Visual Saliency Weighting and Cross-Domain Manifold Ranking for Sketch-Based Image Retrieval

Takahiko Furuya and Ryutarou Ohbuchi

Graduate School of Medicine and Engineering, University of Yamanashi, Japan
{g13dm003,ohbuchi}@yamanashi.ac.jp

Abstract. A Sketch-Based Image Retrieval (SBIR) algorithm compares a line-drawing sketch with images. The comparison is made difficult by image background clutter. A query sketch includes an object of interest only, while database images would also contain background clutters. In addition, variability of hand-drawn sketches, due to "stroke noise" such as disconnected and/or wobbly lines, also makes the comparison difficult. Our proposed SBIR algorithm compares edges detected in an image with lines in a sketch. To emphasize presumed object of interest and disregard backgrounds, we employ Visual Saliency Weighting (VSW) of edges in the database image. To effectively compare the sketch containing stroke noise with database images, we employ Cross-Domain Manifold Ranking (CDMR), a manifold-based distance metric learning algorithm. Our experimental evaluation using two SBIR benchmarks showed that the combination of VSW and CDMR significantly improves retrieval accuracy.

Keywords: sketch-based image retrieval, visual saliency detection, cross-domain matching, manifold ranking.

1   Introduction

Querying modality is an issue central to image retrieval. Query-by-keywords may be the easiest among query modalities for users to perform, but its retrieval accuracy is not satisfactory. Images taken by phones and cameras generally do not have keyword tags, and if they do, specifying image content by keywords alone can be difficult. An alternative is content-based image retrieval, which uses an example image or a user-drawn sketch of some kind to query images. The first option, querying by image example, is not practical as a user often does not have an image appropriate for query. The second option, Sketch-Based Image Retrieval (SBIR), has become a popular modality of late. Earlier SBIR systems, mostly likely using a mouse as an input device, have drawn sketches by using simple geometrical primitives, such as straight lines, rectangles or circles for drawing, possibly combined with color specification. Recently, with prevalence of touch- and pen-based devices, line-drawing sketch has become the most popular querying modality for content-based image retrieval.

The issue in SBIR is an effective comparison between a line drawing sketch and a 2D image. The comparison should be *robust against background clutter* in the 2D

image. It should also be robust against "stroke noise" of the line drawing sketch, that is, wobbling and/or disconnected lines and difference in drawing styles.

A SBIR algorithm uses a common ground representation, either line/edge image [1-5] or gradient field image [6, 7], for sketch-to-image comparison. Then, one or more image features are extracted from the common ground representation for comparison. Saavedra et al. [3] uses a set of local histograms of Canny edge orientations as an image feature, while Eitz et al. [2] combines several orientation sensitive image features including a variant of HoG [14]. The BF-GFHoG by Hu et al. [6][7] employs gradient field image as the common ground and extracts HoG descriptors. Hundreds of HoG descriptors extracted from a gradient field image are integrated into a feature vector per image by using the Bag-of-Features (BF) approach. To our knowledge, BF-GFHoG is one of the best performing SBIR algorithms.

These algorithms perform well if database images do not have any background clutter and if all the strokes in sketch images are stable and well-connected. However, images contain background clutters, and sketch strokes are disconnected or wobbly. For example, images shown in Figure 1a contain leaves and clouds in their backgrounds, producing noisy Canny edge images in Figure 1c. A casual photograph taken by using phones would contain even more clutters. Also, hand-drawn sketches contain stroke noise. Comparison of a sketch having stroke-noise with an image having background-clutter would lead to low retrieval accuracy.

To suppress background clutter, visual saliency map is often employed. A visual saliency map, which approximates visual attention of human, is used for image segmentation [16], object detection [17], and other applications. Recently, Yang et al. [8] proposed a graph-based saliency detection method called Manifold Ranking-based Saliency Detection (MRSD). The MRSD yields better saliency maps than existing methods. Our proposed algorithm uses the MRSD algorithm to emphasize foreground object of an image for background-clutter-free comparison of a sketch with an image.

To perform robust sketch-to-image comparison under the presence of sketch stroke noise, distance metric learning has been employed [9][12][15]. For example, Weinberger et al. [15] compares handwritten digits in a subspace where distances among feature vectors are robust against stroke noise. Recently, we proposed Cross-Domain Manifold Ranking (CDMR) algorithm for the task of sketch-to-3D model "cross-domain" comparison [12]. The CDMR is based on Manifold Ranking proposed by Zhou et al. [9]. It learns distributions, or manifolds, of sketch features and 3D model features to improve distance (or similarity) computation among them. Our proposed algorithm adopts the CDMR algorithm for improved sketch-to-image comparison.

In this paper, we aim for a SBIR algorithm that is *robust against background clutters* in images and *robust against stroke noise* in sketches. Figure 2 illustrates the overall processing pipeline of the proposed algorithm.

