IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received October 4, 2019, accepted October 27, 2019, date of publication October 31, 2019, date of current version November 13, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2950643

Graph Spectral Domain Feature Learning With Application to In-Air Hand-Drawn Number and Shape Recognition

BASHEER ALWAELY, (Student Member, IEEE), AND CHARITH ABHAYARATNE, (Member, IEEE)
Department of Electronic and Electrical Engineering, The University of Sheffield, Sheffield S1 3JD, U.K.
Corresponding author: Charith Abhayaratne (c.abhayaratne@sheffield.ac.uk)

ABSTRACT This paper addresses the problem of recognition of dynamic shapes by representing the structure in a shape as a graph and learning the graph spectral domain features. Our proposed method includes pre-processing for converting the dynamic shapes into a fully connected graph, followed by analysis of the eigenvectors of the normalized Laplacian of the graph adjacency matrix for forming the feature vectors. The method proposes to use the eigenvector corresponding to the lowest eigenvalue for formulating the feature vectors as it captures the details of the structure of the graph. The use of the proposed graph spectral domain representation has been demonstrated in an in-air hand-drawn number and symbol recognition applications. It has achieved average accuracy rates of 99.56% and 99.44%, for numbers and symbols, respectively, outperforming the existing methods for all datasets used. It also has the added benefits of fast real-time operation and invariance to rotation and flipping, making the recognition system robust to different writing and drawing variations.

INDEX TERMS Graph signal processing, graph spectral theory, fully connected graphs, dynamic shape recognition, shape representation, graph spectral feature learning.

I. INTRODUCTION

Graph Signal Processing (GSP) [1]–[3] has attracted great attention in processing, analysis, coding and understanding of data sampled on a non-uniform grid, often referred to as irregular data or graph data. The classical discrete signal processing is not directly applicable on such irregularly sampled data. GSP provides a robust mechanism to represent irregular data in terms of their connectivity to each other when represented as a graph, $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$, where $\mathcal{V}$ is the set of N vertices and data $\mathbf{d} : \mathcal{V} \rightarrow \mathbb{R}$, is associated to the vertices of the graph. The connectivity among these vertices characterizes the global structure of the graph and it does not change after rotation, flipping or mirroring of the graph structure. Therefore this paper aims to explore modelling the connectivity of nodes in a graph structure to learn the dynamic shapes formed by the underlying graph.

Many modern computer vision applications utilise direct observation and analysis of dynamic shapes formed by movement of human hands or the body. The smooth flow and flexibility of the human hand configuration allows generation of very complex and arbitrary shapes drawn in the air. Automated analysis and understanding of these complex shapes benefit real-world applications like video games, human activity recognition, robotics, sign language recognition, and gesture-based human-computer interaction. Recently, in-air hand-drawn number recognition has gained strong interest [4]–[20] due to applications-related importance and the interesting challenges in the problem. These methods are either based on image representation of numbers and shapes [4]–[13] followed by shape matching or the path representation of hand movement [14]–[20] followed by shape matching. Some methods are based on deep learning for shape matching [9]–[13].

However, the current methods suffer from sensitivity to changes in the angle of drawing leading to serious recognition errors of the in-air drawn numbers and symbols. Since these methods use a pixel-based approach, the recognition speeds are found to be too slow for real-time applications. We are motivated by recent advances in psychophysical and

The associate editor coordinating the review of this manuscript and approving it for publication was Zhen Ren.

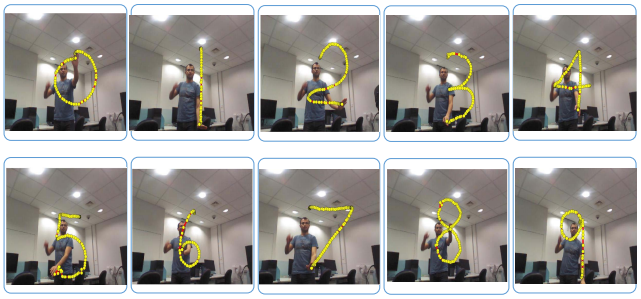

FIGURE 1. Left–right flipped screen shots of in-air drawing of numbers.

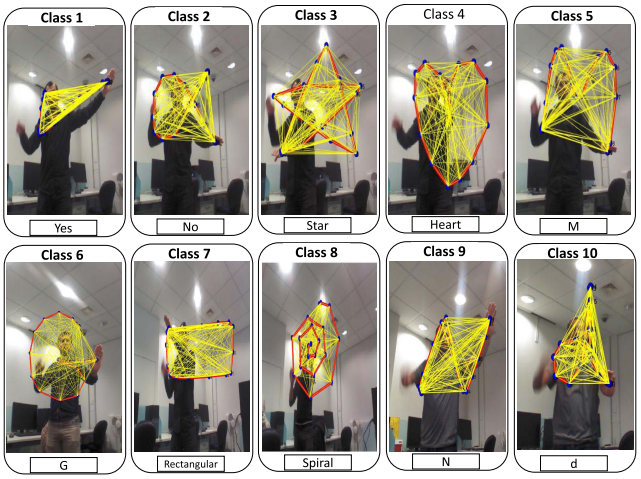

FIGURE 2. Ten classes of in-air drawn other shapes (symbols) with their graph construction.

neuro-physiological studies, which have proposed a hypothesis for a structural representation of shapes in terms of object structures, parts and their positional relationships [21]–[23]. Another study on human vision suggests that the visual cortex perceives and understands shapes by representing the shape boundary as a connected set of nodes [24]. These advances coupled with emerging graph signal processing concepts has led us to propose a novel graph spectral domain representation for shapes as a solution to overcome these issues in dynamic shape recognition in this paper.

This paper proposes novel graph spectral domain features for accurate and fast recognition of dynamic shapes. Since graph spectral representation, as shown in Section III, is driven by the connectivity, the learned features in the graph spectral domain are rotation and scale invariant. Therefore, the proposed method is not sensitive to angles of drawing of the shape. In converting the hand movement paths of the in-air drawn shapes into fully connected graphs, we aim to minimise the number of nodes while keeping the properties of the structure intact. This leads to lowering the complexity without affecting the recognition accuracy rates. The use of the graph matching on vertex domain has been explored for shape matching purposes in the literature [25]–[27]. However, they are not robust to variations in angles of orientation of the shapes and numbers. In our present work, we explore representing the shapes and numbers in the graph spectral domain as opposed to the node domain for feature extraction for recognition. In our proposed work, different shapes are represented based on the connectivity description through the graph spectral domain. A fixed number of features is extracted followed by machine learning for recognition of shapes. The main contributions of the proposed work are:

- Proposal of a new set of graph spectral domain features for dynamic shape recognition.
- Proposal of a new rotation and flip-invariant feature set with real-time operation.

The proposed method is experimented using Kinect sensor for data capturing and real time recognition (as in FIGURE 1 and FIGURE 2) as shown in Section IV.

The rest of the paper is organized as follows: Section II explores the most related work in the field of in-air drawn shape and hand gesture recognition. Section III presents the proposed graph spectral feature learning method for dynamic shape recognition. Section IV shows the results and discussions in terms of the performance of the proposed graph spectral features for an in-air hand drawn number and other shape (symbol) recognition applications followed by the concluding remarks in Section V.

II. RELATED WORK

The present paper proposes the graph spectral domain features for recognition of shapes, such as, in-air drawn numbers and symbols. In general, shape representation work in the literature can be categorized into five different groups: deep-learning methods, model-based methods, view-based methods, feature-based methods, and graph matching methods. In this section, we briefly present the recent work related to graph-based matching, graph-based static hand gesture recognition and related in-air handwritten number recognition work.

A. GRAPH MATCHING

Previous methods on graph matching are primarily based on either bipartite matching [26]–[28] or approximate matching [29]–[36]. They explore either the vertex domain properties or the properties of the graph adjacency matrix. In bipartite graph matching, a set of edges are chosen in such a way that no two edges share the same end point vertex to maximize the matching ratio between two sets of points without increasing the degree of the nodes. Different matching conditions, such as, the shortest edge [26], the convex path inside text images [27] and the largest eigenvalue [28] have been used in bipartite matching. In approximate matching, the maximum probability of correspondence mapping between two patterns through the weight matrix is explored based on various properties of the graph adjacency matrix, such as, polynomial characterization [29], center of clusters [30], eigenvalues [31], spectral relaxation [32], higher order constraints [33] and Kronecker product [34]. Graph matching in terms of Principal Component Analysis (PCA) has also been explored for shape matching [37].

The main limitations of graph matching approaches are the high computational complexity and restrictions on some graph sizes [31], [35]. The former limitation makes them unsuitable for real-time applications.

The GSP offers new opportunities for processing, compression and analysis of spatially non-uniformly sampled data, represented as a graph, by characterizing the global structure on its eigenvalues and eigenvectors of the graph Laplacian matrix [1]. The graph eigenvectors (*e.g.*, basis vectors as in a content adaptive transform) provide an efficient representation of the connectivity and the structure of the graph. The most highlighted basis vector is the eigenvector corresponding to the second smallest eigenvalue, which is known as the Fiedler value [38] or the algebraic connectivity [39]. The sign of Fiedler vector has been explored in analysis tasks, such as, determining the stability of system [40], saliency estimation [41] and image partitioning [42].

