



HAL
open science

Identifying patterns in trunk/head/elbow changes of riders and non-riders: A cluster analysis approach

Imen Trabelsi, Romain Hérault, Héloïse Baillet, Régis Thouvarecq, Ludovic Seifert, Gilles Gasso

► **To cite this version:**

Imen Trabelsi, Romain Hérault, Héloïse Baillet, Régis Thouvarecq, Ludovic Seifert, et al.. Identifying patterns in trunk/head/elbow changes of riders and non-riders: A cluster analysis approach. *Computers in Biology and Medicine*, 2022, 143, pp.105193. 10.1016/J.COMPBIOMED.2021.105193 . hal-03552674

HAL Id: hal-03552674

<https://hal.science/hal-03552674>

Submitted on 22 Jul 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

Identifying Patterns in Trunk/Head/Elbow Changes of Riders and Non-riders: A Cluster Analysis Approach

Imen Trabelsi ^{a,*}, Romain Hérault^a, Héloïse Baillet^b, Régis Thouwarecq^b,
Ludovic Seifert^b, Gilles Gasso^a

^aNormandie Univ., UNIROUEN, UNIHAVRE, INSA Rouen, LITIS, France

^bCETAPS Laboratory, Faculty of Sports Sciences, University of Rouen, Normandie Université, France.

Abstract

Correct rider oscillation and position are the basics for a good horseback riding performance. In this paper, we propose a framework for the automatic analysis of athletes behaviour based on cluster analysis. Two groups of athletes (riders vs non-riders) were assigned to a horseback riding simulator exercise. The participants exercised four different incremental horse oscillation frequencies. This paper studies the postural coordination, by computing the different discrete relative phases of head-horse, elbow-horse and trunk-horse oscillations. Two clustering algorithms are then applied to automatically identify the change of rider and non-rider behaviour in terms of postural coordination. The results showed that the postural coordination was influenced by the level of rider expertise. More diverse behaviour was observed for non-riders. At the opposite, riders produced lower postural displacements and deployed more efficient postural control. The postural coordination for both groups was also influenced by the oscillation frequencies.

Keywords: Cluster analysis, Discrete relative phase, Mechanical horse, Postural coordination, Sport.

*Corresponding author

Email address: imen.beji@insa-rouen.fr (Imen Trabelsi)

1. Introduction

Hippotherapy refers to the use of characteristic movement of a horse to provide sensory feedback and has been used as a treatment tool. It provides multi-sensory input that can improve human walking, postural stability, motor control, breathing, and even having positive mental and social effects [42]. In most studies to date, these effects are evaluated on patients with physical and mental impairments. Shurtleff et al. [40], studied if the hippotherapy improves head and trunk stability in children with spastic diplegia cerebral palsy. Beinotti et al. [4] reported that hippotherapy demonstrated a positive influence in gait training for hemiparetic post-stroke. Menezes et al. [29] completed this proposal conducting an in-depth study on the effects of hippotherapy on the postural control of multiple sclerosis patients. Their study suggested a reduction of balance deficits and a better postural control of the patients. Hippotherapy has also been recognized as a potential tool for rehabilitation in children with cerebral palsy as part of their broader integrated plan of care [30]. Marked improvements in the quality of several motor functions were also noted in young children with Down syndrome when using physical therapy incorporating hippotherapy [31].

Hippotherapy becomes even more accessible with a mechanical reproduction of horse movement. Indeed, mechanical horseback riding presents several advantages over hippotherapy, including lesser cost and no space or weather limitations [27, 19]. The effectiveness of using horse mechanical back riding on balance and spinal geometry in hemiplegic children by comparing its effect with hippotherapy was studied in [14]. The results showed that these simulators showed significant improvements in the spinal geometry and balance and can be an efficient alternative method for real hippotherapy. In [36], the authors provide evidence of the long-term effect of mechanical horse back riding on postural control in children with spastic cerebral palsy. Results indicated that mechanical horse back riding may be an important treatment for improving postural muscles size and associated static and dynamic stability. The use of mechanical horse back riding was also explored for healthy people. An up-to-date research

analysis on horse riding simulation exercise for healthy older adults was provided in [20]. The main outcome of this study was that simulated hippotherapy might improve balance, mobility gait ability and muscle strength. Baillet et al. [3] analyzed the postural coordination of riders and non-riders during an exercise protocol on the mechanical horse at four different oscillation frequencies. A statistical analysis, based on one-way Analysis of Variance (ANOVA) [33] with Bonferroni correction, was performed to understand the postural coordination of the subjects. The findings of this study provided up-to-date insights about the impact of mechanical horse on postural dynamics correlated with the increasing oscillation frequencies.

So far, postural coordination have most often been analyzed in a standardized way, based on descriptive analysis [13, 18, 15, 2, 34, 1].

The effectiveness of mechanical horse riding simulators on postural coordinations was investigated in [13], employing a diversity of statistical scales. The authors [34] studied the effects of horseback riding on postural control of expert riders vs non-athletes using a statistical analysis. In [15], the study evaluated the effects of horses' movements in rider posture by examining kinematic and kinetic measurements. In [2], the authors have used the Haken-Kelso-Bunz (HKB) model in the analysis of interpersonal coordination and other joint-action tasks. HKB model is considered as the most extensively tested quantitative model in human movement behaviour. In [1], the authors studied the inter-limb coordination based on the physiological finding of central pattern generators and muscle synergy.

However, while these approaches constitute useful first approximations, they can't take into account the different interactions between the different variables. They are based on numerical summary of data. In that regard, we avoid such descriptive analysis but instead, we explore specific cluster distance measures to measure similarity between postural coordination and to characterize behavioral patterns across the users.

In this context, the current work, a continuation of the one led by Baillet et al. [3], aims to consolidate the previous results by adopting an automatic

analysis of postural coordination. In fact, Baillet et al. [3] included information beforehand on the type of participant in their experimental analysis. This paper, on the other hand, consists of providing automated tools to find underlying
65 structures (clusters) from untagged data, without anterior knowledge about the level of riding expertise of the participants (neophytes vs experts) and any statistical a priori information. These formed clusters are then used as the basis for postural coordination analysis. This study is thus another step in a larger project to develop rehabilitation protocols using the mechanical horse.

70 The current study has two objectives: (i) to automatically identify the differences between the performance of riders and non-riders after a mechanical horse back riding. This automatic analysis is unsupervised and may provide some representative cluster prototypes that can be used as the basis for future work either for sportsmen with the aim of improving their posture coordination
75 in competition or for people with disabilities, (ii) to understand and analyze the impact of increasing horse oscillation frequencies on the postural dynamics of subjects. Indeed, this is useful to see how the participants follow the horse by adopting their movement to horse's oscillations.