To gain robustness against background clutters, the algorithm tries to emphasize a region presumed to be foreground, *i.e.*, the object sought for by the sketch query. The emphasis, called Visual Saliency Weighting, is done by multiplying, pixel-by-pixel, a Canny edge image with the visual saliency map computed by using the MRSD algorithm [8]. Figure 1d shows the examples of visual-saliency-weighted edge images. Background clutter, i.e., edges of leaves and clouds, are effectively suppressed.

To gain robustness against stroke noise in sketch, the algorithm employs data-adaptive distance metric learning. Relevance values from a sketch to images in the database are computed by using the CDMR algorithm [12]. The CDMR uses a Cross Domain Manifold (CDM) constructed by using similarity values due to multiple, heterogeneous features, each of which is may be optimized for a given comparison task. A manifold of database images is computed (without edge detection) by using the densely sampled SIFT [11] features. A manifold of sketches is computed by using the BF-fGALIF [12] feature. Then these two manifolds in two different domains are coupled into a CDM by using the BF-fGALIF feature. Once the CDM is constructed, relevance values from a sketch to images in a database are computed as diffusion distances over the CDM. The CDMR may be used in either semi-supervised, supervised or unsupervised mode. In this paper, we use the CDMR in unsupervised mode.

Note that the CDMR algorithm has a built-in ability to perform automatic query expansion if an unlabeled (or labeled) corpus of sketches is available. Relevance diffused from a sketch query turns sketches similar to it into secondary sources of relevance diffusion, creating an expanded set of queries. Such a corpus of sketches may be collected beforehand, or collected online over time as sketch queries are made.

We experimentally evaluated the proposed algorithm by using two sketch-based image retrieval benchmarks by Hu et al., that are, the Flickr160 [6] and the Flickr15k [7]. Small but consistent improvement in retrieval accuracy is observed for both benchmarks when the VSW is applied. The CDMR improved retrieval accuracy very significantly. For example, for the Flickr160 benchmark, the combination of the VSW and the CDMR produced MAP score of 72.3 %, which is about 18 % higher than 54.0 % of the BF-GFHoG reported in [6].

Contribution of this paper can be summarized as follows;

- Proposal of a novel sketch-based image retrieval algorithm. It employs Visual Saliency Weighting (VSW) to suppress background clutter in images. The features extracted from edge images processed by VSW are compared against the feature of a sketch query by using the Cross-Domain Manifold Ranking (CDMR), a distance metric learning algorithm adept at comparing heterogeneous feature domains.
- Experimental evaluation of the proposed algorithm using multiple benchmarks, which showed effectiveness of the proposed algorithm.
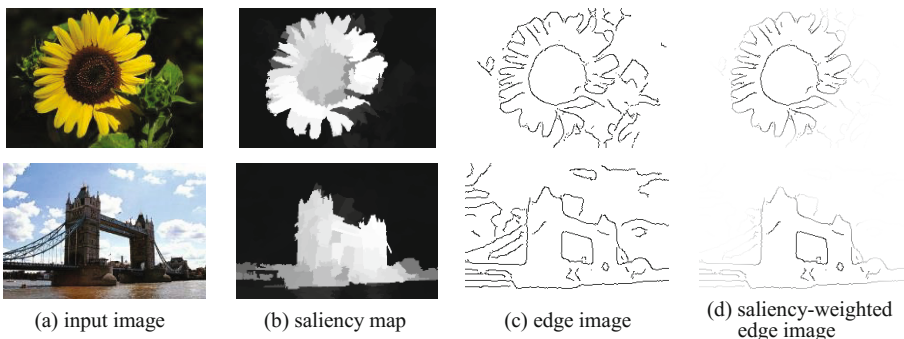

(a) input image   (b) saliency map   (c) edge image   (d) saliency-weighted edge image

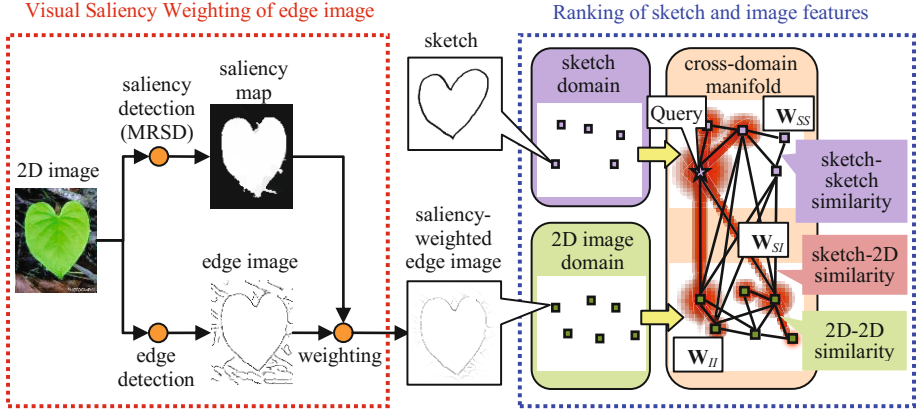**Fig. 1.** Examples of visual saliency weighting

