Proc. VIIth Digital Image Computing: Techniques and Applications, Sun C., Talbot H., Ourselin S. and Adriaansen T. (Eds.), 10-12 Dec. 2003, Sydney

Visual Investigation Using Circular Partitioning of Abstract Images

Abdolah Chalechale[1], Golshah Naghdy[1], and Alfred Mertins[2]

[1] University of Wollongong,
Wollongong, Australia
{ac82,g.naghdy}@uow.edu.au
[2] University of Oldenburg
Oldenburg, Germany
alfred.mertins@uni-oldenburg.de

Abstract. This paper presents a novel approach for sketch-based image retrieval based on low-level features. The approach enables measuring the similarity between a full color image and a simple black and white sketched query and needs no cost intensive image segmentation. The proposed method can cope with images containing several complex objects in an inhomogeneous background. Two abstract images are obtained using strong edges of the model image and thinned outline of the sketched image. Circular-spatial distribution of pixels in the abstract images is used to extract new compact and effective features. The extracted features are scale and rotation invariant and tolerate small translations. The major contribution of the paper is in rotation invariance property of the proposed approach. A collection of paintings and sketches (ART BANK) is used for testing the proposed method. The results are compared with three other well-known approaches within the literature. Experimental results show signi£cant improvement in the Recall ratio using the proposed features.

1 Introduction

Advancements in digital imaging technology especially in cameras and scanners, have resulted in the availability of constantly increasing digital images. Image retrieval from multimedia databases is still an open problem. Traditional textual methods have been shown to be inef£cient and insuf£cient for searching in visual data. Consequently, image and video content-based indexing and retrieval methods have attracted many new researchers in recent years [1, 2].

Representative content-based systems are QBIC [3], Photobook [4], FourEyes [5], MetaSEEk and VisualSEEk [6, 7]. MPEG-7 standard de£nes descriptors for three main image content features that are color, texture and shape [8, 9]. VisualSEEk is also using object layout as another content key.

User interaction is one of the most important aspects of any multimedia system. A simple user interface makes the system more attractive and applicable. In sketch-based image retrieval, where the query example is a rough and simple black and white hand-drawn sketch, color and texture lose their original ability to serve as the content key. Visual shape descriptors are useful in sketch-based image retrieval only when the model and the query images contain one object in a plain background [10]. In multiple-object scenes, object layout is a powerful tool, but object extraction and segmentation costs and rotation variance are the drawbacks.

We focus our discussion on the problem of £nding image features, invariant to scale and rotation, which can be used ef£ciently in sketch-based retrieval where the images have several objects in an inhomogeneous background. To our knowledge, the work of Hirata and Kato, Query by Visual Example (QVE) [11], is the only one that addresses this problem. IBM has adopted a modi£ed version of the approach in its QBIC system [3]. In this approach the query and the target images are resized to 64×64 pixels £rst and then measuring the similarity is performed using a block correlation scheme. The approach does not allow indexing and is computationally expensive. Although the method can tolerate small local rotations, it is not rotation invariant and does not allow for large global rotations.

The edge histogram descriptor (EHD) was proposed in the visual part of the MPEG-7 standard [12, 13]. It originally consists of an 80-bin histogram. A given image is divided into 16 sub-images (4×4) £rst, local edge histograms are then computed for each sub-image based on the strength of £ve different edges (vertical, horizontal, $45°$ diagonal, $135°$ diagonal, and isotropic). S. Won et al. proposed the ef£cient use of the descriptor by extending the histogram to 150 bins [13]. The extended histograms are obtained by further grouping the image blocks into 13 clusters (4 vertical, 4 horizontal and 5 square). A 65-bin semi-global histogram and a 5-bin global histogram are added to make a 150-bin histogram. The EHD is basically intended for gray-image to gray-image matching but changing the internal parameter Th_{edge}, a threshold for edge extraction, could make it applicable for black and white queries in sketch-based retrieval.

Using histograms of edge directions for representing image information is one of the well-known methods in the image retrieval £eld. The feature is appropriate also for sketch-based image matching as it compares the distribution of edge points in the edge map of the query image with the corresponding information of the model image. M. Abdel-Mottaleb [14] uses the approach by applying the Canny edge operator to £nd strong edges in an image and then quantizes them into 4 directions. Jain and Vailaya [15] also proposes edge directions as an image attribute for shape description and apply the method in a trademark registration process [16].