B. GRAPH-BASED STATIC HAND GESTURE (POSE) RECOGNITION

Some of the previous work on graph-based hand gesture recognition has been restricted to static hand pose recognition rather than dynamic hand movements [43]–[48]. They are also mainly based on graph theory rather than spectral domain features. A few of the methods include edge-based hand features [43]–[46], tree-based representation [47] and a combination of hand appearance and graph features and graph eigenvalues [48].

C. IN-AIR HANDWRITTEN NUMBER RECOGNITION

Recent work on in-air handwritten number recognition can be classified into two groups: image-based representation [4]–[13] and node-based representation [14]–[20]. In image-based representation, the numbers are saved as an image with two types of pixels: number path pixels and background pixels, as in static number recognition applications. Thus, a large amount of data is required to represent the numbers, which is identified as a limitation of this approach. In one approach, number images are normalized into a binary table where 1 and 0 referring to the hand writing path and the background, respectively [6]. In order to eliminate the effect of too many zeros in the table, order code with shape code have been used [4]. Later, further improvements have been achieved by normalizing the hand-path by picking out a specific number of points [5]. Deep learning-based techniques employ one dimensional convolutional neural networks to learn the features from image-based representations [9]–[13]. It must be also noted that some methods were concerned with recognizing words written on a touch pad using finger strokes [9]–[11], as opposed to in-air drawing as in robotic applications.

In node representation, the path of the hand movement is saved as a set of coordinates, which are captured by different ways, such as, tracking the hand node in Kinect skeleton representation [14]–[18], tracking the position and the orientation of the user's hand using a Wii remote [19] and using supervised learning to detect hand path [20]. These number recognition methodologies are sensitive to any slight change in the angle of drawing and that causes a serious mismatching of the samples. More recent works [17], [18] in this category consist of high computational complexity pre-processing steps and extracting a set of features based on distance and angles followed by classification. They are sensitive to rotational changes of the shapes and not suitable for real time operation. In order to maintain the invariance to rotation and scaling changes, Fourier descriptors have been used for some shape representation [49], [50]. Also, these low-level features are extracted from pixels, requiring a large amount of data to achieve the optimal level of accuracy. This has hindered their ability to operate in real-time.

In our present work, we propose novel graph spectral domain feature extraction and learning for accurate and fast shape recognition and demonstrate its use on in-air hand-drawn number and other shape recognition that can be useful for hand movements in human-machine (robotic) communication applications. GSP enables reducing the number of nodes for processing while maintaining the specification of the shape, as the features are extracted on spectral domain as opposed to the vertex domain. This leads to real-time operation of the recognition system. The Graph Laplacian matrix, which forms the Graph Fourier Transform's basis vectors, can characterize the global structure of the shape leading to high recognition accuracy rates. Our proposed work is the first approach that explores graph spectral feature learning for dynamic shape recognition and in-air hand path movement recognition. Early results of our work were presented as conference publication in [51]. The current paper extends this early work by proposing more efficient hand crafted graph spectral features, reporting comprehensive evaluation of performance, extending the evaluation into arbitrary shapes using commonly used datasets and a newly generated dataset. It must be noted that the proposed methodology can be applied on recognition of static numbers or image-based representation by performing preprocessing to convert the image-based representation into node (graph) based representation. As our primary motivation is to explore graph spectral representation for shape recognition, we have used in-air drawn numbers and symbols for evaluating the proposed methodology in this paper

III. THE PROPOSED GRAPH SPECTRAL DOMAIN SHAPE LEARNING

The proposed system (FIGURE 3) can be divided into three steps: pre-processing; graph spectral feature extraction; and classification. The main originality in the work presented in this paper is in graph spectral feature extraction and classification steps. Details of pre-processing are included in this section for completeness. However, it should be noted that the graph spectral feature extraction and classification methodology can be applied on any dynamic shape that can be represented as a set of nodes or a movement path.

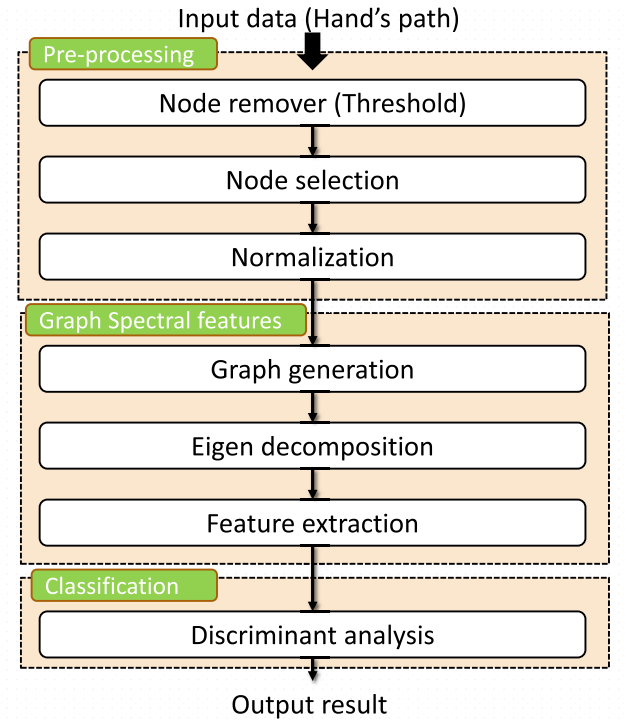

FIGURE 3. Flowchart of the proposed method.

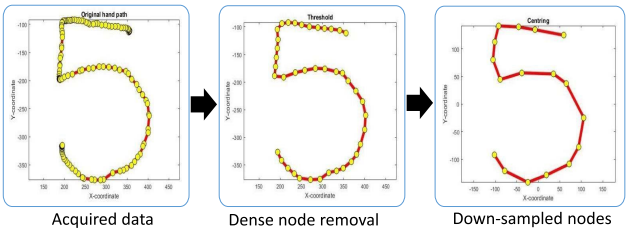

FIGURE 4. Data pre-processing steps: (Left) data acquired from the original hand path; (Middle) dense nodes removal; (Right) node down-sampling.

A. PRE-PROCESSING

The input data, represented as a set of node samples, may be densely sampled as can be seen in the left sub-figure in FIGURE 4. For each point, the (x, y) coordinates on the vertical plane and the depth z are recorded to form the 3D measurement space in (x, y, z). In order to remove these unwanted samples, starting from the first sample, for each sample, all the samples captured within a distance less than a chosen fixed threshold (τ) are removed. This process eliminates the extra nodes that are created when the user stops moving their hand at any point during drawing or at the end of drawing. The resulting in-air drawn path, P, is a smooth curve with T number of nodes as can be seen in the middle sub figure of FIGURE 4. To reduce the complexity of the subsequent graph spectral decompositions, we choose N number of nodes, where $N < T$, to form a new down-sampled path, $\hat{P}$, as follows (as in the right sub figure of FIGURE 4):

$$\hat{P}(k) = P\left(\left\{\frac{Tk}{N}\right\}\right), \tag{1}$$

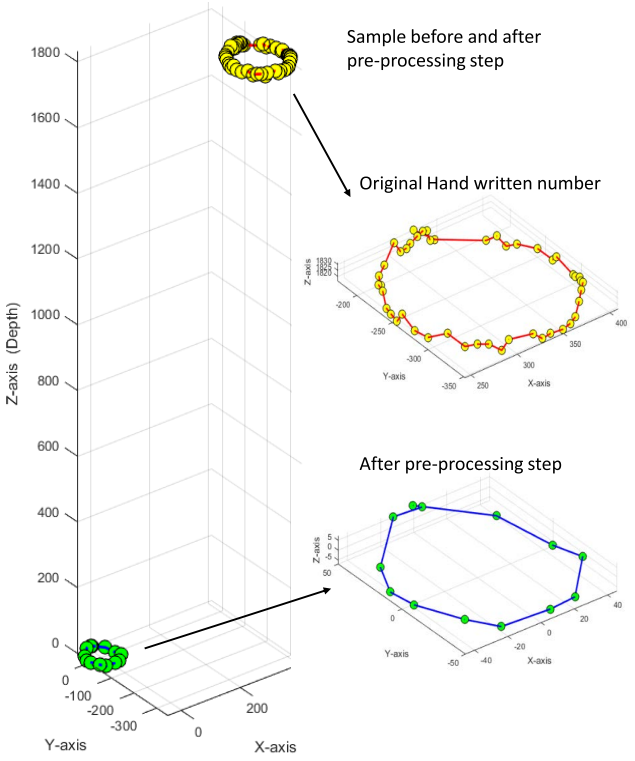

FIGURE 5. The hand moving path pre-processing stages: starts with the acquired nodes (in yellow); down-sampled nodes fitted into the bounding cuboid (in yellow and red); centered around (0,0,0) (in blue and green); mapped to $m \times m$ on the x-y plane (in green).

where $k = 0, 1, \ldots, N-1$ is the new node index and $\{\}$ is the rounding to the nearest integer. This is followed by normalizing the down-sampled path. Normalizing is achieved by considering the minimum and maximum bounds for x,y,z, coordinates to create a bounding cuboid for the shape followed by centering the shape around (0,0,0) by shifting all coordinates by shifting coordinates accordingly as shown in FIGURE 5. Then, x, and y coordinates are re-scaled to fit into $m \times m$ bounding box on the x-y plane. Finally, the node k represented with its coordinates (x_k, y_k, z_k).