Fig. 2. Outline of the proposed method that employs Visual Saliency Weighting (VSW) of Canny edge images and Cross-Domain Manifold Ranking (CDMR) of image features

2   Proposed Method

2.1   Visual Saliency Weighting of Edge Image

Our algorithm converts (2D) images in a database into saliency-weighted edge images for comparison with sketches. The algorithm first computes Canny edge image from the database image. We used the parameters found in [2] for Canny edge detection; low threshold and high threshold are set to 0.05 and 0.2, respectively. Then, edges due to background clutters are suppressed by using *Visual Saliency Weighting* (*VSW*), which multiplies, pixel-by-pixel, a visual saliency map with the Canny edge image. Visual saliency map is computed by using the MRSD algorithm by Yang et al. [8]. The MRSD algorithm computes visual saliency in two steps, assuming periphery of an image as background.

In the first step, "background-ness" is propagated from image periphery at four sides of the image toward the center. The propagation is performed by using Manifold Ranking (MR) [9] algorithm over a graph that connects, conceptually, all the pixels of the image. However, MR on such a large graph would be too expensive. Thus, the propagation is done on a simplified graph connecting superpixels, a cluster of neighboring pixels having similar color property. An edge of the graph connecting a pair of superpixels is weighted by the similarity of the pair of superpixels. The similarity is computed from CIE LAB color features of each superpixel. We set the number of superpixels per image to 200, the number used in [8]. After the first step, the nodes having "background-ness" value lower than a threshold becomes foreground.

In the second step, "foreground-ness" is propagated from the foreground regions over the graph of superpixels. After the second step, "foreground-ness" of each superpixel becomes saliency value for pixels belonging to the superpixel. We blur the saliency map before multiplying it with the Canny edge image.

Figure 1d shows examples of saliency-weighted edge images. Edges on background (e.g., edges on leaves and edges on clouds) are effectively suppressed.

2.2   Cross-Domain Manifold Ranking for Sketch-Based Image Retrieval

Cross-Domain Manifold Ranking (CDMR)

The CDMR [12] consists of two stages; Cross-Domain Manifold (CDM) generation stage and relevance diffusion stage.

In the CDM generation stage, a CDM matrix $\mathbf{W}$ is generated. $\mathbf{W}$ is a graph whose vertices are the features from a sketch domain and an image domain. Given the number of sketches N_S and the number of images N_I, the matrix $\mathbf{W}$ has the size $(N_S + N_I) \times (N_S + N_I)$.

$$\mathbf{W} = \begin{pmatrix} \mathbf{W}_{SS} & \mathbf{W}_{SI} \\ \mathbf{W}_{IS} & \mathbf{W}_{II} \end{pmatrix} \tag{1}$$

The submatrix $\mathbf{W}_{SS}$ having size $N_S \times N_S$ is the manifold graph of sketch features generated by linking features of sketches produced by the BF-fGALIF [12] algorithm. We will describe the BF-fGALIF in the following section. An edge of the graph $\mathbf{W}_{SS}$ connecting vertices i and j is undirected and has a weight, which is a similarity $w(i, j)$ among the vertices. The similarity $w(i, j)$ is computed by using the equation (2) after normalizing the distance $d(i, j)$ of features i and j to range [0,1].

$$w(i, j) = \begin{cases} \exp\left(- d(i, j)/\sigma\right) & if \quad i \neq j \\ 0 & otherwize \end{cases} \tag{2}$$

The submatrix $\mathbf{W}_{II}$ having size $N_I \times N_I$ is a manifold graph of image features. It is created in a similar manner as $\mathbf{W}_{SS}$. Features for image-to-image comparison are computed by using BF-DSIFT [13] algorithm.

The submatrix $\mathbf{W}_{SI}$ of size $N_S \times N_I$ couples two submanifolds $\mathbf{W}_{SS}$ and $\mathbf{W}_{II}$ that lie in different domains, that are, sketch feature domain and image feature domain. To compute a feature similarity $w(i, j)$ between a sketch i and an image j, BF-fGALIF features are extracted from a sketch image i and a saliency-weighted edge image generated from an image j. Finally, feature similarity $w(i, j)$ is computed by using equation (2).

Submatrix $\mathbf{W}_{IS}$ of size $N_I \times N_S$ is a zero matrix as we assume no diffusion of relevance occurs from image features to sketch features.