Starting from acquired signals from riders and non-riders during an exercise
80 protocol on the mechanical horse, we perform the following steps: elbow, trunk and head oscillating cycles are extracted from their corresponding signals. To better understand the behavior of the subjects, we need to organize the extracted cycles in distinct clusters with similar observations within each cluster. Cluster analysis, primitive exploration with little or no prior knowledge, is used
85 here to perform data exploration [43]. The outcomes of cluster analysis in this research work allow to draw conclusions about the impact of horse back riding in the behavior change of neophyte and expert practitioners. Moreover, our empirical evidences show that the investigated methods can automatically identify different riding behaviours without a priori information on rider level.

90 There can be several motivations behind the choice of a clustering algorithm. A clustering algorithm can be used to improve existing baselines or to merely provide a point of comparison and validation of existing works. Since the pro-

posed work provides a point of comparison and validation of the work of Baillet et al. [3], hence, even though clustering algorithms more sophisticated than K -means exist, we use the K -means algorithm [22, 28, 17] owing to its simplicity and known limitations. This clustering algorithm provides a good trade-off between the quality of the solution obtained and its computational complexity. It proceeds by minimizing the within cluster dissimilarity based on the distances of the assigned cycles to the cluster prototypes. Hence, distance functions in K -means play an important role. Two distinct metrics are used to measure the distance between the cycles: Euclidean distance and Dynamic Time Warping (DTW) distance [25]. Euclidean distance is the generic metric of K -means. However, considered that recorded cycles are temporal signals, a more appropriate metric, DTW, is also investigated. Although more sophisticated clustering methods can be applied, the key point in our approach is the study of temporal signals which deserve specific distance measure. Therefore we combine a simple clustering algorithm, namely K -means, with a highly elaborated metric, the DTW, to cope with the peculiarities of the task at hand.

2. Materials and methods

2.1. Experimental Protocol

2.1.1. Participants

The study group was comprised of 22 students within the faculty of sports sciences. Only half of them were horseback riders, whereas the non-riders have also a strong background in sports. None of the subjects involved in the experiments had a history of physical disability or balance disorders, and all were nonsmokers and non-gymnasts. Some studies have proved that gymnasts might have increased postural stability compared with other athletes [21, 16]. Therefore, gymnasts were not involved for the experiments. Information about participants are shown in Table 1.

Table 1: Participant descriptive characteristics

	Characteristics
Gender	Female
Number of riders	11
Number of non-riders	11
Age, mean \pm SD	24.3 ± 4.6
Number of years riding (for riders)	≥ 5
Number of training hours/week, mean \pm SD (for non-riders)	3.7 ± 3.9
Body mass, mean \pm SD (kg)	59.7 ± 7.3
Body height, mean \pm SD (cm)	167.7 ± 4.8

120 2.1.2. Data Acquisition Protocol

The mechanical horse, with a length of 174 cm, provided two dimensional movements: an anterior/posterior movement with an amplitude equal to ± 0.3 cm and an upward/downward movement, with an amplitude equal to ± 5.5 cm. Four horse oscillation frequencies (V_0 , V_1 , V_2 , V_3) were considered. The oscillations ranged from 50% to 100%. The subjects warmed up for 3 minutes on the mechanical horse at V_0 (57.7 osc/min, 0.96 Hz). The procedure continued with 6 minutes ride at V_1 (88.2 osc/min, 1.47 Hz), followed by another 6 minutes at speed V_2 (103.4 osc/min, 1.72 Hz), and ending with 6 minutes exercise at V_3 (150 osc/min, 2.5 Hz).

130 2.1.3. Postural coordination

During the test, the different participants held the reins to simulate the real posture. An optokinetic system (OptiTrack) [7] was used to record the postural coordination between them and the mechanical horse.

Optitrack is composed of ten cameras, to recognize fifteen reflective markers. Fourteen markers were fixed on the participants (i.e., head, vertebral level C2, high spinal cord injury T8, primary sensory cortex S1, left and right acromia, right elbow, right wrist, left and right anterior iliac spine, left and right posterior iliac spine, right knee, right ankle and right toe) and one on the mechanical horse

(behind the saddle).

140 Displacements of the participants' head, trunk and elbow were determined by these markers, as shown in Figure 1. As the markers were placed only on the right side of riders, only right elbow was analyzed. Baillet et al. [3] provide additional information about the experimental protocol.

Three primary methods are commonly used to evaluate the postural co-
145 ordination of coupling signals: (i) discrete relative phase (temporal difference between two similar occurrences from two oscillators, reported on the basis of the period of one cycle as a reference) [35]; (ii) continuous relative phase (spatio-temporal difference based on the phase planes generated from the angular position and angular velocity of the oscillators) [26] and (iii) vector coding
150 (a spatial difference based on an angle-angle plot) [9]. To evaluate the postural coordination, discrete relative phase (DRP) is used in this study. DRP computed the temporal difference between two oscillating segments, reported on the basis of the period of one cycle as a reference [11].

$$DRP(t_i) = \frac{\phi_1(t, i) - \phi_2(t, i)}{\phi_1(t, i + 1) - \phi_2(t, i + 1)} \quad (1)$$

where ϕ_1 is the maximum rotation of segment 1, ϕ_2 is the maximum rotation
155 of segment 2, t is time and $DRP(t, i)$ is the phase difference during the cycle i .

Three discrete relative phases, corresponding to the angles of trunk, elbow and head, estimated on the basis of the horse's oscillator were computed.

The first two angles, called segmental angles, represent the head angle and the trunk angle. The head angle (head/horse relative phase) is estimated by
160 the point coordinates of the highest point on the skull and the second cervical vertebrae markers, expressed as a function of vertical axis ($\phi_{head} - \phi_{horse}$). The trunk angle (trunk/horse relative phase) is estimated by the point coordinates of the seventh cervical vertebrae and the fifth sacral vertebrae markers, expressed as a function of vertical axis ($\phi_{trunk} - \phi_{horse}$). The third angle, called articular
165 angle, represents the elbow angle (elbow/horse relative phase). It is estimated between the point coordinates of the right acromion, elbow and wrist markers ($\phi_{elbow} - \phi_{horse}$).

