International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:3, No:1, 2009

Salient Points Reduction for Content-Based Image Retrieval

Yao-Hong Tsai

***Abstract*—**Salient points are frequently used to represent local properties of the image in content-based image retrieval. In this paper, we present a reduction algorithm that extracts the local most salient points such that they not only give a satisfying representation of an image, but also make the image retrieval process efficiently. This algorithm recursively reduces the continuous point set by their corresponding saliency values under a top-down approach. The resulting salient points are evaluated with an image retrieval system using Hausdoff distance. In this experiment, it shows that our method is robust and the extracted salient points provide better retrieval performance comparing with other point detectors.

***Keywords*—**Barnard detector, Content-based image retrieval, Points reduction, Salient point.

I. Introduction

Digital imagery has become increasingly widespread due to the popularity of digital cameras and online photo-sharing services such as Picasa, the Google image search. By traditional technology, many images are made available via keyword-based search and indexes from billions of images. Content-based image retrieval (CBIR) is the technology that in principle helps to organize digital picture archives by their visual content [1], [2]. Over the past years, many techniques have been proposed for solving the problem of CBIR. Some papers [3] provide an exhaustive survey and the comparison of performance measures over many techniques for image retrieval.

Traditionally, a feature vector is computed to represent an image in CBIR. For a given query image, its feature vector is first computed, then compared to the stored feature vectors of the images in the image database. Afterwards, the system returns the images that are most similar to the query one. For measuring the image similarity, those feature vectors often come from the low-level features such as color, texture and shapes [4], [5] to handle all parts of the natural images which may have different characteristics. Image retrieval approaches based on salient points have been proved to be the effective features to describe the image [6], [7]. It is quite easy to understand that using a subset of the image pixels, the salient points, instead of the all image reduces the amount of data to be processed. In addition, salient points have also been combined with other features in solving CBIR and image matching problems [8]–[11].

Y. H. Tsai is with the Department of Computer Science, Hsuan Chuang University, Hsinchu, 300 Taiwan (corresponding author to provide phone: 886-3-5302255 ext. 2160; fax: 886-3-5391244; e-mail: tyh@ hcu.edu.tw).

Classic corner detectors [12], [13] can also be used to find the salient points. These detectors can detect points that are visually corners of the image, and features around these points can be computed to obtain the image's local properties. However, they have some drawbacks when applied to various natural images for image retrieval, because visual features need not be corners and corners may gather in small regions. From the above reasons, the number of points is usually too large such that it need to keep the local most salient ones and eliminate those redundant points.

One way to reduce the number of salient points is to fix the number of points, say P. It can be implemented by selecting the first N points with maximum saliency values. By this approach, some of the selected points will still joint together and it may need different number of salient points for different input images. For example, a small number N usually can not extract enough information for a large image. Another way is to select a fixed threshold, T, any point in the original set will be extracted if its saliency value is above T. However, some drawbacks are similar as the first approach.

Loupias *et al*. [7] proposed a detecting algorithm for salient points using the representation of multi-resolution of wavelet transform. Since local absolute maximum values of the wavelet coefficients reflect a signal's singularity points, they reflect the edges or boundary of the objects in an input image. Hence, the algorithm extracts salient points where variations occur in the image. Song *et al*. [14] provide a scheme to extract color salient points according to the wavelet transform and the Barnard Detector. To describe different parts of the image, the set of salient points should not be clustered in few regions. A point reduction method is used to get rid of the continuous points that are not so salient.

In this paper, a reduction algorithm that extracts the local most salient ones is proposed. It can give a satisfying representation of an image for image retrieval. The proposed algorithm recursively subdivides the whole image into four equal sized subimages and computes the threshold for each one dynamically. Such that we can extract the local most salient points in every part of the input image. By the way, it does not need a pre-specified or fixed global threshold to choose salient points. The resulting salient points are evaluated by the application of image retrieval with a system using the Hausdoff distance. In this experiment our method provides better retrieval performance comparing with other point detectors.

The outline of the paper is listed as follows. Section 2 gives a brief overview of the salient point detector including the Barnard detector and the wavelet saliency value. In Section 3, we describe the proposed the reduction technique for salient points. Experimental results are presented in Section 4. Section 5 concludes this paper.

