A Normalized Generalized Curvature Scale Space for 2D contour representation

Ameni BENKHLIFA and Faouzi GHORBEL

Grift group, Cristal Laboratory, National School of Computer Sciences
Université de la Manouba, 2010 Manouba, Tunisia
ameni.benkhlifa@ensi-uma.tn , faouzi.ghorbel@ensi.rnu.tn
http://www.ensi.rnu.tn/

Abstract. Here, we intend to propose a discrete normalization of the Generalized Curvature Scale Space (GCSS). The GCSS is an Euclidean invariant planar contour descriptor. It consists on the convolution of the contour by Gaussian functions with different scales. The points having the same curvature values as the selected extremums are the considered key points. This representation implies different number of descriptors from a shape to another. Thus, a step of redistribution of the key points is requested. Therefore, a discrete normalization approach is proceeded. The descriptor is composed by the curvature variation of the key points at the smoothed curve. Several datasets were used to carry on the experiments and to verify the accuracy, the stability and the robustness of the novel description. The Dynamic Time Warping distance is the similarity metric used. Experimental results show that considerable rates of image retrieval are reached comparing to the state of the art.

Keywords: 2D shape description, curvature scale space, iso-curvature, shape classification

1 Introduction

The wide range of applications in computer vision demonstrates the importance of its associated algorithms in many disciplines such as digital medicine, biology, multimedia, remote sensing, robotics. Thus, many benchmarks, created by expert groups of standardization to test and verify these algorithms, witness the importance of the considered problem. Hence, the interest of the classification of shapes is no longer to be proved. However, it is well-known that the problem of 2D shape description is difficult. The shape is a subject to many nonlinear deformations like noise and occlusion, or Euclidean or affine geometric transformations caused by different poses. Such description have often to verify at least the following properties: the efficiency, the stability, the complementeness and the invariance with respect to any transformation belongs to planar Euclidean transformations group $E(2)$.
In this context, many approaches were proposed. These methods could be classified into two major classes: region-based and contour-based ones.

In the first category, we find methods that characterize the shape content. They exploit the information that contain its pixels. Such kind of methods consider the details of the image. Hence, they are reliable for the description of complex shapes like logos, trademarks as mentioned in the work of Kim and al. [1]. Many works were proposed in this context, such as the 2D Zernike moments based methods as complex moments [1],[2]. These methods are very sensitive to local changes such as occlusion or overlapping objects. There are also the 2D Fourier descriptors based methods such as the generic Fourier descriptor of Zang et al. [3] which applied 2-D Fourier transform on polar raster sampled shape image. They also proposed the Enhanced generic Fourier descriptor in [4] which derived the generic Fourier descriptor from the rotation and scale normalized shape. The multi-scale Fourier-based descriptor proposed by Direkoglu et al. [5] represented the shape using its boundary and its content using the Gaussian filter in many scales. It is $E(2)$-invariant and robust to noise. Ghorbel et al. [6] proposed the analytical Fourier Mellin transform which gave also an invariant description for region. In the same context, we find the approach presented by Hong et al. [7] which is based on a kernel descriptor that characterizes local shape.

The second class contains the boundary based methods. There are the Fourier descriptors applied in [8], [9], [10] and [11]. They extracted the global features of the contour.