B. PROPOSED GRAPH SPECTRAL FEATURES

We can now represent the nodes in the hand path, $\hat{P}$, as an undirected graph, $\mathcal{G} = \{\mathcal{V}, \mathcal{E}, \mathbf{A}\}$, where $\mathcal{V}$ is the set of N vertices (defined by the nodes in $\hat{P}$), $\mathcal{E}$ is the set of edges and $\mathbf{A}$ is the adjacency matrix with edge weights. We consider $\mathcal{G}$ as a fully connected graph, which means each vertex has $(N-1)$ connected edges. We have found this type of connectivity is suitable for representing shapes that have an outline without much local variation along its outline. This is the case for in-air drawn numbers and symbols. In our recent work [52], [53], we have shown that for more complex shapes with local variations along their outlines, graph connections have to be adaptively formed by considering some criteria leading to a complex pre-processing step. We define the weight, $\mathbf{A}_{i,j}$ corresponding to an edge, $e_{i,j}$ connecting vertices

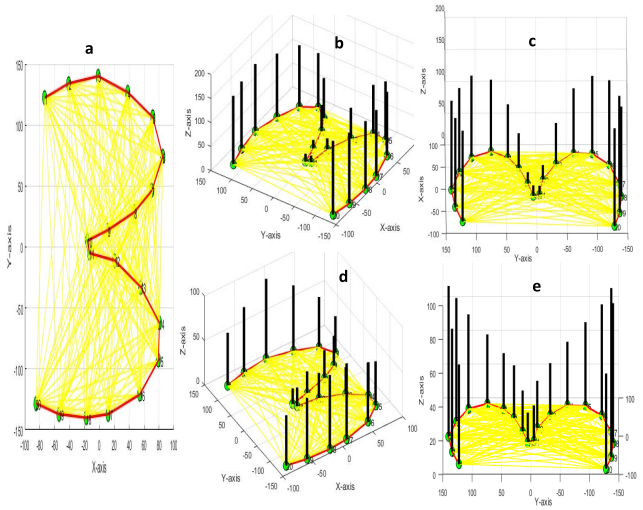

FIGURE 6. Graph construction above the hand path (red lines) with its vertices (green points), fully connected node edges, $\mathcal{E}$, (yellow lines) and the node values (black lines), where (a): Top view of the graph, (b): r values (Side view), (c): r values (45° view), (d): θ values (Side view), (e): θ values (45° view).

i and j is as follows:

$$\mathbf{A}_{i,j} = \frac{|e_{i,j}|}{\frac{1}{N}\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}|e_{(i,j)}|}, \tag{2}$$

which is the Euclidean distance $e_{(i,j)}$ between the vertices, i and j, normalized with the average edge length for a node.

We define the signal $\mathbf{r} : \mathcal{V} \rightarrow \mathbb{R}$, where i^{th} component represents the Euclidean distance from the centre (0,0,0) to the vertex i in $\mathcal{V}$ as follows:

$$\mathbf{r}_i = \sqrt{x_i^2 + y_i^2 + z_i^2}, \tag{3}$$

An example of a $\mathcal{G}$ and its signal $\mathbf{f}$ is shown in FIGURE 6. We also define the signal $\theta : \mathcal{V} \rightarrow \mathbb{R}$, where i^{th} component represents the angle of the vertex i in $\mathcal{V}$ with respect to the centre (0,0,0) as follows:

$$\theta_i = tan^{-1}\left(\frac{|y_i|}{|x_i|}\right). \tag{4}$$

The absolute value of x_i and y_i keeps the range of the angle between ($0° - 90°$). For example, the points (3,4), (-3,4), (3,-4) and (-3,-4) have the same angle value, which is equal to 53.13°. An example of a $\mathcal{G}$ and its signal θ is shown in FIGURE 6.

The combinatorial graph Laplacian matrix, $\mathbf{L}$, is defined as

$$\mathbf{L} = \mathbf{D} - \mathbf{A}, \tag{5}$$

where $\mathbf{D}$ is the diagonal matrix of vertex degrees, whose diagonal components are computed as follows:

$$\mathbf{D}_{(i,i)} = \sum_{j=0}^{N-1}\mathbf{A}_{(i,j)}, \qquad i = 0, 1, \ldots, N-1. \tag{6}$$

As the shapes form non-regular graphs, we consider the symmetric normalized Laplacian matrix, ($\mathcal{L}$), computed as follows:

$$\mathcal{L} = \mathbf{D}^{-\frac{1}{2}}\mathbf{L}\mathbf{D}^{-\frac{1}{2}}. \tag{7}$$

Although, $\mathbf{L}$ has been widely used in image processing related applications [2], [54], the symmetric normalized Laplacian, $\mathcal{L}$, is thought to be more appropriate for representing non-regular graphs in the literature [3], [41]. Therefore, in this work our primary focus is on the symmetric normalized Laplacian.

Since, the Laplacian matrices, $\mathbf{L}$ and $\mathcal{L}$, are symmetric positive semidefinite matrices, from spectral projection theorem, there exists a real unitary matrix, $\mathbf{U}$, that digonalizes $\mathcal{L}$, such that $\mathbf{U}^t\mathcal{L}\mathbf{U} = \Lambda = diag\{\lambda_\ell\}$ is a non-negative diagonal matrix [3], leading to an eigenvalue decomposition of $\mathcal{L}$ matrix as follows:

$$\mathcal{L} = \mathbf{U}^t\Lambda\mathbf{U} = \sum_{\ell=0}^{N-1}\lambda_\ell\mathbf{u}_\ell\mathbf{u}_\ell^t, \tag{8}$$

where $\mathbf{u}_\ell$, the column vectors of $\mathbf{U}$, are the set of orthonormal eigenvectors of $\mathcal{L}$ with corresponding eigenvalues, $0 = \lambda_0 < \lambda_1 \leq \lambda_2 \ldots \leq \lambda_{N-1} = \lambda_{max}$ [1]. These eigenvectors provide the best set of basis functions by considering the connectivity of vertices of the graph as opposed to the signal values measured on graph nodes as generally used in PCA for computing the best basis.

These eigenvectors have been used in analysing graph spectra both algebraic and analytic wise [55], [56]. It has been shown in [55], that given a graph with no isolated vertices and $\mathcal{L}\mathbf{u} = \lambda\mathbf{u}$, *i.e.*, $\mathbf{u}$ is an eigenvector of $\mathcal{L}$, then the corresponding harmonic eigenvector, $\mathbf{y}$, associated with the eigenvalue λ is $\mathbf{D}^{-\frac{1}{2}}\mathbf{u}$. Thus, given the harmonic eigenvector for $\lambda = 0$ defined as $\frac{1}{\sqrt{N}}\mathbf{1}$, where $\mathbf{1}$ is the constant 1 vector and N is the number of nodes, the first eigenvector, $\mathbf{u}_0$, is defined as

$$\mathbf{u}_0 = \mathbf{D}^{\frac{1}{2}}\frac{1}{\sqrt{N}}\mathbf{1}. \tag{9}$$

The remaining eigenvectors are orthogonal to $\mathbf{u}_0$, since $\mathcal{L}$ is symmetric. As can be seen from Eq. (6) and Eq. (9), the eigenvector $\mathbf{u}_0$ of $\mathcal{L}$, corresponding to the lowest eigenvalue, $\lambda_0 = 0$ captures important details of the global structure of the graph through the node degree. Because the graphs used in this paper are fully connected, the node degree represents the geometric location of the nodes in relation to the other nodes by measuring how close or far to the rest of the nodes. Therefore, in our proposed method, we explore the use of $\mathbf{u}_0$ as part of features for shape recognition. We propose a feature vector comprising of the following 3 components :

1) The first part of the feature vector addresses translation invariance of the shape analysis of $\mathbf{u}_0$ by computing the second moment components about the mean as follows:

$$\mathbf{F}_1 = \mathbf{M}_1.\mathbf{u}_0, \tag{10}$$

where modulation matrix, $\mathbf{M}_1$, is a diagonal matrix with diagonal elements computed as follows:

$$\mathbf{M}_{1_{(\alpha,\alpha)}} = (\alpha + 1 - (N+1)/2)^2 + 1, \quad (11)$$

where $\alpha = 0, 1, \ldots, N-1$.

2) The second part of the feature vector modulates $\mathbf{u}_0$ with the distance to each node with respect to the origin, (0,0,0), as follows:

$$\mathbf{F}_2 = \mathbf{M}_2.\mathbf{u}_0, \quad (12)$$

where the modulation matrix, $\mathbf{M}_2$, is a diagonal matrix with diagonal elements formed by the magnitude signal, $\mathbf{r}$, computed by Eq. (3).

3) The final part of the feature vector modulates $\mathbf{u}_0$ with the angle to each node with respect to the origin as follows:

$$\mathbf{F}_3 = \mathbf{M}_3.\mathbf{u}_0, \quad (13)$$

where the modulation matrix, $\mathbf{M}_3$, is a diagonal matrix with diagonal elements formed by the angle signal, θ, computed by Eq. (4).

Overall, the feature vector is formed by concatenating the three vectors, $\mathbf{F}_1$, $\mathbf{F}_2$ and $\mathbf{F}_3$, resulting in a feature vector with a total length of $3N$. The components of the feature vectors for numbers 0 to 9 are shown in FIGURE 7. Similarly, FIGURE 8 shows the feature vector components for the symbols included in our data set, which is shown in FIGURE 2.