II. Salient Point Detection

First, we introduce the Barnard detector [15] for extracting the salient points from the original image. For a given image *f*(*m*,*n*) and the non-overlapped processing window *W*, the operator *t*(*m*,*n*) is defined to measure the strength of the center point from the difference in horizontal, vertical and diagonal directions. For each pixel in the processing window, we calculate the operator *t*(*m*,*n*) by the following formula.

$$t(m,n) = \min\{H, V, L, R\},$$

where

$$H = \left[f(m,n) - f(m-1,n)\right]^2 + \left[f(m,n) - f(m+1,n)\right]^2$$

$$V = \left[f(m,n) - f(m,n-1)\right]^2 + \left[f(m,n) - f(m,n+1)\right]^2$$

$$L = \left[f(m,n) - f(m+1,n-1)\right]^2 + \left[f(m,n) - f(m-1,n+1)\right]^2$$

$$R = \left[f(m,n) - f(m+1,n+1)\right]^2 + \left[f(m,n) - f(m-1,n-1)\right]^2.$$

Then, find a pixel *f*(*m*', *n*') to make *t*(*m*',*n*') maximum in window *W* by the following equation.

$$t(m', n') = \max_{m,n \in W}\{t(m,n)\}.$$

Note that the differences in gray value of *f*(*m*',*n*') and its neighbors in window *W* are large.

Secondly, two-dimensional wavelet transform need three different wavelet functions, which reflect three spatial orientations (horizontal, vertical and diagonal). The wavelet representation of an image of size $M \times N$ is the set of wavelet coefficients. The wavelet coefficient can be expressed as the form, $W_{2^j} f(m,n)$ at scale 2^j, where $j \in Z$ and $j \leq -1$. The children set of wavelet coefficient is as follows:

$$C(W_{2^j} f(m,n)) = \{W_{2^{j+1}} f(k,l)\},$$

where $2m \leq k \leq 2m+2p-1$, $2n \leq l \leq 2n+2p-1$, p is the wavelet regularity. Loupias *et al.* [7] select salient points with the highest gradient from $2p$ points. The saliency value is as the sum of the absolute value of the wavelet coefficients.

$$S = \sum_{k=1}^{-j} \left| C^{(k)}(W_{2^j} f(m,n)) \right|$$

For a point in an image, its corresponding saliency value is computed for every wavelet coefficient. We then have to threshold the saliency value, in relation to the desired number of salient points.

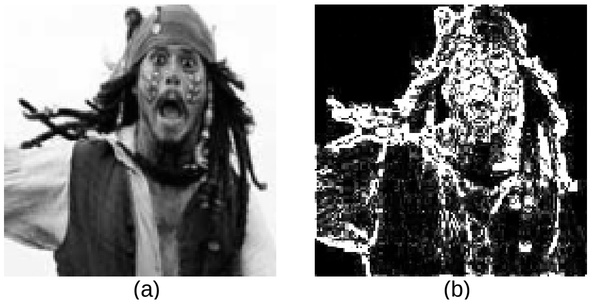

Fig. 1. (a) the original image (b) the points obtained from Barnard detector

Fig. 1 (b) shows the result of Barnard detector from the input image in Fig. 1(a). There are total 2669 potential salient points in Fig. 1 (b). It is obvious that this result contains many continuous point sets. There are still some problems about the point extraction above. The number of points is usually too large and there exists some continuous point sets. Since it is believed that not all of them are needed, we try to reduce the set of points to extract the local most salient ones.

III. Salient Points Reduction

In Song's algorithm [14], there are two main approaches to improve the previous algorithm [7]. First, they choose a threshold varies according to the content of the input image. Calculate the sum of the saliency value of all the potential salient points, and the threshold is set to a given percentage threshold *p* of the sum. Extract the most salient points continuously until the sum of their saliency values reaches the threshold. Second, for a chosen point *f*(*m*,*n*), check the eight neighbors of it, if there are still other salient points, decide whether the saliency value of *f*(*m*,*n*) is the maximum, if not, get rid of it. After applying this method to all the salient points, there are no two salient points join together.

In the proposed algorithm, we subdivide the original image into four equal sized ones recursively and select salient points from the above generated set to describe different parts of the image. Suppose the total sum of the saliency value is *S*, then S_i, where *i*=1,2,3,4, represents the corresponding saliency value for each subimage. The recursive relation among saliency values is

$$\left\{ S = \sum_{i=1}^{4} S_i \right\}, \left\{ S_j = \sum_{i=1}^{4} S_{ji}, j = 1,2,3,4 \right\}, \ldots .$$

For the efficiency of CBIR, the number of salient points is fixed, since it takes a long time to perform the point matching. Let N be the total number of salient points that we specify for further image retrieval. The proposed algorithm controls the number of salient points in each subimage recursively by its corresponding saliency value until the subimage that contains only one salient point is reached. The equations is

$$\left\{ N_i = N \times \frac{S_i}{S} \right\}, \left\{ N_{ij} = N_i \times \frac{S_{ij}}{S_i}, \right\}, \ldots .$$

For real application, we can set a threshold for the sum of saliency values to reduce the computation time. If the sum is smaller than the threshold, the subdivision of the image is stopped.

