. Advances in Neural Information Processing Systems 18 (2006), pp. 1473-1480

[12] Davis, J.V., Kulis, B., Jain, P., Sra, S., Dhillon, I.S.: Information-theoretic metric learning. In: ICML '07: Proceedings of the 24th international conference on Machine learning, New York, NY, USA, ACM (2007), pp. 209-216.

[13] Jacob Goldberger, Shiri Gordon, Hayit Greenspan, “An efficient Image similarity Measure based on a Approximations of KL divergence Between Two Gaussian Mixtures”,In proc. of ICCV 2003, Nice, October 2003,vol.1,pp.487-493.

[14] Shapiro, Linda; George Stockman (2001). Computer Vision. Upper Saddle River, NJ: Prentice Hall. ISBN 0-13-030796-3.

[15] Q. Lv, et al., “Image similarity search with compact data structures” CIKM ’04: Proceedings of the thirteenth ACM international conference on Information and knowledge management, pp. 208–217, 2004.

[16] M. S. Charikar, “Similarity estimation techniques from rounding algorithms”, STOC 02: Proceedings of the thiry-fourth annual ACM symposium on Theory of computing, pp. 380–388, 2002.

[17] Q. Lv et al., “Efficient filtering with sketches in the ferret toolkit”, MIR ’06: Proceedings of the 8th ACM international workshop on Multimedia information retrieval, pp. 279–288, 2006.

[18] Z. Wang, et al., “Sizing sketches: a rank-based analysis for similarity Search”, SIGMETRICS ’07: Proceedings of the 2007 ACM SIGMETRICS international conference on Measurement and modeling of computer systems, pp.157–168, 2007.

[19] Wei Dong et al., “Asymmetric Distance Estimation with Sketches for Similarity Search in High-Dimensional Spaces”, ACM SIGIR’08, Jul 2008.

[20] Hieu et al., “Cosine Similarity Metric Learning for Face Verification”, Conference Paper · November 2010, DOI: 10.1007/978-3-642-19309-5_55 · Source: DBLP

[21] E. Arias-Castro and D.L. Donoho, "Does median filtering truly preserve edges better than linear filtering?", Annals of Statistics, vol. 37, no. 3, pp. 1172–2009.

[22] Moeslund, T. (2009, March 23). Canny Edge Detection. Retrieved December 3, 2014.

[23] Zhou et al., “An Improved Canny Algorithm for Edge Detection” Journal of Computational Information Systems, 7(5), pp1516-1523, 2011.

[24] Otsu N, “A threshold selection method from gray-level histograms”, IEEE Trans Systems, Man and Cybernetics, vol9(1), pp. 62-66,1979.

[25] Jhansi Y, Reddy E S. “A methodology for sketch based image retrieval based on score level fusion”, In International journal of computer Applications, 2015, 109(3), pp. 9-13.

[26] Y. Jhansi et al., “An Efficient Sketch Based Image Retrieval using Cross-correlation”, IJCSIS, Vol 14. No 12, pp 445-451, Dec 2016.

[27] Dalal, N. and Triggs, B., “Histograms of Oriented Gradients for Human Detection,” IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005, San Diego, CA, USA.

[28] EUGENIO DI SCIASCIO et al., “Content-based Image Retrieval over the Web using Query by Sketch and Relevance Feedback”, Proceeding Springer-verlag, VISUAL '99 Proceedings of the Third International Conference on Visual Information and Information Systems Pages 123-130.

[29] E. Di Sciascio et al., “Query by Sketch and Relevance Feedback for Content-Based Image Retrieval over the Web”, Journal of Visual Languages & Computing Volume 10, Issue 6, Pp 565-584, December 1999

[30] Jhansi Y, Reddy E S.,” A new approach for Sketch Based Image Retrieval using FISH SWARM optimization with the aid of Optimal score level fusion,” In Indian Journal of Science and Technology , Vol 9,issue 41,November 2016,pp.1-9.