International Journal of Computer Applications (0975 – 8887)
Volume 146 – No.12, July 2016

A Sketch based Image Retrieval with Descriptor based on Constraints

Dipika R. Birari
Department of Computer Engineering
Late G. N. Sapkal CoE, Nashik

J. V. Shinde
Department of Computer Engineering
Late G. N. Sapkal CoE, Nashik

ABSTRACT
To match sketch and real images some processing is required, because it is very difficult to directly match sketch and real image. Real images contain noise due to many reasons, which makes it very difficult for matching. For this descriptor is designed so that it can give best match by finding relationships between edges and line segments. By applying edge length as constraint, the retrieval performance is increases. Proposed framework tested on public datasets and results shows that proposed method improves SBIR performance significantly.

Keywords
Descriptor, sketch, real images, edge based, histogram, line relationship, shaping edges.

1. INTRODUCTION
Image Retrieval based on sketches is very valuable searching method, which provide ease of interface for searching about images. SBIR can act as a communication tool between peoples having different languages. Sketch provides more information and visualization as compare to keyword based searching. Such Sketch Based Image Retrieval (SBIR) [1] contains sketch as input and real images matching with sketches as output. But matching between sketch and real image is very difficult because their appearances are different in terms of features, because sketch contains user's imagination in terms of only boundary line that forms some kind of shape and real images contains reality such as color, texture and some shape. For matching such images, the extraction of edges in real images is required. After this, the strong edges are represented to make both, sketch and real image comparable. Some valuable work [1-3] gives this approach of edge extraction.

2. RELATED WORKS
In Sketch Based Image Retrieval, many descriptors are designed for matching. Work proposed by J. canny [2] is to reduce the amount of data in an image with preservation of structural properties that can be used for later image processing. It maximizes signal-to-noise ratio. Eitz et. al. [3] contains a method which divides an image into fixed number of cells and each cell represents to a tensor descriptor. Database index structure is not used, algorithm must scan whole database for each query. Bozas et. al.[4] assumes that users choice is spatially consistent with query, means if they draw a sunset sketch and sun is placed at top left then image in results with same spot for sun will be preferred. Therefore system is designed to retrieve near duplicate images. Min Hash estimates similarity between sets and applied on text and image domains contains pre-processing, feature extraction, index construction. Chen et. al. [7] two types of region are used, main region and ROI. To tackle the situation one image only contains one object similar to query but different in size and position. And ROI deals with one object similar to query saliently appears in a complicated background. Hu and Collomosse [9] used HOG descriptor which commonly applied in object recognition and detection tasks. HoG is window based descriptor which detects gradient and magnitude. For that they divide the image into number of cells and compute histogram of gradient direction and edge orientations. By grouping adjacent cells forms blocks and normalization of such blocks shows HoG Descriptor. But BoVW is challenging in this as sketch contains little or no texture and sketch defined by relative positions.

3. PROPOSED SYSTEM
Proposed system contains edge extraction using canny edge detection method and descriptor designing and selection of strong edges based on edge length. Proposed system is executed with Flickr15K dataset. Proposed system contains:

1. Edge Extraction
2. Designing of descriptor
3. Selection of edges

Proposed Architecture is shown in fig. 1, in which preprocessing contains edge extraction. Strong edges are extracted by applying canny edge detector for each real image in database and then convert it into RS image. Descriptor designing provides flexibility to selection or removal of edges by setting relative part to certain value and captures the line level features. Proposed work is in edge selection, in which edge length is applied as constraint to select the shaping edges and to reduce noisy edges.

