ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) | 978-1-7281-7605-5/20/$31.00 ©2021 IEEE | DOI: 10.1109/ICASSP39728.2021.9414144

SHAPELET BASED VISUAL ASSESSMENT OF CLUSTER TENDENCY IN ANALYZING COMPLEX UPPER LIMB MOTION

Shreyasi Datta[1]*, Chandan Karmakar*[2]*, Punit Rathore*[3] *and Marimuthu Palaniswami*[1]

[1]Department of Electrical and Electronic Engineering, University of Melbourne, Australia
[2]School of Information Technology, Deakin University, Melbourne, Australia
[3]Senseable City Lab, Massachusetts Institute of Technology, Cambridge, USA

ABSTRACT

The evolution of ubiquitous sensors has led to the generation of copious amounts of waveform data. Human motion waveform analysis has found significance in clinical and home-based activity monitoring. Exploration of cluster structure in such waveform data prior to developing learning models is an important pattern recognition problem. A prominent category of algorithms in this direction, known as *Visual Assessment of (cluster) Tendency* (VAT), employs visual approaches to study cluster evolution through heat maps. This paper proposes *shape-iVAT*, a new relative of an improved VAT model, that captures *local* time-series characteristics through representative subsequences, known as *shapelets*, to identify interesting patterns in motion data. We propose an unsupervised method for shapelet extraction using maximin shape sampling and shape-based distance computation for selecting key shapelets representing characteristic motion patterns. These shapelets are used to transform waveform data into a dissimilarity matrix for VAT evaluation. We demonstrate that the proposed method outperforms standard VAT with global distance measures for identifying complex upper limb motion captured using a camera-based motion sensing device. We also show that our method has significance in efficient and interpretable cluster tendency assessment for anomaly detection and continuous motion monitoring.

***Index Terms*—** Cluster tendency, distance measures, shapelet, time-series, upper limb motion.

1. INTRODUCTION

Widespread realization of the Internet-of-Things technology, powered by smart sensors in body-worn, mobile and static devices generates huge amounts of human-centric waveform data [1, 2]. Analysis of human motion data has received particular attention for the development of smart health and activity monitoring applications [2, 3]. Automated knowledge extraction from such data is essential for timely detection of interesting events and continuous monitoring [4]. Clustering is often used to extract information from such evolving data for event and anomaly detection [1]. Consider a set of N objects $O = \{o_1, o_2, ..., o_N\}$, that can be partitioned into $c \in \{1, ..., N\}$ subsets, where each object is defined by a p-dimensional feature vector, $\mathbf{x_i} \in \mathbb{R}^p$ forming the dataset $X = \{\mathbf{x_1}, ..., \mathbf{x_N}\}$. This dataset may be presented in the form of a dissimilarity matrix $D_N = [d_{ij}]$, where d_{ij} represents dissimilarity between o_i and o_j. Before applying a clustering algorithm to this dataset, it must be decided whether it contains meaningful clusters or *non-random structures* through the assessment of cluster tendency [1, 5]. This is an important issue with unsupervised learning, because, clustering methods will estimate arbitrary c ($1 \leq c \leq N$) clusters, even when there are no apparent clusters, to satisfy the constraints of the algorithm. This problem can be addressed by a simple and intuitive visual approach called the *Visual Assessment of (cluster) Tendency* (VAT) [5, 6].

VAT uses a variant of Prim's algorithm to perform reordering of the input pairwise dissimilarity matrix such that, the reordered dissimilarity matrix, when viewed as a monochrome image or *cluster heat map*, shows possible clusters in the dataset by dark blocks along the diagonal [6]. This technique has been extended to develop an improved VAT (iVAT) to obtain sharper heat maps for complex clusters by combining a path based distance transform with VAT [7]. To use VAT/iVAT for waveform datasets, the elements of the dissimilarity matrix D_N would comprise a pairwise similarity or *distance measure* between each time-series [8]. Point-to-point distance measures, such as Lp norms, are unable to handle local time and amplitude shifts in time-series [9]. Though Dynamic Time Warping (DTW) overcomes this problem, it is computationally expensive, especially with increasing data length [10]. *Shape* based distance measures that use time-series correlation, can avoid this limitation associated with DTW [10]. Recently, VAT has been evaluated using a variety of distance measures including Lp norms, DTW, and shape-based distance to cluster waveform datasets [11, 12].