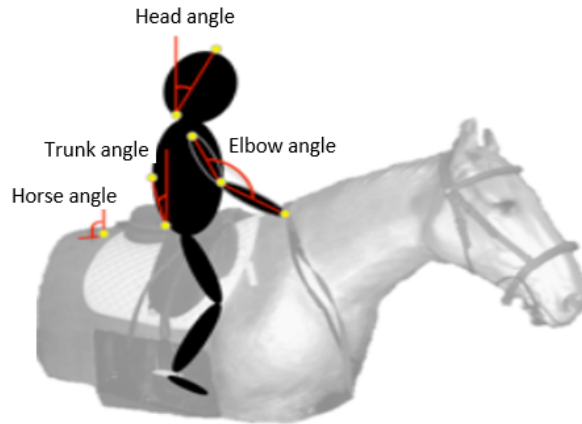


Figure 1: Postural angle representation, see [3].

2.2. Clustering analysis of postural coordination

In this Section, we describe our method to build riding behaviour models
 170 from their postural coordination, represented through the different angles from
 head, trunk and elbow. This method, illustrated in Figure 2, consists of two
 major components: cycle detection and cycle clustering analysis. In the first
 one, the different angles are separated into cycles of same length. The overall
 procedure is complemented by Algorithm 1 in Appendix summarizing the main
 175 steps of the approach. In the second one, two clustering algorithms are designed
 according to the characteristics of the problem, as showed in Algorithm 2.

2.2.1. Cycle detection

Cycle detection component is composed of two different steps. The first one
 aims at finding the local peaks from the horse signal which by design follows a
 180 sinusoidal shape. The detected peaks are required to be separated by more than
 a minimum peak distance to ignore the smaller peaks (related to undesirable
 measurement artifacts). The minimum peak value is fixed according to the
 oscillation frequency. Peak detection was used rather than pattern matching as
 the mechanical horse has variable frequency and phase.

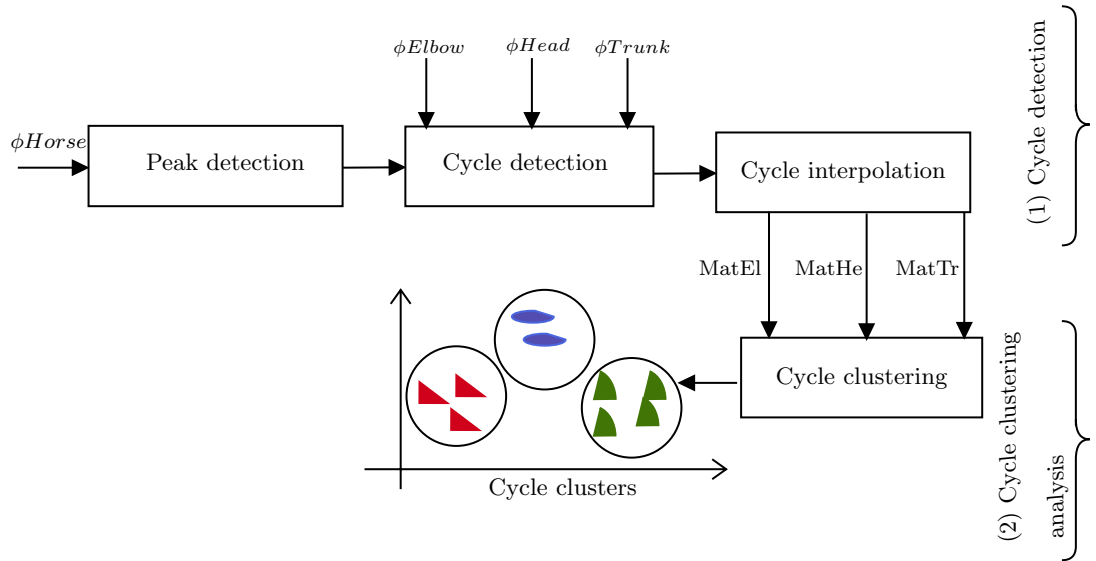


Figure 2: Proposed system components. MatTr, MatHe and MatEl stand respectively for cycle matrix of trunk, cycle matrix of head and cycle matrix of elbow.

185 Figure 3 illustrates the process of cycle detection from the different angles. It can be clearly seen that only the oscillatory movement of mechanical horse is sinusoidal, allowing to determine easily the peaks. From the indices at which these peaks occur, we process to the cell cycles detection on head, trunk and elbow signals.

190 As a remark, Figure 4 shows the power spectrum of these cycles using the Fast Fourier Transform (FFT). We then calculate the correlation between the obtained results (Figure 5). Postural responses occurred with a notable amplitude at the same two frequencies for the different signals, with prominent low-frequency as a common phenomenon. Fast head oscillations are represented
 195 by several low harmonic frequencies, which reflects the rapid movement of the head. The trunk is the most correlated with the horse. The corresponding signal follows almost the same sinusoidal movement as the mechanical horse.

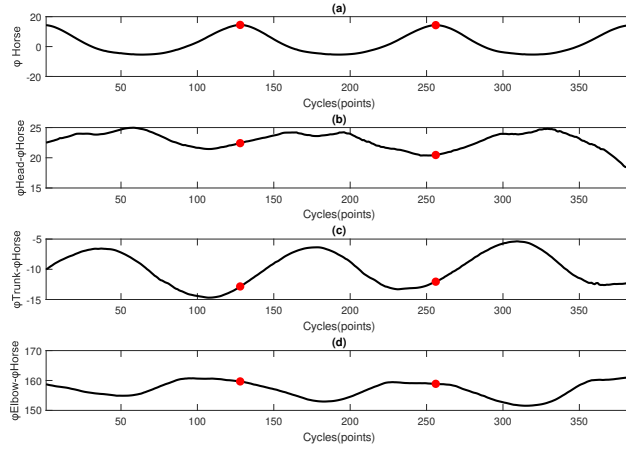


Figure 3: Cycle detection : (a) 3 sinusoidal horse cycles and their corresponding cycles from (b) head, (c) trunk and (d) elbow.

The higher value of correlation at lag=0 for head and elbow suggests that they follow better the horse motion at the beginning of the cycles (Figure 5).

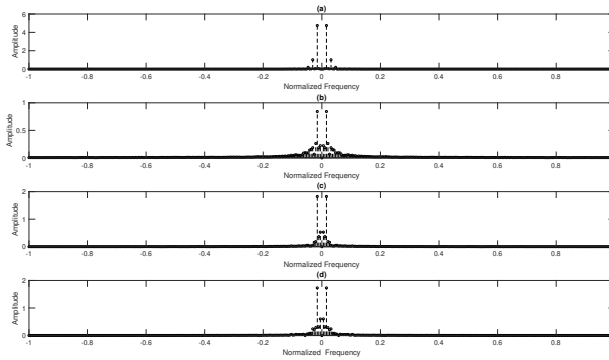


Figure 4: Power spectrum of (a) horse, (b) head, (c) trunk and (d) elbow cycle signals of Figure 3, using the centered FFT function.

