Time series clustering to examine presence of decrement in finger-tapping bradykinesia

Zhibin Zhao[1], Hui Fang[2], Stefan Williams[3], Samuel D. Relton[3], Jane Alty[4], and David C. Wong[5]

***Abstract*— Parkinson's disease is diagnosed base on expert clinical observation of movements. One important clinical feature is decrement, whereby the range of finger motion decreases over the course of the observation. This decrement has been assumed to be linear but has not been examined closely.**

We previously developed a method to extract a time series representation of a finger-tapping clinical test from 137 smartphone video recordings. Here, we show how the signal can be processed to visualize archetypal progression of decrement. We use k-means with features derived from dynamic time warping to compare similarity of time series. To generate the archetypal time series corresponding to each cluster, we apply both a simple arithmetic mean, and dynamic time warping barycenter averaging to the time series belonging to each cluster.

Visual inspection of the cluster-average time series showed two main trends. These corresponded well with participants with no bradykinesia and participants with severe bradykinesia. The visualizations support the concept that decrement tends to present as a linear decrease in range of motion over time.

***Clinical relevance*— Our work visually presents the archetypal types of bradykinesia amplitude decrement, as seen in the Parkinson's finger-tapping test. We found two main patterns, one corresponding to no bradykinesia, and the other showing linear decrement over time.**

I. INTRODUCTION

Bradykinesia is a pathological slowness of movement and is the core clinical feature of Parkinson's disease [1]. Clinically, it is assessed by a specialist clinician observing the patient repetitively tapping index finger and thumb together as 'wide and fast as possible' over ten taps. Key indicators of bradykinesia are reduction in tapping amplitude and/or impairment of rhythm (interruptions or hesitations). A characteristic decrement is frequently seen, whereby the amplitude reduces over the duration of the finger-tapping test.

Decrement is currently described within two clinically validated scales. In the Universal Parkinson's Disease Rating Scale (UPDRS), decrement is categorised how soon in the test the reduction in amplitude begins, but no minimum reduction is defined [2]. In the Modified Bradykinesia Rating Scale (MBRS), amplitude (which includes decrement) is defined on a scale of 0 (normal) to 4 (severe). For example, a score of 1 occurs when 'Mild reduction in amplitude in later performance, most movements close to normal'[3].

Previous attempts to quantify decrement more precisely in the clinical literature have primarily assumed linear decrease in amplitude over time. For instance, both Bank et al. and Martinez-Manzanera et al quantified decrement by computing a linear regression of the frequency amplitude and velocity over the course of a finger tapping task [4], [5]. However, our clinical experience has been that finger-tapping amplitude in Parkinson's is sometimes observed to increase after an initial decrease.

We hypothesize that there may be archetypal trends in decrement that are both non-linear, and inadequately described by the current clinical scales. In this work, we use a time-series clustering approach to identify whether such archetypal trends exist and to visualize the results.

II. METHODS

A. Data Collection

Video data recordings of 137 finger-tapping examination were recorded from 69 participants at Leeds Teaching Hospitals Trust using an iPhone smartphone at 60 fps. 39 participants had previously diagnosed idiopathic Parkinson's disease; the remaining 30 were health controls. All participants had recording of both left and right hands, but one video was rejected for quality reasons. In our case, the primary unit of analysis was the number of finger-tapping examinations, as we are not concerned here with per-participant outcomes.

The study cohort described here is an expanded sample of those recorded previously by Wong *et al.* [6], who report additional details of the data collection method. Videos were then processed to output pixel coordinates of the participant's thumb and index fingertips at each frame. We achieved this using the Deeplabcut software package [7], [8]. A neurologist

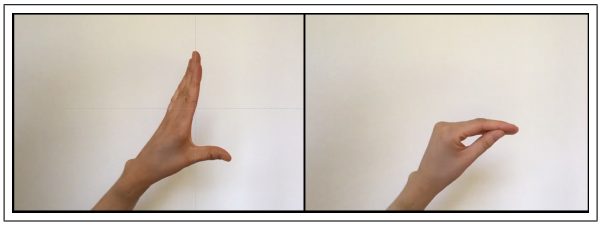

Fig. 1. Example frames from the smartphone video capture showing the hand and forearm in frame in an open (left) and closed (right) finger-tapping test position

[1]Zhibin Zhao is with the School of Mechanical Engineering, Xi'an Jiaotong University, Xi'an, China zhibinzhao1993@gmail.com

[2] H. Fang is with the Department of Computer Science, Loughborough Univerity, Loughborough, UK h.fang@lboro.ac.uk