However, these realizations of VAT/iVAT for waveform datasets employ *global* distance measures, that fail to capture *local* characteristics. Local templates and shape characteristics that recur in human motion time-series can be representative of upper limb activities [3, 13], motor impairments [14] as well as seizures [15]. Instead of using distances between entire time-series, distance to such patterns within the series can be used as local features for cluster exploration in such cases. This can be realized by *shapelets*, representative subsequences of a time-series, that offer domain based interpretability [16, 17]. Use of shapelets for waveform data mining includes generation of the set of all candidate shapelets, evaluation of their quality in representing each class of data and computing the similarity of each time-series with the shapelets to assign it to a particular class. Usually, *information gain* on class values for each possible split of the dataset into two parts, one containing and the other devoid of the shapelet, is used as the quality measure [16]. In [18], *Shapelet Transform* was introduced, to extract k best shapelets, and use them to transform a dataset by calculating the distances of each time-series to each shapelet, creating a feature space for time-series classification [19]. Further, a mathematical formalization of shapelet discovery has been proposed as an optimization problem to *learn* near-to-optimal shapelets instead of exhaustive search for the best shapelets [20].

In this paper, we propose *shape-iVAT*, shapelet based visual assessment of cluster tendency, to identify movement patterns during complex upper limb motion. Unsupervised cluster tendency assessment through VAT using shapelets is challenging because, shapelet extraction is inherently a supervised approach that requires class information to evaluate shapelet quality. Though an unsupervised method for shapelet extraction was proposed in [21], it is unsuitable for exploration of cluster tendency as it uses cluster validation indices (after clustering) during shapelet selection. Given the minimum and maximum length of shapelets to search

for, our proposed method automatically extracts multiple variable-length shapelets to represent motion patterns for exploring cluster tendency (without actually clustering the dataset). We first use shape-based distance [10] and *maximin* sampling [22] to extract *shape-distinct* candidate subsequences representing different classes of motion. Then, we evaluate their quality based on their ability to maximize a *gap-score* [21] to segment the entire dataset without the requirement of class labels. Selected shapelets are then used to transform the dataset into a feature space comprising distances indicating the likelihood of their occurrence in each motion time-series, which is used to construct the dissimilarity matrix for shape-iVAT. Using a publicly available dataset [13] comprising upper limb trajectories during complex 2D and 3D motion, we show that shape-iVAT can identify interesting motion patterns as well as anomalous movements through cluster heat maps.

2. METHODS

In this section, we first introduce some definitions and then explain our proposed unsupervised shapelet extraction method and shape-iVAT algorithm for waveform datasets.

Definition 2.1. A waveform dataset $W^{N\times n}$ comprises N time-series, each of length n, represented by an ordered set of real-valued observations $T \equiv (t_i, t_{i+1}, ..., t_n)$, where t_i denotes the i^{th} observation. Each time-series is z-normalized with respect to its mean and standard deviation before further processing.

Definition 2.2. Given a time-series T of length n, a time-series subsequence or shapelet S of length $l(\leq n)$ is a contiguous set of samples $S \equiv (t_p, t_{p+1}, ..., t_{p+l-1})$, for $1 \leq p \leq n–l + 1$.

Definition 2.3. The distance between a shapelet S of length l and time-series T is the minimum distance between S and each subsequence $T^j \equiv (t_j, ..., t_{j+l-1})$ of length l of T, given by (1). Here, *dist(.)* denotes the distance measure, for which we can use Lp norms or a shape-based distance.

$$d_{T,S} = \min_{j=1,...,n-l+1} dist(T^j, S) \tag{1}$$

Definition 2.4. The dataset $W \in \mathbb{R}^{N\times n}$ can be transformed into a new representation $W_S \in \mathbb{R}^{N\times k}$ using distances of each time-series to k best shapelets representing the dataset. This transformation reduces the dimensionality of the original dataset (as typically $k \leq n$) while retaining phase-independent shape characteristics of each series [18].

2.1. Unsupervised Shapelet Extraction

Finding representative shapelets in a dataset involves generating an exhaustive set of candidate subsequences and selecting those that best partition the dataset. Suppose, all subsequences of length l for the i^{th} time-series be $S_{i,l}$ and the set of all subsequences of length l for dataset $W^{N\times n}$ be $S_l = \{S_{1,l}, S_{2,l}, ..., S_{N,l}\}$. Then, the set of all candidate subsequences in this dataset will be $S = \{S_{L_{min}}, S_{L_{min}+1}, ..., S_{L_{max}}\}$, where $L_{min} \geq 3$ and $L_{max} \leq n$. This set of candidate subsequences is very large and needs pruning before further processing. In [18], self similar shapelets, defined as those which come from the same time-series and have overlapping indices, are removed from analysis. But this might not be completely justified, as multiple shapelets with overlapping indices can be representative of a particular class of data. We propose Maximin Shape Sampling (MMSS) to prune S and select distinct subsequences with respect to their global shapes. Then we evaluate the quality of each candidate in this pruned subset using an unsupervised approach to extract the best shapelets. These steps are explained in the following subsections along with the shape-based distance measure used.