C. CLASSIFICATION

For the multi-class classification problem and the length of the feature vectors proposed, Discriminant Analysis classifier is expected to work well. Several classifiers were tested as detailed in Section IV-B. The Quadratic Discriminant Analysis (QDA) function, and the Linear Discriminant Analysis (LDA) function are similar in terms of the function and classification rules except the way covariance matrix is computed separately for each class (*i.e.*, varying, not identical). As a result, QDA tends to fit the data better than LDA.

IV. PERFORMANCE EVALUATION

In this section, we evaluate the performance of the proposed graph spectral feature learning doe dynamic shape representation in the context of in-air hand-drawn number and other shape recognition. The experimental set up includes creation of a new dataset for both numbers and other shapes using in-air hand drawing by several users captured by a Kinect sensor as detailed in Section IV-A. The evaluation process includes evaluating the effect of different classifiers for the proposed features in order to find the optimal classifier, the effect of number of nodes, N, for graph formulation, the impact of the threshold value τ and the choice of eigenvectors (from both normalized and combinatorial Laplacian matrices) for feature vector formation. The performances in terms of average recognition rates, confusion matrices and execution times are reported.

A. EXPERIMENTAL SET UP AND THE DATASETS

Currently available datasets mostly consist of numbers recorded as images leading to image-based representation. The challenge in such cases is to identify the hand movement path, *i.e.*, the dynamic features of the shape and its starting and ending points. We evaluate the proposed system using three different existing datasets, which provide a sequence of the hand writing (*i.e.*, hand path movement sequence is provided as a vector): The dataset presented in [5] consists of 1000 training samples and 2300 testing samples, while the dataset presented in [20] consists of 100 samples per number (0-9) captured using a PrimeSense 3D camera. The dataset in [18] includes 100 samples per number (0-9) captured using SoftKinetic DepthSense DS325.

In order to evaluate our proposed method using a larger dataset, we have created a new in-air hand-drawn numbers and symbols dataset, details of which can be found in [57]. In creating the dataset, we have also developed a system that can acquire live hand path data streams of the in-air hand-drawn dynamic shapes using the Kinect sensor v1.0 (640 × 480 pixels). It acquires depth data based on skeleton tracking (As shown in FIGURE 9). Users have to stand in front of Kinect around 1 to 3 metres away [58]. The right hand is used to draw the shapes in-air. Then, users can raise their left hand higher than the shoulder to end the movement as shown in FIGURE 1. By doing so, users will have unlimited time to draw digits in-air. The depth values acquired using the Kinect skeleton tracking is not accurate when the hand movement is fast. Therefore, several methods have been proposed in the literature to get more accurate depth information, such as, 3D analysis [59], combining of colour and depth information [60] and considering the closest object to the camera in the scene [61]. In creating this dataset, hand contour searching is performed based on the assumption that the hand is the closest object to the camera in the scene. The search takes place in a block of 20×20 pixels. In this case, the block centre represents the right hand position. The left hand tracking is based on the Kinect skeleton tracking because this technique is fast and no depth information is used for the left hand operation.

The samples in our dataset provide the direct observation of sequence of hand movements (*i.e.*, starting and ending points coordinates). The data set can be divided into two parts: numbers FIGURE 1 and symbols FIGURE 2. The number sub-dataset includes 500 instances per each number 0 to 9, resulting in a total of 5000 samples of in-air hand-drawn numbers. Similarly, the symbol sub-dataset also includes 500 samples per each of 10 different symbols drawn in 3D space, resulting in a total of 5000 samples. Our numbers and symbols datasets provide X, Y, Z coordinates of the users writing hand movement. The dataset provides raw sampled hand path with original sampling rates and without any normalizing nor smoothing, resulting in various numbers of data points within a given hand path. Our approach for pre-processing was outlined in Section III-A.

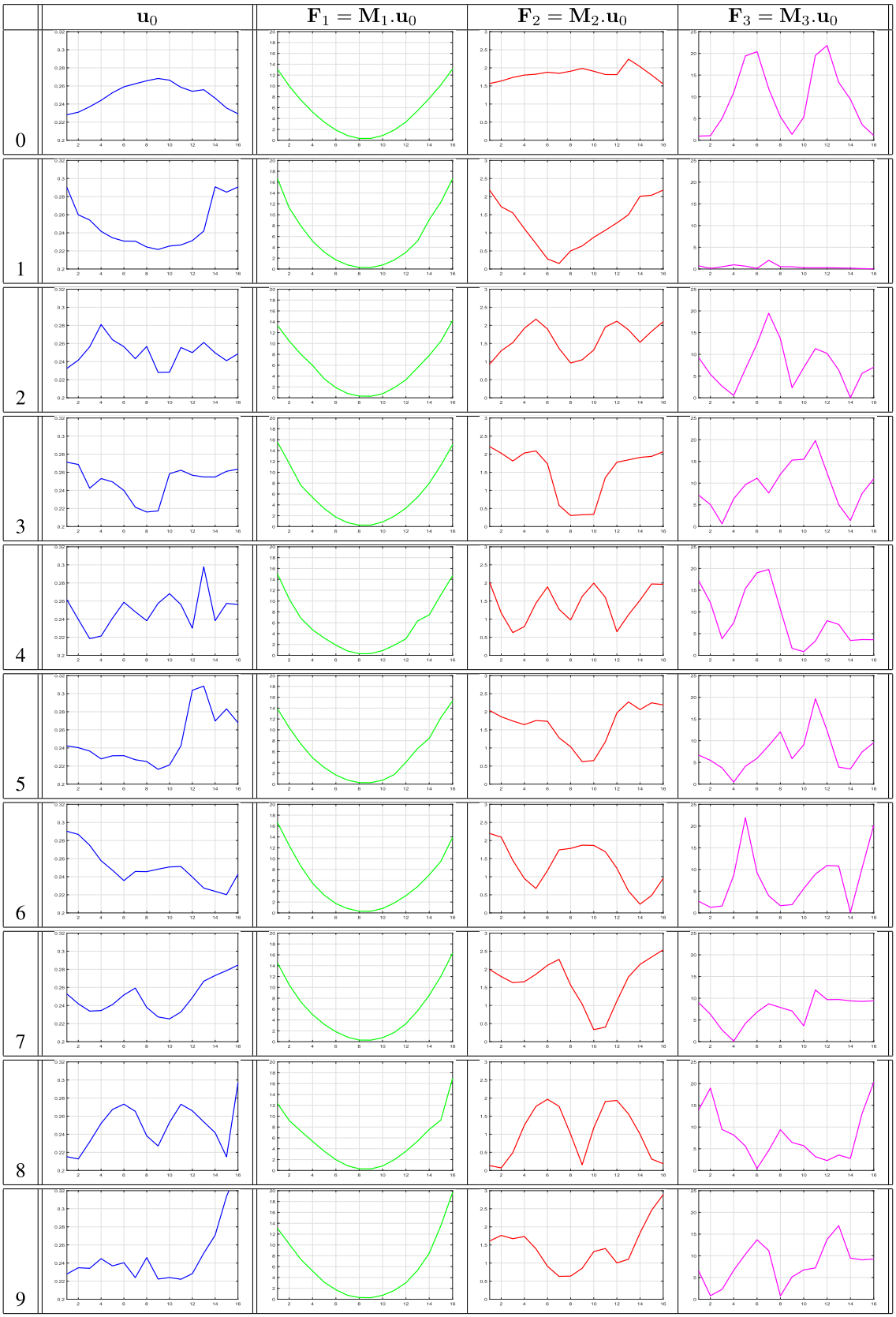

FIGURE 7. The proposed feature vectors for the numbers 0 to 9 (shown in each row): Column 1: Shows the eigenvector $\mathbf{u}_0$; Column 2 - Column 4 show the frature vector parts $\mathbf{F}_1$, $\mathbf{F}_2$ and $\mathbf{F}_3$, respectively.

