[Array 7 (2020) 100027](https://doi.org/10.1016/j.array.2020.100027)

|  |  |  |
| --- | --- | --- |
|  |  |  |
|  | Contents lists available at ScienceDirect |  |
| Array |
| journal homepage: www.elsevier.com/journals/array/2590-0056/open-access-journal |
|  | | |

Chordiogram image descriptor based on visual attention model for image retrieval

S. Sathiamoorthya, A. Saravananb,\*, R. Ponnusamyb

a Tamil Virtual Academy, Chennai, India   
b Division of Computer and Information Science, Annamalai University, Annamalai Nagar, India

|  |  |
| --- | --- |
| A R T I C L E I N F O | A B S T R A C T |
| Keywords:  Chordiogram image descriptor Edge map  Saliency map  Salient edges | A novel shape-based image retrieval is presented in this study. The foreground and background contents of images are strongly concealed, so they are represented individually to reduce their influence on each other in the proposed approach. The Otsu method is employed for segmenting the foreground from the background, and the saliency map and edge map are then clearly identified. Saliency reduces the time cost for feature computation, so salient edges are computed for the foreground and background images based on the selective visual attention model. |

Autocorrelation-based chordiogram image descriptors are computed separately for the foreground and background images, which are then combined in a hierarchical manner to form the proposed new descriptor. This approach avoids the concealment of foreground and background information, and the new descriptor is rich in geometric and its underlying texture, structure and spatial information. The proposed novel shape-based descriptor performs considerably better than conventional descriptors at content-based image retrieval. The proposed shape descriptor were extensively tested at image retrieval based on the Gardens Point Walking, St Lucia, University of Alberta Campus, Corel 10 k, and self-photographed image data sets. The precision and recall values were compared for the proposed and state-of-the-art-approaches when applied for shape-based image retrieval from these databases. The proposed shape descriptor provided satisfactory retrieval results in the experiments.

|  |  |
| --- | --- |
| 1. Introduction | Feature extraction is characterized by the utilization of the color, shape, |

and texture of images, where it must be able to differentiate among im-

At present, the number of digital images is increasing rapidly due to the use of various photography devices, such as webcams, mobile phones, and closed-circuit television (CCTV) cameras, and thus the sizes of image databases are growing greatly. Hence, storage, retrieval, and mainte-nance are important tasks for image databases. Image retrieval is broadly divided into two groups comprising (1) text-based image retrieval (TBIR) [1] and (2) content-based image retrieval (CBIR) [2–4]. TBIR was first introduced in the early 1970s and it uses manually annotated words to describe images, which is a difficult, time-consuming, and tedious task when the size of the image database is large, and this method is also subject to problems related to visual perception [5,6]. To resolve the

ages from the same and other classes with very small differences [8–11]. Furthermore, the features must be robust to geometric changes such as rotation, scaling, and translation, and photometric changes including differences in illumination and occlusion. Images are characterized using (1) global and (2) local approaches. In global approaches, images are characterized by ignoring the local and spatial information in the picture elements. The global approaches are computationally efficient and robust to noise to some extent, but they are not satisfactory at handling issues such as variations in illumination and occlusion. However, all of the problems with global approaches can be addressed by using local ap-proaches where features are computed based on local patches, regions, or

issues related to TBIR, CBIR was introduced in 1992 by Kato [7] and selected key points.

research in this domain of computer vision has continued for more than three decades. The need for more effective CBIR systems with high ac-curacy and low time costs has stimulated the development of improved CBIR systems. CBIR allows the user to retrieve images more efficiently from image databases by employing feature extraction and matching.

\* Corresponding author.

Shape is an important component used in image recognition and matching. In the present study, we propose a method based on shape information, and thus we focus on previous methods based only on shape information in the following. Several shape characterization and matching approaches have been proposed in previous studies. In general,

E-mail addresses: [ks\_sathia@yahoo.com](mailto:ks_sathia@yahoo.com) (S. Sathiamoorthy), [sjpramoth@gmail.com](mailto:sjpramoth@gmail.com) (A. Saravanan).

<https://doi.org/10.1016/j.array.2020.100027>  
Received 14 July 2019; Received in revised form 20 April 2020; Accepted 22 April 2020   
Available online 1 May 2020   
2590-0056/© 2020 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

S. Sathiamoorthy et al. Array 7 (2020) 100027

shape features are computed with: (1) boundary-or contour-based ap-proaches, and (2) region-based approaches, where the former charac-terize the details of the shapes in images using the contours of an object and the latter use all of the pixels in a region [12–15]. We focus our discussion on boundary-based shape characterization methods.

Boundary-based shape characterization using chain code histogram [15,16] provides compact representations, translation invariant, preser-ving all of the morphological information. Shape signatures [17,18] such as cumulative angle, centroidal profile and chord lengh, and global shape descriptors such as the area, circularity, eccentricity, bending energy, convexity, major axis orientation, ratio of the principal axis, circular variance, and elliptic variance, and shape invariants were also employed in previous studies [19–21]. The contour point distribution histogram [22] comprises the distribution of points on an object contour under polar coordinates and it is a suitable approach for describing shapes with closed contours but not for images with multiple connected regions. Boundary moments were considered in some previous studies [23,24]. The curvature scale space approach employs a corner point detector to search for the curvature maxima or inflection points on the edges detected using the Canny approach [25]. Elastic matching involves an optimization problem based on the pixel to pixel correspondence be-tween two images and it is robust to geometric deformation [26,27]. Eigen values are also invariant under rigid motion and scaling [28]. A shape-based non-redundant local binary pattern was presented by Yao and Chen [29]. The local maximum edge binary pattern suggested by Murala et al. [8] is computed based on the local differences among the center and eight neighborhood pixels, and it is combined with the Gabor transform to ensure its effectiveness. The local edge pattern proposed by Yao and Chen [30] uses the Sobel edge to compute the local edge pattern for segmentation and the local edge pattern for retrieval. Murala and Wu proposed peak valley edge patterns [31], where they obtained the local mesh peak valley edge pattern by including the first order derivative in the local mesh patterns [32]. The edge histogram descriptor reported by Jain and Vailya involves the distribution of the orientations of edgels, and the edges are computed using the Canny approach [33]. Another histogram method based on the edge distribution at the local level uses the Sobel operator for edge detection [34], although it obtains sub-standard retrieval results. Thus, a method was developed that uses the absolute locations of edge and the global composition to enhance the retrieval rate [35]. Jiebo et al. [36] reported a color edge co-occurrence histogram that calculates the distribution of the separation between pairs of color edges. A method based on the color distributions on directional and non-directional edges was suggested by Shim and Choi [37]. The block variation of local correlation coefficients method presented by Chen et al. identifies the edges and valleys in an image, which are characterized by first order moments [38,39], and a method based on the distributions of the edges and valleys was then proposed [40]. The edge orientation autocorrelogram method [41] computes the correlations among the edgels based on their orientation, and this method is robust to differences in illumination, viewpoint, translation, and small amounts of rotation. An enhanced version based on the edge histogram descriptor

intensity and distance among each pair of edgels [45]. Moreover, a bi-nary coherent edge descriptor was presented that characterizes the co-ordinates and orientation of each edgel, and the length of an edge that passes through the edgels [46].