2.1.1. Shape-based Distance

For two time-series X and Y, each of length n, *Shape-based Distance* (SBD) is computed from their normalized cross-correlation [10], given by the inner product for each time-shift of X over a static Y. Considering all possible shifts, the cross-correlation sequence $CC_w(X, Y) = (c_1, c_2, ..., c_w)$ of length $2n$-1 is produced, given by $CC_w(X, Y) = R_{w-n}(X, Y), w \in 1, 2, ..., 2n - 1$ where $R_{w-n}(X, Y)$ is computed by (2). SBD is computed from the optimal shift w at which $CC_w(X, Y)$ is maximized by (3).

$$R_k(X, Y) = \begin{cases} \sum_{l=1}^{n-k} x_{l+k}.y_l, & k \geq 0 \\ R_{-k}(Y, X), & k < 0 \end{cases} \tag{2}$$

$$SBD(X, Y) = 1 - \max_w \left(\frac{CC_w(X, Y)}{\sqrt{R_0(X, X)R_0(Y, Y)}}\right) \tag{3}$$

Computation of $CC_w(X, Y)$ for all values of w requires $\mathcal{O}(n^2)$ time. Cross-correlation can also be computed as the convolution of two time-series (represented by $*$), if one series is first reversed in time, obtained from the complex conjugate in the frequency domain as

$$CC(X, Y) = \mathcal{F}^{-1}\{\mathcal{F}(X) * \mathcal{F}(Y)\} \tag{4}$$

where $\mathcal{F}(\cdot)$ and $\mathcal{F}^{-1}(\cdot)$ denote the Discrete Fourier Transform (DFT) and Inverse Discrete Fourier Transform (IDFT) operations respectively. Though, DFT and IDFT still require $\mathcal{O}(n^2)$ time, by using Fast Fourier Transform (FFT) algorithm, the time complexity of SBD computation can be reduced to $\mathcal{O}(n \log n)$ [10]. This efficient implementation of SBD has been used in this paper.

2.1.2. Maximin Shape Sampling

Maximin Shape Sampling (MMSS) selects shape-distinct subsequences from the candidate of all subsequences generated from a waveform dataset for further processing. MMSS combines maximin sampling scheme [22] with SBD to extract k' (an overestimate of the number of distinct shapelets in a time-series) subsequences that are maximally apart with respect to their shapes. Given pairwise SBDs in the matrix $R^{K\times K}$ between K candidate subsequences in a dataset, MMSS returns the indices $(m_1, m_2, ..., m_{k'})$ of the k' distinguished shapelets with minimal pairwise distance between them maximized. It randomly initializes m_1 and the search array $d = (d_1, ..., d_K) = (R_{m_1,1}, ..., R_{m_1,m_K})$. Then, for each iteration $t : 2, ..., k'$, d is updated with $(min\{d_1, R_{m_{t-1},1}\}, min\{d_2, R_{m_{t-1},2}\}, ..., min\{d_K, R_{m_{t-1},K}\})$, followed by the selection of the next maximin index $m_t = argmax_{1\leq j\leq K}\{d_j\}$. MMSS therefore reduces the search space of the number of possible shapelet candidates based on their global shape similarity. Initially, MMSS is applied separately to $S_{i,l}$ for each $i \in \{1, N\}$ to get the distinct k' shapelets $S^s_{i,l} \in S_{i,l}$ from each series thereby forming $S^s_l = \{S^s_{1,l}, S^s_{2,l}, ..., S^s_{N,l}\}$. However, it is expected that multiple time-series of the same class will contain similar shapelets with the same l, and hence MMSS is re-applied to S_l to get the most 'shape-distinct' $k^{s'}$ shapelets $S^{s'}_l$ for each l. The pruned candidates are therefore the maximin shape samples, denoted as $S^{s\prime} = \{S^{s'}_{L_{min}}, S^{s'}_{L_{min+1}}, ..., S^{s'}_{L_{max}}\}$.