[3] S.D. Relton and S. Williams are with the Leeds Institute of Health Science, University of Leeds, UK s.d.relton@leeds.ac.uk, stefan.williams2@nhs.net

[4] J. Alty is with Leeds Teaching Hospitals NHS Trust, UK and University of Tasmania, Australia j.alty@utas.edu.au

[5] D.C. Wong is with the Department of Computer Science and the Centre for Health Informatics, University of Manchester, Manchester, UK corresponding author: david.wong@manchester.ac.uk,

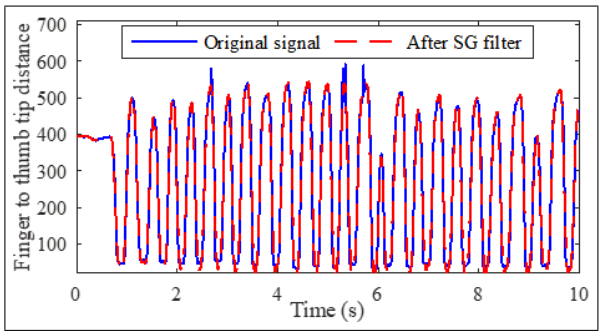

Fig. 2. One example of the original signal and the SG filtered signal

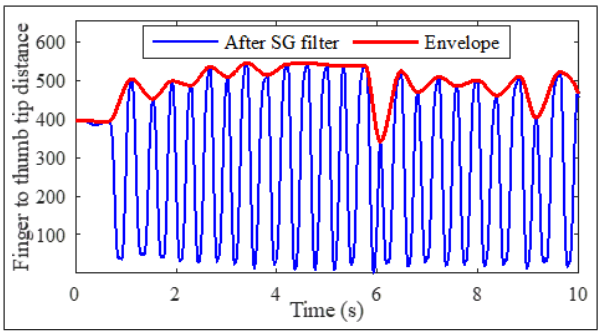

Fig. 3. One example of the envelope extraction output

(SW) manually labelled 20 frames from each 660 frame videos. The salient points were automatically extracted for the remaining 97% of frames by the algorithm. Visual inspection of the salient points aligned to the original video were used to confirm the accuracy of the algorithm output.

Each video was independently assessed for bradykinesia by a random selection of 22 Parkinson's specialists, using the MBRS scale to provide a score between 0 and 4. Each neurologist assessed 30 videos. Videos were labelled using the modal score.

During the examination, participants rested their elbow on a chair arm, and the video camera was adjusted such that the hand and forearm were in frame (see Fig. 1). The lateral surface of the hand faced the camera. Participants were told to tap their thumb and index finger together as wide and as fast as possible. Each video sample was restricted to 11 seconds.

B. Preprocessing

Using the output pixel coordinates of the thumb and index fingertips, we generated time series data of the finger to thumb tip distance of each participant. One example is shown in Fig. 2, where one may observe outliers (large, sudden transient label 'jumps') in the original signal due to mistakes in labelling the index and middle fingers by deeplabcut. To remedy this, we applied a Savitsky-Golay (SG) filter with polynomial order 3 and frame length 11 to all the time series. Fig. 2 also shows the output of the filtered time series.

C. Envelope Extraction

After performing the SG filter of each original signal, we extracted the upper envelopes of the filtered signals. This represents the maximum distance between thumb and index finger tip over time. The designed envelope analysis mainly consists of two steps: finding the peak values and interpolating the envelope with those peak values.

Peak detection was undertaken using MATLAB's findpeaks function [9]. We restricted acceptable peak-to-peak intervals to be greater than 0.3 seconds to avoid false peaks which are very close to each other. The threshold was determined by visual inspection, using the average interval of adjacent finger taps as a guide. To ease the undulation and maintain the wave shape, we performed shape-preserving piecewise cubic interpolation of values at neighboring peak points.

Fig. 3 shows one example of the envelope analysis in which the extracted envelope accurately reflects the variance of peak values. Finally, we normalized all the extracted envelopes through Z-score normalization using the mean and standard deviation of each envelope.

D. Clustering and Visualization of Decrement

Clustering is an unsupervised learning approach for finding intrinsic patterns in data. Typically, similarity of dataset features are compared using a distance metric, and similar features are assigned to the same cluster. Multiple approaches have been considered for time series clustering, including Gaussian process clustering [10] and Hidden Markov models [11]. Here, we implement two related methods for time series clustering based on K-means clustering.