Edge Pixel Neighborhood Information (EPNI) method [17] is employing neighborhood structure of the edge pixels to make an extended feature vector . The vector is used ef£ciently for measuring the similarity between sketched queries and arbitrary model images. The semantic power of the method is examined in [18]. Although the method is scale and translation invariant it does not exhibit rotation invariance property.

In this paper we present a novel approach of feature extraction for matching purposes based on circular partitioning of abstract images. Abstract images are obtained from the model image and from the query image by two different procedures. Circular-spatial distribution of pixels in the abstract images is then employed as the key concept for feature extraction. The extracted features are scale and rotation invariant and also tolerate small translations. The major contribution of the paper is in rotation invariance property as, based on the existing literature, there is no rotation invariant approach dealing with arbitrary images in the sketch-based retrieval.

An art image experimental database, called ART BANK, has been developed and used in this study. The ART BANK consists of two parts. The £rst part is a collection of artistic full color paintings and images. The second part contains several rough, black and white sketches similar to arbitrary candidates from the £rst part and their rotated versions. These are served as sketched queries while the images in the £rst part are used as the image models. The proposed algorithm and three other well-known approaches are implemented and examined on the ART BANK. Experimental results show signi£cant improvement in the Recall ratio using the new approach.

The details of the proposed approach are discussed in the next section. Section 3 exhibits experimental results and evaluation. Section 4 concludes the paper and poses some new directions.

2 Circular Partitioning of Abstract Image (CPAI)

The main objective of the proposed approach is to transform the image data into a new structure that supports measuring the similarity between a full color image and a black and white hand-drawn simple sketch. The proposed algorithm is scale and rotation invariant and also tolerates small translations.

The edge map of an image carries the solid structure of the image independent of the color attribute. Its applicability is well known in computer vision, pattern recognition and image retrieval. Edges are also proven to be a fundamental primitive of an image for the preservation of both the semantics and the perceived attribute [19]. Furthermore, in sketch-based image retrieval, it is the most useful feature that can be employed for matching purposes [17, 10, 20].

According to the assumption that sketched queries are more similar to the edge maps which contain only the perceptive and vigorous edges, we obtain two abstract images through the strong edges of the model image and the thinned version of the query image. The proposed features are then extracted from the abstract images.

2.1 Abstract Images

Two abstract images are obtained for the model image and for sketched query. The full color model image is initially converted to a gray intensity image I, by eliminating the hue and saturation while retaining the luminance. The edges are then extracted using the Canny operator [21] with $\sigma = 1$ and Gaussian mask of size = 9 using the following procedure for depicting the most perceived edges.

The values of high and low thresholds for the magnitude of the potential edge points are automatically computed in such a way that only the strong edges remain. This improves the general semblance of the resulted edge map and the hand drawn query. In order to depict strong edges, let G be the Gaussian 1-D £lter and let g be the derivative of the Gaussian used in the Canny edge operator. Then

$$H(k) = \sum_i G(i)g(k+1-i)$$

is the 1-D convolution of the Gaussian and its derivative.

$$X(\epsilon, \xi) = [\sum_{j=1}^{\Xi} I'(\epsilon, j) H(\xi - j)]'$$

$$Y(\epsilon, \xi) = \sum_{i=1}^{\Upsilon} I(i, \xi) H(\epsilon - i)$$

for $\epsilon = 1, 2, 3, \ldots \Upsilon$ and $\xi = 1, 2, 3, \ldots \Xi$ are the vertical and horizontal edge maps respectively, where Υ is the number of rows and Ξ is the number of columns in the image I. The magnitude of the edge points is then obtained as

$$\Gamma(\epsilon, \xi) = \sqrt{X(\epsilon, \xi)^2 + Y(\epsilon, \xi)^2}.$$

For ef£cient selection of the high and low thresholds, we then make a 64-bin cumulative histogram of the $\Gamma(\epsilon, \xi)$ values and £nd the minimum index ι in this cumulative histogram that is grater than $\alpha * M * N$, where α denotes the percentage of non edge points in the image ($\alpha = 0.7$ is an adequate choice for many images). To retain strong edges of the image, $\beta * \iota$ is selected as the high threshold value and $0.4\beta * \iota$ is used for the low threshold value in the Canny edge operator. β is a parameter that controls the degree of the strength of the edge points. Higher β's lead to lower number of edge points but more perceptive ones (see Figure 1). Consequently, the gray image I is converted to edge image P using the Canny edge operator exploiting the above automatic extracted thresholds.