2.1.3. Shapelet Selection

The original shapelet related work use *information gain* to determine the quality of a shapelet [16, 17]. First, all distances between a candidate shapelet S and all time-series in $W^{N\times n}$ are computed to generate a list of N distances $D^S =< d_{1,S}, d_{2,S}, ..., d_{N,S} >$ using (1). Information gain is then computed on the class values for each possible binary split of the sorted D^S list. The candidate resulting in the highest information gain is assessed to be the best shapelet. Alternatively, significance values

evaluated from statistical tests have also been used to estimate shapelet quality [23]. Multiple shapelets can be selected through the use of F-statistic of ANOVA to assess differences in class distributions for supervised shapelet transformation for multiple classes [18]. However, since our proposed approach relies on VAT, which is an unsupervised method to explore possible clusters, we cannot use class labels to assess the quality of the candidate shapelets.

We leverage the principle of finding unsupervised shapelets [21], in which each shapelet can separate a subset of time-series from the rest of the dataset based on whether that subset contains the particular shapelet. A good shapelet S should be able to divide the waveform dataset W into two groups of time-series, T_A and T_B, such that, T_A consists of the time-series that have subsequences similar to S while T_B contains the other time-series in W. Therefore, the mean value of $D^A =< d_{T,S} >, \forall T \in T_A$ is expected to be lesser than the mean value of $D^B =< d_{T,S} >, \forall T \in T_B$ ($D^A, D^B \in D^S$). Larger the gap between the means of the distance vectors D^A and D^B, the better cluster formation can S lead to. We use a greedy search algorithm [21] to maximize the separation gap between two subsets of W given by (5), where, μ_A and μ_B denote the means and σ_A and σ_B denote the standard deviations of D^A and D^B respectively.

$$gap_S = (\mu_B - \sigma_B) - (\mu_A + \sigma_A) \tag{5}$$

For each candidate S, this separation measure is iteratively computed for all ($N-1$ for N time-series in the dataset) possible splits of the dataset to obtain the maximum gap score *maxgap* for S. In order to exclude subsequences that result in all but one time-series to belong to either T_A or T_B (undesirable clusters: single outlier or universal pattern) from getting selected, we ensure the ratio $|D^A|/|D^B|$ lies within the range $1/c < |D^A|/|D^B| < (1-1/c)$, where c is a guess on the number of clusters. Once we know *maxgap* for all $k^{s'}$ candidates in $S^{s'}$, we retain the top k shapelets for which *maxgap* is above a preset threshold.

2.2. shape-iVAT

Once the k best shapelets have been found, a transformation (Definition 2.4) is carried out, where distance between each time-series T in W and each selected shapelet $S_j^{s'} \in S^{s'}, j \in \{1,k\}$ is calculated to estimate the presence of a similar subsequence in the time-series, thereby forming $W_S \in \mathbb{R}^{N \times k}$, on which we now employ VAT. W_S is represented by an undirected graph $G(\mathbb{V},\mathbb{E})$, where $\mathbb{V}$ denotes the N time-series as nodes so that $|V| = N$, and the edge set $\mathbb{E}$ denotes the distance between each pair of transformed time-series in $\mathbb{V} \times \mathbb{V}$, so that $|E| = N(N-1)/2$. The edge weight between the i^{th} and j^{th} transformed series is the $L2$-norm $D_{N\,ij}$ between them. A Minimum Spanning Tree (MST) of $G(\mathbb{V},\mathbb{E})$ is then constructed by reordering the dissimilarity matrix D_N to D_N^* using a modified Prim's algorithm. The reordering is done with respect to edge insertion ordering of the vertices added to the MST and initial vertex for MST formation is either end of the longest edge [6]. Each pixel intensity in image $I(D_N^*)$ obtained from the reordered distance matrix D_N^* reflects the dissimilarity between the corresponding time-series. When displayed as an intensity image in gray levels, $D_{N\,ij}^* = 0$ or minimum dissimilarity corresponds to pure black and $D_{N\,ij}^* = max(D_{N\,ij}^*)$ across all i,j corresponds to pure white. All other values of $D_{N\,ij}^*$ correspond to intermediate values of gray levels. Dark blocks appearing along the diagonal of D_N^* represent potential (say, c) clusters. We use the iVAT algorithm [7] to create sharper images by replacing input distances $[d_{ij}]$ in D_N by $D'_N = [d'_{ij}]$ calculated as (6) prior to VAT reordering, where $r \in P_{ij}$ is an acyclic path in the set of all acyclic paths between the transformed i^{th} and j^{th} time-series.

$$d'_{ij} = \min_{r \in P_{ij}} \max_{1<h<|r|} D_{N\,r[h]r[h+1]} \tag{6}$$

Single-linkage clusters [8] are always diagonally aligned in VAT/iVAT images and these can be obtained by cutting the largest $c-1$ edges (given by the MST cut magnitude order) in the MST [1]. We use Partition Accuracy (PA%) to evaluate the goodness of these output partitions, calculated as the ratio of the number of samples with matching ground truth labels and output labels to the total number of samples in the data [1]. Before the PA is calculated, it is ensured that the output labels obtained from the single-linkage clusters correspond to the same subsets in the ground truth.