Wang et al. proposed a new variant of CID that collects the distribu-tions of the chord details for patches in images. The image is divided into a number of non-overlapping rectangular patches to reduce the influence of lighting. The CID method is robust to edge detector and it reduces the computational cost by employing predominant edgels. Statistical tests are conducted to identify the predominant edgels in each patch [47] and for every pair of predominant edgels in each patch, it is necessary to compute the distance among every pair of predominant edgels, the ori-entations of each predominant edgel in a pair, and the degree of the angle between a line segment among pair predominant edgels and the hori-zontal axis, before combining these geometric details in a local edgel chordiogram (LEC) [47]. The ordered collection of LECs for all patch images comprises the CID. The CID is robust to differences in illumina-tion, translation, and in-plane rotation, but it is affected considerably by noise. Thus, the patches with noise are eliminated during the matching operation by avoiding higher values in the similarity results obtained between the corresponding patches in the query and target images. CID is an appropriate method for place recognition with illumination changes, while the time cost is low and it can avoid fake edgels because edge detectors are used for edge identification instead of segmentation.

In a previous study, we enhanced the efficiency of this method by computing the CID using an autocorrelation function to obtain the autocorrelation-based CID (ACID) [48]. The ACID exploits the spatial correlation among identical predominant edgels at distance d, the orientation details for each predominant edgel in a pair of identical predominant edgels at distance d, and the degree of the angle along the line segment between a pair of identical predominant edgels and the horizontal axis. Our method neglects the length between a pair of pre-dominant edgels because the length is always 1 in our approach. We demonstrated that ACID performs better than the conventional CID.

All of the previous approaches mentioned above compute the shape details for either a whole image or objects segmented from an image. However, previous studies have shown that the background and fore-ground details in images are concealed by both the global and local features [49], thereby resulting in poor retrieval performance because the user may be focused on objects in the background or foreground, or both. However, pinpointing the interests of users such as the background or foreground or even a specific object in the foreground is a challenging problem for the current CBIR approaches. Furthermore, separating the objects in the foreground and comparing them with the corresponding object in a target image is still a difficult issue for the existing CBIR ap-proaches. At present, these problems are resolved using a relevance feedback approach where users are permitted to choose the images from the retrieval results obtained by a query and the selected images are then jointly employed to refine the query image until it corresponds subjec-tively to a user’s needs in a particular search, and thus this process continues until the user is satisfied with the results [50]. Recently, ma-

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| and | edge | orientation | autocorrelogram | methods | [42,43] | employs | chine learning has been combined with the relevance feedback approach |

extremely minute and fine edges using a framework based on the full range autoregressive model for grayscale and color images, where the edges of color images are computed in the HSV space in order to avoid

to enhance the retrieval rate, but this approach also fails because of the low number of training images and the unwillingness of users to partic-ipate in the relevance feedback approach for a lengthy period of time

missing minute and fine edges due to changes in spectral and chromatic [50].

details. Recently, Feng et al. [50] presented a CBIR where the salient edges

Recently, Toshev et al. [44] proposed a chordiogram image descriptor (CID) that computes the geometric details for a selected set of edgels obtained by segmentation. This method then employs geometric details comprising the distances among pairs of edgels, orientations of pairs of edgels, and the degree of the angle connecting pairs of edgels and the horizontal axis. However, the boundaries computed by segmentation might include fake edgels and this can affect the accuracy of results. Further computing the geometric details for every pair of edgels greatly increases the time cost. A subsequently developed method computes the

2

S. Sathiamoorthy et al. Array 7 (2020) 100027



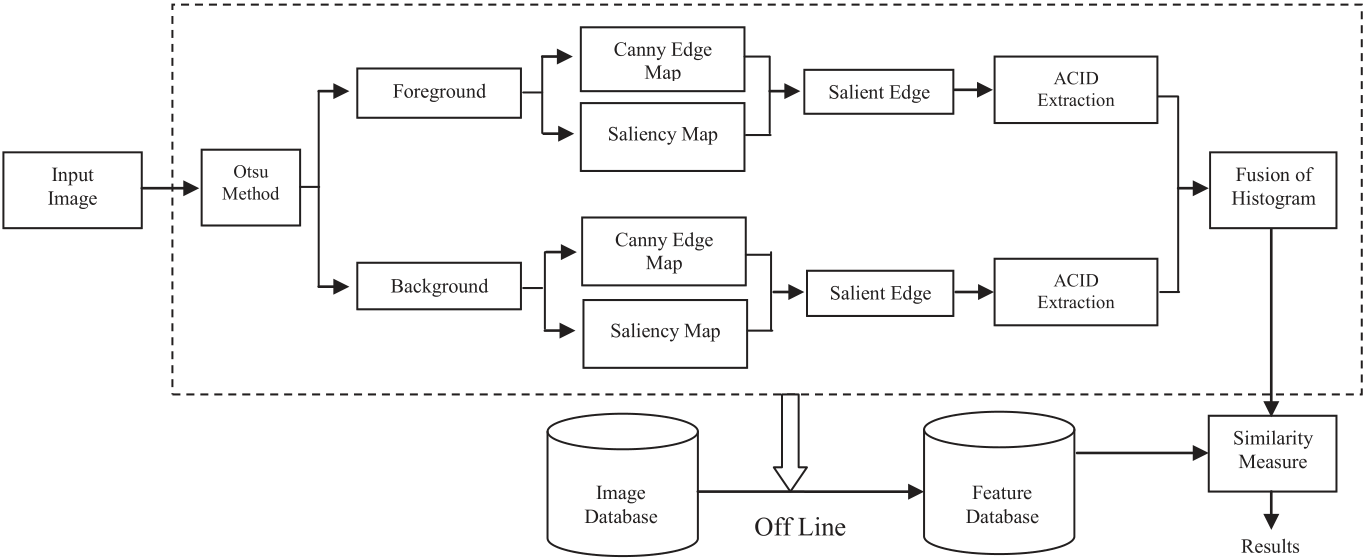


Fig. 1. Architecture of the proposed image retrieval approach.

computed in both the foreground and background in order to avoid them influencing each other, thereby obtaining hierarchical feature de-scriptions and achieving more accurate image matching.