However other methods treat local features. We find the descriptor of Hoffman et al. [12] who partitioned the curve into parts at negative curvature minima which enhanced the object recognition. Xu et al. [14] proposed another method called contour flexibility which represents the deformable potential at each point of the contour. Klassen et al. [16] presented a differential geometric curve representation using its direction and curvature functions. Shu et al [17] proposed a descriptor named contour points distribution histogram which is based on the distribution of points on object contour under polar coordinates. In the work of Sebastian et al. [18], the contour was characterized by two intrinsic properties: its length and the curvature variations and use them for registration and matching. Their method is called Curve Edit. There is also the descriptor of Belongie et al. [19]. It consists of an algorithm called the Shape Context. At each reference point of the contour, they captured the distribution of the remain points. For two similar contours, the corresponding points had similar shape contexts. This correspondence gives an optimal registration. A new distance called the Inner-Distance was suggested by Ling et al. [20]. It is defined as the length of the shortest path between feature points. This distance can replace the Euclidean distance for complex shapes. It was combined with several methods such as the Shape context [19]. There is also the work of Laiche et al. [21] which is a part-based approach for contour description called Curve Normalization. They represented the shape boundary by an ordered sequence of parts. Then, they associated each part with the cubic polynomial curve using the Least Squared method. A multiscale approaches also were developed. The Angle Scale Descriptor was proposed by Fotopoulou et al. [15]. It consisted of computing the angles between points of the contour in different scales. Another multiscale method of Mokhtarian et al. [22]. It is the curvature scale space which is based on the computation of the maxima of the curvature of the smoothed shape by Gaussian functions in different scales.

In this paper, we intend to propose a discrete normalization of the Generalized Curvature Scale Space proposed in [23]. It is a contour-based descriptor, that we nominate Normalized Generalized Curvature Scale Space (N-GCSS). Our approach is based on the iso-curvature parameterization [24] which is invariant to Euclidean transformations $E(2)$ and CSS [22] method that extracts the extremums of the shape in different scales. In a chosen set of scales, we extract the local extremums of the curvature. We select only those superior to a given threshold. Our descriptor is formed by the points of the contour having the selected curvature levels. Besides, the number of interest point obtained is not the same to all the shapes. Therefore, a novel discrete normalization is proceeded. N-GCSS gives an $E(2)$-invariant non uniform parameterization of the contour since it is constructed by the curvature values of the selected points.

The following paper is organized as follows: we describe the steps of our approach in the second section. In the third section, we expose and discuss the results of the application of our approach using the MPEG7 Set A Part-A1, MPEG7 CE SHAPE-1 Part-B [26] and HMM GPD datasets.

2 Normalized-Generalized Curvature Scale Space

In this section, a detailed description of the GCSS [23] and the discrete normalization will be presented .

2.1 Generalized Curvature Scale Space

The Generalized Curvature Scale Space of [23] corresponds to a set of finite and $E(2)$ invariant points. This representation gives a set of points in the strong variation regions. The GCSS deals with an injective closed contour denoted by C. For a given set of scales $\sigma \in \Sigma$, they applied the Curvature Scale Space proposed by [22]. For each σ the smoothed contour C_σ is given as follows:

$$\begin{aligned} C_\sigma :& [0,1] \to \mathbb{R}^2 \\ & t \mapsto [x(t,\sigma), y(t,\sigma)]^t \end{aligned} \tag{1}$$

In order to extract the key points, the curvature $\kappa(t,\sigma)$ at each point of the smoothed contour C_σ is computed as follows 2:

$$\kappa(t,\sigma) = \frac{x_t(t,\sigma) y_{tt}(t,\sigma) - y_t(t,\sigma) x_{tt}(t,\sigma)}{(x_t^2(t,\sigma) + y_t^2(t,\sigma))^{3/2}} \tag{2}$$

$x_t(t,\sigma)$, $y_t(t,\sigma)$, $x_{tt}(t,\sigma)$ and $y_{tt}(t,\sigma)$ are respectively the first and the second derivatives of $x(t,\sigma)$ and $y(t,\sigma)$. The extremums of $\kappa(t,\sigma)$ are stored in ℓ_σ. A chosen threshold τ was fixed in order to remove the low curvature extremums.

$$\ell_\sigma(\tau) = \{\ell_\sigma \qquad ; \ell_\sigma > \tau\} \tag{3}$$

Since the curvature κ_σ is not a bijective function, the selected points consists of the reciprocal image of $\kappa_\sigma^{-1}(\{\ell_\sigma^j(\tau)\})$.

$$\kappa_\sigma^{-1}(\{\ell_\sigma^j(\tau)\}) = \{t_i^j \qquad /\kappa_\sigma(t_i^j) = \ell_\sigma^j(\tau)\} \tag{4}$$

