Proceedings Paper:
Alwaely, B. and Abhayaratne, C. orcid.org/0000-0002-2799-7395 (2018) Graph spectral domain shape representation. In: 2018 26th European Signal Processing Conference (EUSIPCO). 2018 26th European Signal Processing Conference (EUSIPCO), 03-07 Sep 2018, Rome, Italy. IEEE , pp. 598-602. ISBN 978-9-0827-9701-5

10.23919/EUSIPCO.2018.8553054

Graph Spectral Domain Shape Representation

Basheer Alwaely* , Charith Abhayaratne†
Department of Electronic and Electrical Engineering
The University of Sheffield
Sheffield, S1 3JD, United Kingdom
Email: *b.alwaely@sheffield.ac.uk, †c.abhayaratne@sheffield.ac.uk

***Abstract*—One of the major challenges in shape matching is recognising and interpreting the small variations in objects that are distinctly similar in their global structure, as in well known ETU10 silhouette dataset and the Tool dataset. The solution lies in modelling these variations with numerous precise details. This paper presents a novel approach based on fitting shape's local details into an adaptive spectral graph domain features. The proposed framework constructs an adaptive graph model on the boundaries of silhouette images based on threshold, in such a way that reveals small differences. This follows feature extraction on the spectral domain for shape representation. The proposed method shows that interpreting local details leading to improve the accuracy levels by 2% to 7% for the two datasets mentioned above, respectively.**

***Index Terms*—Graph spectral analysis, Graph spectral features, Shape matching, Adaptive graph connectivity.**

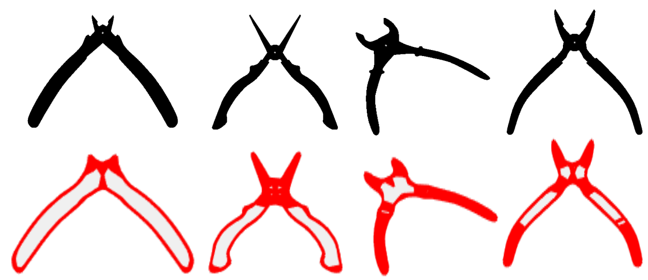

Fig. 1. Challenging objects with conceptual high similarity (top row) and associated graph connectivity (bottom row).

I. Introduction

Understanding the content of visual media and captured images receives a lot of interest in computer vision [1] and security [2]. The human eye can easily discriminate patterns in images. However, a manual classification process of these patterns is a challenging task for researchers. For example, objects may be classified in the same class, although they have different types of geometric structures as shown in Fig. 1. This is because the similarities between the patterns are high in terms of the global structure, while few differences are noticed. As a result, the interpretation of small details is an important factor for distinguishing between different shapes. This challenge motivates us to exploit differences in shape in terms of protrusions and fine details as well as the global shape.

Many strategies have been introduced to design the optimal model. Shape matching studies can be considered from two different perspectives: feature-based methods [3], [4] and model-based methods, which are based on either a shape-skeleton [5], [6] or shape-contour [7], [8]. For the model-based studies, on the one hand, skeleton-based studies have mainly constructed a tree model using the object edges to form a shape descriptor, where the similarity measurement is based on the tree matching approaches. For example, different methods are implemented by creating a shape descriptors prototype using: points corresponding [5], part decomposition [6], shortest path [9] and relative measurement [10].

On the other hand, several studies have relied on the boundaries of silhouette images. These edges efficiently characterise the global structure of the object with a single closed curve, if there are no holes in the object. An early study was proposed in this regard by Zahn and Roskies [11] using Fourier descriptors to represent the shape, while the latest studies are based on: circle view shape signature (CVs) with multi view [7] and applying tangent PCA for variation mode extraction [12].

Shapes are also recognized using feature-based representation that typically requires more than one feature to describe a complex structure. Such features may include a scale invariant feature transform (SIFT) [13], tree union [3], local phase [14], distance and the central point [4], contour features and distance [15]. With regards to the graph-based studies, using graph for 2D shape matching [16], [17] mainly relied on bipartite matching.