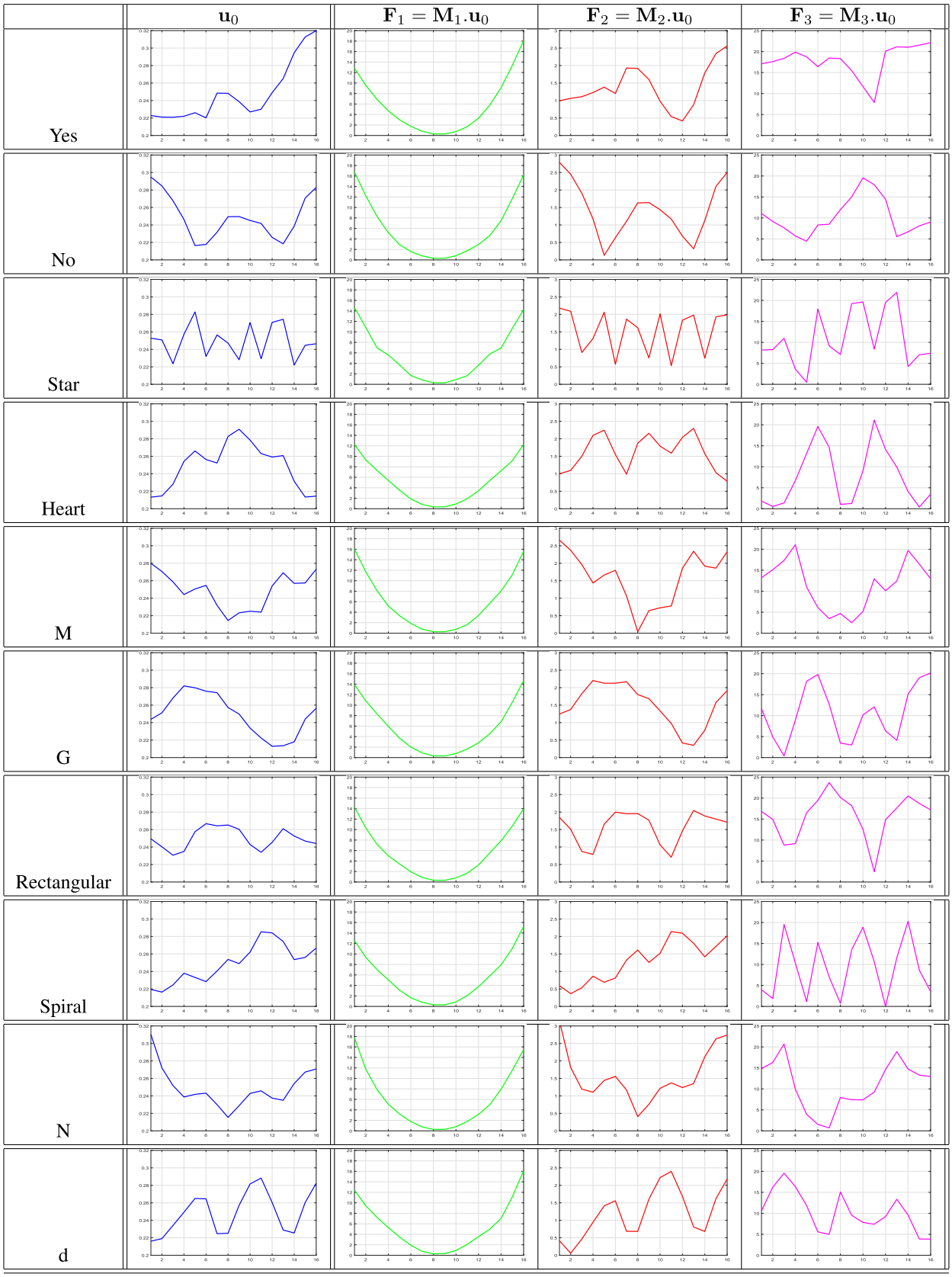

FIGURE 8. The proposed feature vectors for the symbols (shown in each row): Column 1: Shows the eigenvector $\mathbf{u}_0$; Column 2 - Column 4 show the feature vector parts $\mathbf{F}_1$, $\mathbf{F}_2$ and $\mathbf{F}_3$, respectively.

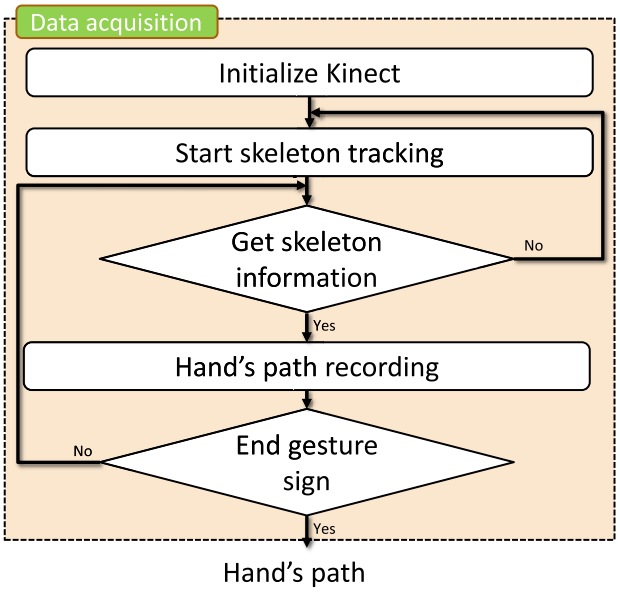

FIGURE 9. The data acquisition flow.

TABLE 1. Average recognition rates (%) for different classifiers.

Classifier	QDA	QSVM	CT	KNN
Mean accuracy (%)	99.5	97.8	98.6	99.18

B. EVALUATION OF DIFFERENT CLASSIFIERS

In order to test the entire dataset instead of random partitioning, a 5-fold cross-validation procedure is implemented to find the optimal classifier. TABLE 1 shows the mean accuracy for each classifier. Quadratic Support Vector Machine (QSVM) records relatively lower degree of accuracy than the other classifiers, whereas the Quadratic Discriminant Analysis (QDA) shows the highest level of recognition rate with more than 99%, while both classification tree (CT) with 30 learners using the bag-ensemble method, and Nearest Neighbour (KNN) with K=1, provide nearly 99% level of accuracy. Thus, QDA is used for the rest of the experiments.

C. EVALUATION OF DIFFERENT NUMBER OF GRAPH NODES (N) AND THE THRESHOLD VALUES (τ)

In this experiment, the pre-processing steps were repeated for various values of N. We start the test using $N = 3$, which is the minimum number of nodes to form the graphs. As can be seen in FIGURE 10, it is clear that too few number of nodes (*i.e.*, $N \leq 10$ nodes) is not suitable for accurately representing the samples. The level of accuracy is then relatively stable using $12 \leq N \leq 24$ nodes. Note that 24 nodes is the minimum length among all samples in our dataset.

Similarity, the pre-processing steps were also repeated for various values of τ. As can be seen in FIGURE 11, it is evident that the optimal range lies in 1 to 8 unit distance. Then the accuracy falls dramatically using $\tau \geq 12$.

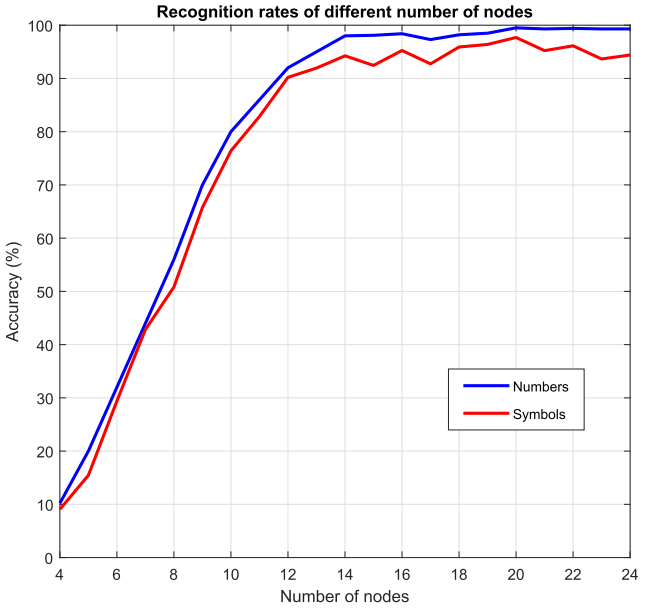

FIGURE 10. Recognition rates for different values of N at $\tau = 1$ using our dataset.

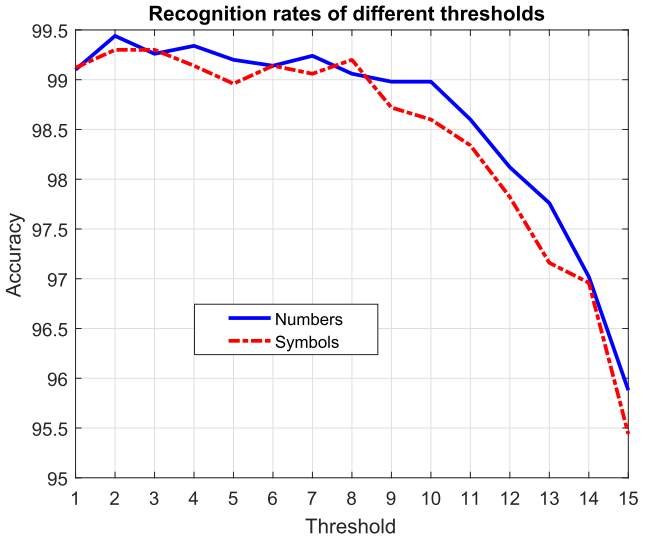

FIGURE 11. Recognition rates for different values of τ at $N = 17$.

D. EVALUATION OF DIFFERENT EIGENVECTOR OF L AND $\mathcal{L}$ AS THE FEATURE VECTOR

In this experiment, we evaluate the individual eigenvectors from $\mathbf{L}$ and $\mathcal{L}$, as the feature vector and show the average recognition rates in FIGURE 12. The eigenvector indexes correspond to the order when they are arranged according to the increasing eigenvalues. It can be seen that the $\mathbf{u}_0$ of $\mathcal{L}$ provides the highest recognition rate of 99.56%, as it captures details of the structure of the graph. The second best result is from the $\mathbf{u}_{N-1}$ of $\mathcal{L}$, which corresponds to the maximum eigenvalue (λ_{max}). For most eigenvectors, those from $\mathcal{L}$ appear to outperforming those from $\mathbf{L}$.

