|| Volume 3 || Issue 3 || March 2018 || ISSN (Online) 2456-0774
INTERNATIONAL JOURNAL OF ADVANCE SCIENTIFIC RESEARCH AND ENGINEERING TRENDS

Image Retrieval Through Sketches Based on Descriptor

Dipika Birari[1], Kushal Mandge[2]
Assistant Professor, Computer Engineering, PGMCOE, Pune, Maharashtra[1]
Assistant Professor, E&TC, DYPSOE, Pune, Maharashtra[2]
dipikabirari001@gmail.com [1],mandgekushal@gmail.com [2]

Abstract: Sketch-based communication is no newer; it is the oldest form of writing. In which sketch shows rough or unfinished drawing of object. The photorealistic image shows attention to realistic detail. There is a large appearance gap in user sketches and photorealistic images therefore to bridge this gap is fundamental challenge in SBIR. The existence of noisy edges on photo realistic image degrades retrieval performance and to bridge this gap there is framework consisting of line segment descriptor and noise impact reduction algorithm. Proposed descriptor extracts edges and captures the relationship between those edges. Object boundary selection algorithm used to reduce the noisy edges for which the hypothesis is used to maximize retrieval score, for which multiple hypotheses are generated.

Keywords: *descriptor, sketch retrieval, edge based, histogram, line relationship.*

I INTRODUCTION

A sketch is swiftly accomplished freehand picture which serves various purposes, it might trace something that artist visualize, it might trace or increase an idea for later use or it might also be used as a rapid means of graphically representing an idea and an image. A sketch is rough or unfinished drawing, often made to assist in making a more finished picture. A style of painting that contains an attention to realistic detail. Although edge extraction can bridge the appearance gap between sketches and photo-realistic images to some extent, it is quite common for noisy edges from background clutter, object detail and texture to be extracted with the object shaping edges. These noisy edges usually widen the appearance gap and degrade retrieval performance. Therefore, retrieval performance can be enhanced if the impact of noisy edges is reduced. Retrieval performance of the human visual system is not sensitive to these noisy edges since humans are able to distinguish object boundaries or contour from noisy edges based on their inference ability. Using this fact, algorithm can select the object boundaries from all extracted edges, the appearance gap can be filled and the performance of SBIR can be improved. This motivation provides with a new pathway to improved performance, which imposes a new requirement, i.e., that sketches/extracted edges should be treated as a set of lines, and the descriptors should be able to capture line-level features. This is because line- based descriptors give the flexibility to achieve edge selection or removal by setting the corresponding parts of the feature vector to a certain value, which is critical for boundary selection. Beside the need to solve the noise problem, that an effective descriptor for SBIR should be designed to describe lines and their relationships, rather than describing image patches, since a sketch/object boundary is essentially composed of lines (strokes) and the shape is determined by the relationships between these lines.

II LITERATURE SURVEY

Eitz et al.[4] performed random sampling on images and then proposed the SHoG descriptor to describe each sampling point. Only the gradient value near the most dominant edge line is retained in SHoG. Hu and Collomosse[5] introduced dense gradient field on which they computed a multi-scale HOG feature (GF-HOG). GF-HOG is also utilized to describe regions which are generated by hierarchical image segmentation. Bozas [2] divided an image into overlapping patches and computed a HOG feature for each patch. In addition to HOG-based descriptors, Eitz utilized shape context to perform retrieval. Contour consistency filtering based on shape context descriptor was performed by Chen et al. Chalechale et al. performed angular partitioning on the edge image. Fourier transform was applied to achieve rotational invariance. Eitz [10] proposed a descriptor known as structure tensor, which was designed to find a single vector that is closest to the parallel direction of the majority of the edges in a local region. The MinHash method was used to build an index structure. In addition to HOG-based descriptors, Visual Saliency Weighting (VSW) was employed by Furuya and Ohbuchi [9] to emphasize the object of interest. Saavedra and Bustos represented sketches by six types of key shape. Zhou et al. extracted multi-scale features on candidate regions and built a hierarchical index structure to achieve coarse-to-fine retrieval.

III PROPOSED SYSTEM

Proposed system contains pre-processing of image from database. In pre-processing the photorealistic image is converted into edge image. And this edge image is then

|| Volume 3 || Issue 3 || March 2018 || ISSN (Online) 2456-0774
INTERNATIONAL JOURNAL OF ADVANCE SCIENTIFIC RESEARCH AND ENGINEERING TRENDS

converted into PLS image. Line segments and relationship between them shows the content of image. This conversion is used to bridge the gap between sketches and photorealistic images, for this the Canny Edge Detector is used on photorealistic images to extract strong edges. Figure 1 shows the edge extraction using canny detection. After finding strong edges the descriptor is designed to find the line relationship.

