Sketch-based Image Retrieval using Rotation-invariant Histograms of Oriented Gradients

Y.Jhansi[1], Dr.E.Sreenivasa Reddy[2]

[1]*Research Scholar, ANU college of Engineering, Acharya Nagajuna University, GUNTUR*
[2]*Professor and Prinicipal, ANU college of Engineering, Acharya Nagajuna University, GUNTUR*

Abstract:

This paper presents a novel approach for sketch-based image retrieval based on Rotation Invariant. The approach enables measuring the similarity between a full color image and a simple black and white sketched query. We present efficient feature representation method namely rotation-invariant histograms of oriented gradients (Ri-HOG) for image retrieval. Most of the existing HOG techniques are computed on a dense grid of uniformly-spaced cells and use overlapping local contrast of rectangular blocks for normalization. However, we implement annular spatial bins type cells and apply radial gradient to attain gradient binning invariance for feature extraction. In this way, it significantly improves HOG in regard to rotation invariant ability and feature descripting accuracy. In experiments, the proposed method is evaluated on wang datasets. The experimental results demonstrate that the proposed method is much more effective than many existing image feature descriptors for sketch based image retrieval.

Keywords: *Image retrieval, rotation, HOG*

I. Introduction:

The explosive increase of digital images on the web has significantly increased the need of a correct, efficient and user-friendly large scale image retrieval system. With the rising popularity of touch-based smart computing devices and the consequent ease and simplicity of querying images via hand-drawn sketches on touch screens [1]. Content-based image retrieval (CBIR) is the application of computer vision to the image retrieval trouble, so as to, the difficulty of searching for digital image in huge databases. "Content-based" means with the purpose of search will consider the actual contents of the image. The term 'content' in this context might submit to shade, figure, surface, or some other information that can be derived from the image itself [2]. In several cases if we want to search efficiently some data have to be memorized. The human is capable to remember pictorial information more easily for example the shape of an object or arrangement of colors and objects. Sketch-based image retrieval, on the other hand, being a far more significant way of image search. In the sketch based image retrieval methodology the hand drawn sketches being a natural way of representing and exploring a synthetic query. Finding similarities between a binary sketch and a database of colored images taken under subjective conditions consists a challenging problem. Images and binary sketches do not share many common modalities. Images contain rich information in domains such as color and texture, while sketches can be described only by their shape and spatial configuration, therefore traditional CBIR methods relying on texture and color cannot be inherited in SBIR.[3]

In this paper, sketch based image retrieval with rotation invariant features is implemented. Currently, the dominant technique for rotation-invariant image matching is to find corresponding local image features, each of which is rotation invariant.

II. Related work:

A number of approaches have been considered in the literature for unfolding shapes and measuring their similarity. Eitz et al. (2009) decompose an image or sketch into different spatial regions and measure the correlation between the direction of strokes in the sketch and the direction of gradients in the image by proposing two types of descriptors, viz. an Edge Histogram Descriptor and a Tensor Descriptor. Histograms of prominent gradient orientations are encoded in the Edge Histogram Descriptor, whereas the Tensor Descriptor determines a single representative vector per cell that captures the main orientation of the image gradients of that cell. Descriptors at corresponding positions in the sketch and the image are correlated for matching. Due to this strong spatial assumption, they fail to retrieve images if the sketched object is present at a different scale, orientation and/or position[4]. To determine similar images corresponding to a sketch, a linear scan over all database images is performed.

This further limits the scalability of the method for large databases.

The Query by Visual Example (QVE) [5] meaasures correlation similarities between two abstract images based on bitwise operations. The correlation performs both horizontal and vertical shifts achieve invariance to translation. This method is not supporting indexing and expensive and not invariant to rotation .

In Angular partitioning of abstract images the key feature is angular distribution of the pixel in the abstract image. In this method the whole image is partitioned into slices and considers the edges of each slice. It forms a feature vector from these edges and then apples Fourier Transformation to achieve rotation invariance and not fully invariant to scaling, rotation, and translation. The effectiveness of these on image retrieving from data base is less [6].

From the above methodologies highlighted in the above relevant works, it can be identified that the rotation invariant in most of the cases is not maximized. In this paper, we develop a system for large scale sketch-based image retrieval that can handle rotation (similarity) variations. The proposed feature is based on histogram descriptors. It is well known that histogram is a useful tool for image feature representation, but the robustness of many algorithms based on histogram descriptor does not reach maturity. In order to address this problem, this paper presents a new method of feature representation for SBIR, i.e. rotation-invariant histograms of oriented gradients for image retrieval.