In the relevance diffusion stage, ranking of images to a query sketch is done by diffusing relevance value from the query to the images over the CDM by using MR [9]. We normalize $\mathbf{W}$ for $\mathbf{S}$ by using the following equation;

$$\mathbf{S} = \mathbf{D}^{-1/2} \mathbf{W} \mathbf{D}^{-1/2} \tag{3}$$

where $\mathbf{D}$ is a diagonal matrix whose diagonal element is $\mathbf{D}_{ii} = \sum_j \mathbf{W}_{ij}$. We use the following equation to find rank values in $\mathbf{F}$ given initial value, or "source" matrix $\mathbf{Y}$;

$$\mathbf{F} = \left(\mathbf{I} - \alpha \mathbf{S} \right)^{-1} \mathbf{Y} \tag{4}$$

$\mathbf{Y}$ is a diagonal matrix of size $(N_S + N_I) \times (N_S + N_I)$ that defines source(s) of relevance value diffusion. If a vertex i is the source of diffusion $\mathbf{Y}_{ii} = 1$, and, if not, $\mathbf{Y}_{ii} = 0$. In our case, the vertex corresponding to the query sketch becomes the source of diffusion. $\mathbf{F}_{ij}$ is the relevance value of the image j given the sketch i. The higher the relevance value $\mathbf{F}_{ij}$, the higher the rank of the image j in the retrieval result.

The ranking is very robust, as diffusion from the query sketch to the images occurs via multiple paths. For example, the relevance value may first diffuse quickly through sketches similar to the query before reaching to the images. In such a case, the CDMR embodies a form of query expansion.

The parameter σ in equation (2) controls diffusion of relevance value across the CDM. We use different values σ_{SS}, σ_{II}, and σ_{SI} for each of the submatrices $\mathbf{W}_{SS}$, $\mathbf{W}_{II}$, and $\mathbf{W}_{SI}$ as optimal value of σ depends on each submatrix. The parameter $\alpha = [0,1)$ in equation (4) controls regularization.

Computing Feature Similarities for Cross-Domain Manifold

Generation of the CDM $\mathbf{W}$ requires computation of similarities in the submatrices $\mathbf{W}_{SS}$, $\mathbf{W}_{II}$, and $\mathbf{W}_{SI}$. In this section, we describe features used for computing each submatrix.

Computing Similarities for $\mathbf{W}_{SI}$

Figure 3a shows a feature comparison pipeline for $\mathbf{W}_{SI}$. Given an image in a database, it is resized so that the shorter edge of the image becomes 256 pixels. A saliency-weighted edge image is generated from the resized image by the VSW described in Section 2.1.

From each of the saliency-weighted edge images and the sketch images, we extract BF-fGALIF feature, which is shown to be among the best methods for sketch-based 3D shape retrieval [12]. GALIF features [10] are extracted densely at regular grid points on the image. GALIF feature captures orientation of lines and intensity gradient in the image by using Gabor filters. Parameters for GALIF we used in this paper are the same as those in [12]. 1,000 to 1,500 GALIF features are extracted per image.

The set of GALIF features extracted from an image is integrated into a feature vector per image by using a standard bag-of-features (BF) approach. This integration reduces cost of image-to-image matching significantly compared to directly comparing a set of features to another set of features. We used vocabulary size k=3,500 for the experiments. We used k-means clustering to learn the vocabulary, and used kd-tree to accelerate vector quantization of GALIF features into words of vocabulary.

A BF-fGALIF feature of a sketch image is compared against a BF-fGALIF feature of a saliency-weighted edge image by using Cosine distance.

Computing similarities for $\mathbf{W}_{SS}$

Similarities for the submatrix $\mathbf{W}_{SS}$ is computed, again, by the BF-fGALIF algorithm (Figure 3b). For each sketch image, BF-fGALIF feature is extracted after the sketch

image is resized down to 256×256 pixels. A distance between BF-fGALIF features is computed using Cosine distance. All the parameters for the BF-fGALIF (i.e., parameters for Gabor filter, number of GALIF features per image, and vocabulary size) are the same as those used to compute $\mathbf{W}_{SI}$.

Computing Similarities for $\mathbf{W}_{II}$

Similarities for the submatrix $\mathbf{W}_{II}$ is computed by the BF-DSIFT algorithm [13] (Figure 3c). BF-DSIFT is extracted from grayscale images without edge detection. Given an image, it is resized so that the shorter edge of the image becomes 256 pixels as with computing $\mathbf{W}_{SI}$.