200 The second step of cycle detection components aims to interpolate the different cycles using spline filters, to have the same signal length, as it can be seen from Figure 6. The cycle should be clustered according to the shape of the

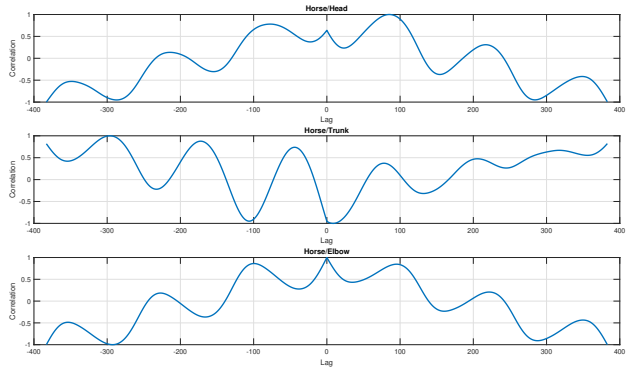


Figure 5: The correlation between horse/head, horse/trunk and horse/elbow.

pattern whatever their size is. Two similar shapes at different scale should be categorized in the same cluster.

205 Indeed, because oscillation frequencies are changing throughout the exercise, the duration of the cycles is variable. For each angle, we obtain a matrix $\text{Mat} \in \mathbb{R}^{N \times M}$ composed of all cycles from all the participants, where N is the total number of extracted cycles and M is the fixed length of the cycles. In the sequel, M is fixed to 128. Each row corresponds to a single cycle. The total
 210 number of cycles (N) for each angle is equal to 19387.

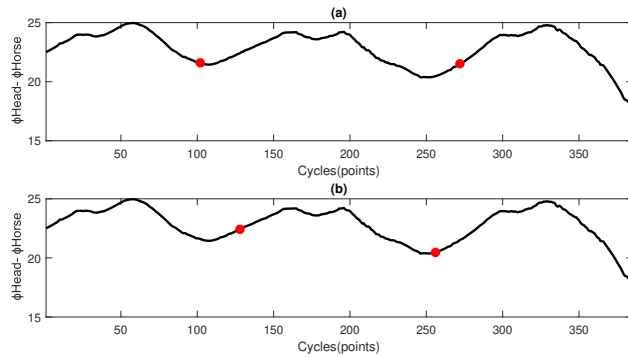


Figure 6: Cycle interpolation : (a) the top panel corresponds to the three cycles before interpolation. (b) The panel below corresponds to the three cycles after interpolation (a fixed length of 128 points).

2.2.2. Cycle clustering

In order to divide postural coordination into homogeneous clusters, K -means algorithm, one of most popular and simplest unsupervised techniques is used. Let $\{x_i\}_{i=1}^n$ to be the set of n d -dimensional points to be clustered into K clusters of $\{C_k\}_{k=1}^K$ centers. The aim of the K -means clustering is to minimize the objective function:

$$J = \sum_{k=1}^K \sum_{i=1}^n dist(x_i, C_k) \quad (2)$$

where $dist$ is a distance function. Two well known metrics are Euclidean distance [24] and Dynamic Time Warping (DTW) distance [31]. The first one is generic for continuous data x_i in which are deemed to be sampled according to a conditional d -dimensional Gaussian distribution for each cluster C_k . The second one is more suitable for measuring similarities between time series, therefore more adapted to the cycles. In our study, we compare the obtained clustering results using both Euclidean and DTW distances.

K -means with Euclidean distance:

It is based on the classical Euclidean distance defined as:

$$dist(x, C) = \|x - C\|_2^2 \quad (3)$$

where $\|x - C\|_2^2$ is the Euclidean distance between the data point x and the cluster center C . A set of cluster centers C_k is chosen randomly, and then it is updated repeatedly until convergence. Each update is followed by the assignment of the samples x_i to the closest cluster, according to (Eq.2). Formally the main steps of the procedure are as follows [5]:

- Assignment step: for all samples x_i , assign x_i to its closest cluster k such that $k = \operatorname{argmin}_{\ell=1, K} \|x_i - C_\ell\|_2^2$
- Cluster center update: $C_k = \frac{1}{n_k} \sum_{i \in \mathcal{C}_k} x_i$ for all $k = 1, K$ where \mathcal{C}_k stands for cluster k and n_k its cardinality.

***K*-means with DTW distance:**

One of the eminent techniques to search the best alignment that matches two time series is DTW [25]. It captures flexible similarities by aligning the coordinates inside both sequences. DTW adopts a generalization of Euclidean distance which allows a non-linear mapping of one time series to another by minimizing the distance between them. The distance $dist(x, C)$ between x and C is calculated as follows: first, we create an n -by- m matrix, where every (i, j) entry is equal to $\|x_i - C_j\|^2$. All the possible cumulative distances between the sequences are found for every possible path, with the aim of minimization of the mapping cost:

$$dist(x, C) = \min_{w \in P} \sqrt{\sum_{l=1}^L d_{w_l}} \quad (4)$$

where P is the set of all possible warping paths, w_l is the position (i, j) at the l -th observation of a warping path, and L is the length of the warping path.

In our task, cycles have the same length. In practice we use the soft-DTW, a differentiable version of DTW proposed in [10]. As a result, the center of a cluster is computed as the Fréchet Mean [44, 8], instead of of the classical mean, of time series. Specifically for cluster \mathcal{C}_k its center C_k is obtained by solving [10] $\min_C \frac{1}{n_k} \sum_{i \in \mathcal{C}_k} \text{Soft-DTW}(x_i, C)$ using classical gradient descent optimization tools [32]. Notice that each clustering assignment is done according to the Soft-DTW distance.

The cycle clustering was performed thanks to Python (version 3.5.0) software [41], and especially the scikit-learn (version 2.4.8) library [37]. Both algorithms are sensitive to parameters initialization and require to select the number of clusters.