In the method proposed in the present study, we computed the ACID based on the saliency edges for the foreground and background images, and obtain hierarchical feature descriptions by using the ACID to reduce

as a linear combination of the intensity contrasts in the Gaussian image pyramid. They reported that their approach can obtain more precise details from an image. Thus, let us assume that the image I is in the RGB color space and a mask with a size of 3 � 3 is centered on a given pixel p in an image with M � N dimensions in order to compute the saliency value at pixel p based on its adjacent neighbors as follows [50]:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| the effects on each other of the foreground and background features. This | SVðpÞ ¼ | X | p2LxL X ωcSl cðp; qÞ þ ωoSl oðp; qÞ | (1) |
| method captures more geometric and its underlying texture, structure, |
| and spatial details by considering a higher number of more responsive |
| salient edgels (25% of the total number of edgels) than the conventional |

method (15% of the total number of edgels). The proposed method se-lects more salient edgels than the conventional approach in order to capture the rich underlying texture and structure information. By contrast, considering less salient edgels will reduce the computational cost, but this method fails to capture much of the underlying texture and structure information among the salient edgels. We experimentally evaluated the proposed approach by considering various subsets of salient edgels where each subset varied in terms of the numbers of salient edgels, and the response strengths of the salient edgels were considered when selecting them for a subset. The experimental results demonstrated that considering 25% of the salient edgels for extracting the rich un-derlying texture, structure and spatial information could obtain more accurate results, whereas reducing the number of salient edgels signifi-cantly reduced the cost but it yielded less accurate retrieval results. Our proposed method employs the Otsu algorithm [49] to segregate the image into foreground and background details, while the Canny operator is used for edge detection and the selective visual attention model [50] to exploit the salient edges. We comprehensively tested the proposed approach based on benchmark databases and the results were compared with those obtained using CID [47], ACID [48], and SEH [50]. The proposed approach obtained more accurate results than CBIR. The pro-posed retrieval approach is an enhanced version of our previously re-ported method [48].

The remainder of this paper is organized as follows. In Section 2, we describe the approaches incorporated in the proposed CBIR. The exper-imental results and discussion are presented in Section 3. Finally, we give our conclusions in Section 4.

2. Feature extraction techniques

In the following, we provide overviews of the selective visual atten-tion model, CID and ACID techniques, proposed CBIR, and feature descriptor matching method. The architecture of the proposed retrieval approach is illustrated in Fig. 1.

2.1. Selective visual attention model

Recently, Feng et al. [50] enhanced the saliency model and defined it

3

S. Sathiamoorthy et al. Array 7 (2020) 100027

Table 1   
The representation of autocorrelation of identical

|  |  |
| --- | --- |
| predominant edgels at d ¼ 1. Edgel Valu | Distance (d) |

D ¼ 1

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 255 | | | | 0.019 | | | | |
| 254 | | | | 0.075 | | | | |
| 253 | | | | 0.102 | | | | |
| . | | | | . | | | | |
| . | | | | . | | | | |
| . | | | | . | | | | |
| . | | | | . | | | | |
|  | | | | |
| SA 0 @eI | 1 A ¼ | X | p2Θpin X | SV | �� �  p | L�eI i | � ; | (4) |
| Where Θpi nrepresents a 3 � 3 mask centered at pixelpi nand SV(p) is the  saliency value of pixel p. After computing all of the saliency values for the | | | | | | | | |

edges, the following threshold operation is performed [50].

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| TE ¼ max The final salient edge is then described as [50]:�SA�eI i���4 | | | | | | (5)  (6) |
| ΘSE ¼ | �eI i | ��SA�  eI i | � | > TE; i ¼ 1; :::; N | � |
| 2.3. CID | |

A given image is divided into several non-overlapping rectangular patches numbered from 1, 2, 3,…, N. Each patch is characterized using the LEC. To compute the LEC, the edges are identified using any edge operator and the prominent edgels are identified for each patch by applying a statistical hypothesis test, before the local geometric features are computed based on the prominent edgels. The LECs obtained for all patches are then collected in an order to construct a CID. The chord details for an image patch are computed based on each pair of prominent edgels coordinated at p and q as follows [47]:

Cipq ¼�ℓpq; ϕpq; θp; θq� ; (7)

where ℓpq;ϕpq;θp, and θqdenote the distance among predominant edgels p and q, the angle between the line connecting p and q and horizontal plane, and the degrees of orientation for predominant edgels p and q about the normal directions, respectively. The values ofℓpqrange from 0 to the diameter of an image patch, and ϕpq;θp, andθqrange from 0 to 2Π. The values of ℓpqare discretized in the logarithmic space and they are divided into ηd bins where d is the distance, which is fixed to 4 [47]. ϕpq; θp, andθqare each quantized into eight bins [47]. Thus, CID is a four-dimensional histogram that is normalized as described previously [47], and it encompasses the chord details for every pair of prominent edges in an image patch.

2.4. ACID

We propose a novel characterization approach for an image using ACID, which represents an improved version of the approach described by Wang et al. [47]. ACID computes the spatial correlations among the identical predominant edgels and explores the variations in the correla-tions at a distance d for each pair of the predominant edgels in a sub-image with a size of 3 � 3 (one is at the center of the sub-image and other is an identical edgel at distance d from the center of the mask). We also compute the orientation of each predominant edgel and the degree of the angle along the line segment between the pair of predominant edgels and the horizontal plane. The autocorrelations among the iden-tified identical predominant edgels are depicted in a table and indexed

4

S. Sathiamoorthy et al. Array 7 (2020) 100027

the proposed spatial correlations among identical predominant edgels at distance 1, where the first and second columns show the edgels with

the proposed approach avoids the concealment of details in the fore-ground and background, and the computational cost is also significantly

value i and the probability of finding an edgel with value i at correlation reduced.

distance 1 from an edgel with value i, respectively. The autocorrelations

among identical predominant edgels are described as follows [48].