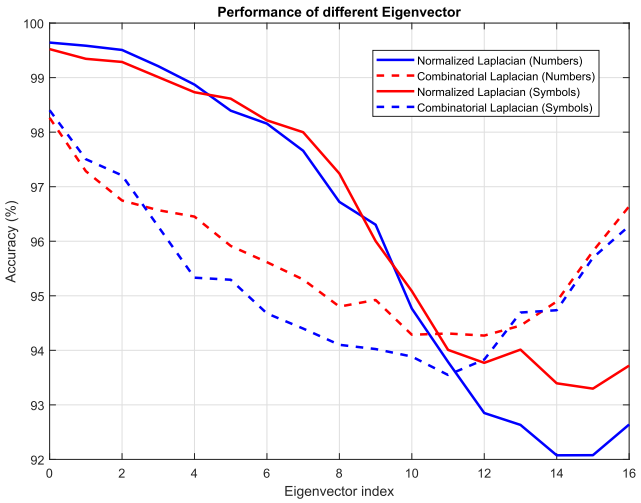

FIGURE 12. Accuracy rate of individual eigenvectors normalized $\mathcal{L}$ and combinatorial L graph Laplacian matrix.

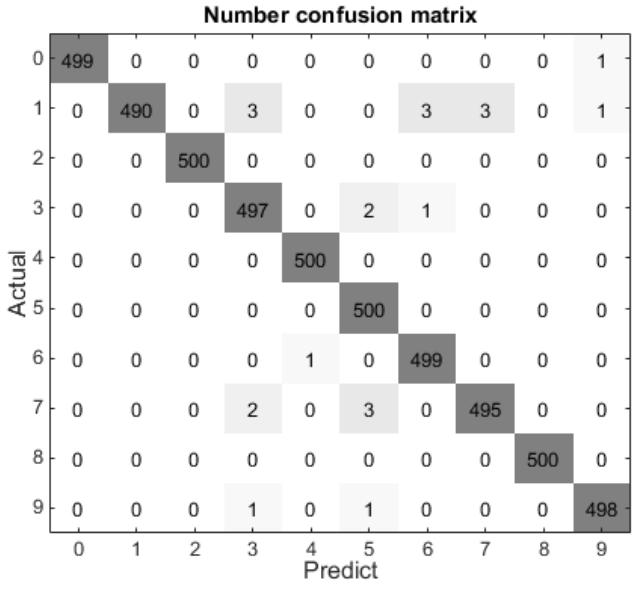

FIGURE 13. Confusion matrix for individual numbers using our dataset (overall average recognition rate is 99.56%).

E. PERFORMANCE OF THE PROPOSED METHOD

From the above experiments, we use ($N = 20$, $m = 15$, $\tau = 2$) for numbers and ($N = 20$, $m = 15$, $\tau = 3$) for symbols to construct the graph. This is followed by using $\mathbf{u}_0$ of $\mathcal{L}$, for generating the feature vector components $\mathbf{F}_1$, $\mathbf{F}_2$ and $\mathbf{F}_3$ and the QDA classifier to evaluate the performance of the proposed method using our dataset. The corresponding confusion matrices for numbers and symbols are shown in FIGURE 13 and FIGURE 14, respectively. The proposed method has achieved average recognition rates of 99.56% and 99.44% for numbers and symbols, respectively.

In order to compare the proposed method with the state of the art methods in the literature, we first evaluate our method for the existing datasets in [5], [18] and [20]. For the

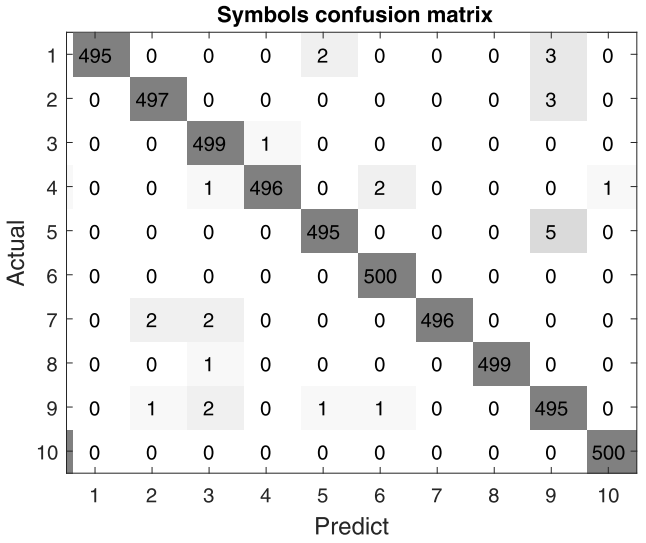

FIGURE 14. Confusion matrix for individual symbols shown in FIGURE 2 using our dataset (overall average recognition rate is 99.44%).

three datasets, the parameters, ($N = 9$, $m = 35$, $\tau = 3$), ($N = 12$, $m = 1$, $\tau = 0.03$) and ($N = 24$, $m = 10$, $\tau = 2$) were used respectively for forming graphs. The average recognition accuracy rates are shown in TABLE 2. The proposed method has achieved average accuracy values of 97.39% 99.1% and 99.2% for datasets in [5], [18] and [20], respectively. The proposed method has outperformed the corresponding existing methods reported in those works [5], [18] and [20], which reported accuracy rates of 96.82% 99% and 98.63%, respectively. The work in [18] does not report accuracy rates for individual number recognition. Hence a comparison column is not shown in TABLE 2. Only the overall average accuracy for the method [18] is reported. Similarly, the proposed method has outperformed all methods [4]–[6] reported for the dataset [5]. Overall, the proposed method has achieved the best results compared to the existing methods in TABLE 2.

All experiments were implemented using Matlab R2015b on a PC with Intel processor, CPU@3.6GHz and RAM 16GB. The time requirement of operating number and symbol recognition in a Kinect-based real-time system is about 4.127 ms per sample, which is suitable for a real-time in-air hand-drawn number and symbol recognition system.[1] The breakdown of the average times for each step of the algorithm is shown in TABLE 3. It also shows the computational complexity for each step. The total computational complexity is $\mathcal{O}(T^2 + 5N^2)$. Since $T >> N$, the total complexity can be expressed as $\mathcal{O}(T^2)$. If the pre-processing step is ignored and only the proposed graph spectral feature extraction is considered the overall complexity becomes $\mathcal{O}(N^2)$.

Since the graph adjacency matrix was defined based on the connectivity, it is insensitive to the rotation and flip changing.

[1] A recorded video of the real-time working recognition system available at https://youtu.be/cjY7UKTI-i4

TABLE 2. Average recognition accuracy rate (%) for sign numbers.

Method	[6]	[4]	[5]	**Proposed method**				[20]	[51]
Dataset	[5]	[5]	[5]	[5]	Ours	[18]	[20]	[20]	Ours
Zero	92.61	97.83	97.83	97.82	99.8	100	100	98.5	99.62
One	76.09	94.78	91.3	93.47	98	100	98	99.5	93.85
Two	86.96	95.65	96.09	99.56	100	99	100	98.8	96.15
Three	86.96	93.48	97.39	96.95	99.4	99	99	99	99.62
Four	91.74	96.09	97.39	99.56	100	99	98	98.7	100
Five	75.96	89.13	98.26	95.21	100	96	100	97.5	94.23
Six	86.96	95.65	97.83	98.26	99.8	99	99	99	97.69
Seven	91.3	94.35	97.83	96.52	99	100	99	99	95.38
Eight	87.83	93.91	95.22	99.56	100	100	100	97.9	100
Nine	89.57	95.22	99.13	96.95	99.6	99	99	98.6	98.46
Average	90.8	94.6	96.82	**97.39**	**99.56**	**99.1**	**99.2**	98.63	97.53

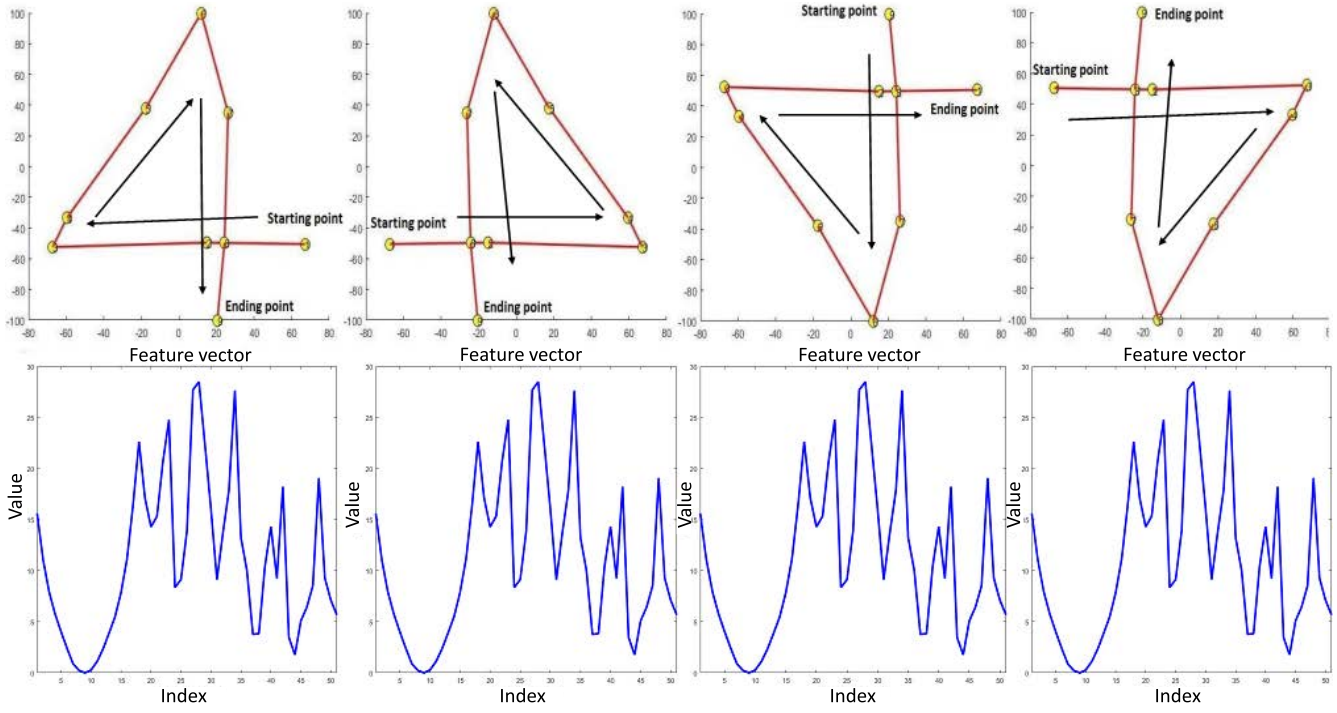

FIGURE 15. Hand drawing in different rotating angles (top row) with corresponding features (bottom row).