The first approach was to apply K-means clustering directly to the time series. K-means involves iteratively assigning a time series to the cluster centre with the lowest Euclidean distance and updating the cluster centre according to the centroid of its members. The Euclidean distance, d, between two time series, A and B, of equal length, n, is computed as:

$$d(A,B) = \sqrt{\sum_{t=1}^{n} (A_t - B_t)}$$

We used a predetermined $K = 5$ to match the number of distinct classifications within the MBRS clinical rating score. From these clusters, we sought to find the archetypal time series by extracting the arithmetic mean of the within-cluster time series. The arithmetic mean is the standard approach used to calculate an average sequence when all the sequences are consistent with each other. However, this approach risks losing shape information in instances where similar sequences are shifted in time. Fig. 4 shows one example, in which the shape of the signals in Fig. 4(a) are not preserved by the arithmetic mean in 4(b).

The second approach used an alternative distance measure named dynamic time warping (DTW). DTW has been a standard metric for a large range of time series analysis applications, such as speech recognition [12], gene expression [13] and biomedical signals [14]. DTW aims to align

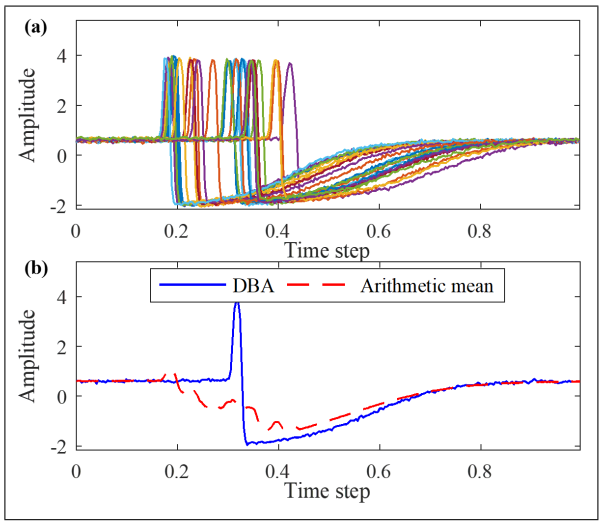

Fig. 4. One example of time series averaging. (a) the set of time series data (the first class in the Trace dataset) and (b) the average signals extracted by arithmetic mean and DBA.

different time sequences to find the optimal alignments. It allows for both time shifts and non-linear distortions, and it does not require that the length of two sequences should be equal. The DTW metric approaches zero as two sequences become more similar.

Multiple related studies have attempted to address the problem of approximating an average time series [15], [16]. Among them, dynamic time warping barycenter averaging (DBA) proposed by Petitjean et al. is one relatively robust approach [17].

The DBA algorithm is an iterative procedure which consists of the following two steps: (1) Calculate the DTW metric between the temporary average sequence and each sequence in the predefined set and find the relationships (similarities) between elements of each sequence in the predefined set and elements of the temporary average sequence; (2) Update each element of the average sequence using the barycenter of elements which are related to it corresponding to the DTW metric in the above step. The convergence property, and demonstration of how DBA can be applied alongside existing clustering methods like K-means has been proven empirically and theoretically [18].

Fig. 4(b) shows that, compared with a simple arithmetic mean, DBA can preserve the local structure of a time series. We denote the clustering method used in this paper as K-DBA, signifying the use of K-means with DTW as the distance measure alongside DBA as the averaging approach.

III. RESULTS

For K-DBA, the number of centroid seeds is 10 and the number of iterations for the DBA computation is 20. The results of traditional K-means and K-DBA, in which we show the average time series (in bold red) alongside the individual time series belonging to the cluster (grey), are shown in Fig. 5 and Fig. 6, respectively. The membership of each archetypal cluster, grouped by clinically assessed MBRS decrement score, is presented in Tables I and II for K-DBA and K-means respectively.

The figures show two main time series trajectories. Fig. 5(c) and 6(a) mainly represent the participants with no bradykinesia decrement; the amplitude throughout the archetypal time series is relatively stable. As expected, the K-DBA derived signal contains greater local structure, but the two approaches provide a similar overall shape. The tables show that these clusters correspond well with those with an MBRS decrement score of 0. For 5(c) and 6(a), 37/58 and 27/38 time series were rated with an MBRS = 0, respectively.

Figs. 5(d) and 6(c) represent clusters in which decrement is clearly visible. For these clusters, the amplitude envelope appears to decrease linearly over time. Both clusters were associated with participants assessed with more severe bradykinesia. The remaining representative clusters show large and rapid changes in amplitude at the start and end of the time series. These changes are visible in the original envelope signals but were not usually present in the raw periodic finger-tapping signal. Therefore, the pattern is most likely due to artefact introduced at the piecewise interpolation step.