From each resized image, about 3,000 SIFT [11] features are extracted at densely and randomly placed feature points on the image. SIFT has invariance against scaling, rotation, illumination changes and minor changes in viewing direction. The set of about 3,000 SIFT features are integrated into a feature vector per image by using a BF approach. We used vocabulary size k=3,500 for the experiments. A distance between BF-DSIFT features is computed using symmetric version of Kullback-Leibler Divergence.
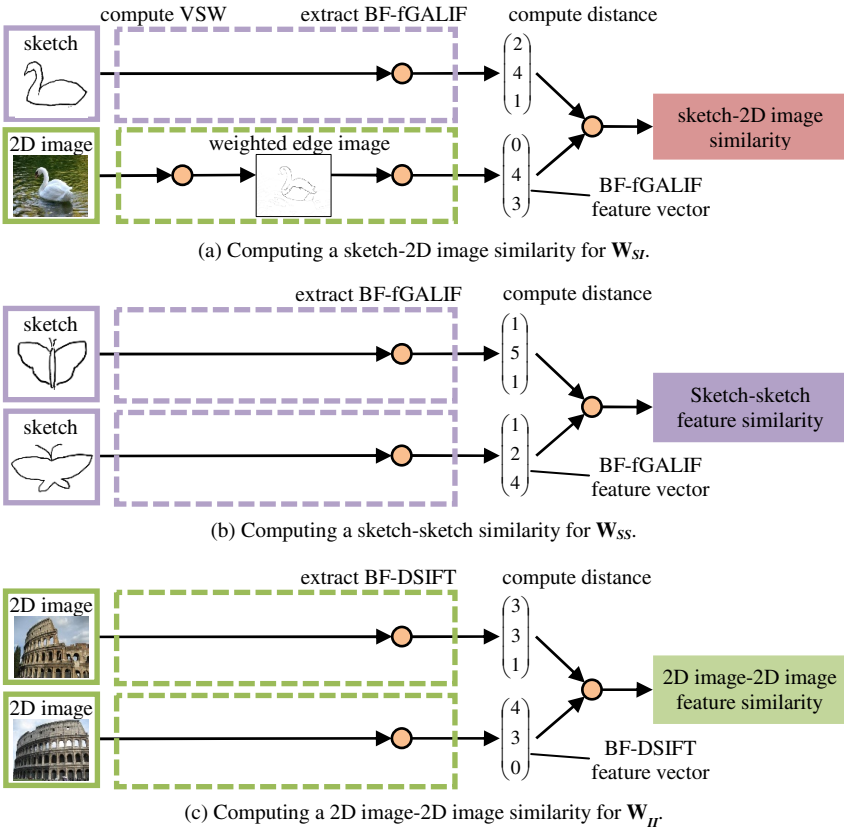

(a) Computing a sketch-2D image similarity for $\mathbf{W}_{SI}$.

(b) Computing a sketch-sketch similarity for $\mathbf{W}_{SS}$.

(c) Computing a 2D image-2D image similarity for $\mathbf{W}_{II}$.

Fig. 3. Feature comparison methods for generating CDM

3   Experiments and Results

We experimentally evaluated effectiveness of weighting by the VSW of Canny edge images and ranking by the CDMR for sketch-based image retrieval. We used two sketch-based image retrieval benchmark databases; the Flickr160 [6] and the Flickr15k [7], both by Hu et al. Figure 4 shows examples of sketch queries and retrieval target images for the two benchmarks.

The Flickr160 consists of a set of 25 sketch queries and a set of 160 retrieval target images. Each of the set of sketch queries and the set of target images is partitioned into 5 categories. The Flickr15k, which is a larger-scale version of the Flickr160, consists of a set of 330 sketch queries and a set of 14,660 retrieval target images. The set of sketch queries is partitioned into 33 shape categories (e.g., "round", "heart-shape", etc.). The set of retrieval target images is partitioned into 60 semantic categories (e.g., "pyramid", "bicycle", etc.). A query in the Flickr15k could belong to multiple semantic categories, for example, a round-shaped sketch query belongs to such semantic categories as "moon", "fire_balloon", and "london_eye".

We used our own implementations for the CDMR and the BF-fGALIF. The MRSD [7] and BF-GFHoG [8] are computed by using original codes downloaded from respective authors' websites.

Parameters σ_{SS}, σ_{II}, σ_{SI} and α for the CDMR are determined through a set of preliminary experiments so that the retrieval accuracy is the highest among the combinations of parameters we tried. Table 1 summarizes parameters for the CDMR used in the experiments below.

We used Mean Average Precision (MAP) [%] and Recall-Precision plot for quantitative evaluation of retrieval accuracy.

Table 1. Parameters for the CDMR

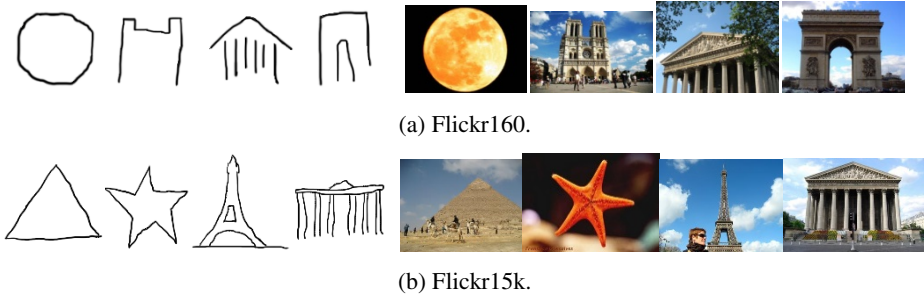| algorithms	Flickr160				Flickr15k			
	σ_{SS}	σ_{II}	σ_{SI}	α	σ_{SS}	σ_{II}	σ_{SI}	α
BF-GFHoG	0.0075	0.02	0.020	0.95	0.04	0.0075	0.05	0.7
BF-GALIF	0.0075	0.02	0,075	0.95	0.04	0.0075	0.05	0.7
BF-GALIF(w)	0.0075	0.02	0.075	0.95	0.04	0.0075	0.05	0.7

(a) Flickr160.