Let I be an n � n image and the predominant edgels in I are e1;e2;:::;em. For an predominant edgel E ¼ ðx;yÞ 2 I, let IðEÞrepresent its edge value. Let Ie¼ IðEÞ ¼ distance among predominant edgels [48], i.e., for predominant edgels ΔfEjIðEÞ ¼ eg Hence, the notation E 2 Ieis synonymous withE 2 I; e. For convenience, we use theL∞norm to assess the

E1 ¼ ðx1; y1Þ; E2 ¼ ðx2; y2Þ, we define jE1 � E2j¼ Δmaxfjx1 � x2j; jy1 �y2jgWe denote the set f1; 2;:::;ngbyjnj.

Definitions. The histogram h of I is expressed for i 2 jmjas follows.

heiðIÞ¼ Δ n2PrjE 2 Ieij E2I (8)

For any predominant edgel in the image, heiðIÞ=n2is the probability that the value of predominant edgel is ei.

Let a distance d 2 jnjbe fixed a priori. Then, the correlation of I is expressed for i;j 2 jmj;k 2 jdjas follows.

γðkÞ ei;ej¼ Δ

E12Iei ;E22I Pr �E2 2 IejjjE1 � E2j ¼ k� (9)

Given any predominant edgel of value eiin the image, γðkÞ ei;ejðIÞis the probability that a predominant edgel at distance k from the given pre-dominant edgel is of value ej. The autocorrelation of I only exploits the spatial correlation among identical predominant edgels, and it is defined as [48]:

αðkÞ c¼ Δ γðkÞ e;eðIÞ (10)

image from left to right and then from the topmost left corner of the To estimate the ACID, a 3 � 3 non-overlapping mask is moved over an

image for each identified predominant edgel value. For instance, as shown in Fig. 2, when we move the 3 � 3 mask over an image to compute the autocorrelation value of the predominant edgel with a value of 255, we obtain an identical predominant edgel with a value of 255 at the center of the first 3 � 3 sub-image and it has two identical predominant edgels at distance 1. Thus, we estimate the autocorrelation using Eq. (10) and then estimate the orientation of each predominant edgel in every pair of predominant edgels (center and its identical predominant edgel) as θ1and θ2 then compute the degree of the angle of the line segment between each pair of predominant edgels and the horizontal axis as ϕ, as depicted in Fig. 2(b). Next, the 3 � 3 mask is moved to the left and there is no predominant edgel with a value of 255 at the center of the mask. Thus, the mask is moved further. In the third 3 � 3 sub-image, there is also no center edgel with a value of 255 but the fourth contains a pre-dominant edgel in the center of the sub-image with a value of 255 and it has one identical predominant edgel at distance 1. Therefore, we esti-mate the autocorrelations among the identical predominant edgels for the fourth sub-image using Eq. (10), and we compute the values of θ1,θ2, andϕ[48]. This process continues for each predominant edgel value identified in the image. Finally, we obtain the autocorrelations of pre-dominant edgels, as shown in Table 1, and the set of θ1,θ2, and ϕvalues are represented by separate histograms.

In our previous study [48], we demonstrated that ACID obtains a better retrieval rate because it exploits geometric and its underlying texture, structure, and spatial details and the computational cost is equivalent to that of the CID approach [47]. We also normalized the histogram for ACID to obtain more lighting variations [48].

Therefore, in the method proposed in the present study, we employ four-dimensional ACID for CBIR. In the proposed CBIR method, an image is segregated into foreground and background images, and ACID is then computed for the whole foreground and background images instead of computing it based on the patches in the whole image [47,48]. Therefore,