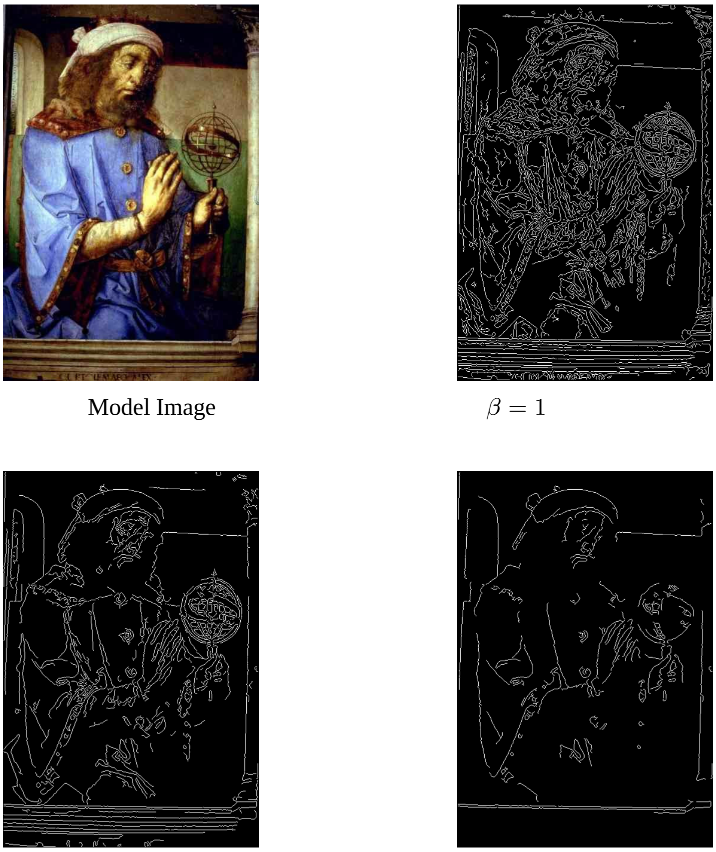
For the query images, the procedure of black and white morphological thinning [22] is applied to extract a thinned version of the sketched image. This image, namely Q, shows an outline of the query and contains the main structure of the user request. It contains spatial distribution of pixels similar to the strong edge map of the model image P.

Comparable P and Q images are then normalized to $J \times J$ pixels, using nearest neighbor interpolation. The proposed normalization of P and Q images ensures the scale invariance and translation robustness properties. The resulted image is called *abstract image* Ω and used for the next feature extraction scheme.

2.2 Circular Partitioning

Based on the fact that any rotation of a given image, with respect to its center, moves a pixel at (ρ, θ) to a new position at $(\rho, \theta + \tau)$, we de£ne a circle of radius r at the center of the abstract image Ω and also $N - 1$ circular strips around the circle with width r. A typical strip S_i has the boundary concentric circles with radii ir and $(i+1)r$ respectively, for $i = 1, 2, 3 \ldots N - 1$. The largest strip has the boundary circles with the radios of $R - r$ and R, where R is the radius of the surrounding circle of image Ω (see Figure 2).

The number of edge points in the central circle and in each circular strip of Ω are chosen to represent the image. The scale and rotation invariant image features are then $\{f(k)\}$ where



Model Image $\beta = 1$

$\beta = 2$ $\beta = 3$

Fig. 1. The effect of β parameter on the Canny edges at one of the art work images

$$f(k) = \sum_{\rho=(k-1)R/N}^{kR/N} \sum_{\theta=0}^{\theta=2\pi} \Omega(\rho, \theta)$$

for $k = 1, 2, 3 \ldots N$.

The features extracted above are rotation invariant because the edge points in any circular strip S_i remain in the same strip with any rotation. The feature can also tolerate small translations as the inner edge points remain in the same circular strip with small translations. The strip width $r = R/N$ is a discretionary variable that controls the translation invariance property. Figure 3 shows an example of Ω image, extracted from a sketched query, and its circular partitions.