3. Results

We analyse the postural coordination of the different subjects by clustering the extracted cycles into K homogeneous groups. For this sake, we apply K -means based respectively on Euclidean and DTW distances. In order to assess

the influence of simulated hippotherapy on the postural coordination, the clustering is applied on the discrete relative phase cycles of head, trunk and elbow separately. These segments are not merged into one cluster model, because the aim of this paper is to independently study the coordination between each of these variables with the horse.

In the remainder, we present how the cluster number K is selected in each experiment. The obtained results are then highlighted according to the cycles signals related to riders and non-riders assigned to a cluster in order to explore the behaviours shared by all subjects or specific to riders and non-riders. Finally, we illustrate the influence of horse oscillation frequency on the clustering outcomes.

3.1. Number of clusters

In order to find the optimal number of clusters, silhouette score analysis was performed [38]. This method has been commonly used in cluster analysis for finding the optimal number of clusters, as well as for clustering validation. The silhouette value is a measure of how close a sample x_i is to its own cluster compared to its neighboring clusters. Let assume x_i is assigned to cluster \mathcal{C}_k . The silhouette score is defined as:

$$s_i = \frac{b_i - w_i}{\max\{b_i, w_i\}} \quad (5)$$

$$b_i = \min_{\ell \neq k} \{B(i, \ell)\} \quad (6)$$

where w_i represents the average within distance from x_i to the samples in the same cluster \mathcal{C}_k . $B(i, \ell)$ represents the average distance between x_i and the samples of cluster $\mathcal{C}_\ell, \ell \neq k$ it does not belong to. The higher the silhouette value, better is the cluster configuration. The blue lines show the silhouette score for head, elbow and trunk in Figure 7.

The silhouette plot suggests an optimal of 5 clusters for trunk and elbow signals, and an optimal of 4 clusters for head signal. We use the same values of K proposed by the K -means with Euclidean distance for the postural coordination

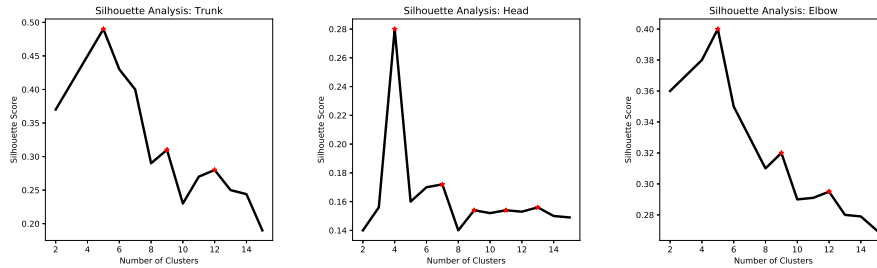


Figure 7: Selecting number of clusters by Silhouette Analysis: K -means with Euclidean distance. The red dots correspond to optimal cluster number.

clustering by K -means with DTW distance, to allow comparison between the results of the two algorithms.

3.2. Clustering elbow, trunk and head cycles

Here after, we report the results of the clustering of postural coordination, represented through the extracted cycles from the different angles (head, elbow and trunk). To facilitate the interpretation of the results, we report the cycles assigned to each cluster according to the level of the related subjects (riders and non-riders)

Table 2 summarizes the number of cycles, given by K -means with DTW distance and K -means with Euclidean distance for head signal in each cluster. The greatest number of head cycles is present in $K1$ for riders. $K1$ is also the most representative for non-riders but at a lower percentage compared to riders, for both clustering algorithms. The distribution of cycles between clusters for non-riders is relatively more balanced than that of riders, which means that head variability is lower for the expert riders compared to the non-riders. This finding is more emphasized for K -means with DTW distance where the distribution for $K1$ to $K4$ are respectively of 39, 16, 26 and 17% approximately. Moreover, head cycles for riders are distributed among only 3 clusters ($K1$, $K2$, $K3$). For both algorithms, few samples appears to be in $K4$. Most notably, all the $K4$ riders samples belongs to head data are only from two riders.

Table 2: Head cycles distribution between the different clusters, expressed in (%): K -means with Euclidean distance vs K -means with DTW distance. Clusters for both algorithms are named: $K1$, $K2$, $K3$ and $K4$.

	Clusters	$K1$	$K2$	$K3$	$K4$
K -means with Euclidean distance	Riders	54.85	20.93	15.25	8.97
	Non-riders	45.59	23.34	20.23	10.84
K -means with DTW distance	Riders	49.0	16.75	26.14	8.08
	Non-riders	39.22	16.21	26.66	17.89

The distribution of elbow cycles are highlighted in Table 3. For riders, elbow cycles are distributed mainly among three clusters ($K2$, $K3$ and $K5$), and this for K -means with Euclidean distance and in ($K2$, $K3$ and $K4$) for K -means with DTW distance. It is pertinent to mention that cluster $K1$ in K -means with Euclidean distance contains elbow data only from two riders. For K -means with DTW distance, $K5$ contains only cycles from one rider. For non-riders, the distribution of cycles over the five clusters is more dynamic with all the clusters represented for K -means with DTW distance and K -means with Euclidean distance.

The distribution of trunk cycles between the different clusters are highlighted in Table 4. For K -means with Euclidean distance, trunk cycles are mostly presented in ($K1$, $K3$, $K4$ and $K5$) for riders. For non-riders, the trunk cycles are more distributed between the 5 clusters. K -means with DTW distance induces the same behaviour as K -means with Euclidean distance for riders and non-riders.

Thus, the distribution of the cycles for the angles of head and elbow are distributed between three clusters. Trunk cycles, are divided into 4 clusters. On the other hand, non-riders display more diverse behaviours entering all clusters. The curves of the different cluster centers for elbow, trunk and head are drawn in Figures 8 and 9. The curves of different clusters are well separated for both algorithms. On the basis of current experiments, we are not able to evaluate or compare the two algorithms in the context of unsupervised clustering. K -means

Table 3: Elbow cycles distribution between the different clusters, expressed in (%): K -means with Euclidean distance vs K -means with DTW distance. Clusters for both algorithms are named: $K1$, $K2$, $K3$, $K4$ and $K5$.

	Clusters	$K1$	$K2$	$K3$	$K4$	$K5$
K -means with Euclidean distance	Riders	14.83	28.65	20.89	1.89	33.72
	Non-riders	15.58	24.11	15.79	6.78	37.72
K -means with DTW distance	Riders	1.49	51.42	25.67	16.67	4.72
	Non-riders	13.66	40.11	18.65	22.71	4.82

Table 4: Trunk cycles distribution between the different clusters, expressed in (%): K -means with Euclidean distance vs K -means with DTW distance. Clusters for both algorithms are named: $K1$, $K2$, $K3$, $K4$ and $K5$.

