A SKETCH IMAGE RETRIEVAL SYSTEM USING DIRECTIONAL PROJECTION: DPSIR

Zhihua He, Maja Bystrom

ECE Dept. Boston University, 8 St. Mary's St. Boston, MA 02215, zhihua,bystrom@bu.edu

ABSTRACT

For sketch-based retrieval of images, characterization and comparison of the geometrical structure (shape) information is of great importance. In this paper, a directional-projection-based sketch image retrieval system (DPSIR) is proposed. In the DPSIR system, efficient filter banks are adopted to decompose the image into directional subbands in order to capture edges at different orientations. Through projection of each subband into a pair of orthogonal profiles, the shape information is compactly captured by identifying the structure of each one-dimensional projection profile. By comparing the projection profiles of the sketched image with those in an image database, a set of images with shapes similar to the sketched images are retrieved. Experiments using DPSIR system on an MPEG-7 shape database yield high retrieval rates.

I. INTRODUCTION

Driven by the development of internet and digital imaging devices, easy access to and management of mass image information has become an important challenge. In particular, interest in performing efficient retrieval of images from large databases has grown with use and size of databases. We focus on the retrieval of images on the basis of sketches of major image structures, and propose a directional-projection-based sketch image retrieval system: DPSIR.

Image retrieval has been an active and rapidly advancing research field with contributions in both theory and practice. A thorough review of content-based image retrieval (CBIR) is given in [1], and lists of existing commercial and research CBIR systems are given in [2], [3]. Retrieval systems make use of a variety of image features to describe image information. Depending on the application, users might be interested in image *textures, shapes, colors* or *semantic* content. Thus, various image features must be identified.

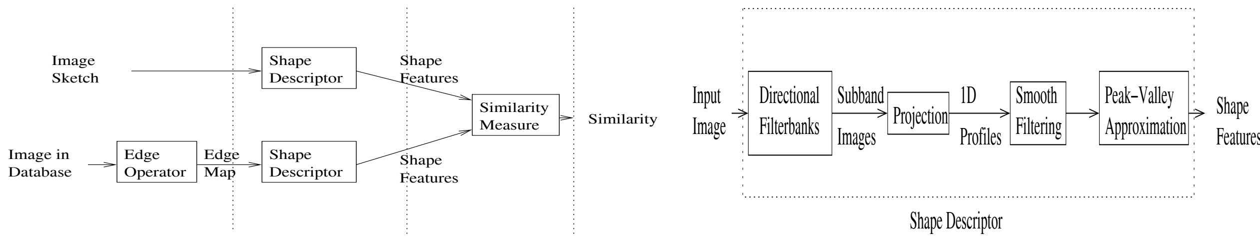
While there are many methods employing different features (e.g. [4]–[10]), we focus on retrieval using sketches, and thus only shape features are utilized. In many shape-oriented image retrieval systems, including DPSIR system, shapes are described as sets of edges, approximated by line segments. Projection-based methods such as the Hough transform, can efficiently capture line features [11]–[13]. However, the Hough transform treats points on integration lines equally, and thus reduces the spatial information of the shapes. Alternatively, Kadyrov and Petrou [14] proposed a trace transform as the generalization of the Hough transform, where selected functionals along the straight lines are utilized. In [15], histograms of edge segments in four orientations of a block of pixels are used as the shape features for retrieval of image sketches. Tang and Wang [16] approached the face sketch recognition problem by first transforming face photos into pseudo-sketches, and then comparing the query sketches with pseudo-sketches in database using an eigenface method.

In the proposed DPSIR system, the image sketch and the edge map of each image in the database are first decomposed into directional subbands through use of directional filter banks [17]; thus, edges with different orientations are captured in different subbands. By projecting each subband onto the principle and orthogonal axes, edges with different orientations, locations and strengths are reduced to a set of 1D projection profiles. Comparison of these profiles determine similarities between the sketches and image edge maps. A similar projection-based scheme is proposed by Rose and Shah in [18], in which spatial domain projections of the edge map along four orientations are used to construct the feature space directly. In contrast, the proposed DPSIR system uses directional filterbanks to isolate edges and thus reduce noise.

The paper is organized as follows. In Section II, a detailed description of the proposed DPSIR system is presented, followed by some experimental results in Section III. Conclusions and suggested future work are given in Section IV.

II. RETRIEVAL SYSTEM: DPSIR

The proposed directional-projection-based sketch image retrieval system consists of two parts, a shape descriptor and a similarity measure, as shown in Fig. 1 (a). The query image sketch and edge maps of images in database are first separately passed through shape descriptors, then the shape features of the two are compared to obtain similarity values.