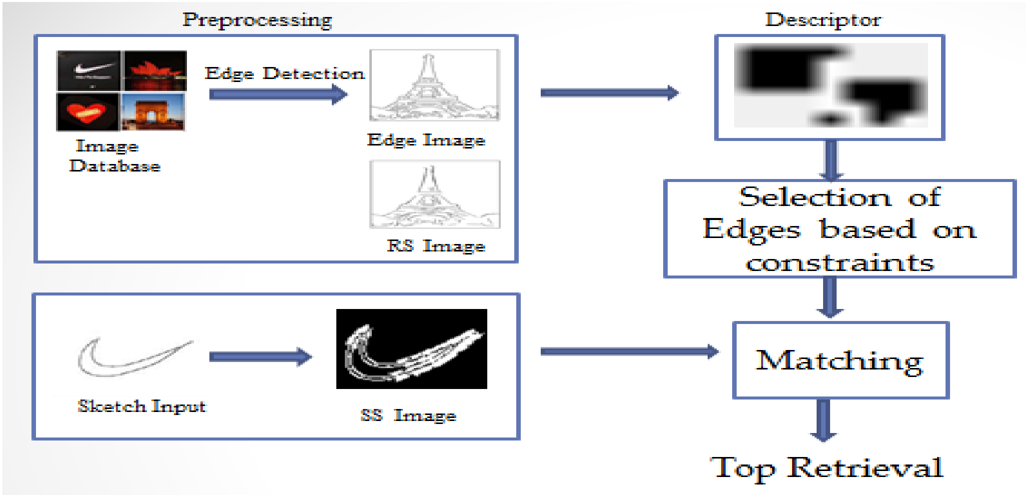

Fig 1: Block Diagram of proposed model

3.1 Edge Extraction

For strong edge extraction the Canny edge detector [4] is used on real images and then convert it into line segments. In this, convert sketch image into Sketch line segment (SS) image and real image into Real line segment (RS) image. Line segments and relationship between them shows the content of image. Fig. 2 shows that strong edges are extracted by applying canny edge detector on real image and Fig. 3 shows that strong edges are then converted into a set of line segments.

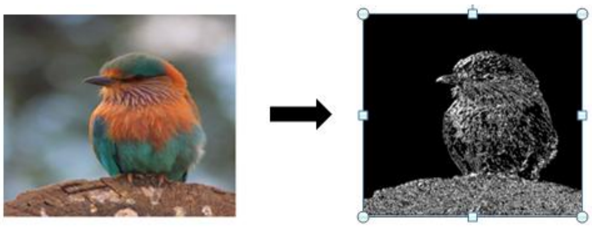

Fig 2: Image Preprocessing

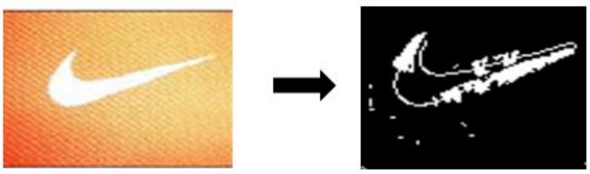

Fig 3: Edge Detection

3.2 Descriptor

A descriptor is designed such that it can find the relationship between line segments. Descriptor designing is block structure, in which blocks are define such that it divides area into four parts i.e. upper, lower, left, right. And next four blocks cover two points of middle line as well as cover block boundaries. Therefore this descriptor represents the objects boundaries and ensures that all adjacent line segments are covered. In following Fig. 4 shows block which divides neighboring area in left part. And red line shows middle (central) line of descriptor and green area shows the block definition.

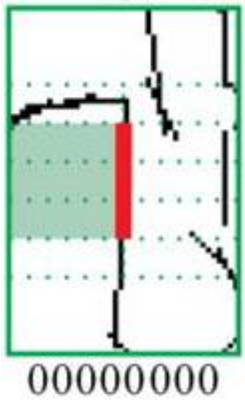

Fig 4: Descriptor structure

3.3 Selection of shaping edges

Because of noisy edges the retrieval performance get decreases, but human visual system is not affected by these noisy edges as human can differentiate between noisy edges and real edges. As an example shown in Fig 5, the input sketch is square and extracted edges form square with many noisy edges. It is very difficult for computer visual system to differentiate between them. As we are human can recognise the best match is (c) and others are noisy edges i.e. (a), (b) and (d) from following subset of images. Therefore the edge selection method is used to extract the shaping edges, which reduces the impact of noise.