Where i is the index of the point on C_σ and j is the index of the level. The selected points at each scale σ are saved in a $\digamma_c(\sigma)$. They can be described as follows.

$$\digamma_c(\sigma) = \{C(t_i^j, \sigma) \qquad ; t_i^j = \kappa_\sigma^{-1}\{\ell_\sigma^j(\tau)\}\} \tag{5}$$

The obtained descriptors constitute the following $\digamma_c$.

$$\digamma_c = \bigcup_{\sigma \in \Sigma} \digamma_c(\sigma) \tag{6}$$

Therefore, our descriptor is composed by the curvature values of these key points in the selected scales:

$$\digamma = \bigcup_{\sigma \in \Sigma} \digamma(\sigma) \qquad ; \digamma(\sigma) = \kappa_\sigma(\digamma_c(\sigma)) \tag{7}$$

The steps of the GCSS are described in Figure.1.

2.2 Discrete normalization of the key points

GCSS could be seen also as a new parameterization of the contour. Such parameterization is $E(2)$-invariant as the curvature. However, in the discrete case, $\digamma_c$ is a set of unordered points of C because it is a resampling procedure of the points. In the last step, the obtained set of points is ordered. Such set is distributed non- uniformly and we have more points in strong curvature areas. In order to make the number of points in $\digamma_c$ the same for all contours in the dataset, a discrete normalization step is proceeded.
We denote by N the number of key points from GCSS of a given curve ie $N = card(\digamma_c)$. However, the wished number of points is N_w. Let $\digamma_c^*$ be the normalized set. In order to obtain N_w interest points from the total N, we start by computing the cumulative distance between the starting point P_1 and P_i where i the i^{th} point. This procedure is equivalent to define a finite function from $1..N$ to an interval $[0, a]$ from $\mathbb{R}$. We consider $S(P_i)$ defined as follows:

$$S(P_i) = \int_{\widehat{P_1, P_i}} ||C'(t)||dt \tag{8}$$

Hence, we resample regularly the abscissa vector $[P_1..P_N]$ into $[P_1^*..P_{N_w}^*]$. We compute $S(P_i^*)$ and we search the nearest points $S(P_j)$.

$$\underset{j}{\operatorname{argmin}} ||S(P_i^*) - S(P_j)|| \tag{9}$$

Figure 2 gives an illustration of the proposed normalization procedure for $N = 28$ and $N_w = 10$.