The similarity between the model and the sketched images is measured by the ℓ_1 (Manhattan) distance between the two feature vectors. Experimental results (Section 3) con£rm the robustness and ef£ciency of the method.

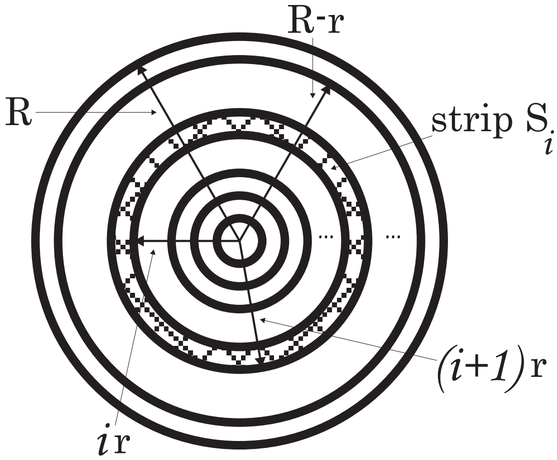


Fig. 2. The central circle and the $N - 1$ circular strips are used for circular partitioning

3 Experimental Results and Ef£cacy Evaluation

This section presents experimental results using the new proposed approach in comparison with three other methods known from the literature. We made a collection of different model and query images called ART BANK. Currently, it contains 365 full color images of various sizes in the model part and 180 sketches in its query part. Images in the model part are mostly art works obtained from the World Art Kiosk at California State University. Images in the query part are black and white images which are draft sketches, drawn by three different subjects similar to 45 arbitrary candidates from the model part and their rotated versions ($90°$, $180°$ and $270°$). Figure 4 shows an example of sketched images.

The ART BANK was used as the test bed for the following four approaches. The proposed method (CPAI), the QVE approach, as it used in the QBIC system [3], MPEG-7 edge histogram descriptor (EHD) [12, 13], and the edge direction histogram (EDH) approach [16]. All the methods were tested using the same database.

In the CPAI method (Section 2), we applied $\beta = 3$, $J = 129$ and $N = 90$, resulting in a 90-entry feature vector f. For the EHD method, *desired_num_of_blocks* was set to 1100 and Th_{edge} set to 11 (the default values) for the model images and Th_{edge} was set 0 for the queries since they are binary images. A 150-bin histogram was obtained in this approach employing local, semi-global and global histograms. We used $k = 1$ in the EDH method, resulting in a 70-entry feature vector. The quantization stage in the EHD method was ignored to put all methods in the same situation.

The ℓ_1 distance was used for measuring the similarity between image features of the edge direction histogram (EDH) approach and of the proposed CPAI method. For MPEG-7 edge histogram descriptor (EHD) method a weighting factor of 5 for global bins, as recommended in [13], was also applied. Global correlation factor was employed for measuring the similarity between images in the QVE method.

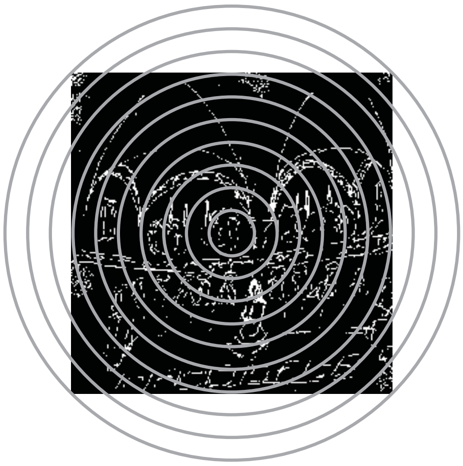


Fig. 3. An example of abstract image Ω and the circular partitions used for the feature extraction

The queries were depicted from the following sets (each contains 45 images): the original sketches (Q_0), their $90°(Q_{90})$, $180°(Q_{180})$ and the $270°(Q_{270})$ rotated versions. This is to simulate different vertical and horizontal directions when posing a sketched query to the retrieval system.