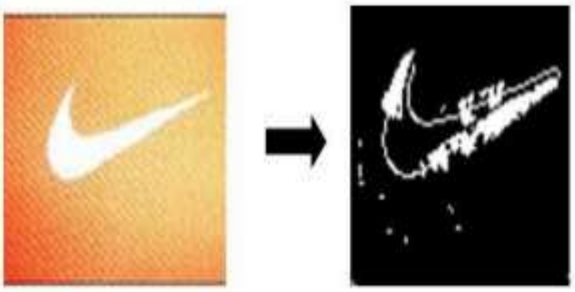

Figure 1: Preprocessing of image

Using this descriptor, it is able to selectively capture a subset of neighboring line segments rather than all of them, which makes it quite flexible and serve as the basis for noise impact reduction. In Object boundary selection the removing operation is get performed to remove certain neighboring edges by setting the corresponding bins to zero. Images after these operations are stored in knowledge base. Then hypothesis are retrieved. And matching is performed in between sketch image and hypothesis retrieved. Top retrievals are returned. Figure 2 shows architecture of proposed system.

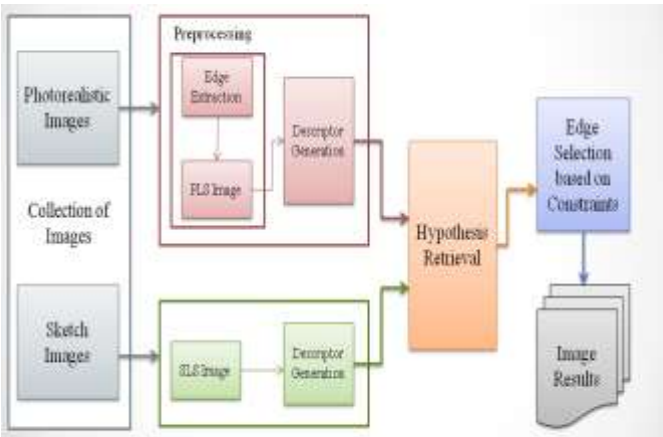

Figure 2: Block diagram of proposed system

A. Preprocessing

The aim is to develop detection that is optimal with criteria like Detection in which the probability of detecting real edge points should be maximized while the probability of falsely detecting non-edge points should be minimized. This corresponds to maximizing the signal-to-noise ratio. In localization the detected edges should be as close as possible to the real edges. In number of responses in which one real edge should not result in more than one detected edge. The Canny[8] edge detection contains

Smoothing: Whenever images are taken from a camera, it will contain some amount of noise. To prevent that noise, the noise must be reduced or minimized [11]. For that by applying a Gaussian filter, the image is first smoothed. The kernel of a Gaussian filter with a standard deviation of = 1.4 is used.

$$B = \frac{1}{159}\begin{bmatrix} 2 & 4 & 5 & 4 & 2 \\ 4 & 9 & 12 & 9 & 4 \\ 5 & 12 & 15 & 12 & 5 \\ 4 & 9 & 12 & 9 & 4 \\ 2 & 4 & 5 & 4 & 2 \end{bmatrix} \quad \dots (1)$$

Finding Gradients: [11] Where the grayscale intensity of the image changes the most, the Canny detector basically finds edges. For that the gradients of the image are calculated. Equation (2) and (3) shows kernel gradient with respect to x and y.

$$K_{GX} = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \quad \dots (2)$$

$$K_{GY} = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} \quad \dots (3)$$

The gradient magnitudes can be calculated as a Euclidean distance measure by applying the law of Pythagoras as shown in (4). G_x And G_y are the gradients in the x- and y directions respectively. Figure 3 shows gradient magnitude of image.