3. RESULTS AND OBSERVATIONS

Experiments are performed on the Complex Upper-Limb Movements (UL) database [13] from the Physiobank repository [24]. UL comprises 3D (x-, y-, z- co-ordinates) hand trajectory data collected using a camera-based motion capture system (VICON) from ten healthy subjects during different motor tasks. We used 261 time-series samples across six classes of movements in our study: 2D movements of drawing *circle* (C1), *flower* (C2), *supermegacloud* (C3) and *spiral*(C4), and 3D reach-and-transport (a target) (C5) and random (C6) movements. For 2D (and 3D) motion analysis, x- and y- co-ordinate (and all three co-ordinate) data were concatenated respectively. Fig. 1 shows two examples of recurrent patterns from each class of 2D motion, that are of our interest. It can be clearly observed that for the same class, these patterns vary in temporal location and magnitude but retain their characteristic shapes. Prior to presenting the results on UL, we use a synthetic dataset *Trace*, from the UCR time-series classification/clustering repository [25] to illustrate our method. *Trace* comprises 200 time-series samples across 4 classes. All computations are performed using MATLABR2019b [26].

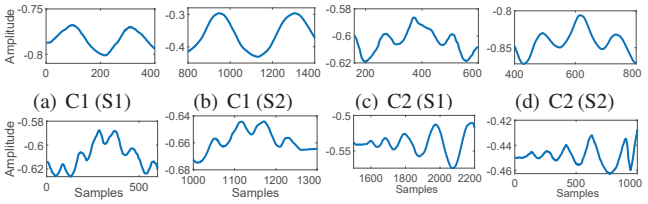
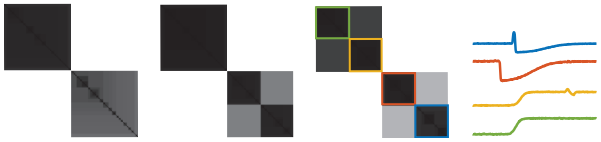


Fig. 1. Examples of recurrent subsequences in UL data (distance in meters) from 4 (C1-C4) 2D motion classes across 10 subjects (S1-S10).

Fig.2 shows cluster visualizations and PA using shape-iVAT along with iVAT using global $L2$-norm ($L2$-iVAT) and SBD (SBD-iVAT) for *Trace*. Here, shape-iVAT was constructed using L_{min} and L_{max} respectively set to 5% and 50% of the length of each time-series, which resulted finally in the selection of $k = 9$ representative shapelets for transforming W into W_S using SBD based $dist$ (Eq.1) computation. The proposed method (Fig. 2(c)) shows four crisp clusters (PA = 100) corresponding to the four classes, indicated by the dark blocks along the diagonal. Global SBD (Fig. 2(b)) improves the iVAT image compared to using $L2$-norm (Fig. 2(a)), but still fails to capture local shape details. But shape-iVAT captures the local pattern difference between the yellow and green classes (Fig.2(d)), though they have the same global shape.

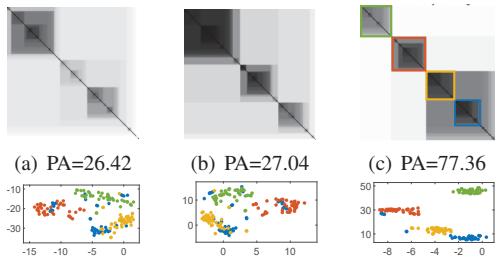
Figs.3 and 4 show cluster visualizations for 2D and 3D motion from UL respectively. Here, we also present the t-distributed Stochastic Neighbor Embedding (t-SNE) images [12, 27] for 2D visualizations of the high dimensional dataset W or W_S corresponding to iVAT/shape-iVAT images. Here, shape-iVAT was constructed using L_{min} and L_{max} respectively set to 5% and 20% of the length of each time-series, and $k = 16$ and $k = 6$ representative shapelets were finally selected for 2D and 3D motion data respectively. Global $L2$-norm produces overlapping clusters on the iVAT image in both cases as observed from Figs. 3(a)



(a) PA=50 (b) PA=75 (c) PA=100 (d)

Fig. 2. Evaluation on *Trace* with PA(%): (a) $L2$-iVAT, (b) SBD-iVAT and (c) shape-iVAT. Coloured boxes in (c) represent cluster boundaries for corresponding classes, example data from each class shown in (d).