However, the majority of existing works takes into account the global shape to design their model, while they do not pay attention to fine details. Interpreting these areas greatly enhances the detection process, especially in high-similarity shapes as in the Tool dataset Fig. 1, which has not been evaluated much compared to the other datasets. Therefore, the optimal shape-matching system must have the ability to interpret and understand the protrusions of the boundary along with the overall structure containment. This paper proposes a new method for 2D shape recognition based on graph adaptive connectivity. To the best of our knowledge, this is the first study to recognize shapes based on their local details.

The main contributions of this paper are:

- Methodology for forming an adaptive graph connectivity to capture the local protrusions of the shapes.
- Proposing a new set of features based on the graph spectral domain.

This paper is structured as follows: a comprehensive explanation of the graph concepts, graph connectivity, graph basis and graph spectral features are shown in Section II. Then, Section III will evaluate and discuss the proposed system based on different classifiers and parameters. Finally, the work will be concluded in Section IV.

II. The Proposed Method

The objective of this paper is to identify shapes using both global and local details. Fig. 2 shows an overview of the proposed shape-matching algorithm that matches the corresponding graph spectral features of the shape via machine learning.

For a given 2D binary shape S, we extract its contour (x, y) from the input image silhouette using an edge detector filter such as Sobel operator. The 2D path P of the shape's contour with length N is then moved in such a way that the point $(0, 0)$ will be in the center box of the shape. Since this paper utilizes machine learning for classification step, a fixed number of pixels is needed to generate the same length features. Therefore, $n < N$ pixels are selected from P to be a candidate to form a new path $\hat{P}$ as shown in (1) and (2).

$$\hat{P}(i) = P(i * H), \qquad i = 0, \ldots, n-1. \tag{1}$$

and,

$$H = \frac{N-1}{n-1}. \tag{2}$$

This process is similar to signal down-sampling processing, where n-samples are uniformly selected from N-samples. Then, $\hat{P}$ is used to generate a graph, which represents the shape structure. Further information can be found in the next subsections to describe the graph concepts, graph connectivity, graph basis and finally the proposed features.

A. Graph concepts

This work generates an undirected graph G=$\{\varepsilon, v\}$, which comprises vertices v or nodes. These nodes are connected by edges (ε) that represent the Euclidean distance between nodes in the graph adjacency matrix $A \in \mathbb{R}^{N\times N}$ as shown in (3).

$$A_{i,j} = \begin{cases} \varepsilon_{(i,j)}, & \text{if node i and j are connected,} \\ 0, & \text{otherwise.} \end{cases} \tag{3}$$

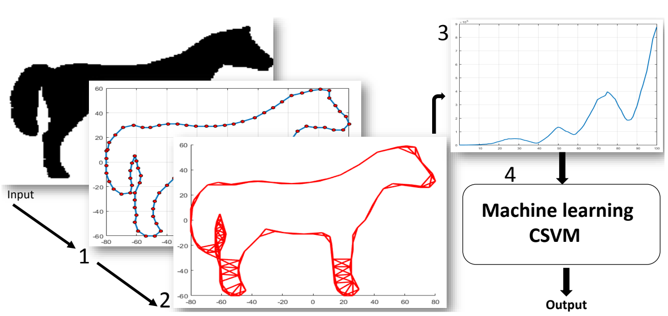

Fig. 2. The proposed method includes: 1-contour extraction, 2- graph generation, 3-feature extraction and 4-classification process.

Then, the non-normalised graph Laplacian matrix is obtained as $L := D - A$, where $D_{(i,i)} = \sum_{j=1}^{N} A_{(i,j)}$ $i = 1, 2, ..., n$. D has zeros in its off-diagonal elements and the diagonal element corresponding to the degree of each vertex is the summation of the weights of all its connected edges.