(a) (b)

Fig. 1. DPSIR system. (a) Overall DPSIR system. (b) Shape Descriptor in DPSIR

II-A. Shape Descriptor

The shape descriptor as shown in Fig. 1 (b) consists of four parts, namely, directional filtering, directional projection to form 1D profiles, lowpass filtering of the 1D profiles, and peak-valley approximation of the 1D profiles.

1) *Directional filtering.*
   In the DPSIR system, an efficient iterative two-level directional filter bank [17] is used to decompose each sketch or image edge map into local directional expansions, so that edges with different orientations are enhanced and captured in different directional subbands. As suggested in [17], the directional filtering resembles local Radon transform; this is conceptually consistent with image retrieval systems using the Hough or Radon transforms. Fig. 2(b) gives an example of the horizontal subband of the edge map Fig. 2(c) of the database image of Fig. 2(a).
2) *Directional projection.*
   To efficiently capture edges at different orientations, projections along each subband's principle and orthogonal directions are performed to yield sets of principle and orthogonal profiles, respectively. This process captures the location and length of edges at different orientations. Fig. 2(b) shows an example of the projection of the horizontal subband in the horizontal and vertical directions. These profiles are also shown in Fig.'s 3(a) and (b).
3) *Lowpass filtering of 1D profiles.*
   Edge isolation through directional filtering is not ideal; thus, any edges not in the direction of a subband contribute noise to the 1D profiles. Both a Gaussian low-pass filter and a fuzzy median filter [19] are adopted to remove the noise, while preserving the true peak information. Fig.'s 3(c) and (d) show the filtered versions of the profiles of Fig.'s 3(a) and (b).
4) *Peak-valley approximation of 1D profile.*
   In order to efficiently represent this shape information, the 1D projection profiles are approximated by linear fitting, so that only the locations and heights of peaks and valleys are preserved. The peak-valley approximation representation is efficient and convenient for both representation of edges and for computation of the similarity metrics. Peaks in the principle projection profiles reflect the locations of edges of the corresponding orientation, while peaks in orthogonal projection profiles give the strength/length of those edges. Note that peaks with relatively small height and width considered to be noise and are removed. Fig.'s 3(e) and (f) give the peak-valley approximation of Fig.'s 3(a) and (b).

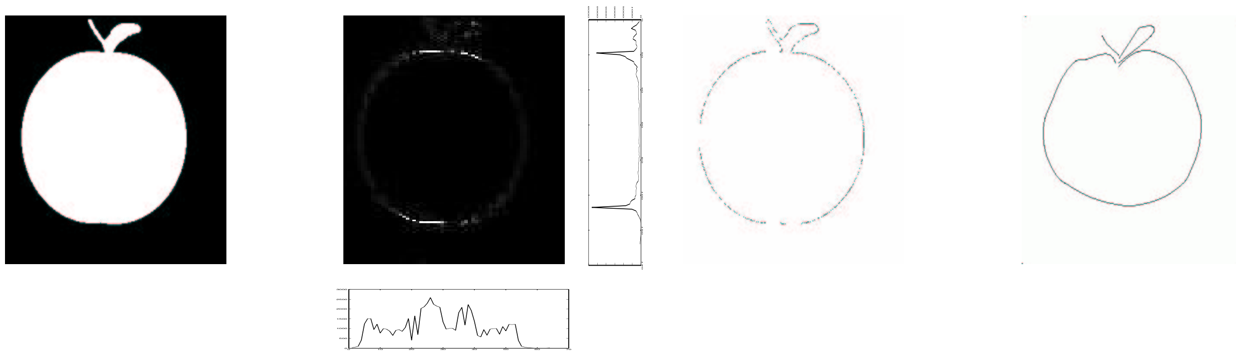
II-B. Similarity Measure

The shape features of the query image sketch are compared with those of each candidate image in the database. The similarity between two images is defined as the distance between the location and strength of edges from the two images, a measure consistent with the human visual system. First, we determine the correspondence between the edges in two images. In the feature space of DPSIR system, we determine the correspondence between the shape features, i.e., the locations of peaks and valleys. The profile with smaller number of peaks is used as an anchor, a, and the target profile, m, is aligned with the anchor by minimizing the Euclidean distance between the peaks in the anchor and those in the target profile.