and 4(a) and the corresponding t-SNE images in Figs. 3(d) and 4(d). The use of global SBD improves cluster partitions for the four classes of 2D motions as observed from Figs. 3(b) and 3(e). However, shape-iVAT (Figs. 3(c) and 4(c)) outperforms the other iVAT images in both cases with clear clusters and significantly higher PAs. This is also evident from the better separability observed from the t-SNE images constructed on W_S, in Figs. 3(f) and 4(f). The coloured boxes in the shape-iVAT image of Fig. 3(c) shows C4 to be far apart than the rest of the clusters. This follows from Fig. 1, where it can be clearly seen that C4 has the most distinct shape of all, while C1, C2 and C3 are closely related. We computed $dist$ using $L2$-norm and SBD for 3D and 2D motions respectively. This is because, while each class of 2D motion has a characteristic subsequence, 3D random motion is actually composed of several reach-and-transport tasks, therefore, possessing similar component shape patterns [13]. In such a case, $L2$-norm was found to be more accurate in computing the similarity of a shapelet to each time-series. In all our experiments k' and $k^{s'}$ were set to 10 and 30 respectively, the threshold for $maxgap$ was set to the mean $maxgap$ across all candidate shapelets and c for avoiding undesirable shapelets while computing gap_S was set to 5 ($\sim$number of clusters [21] across all datasets). The choice of k' and $k^{s'}$ were made such that MMSS produced shape-distinct subsequences, while being over-estimates of the number of actual clusters in each dataset (following our previous work [1]) to ensure representative shapelets from each class. Increasing these values beyond the ones selected did not improve our results. Objective selection of these parameters remains a future scope of this study.

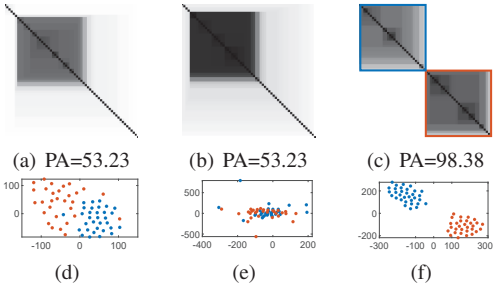


(a) PA=26.42 (b) PA=27.04 (c) PA=77.36

(d) (e) (f)

Fig. 3. Evaluation on 4 classes (C1 C2 C3 C4) of 2D movements in UL with PA(%): (a) $L2$-iVAT, (b) SBD-iVAT and (c) shape-iVAT and corresponding t-SNE images with (d) $L2$-norm, (e) SBD and (f) $L2$-norm on transformed dataset W_S.

Finally, we show an application where shape-iVAT can also be used to visualize anomalous movements based on a previously developed method [1, 28]. When single linkage clusters are obtained by cutting the MST given by edge distances d_n, the cluster boundaries defined by indices t which satisfy (7) are designated to be far apart. Here, $\alpha \geq 1$ is a parameter to control distance between two groups of time-series to be considered as separate clusters (smaller values of α represent tighter cluster boundaries). The t^{th} subset C_t of datapoints cut by the MST is considered anomalous if it satisfies (8). Here, $0 \leq \beta \leq 1$ is a parameter to control the size of a subset to be considered anomalous (smaller values



(a) PA=53.23 (b) PA=53.23 (c) PA=98.38

(d) (e) (f)

Fig. 4. Evaluation on reach-and-transport ( ) and random ( ) 3D movements in UL with PA(%): (a) $L2$-iVAT, (b) SBD-iVAT and (c) shape-iVAT and corresponding t-SNE images with (d) $L2$-norm, (e) SBD and (f) $L2$-norm on transformed dataset W_S.

of β will ensure only isolated/few data points remain anomalous).

$$d_{nt} > \alpha \times mean(d_n) \tag{7}$$

$$|C_t| < \beta \times N \tag{8}$$

We illustrate two examples in Fig. 5, when a data sample from random motion (C6) is added as an anomaly to (E1): 2D motion (C2) in the ratio 1:40 and (E2): 3D reach-and-transport motion (C5) in the ratio 1:32. α is chosen to be 4 and 2.5 in former and latter case respectively, while β is chosen as 0.03 in both cases following [1, 28]. Two dark blocks (one big and one tiny) along the diagonal represent two clusters in each shape-iVAT image. The lower plots show d_n, with a red horizontal line marking the cut threshold $\alpha \times mean(d_n)$. The group of time-series, for which d_n is above this threshold, are possible candidates for anomalies. The cluster with size less than $\beta \times N$ among them is anomalous. Therefore, in each case, the big dark block corresponds to a normal cluster (2D motion from C2 in Fig. 5(a) and 3D motion from C5 in Fig. 5(b)) and the tiny dark block, containing only one data point (circled for visibility), corresponds to the anomalous random motion.