TABLE 3. Average time requirement to perform different steps in the proposed system.

Step	Average time performance (ms)	Computational complexity
Hand's path detection	2.729	$\mathcal{O}(T^2)$
Graph spectral decomposition	0.055	$\mathcal{O}(N^2)$
Feature extraction	0.195	$\mathcal{O}(3N^2)$
Classification	1.148	$\mathcal{O}(N^2)$
In total	4.127	$\mathcal{O}(T^2)$

In other words, it does not matter what the angle or direction of writing is as long as it follows the rules of starting and ending point. For example, all the cases shown in FIGURE 15 are detected as number four. Due to this reason, our proposed method is invariant to the rotation or flipping or orientation angle variation of the in-air drawn samples.

V. CONCLUSION

In this paper, we have presented novel graph spectral features for dynamic shape recognition and demonstrated its use on an in-air hand-drawn number and symbol recognition applications. The proposed method includes pre-processing for converting the input data, such as, hand path movement, into a fully connected graph followed by analysis of the eigenvectors of the normalized Laplacian of the graph adjacency matrix for forming the representative features. We have

utilised the eigenvector, $\mathbf{u}_0$, as it captures the details of the structure of the graph as the main representative feature. The proposed method has resulted in the highest performance with accuracies of 99.56% and 99.44%, for numbers and symbols, respectively, outperforming the existing methods for three different datasets. The proposed method also has added benefits of fast operation and invariance to rotation and flipping.

ACKNOWLEDGMENT

The authors like to thank Dr. C. Wang and Dr. S. Berman for providing the datasets of [5] and [20], respectively.

REFERENCES

[1] D. I. Shuman, S. K. Narang, P. Frossard, A. Ortega, and P. Vandergheynst, "The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains," *IEEE Signal Process. Mag.*, vol. 30, no. 3, pp. 83–98, May 2013.

[2] D. K. Hammond, P. Vandergheynst, and R. Gribonval, "Wavelets on graphs via spectral graph theory," *Appl. Comput. Harmon. Anal.*, vol. 30, no. 2, pp. 129–150, Mar. 2011.

[3] F. R. K. Chung, *Spectral Graph Theory*. Providence, RI, USA: American Mathematical Society, 1997.

[4] F.-A. Huang, C.-Y. Su, and T.-T. Chu, "Kinect-based mid-air handwritten digit recognition using multiple segments and scaled coding," in *Proc. Int. Symp. Intell. Signal Process. Commun. Syst. (ISPACS)*, Nov. 2013, pp. 694–697.

[5] C. Wang, C.-Y. Su, and C.-L. Lin, "A novel recognition system for digits writing in the air using coordinated path ordering," in *Proc. Int. Conf. Intell. Inform. Biomed. Sci. (ICIIBMS)*, Nov. 2015, pp. 244–249.

[6] T.-T. Chu and C.-Y. Su, "A Kinect-based handwritten digit recognition for TV remote controller," in *Proc. Int. Symp. Intell. Signal Process. Commun. Syst. (ISPACS)*, Nov. 2012, pp. 414–419.

[7] J. Singha and R. H. Laskar, "Self co-articulation detection and trajectory guided recognition for dynamic hand gestures," *IET Comput. Vis.*, vol. 10, no. 2, pp. 143–152, Mar. 2016.

[8] L. Kane and P. Khanna, "Vision-based mid-air unistroke character input using polar signatures," *IEEE Trans. Human-Mach. Syst.*, vol. 47, no. 6, pp. 1077–1088, Dec. 2017.

[9] X. Qu, W. Wang, K. Lu, and J. Zhou, "In-air handwritten Chinese character recognition with locality-sensitive sparse representation toward optimized prototype classifier," *Pattern Recognit.*, vol. 78, pp. 267–276, Jun. 2018.

[10] J. Gan and W. Wang, "In-air handwritten English word recognition using attention recurrent translator," *Neural Comput. Appl.*, vol. 31, no. 7, pp. 3155–3172, 2019.

[11] J. Gan, W. Wang, and K. Lu, "A new perspective: Recognizing online handwritten Chinese characters via 1-dimensional CNN," *Inf. Sci.*, vol. 478, pp. 375–390, Apr. 2019.

[12] R. Aggarwal, S. Swetha, A. M. Namboodiri, J. Sivaswamy, and C. V. Jawahar, "Online handwriting recognition using depth sensors," in *Proc. Int. Conf. Document Anal. Recognit. (ICDAR)*, Aug. 2015, pp. 1061–1065.

[13] S. Mukherjee, S. A. Ahmed, D. P. Dogra, S. Kar, and P. P. Roy, "Fingertip detection and tracking for recognition of air-writing in videos," *Expert Syst. Appl.*, vol. 136, pp. 217–229, Dec. 2019.

[14] J.-S. Sheu and Y.-L. Huang, "Implementation of an interactive TV interface via gesture and handwritten numeral recognition," *Multimedia Tools Appl.*, vol. 75, no. 16, pp. 9685–9706, 2016.

[15] T. Murata and J. Shin, "Hand gesture and character recognition based on Kinect sensor," *Int. J. Distrib. Sensor Netw.*, vol. 10, no. 7, Jul. 2014, Art. no. 278460.

[16] O. Mendels, H. Stern, and S. Berman, "User identification for home entertainment based on free-air hand motion signatures," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 44, no. 11, pp. 1461–1473, Nov. 2014.

[17] W. H. Khoh, Y. H. Pang, and A. B. J. Teoh, "In-air hand gesture signature recognition system based on 3-dimensional imagery," *Multimedia Tools Appl.*, vol. 78, no. 6, pp. 6913–6937, 2019.

[18] J. Huang, P. Jaiswal, and R. Rai, "Gesture-based system for next generation natural and intuitive interfaces," *AI EDAM*, vol. 33, no. 1, pp. 54–68, 2019.

[19] M. Chen, G. AlRegib, and B.-H. Juang, "Air-writing recognition—Part I: Modeling and recognition of characters, words, and connecting motions," *IEEE Trans. Human-Mach. Syst.*, vol. 46, no. 3, pp. 403–413, Jun. 2016.

[20] D. Frolova, H. Stern, and S. Berman, "Most probable longest common subsequence for recognition of gesture character input," *IEEE Trans. Cybern.*, vol. 43, no. 3, pp. 871–880, Jun. 2013.

[21] C. E. Connor, S. L. Brincat, and A. Pasupathy, "Transformation of shape information in the ventral pathway," *Current Opinion Neurobiol.*, vol. 17, no. 2, pp. 140–147, 2007.

[22] S. L. Brincat and C. E. Connor, "Underlying principles of visual shape selectivity in posterior inferotemporal cortex," *Nature Neurosci.*, vol. 7, no. 8, pp. 880–886, 2004.

[23] I. Reppa and E. C. Leek, "Surface diagnosticity predicts the high-level representation of regular and irregular object shape in human vision," *Attention, Perception, Psychophys.*, vol. 81, no. 5, pp. 1589–1608, 2019.

[24] S. O. Murray, B. A. Olshausen, and D. L. Woods, "Processing shape, motion and three-dimensional shape-from-motion in the human cortex," *Cerebral Cortex*, vol. 13, no. 5, pp. 508–516, 2003.

[25] D. Fernández-Mota, J. Lladós, and A. Fornés, "A graph-based approach for segmenting touching lines in historical handwritten documents," *Int. J. Document Anal. Recognit.*, vol. 17, no. 3, pp. 293–312, 2014.

[26] A. Shamaie and A. Sutherland, "Graph-based matching of occluded hand gestures," in *Proc. 30th Appl. Imag. Pattern Recognit. Workshop*, Oct. 2001, pp. 67–73.

[27] P. Riba, J. Lladãs, and A. Fornés, "Handwritten word spotting by inexact matching of grapheme graphs," in *Proc. 13th Int. Conf. Document Anal. Recognit. (ICDAR)*, Aug. 2015, pp. 781–785.