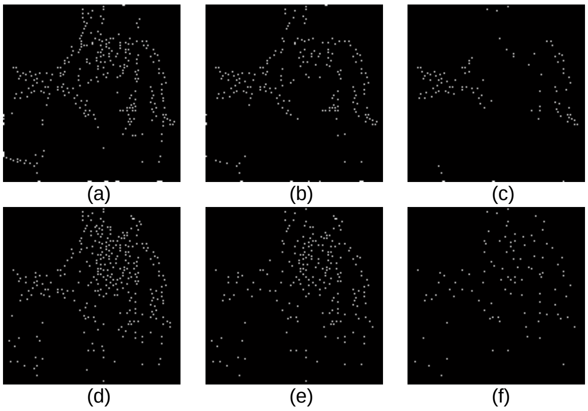

Fig. 2. (a)-(c) the salient points extracted by Song's algorithm (d)-(f) the salient points extracted by the proposed algorithm

The first (second) row in Fig. 2 shows the salient points extracted by Song's algorithm (the proposed algorithm), respectively. From the leftmost column, we set the upper bound of the number of salient points to be 300, 200, and 100. It is obvious that the result of the proposed algorithm successfully reduces the redundant points and gives a satisfactory description for different parts of the input image.

IV. EXPERIMENTATIONS

We test the extracted salient points by image retrieval on the test image database. A simplified CBIR is developed by Hausdorff distance [16], since the Hausdorff distance can be calculated without the explicit pairing of points in their respective data sets, and can be extended to allow partial matching. Hence, Hausdorff distance is widely used in the field of computer vision applications, although it was originally proposed for binary image comparison. Because of this desirable property, we use it to measure the proposed algorithm. Given two finite points sets $A=\{a_1, a_2, \ldots ,a_m\}$ and $B=\{b_1, b_2, \ldots ,b_n\}$, the Hausdorff Distance is defined as

$$H(A,B) = \max\{h(A,B), h(B,A)\},$$

where

$$h(A,B) = \max_{a \in A} \min_{b \in B} \|a - b\|$$

is called the directed Hausdorff distance from A to B. Since the measure by Hausdorff Distance is very sensitive to the degradations caused by noise and occlusions, many improved methods [17]–[19] have been proposed to solve real problems in computer vision .

A fish image database was collected from the website of the Fisheries Agency, Council of Agriculture, Executive Yuan (http://www.fa.gov.tw/chn/index.php). It contains 375 images and is used to test the performance of the proposed algorithm and comparison with Song's algorithm. For the database, it was classified into 10 categories of images such as fish, crab, shrimp, etc. Our experimental platform is a personal computer with an Intel P4 2.8GHz CPU.

The comparison between the proposed method and the existing methods is illustrated in Fig. 3 and Fig.4. The query result by salient points generated by Song's algorithm (the proposed algorithm) is shown in Fig. 3 (Fig. 4). The layout of Fig. 3 and Fig. 4 is described as follows. The query image is shown in the upper left corner and we can see the image of the extracted salient points, denoted with white spots, below the query image. The upper right area in the figure shows 12 query results of the experimentation. It only shows a small part of the comparison results for the limitation of the literature. For more detailed information, the figure also lists the small 100 values of Hausdorff distance in the experimentation. The blue (green) curve presents the Hausdorff distance from the query image to 100 images in the database by the salient points of Song's algorithm (the proposed algorithm).

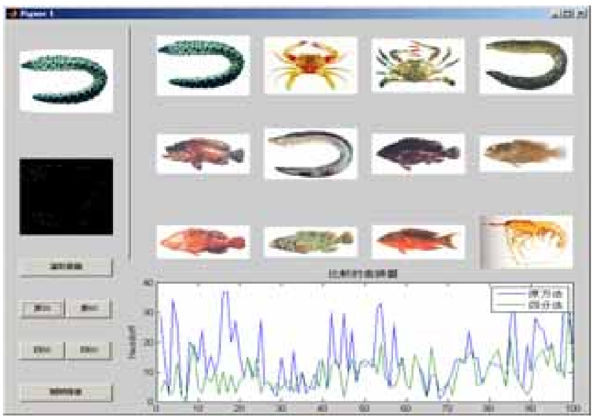

Fig. 3. Results of the CBIR using salient points generated by Song's algorithm.

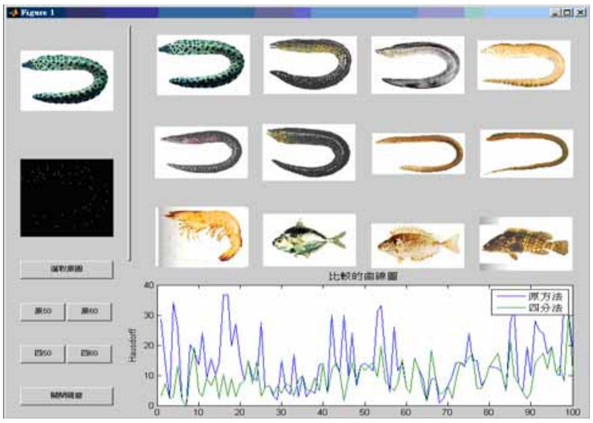

Fig. 4. Results of the CBIR using salient points generated by the proposed algorithm.