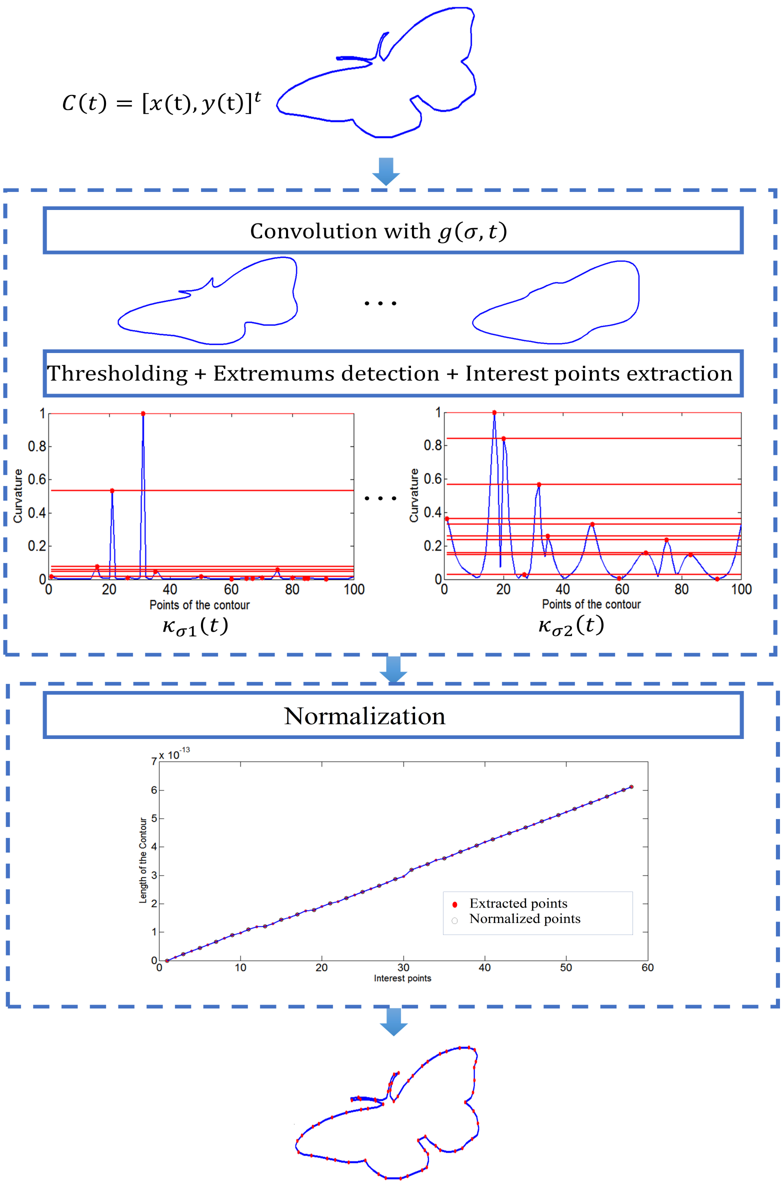

Fig. 1. The Block Diagram of the Normalized Generalized Curvature Scale Space.

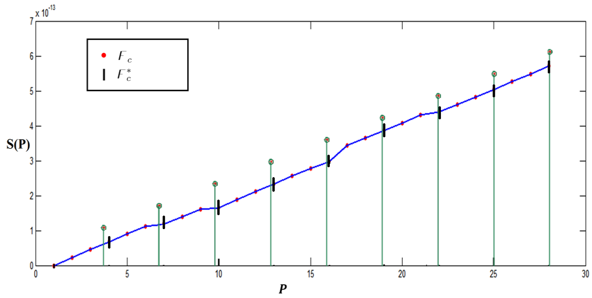

Fig. 2. A discrete normalization example from $N = 28$ to $N_w = 10$

2.3 The invariance to $E(2)$

The Normalized Generalized Scale Space descriptor is based on the computation of the curvature of the smoothed contour in given scales. As the curvature behavior is the same whatever is the transformation applied to the contour: rotation, translation or scale, the obtained set of points of C and $g(C)$ is the same. Where g is an $E(2)$ transformation. The problem of the starting point is resolved by the use of the Dynamic Time Warping [28] as similarity metric. Figure 3 demonstrates well the distribution of the key points and their invariance under $E(2)$.

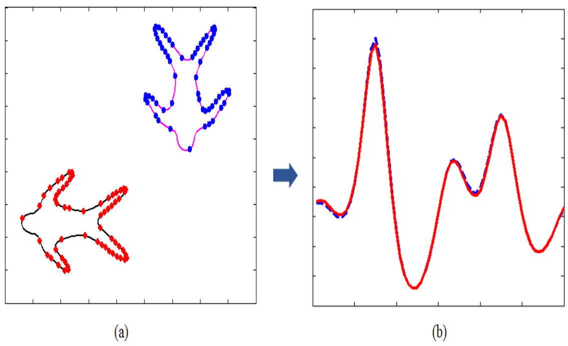

Fig. 3. The invariance under $E(2)$ transformation (a) point of interest on the original curves (b) the curvature variation of the two signatures.

3 Experiments and results

The performance of the N-GCSS is tested on three datasets and evaluated in terms of shape retrieval efficiency and precision-recall curves. The datasets used for the experimentations are: HMM GPD and MPEG7 CE Shape-1 Part-B [26] and MPEG7 Part-A1. For the recognition, each object is compared to all the shapes in the dataset using the Dynamic Time Warping [28] algorithm and matched to the closest one.