(b) Flickr15k.

Fig. 4. Examples of sketch queries and retrieval target images

3.1   Effectiveness of Edge Weighting by VSW

Figure 5 shows the relationship between vocabulary size k and retrieval accuracy for the Flickr160 and the Flickr15k benchmarks. In the figure, "BF-fGALIF" means BF-fGALIF is extracted directly from Canny edge images, while "BF-fGALIF(w)"   with "(w)" means that the feature is computed from saliency-weighted edge images generated by the VSW. For both benchmarks, at almost all the vocabulary size, the BF-fGALIF(w) with saliency weighting produced MAP scores about 2 % higher than the BF-fGALIF without saliency weighting.

It can be concluded that edges due to background clutter are suppressed to certain degree by the VSW, resulting in small but consistent improvements in retrieval accuracy.
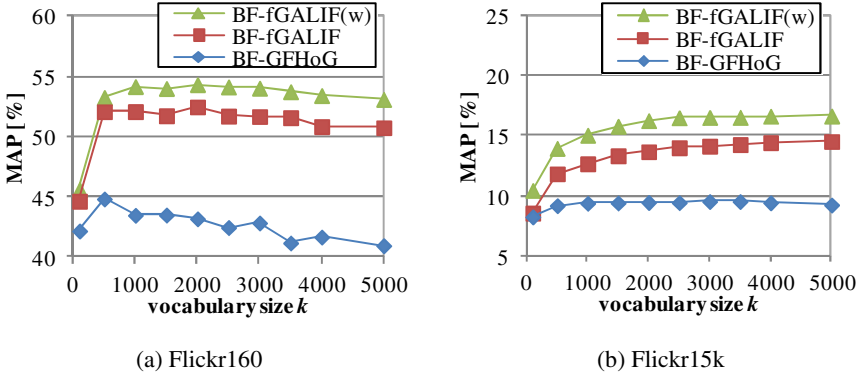

Fig. 5. Vocabulary size and retrieval accuracy. (Please note the difference in MAP scales.).

3.2   Effectiveness of Ranking by CDMR

Table 2 shows ranking performance of the CDMR for the Flickr160 and the Flickr15k benchmarks. In this experiment, the sketch-to-image comparison algorithm for the submatrix $\mathbf{W}_{SI}$ of the CDMR is selected from the following three; the BF-GFHoG, the BF-fGALIF, and the BF-fGALIF(w). We fixed the sketch-to-sketch comparison algorithm for the submatrix $\mathbf{W}_{SS}$ to the BF-fGALIF. We also fixed the image-to-image comparison algorithm for the submatrix $\mathbf{W}_{II}$ to the BF-DSIFT.

For all the three sketch-to-image comparison algorithms we compared, that are, the BF-GFHoG, the BF-fGALIF, and the BF-fGALIF(w), the CDMR significantly improved their retrieval accuracy for the two benchmarks. In case of the Flickr160, the BF-fGALIF(w) feature using the CDMR produced the highest MAP score of 72.3 %. It is about 19 % better than the BF-fGALIF(w) feature without using the CDMR. In case of the Flickr15k, the BF-fGALIF(w) with the CDMR yielded MAP score of 22.4 %. This score is 6 % higher than the BF-fGALIF(w) without the CDMR.

Overall, gain in retrieval accuracy due to the CDMR is quite significant. Diffusion of relevance via multiple paths through the sketches and images of the CDM makes

Table 2. Feature selection and ranking accuracy by the CDMR

algorithms	Flickr160		Flickr15k	
	without CDMR	with CDMR	without CDMR	with CDMR
BF-GFHoG	44.8	64.1	9.5	15.4
BF-fGALIF	51.6	69.4	14.3	20.6
BF-fGALIF(w)	53.7	72.3	16.5	22.5

the ranking robust against stroke noise and other variations in sketches. Also, diffusion of relevance from the query sketch via sketches similar to the query toward target images pushes retrieval accuracy up, as it works like an automatic query expansion.

3.3   Comparison with Other Algorithms

Table 3 shows comparison of retrieval accuracy using the Flickr160 and the Flickr15k benchmarks. The table also lists MAP scores for algorithms reported in Hu et al. [6][7] for comparison. Figure 7 shows recall-precision plots for the 6 algorithms we compared. Every algorithm listed in Table 2 employs a set of local features to compare images with sketches. Note that, in Table 2, MAP scores of the BF-GFHoG are different among our own experiments and theirs found in [6][7]. This discrepancy is due probably to the difference in distance metric; we used Cosine distance for feature comparison, while Hu et al. used Histogram Intersection.