5

S. Sathiamoorthy et al. Array 7 (2020) 100027

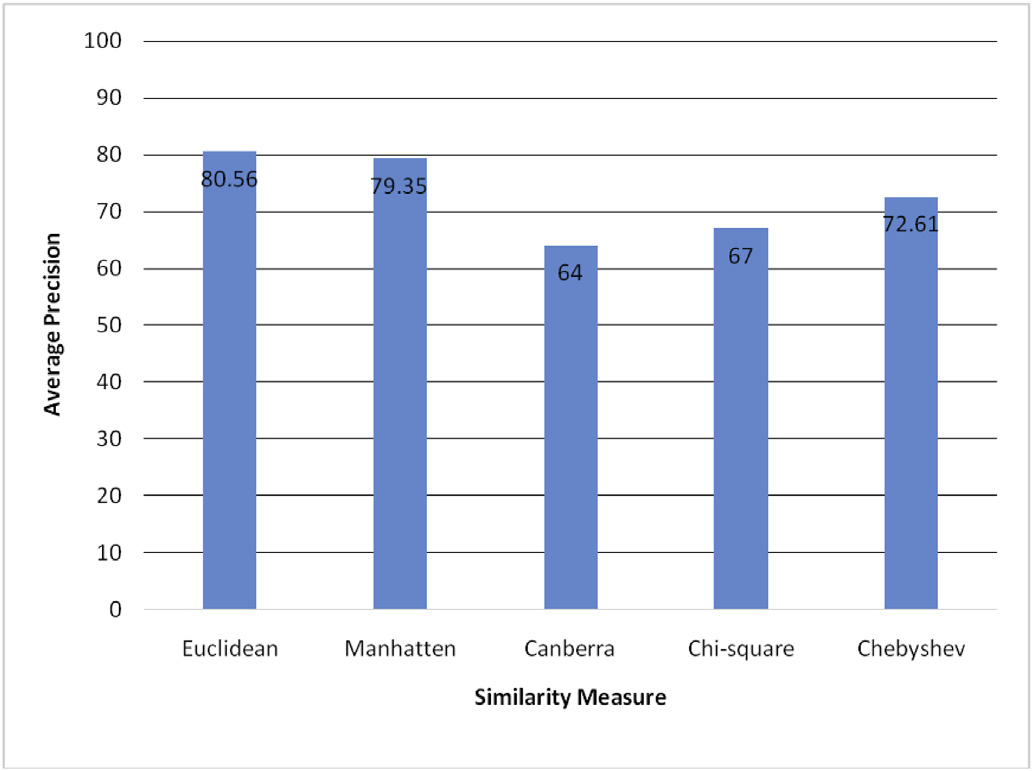


Fig. 3. Average retrieval accuracy attained by various distance measures for the

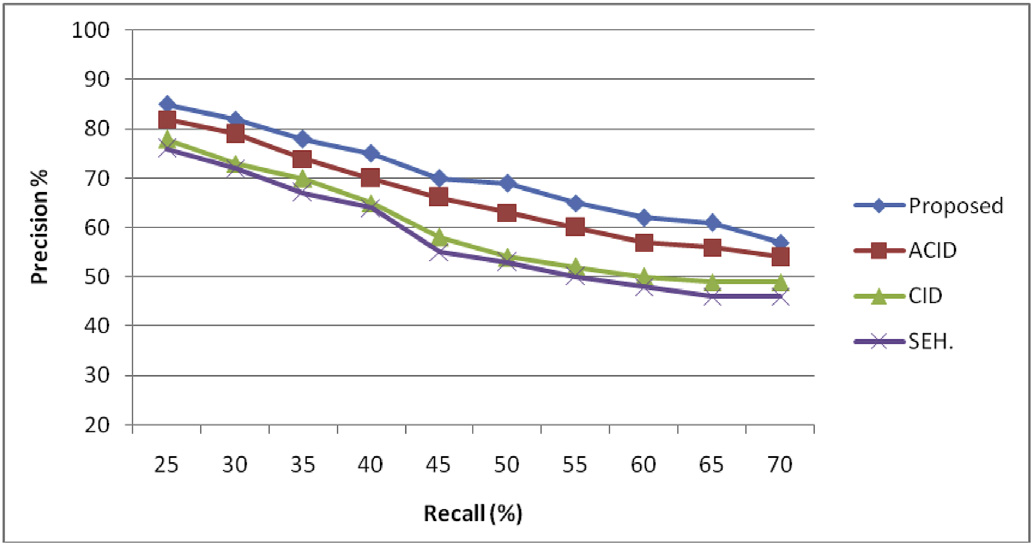


Fig. 4. Precision Vs recall of the proposed and existing approaches for Gardens Point Walking database.

image for a given query and it increases as the number of retrieved im-ages increases. The recall is defined as follows [48].

|  |  |  |
| --- | --- | --- |
| proposed feature descriptor. | RecallRTotal No: of correct images retrieved from database |  |
| images are related to each other when the similarity between the two | ðÞ ¼ | (13) |
| images is a small value. Various similarity metrics can be employed to |

compute the similarity of two images, but the Euclidean method is more familiar and it is used extensively for CBIR. In the proposed method, after computing the features locally in the form of four-dimensional histo-grams, the similarity between the query and target images is computed

|  |  |  |
| --- | --- | --- |
| with the following equation [50–52]: | | (11) |
| SðQ; TÞ ¼ | v u u t ffiffiffiffiffiffiffiffiffiffiffiffiffiffiffiffiffiffiffiffiffiffiffiffiffiffiffiffiffiffiffi ðjQi � TijÞ2 ; |

where T and Q denote the input image and target image feature descriptor, respectively, and N is the number of descriptors in the image feature vector. We also compared the performance of other distance measures [50–52] such as the Manhattan, Canberra, Chi-square, and Chebyshev metrics in terms of the average precision, and the results demonstrated that worst performance was obtained with the Canberra metric and the highest performance using the Euclidean distance metric. The Manhattan metric performed significantly worse than the Euclidean metric. Thus, the Euclidean distance is used in our proposed CBIR approach. Fig. 3 shows the average recall results with various distance measures.

3. Experimental results and discussion

In the experiments, we compared selected methods comprising CID [47], ACID [48], and SEH [50] with the proposed approach. We evalu-ated the performance of these approaches based on the Gardens Point Walking [47], University of Alberta (UA) Campus [47], St Lucia [47], Corel 10 k [54], and self-photographed image databases. The superior performance of the proposed approach was validated by evaluating the precision and recall, and based on comparisons with the other ap-proaches with all five image databases. The precision defines the rela-tionship among the total number of related images retrieved for a given input image and the total number of images retrieved from the database, which gradually decreases as we increase the number of retrieved im-ages. The precision is defined as follows [48].

PrecisionðPÞ ¼Total No: of correct images retrieved from database (12)

Another common measure used for computing the accuracy is the recall, which is defined as the probability of retrieving a correct related

6

S. Sathiamoorthy et al. Array 7 (2020) 100027

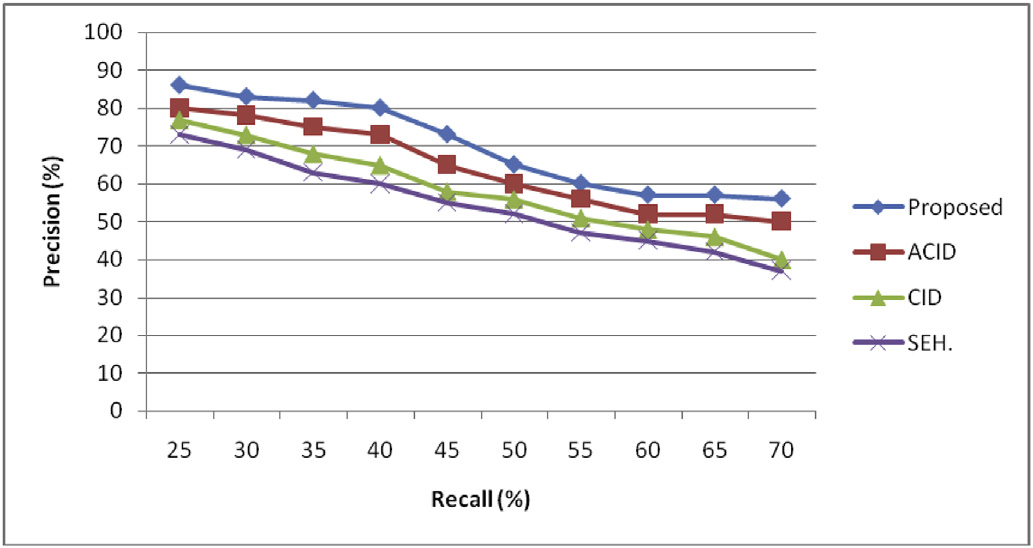


Fig. 5. Precision Vs recall of the proposed and existing approaches for UA database.