	Clusters	$K1$	$K2$	$K3$	$K4$	$K5$
K -means with Euclidean distance	Riders	10.75	0.58	37.19	40.9	10.55
	Non-riders	18.81	10.25	29.12	24.81	16.98
K -means with DTW distance	Riders	37.26	0	11.07	10.7	40.86
	Non-riders	29.14	10.26	17.02	18.75	24.81

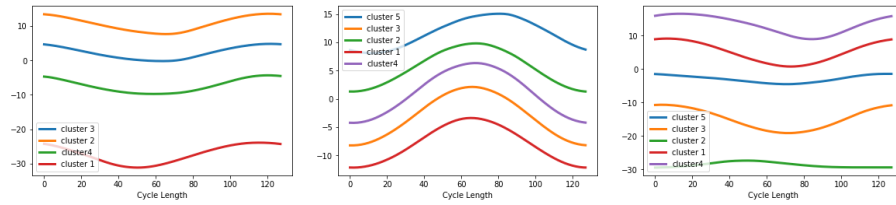


Figure 8: Cluster centers curves obtained with K -means with Euclidean distance. From left to right: head, trunk and elbow.

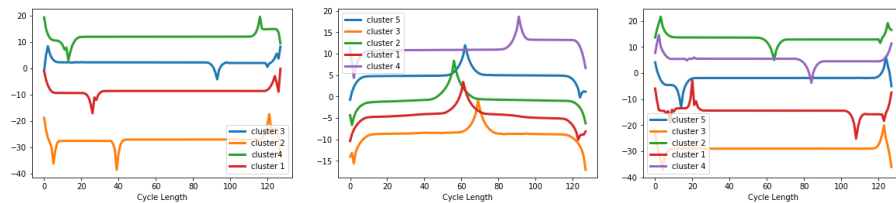


Figure 9: Cluster centers curves obtained with K -means with DTW distance. From left to right: head, trunk and elbow.

with DTW distance appears to identify clearly the different cluster centers. This is the case for trunk cycles (middle panel of Figure 9). In addition, since no
 335 work previously exists that performs clustering on postural coordination, we decided to choose K -means with DTW distance for the rest of the experiments.

3.3. Impact of horse oscillation frequencies

The impact of increasing horse oscillation frequencies on the clustering for
 340 head, elbow and trunk can be highlighted in Tables 5, 6 and 7. This analysis aims to see the temporal representation of the different clusters and also to understand how the different subjects control their oscillation depending on the horse oscillation frequency ($V0$, $V1$, $V2$ and $V3$). Each column represents the distribution of each cluster at the different frequencies. The sum of the values
 345 at each column is equal to 100%. Each row illustrates the distribution of the cycles at each frequency through the clusters.

Table 5 shows the distribution of the different clusters for head over the

Table 5: Impact of oscillation frequencies on cluster distribution, expressed in (%), for head.

Riders					
Frequency	Cluster	<i>K1</i>	<i>K2</i>	<i>K3</i>	<i>K4</i>
	<i>V0</i>		7.09	1.26	1.59
<i>V1</i>		35.30	4.85	11.49	39.29
<i>V2</i>		34.98	16.87	34.79	8.64
<i>V3</i>		22.62	76.98	51.76	0.67
Non-riders					
Frequency	Cluster	<i>K1</i>	<i>K2</i>	<i>K3</i>	<i>K4</i>
	<i>V0</i>		3.03	0.89	1.67
<i>V1</i>		33.17	2.29	11.26	44.77
<i>V2</i>		40.64	11.18	31.38	17.29
<i>V3</i>		22.98	85.59	55.64	4.84

different frequencies. Riders' heads oscillated weakly at *V0* (*K1*, 7%) and moved quickly at *V1*, *V2* and at *V3*. Except at the first frequency (*V0*), the riders followed a tri-dimensional behaviour. *K4*, highlighted as blue, has not been taken into consideration as it's poorly represented among riders (only 8%, Table 2), and that's just for two riders. The non-riders have a more stable head at *V0*: the presence of the *K4* cluster (33.04%), which is related to a distinct type of postural coordination created from the first oscillations of the horse, is clearly distinguished. However, the non-riders have then increased their variability at *V1*, *V2* and at *V3*.

Table 6 shows the distribution of the five clusters for elbow over the different frequencies. A significant effect of frequency is found for elbow segment variability: the elbow variability increases when the frequency increases especially for non-riders at *V3* (*K4*: 60.41%, *K5*: 75.25%). The riders have a more stable elbow movement represented by 3 different patterns at all frequencies.

Table 6: Impact of oscillation frequencies on cluster distribution, expressed in (%), for elbow.

Riders						
Frequency \ Cluster	Cluster	<i>K1</i>	<i>K2</i>	<i>K3</i>	<i>K4</i>	<i>K5</i>
	<i>V0</i>		0.06	12.34	4.91	4.52
<i>V1</i>		2.31	26.84	30.57	15.70	13.63
<i>V2</i>		53.21	31.61	26.59	20.74	51.09
<i>V3</i>		44.41	29.20	37.92	59.03	34.89

Non-riders						
Frequency \ Cluster	Cluster	<i>K1</i>	<i>K2</i>	<i>K3</i>	<i>K4</i>	<i>K5</i>
	<i>V0</i>		7.39	12.27	6.04	4.26
<i>V1</i>		25.69	30.05	32.51	12.48	3.45
<i>V2</i>		26.42	31.24	35.84	22.84	20.77
<i>V3</i>		40.43	26.41	25.59	60.41	75.28

Table 7: Impact of oscillation frequencies on cluster distribution, expressed in (%), for trunk.

Riders						
Frequency \ Cluster	K1	K2	K3	K4	K5	
	V0	19.45	-	0	44.33	1.04
V1	38.81	-	0	22.70	19.53	
V2	18.63	-	3.06	16.54	35.93	
V3	23.09	-	96.9	16.41	43.48	

Non-riders						
Frequency \ Cluster	K1	K2	K3	K4	K5	
	V0	7.54	0.31	37.84	41.98	2.98
V1	35.71	0.22	28.21	2.66	14.54	
V2	27.64	0.69	33.01	1.37	54.44	
V3	29.09	98.75	0.94	53.97	28.08	

Table 7 shows the distribution of the five clusters for trunk over the different frequencies. At low frequencies ($V0$), riders adopt a bi-dimensional behaviour ($K1$ - $K4$) but at higher frequencies ($V1$, $V2$ and $V3$), the number of coordination increases, reaching four different patterns at $V3$. For both groups, a new pattern (i.e the emergence of a new coordination) appears at $V3$ ($K3$ for riders and $K2$ for non-riders). As well, the trunk variability increases when the frequency increases for non-riders at all frequencies. To further illustrate the behavioral difference between riders and non-riders, Figures 10 and 11 provide an example of the temporal representations of the clusters for two subjects.