$$\{L_m\} = arg\min_{L'_m}\sum_{k=1}^{K} D(L_a(k), L'_m(k)), \qquad (1)$$

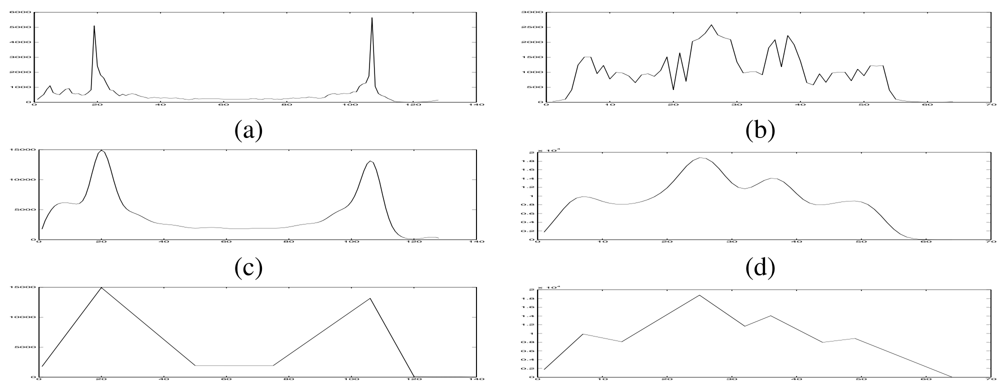
where K is the total number of peaks in the anchor profile. The location of the k^{th} peak of the anchor profile is denoted as $L_a(k)$, $L'_m(k)$ is the choice of the k^{th} peak in a subset, L'_m, of the peaks of target profile, and $D(.)$ denotes Euclidean distance. Note that each locations and heights for a given profile are first normalized by the largest location and height, respectively. The unmatched peaks in the non-anchor profile, denoted as s, also provide information about the distance between two profiles.

Given matched profiles, distance metric for the j^{th}



(a) (b) (c) (d)

Fig. 2. Experimental images. (a) Original image. (b) Horizontal subband image with principle horizontal projection and orthogonal vertical projection. (c) Edge map. (d) Example of image sketch.



(a) (b) (c) (d) (e) (f)

Fig. 3. Image profiles of Fig. 2(b). (a) Profile 1: projecting along principle direction, 0, to subband. (b) Profile 2: projection along orthogonal direction, 90°, of subband. (c) lowpass filtered (a). (d) lowpass filtered (b). (e) peak-valley approximation of (a). (f) peak-valley approximation of (b)

projection of the i^{th} subband is defined as

$$\begin{aligned} E_{ij} &= ((3-j)*L_P+1)\times(L_V+1)\times(H_P+1) \\ &\quad \times(H_V+1)\times(s+1)-1 \quad . \end{aligned}$$

The parameters L_P, L_V, H_P and H_V are the differences between the location of peaks, the location of valleys, the height of peaks and the height of valleys of two matched profiles, La and L_m, respectively. The final similarity measure for the i^{th} subband is $E_i = 2*E_{i1}+E_{i2}$. Because the locations of edges are more salient and robust than the strengths of edges in sketch image retrieval, a higher weight is given to the profile along principle direction, the first/primary projection.

Finally, the distances from each subband are combined together to yield the final distortion value $E = \prod_{i=1}^{\alpha N} E_i^{'}$, where the subband feature distances E_i are ordered from largest to lowest to form the set $E_i^{'}$. The parameter N is the total number of directional subbands, and only α of the lowest distances are used. The similarity value is the inverse of distortion value.

III. EXPERIMENTAL RESULTS

The MPEG-7 Core Experiment CE-Shape-1 part B database [20] is adopted in experiments using the DPSIR system for retrieval of images based on sketches. This database consists of 70 classes of shapes, with 20 images in each class. For each class of shapes, an image sketch is drawn, thus, 70 image sketches are used as queries. The retrieval rate is defined as $R_M = N_M/70$, where N_M is the number of successful retrievals when only M candidate images are preserved. In experiments, all the image sketches and images in database are first resized to 256×256. The Canny edge detector is adopted to generate the edge maps for the database images. Directional filter banks with orientation precision $\pi/8$, yielding 8 directional subbands, are employed, and the distortion parameter $\alpha = 0.75$ is adopted. An example of the original image, directional

Table I. Retrieval results using DPSIR system for MPEG-7 shape database.

Methods	R_{20}	R_{15}	R_{10}	R_5
DPSIR	0.94	0.94	0.94	0.90
Spatial-domain histogram [15]	0.81	0.81	0.76	0.71

subband image, edge map and image sketch are given in Fig. 2. The retrieval results using proposed DPSIR system and spatial-domain histogram method [15] are compared in Table I. The proposed DPSIR system provides retrieval rates which are consistently over 90%, and are higher than those using the method of [15].