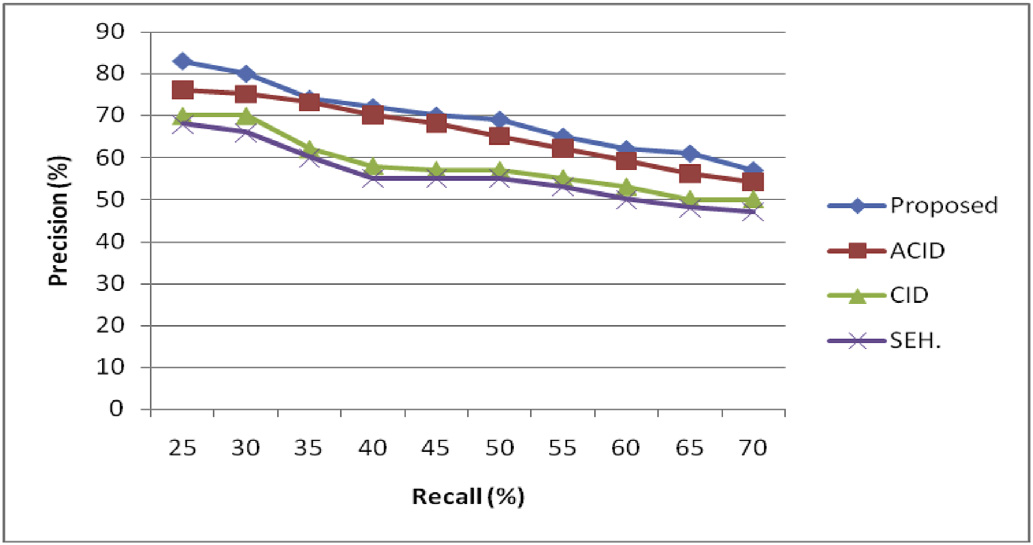


Fig. 6. Precision Vs recall of the proposed and existing approaches for St.Lu-cia database.

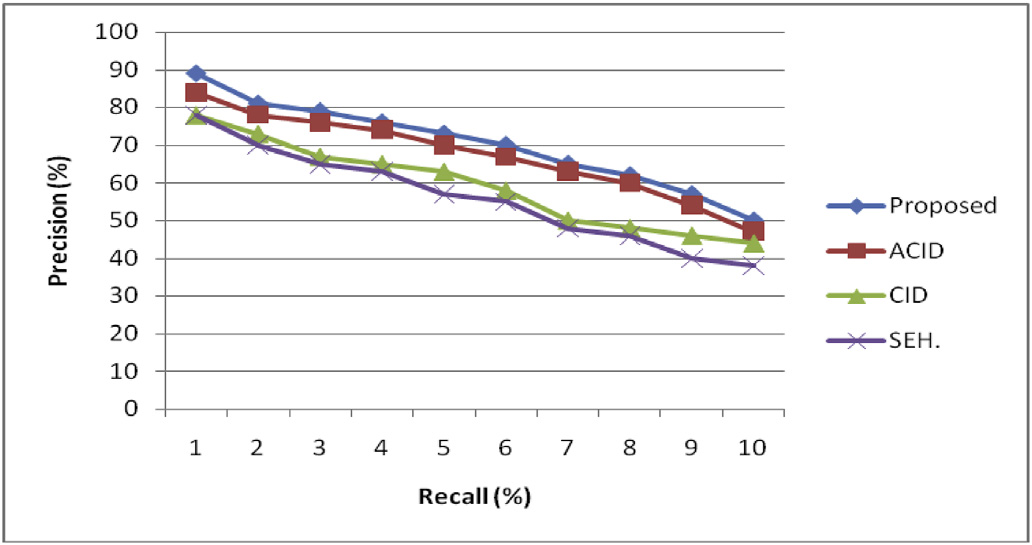


Fig. 7. Precision Vs recall of the proposed and existing approaches for Corel 10 k database.

day, i.e., at 06:20, 16:40, and 22:15, where each sequence contained 630 frames [47]. The results obtained by applying CID, ACID, SEH, and the proposed ACID based on the selective visual attention model to the foreground and background images in a hierarchical manner were rela-tively similar to the results obtained with the Gardens Point Walking database. The proposed approach performed better than ACID, while the performance of CID was intermediate and that of SEH was again the worst. The corresponding precision and recall plots are depicted in Fig. 5.

Next, we compared the performance of the different descriptors based on the St. Lucia database. The precision and recall curves obtained for the proposed and existing approaches are depicted in Fig. 6. The St Lucia database contains images acquired at five different times between the early morning and late afternoon during the day, and during the day after two weeks, where it comprises 10 groups of images. In the experiments, the worst performance was obtained using SEH and the best performance

7

S. Sathiamoorthy et al. Array 7 (2020) 100027

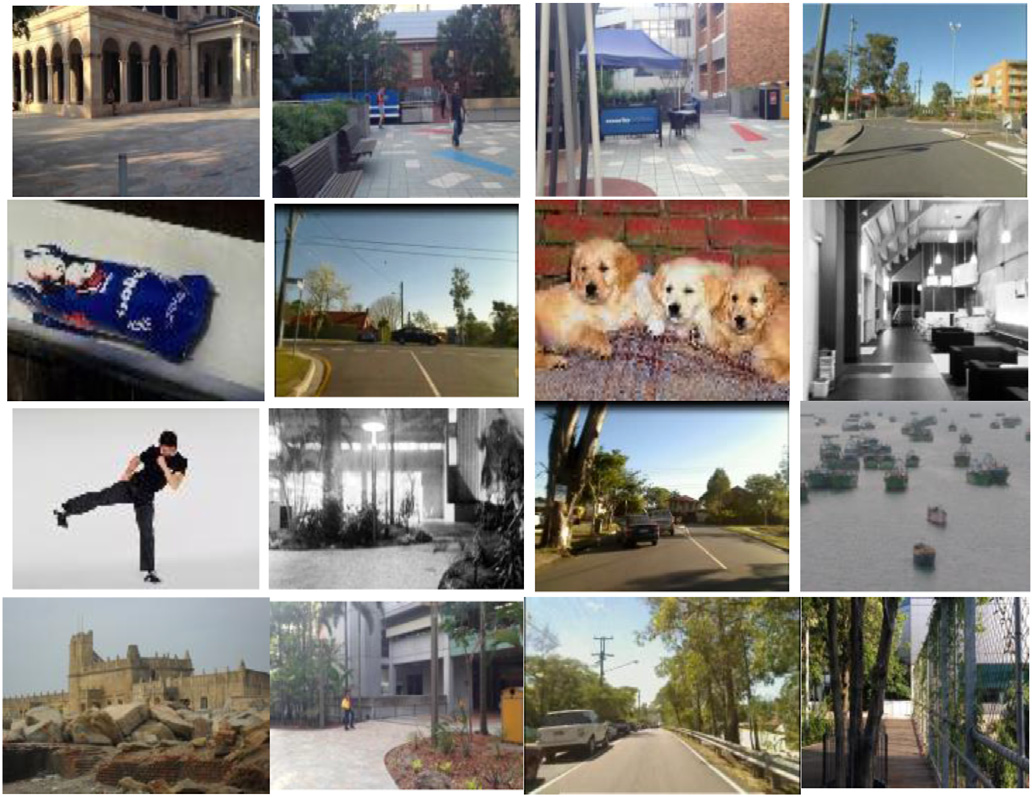


Fig. 9. Few sample images of experimental databases.