. It is obviously that the results of the proposed algorithm are very similar to the query image. On the other hand, we can see that the green curve is almost under the blue one, i.e., the selected top 100 images is more similar to the query one. From the experimental results, the proposed method has better performance than Song's method.

V. Conclusion

Salient points are frequently used to represent local properties of the image in content-based image retrieval. In this paper, we present a reduction algorithm that extracts the local most salient points such that they can give a satisfying representation of an image for applications in computer vision. The resulting salient points are evaluated by image retrieval with a retrieval system using the measure of Hausdoff distance. In this experiment, our method provides better retrieval performance comparing with other point detectors**.**

References

[1] M. Fiala, “Using normalized interest point trajectories over scale for image search,” *Proceedings of the 3rd Canadian Conference on Computer and Robot Vision*, CRV 2006. pp. 58 – 58.

[2] Z.H. Zhang Y. Quan, W.H. Li, W. Guo, “A New Content-Based Image Retrieval”, *Proceedings of the Fifth International Conference on Machine Learning and Cybernetics*, Dalian, 13–16 August 2006. pp. 385–388.

[3] R. Datta, D. Joshi, J. Li, and J. Z. Wang, “Image retrieval: ideas, influences, and trends of the new age,” *ACM Computing Surveys*, vol. 40, no. 2, Article 5, 2008.

[4] H. J. Lin, Y. T. Kao, S. H. Yen, and C. J. Wang, “A study of shape-based image retrieval,” *Proceedings of the 24th International Conference on Distributed Computing Systems Workshops*, ICDCSW 2004. pp. 118–123.

[5] L. Birgale, M. Kokare, and D. Doye, “Colour and texture features for content based image retrieval,” *Proceedings of the International Conference on Computer Graphics, Imaging and Visualisation*, CGIV 2006. pp. 146–149.

[6] J. W. Han and L. Guo, "New image retrieval approach based on interest points", Proceedings of SPIE, vol. 4862, 2002, pp. 197–197.

[7] E. Loupias, N. Sebe, S. Bres, and J.-M. Jolion, “Wavelet-based salient points for image retrieval,” *Proceedings of International Conference on Image Processing*, vol. 2, 2000, pp. 518–521.

[8] H. Ling, and D. W. Jacobs, “Deformation invariant image matching,” *Proceedings of the Tenth IEEE International Conference on Computer Vision*, ICCV 2005, vol. 2, pp. 1466–1473.

[9] J. Wang, H. Zha, and R. Cipolla,“Combining interest points and edges for content-based image retrieval,” *Processings of the ICIP in IEEE International Conference*, 2005. pp. III - 1256-9.

[10] Md. M. Rahman, B. C. Desai, and P. Bhattacharya, “Visual keyword-based image retrieval using latent semantic indexing, correlation-enhanced similarity matching and query expansion in inverted index,” *10th International Database Engineering and Applications Symposium*, IDEAS 2006, pp. 201–208.

[11] G. Ding, Q. Dai, W. Xu, and F. Yang, “Affine-invariant image retrieval based on Wavelet interest points,” *Processings of the IEEE Multimedia Signal*, 2005, pp. 1–4.

[12] C. Schmid and R. Mohr, “Local gray value invariants for image retrieval,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 5, pp. 530–535, 1997.

[13] P. B. Albee and G. C. Stockman, “Interest points from the radial mass transform,” *IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 2, pp. 261–264, 2005.

[14] H. Song, B. Li and L. Zhang, “Color salient points detection using wavelet,” *Proceedings of the 6th World Congress on Intelligent Control and Automation*, Dalian, China, June 21–23, 2006, pp. 10298–10301.

[15] S. T. Barnard, and W. B. Thompson, “Disparity analysis of images,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 333–340, 1980.

[16] D. P. Huttenlocher, G. A. Klanderman, and W. J. Rucklidge, “Comparing images using the Hausdorff distance,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 15, no. 9, pp. 850–863, 1993.

[17] B. Takacs, “Comparing face images using the modified hausdorff distance,” *Pattern Recognition*, vol. 31, no. 12, pp. 1873–1881, 1998.

[18] Y. Wang, and C. Chua, “Robust face recognition from 2D and 3D images using structural Hausdorff distance”, *Proceedings of the ICARCV in International Conference*, 2006. pp. 502 – 507.

[19] H. Tan, Y. Zhang, “A novel weighted Hausdorff distance for face localization,” *Image and Vision Computing*, vol. 24, no. 7, 2006. pp. 656-662.