A complete set of orthonormal eigenvectors χ_ℓ of L and their associated real eigenvalues λ_ℓ for $\ell = 0, ..., n-1$ are calculated. The eigenvalues of the non-normalised graph Laplacian matrix are ordered as $0 = \lambda_0 < \lambda_1 \leq \lambda_2 ... \leq \lambda_{n-1} = \lambda_{max}$.

B. Graph connectivity

Connectivity has crucial rules in generating graphs because the way to connect nodes has a direct effect on the spectral basis. For global shape prediction, usually full connected graph, where each node has $(n-1)$ connections or $\Phi_i = n-1)$ provides an efficient representation [18]. However, in this case local details are not reliably detected.

This work therefore applies a conditional connectivity to reveal local details in shape's contour. Conditional connectivity means that each node is connected to other nodes that fall in less than a certain distance as shown in Fig. 3 as an example. A certain threshold T is used to determine the distance to link pixels in each shape. T can also be defined as the minimum distance that keeps all nodes connected, as will be shown in Section II-C. The next subsection will shown in details how to compute T.

C. The graph eigenvalues.

For better understanding of the connectivity effects on the graph basis, Fig. 4 illustrates a single shape with four different values of T and the graph eigenvalues of these shapes are listed below:

λ_ℓ of A= $[0,\ 0,\ 0,\ 0, ..., 0.50,\ 1.86, ..., 24.97]$.

λ_ℓ of B= $[0,\ 0,\ 0,\ 0,\ 0.01,\ 0.06, ..., 34.54]$.

λ_ℓ of C= $[0,\ 0.030,\ 0.050, ..., 49.94]$.

λ_ℓ of D= $[0,\ 0.031,\ 0.054, ..., 60.49]$.

From the graph eigenvalues, we can find that:

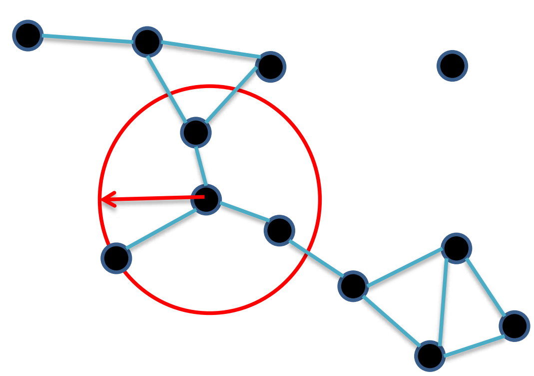

Fig. 3. Nodes with certain distance are connected.

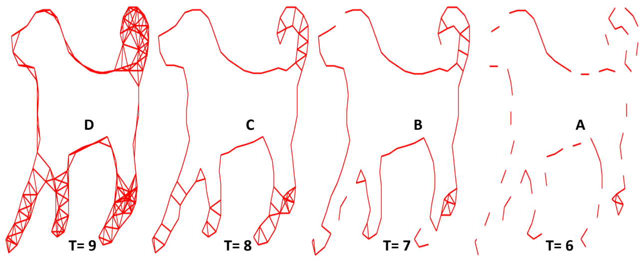

Fig. 4. Graph construction with different value of T.

1) The shape's density.
   As can be seen in λ_ℓ of A, B, C and D, the last eigenvalue (λ_{n-1}) reflects the density of the shape. For example, λ_{n-1} has a large value in high density shapes in Fig. 4 D compared with low density in Fig. 4 A.
2) The number of clusters.
   The graph eigenvalue can be also used as an indicator to the number of existing clusters. For example, in shape A, $\lambda_0 \to \lambda_{38}$ are equal to zeros, which means that there are 39 groups in Fig. 4 A. These groups can be made up of connected nodes or an individual non-connected node. For shape B, $\lambda_0 \to \lambda_3$ are equal to zeros, which means that there are four groups as can be seen in Fig. 4 B. In both shapes C and D, only λ_0 is equal to zeros, and that means all the existing nodes are connected as a single group.

The graph eigenvalue confirms the importance of connectivity parameters. Based on this interpretation, the number of zeros in the graph eigenvalues is used to check whether the graph is connected as a single group or not.