Recall ratio R_n [11] was used for the evaluation of retrieval performance. It shows the ratio to retrieve the original full color model image in the best n-candidates. That is

$$R_n = \frac{\text{queries retrieved the target image in the top } n \text{ retrievals}}{\text{total number of queries}} * 100$$

The R_n was obtained for each approach using the four different query sets. Figure 5 exhibits the supremacy of the proposed CPAI method over the other approaches. For the queries with the same direction as the model images (Q_0 set), the retrieval performance of QVE, EHD and EHD methods decline respectively compared to the CPAI method for all n's. The R_n of the CPAI method reaches to 100% at $n = 10$ in sets Q_0 and Q_{90} and at $n = 20$ in sets Q_{180}. It shows 97.5% at $n = 20$ in set Q_{270}. The EHD method has better Recall ratio R_n than the QVE method in $n < 20$ using set Q_{90}. These two methods possess close retrieval performances in Q_{180} and Q_{270} test data. EDH turned to have the worst Recall ratio in most cases.

4 Conclusion

The approach presented in this paper (CPAI) enables measuring the similarity between a full color model image and a simple black and white sketched query. The images are arbitrary and may contain several complex objects in inhomogeneous backgrounds. The approach deals directly with the whole image and needs no cost intensive image segmentation and object extraction. Abstract images were de£ned based on the strong

Fig. 4. An example of sketched images

edges of the model image and the thinned outline of the query image. Circular partitioning of the abstract image is used to extract features that are scale and rotation invariant and also tolerate small translations. Experimental results, using CPAI approach and the ART BANK as the test bed, show signi£cant improvement in the Recall ratio over three other approaches known from the literature.

The partitioning scheme could be re£ned to improve retrieval performance. We also intend to investigate sub-image search by dividing the image into several region and applying the approach to each region successively.

Acknowledgments. The authors would like to acknowledge the World Art Kiosk at California State University for providing paintings used in the ART BANK. We also thank Dr. Nargess Yasini, Andrew Mani and Khadijeh Moosavian who kindly help us producing sketched queries. The £rst author is £nancially supported by the Ministry of Science, Research and Technology of I.R. Iran.

References

1. Smeulders, A., Worring, M., Santini, S., Gupta, A., Jain, R.: Content-based image retrieval at the end of the early years. IEEE Trans. Patt. Anal. and Mach. Intell. **22** (2000) 1349–1380
2. Rui, Y., Huang, T., Chang, S.: Image retrieval: current techniques, promising directions, and open issues. Journal of Visual Commun. and Image Representation **10** (1999) 39–62
3. Niblack, W., Barber, R., Equitz, W., Flickner, M., Glasman, E., Petkovic, D., Yanker, P., Faloutsos, C., Taubin, G.: The QBIC project: querying images by content using color, texture, and shape. In: Proc. of Spie. Volume 1908., USA (1993) 173–187
4. Pentland, A., Picard, R.W., Sclaroff, S.: Photobook: content-based manipulation of image databases. Int. Journal of Comput. Vision **18** (1996) 233–254
5. Minka, T.P., Picard, R.W.: Interactive learning with a ”society of models”. Patt. Recog. **30** (1997) 565–581