3.1 The datasets

The HMM GPD is composed of four sub datasets as shows the following table 1: bicego-data [25] , plane-data , mpeg-data and car-data [27]. We form an other set of shapes using the four sub datasets (bicego, plane, car and mpeg) by picking up the 20 first elements of each class. The MPEG7 CE Shape-1 Part-B dataset

Table 1. HMM GPD sub-datasets

Sub-dataset	Number of objects
Bicego	140
Plane	210
Car	120
Mpeg	120
HMM	480

[26] is a well-known dataset. It is composed of 1400 elements that are grouped in 70 classes. Each class contains 20 images.
The MPEG7 Part-A1 is composed of 420 objects grouped in 70 classes. It is used to test the performance of the descriptor under scales transformations

3.2 Results

In this paragraph, the experiments on the above datasets are carried on. The parameters of N-GCSS are chosen empirically: $\sigma \in 5, 6$ and $\tau = 10^{-3}$ in order to eliminate the local extremums having very low curvature. We choose $N_w = 100$ the number of key points for each contour.
We proceeded the k nearest neighbors (k-NN) algorithm in order to compute the pairwise shape matching scores in the recognition step. For each shape, the distance (Dynamic time warping) is computed from all the other shapes in the dataset and the knearest neighbors are selected.
To evaluate the performance of our representation and to compare it with other techniques from the state of the art. Table.2 lists the retrieval results of our descriptor N-GCSS on HMM GPD dataset using 1NN algorithm. We reach very high score for Mpeg (100%) and Plane (98.57%) datasets. Although the Car sub-dataset contains bad quality contours, N-GCSS outperforms the CSS descriptor [22] and reaches 73.33%. This demonstrates well the robustness of our descriptor to numerical approximation. This robustness is due to the use of the multiscale approach in the construction of the proposed approach.

Table 2. Retrieval results on HMM dataset using 1NN algorithm for : N-GCSS and CSS [22]

	Rate N-GCSS (%)	Rate CSS (%)
Bicego	94.29	90.00
Plane	98.57	79.52
Car	73.33	55.00
Mpeg	100	95.83
HMM	85.42	75.62

Our descriptor was compared also to the Contour Points Distribution Histogram (CPDH) [17], Fourier Descriptor (FD) [11] and Curvature Scale Space (CSS) [22] for MPEG7 CE Shape-1 Part-B [26]. Figure 4 shows the precision-recall curves of the mentioned descriptors on the MPEG7 CE Shape-1 Part-B [26]. The proposed representation gives higher precision rates than the above-mentioned descriptors. By computing the retrieval rate using the 1NN algorithm for this dataset [26], we reached 91.27%.
The performance of the proposed descriptor is compared also with other approaches in the literature. The retrieval rates are measured with the Bull's eyes algorithm. Each shape is considered as a query and we count how many objects within the 40 most similar objects belong to the class of the query. Table 3 lists the Bull's eye scores of some algorithms. We remark that our method gives a competitive score comparing to the state of the art.

Table 3. Bull's eye MPEG7 CE SHAPE-1 Part-B [26]

IDSC [20]	85.40%
N-GCSS	77.20%
CPDH [17]	76.56%
SC [19]	76.51%
CSS [22]	75.44%
ASD [15]	70.51%
Curve normalization [21]	50.76%

We used also the MPEG7 Part-A1 set in order to test the accuracy of the proposed descriptor under scale transformation. N-GCSS outperforms the Fourier Descriptor (FD) [11] and Curvature Scale Space (CSS) [22]. The precision-recall curves are illustrated in Figure 5. The results prove the accuracy, stability and invariance to $E(2)$ properties of the proposed representation.