[28] M. Leordeanu and M. Hebert, "A spectral technique for correspondence problems using pairwise constraints," in *Proc. IEEE 10th Int. Conf. Comput. Vis. (ICCV)*, vol. 1, Oct. 2005, pp. 1482–1489.

[29] R. Horaud and H. Sossa, "Polyhedral object recognition by indexing," *Pattern Recognit.*, vol. 28, no. 12, pp. 1855–1870, 1995.

[30] M. Carcassoni and E. R. Hancock, "Alignment using spectral clusters," in *Proc. Brit. Mach. Vis. Conf. (BMVC)*, 2002, pp. 1–10.

[31] S. Umeyama, "An eigendecomposition approach to weighted graph matching problems," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 10, no. 5, pp. 695–703, Sep. 1988.

[32] T. Cour, P. Srinivasan, and J. Shi, "Balanced graph matching," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 19, 2007, pp. 313–320.

[33] O. Duchenne, F. Bach, I. S. Kweon, and J. Ponce, "A tensor-based algorithm for high-order graph matching," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 12, pp. 2383–2395, Dec. 2011.

[34] F. Zhou and F. De la Torre, "Factorized graph matching," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 9, pp. 1774–1789, Sep. 2016.

[35] L. S. Shapiro and J. M. Brady, "Feature-based correspondence: An eigenvector approach," *Image Vis. Comput.*, vol. 10, no. 5, pp. 283–288, 1992.

[36] B. W. Miners, O. A. Basir, and M. S. Kamel, "Understanding hand gestures using approximate graph matching," *IEEE Trans. Syst., Man, Cybern. A, Syst. Humans*, vol. 35, no. 2, pp. 239–248, Mar. 2005.

[37] X. Zhu, X. Li, S. Zhang, Z. Xu, L. Yu, and C. Wang, "Graph PCA hashing for similarity search," *IEEE Trans. Multimedia*, vol. 19, no. 9, pp. 2033–2044, Sep. 2017.

[38] M. Fiedler, "A property of eigenvectors of nonnegative symmetric matrices and its application to graph theory," *Czech. Math. J.*, vol. 25, no. 4, pp. 619–633, 1975.

[39] M. Saerens, F. Fouss, L. Yen, and P. Dupont, "The principal components analysis of a graph, and its relationships to spectral clustering," in *Proc. Eur. Conf. Mach. Learn.* Berlin, Germany: Springer, 2004, pp. 371–383.

[40] Y. Kim and M. Mesbahi, "On maximizing the second smallest eigenvalue of a state-dependent graph Laplacian," *IEEE Trans. Autom. Control*, vol. 51, no. 1, pp. 116–120, Jan. 2006.

[41] F. Perazzi, O. Sorkine-Hornung, and A. Sorkine-Hornung, "Efficient salient foreground detection for images and video using Fiedler vectors," in *Proc. WICED*, 2015, pp. 21–29.

[42] J. Shi and J. Malik, "Normalized cuts and image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 8, pp. 888–905, Aug. 2000.

[43] J. Triesch and C. von der Malsburg, "Robust classification of hand postures against complex backgrounds," in *Proc. 2nd Int. Conf. Autom. Face Gesture Recognit.*, Oct. 1996, pp. 170–175.

[44] P. P. Kumar, P. Vadakkepat, and L. A. Poh, "Graph matching based hand posture recognition using neuro-biologically inspired features," in *Proc. 11th Int. Conf. Control Autom. Robot. Vis. (ICARCV)*, Dec. 2010, pp. 1151–1156.

[45] Y.-T. Li and J. P. Wachs, "Recognizing hand gestures using the weighted elastic graph matching (WEGM) method," *Image Vis. Comput.*, vol. 31, no. 9, pp. 649–657, 2013.
[46] Y.-T. Li and J. P. Wachs, "HEGM: A hierarchical elastic graph matching for hand gesture recognition," *Pattern Recognit.*, vol. 47, no. 1, pp. 80–88, 2014.
[47] H. Hamer, K. Schindler, E. Koller-Meier, and L. Van Gool, "Tracking a hand manipulating an object," in *Proc. 12th Int. Conf. Comput. Vis.*, Sep./Oct. 2009, pp. 1475–1482.
[48] M. Avraam, "Static gesture recognition combining graph and appearance features," *Int. J. Adv. Res. Artif. Intell.*, vol. 3, no. 2, pp. 1–4, 2014.
[49] T. Ai, X. Cheng, P. Liu, and M. Yang, "A shape analysis and template matching of building features by the Fourier transform method," *Comput., Environ. Urban Syst.*, vol. 41, pp. 219–233, Sep. 2013.
[50] Z. Xu, Y. Xie, Z. Chen, and L. Wu, "Shape similarity measurement model for holed polygons based on position graphs and Fourier descriptors," *Int. J. Geographical Inf. Sci.*, vol. 31, no. 2, pp. 253–279, 2017.
[51] B. Alwaely and C. Abhayaratne, "Graph spectral domain feature representation for in-air drawn number recognition," in *Proc. 25th Eur. Signal Process. Conf. (EUSIPCO)*, Aug./Sep. 2017, pp. 370–374. [Online]. Available: https://youtu.be/cjY7UKTI-i4
[52] B. Alwaely and C. Abhayaratne, "Graph spectral domain shape representation," in *Proc. Eur. Signal Process. Conf. (EUSIPCO)*, Sep. 2018, pp. 598–602.
[53] B. Alwaely and C. Abhayaratne, "Adaptive graph formulation for 3D shape representation," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2019, pp. 1947–1951.
[54] S. Zhong, Y. Hu, and J. Lu, "A new geometric-transformation robust and practical embedding scheme for watermarking 2D vector maps in the graph spectral domain," in *Proc. Int. Conf. Commun., Circuits Syst.*, vol. 1, Jun. 2006, pp. 24–30.
[55] S. Butler, "Algebraic aspects of the normalized Laplacian," in *Recent Trends in Combinatorics*. Cham, Switzerland: Springer, 2016, pp. 295–315.
[56] U. von Luxburg, "A tutorial on spectral clustering," *Statist. Comput.*, vol. 17, no. 4, pp. 395–416, 2007.
[57] B. Alwaely and C. Abhayaratne. (Nov. 2018). *In-Air Hand-Drawn Numbers and Shapes Dataset*. [Online]. Available: https://figshare.shef.ac.uk/articles/In-Air_Hand-Drawn_Number_and_Shape_Dataset/7381472/1
[58] K. Khoshelham, "Accuracy analysis of Kinect depth data," in *Proc. Laser Scanning Workshop ISPRS*, 2011, vol. 38, no. 5, p. W12.
[59] M. Asad and C. Abhayaratne, "Kinect depth stream pre-processing for hand gesture recognition," in *Proc. 20th IEEE Int. Conf. Image Process.*, Sep. 2013, pp. 3735–3739.
[60] Z. Ren, J. Yuan, J. Meng, and Z. Zhang, "Robust part-based hand gesture recognition using Kinect sensor," *IEEE Trans. Multimedia*, vol. 15, no. 5, pp. 1110–1120, Aug. 2013.
[61] G. Plouffe and A.-M. Cretu, "Static and dynamic hand gesture recognition in depth data using dynamic time warping," *IEEE Trans. Instrum. Meas.*, vol. 65, no. 2, pp. 305–316, Feb. 2016.

BASHEER ALWAELY received the B.E. and M.Sc degrees from Baghdad University, Baghdad, Iraq, in 2007 and 2012, respectively, all in mechatronic engineering. Since September 2015, he has been a Ph.D. student advised by Dr. Charith Abhayaratne with The University of Sheffield, Sheffield, U.K. His research interests include the field of hand gesture recognition, shape representation, human–computer interaction, computer vision, image processing, signal processing, and machine learning. He is a member of the European Association for Signal Processing (EURASIP).

CHARITH ABHAYARATNE (M'98) received the B.E. degree in electrical and electronic engineering from The University of Adelaide, Australia, in 1998, and the Ph.D. degree in electronic and electrical engineering from the University of Bath, U.K., in 2002. He is currently a Lecturer with the Department of Electronic and Electrical Engineering, The University of Sheffield, U.K. He has contributed to the ISO/IEC Moving Picture Expert Group on scalable video coding standardization and wavelet video exploration, from 2004 to 2006. He has published more than 80 peer-reviewed articles in leading journals, conferences, and book editions. His research interests include multidimensional signal processing, image and video compression, visual content understanding, multimedia content security, and forensics. He is a committee member of the British Standards Institution (BSI) Group IST/037 coding of picture, audio, multimedia, and hypermedia information. He was a recipient of the European Research Consortium for Informatics and Mathematics (ERCIM) Postdoctoral Fellowship, from 2002 to 2004, to carry out research with the Centre of Mathematics and Computer Science (CWI), The Netherlands, and the National Research Institute for Computer Science and Control (INRIA), Sophia Antipolis, France. He is the U.K. Liaison Officer for the European Association for Signal Processing (EURASIP). He is an Expert Reviewer for the U.K. Engineering and Physical Science Research Council (EPSRC), European Commission Framework Program, Hong Kong Research Council, leading book publishers and leading the IEEE/IET/ACM/SPIE journals and conferences. He is also an Associate Editor of the IEEE TRANSACTIONS ON IMAGE PROCESSING and Elsevier *Journal of Information Security and Applications* (JISA). He has been in technical program committees for several leading international conferences and served as the Guest Editor of journal special issues.