To select such shaping edges, multiple hypotheses are generated for each descriptor by selecting certain adjacent edges as shaping edges and then find the best match with sketch query. Hypothetical word is generated for each hypothesis by removing the noisy edges rather than selecting the shaping edges, because generally observation is that noisy edges are often less than the shaping edges. The remaining edges are selected as shaping edges, once the noisy edges are removed.

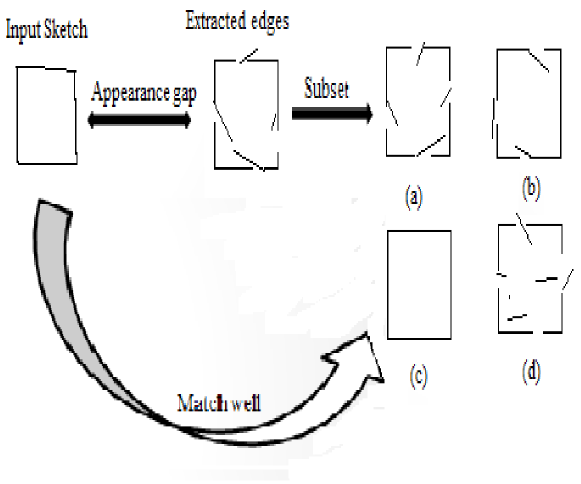

Fig 5: Edge Selection

To detect the false matches the constraints are given as follows:

(a) Constraint I

Define a word $w^j{}_s$ in SS image I_s and a word $w^j{}_r$ in RS image I_r as pair of corresponding words. Where j is mapped word id. For each corresponding words, the transformation parameter of scale and location $\theta(\hat{p},\hat{q},\hat{l}\,)$ is calculates as,

$$\hat{l} = \frac{l_s}{l_r} \quad \ldots(1)$$

$$\hat{p} = p_s - p_r\,\hat{l}\,,\ \hat{q} = q_s - q_r\,\hat{l} \quad \ldots(2)$$

Where, p and q are word coordinates, and *l* is the length of word. If the relationship between $w^i{}_s \sim w^j{}_s$ and $w^i{}_r \sim w^j{}_r$ are similar then their corresponding transformation parameter θ_i and θ_j must be similar. This transformation is 3 dimensional which forms small cubes. Then best cube C is selected from number of cube c_i according to weight of word w_i. Words in the best cube C have similar transformation parameters.

(b) Constraint II

The above constraint is cube based in which similar transformation is considered as best cube. Now in second constraint it is edge length based. In which edge length are compared in SS and RS images. Their directions are also matched. If matching done within threshold then that edges are selected.

Algorithm: Edge selection

```
for each edge in SS
int matched=infinity
int edgeId=-1
for each edge1 in RS
 if |Location(edge)-Location(edge1)| < threshold
 && |length(edge) - length(edge1)| <threshold
   if direction(edge) == direction(edge1) then
     if matched>|length(edge) - length(edge1)|
        matched=|length(edge) - length(edge1)|
        edgeId=edge1
     end if
   end if
 end if
end for
if edgeId!= -1 then
   countMatch++
end if
countTotal++
end for
```

4. EXPERIMENTAL RESULTS

A. Dataset

Two dataset are available for performance evaluation of proposed work those of Eitz [8] and Hu [9] (Flickr15K) and Wang[1] (Caltech256)[22].

B. Evaluation Metric

Proposed system use widely accepted evaluation metric Precision-Recall. For comparison Precision-Recall table and graph is provided. Fig. 6 shows graph of Precision-Recall of proposed and existing system.

