A Novel Approach for Sketch Based Image Retrieval with Descriptor

Dipika R. Birari¹, Prof. J. V. Shinde²

¹Department of Computer Engineering, Late G. N. Sapkal COE, Nashik, dipikabirari001@gmail.com
²Department of Computer Engineering, Late G. N. Sapkal COE, Nashik, jv.shinde@rediffmail.com

Abstract- A sketch is a freehand picture which serves various purposes such as tracing something that the artist can visualize so that they quickly explore concepts. Sketch-based communication is no newer, it is the oldest form of writing, and therefore Sketch-Based Image Retrieval (SBIR) can be a very valuable information search tool. Although sketch is a good way to express people’s ideas, but sketches and photorealistic images are very different in the term of appearance. The photorealistic image shows attention to realistic detail. There is a large appearance gap in user sketches and photorealistic images, when people sketch, they usually focus on main structure or shape of object and only draw the semantic contour boundary. On the other hand the color, texture and shape of an object are main elements of photorealistic images, which makes it very difficult to directly match a sketch and the corresponding photo-realistic image. 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 a framework consisting of line segment descriptor and noise impact reduction algorithm. Proposed descriptor extracts edges and captures the relationship between the 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 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 contours 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 converted into PLS image as shown in Fig. 1. 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. Following figure shows the edge extraction using canny detection.

![Figure 1. Edge Extraction](image_url)

After finding strong edges the descriptor is designed to find the line relationship. 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.
3.1 Preprocessing

The Canny[8] edge detection contains following:

- **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 \( \sigma = 1.4 \) is used as shown in (3.1).

\[
B = \frac{1}{159} \begin{bmatrix} 2 & 4 & 5 & 4 & 2 \\ 4 & 9 & 12 & 9 & 4 \\ 5 & 12 & 15 & 12 & 5 \\ 4 & 9 & 12 & 9 & 4 \\ 2 & 4 & 5 & 4 & 2 \end{bmatrix} \quad (3.1)
\]

- **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 (3.2) and (3.3) shows kernel gradient with respect to x and y.

\[
K_{Gx} = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \quad (3.2)
\]

\[
K_{Gy} = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} \quad (3.3)
\]

The gradient magnitudes can be calculated as a Euclidean distance measure by applying the law of Pythagoras as shown in (3.4) and (3.5). \( G_x \) And \( G_y \) are the gradients in the x- and y directions respectively.

\[
|G| = \sqrt{G_x^2 + G_y^2} \quad (3.4)
\]

\[
|G| = |G_x| + |G_y| \quad (3.5)
\]

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 (3.6).

$$\Theta = \arctan \left( \frac{|G_x|}{|G_y|} \right)$$  \hspace{1cm} (3.6)

- **Nonmaximum suppression:** This step is used to convert the “blurred” edges in the image of the gradient magnitudes to “sharp” edges. The algorithm is for each pixel in the gradient image[11]:
  1. Round the gradient direction to nearest 45°, 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.

3.2 **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 threshold , 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.

3.3 **Descriptor Design and Object Boundary Selection**
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. We will define the blocks which divides the neighboring area into four areas that ensures that no neighboring line segment is missing so that we can capture the line relationship. We will define extra blocks that will lie near two endpoints of the central line and covers other block boundaries. In object boundary selection, we will remove the noisy edges using removing operation. It will imitates the mechanism of human visual system in which it will focuses on shaping edges and ignores the noisy edges and then will predict the shaping edges.

**IV. RESULTS**

In figure 4 user browsed the image and start with further process. The edges are extracted using canny edge detector. In next figure 5, input image is converted into Gaussian filtered image, then further converted into non max suppressed image, then extract strong and weak edges and finds canny edges.

![Figure 4. Browsed image](image-url)
CONCLUSION

The proposed system extracts the edges using canny edge detector and then by applying descriptor, it enhance the performance of the image retrieval. A systematic approach that bridges the appearance gap for SBIR by considering sketches and extracted edges from a new angle, i.e., treating them as a set of line segments, laying the foundation for better sketch and extracted edge description and noise impact reduction using canny edge detector. Further by using the object boundary selection and applying constraints, the proposed system will efficiently retrieve the images with strong edges.

REFERENCES