TABLE I
K-MEANS CLUSTER MEMBERSHIP GROUPED BY MBRS AMPLITUDE (DECREMENT) SCORE

Cluster	**MBRS Amplitude (Decrement Score**				
	0 (normal)	***1***	***2***	***3***	***4 (severe)***
(a)	8	3	3	1	1
(b)	5	2	1	4	0
(c)	37	10	8	3	0
(d)	9	9	5	7	2
(e)	4	2	4	4	1

TABLE II
K-DBA CLUSTER MEMBERSHIP GROUPED BY MBRS AMPLITUDE (DECREMENT) SCORE

Cluster	**MBRS Amplitude (Decrement Score**				
	0 (normal)	***1***	***2***	***3***	***4 (severe)***
(a)	27	6	4	1	0
(b)	6	5	2	5	1
(c)	8	7	5	9	1
(d)	19	7	9	1	3
(e)	3	1	1	0	2

IV. DISCUSSION

We have developed a method to visualize bradykinesia decrement directly from smartphone video recordings of a finger-tapping clinical assessment. We have also used a shape-preserving time series clustering approach to extract common archetypal decrement trends in a cohort of 133 videos.

Although we selected five cluster centres a priori based on the gradation of the MBRS decrement scale, only two clearly distinct time series patterns were observed - the remaining

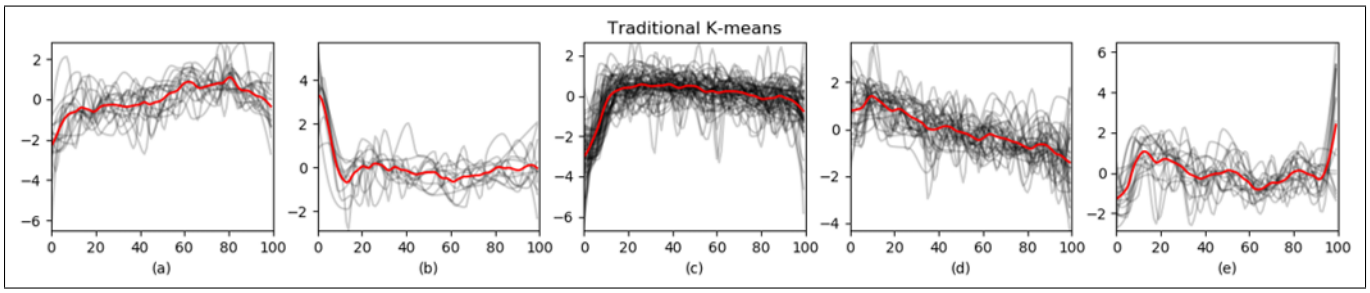

Fig. 5. The visualization results of Traditional K-means (red lines are average sequences of five class centers).

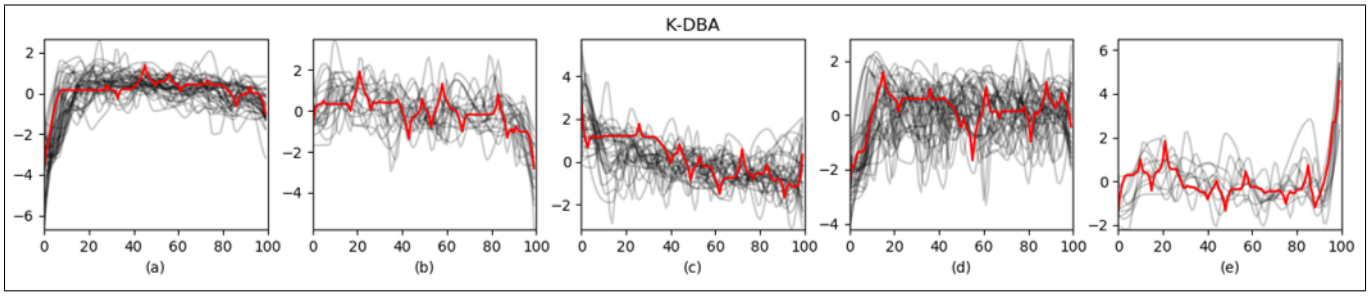

Fig. 6. The visualisation results of K-DBA (red lines are average sequences of five class centers).

three clusters appear to have similar characteristics. These two distinct archetypes corresponded well with participants with no bradykinesia and with high bradykinesia.

Mild grades of bradykinesia (MBRS=1-2) were not clearly associated with any cluster. It is unclear why this is the case. Perhaps mild grades of decrement are indistinguishable from low levels (or no) decrement. However, it is more likely that the decrement patterns do not constitute an independent cluster because the data set contains relatively few examples of medium decrement. Future work should therefore include repetition on a larger, balanced data set.