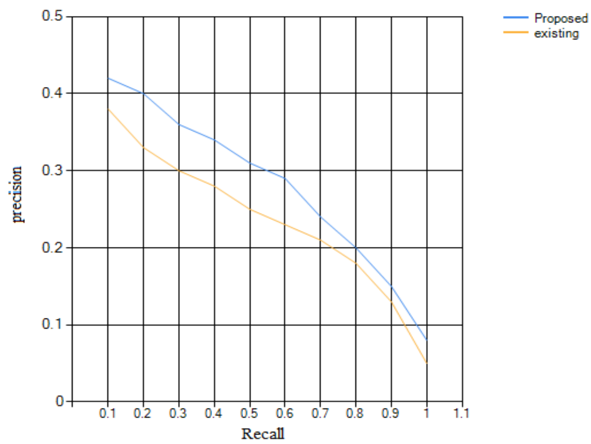

Fig 6: The precision and recall graph

C. Retrieval Time

In following table 1 shows the average retrieval time from which it can be seen that proposed retrieval speed is much faster than existing system[1].

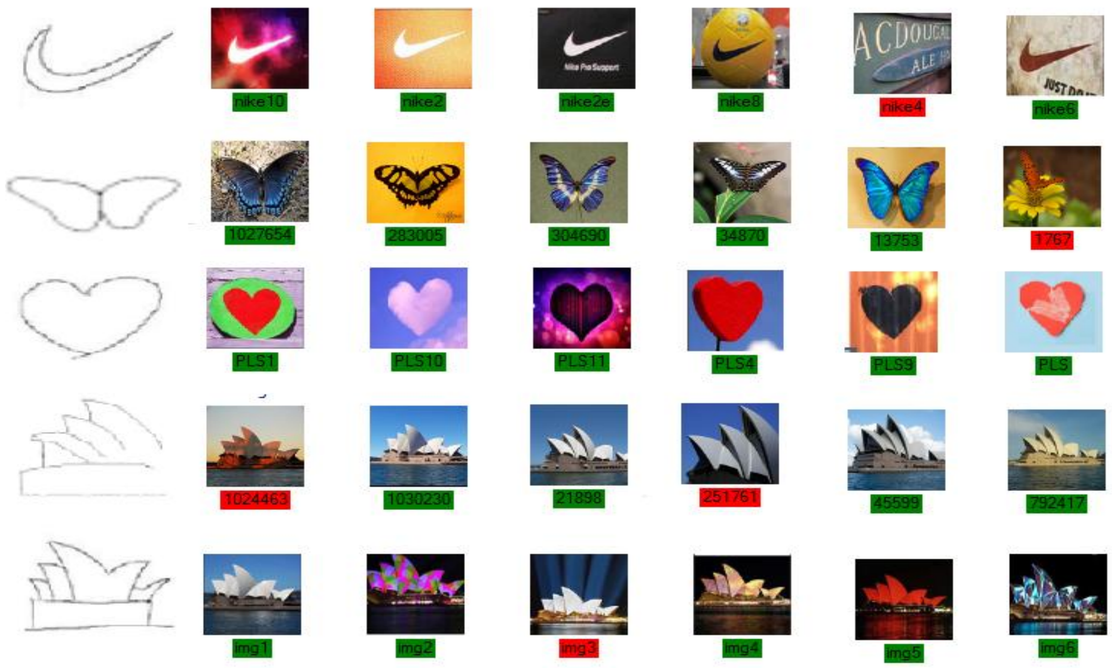

Fig 7: Examples of retrieval results. Image marked with green shows true matchingAnd failure cases are marked with red.

Table 1. Comparison of Retrieval Time

	Existing	**Proposed**
Time(s)	3.09	2.89

5. CONCLUSION

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 select the boundaries and to detect false matches edge length based constraint is applied which reduce the impact of noisy edges. To reduce the false matches, constraints are applied which improves the retrieval performance significantly.

This system can be extended to detect face sketches in crime. This will address in future work.

6. REFERENCES

[1] Shu Wang, Jian Zhang,Ding, "Sketch-Based Image Retrieval Through Hypothesis-Driven Object Boundary Selectionwith HLR Descriptor. ieee transactions on multimedia, vol. 17, no. 7, july 2015.