IV. CONCLUDING REMARKS

A directional-projection-based sketch image retrieval system, DPSIR, is described in this paper. Shape information is captured through directional filtering and directional projection. A similarity measure is developed and results using an MPEG-7 shape database are shown. Higher retrieval rates are obtained using DPSIR system than using a spatial-domain sketch retrieval method of [15]. In cases of distorted query images, the authors have been working on translation and scaling invariance through alignment of 1D projection profiles; similarly, rotation invariance can approached through alignment of the profiles from different directional subbands. In cases of images with deformations, with the assumptions of the distribution for location and strength of edges, corresponding distributions of the locations and heights of peaks in profiles can be used to perform hypothesis testing.

V. REFERENCES

[1] A. W. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, "Content-based image retrieval at the end of the early years," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 22, pp. 1349–1380, Dec. 2000.

[2] B. Johansson, "A survey on: Contents based search in image databases," *http://www.isy.liu.se/cvl/Projects/VISIT-bjojo/survey/surveyonCBIR/surveyonCBIR.html*.

[3] R. C. Veltkamp and M. Tanase, "Content-based image retrieval systems: A survey," *http://www.aa-lab.cs.uu.nl/cbirsurvey/cbir-survey/cbir-survey.html*, 2001.

[4] I. Cox, M. Miller, T. Minka, T. Papathomas, and P. Yianilos, "The bayesian image retrieval system, pichunter: theory, implementation, and psychophysical experiments," *IEEE Trans. Image Process.*, vol. 9, no. 1, pp. 20–37, Jan. 2000.

[5] A. Natsev, R. Rastogi, and K. Shim, "Walrus: A similarity retrieval algorithm for image databases," *IEEE Trans. on Knowledge and Data Engineering*, vol. 16, pp. 301–316, March 2004.

[6] Y. D. Chun, S. Y. Seo, and N. C. Kim, "Image retrieval using bdip and bvlc moments," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 13, pp. 951–957, Sept. 2003.

[7] T. Gevers and A. Smeulders, "Pictoseek: Combining color and shape invariant features for image retrieval," *IEEE Trans. Image Process.*, vol. 9, pp. 102–119, Jan. 2000.

[8] C. Grigorescu and N. Petkov, "Distance sets for shape filters and shape recognition," *IEEE Trans. Image Process.*, vol. 12, pp. 1274–1286, Oct. 2003.

[9] S. Berretti, A. D. Bimbo, and P. Pala, "Retrieval by shape similarity with perceptual distance and effective indexing," *IEEE Trans. on Multimedia*, vol. 2, pp. 225–239, Dec. 2000.

[10] E. Petrakis, Diplaros, and E. Milios, "Matching and retrieval of distorted and occluded shapes using dynamic programming," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 24, pp. 1501–1516, Nov. 2002.

[11] P. Tipwai and S. Madarasmi, "Image search using deformable contours," *IEEE International Conference on Image Processing*, vol. 1, pp. 22–25, Sept. 2002.

[12] P. Franti, A. Mednonogov, and H. Kalvianen, "Hough transform for rotation invariant matching of line-drawing images," *15th International Conference on Pattern Recognition*, vol. 4, pp. 389–392, Sept. 2003.

[13] P.-F. Fung, W.-S. Lee, and I. King, "Randomized generalized hough transform for 2-d gray scale object detection," *Proceedings of the 13th International Conference on Pattern Recognition*, vol. 2, pp. 511–515, Aug. 1996.

[14] A. Kadyrov and M. Petrou, "The trace transform and its applications," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 23, pp. 811–828, Aug. 2001.

[15] M. Abdel-Mottaleb, R. Phillips, and B. Manor, "Image retrieval based on edge representation," *IEEE International Conference on Image Processing*, vol. 3, pp. 734–737, Sept. 2000.

[16] X. Tang and X. Wang, "Face sketch recognition," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 14, pp. 50–57, Jan. 2004.

[17] M. Do and M. Vetterli, "The contourlet transform: An efficient directional multiresolution image representation," *IEEE Trans. Image Process.*, to appear.

[18] J. Rose and M. Shah, "Content-based image retrieval using gradient projections," *Southeastcon'98 Proceedings, IEEE*, pp. 118–121, Apr. 1998.

[19] Y. Nie and K. Barner, "Fuzzy weighted median filters," *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, vol. 2, pp. 1125–1128, 2002.

[20] I. J. Latecki, R. Lakamper, and U. Eckhardt, "Shape descriptors for non-rigid shapes with a single close contour," *IEEE Conf. Computer Vision and Pattern Recognition*, pp. 424–429, 2000.