No clear movement was observed from head, trunk and elbow for the rider, resulting in no noticeable clustering changes. For the non-rider, the distribution of cycles is more dynamic. He displayed more diverse behaviours. Furthermore, the variability increases when the frequency increases, especially for the non-rider. But in both cases, we observe that the postural patterns are not constant

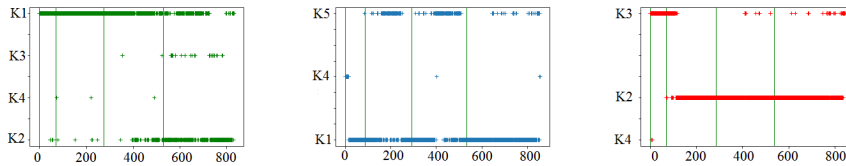


Figure 10: Impact of oscillation frequencies on cluster distribution for a rider subject. From left to right: head, trunk and elbow. The vertical bars represent the change in horse oscillation frequencies. The abscissa corresponds to the number of cycles and the ordinate to the cluster name.

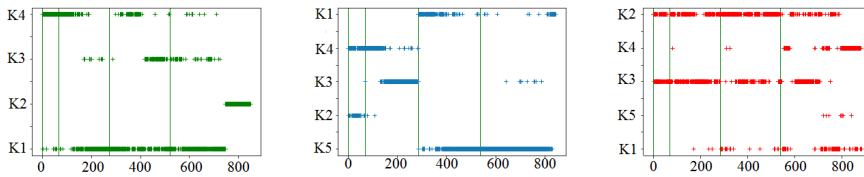


Figure 11: Impact of oscillation frequencies on cluster distribution for a non-rider subject. From left to right: head, trunk and elbow. The vertical bars represent the change in horse oscillation frequencies. The abscissa corresponds to the number of cycles and the ordinate to the cluster name.

even in a short period of time.

4. Discussion

The main issues addressed in this study aimed to explore the differences between the postural changes of riders and non-riders after a mechanical horse riding by using an unsupervised clustering of their postural coordination and to analyze the impact of horse oscillation frequencies on the postural dynamics for both of them.

Riders produce lower postural displacements and deploy more efficient postural control by following a tri-dimensional behaviour for head and elbow. At the opposite, for non-riders, the postural coordination was distributed between more than four clusters, showing a greater dynamics. The postural coordination differed between the two groups, as already confirmed by Baillet et al. [3]. In fact, an optimally adapted postural coordination is influenced by the level of expertise [18]. The paper endorsed this assumption. Riders' postural adaptations also confirmed the findings of Byström [6]. The authors demonstrated that practice and training allow riders to better control and instantly adapt their postural changes.

A student's test for unpaired data was then used to compare postural coordinations between riders and non-riders and validate the results that are obtained. The t -test showed a p -value <0.001 for the three different angles, resulting in a significant difference between the two groups.

The study also showed changes in the postural dynamics with the increasing oscillation frequencies. Indeed, all subjects have increased their ability to master the activity at higher oscillation frequencies and have presented different patterns over time. Furthermore, we can see that the frequency changes affects more the non-rider posture than rider posture for elbow and trunk segments. The main reason behind this observation could be explained by the fact that the group of riders under experimentation, were expert riders. Thus, their expertise level permits a more controlled and predictable behaviour.

405 The non-riders' stability was only better than that of riders for head, thus
confirming the results obtained by Baillet et al. [3]. This finding suggests that
postural stability at low frequency of the head is not a signature of expertise in
horse back riding. One possible explanation for this result might be that the
variability of the mechanical horse is more predictable compared to a real horse,
410 and that the non-riders can adapt more easily their heads.

Several works have investigated cluster analysis in interpreting postural co-
ordination: for healthy people after 6-week training program of supervised slack-
line exercise [39], for children with unilateral cerebral palsy [12], and for children
with and without developmental coordination disorder [23].

415 To the best of our knowledge, cluster analysis has not been used in the
investigation of postural coordination in sports. This is the first study that aims
to understand postural coordination on the mechanical horse through cluster
analysis. Since we have no previous results, this work can be used to establish
benchmarks for future comparisons.

420 Finding specific behaviours (clusters) in the postural coordination validates the
previous studies, based on statistical analysis and quality measures.

The conclusions drawn in this paper provide new insights into horseback
riding assets and methodological clues to assess the impact of sport practice.
The comparison of the coordination between riders and non-riders helps us to
425 understand the interaction of athletes according to their level of expertise. This
can be explored in several future applications. Design a neck brace or knee
brace for non-riders to help them for adjusting their body similar to riders or
help athletes to improve their posture is of great interest.

430 However, some limitations should be noted. First, the comparison was be-
tween riders and non-riders, both associated with strong background in sports,
which may bias the comparison results between the two groups. Second, we
used a mechanical horse. As its behaviour is more predictable and stable than
a real horse, this may also affect the results. Third, only women participants
435 have been examined. This selection has been done to prevent a possible bias

related to the impact of gender on postural coordination.

5. Conclusion

The present paper investigated the impact of mechanical horse movements on the postural coordination of 22 sportswomen, constituting two distinct groups according to expertise level: riders and non-riders. An unsupervised learning with no previous information about the level of rider expertise was performed on the data, namely clustering by K -means with Euclidean distance and K -means with DTW distance. Analyzing the temporal distribution of the clusters, we clearly saw that riders developed postural patterns different from the non-riders, allowing a more adapted posture according to the increase of oscillation frequencies.

Note that this work remains descriptive and thus preliminary, but one can consider it as a start for a further predictive analysis. We call to mind that all clustering algorithms are subjective concerning how they involve the grouping of data points or which metric they use. One should consider as good results any helpful perspective hidden in, the raw data. Thus, we worked closely with sport experts, to analyse and improve the results, by using the continuous relative phase to evaluate the postural coordination.

Although it is well known that exercise-induced fatigue may be responsible for a decrease in performance, its effects on postural coordination for athletes were not explored in our study. The response to fatigue in terms of postural coordinate, level of expertise (rider vs non-rider) and time will constitute a possible direction for future work.