III. Proposed Method:

We present the proposed methodology for sketch based image retrieval with rotation invariant which consists of three stages such as preprocessing, Feature Extraction, Image matching and Similarity Analysis. In the preprocessing stage Query sketch and all the images in database are normalized to 256× 256 size, the purpose of normalization is to explore all the sketch and images to be of same size.. In the second stage feature extraction technique is applied: Rotation invariant Histogram Of oriented Gradients computed at a circular support area and uses an annular binning to achieve orientation invariance [7] and the obtained feature space is stored in the respective feature template database. After feature extraction, the given query sketch feature space will be compared with those in the feature templates one by one and the similar images are retrieved.

1	**2**
3	**4**

3	**1**
4	**2**

Fig.1 (a) The original block-size image (b) Image (a) rotated clockwise 90^0

HOG are feature descriptors, which are computed on a dense grid of uniformly-spaced cells and use overlapping local contrast normalization for improved accuracy. This features set based on cells and blocks is widely used as object feature descriptors, especially the descriptors in human detection task [8]. A HOG is not a rotation invariant depiction. Therefore, when used in image retrieval for rotated images, it can only handle images that are observed at a certain orientation. To overcome this limitation, a new variant of the HOG descriptor was recently proposed. The so called Rotation-Invariant Fast Feature (RIFF) descriptor [7] is based on a HOG computed at a circular support area and uses an annular binning to achieve orientation invariance. See Fig. 1 for an example. Suppose Fig. 1(a) is a HOG block-size image, there are 4 cells in the block. Fig. 1(b) is an image of Fig. 1(a) after making a quarter turn. HOG features are extracted from the two images individually. If the histogram of oriented gradients obtained from the regions 1, 2, 3, and 4 are respectively denoted as x1, x2, x3, x4, then, the HOG features extracted from Fig. 1(a) and Fig. 1(b) are (x1, x2, x3, x4) and (x3, x1, x4, x2) respectively. This means that the rotation of image accompanies with the change of its HOG descriptors. Hence, we have to substantially enhance the robustness of HOG descriptors.

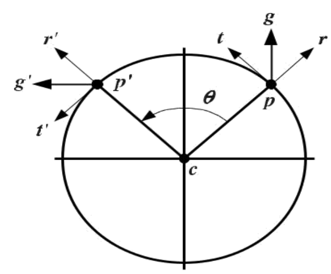

Fig.2. Illustration of radial gradients

A. Rotation invariance

The Rotation Invariant Fast Features (RIFF) descriptor [7] is a recent approach that leverages on the proven methods of SIFT [9] and HOG [9] and provides robustness and rotation invariance. To make gradient binning invariant we apply an invertible, spatially-varying transform. By rotating the gradients to the proper angle, we achieve rotation invariance with no loss of information, yielding the Radial Gradient Transform (RGT).

As shown in Fig.2, we choose two orthogonal basis vectors to provide a local, polar reference-frame for describing the gradient. These basis vectors, r and t, are the radial and tangential directions at a point p, relative to the center of the patch, c. The gradient, g, is projected onto a local, radial coordinate system (r, t). Decompose vector g into its local coordinate system as ($g^T r$, $g^T t$), by projecting g into the r and t orientations. 0We define R_θ as the rotation matrix for angle θ, yielding

$$r = \frac{p-c}{|p-c|} \quad , \quad t = R\frac{\pi}{2}r \qquad (1)$$

Now that we have discussed the arrangement of blocks and how to get the feature descriptors from cells, we summarize the four steps of rotation-invariant HOG features extraction:

1. Subdivide the local patch into annular spatial cells as shown in Fig. 2.

2. Calculate ($g^T r$, $g^T t$) of each pixel in the cell;

3. Calculate the radial gradient magnitude (MGR) and its orientation (θ) on location (x, y), using the Eq. 2

$$M_{GR} = \sqrt{(\mathrm{g\,T\,r})^2 + (\mathrm{gT\,t})^2}$$

$$\theta(x, y) = arctan\frac{\mathrm{gT\,t}}{\mathrm{gT\,r}} \qquad (2)$$

4. Accumulating the gradient magnitude of radial gradient for each pixel over each annular spatial cells into 8 bins.

As shown in Fig. 2 assume that the patch has been rotated about its center by some angle, θ. This yields a new local coordinate system and gradient

$$R_\theta P = P', R_\theta r = r', \; R_\theta t = t', R_\theta g = g',$$

The coordinates of the gradient in the local frame are invariant to rotation, which is easily verified by

$$\begin{aligned} g^T r', g^T t' &= ((R_\theta g)^T R_\theta r, (R_\theta g)^T R_\theta t) \\ &= (g^T R_\theta{}^T R_\theta r, g^T R_\theta{}^T R_\theta t) \\ &= (g^T r, g^T t) \qquad (3) \end{aligned}$$

All gradients are rotated by the same angle and R_θ is a one-to-one mapping. Thus, the set of gradients on any given circle centered around the patch is invariant to rotation.