D. The boundary scale of T measurement.

In this paper, we refer to the number of connected nodes at each pixel with ϕ. For a given S_ℓ, we compute T_ℓ, which is the minimum distance that makes $\lambda_1 \neq 0$. Then, T is used in graph generation process to compute the spectral basis. However, the different value of T produces a different number of nodes connected to each pixel. Fig. 5 shows the nodes connectivity of random shape with 40 pixels. In this example, T locates the number of connected nodes in the range between upper and lower limits. We can determine the boundary scale of Φ at each node according to :

1) Upper boundary: each pixel reaches the maximum value at $n-1$, which is 39 in this example using high value of T.

2) Lower boundary is the minimum distance that keeps all nodes are connected as a one group.

E. Graph Spectral Features (GSF)

A novel set of features is proposed for shape matching using the graph spectral representation of the adaptive graph connectivity. The proposed feature captures the local F_L and the global details F_G. Therefore, this paper uses:

1) Local features (F_L) are presented by number of connected elements at each pixel, Φ_i.
$$F_L = \Phi_i, \qquad i = 0, \ldots, n-1. \tag{4}$$
2) Global shape features (F_G) are presented by scaling the graph eigenvalues by the distance ($\ominus$). This combination results in high distinctive features.
$$F_G = \ominus_i \lambda_i, \qquad i = 0, \ldots, n-1, \tag{5}$$
   where $\ominus$ equals to to the distance of each node to the point (0,0) :
$$\ominus_i = \sqrt{x_i^2 + y_i^2}, \qquad i = 0, \ldots, n-1. \tag{6}$$

The final features with length $2n$ are:
$$GSF = [F_L, F_G] \tag{7}$$

F. Machine learning

Based on several experiments conducted to select the best classifier using four datasets, Support Vector Machine with a cubic form as a kernel function shows better performance compared to other classifiers, as will be shown in Section III. The experiments include: a Support Vector Machine with a cubic form as a kernel function (CSVM), Nearest Neighbour (KNN), Support Vector Machine with a Quadratic form as a kernel function (QSVM), Classification Tree (CT), and Decision tree (DT).

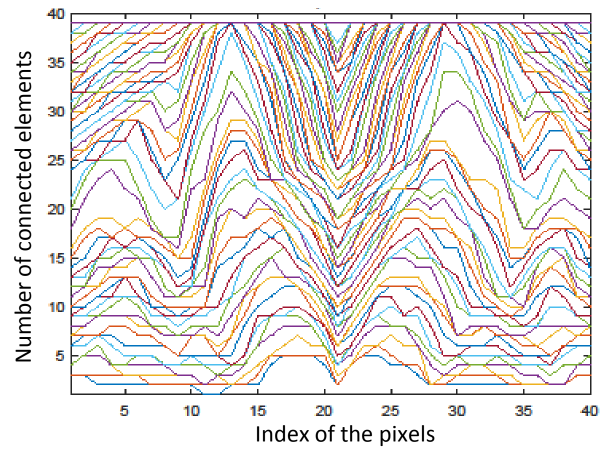

Fig. 5. Number of connected node at each pixel using different values of threshold.

III. Experimental validation

To evaluate the proposed features for shape matching, a large number of experiments is implemented using many sets of 2D shapes. 10-fold cross validation scheme is utilised to train and test all the datasets. N=100 is used to generate graphs.

TABLE I
AVERAGE RECOGNITION SCORE (%) OF 12 EXPERIMENTS.