Our proposed CDMR-BF-fGALIF(w), which employs edge weighting by the VSW and ranking by the CDMR, did the best among 12 methods listed in Table 3, for both benchmarks. Our best performer, the CDMR-BF-fGALIF(w), produced MAP=72.3% for the Flickr160 benchmark. In comparison, the MAP score of the BF-GFHoG is 54% [6]. When compared using the Flickr15k benchmark, the CDMR-BF-fGALIF(w) produced MAP score of 22.5 %. This score is about 10 % higher than that of BF-GFHoG, 12.2%, reported in [7]. Most of the gain in retrieval accuracy comes from the CDMR ranking, but contribution due to visual saliency weighting is consistent.

Recall-precision curves of Figure 6 also shows the advantages in retrieval accuracy of the CDMR-BF-fGALIF(w) over the other algorithm.

Table 3. Comparison of MAP scores [%] among several algorithms

algorithms	Flickr160	Flickr15k
BF-GFHoG	44.8	9.5
BF-fGALIF	51.6	14.3
BF-fGALIF(w)	53.7	16.5
CDMR-BF-GFHoG	64.1	15.4
CDMR-BF-fGALIF	69.4	20.6
CDMR-BF-fGALIF(w)	72.3	22.5
BF-GFHoG [6][7]	54.0	12.2
BF-HoG [6][7]	42.0	10.9
BF-SIFT [6][7]	41.0	9.1
BF-SelfSimilarity [6][7]	42.0	9.6
BF-ShapeContext [7]		8.1
BF-StructureTensor [7]		8.0