With respect to our initial hypothesis, that there may exist non-linear decrement trajectories, our analysis found no evidence to lend support. This means that current quantitative metrics using linear regression are likely to be sufficient in clinical practice. The method shown here is flexible and extensible to clinical time series data more generally. It allows grouping of similar time series trajectories to identify common trend patterns.

REFERENCES

[1] A. Lees. "Parkinson's disease," Practical Neurology, vol 10, no. 4, pp. 240–246, 2010.

[2] P. Martinez-Martin, A. Gil-Nagel, L. Morlan Gracia, J. Balserio Gomez, J. Martines-Sarries, F. Bermejo. "Unified Parkinson's disease rating scale characteristics and structure," Movement Disorders. Vol. 9, no. 1, pp. 76–83, 1994.

[3] D. Heldman *et al.* "The modified bradykinesia rating scale for Parkinson's disease: Reliability and comparison with kinematic measures," Movement Disorders, vol. 26, no. 10, pp 1859–1863, 2011.

[4] P.J. Bank, J. Marinus, C.G. Meskers, J.H. de Groot, J.J. van Hilten. "Optical Hand Tracking: A Novel Technique for the Assessment of Bradykinesia in Parkinson's Disease," Movement disorders clinical practice, vol. 4, no. 6, pp. 875-83, 2017

[5] O. Martinez-Manzanera, E. Roosma, M. Beudel, R.W. Borgemeester, T. van Laar, N.M. Maurits. "A method for automatic and objective scoring of bradykinesia using orientation sensors and classification algorithms," IEEE Trans. on Biomed Eng. Vol. 63, No. 5, pp.1016-24, 2015.

[6] D.C Wong, S.D. Relton, H. Fang, R. Qhawaji, C.D. Graham, J. Alty, S. Williams. "Supervised classification of bradykinesia for Parkinson's disease diagnosis from smartphone videos," Proc. IEEE 32nd Int Symp on Computer-Based Medical Systems (CBMS) 2019 Mar 29.

[7] A. Mathis *et al.* "DeepLabCut: markerless pose estimation of user-defined body parts with deep learning," Nature Neuroscience, vol 21, pp. 1281-1289, 2018.

[8] S. Williams, Z. Zhao, A. Hafeez, D.C. Wong, S.D. Relton, H. Fang, J. Alty. "Visual artificial intelligence can measure Parkinson's bradykinesia correlation with 22 neurologists," J. Neurology, submitted for publication.

[9] MATLAB and Statistics Toolbox Release 2018a, The Mathworks, Inc., Natick, Massachusetts, United States.

[10] M.A. Pimentel, D.A. Clifton, L. Tarassenko. "Gaussian process clustering for the functional characterisation of vital-sign trajectories," Proc. - 2013 IEEE Int. Workshop on Machine Learning for Signal Process. (MLSP) pp. 1-6, 2013.

[11] E.B. Fox, E.B. Sudderth, M.I. Jordan, A.S. Willsky, "Sharing Features among Dynamical Systems with Beta Processes," Proc. - Neural Information Processing Systems, Vancouver, Canada December 2009.

[12] H. Sakoe "A dynamic programming approach to continuous speech recognition," Proc. – Int. Congress of Acoustics, Budapest, Hungary, 1971.

[13] J. Aach, G.M. Church. "Aligning gene expression time series with time warping algorithms," Bioinformatics, vol. 17, no. 6, pp. 495-508, 2001.

[14] D. Wong, I Strachan, L Tarassenko. "Visualisation of high-dimensional data for very large data sets," Proc. - Workshop Mach. Learn. Healthcare Appl., 25th Int. Conf. Mach. Learn., Helsinki, Finland, 2008.

[15] L. Gupta, D.L. Molfese, R. Tammana, P.G. Simos. "Nonlinear alignment and averaging for estimating the evoked potential," IEEE Trans. Biomed. Eng, vol. 43, no.4, pp.348-56, 1996.

[16] V. Niennattrakul, C.A. Ratanamahatana. "Inaccuracies of shape averaging method using dynamic time warping for time series data," Proc. – Int. Conf. Computational Science, Berlin, May 2007.

[17] F. Petitjean, A. Ketterlin, P. Gançarski. "A global averaging method for dynamic time warping, with applications to clustering," Pattern Recognition, vol. 44, no.3 pp. 678-693, 2011.

[18] F. Petitjean, G. Forestier, G.I. Webb, A.E. Nicholson, Y. Chen, E. Keogh. "Dynamic time warping averaging of time series allows faster and more accurate classification," Proc. – Int. Conf. Data Mining, pp. 470-479, 2014.