Dataset	CSVM	KNN	QSVM	CT	DT
ETU10 silhouette	99.1	99.24	90.7	93.17	90
Tool	97.14	97.14	97.14	96.66	96.90
Kimia 99	95.87	95.87	94.44	86.5	90.88
Kimia 216	94.75	93.75	91.12	86.45	90.08

Initially, we test different classifiers to determine the optimal classifier for recognition. The mean accuracy of 12 experiments is shown in Table I for all the datasets. It is clear that CSVM and KNN show better performance than the other classifiers. Therefore, we select a CSVM to evaluate the proposed features based on four public and well-known datasets, which are:

1) ETU10 silhouette dataset.
   The ETU10 database has 10 classes×72 shapes in each class = 720 total images. Sample silhouettes from each class are shown in the top two rows of Fig. 6A. The bottom row shows different angles of the object. The evaluation results exceed the state-of-the-art range by recording 99.31% mean accuracy as shown in Fig. 7. The ten classes in the confusion matrix are corresponding to the Bed, Bird, Fish, Guitar, Hammer, Horse, Sink, Teddy, Television and Toilet respectively.
2) Tool dataset.
   The tool dataset in Fig. 6B contains 35 articulated shapes, which are classified into four classes: 10 scissors, 15 pliers, 5 knives and 5 pincers respectively. The mean accuracy of 12 experiments achieves to 97.14% as shown in Fig. 7.
3) Kimia 99 dataset.
   The Kimia 99 dataset consists of 9 classes $\times$ 11 samples = 99 images as shown in Fig. 6C. The mean accuracy of 12 experiments is 96% as shown in Fig. 7. There is no case that causes serious confusion in this dataset. The nine classes in the confusion matrix correspond to the Fish, Hand, Human, Aeroplane, Ray, Rabbit, Misk, Spanner and Dog respectively.
4) kimia 216 dataset.
   The kimia 216 dataset consists of 18 classes $\times$ 12 samples = 216 images as shown in Fig. 6D. The mean accuracy of 12 experiments is 95.37% as shown in Fig. 7. The 18 classes in the confusion matrix correspond to the Bird, Bone, Brick, Camel, Car, Children, Classic, Elephant, Face, Fork, Fountain, Glas, Hammer, Heart, Key, Misk, Ray and Turtle respectively.

Although samples in these datasets have different angles of views, our proposed feature provides a unique description for each sample. In the literature, there are many methods to evaluate shape matching studies such as precision-recall curve, retrieval table and machine learning.

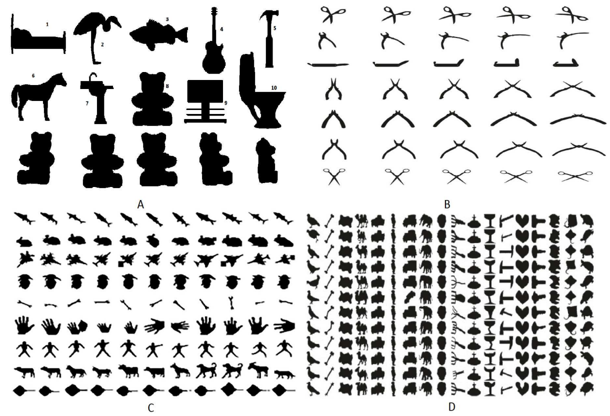

Fig. 6. A: ETU10 silhouette Dataset, B: Tools Dataset, C: Kimia99 dataset and D: Kimia216 dataset.

This paper relies on a machine learning using support vector machine classifier to recognize samples based on its features. SVM classifier recognizes samples based on the available trained classes, which means that SVM is a class-based classifier rather than samples. Therefore, finding the closest features or sample to create a retrieval table is not possible.

To compare with the state-of-the-art performance, Table II shows the recognition score of the proposed GSF and the maximum score of the existing works for the four datasets. The proposed features perform better than the existing work using ETU10 silhouette and Tool datasets. kimia 99 and kimia 216 have been used in many studies and satisfactory results are provided using graph spectral features, comparable with other works.

IV. CONCLUSIONS

Recognizing the content of images is an important issue in computer vision field. This paper proposes graph spectral features for 2D shape matching based on the shape contour. A graph model based on adaptive connectivity has been created to cover the variation in shapes. Based on this connectivity, a novel set of features have been presented to interpret the local and global details of the shapes. This results in high recognition level even with high similarity shapes. Four public and well-known datasets are used for evaluation: ETU10 dataset, Tool dataset, kimia 99 and kimia 216 datasets. The evaluation process shows that the proposed features are robust in detecting different kinds of complex shapes and with satisfactory results compared with the state-of-the-art studies.