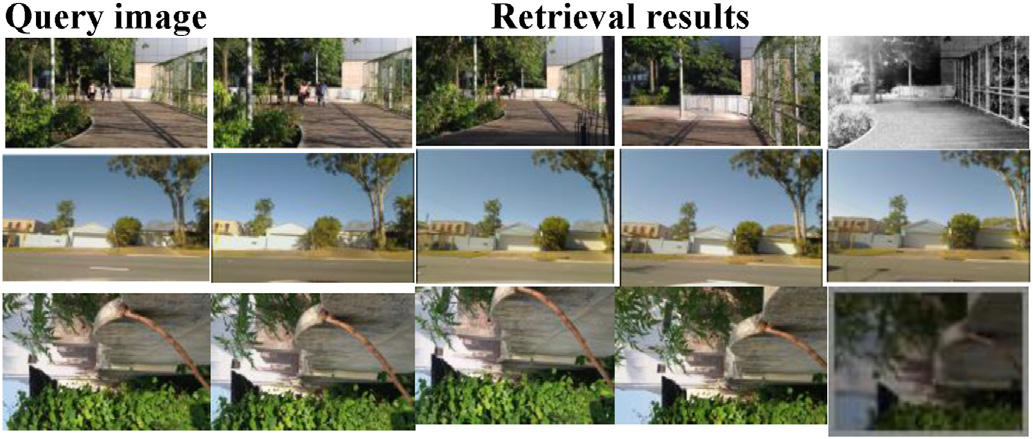


Table 3   
Computational complexity of proposed, ACID, CID and SEH..

|  |  |
| --- | --- |
| Methods | Time in seconds |
| SEH  CID  ACID  Proposed | 47  56  45  52 |

edgels were selected initially, and we then increased P by 5% incre-mentally. For each value of P, we computed the proposed feature and measured the performance in terms of the accuracy and time cost. The best performance was obtained when P ¼ 25%. In the experiments, we changed the value from P ¼ 5%–100% and the accuracy increased gradually to P ¼ 25%, but there were no further changes in the perfor-mance with higher values when we increased P, although the computa-tional cost increased, as shown in Fig. 11. Therefore, the results clearly showed that the proposed approach could extract rich geometric and its underlying texture, structure and spatial information when P ¼ 25%. Any descriptor must perform effectively and the computational cost should also below. Therefore, the central processing unit (CPU) times required by the proposed method, CID, SEH, and ACID were computed by extracting the features offline (to create a feature database) and online (retrieval). The average CPU times required by the proposed method and the CID, SEH, and ACID descriptors are shown in Table 3. SEH required the least CPU time whereas the proposed approach consumed the most CPU time, and it was slightly higher than that by ACID and slightly lower than that by CID. However, the fairly high time cost is acceptable because of the high accuracy of the proposed method. Thus, the proposed approach exhibits efficient retrieval performance and it is robust to dif-ferences in illumination and occlusion.

|  |  |
| --- | --- |
| Fig. 10. For instance, top 4 retrieval results for Gardens point walking, St.Lucia | 4. Conclusion |

and self photographed image databases.

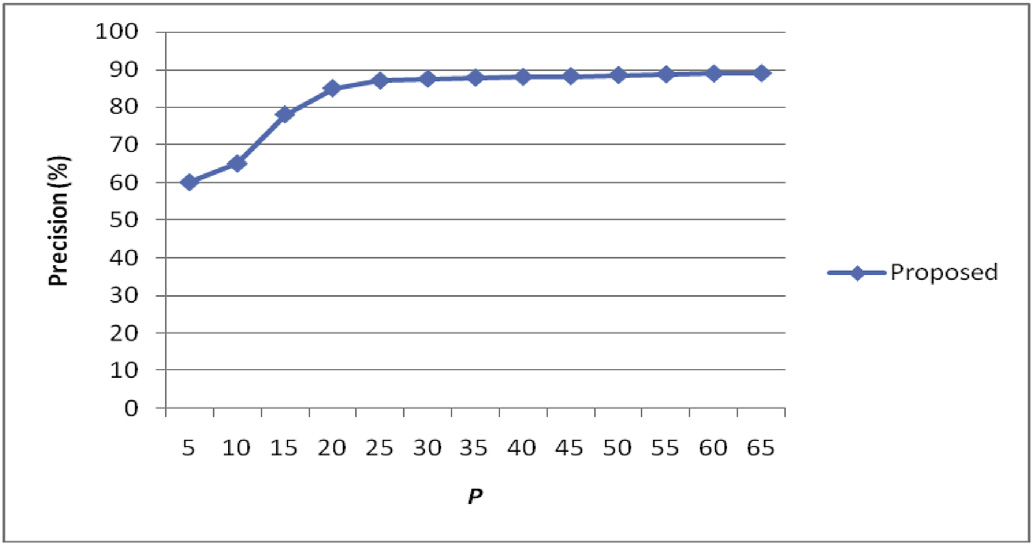


Fig. 11. Average precision Vs various P% of salient edgels.

Table 2 show the average precision and recall values using the proposed descriptor and the ACID, CID, and SEH approaches based on the Gardens Point Walking, St Lucia, UA Campus, Corel 10 k, and self-photographed images databases. Examples of images in all of the benchmark databases are shown in Fig. 9. In addition, examples of the retrieval results obtained by the proposed approach based on the Gardens Point Walking, St Lucia, and self-photographed images databases are presented in Fig. 10.

In the experiments, we also computed the proposed feature using various subsets of salient edgels, where each subset (S) varied in terms of the number of salient edgels. In particular, if N salient edgels are present in an image, then the subset S contains P � N salient edgels, where P is the proportion used to determine the size of the subset. In the experi-ments, we started with P ¼ 5%, i.e., 5% of the stronger responsive salient

8

S. Sathiamoorthy et al. Array 7 (2020) 100027

Declaration of competing interest

The authors declare that they have no known competing financial

interests or personal relationships that could have appeared to influence

the work reported in this paper.