$$|G| = \sqrt{G_x^2 + G_y^2} \quad \dots (4)$$

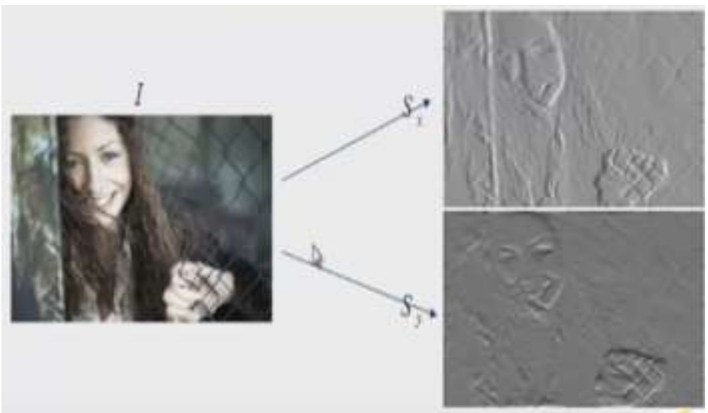

Figure 3: Gradient magnitude with respect to x and y

However, the edges do not indicate exactly where they are because of the edges are typically broad. To make it possible to determine this, the direction of the edges must be determined and stored as shown in (5). Figure 4 shows gradient magnitude of image.

Figure 4: The Gradient magnitude in smoothed image

Non-maximum suppression: This step is used to convert the blurred edges in the image of the gradient magnitudes to sharp edges. Figure 5 and 6 shows non maximum suppression. The algorithm is for each pixel in the gradient image [11]:

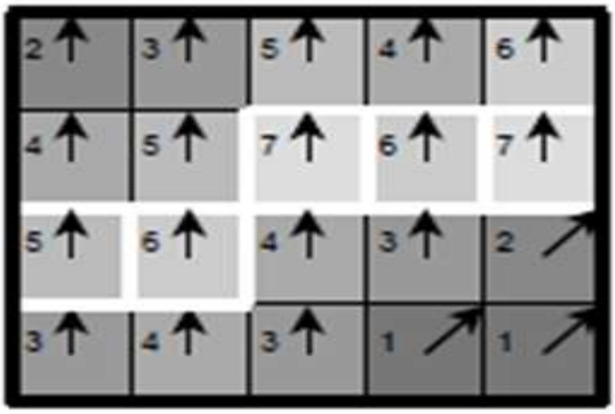

Figure 5: Illustration of non-maximum suppression

1. Round the gradient direction to nearest 45 degree, corresponding to the use of an 8-connected neighborhood.
2. Compare the edge strength of the current pixel with the edge strength of the pixel in the positive and negative gradient direction, i.e. if the gradient direction is north (theta = 90), compare with the pixels to the north and south.
3. If the edge strength of the current pixel is largest; preserve the value of the edge strength. If not, suppress (i.e. remove) the value.

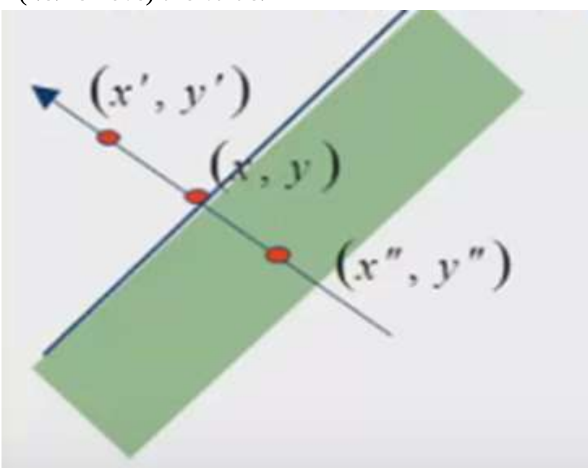

Figure 6: Non-maximum suppression

Double thresholding: The edge-pixels remaining after the non-maximum suppression step are marked with their strength pixel-by-pixel[11]. Here is use a thresh old , so that only edges stronger that a certain value would be preserved. The Canny edge detection algorithm uses double thresholding. Edge pixels stronger than the high threshold are marked as strong; edge pixels weaker than the low threshold are suppressed and edge pixels between the two thresholds are marked as weak. Following figure 7 shows double thresholding.

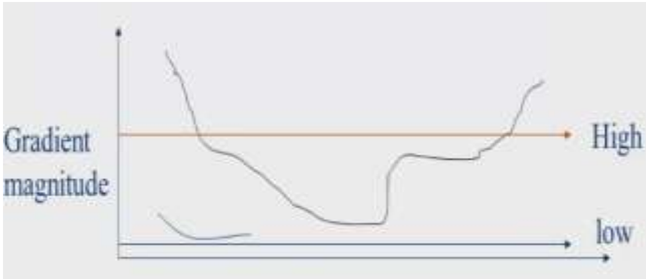

Figure 7: Double Thresholding

B. Descriptor Designing

A histogram of line relationship is centered on each line segment and rotated to the same angle. The size of the HLR is proportional to the length of the line segment. Define the blocks which divide the neighboring area into four areas that ensures that no neighboring line segment is missing so that we can capture the line relationship. Define extra blocks that will lie near two endpoints of the central line and covers other block boundaries. In object boundary selection, we remove the noisy edges using removing operation. It imitates the mechanism of human visual system in which it focuses on shaping edges and ignores the noisy edges and then will predict the shaping edges [14].