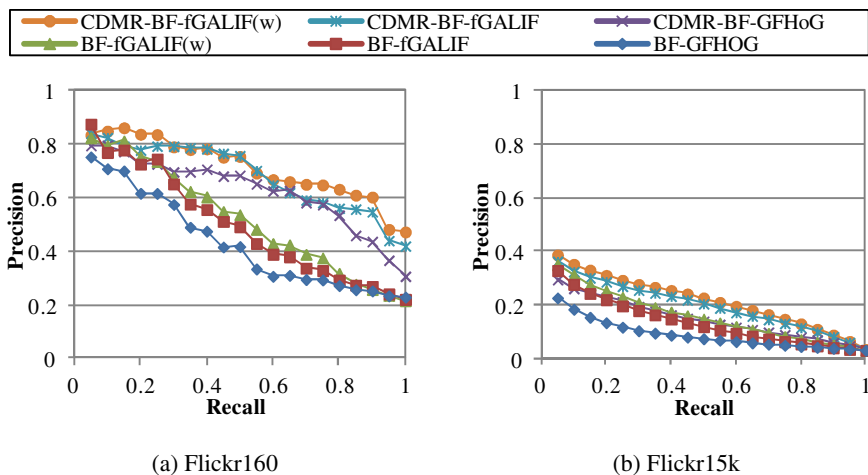(a) Flickr160   (b) Flickr15k

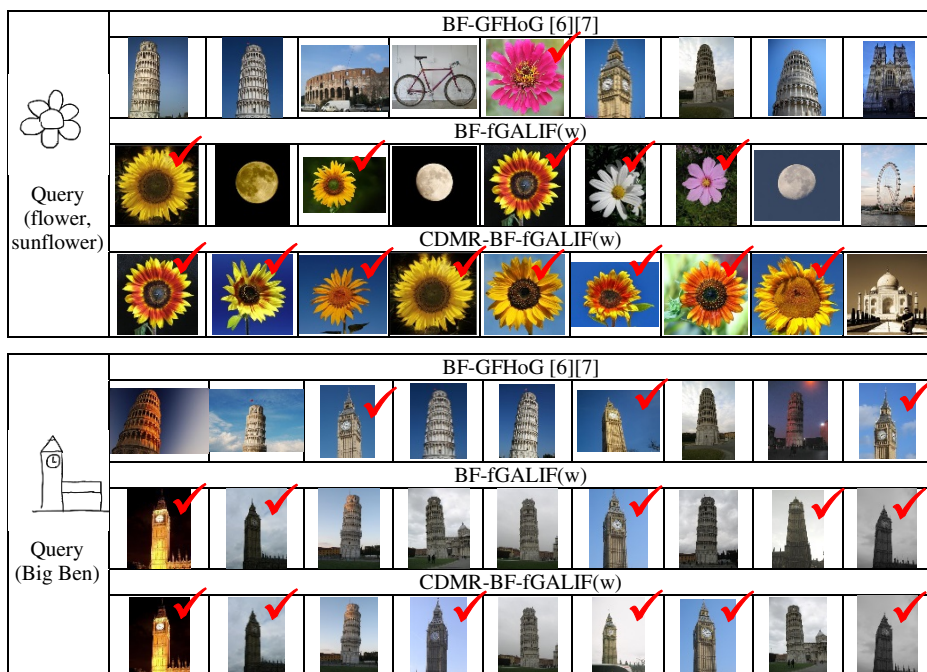**Fig. 6.** Recall-Precision plots

Fig. 7. Retrieval examples for the Flickr15k

4   Conclusion and Future Work

In this paper, we proposed an algorithm for Sketch-Based Image Retrieval (SBIR). A challenge in SBIR is effective comparison of a line drawing sketch and a 2D image. It should be robust against background clutter in the image. It should also be robust against stroke noise in the line drawing sketch. Previous algorithms for SBIR convert a 2D image into an edge image to compare it against a query line drawing sketch. However, unnecessary edges due to background clutter in an image interfere with the feature comparison between a sketch and the 2D image. Stroke noise such as disconnected or wobbly lines also makes the comparison difficult.

Our proposed algorithm first converts an image into edge image, and then performs Visual Saliency Weighting (VSW) to suppress edges due to background clutters. To effectively compare a sketch containing stroke noise with images, we employ a distance metric learning algorithm called Cross-Domain Manifold Ranking (CDMR) [12]. Our experimental evaluation using two SBIR benchmarks showed that the combination of the VSW and the CDMR significantly improves retrieval accuracy.

We are currently looking into the improvement of computational efficiency of the CDMR algorithm, as the CDMR is expensive to compute for a large database.

References

1. Chalechale, A., Naghdy, G., Mertins, A.: Sketch-based image matching using angular partitioning. IEEE Transactions on Systems, Man and Cybernetics 35(1), 28–41 (2005)
2. Eitz, M., Hildebrand, K., Boubekeur, T., Alexa, M.: Sketch-Based Image Retrieval: Benchmark and Bag-of-Features Descriptors. Visualization and Computer Graphics 2011 17(11), 1624–1636 (2011)
3. Saavedra, J., Bustos, B.: An improved histogram of edge local orientations for sketch-based image retrieval. In: Goesele, M., Roth, S., Kuijper, A., Schiele, B., Schindler, K. (eds.) Pattern Recognition. LNCS, vol. 6376, pp. 432–441. Springer, Heidelberg (2010)
4. Cao, Y., Wang, C., Zhang, L., Zhang, L.: Edgel index for large-scale sketch-based image search. In: CVPR 2011 (2011)
5. Bozas, K., Izquierdo, E.: Large scale sketch based image retrieval using patch hashing. In: Bebis, G., et al. (eds.) ISVC 2012, Part I. LNCS, vol. 7431, pp. 210–219. Springer, Heidelberg (2012)
6. Hu, R., Barnard, M., Collomosse, J.: Gradient Field Descriptor for Sketch based Retrieval and Localization. In: ICIP 2010 (2010)
7. Hu, R., Collomosse, J.: A Performance Evaluation of Gradient Field HOG Descriptor for Sketch Based Image Retrieval. In: CVIU 2013 (2013)
8. Yang, C., Zhang, L., Lu, H., Ruan, X., Yang, M.H.: Saliency Detection via Graph-based Manifold Ranking. In: CVPR 2013 (2013)
9. Zhou, D., Bousquet, O., Lal, T.N., Weston, J., Schölkopf, B.: Learning with Local and Global Consistency. In: NIPS 2003 (2003)
10. Eitz, M., Richter, R., Boubekeur, T., Hildebrand, K., Alexa, M.: Sketch-Based Shape Retrieval. ACM TOG 31(4), 1–10 (2012)
11. Lowe, D.G.: Distinctive Image Features from Scale-Invariant Keypoints. IJCV 60(2) (November 2004)

12. Furuya, T., Ohbuchi, R.: Ranking on cross-domain manifold for sketch-based 3D model retrieval. Accepted as Regular Paper, Cyberworlds 2013 (2013)
13. Furuya, T., Ohbuchi, R.: Dense sampling and fast encoding for 3D model retrieval using bag-of-visual features. ACM CIVR 2009, Article No. 26 (2009)
14. Dalal, N., Triggs, B.: Histograms of oriented gradients for human detection. In: CVPR 2005 (2005)
15. Weinberger, K.Q., Saul, L.K.: Distance Metric Learning for Large Margin Nearet Neighbor Classification. In: Journal of Machine Learning Research (JMLR) (2009)
16. Achanta, R., Estrada, F., Wils, P., Süsstrunk, S.: Salient region detection and segmentation. In: Gasteratos, A., Vincze, M., Tsotsos, J.K. (eds.) ICVS 2008. LNCS, vol. 5008, pp. 66–75. Springer, Heidelberg (2008)
17. Hou, X., Zhang, L.: Saliency detection: A spectral residual approach. In: CVPR 2007 (2007)