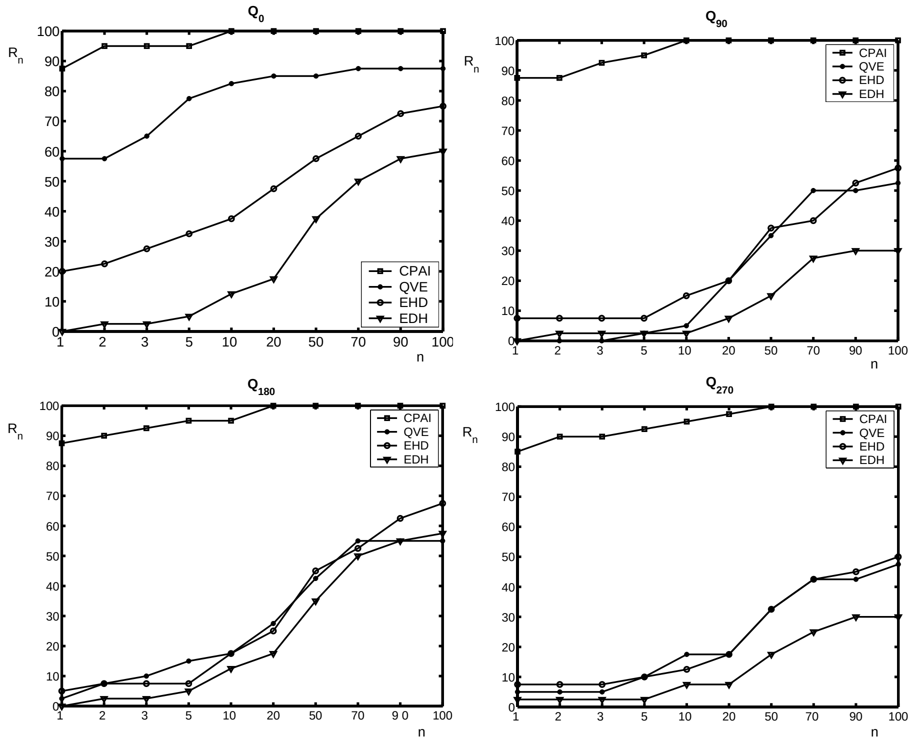


Fig. 5. R_n verses n for sets Q_0, Q_{90}, Q_{180}, and Q_{270} respectively

6. Beigi, M., Benitez, A.B., Chang, S.: MetaSEEk: a content-based meta-search engine for images. In: Proc. of Spie. Volume 3312., USA (1997) 118–128
7. Smith, J.R., Chang, S.: VisualSEEk: a fully automated content-based image query system. In: Proc. ACM Multimedia 96., NY, USA (1996) 87–98
8. Manjunath, B.S., Ohm, J.R., Vasudevan, V.V.: Color and texture descriptors. IEEE Trans. Circ. and Syst. for Video Tech. **11** (2001) 703–715
9. Bober, M.: Mpeg-7 visual shape descriptors. IEEE Trans. Circ. and Syst. for Video Tech. **11** (2001) 716–719
10. Ip, H.H.S., Cheng, A.K.Y., Wong, W.Y.F., Feng, J.: Af£ne-invariant sketch-based retrieval of images. In: proc. IEEE Int. conf. Comput. Graphics. (2001) 55–61
11. Hirata, K., Kato, T.: Query by visual example-content based image retrieval. In: Advances in Database Technology - EDBT '92, Berlin, Germany (1992) 56–71
12. ISO/IEC JTC1/SC29/WG11/N4063: MPEG-7 Visual part of eXperimentation Model Version 10.0, Singapore (2001)
13. Won, C.S., Park, D.K., Park, S.: Ef£cient use of Mpeg-7 edge histogram descriptor. Etri Journal **24** (2002) 23–30
14. Abdel-Mottaleb, M.: Image retrieval based on edge representation. In: Proc. Int. Conf. Image Processing. Volume 3., Piscatway, NJ, USA (2000) 734–737
15. Jain, A.K., Vailaya, A.: Image retrieval using color and shape. Patt. Recog. **29** (1996) 1233–1244

16. Jain, A.J., Vailaya, A.: Shape-based retrieval: a case study with trademark image databases. Patt. Recog. **31** (1998) 1369–1390
17. Chalechale, A., Mertins, A.: An abstract image representation based on edge pixel neighborhood inform. (EPNI). In: Lecture Notes in Comput. Science, EurAsian-ICT 2002:Inform. and Commun. technology. Volume 2510. (2002) 67–74
18. Chalechale, A., Mertins, A.: Semantic evaluation and ef£ciency comparison of the edge pixel neighboring histogram in image retrieval. In: WITSP'02 £rst workshop on the internet, telecommunications and signal processing, Australia (2002) 179–184
19. Atzori, L., Natale, F.D.: Error concealment in video transmission over packet networks by a sketch-based approach. Signal Processing: Image Commun. **15** (1999) 57–76
20. Matusiak, S., Daoudi, M., Blu, T., Avaro, O.: Sketch-based images database retrieval. In: Advances in Multimedia Inform. Syst. 4th Int. Workshop, MIS'98. Proc. (1998)
21. Canny, J.: A computational approach to edge detection. IEEE Trans. Patt. Anal. and Mach. Intell. **PAMI-8** (1986) 679–698
22. Gonzalez, R.C., Woods, R.E.: Digital Image Processing. Addison-Wesley (1992)