References

[1] [Ma H, Zhu J, Lyu MRT, King I. Bridging the semantic gap between image contents](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref1)

[2] [Feng S, Xu D, Yang X. Attention-driven salient edge (s) and region (s) extraction](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref2) [and tags. IEEE Trans Multimed 2010;12(5):462](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref1)[–](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref2)[73](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref1).

[3] [Wang M, Ye Z-L, Wang Y, Wang S-X. Dominant sets clustering for image retrieval.](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref3) [with application to CBIR. Signal Process 2010;90(1):1](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref2)[–](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref3)[15](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref2).

[4] [Murala S, Maheshwari R, Balasubramanian R. Local tetra patterns: a new feature](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref4) [Signal Process 2008;88(11):2843](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref3)[–](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref4)[9](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref3).

[descriptor for content-based image retrieval. IEEE Trans Image Process 2012;21(5):](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref4)

[5] [Long F, Zhang H, Feng DD. Fundamentals of content-based image retrieval. In:](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref5) [2874](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref4)[–](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref5)[86](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref4).

[6] [Zhang X, Zhao X, Li Z, Xia J, Jain R, Chao W. Social image tagging using graph-](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref6)[Multimedia information retrieval and management. Springer; 2003. p. 1](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref5)[–](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref6)[26](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref5).

[based reinforcement on multi-type interrelated objects. Signal Process 2013;93(8):](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref6)

[7] [Androutsos P, Kushki A, Plataniotis KN, Venetsanopoulos AN. Aggregation of color](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref7) [2178](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref6)[–](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref7)[89](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref6).

[and shape features for hybrid query generation in content based visual information](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref7)

[8] [Subrahmanyam M, Maheshwari R, Balasubramanian R. Local maximum edge binary](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref8) [retrieval. Signal Process 2005;85(2):385](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref7)[–](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref8)[93](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref7).

[patterns: a new descriptor for image retrieval and object tracking. Signal Process](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref8)

[9] [Zhang J, Ye L. Local aggregation function learning based on support vector](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref9) [2012;92(6):1467](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref8)[–](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref9)[79](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref8).

[10] [Qian Y, Hui R, Gao X. 3D CBIR with sparse coding for image-guided neurosurgery.](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref10) [machines. Signal Process 2009;89(11):2291](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref9)[–](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref10)[5](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref9).

[11] [Hu MK. Visual pattern recognition by moment invariants. IRE Trans. Info. Theory](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref11) [Signal Process 2013;93(6):1673](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref10)[–](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref11)[83](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref10).

[12] [Lu G, Sajjanhar A. Region based shape representation and similarity measure](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref12) [1962;8:179](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref11)[–](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref12)[87](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref11).

[13] [Zhang D, Lu G. Shape-based image retrieval using generic Fourier descriptor. Signal](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref13) [suitable for content-based image retrieval. Multimed Syst 1999;7(2):165](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref12)–[74](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref12).

[14] [Kim W-Y, Kim Y-S. A region based shape descriptor using Zernike moments. Signal](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref14) [Process Image Commun 2002;17(10):825](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref13)[–](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref14)[48](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref13).

[Process Image Commun 2000;16:95](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref14)–[102](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref14).

[15] Freeman, H., Saghri, A., November 7–10, 1978. Generalized chain codes for planar curves, in: Proceedings of the fourth international joint conference on pattern recog.

[16] [Iivarinen J, Visa. A. Shape recognition of irregular objects. In: Casasent DP, editor.](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref16) Kyoto, Japan, pp. 701[–](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref16)703.

[Intelligent robots and computer vision XV: algorithms, techniques, active vision,](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref16)

[17] [Otterloo PJV. A contour-OrientedApproach to shape analysis. Englewood Cliffs, NJ:](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref17) [and materials handling. Proc; 1996. p. 25](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref16)[–](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref17)[32. SPIE 2904](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref16).

[Prentice-Hall International (UK) Ltd; 1991. p. 90](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref17)[–](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref18)[108](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref17).

[18] [Zhang DS, Lu G, January 22–25. A comparative study of Fourier descriptors for shape representation and retrieval. In: Proceedings of the f](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref18)i[fth asian conference on](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref18)

[19] [Yong J, Bowie Walker, J. An analysis technique for biological shape. Comput Graph](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref19) [computer vision. Melbourne, Australia: ACCV02); 2002. p. 646](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref18)[–](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref19)[51](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref18).

[20] [Peura M, Iivarinen J, May. Eff](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref20)i[ciency of simple shape descriptors. In: Proceedings of](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref20) [Image Process 1974;25:357](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref19)–[70](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref19).

[21] [Huang C-L, Huang D-H. A content-based image retrieval system. Image Vis Comput](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref21) [the third inter. Workshop on visual form. Italy: Capri; 1997. p. 443](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref20)[–](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref21)[51](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref20).

[22] [Shu X, Wu X-J. A novel contour descriptor for 2D shape matching and its](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref22) [1998;16:149](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref21)[–](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref22)[63](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref21).

[23] [Sonka M, Hlavac V, Boyle R. Image processing, analysis and machine vision.](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref23) [application to image retrieval. Image Vis Comput 2011;29(4):286](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref22)[–](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref23)[94](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref22).

[24] [Gonzalez RC, Woods RE. Digital image processing. Reading, MA: Addison-Wesley;](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref24) [London, UK, NJ: Chapman](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref23) & [Hall; 1993. p. 193](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref23)[–](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref24)[242](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref23).

[25] [Mokhtarian F, Abbasi F, Kittler J. Eff](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref25)i[cient and robust retrieval by shape content](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref25) [1992. p. 502](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref24)[–](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref25)[3](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref24).

[through curvature scale space. Int. workshop on Image Databases and Multi-Media](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref25)

[26] [Del Bimbo A, Pala P. Visual image retrieval by elastic matching of user sketches.](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref26) [Search; 1997. p. 51](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref25)[–](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref26)[8](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref25).

[IEEE Trans Pattern Anal Mach Intell 1997;19(2):121](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref26)–[32](http://refhub.elsevier.com/S2590-0056(20)30012-6/sref26).

9