IV. Experimental results

The system has been implemented with a wang database includes colored images, which are stored in JPEG format with size 256×256. Before extracting features all the images in the database are preprocessed and then rotation invariant features are extracted. The sketch image is provided as an input, and then the relevant images in the database are retrieved. The measurement of image content similarity is evaluated by distance metric. Namely, the distance between the query image and the template image in the dataset is calculated by Euclidean distance. The distance metric is one of the simplest approaches, therefore, it can prove the validity of descriptors most directly. The solutions that we attain using the projected scheme are illustrated in the Fig.3.

We use the Precision and Recall to evaluate the performance of the proposed method. These two indices are the most commonly used for evaluating image retrieval performance. Precision is the ratio of the number of retrieved similar images to the number of retrieved images; Recall is the ratio of the number of retrieved similar images to the total number of similar images. They are defined as follows

$$precision = \frac{\text{Number of retrieved images relevant to the query image}}{\text{Total number of images retrieved}} \qquad (4)$$

$$recall = \frac{\text{Number of retrieved images relevant to the query image}}{\text{Total number of relevant images in the database}} \qquad (5)$$

Table1: Precision and recall values

S.No	Retrieval count	Precision	Recall
1	20	0.8646	0.13
2	30	0.733	0.37
3	40	0.56	0.49

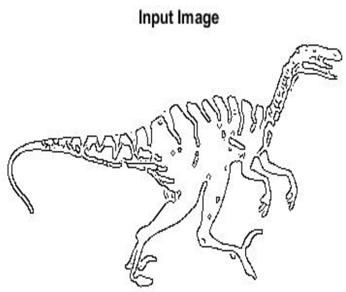

Fig.3. (a) Input sketch

(b) Retrieved images using Ri-HOG

V. CONCLUSION

In this paper, we have proposed a novel feature representation method for sketch-based image retrieval, i.e., rotation invariant histograms of oriented gradients (Ri-HOG). It is derived from Dalal et al.'s HOG yet ameliorated by the theory of polar coordinate. This histogram entirely differs from existing ones based on histograms. Therefore, it is a simple but efficient image retrieval approach. The retrieval performance of this method is shown in Table 1. We showed that Rotation Invariant achieves 86% retrieval accuracy from wang database.

VI. References

[1] Lee, Y.J., Zitnick, C.L., Cohen, M.F.: Shadow draw: real-time user guidance for freehand drawing. In: ACM Transactions on Graphics (Proc. SIGGRAPH). vol. 30, p. 27. ACM (2011)

[2] Samy Ait-Aoudia, Ramdane Mahiou, Billel Benzaid, ‖YACBIR: Yet another Content Based Image Retrieval system‖, IEEE 2010

[3] Bozas B, Konstantinos K, Izquierdo E. Large scale sketch based image retrieval using patch hashing.Advances in Visual Computing; 2012. p.210–19.

[4] Eitz M. Hildebrand K. Boubekeur T, Alexa M . Sketch-Based Image Retrieval: Benchmark and Bag-of-Features Descriptors In IEEE Transactions on Visualization and Computer Graphics 2011, 17(11), pp. 1624 – 1636.

[5] Kato, T., Kurita, T., Otsu, N., Hirata, K.: A sketch retrieval method for full color image database-query by visual example. In: Proc. of the 11th IAPR International Conf. on Computer Vision and Applications, Conf. A: Pattern Recognition, pp. 530–533, 1992

[6] Chalechale, A., Naghdy, G., Mertins, A.: Sketch-based image matching using angular partitioning. IEEE Trans. on Systems, Man and Cybernetics, Part A: Systems and Humans 35(1), 28–41, 2005

[7] Chen, J.; Nakashika, T.; Takiguchi, T.; and Ariki, Y. 2015. Content-based image retrieval using rotation-invariant histograms of oriented gradients. In The ACM, 443–446

[8] 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.

[9]. 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

[10]. Takacs, G., Chandrasekhar, V., Tsai, S., Chen, D., Grzeszczuk, R., Girod, B.: Unified real-time tracking and recognition with rotation-invariant fast features. In: CVPR. (2010)

[11] 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.

[12] Jhansi Y, Reddy E S "An Efficient Sketch Based Image Retrieval using Cross-correlation", IJCSIS, Vol 14. No 12, pp 445-451, Dec 2016.

[13] Rupinder kaur , Navleen kaur,"Content Based Image Mining Technique for Image Retrieval Using Optimized Hybrid Clustering", IJCTT, Vol 11.NO 3,2014, pp 141-143