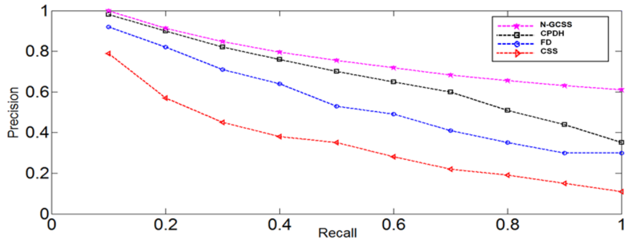

Fig. 4. Retrieval Rate for MPEG7 CE Shape-1 Part-B dataset for :N-GCSS, CPDH, FD and CSS

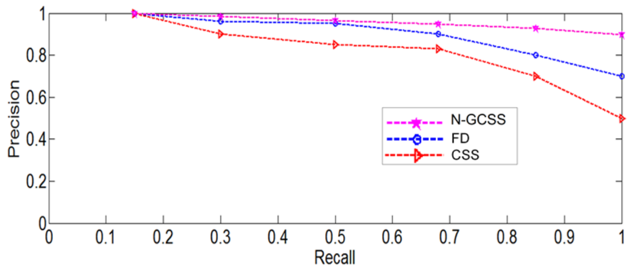

Fig. 5. Precision recall for MPEG7 Part-A1 dataset for : N-GCSS, FD and CSS

4 Conclusion

In this work, we introduced a 2D contour description. It is based on the Generalized Curvature Scale Space of [23]. The GCSS is constructed by a $E(2)$ invariant set of points. This set corresponds to the high curvature zones of a contour. The contribution of this paper lies on the discrete normalization of the obtained key points. This normalization makes comparison between the contours easier. The problem of starting point is overcome by the use of the Dynamic Time Warping [28]. The performance of the normalized GCSS is tested by carrying out many experiments on two well know datasets which are: HMM GPD, MPEG7 CE SHAPE-1 Part-B [26] and MPEG7 Part-A1 and comparing to several descriptors. Results show well the accuracy of our $E(2)$-invariant descriptor and its stability and robustness to numerical approximation.
Many perspectives could be cited to this work. We intend to present a method of choosing the set of scales while considering the complexity of the shape to be analyzed. We look also for a combination of this novel parameterization with other descriptors from the state of the art since the set of key points found can be considered as a non-uniform parameterization of the contour. They are located at strong-curvature zones only. However, the arc-length reparameterization gives a uniform distribution of points which can be a waste of energy. An immigration to the 3D dimension also is among our future work.