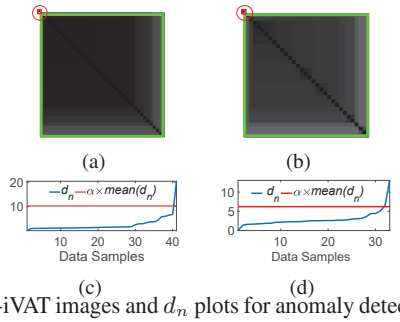


Fig. 5. shape-iVAT images and d_n plots for anomaly detection example E1 in (a) and (c) and E2 in (b) and (d) respectively. Green and red boxes denote normal and anomalous clusters respectively in (a) and (b).

4. CONCLUSION AND FUTURE WORK

In this paper, we have shown that extraction of interpretable patterns, called shapelets, to transform a waveform dataset into a lower dimensional representation significantly improves VAT evaluation. We performed experiments on a waveform dataset comprising complex upper limb motion and showed that the proposed method can identify different classes of 2D and 3D motion, as well as anomalous movements. However, shapelet extraction is computationally expensive and pose challenges for larger datasets. Future work in this direction include extending shape-iVAT (1) from learned shapelets instead of exhaustive search [20], (2) for identifying pathological motion patterns from wearable sensor data in neurological disorders such as stroke and epilepsy [14, 15] and (3) for visualization of evolving clusters [1, 5] through an *incremental* shape-iVAT for streaming data for long-term motion monitoring.

5. REFERENCES

[1] Punit Rathore, Dheeraj Kumar, James C Bezdek, Sutharshan Rajasegarar, and Marimuthu Palaniswami, "Visual structural assessment and anomaly detection for high-velocity data streams," *IEEE Transactions on Cybernetics*, 2020.

[2] L Minh Dang, Kyungbok Min, Hanxiang Wang, Md Jalil Piran, Cheol Hee Lee, and Hyeonjoon Moon, "Sensor-based and vision-based human activity recognition: A comprehensive survey," *Pattern Recognition*, vol. 108, pp. 107561, 2020.

[3] Jenny Margarito, Rim Helaoui, Anna M Bianchi, Francesco Sartor, and Alberto G Bonomi, "User-independent recognition of sports activities from a single wrist-worn accelerometer: A template-matching-based approach," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 4, pp. 788–796, 2015.

[4] Philippe Esling and Carlos Agon, "Time-series data mining," *ACM Computing Surveys (CSUR)*, vol. 45, no. 1, pp. 1–34, 2012.

[5] D. Kumar and J. C. Bezdek, "Visual approaches for exploratory data analysis: A survey of the visual assessment of clustering tendency (vat) family of algorithms," *IEEE Systems, Man, and Cybernetics Magazine*, vol. 6, no. 2, pp. 10–48, 2020.

[6] James C Bezdek and Richard J Hathaway, "Vat: A tool for visual assessment of (cluster) tendency," in *Proceedings of the IEEE International Joint Conference on Neural Networks*, 2002, vol. 3, pp. 2225–2230.

[7] Liang Wang, Uyen TV Nguyen, James C Bezdek, Christopher A Leckie, and Kotagiri Ramamohanarao, "ivat and avat: enhanced visual analysis for cluster tendency assessment," in *Pacific-Asia Conference on Knowledge Discovery and Data Mining*, 2010, pp. 16–27.

[8] Shreyasi Datta, James C Bezdek, and Marimuthu Palaniswami, "Experiments with dissimilarity measures for clustering waveform data from wearable sensors," in *IEEE Symposium Series on Computational Intelligence (SSCI)*, 2019, pp. 588–595.

[9] Amaia Abanda, Usue Mori, and Jose A Lozano, "A review on distance based time series classification," *Data Mining and Knowledge Discovery*, vol. 33, no. 2, pp. 378–412, 2019.

[10] John Paparrizos and Luis Gravano, "k-shape: Efficient and accurate clustering of time series," in *ACM SIGMOD International Conference on Management of Data*, 2015, pp. 1855–1870.

[11] Timothy B Iredale, Sarah M Erfani, and Christopher Leckie, "An efficient visual assessment of cluster tendency tool for large-scale time series data sets," in *IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, 2017, pp. 1–8.