[2] "Canny Edge Detection" March 23, 2009.

[3] 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. 29–36.

[4] K. Bozas and E. Izquierdo, "Large scale sketch based image retrieval using patch hashing," *Adv. Visual Comput.*, vol. 7431, pp. 210–219.

[5] A. Chalechale, G. Naghdy, and A. Mertins, "Sketch-based image matching using angular partitioning," *IEEE Trans. Syst., Man, Cybern. A, Syst., Humans*, vol. 35, no. 1, pp. 28–41, Jan. 2005.

[6] J. M. Saavedra and B. Bustos, "An improved histogram of edge local orientations for sketch-based image retrieval," in *Proc. 32nd DAGM* Conf. Pattern Recogn., 2010, pp. 432–441.

[7] 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. 973–976.

[8] M. Eitz, K. Hildebrand, T. Boubekeur, and M. Alexa, "Sketch-based image retrieval: Benchmark and bag-of-features descriptors," *IEEE Trans. Vis. Comput. Graph.*, vol. 17, no. 11, pp. 1624–1636, Nov.2011.

[9] R. Hu and J. Collomosse, "A performance evaluation of gradient field hog descriptor for sketch based image retrieval," *Comput. Vis. ImageUnderstand.*, vol. 117, no. 7, pp. 790–806, 2013.

[10] M. Eitz, J. Hays, and M. Alexa, "How do humans sketch objects," *ACM Trans. Graph.*, vol. 31, no. 4, pp. 44:1–44:10, Jul. 2012.

[11] C. Ma, X. Yang, C. Zhang, X. Ruan, M.-H. Yang, and O. Coporation, "Sketch retrieval via dense stroke features," in *Proc. Brit. Mach. Vis. Conf.*, 2013, vol. 2, pp. 65.1–65.11.

[12] 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. 3661–3664.

[13] A. Shrivastava, T. Malisiewicz, A. Gupta, and A. A. Efros, "Data-driven visual similarity for cross-domain image matching," *ACM Trans. Graph.*, vol. 30, no. 6, pp. 154:1–154:10, Dec. 2011.

[14] T. Menp and M. Pietikinen, , J. Bigun and T. Gustavsson, Eds., "Multiscale binary patterns for texture analysis," in *Image Analysis*, ser. Lecture Notes Comput. Sci.. Berlin, Germany: Springer-Verlag, 2003, vol. 2749, pp. 885–892.

[15] 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. 471–474.

[16] SET. Chen, M.-M. Cheng, P. Tan, A. Shamir, and S.-M. Hu, "Sketch2photo: Internet image montage," in *Proc. ACM SIGGRAPH Asia*, 2009, pp. 124:1–124:10.

[17] X. Cao, H. Zhang, S. Liu, X. Guo, and L. Lin, "Sym-fish: A symmetryaware flip invariant sketch histogram shape descriptor," in *Proc. IEEE Int. Conf. Comput. Vis.*, Dec. 2013, pp. 313–320.

[18] O. Chum, J. Philbin, J. Sivic, M. Isard, and A. Zisserman, "Total recall: Automatic query expansion with a generative feature model for object retrieval," in *Proc. IEEE Int. Conf. Comput. Vis.*, Oct. 2007, pp. 1–8.

[19] Y.-L. Lin, C.-Y. Huang, H.-J. Wang, and W. Hsu, "3D sub-query expansion for improving sketch-based multi-view image retrieval," in *Proc. IEEE Int. Conf. Comput. Vis.*, Dec. 2013, pp. 3495–3502.

[20] P. Sousa and M. J. Fonseca, "Sketch-based retrieval of drawings using spatial proximity," *J. Vis. Languages Comput.*, vol. 21, no. 2, pp. 69–80, 2010.

[21] 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. 37–49.

[22] G. Griffin, A. Holub, and P. Perona, "Caltech-256 object category dataset," California Inst. Technol., Pasadena, CA, USA, Tech. Rep. CNS-TR-2007-001, 2007.