To capture the relationships between the two lines the descriptor is designed in such way that it divides area into four different parts as left, right, upper and lower. To cover the boundaries of blocks as well as two end points of center line, we have to define next four blocks. Therefore all the block boundaries and all the adjacent line segments are covered using this descriptor. This gives accurate description when line segments falls into block are spatially close to endpoints of middle line because of large angle difference between them and middle line.

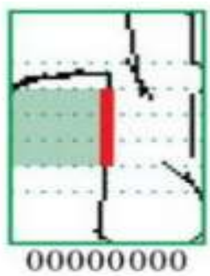

Figure 8: Descriptor Designing

|| Volume 3 || Issue 3 || March 2018 || ISSN (Online) 2456-0774
INTERNATIONAL JOURNAL OF ADVANCE SCIENTIFIC RESEARCH AND ENGINEERING TRENDS

Secondly it is robustness to sketch variation. Because when people sketch, they usually sketch casually, this makes the sketch variation quite significant. Consequently, given the central line, strokes near the block boundaries could fall into any one of the blocks due to sketch variation, resulting in two very similar shapes being described totally differently. The next implementation will be regarding to constraints which we have to apply for selection of boundary. By using edge length based constraints as well as coherent constraints we will find out the strong boundary selection algorithm which finds the shaping edges by removing the noisy edges using removing operation[12].

IV RESULTS

Proposed system use widely accepted evaluation metric Precision-Recall. Precision and recall is calculated in the terms of true positives, true negatives, false positives and false negatives as shown in equation (5) and (6).

Precision: It is nothing but percentage of photo-realistic images that are correctly selected with respect to sketch image that is ratio of true positive to true positive plus false positive.

Recall: It is percentage of photo-realistic images that are incorrectly selected that is ratio of true positive to true positive plus false negative.

$$\text{Precision} = \frac{True\_Positive}{True\_Positive + False\_Positive} \quad \ldots (5)$$

$$\text{Recall} = \frac{True\_Positive}{True\_Positive + False\_Negative} \quad \ldots (6)$$

Following images shows the partial implementation of proposed system containing preprocessing of image using canny edge detection and descriptor generation. Figure 9 shows the GUI of edge extraction. System shows extraction of edges and conversion of images into SLS and PLS images.

Figure 9: Canny edge detection for browsed image

Figure 10 shows the GUI of descriptor generation. System generates descriptor values as shown below and figure 11 shows the GUI of descriptor constructed by clicking on certain values.