TABLE II
COMPARISON WITH THE STATE-OF-THE-ART STUDIES.

Dataset	Proposed graph features (%)	Existing work (%)
ETU10 silhouette	99.31	97.5 [5]
Tool	97.14	90.86 [14]
Kimia 99	96	99 [15]
Kimia 216	95.37	98.15 [4]

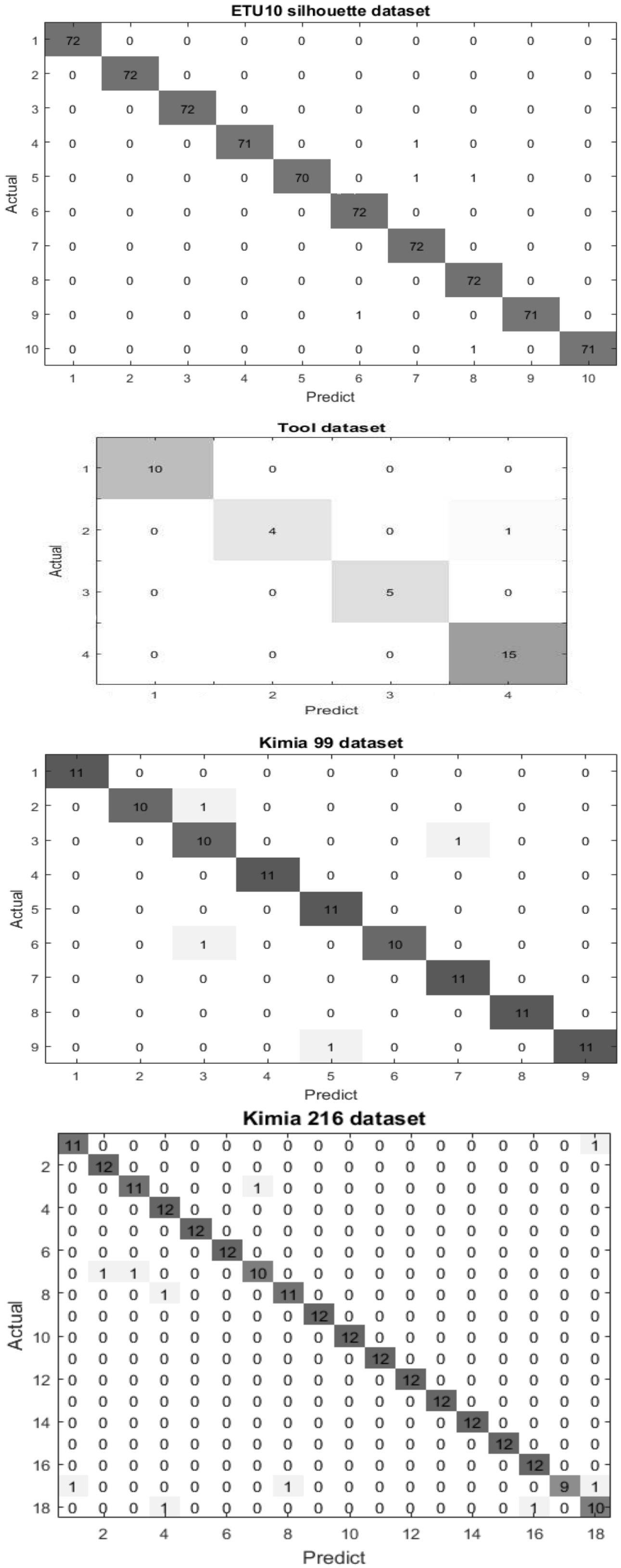

Fig. 7. Confusion matrices of the datasets.

REFERENCES

[1] F. De la Torre and M. J. Black, "Robust principal component analysis for computer vision," in *Proc. International Conference on Computer Vision, ICCV.*, vol. 1. IEEE, 2001, pp. 362–369.