[12] Sara Mahallati, James C Bezdek, Dheeraj Kumar, Milos R Popovic, and Taufik A Valiante, "Interpreting cluster structure in waveform data with visual assessment and dunn's index," in *Frontiers in Computational Intelligence*, pp. 73–101. 2018.

[13] Jose Garcia Vivas Miranda, Jean-François Daneault, Gloria Vergara-Diaz, Ana Paula Quixadá, Marcus de Lemos Fonseca, João Paulo Bomfim Cruz Vieira, Vitor Sotero dos Santos, Thiago Cruz da Figueiredo, Elen Beatriz Pinto, Norberto Peña, et al., "Complex upper-limb movements are generated by combining motor primitives that scale with the movement size," *Scientific reports*, vol. 8, no. 1, pp. 1–11, 2018.

[14] Anne Schwarz, Miguel Bhagubai, Gerjan Wolterink, Jeremia PO Held, Andreas R Luft, and Peter H Veltink, "Assessment of upper limb movement impairments after stroke using wearable inertial sensing," *Sensors*, vol. 20, no. 17, pp. 4770, 2020.

[15] Shitanshu Kusmakar, Chandan Karmakar, Bernard Yan, Ramanathan Muthuganapathy, Patrick Kwan, Terence J O'Brien, and Marimuthu Swami Palaniswami, "Novel features for capturing temporal variations of rhythmic limb movement to distinguish convulsive epileptic and psychogenic nonepileptic seizures," *Epilepsia*, vol. 60, no. 1, pp. 165–174, 2019.

[16] Lexiang Ye and Eamonn Keogh, "Time series shapelets: a new primitive for data mining," in *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*, 2009, pp. 947–956.

[17] Lexiang Ye and Eamonn Keogh, "Time series shapelets: a novel technique that allows accurate, interpretable and fast classification," *Data mining and knowledge discovery*, vol. 22, no. 1-2, pp. 149–182, 2011.

[18] Jason Lines, Luke M Davis, Jon Hills, and Anthony Bagnall, "A shapelet transform for time series classification," in *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*, 2012, pp. 289–297.

[19] Jon Hills, Jason Lines, Edgaras Baranauskas, James Mapp, and Anthony Bagnall, "Classification of time series by shapelet transformation," *Data Mining and Knowledge Discovery*, vol. 28, no. 4, pp. 851–881, 2014.

[20] Josif Grabocka, Nicolas Schilling, Martin Wistuba, and Lars Schmidt-Thieme, "Learning time-series shapelets," in *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, 2014, pp. 392–401.

[21] Jesin Zakaria, Abdullah Mueen, and Eamonn Keogh, "Clustering time series using unsupervised-shapelets," in *IEEE 12th International Conference on Data Mining*, 2012, pp. 785–794.

[22] Richard J Hathaway, James C Bezdek, and Jacalyn M Huband, "Scalable visual assessment of cluster tendency for large data sets," *Pattern Recognition*, vol. 39, no. 7, pp. 1315–1324, 2006.

[23] Jason Lines and Anthony Bagnall, "Alternative quality measures for time series shapelets," in *International Conference on Intelligent Data Engineering and Automated Learning*, 2012, pp. 475–483.

[24] Ary L Goldberger, Luis AN Amaral, Leon Glass, Jeffrey M Hausdorff, Plamen Ch Ivanov, Roger G Mark, Joseph E Mietus, George B Moody, Chung-Kang Peng, and H Eugene Stanley, "Physiobank, physiotoolkit, and physionet: components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. e215–e220, 2000.

[25] Hoang Anh Dau, Eamonn Keogh, Kaveh Kamgar, Chin-Chia Michael Yeh, Yan Zhu, Shaghayegh Gharghabi, Chotirat Ann Ratanamahatana, Yanping, Bing Hu, Nurjahan Begum, Anthony Bagnall, Abdullah Mueen, Gustavo Batista, and Hexagon-ML, "The ucr time series classification archive," 2018.

[26] The Mathworks, Inc., Natick, Massachusetts, *MATLAB version 9.7.0.1296695 (R2019b)*, 2019.

[27] Laurens van der Maaten and Geoffrey Hinton, "Visualizing data using t-sne," *Journal of machine learning research*, vol. 9, no. Nov, pp. 2579–2605, 2008.

[28] Dheeraj Kumar, James C Bezdek, Sutharshan Rajasegarar, Marimuthu Palaniswami, Christopher Leckie, Jeffrey Chan, and Jayavardhana Gubbi, "Adaptive cluster tendency visualization and anomaly detection for streaming data," *ACM Transactions on Knowledge Discovery from Data (TKDD)*, vol. 11, no. 2, pp. 1–40, 2016.