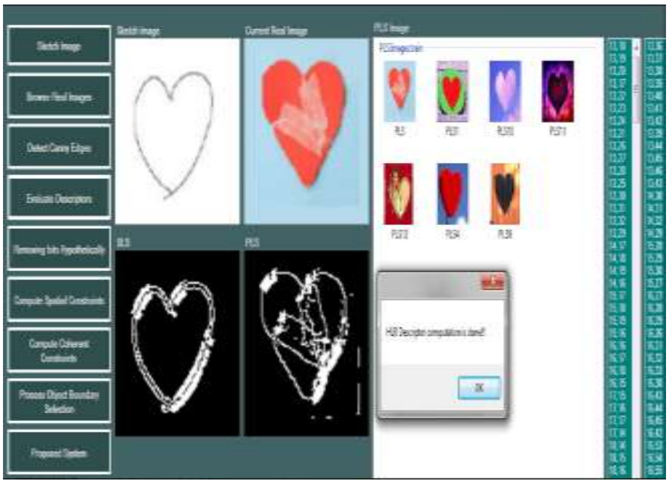

Figure 10: Descriptor generation of preprocessed image

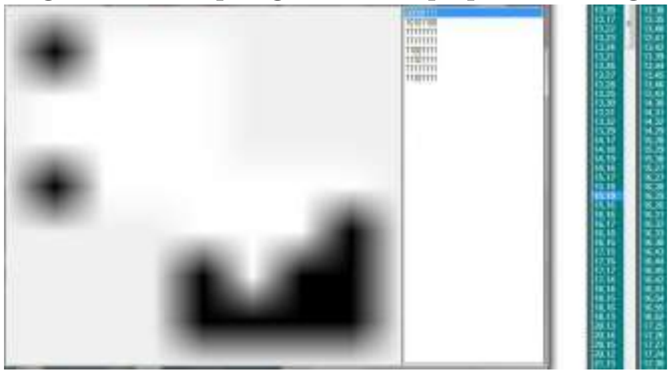

Figure 11: Descriptor structure

VI CONCLUSION

To bridge the large appearance gap in SBIR line segment based descriptor is used to describe the sketches and extracted edges by treating them as set of line segments. Descriptor captures the relationship between one line segment and all neighboring line segments, while in various system descriptor only captures the relationship between two connected line segments. To select the boundaries and to detect false matches edge length based constraint is applied which reduce the impact of noisy edges. The proposed system extracts the strong edges and convert it into line segments and then by applying descriptor, it enhance the performance of the image retrieval by ensuring that line relationship is captured. To reduce the false matches, constraints are applied which improves the retrieval performance significantly.

This proposed system aims to reduce the impact of noisy edges in the SBIR system. But descriptosr that is too small can only capture very local information, which makes the descriptor insufficiently discriminative; too large a descriptor size, however, makes the information captured by the descriptor contains too much variety, which results in loss of generalization. Therefore the system can be extended to decrease quantization errors in descriptor mapping an in future work this system can be extended to detect face sketches which can be useful in crime detection.

ACKNOWLEDGMENT

I would also like to express my appreciation and thanks to all my colleagues and family members who knowingly or unknowingly have assisted and encouraged me throughout my journey.

REFERENCES

1] Shu Wang, Jian Zhang, Tony X. Han, and Zhenjiang Miao, "Sketch-Based Image Retrieval Through Hypothesis-Driven Object Boundary Selection With HLR Descriptor," in IEEE transaction on multimedia, Vol. 17, No. 7, July 2015.

[2] K. Bozas and E. Izquierdo, "Large scale sketch based image retrieval using patch hashing," in Adv. Visual Comput., vol. 7431, pp. 210219, 2012.

[3] R. Zhou, L. Chen, and L. Zhang, "Sketch-based image retrieval
on a large scale database," in Proc. 20th ACM Int. Conf. Multimedia, 2012, pp. 973976.

[4] M. Eitz, K. Hildebrand, T. Boubekeur, and M. Alexa, "Sketch-based image retrieval: Benchmark and bag-of-features descriptors," in IEEE Trans. Vis. Comput. Graph., vol. 17, no. 11, pp. 16241636, Nov. 2011.

[5] R. Hu and J. Collomosse, "A performance evaluation of gradient field hog descriptor for sketch based image retrieval," in Comput. Vis. Image Understand., vol. 117, no. 7, pp. 790806, 2013.

[6] M. Eitz, J. Hays, and M. Alexa, "How do humans sketch objects," in ACMTrans. Graph., vol. 31, no. 4, pp. 44:144:10, Jul. 2012.

[7] R. Hu, T. Wang, and J. Collomosse, "A bag-of-regions approach
to sketch-based image retrieval," in Proc. IEEE Int. Conf. Image Process., Sep. 2011, pp. 36613664.

[8] J. Canny, "A computational approach to edge detection," in IEEE Trans. Pattern Anal. Mach. Intell., vol. PAMI-8, no. 6, pp. 679698, Nov.

[9] T. Furuya and R. Ohbuchi, "Visual saliency weighting and cross-domain manifold ranking for sketch-based image retrieval," in Proc. Int. Conf. Multimedia Modeling, 2014, pp.3749.

[10] M. Eitz, K. Hildebrand, T. Boubekeur, and M. Alexa, "A descriptor for large scale image retrieval based on sketched feature lines," in Proc. 6th Eurograph. Symp. Sketch-Based Interfaces Modeling, 2009, pp. 2936.

[11] "Canny edge detection," March 23, 2009.

[12] S. Salve and K. Jondhale, "Shape matching and object recognition using shape contexts," in Proc. 3rd IEEE Int. Conf. Comput. Sci. Inf. Technol., 2010, vol. 9, pp. 471474.

[13] J. Yao, M. Li, Z. Li, L. Zhang, and W.-Y.Ma, "Natural image retrieval with sketches," in Proc. IEEE Int. Conf. Multimedia Expo, Jul. 2005, pp. 11981201

[14] C. L. Zitnick, "Binary coherent edge descriptors," in Proc. 11th Eur. Conf. Comput. Vis.: Part II, 2010, pp. 170182.