[2] S.-L. Chang, L.-S. Chen, Y.-C. Chung, and S.-W. Chen, "Automatic license plate recognition," *transactions on intelligent transportation systems*, vol. 5, no. 1, pp. 42–53, 2004.

[3] B. Wang, W. Shen, W. Liu, X. You, and X. Bai, "Shape classification using tree-unions," in *Proc. International Conference on Pattern Recognition (ICPR),*. IEEE, 2010, pp. 983–986.

[4] J. C. Figueiredo, F. G. M. Neto, and I. C. de Paula, "Contour-based feature extraction for image classification and retrieval," in *Proc. International Conference of the Chilean Computer Science Society (SCCC).* IEEE, 2016, pp. 1–7.

[5] M. Akimaliev and M. F. Demirci, "Improving skeletal shape abstraction using multiple optimal solutions," *Pattern Recognition*, vol. 48, no. 11, pp. 3504–3515, 2015.

[6] W.-B. Goh, "Strategies for shape matching using skeletons," *Computer vision and image understanding*, vol. 110, no. 3, pp. 326–345, 2008.

[7] H. D. Jomma and A. I. Hussein, "Circle views signature: A novel shape representation for shape recognition and retrieval," *Canadian Journal of Electrical and Computer Engineering*, vol. 39, no. 4, pp. 274–282, 2016.

[8] S. Belongie, J. Malik, and J. Puzicha, "Shape matching and object recognition using shape contexts," *Transactions on pattern analysis and machine intelligence*, vol. 24, no. 4, pp. 509–522, 2002.

[9] X. Bai, L. J. Latecki, and W.-Y. Liu, "Skeleton pruning by contour partitioning with discrete curve evolution," *Transactions on pattern analysis and machine intelligence*, vol. 29, no. 3, 2007.

[10] A. Temlyakov, B. C. Munsell, J. W. Waggoner, and S. Wang, "Two perceptually motivated strategies for shape classification," in *Proc. Conference on Computer Vision and Pattern Recognition (CVPR).* IEEE, 2010, pp. 2289–2296.

[11] C. T. Zahn and R. Z. Roskies, "Fourier descriptors for plane closed curves," *Transactions on computers*, vol. 100, no. 3, pp. 269–281, 1972.

[12] M. Abboud, A. Benzinou, K. Nasreddine, and M. Jazar, "Robust statistical shape analysis based on the tangent shape space," in *Proc. International Conference on Image Processing (ICIP).* IEEE, 2015, pp. 3520–3524.

[13] I. Setitra and S. Larabi, "Sift descriptor for binary shape discrimination, classification and matching," in *Proc. International Conference on Computer Analysis of Images and Patterns*. Springer, 2015, pp. 489–500.

[14] L. Nanni, S. Brahnam, and A. Lumini, "Local phase quantization descriptor for improving shape retrieval/classification," *Pattern Recognition Letters*, vol. 33, no. 16, pp. 2254–2260, 2012.

[15] D. C. G. Pedronette and R. da Silva Torres, "Shape retrieval using contour features and distance optimization." in *VISAPP (2)*. Citeseer, 2010, pp. 197–202.

[16] M. Leordeanu and M. Hebert, "A spectral technique for correspondence problems using pairwise constraints," in *Proc. International Conference on Computer Vision (ICCV'05) Volume 1*, vol. 2, Oct 2005, pp. 1482–1489 Vol. 2.

[17] D. Mateus, R. Horaud, D. Knossow, F. Cuzzolin, and E. Boyer, "Articulated shape matching using laplacian eigenfunctions and unsupervised point registration," in *Proc. Conference on Computer Vision and Pattern Recognition, CVPR.* IEEE, 2008, pp. 1–8.

[18] B. Alwaely and C. Abhayaratne, "Graph spectral domain feature representation for in-air drawn number recognition," in *Proc. European Signal Processing Conference (EUSIPCO)*. IEEE, 2017, pp. 370–374.