References

1. Kim, W.Y., Kim, Y.S., 2000. A region-based shape descriptor using zernike moments. Signal processing: Image communication 16, 95-102.
2. Khotanzad, A., Hong, Y.H., 1990. Invariant image recognition by zernike moments. IEEE Transactions on pattern analysis and machine intelligence 12, 489-497.
3. Zhang, D., Lu, G., 2002b. Generic fourier descriptor for shape based image retrieval, in: Multimedia and Expo, 2002. ICME'02. Proceedings. 2002 IEEE International Conference on, IEEE. pp. 425-428.
4. Zhang, D., Lu, G., 2002a. Enhanced generic fourier descriptors for object-based image retrieval, in: Acoustics, Speech, and Signal Processing (ICASSP), 2002 IEEE International Conference on, IEEE. pp. IV-3668.
5. Direkoglu, C., Nixon, M.S., 2008. Shape classification using multiscale fourier-based description in 2-d space, in: Signal Processing, 2008. ICSP 2008. 9th International Conference on, IEEE. pp. 820-823.
6. Ghorbel, F., 1998. Towards a unitary formulation for invariant image description: application to image coding. Annals of telecommunications 53, 242-260.
7. Hong, B.W., Prados, E., Soatto, S., Vese, L., 2006. Shape representation based on integral kernels: Application to image matching and segmentation, in: Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on, IEEE. pp. 833-840.
8. Ghorbel, F., 1992. Stability of invariant fourier descriptors and its inference in the shape classification, in: Pattern Recognition, 1992. Vol. III. Conference C: Image, Speech and Signal Analysis, Proceedings., 11th IAPR International Conference on, IEEE. pp. 130-133.
9. Persoon, E., Fu, K.S., 1986. Shape discrimination using fourier descriptors. IEEE transactions on pattern analysis and machine intelligence , 388-397.
10. Wallace, T.P., Wintz, P.A., 1980. An efficient three-dimensional aircraft recognition algorithm using normalized fourier descriptors. Computer Graphics and Image Processing 13, 99-126.
11. Rui, Y., She, A. C., Huang, T. S. 1997. A modified Fourier descriptor for shape matching in MARS, Image Databases and Multimedia Search 8 (1998) 165-180.
12. Hoffman, D.D., Richards, W.A., 1984. Parts of recognition. Cognition 18, 65-96.
13. Siddiqi, K., Kimia, B.B., 1995. Parts of visual form: Computational aspects. IEEE Transactions on Pattern Analysis and Machine Intelligence 17, 239-251.
14. Xu, C., Liu, J., Tang, X., 2009. 2d shape matching by contour flexibility. IEEE Transactions on Pattern Analysis and Machine Intelligence 31, 180-186.
15. Fotopoulou, F., Economou, G. 2011. Multivariate angle scale descriptor of shape retrieval. Proc. Signal. Process Appl. Math. Electron. Commun, 105-108.
16. Klassen, E., Srivastava, A., Mio, M., Joshi, S.H., 2004. Analysis of planar shapes using geodesic paths on shape spaces. IEEE transactions on pattern analysis and machine intelligence 26, 372-383.
17. Shu, X., Wu, X. J. 2011. A novel contour descriptor for 2D shape matching and its application to image retrieval. Image and vision Computing, 29(4), 286-294.
18. Sebastian, T. B., Klein, P. N., Kimia, B. B. 2003. On aligning curves. IEEE transactions on pattern analysis and machine intelligence, 25(1), 116-125.
19. Belongie, S., Malik, J., Puzicha, J. 2002. Shape matching and object recognition using shape contexts. IEEE transactions on pattern analysis and machine intelligence, 24(4), 509-522.

20. Ling, H., Jacobs, D. W. 2007. Shape classification using the inner-distance. IEEE transactions on pattern analysis and machine intelligence, 29(2), 286-299.
21. Laiche, N., Larabi, S., Ladraa, F., Khadraoui, A. 2014. Curve normalization for shape retrieval. Signal Processing: Image Communication, 29(4), 556-571.
22. Mokhtarian, F., Abbasi, S., Kittler, J., 1996. Robust and efficient shape indexing through curvature scale space, in: Proceedings of the 1996 British Machine and Vision Conference BMVC, Citseer.
23. Benkhlifa, A., Ghorbel, F. 2016. A Novel 2D Contour Description Generalized Curvature Scale Space. In International Workshop on Representations, Analysis and Recognition of Shape and Motion FroM Imaging Data (pp. 129-140). Springer, Cham.
24. Bannour, M., Ghorbel, F., 2000. Isotropie de la representation des surfaces; application a la description et la visualisation d' objets 3d, in: Conf. proc. RFIA, pp. 275-282.
25. Bicego, M., Murino, V., Figueiredo, M.A., 2004. Similarity-based classification of sequences using hidden markov models. Pattern Recognition 37, 2281-2291.
26. Latecki, L.J., Lakamper, R., Eckhardt, T., 2000. Shape descriptors for non-rigid shapes with a single closed contour, in: Computer Vision and Pattern Recognition, 2000. Proceedings. IEEE Conference on, IEEE. pp. 424-429.
27. Thakoor, N., Gao, J., Jung, S., 2007. Hidden markov model-based weighted likelihood discriminant for 2-d shape classification. IEEE Transactions on Image Processing 16, 2707-2719.
28. Sankoff, D., Kruskal, J. B. 1983. Time warps, string edits, and macromolecules: the theory and practice of sequence comparison. Reading: Addison-Wesley Publication, 1983, edited by Sankoff, David; Kruskal, Joseph B.