Also, the present paper will be further exploited to develop a consistent and safe riding procedure for sportsmen and also for people with disabilities. The machine learning aspect of this study will be considered for improvement. The results showed that even for a short period of time, the postural strategies were not stable. We think this is related to the choice of the clustering algorithm. The results obtained for this research work will be used as a starting

465 point for developing a new clustering algorithm based on Total Variation (TV)
regularization.

Conflict of Interest

The authors declare no competing interests.

Funding Information

470 This work was supported by the DAISI project, co-funded by the European
Union with the European Regional Development Fund (ERDF) and by the
Regional Council of Normandie.

References

- [1] S. Aoi, T. Kondo, N. Hayashi, D. Yanagihara, S. Aoki, H. Yamaura, N. Ogi-
475 hara, T. Funato, N. Tomita, K. Senda, et al., Contributions of phase reset-
ting and interlimb coordination to the adaptive control of hindlimb obstacle
avoidance during locomotion in rats: a simulation study, *Biological cyber-
netics* 107 (2) (2013) 201–216.
- [2] D. Avitabile, P. Słowiński, B. Bardy, K. Tsaneva-Atanasova, Beyond in-
480 phase and anti-phase coordination in a model of joint action, *Biological
cybernetics* 110 (2-3) (2016) 201–216.
- [3] H. Baillet, R. Thouwarecq, E. Vérin, C. Tourny, N. Benguigui, J. Komar,
D. Leroy, Human Energy Expenditure and Postural Coordination on the
Mechanical Horse, *Journal of motor behavior* 49 (4) (2017) 441–457.
- 485 [4] F. Beinotti, N. Correia, G. Christofolletti, G. Borges, Use of hippotherapy
in gait training for hemiparetic post-stroke, *Arquivos de neuro-psiquiatria*
68 (6) (2010) 908–913.
- [5] C. M. Bishop, *Pattern recognition, Machine learning* 128 (9).

- 490 [6] A. Byström, The movement pattern of horse and rider in different degrees of collection, Ph.D. thesis, Department of Anatomy, Biochemistry and Physiology, Swedish University of Agricultural Sciences, 2019.
- [7] B. Carse, B. Meadows, R. Bowers, P. Rowe, Affordable clinical gait analysis: An assessment of the marker tracking accuracy of a new low-cost optical 3D motion analysis system, *Physiotherapy* 99 (4) (2013) 347–351.
- 495 [8] A. D. Chouakria, P. N. Nagabhushan, Adaptive dissimilarity index for measuring time series proximity, *Advances in Data Analysis and Classification* 1 (1) (2007) 5–21.
- [9] T. J. Cunningham, The Clinical Usefulness of Vector Coding Variability in Female Runners With and Without Patellofemoral Pain, Ph.D. thesis, University of Kentucky, 2012.
- 500 [10] M. Cuturi, M. Blondel, Soft-DTW: a differentiable loss function for time-series, in: *Proceedings of the 34th International Conference on Machine Learning-Volume 70*, JMLR. org, 894–903, 2017.
- [11] T. A. Dierks, I. Davis, Discrete and continuous joint coupling relationships in uninjured recreational runners, *Clinical Biomechanics* 22 (5) (2007) 581–591.
- 505 [12] M. Domagalska-Szopa, A. Szopa, Postural pattern recognition in children with unilateral cerebral palsy, *Therapeutics and clinical risk management* 10 (2014) 113.
- 510 [13] J. G. Dominguez-Romero, A. Molina-Aroca, J. A. Moral-Munoz, C. Luque-Moreno, D. Lucena-Anton, Effectiveness of Mechanical Horse-Riding Simulators on Postural Balance in Neurological Rehabilitation: Systematic Review and Meta-Analysis, *International Journal of Environmental Research and Public Health* 17 (1) (2020) 165.

- 515 [14] M. A. Elshafey, Hippotherapy simulator as alternative method for hippo-
therapy treatment in hemiplegic children, *Int. J. Physiother. Res* 2 (2)
(2014) 435–441.
- [15] M. Engell, H. Clayton, A. Egenvall, M. A. Weishaupt, L. Roepstorff, Pos-
tural changes and their effects in elite riders when actively influencing the
520 horse versus sitting passively at trot, *Comparative Exercise Physiology*
12 (1) (2016) 27–33.
- [16] C. Garcia, J. A. Barela, A. R. Viana, A. M. F. Barela, Influence of gymnastics
training on the development of postural control, *Neuroscience letters*
492 (1) (2011) 29–32.
- 525 [17] R. Guo, J. Chen, L. Wang, Hierarchical K-means clustering for registration
of multi-view point sets, *Computers & Electrical Engineering* 94 (2021)
107321.
- [18] R. Herault, D. Orth, L. Seifert, J. Boulanger, J. A. Lee, Comparing dy-
namics of fluency and inter-limb coordination in climbing activities using
530 multi-scale Jensen–Shannon embedding and clustering, *Data Mining and*
Knowledge Discovery 31 (6) (2017) 1758–1792.
- [19] P. Herrero, Á. Asensio, E. García, Á. Marco, B. Oliván, A. Ibarz, E. M.
Gómez-Trullén, R. Casas, Study of the therapeutic effects of an advanced
hippotherapy simulator in children with cerebral palsy: a randomised con-
535 trolled trial, *BMC musculoskeletal disorders* 11 (1) (2010) 71.
- [20] C. Hilliere, D. Collado-Mateo, S. Villafaina, P. Duque-Fonseca, J. A.
Parraga, Benefits of hippotherapy and horse riding simulation exercise on
healthy older adults: A systematic review, *PM&R* 10 (10) (2018) 1062–
1072.
- 540 [21] K. Howerton, A Comparison of Postural Stability in Gymnasts, Volleyball
Players, and Non-Athletes, *Ursidae: The Undergraduate Research Journal*
at the University of Northern Colorado 2 (2) (2012) 12.

- [22] A. K. Jain, Data clustering: 50 years beyond K-means, *Pattern recognition letters* 31 (8) (2010) 651–666.
- 545 [23] L. Jelsma, R. Geuze, B. Smits-Engelsman, Movement Control Strategies in a Dynamic Balance Task in Children With and Without Developmental Coordination Disorder, *Journal of motor behavior* (2019) 1–12.
- [24] T. Kanungo, D. M. Mount, N. S. Netanyahu, C. D. Piatko, R. Silverman, A. Y. Wu, A local search approximation algorithm for k-means clustering, 550 *Computational Geometry* 28 (2-3) (2004) 89–112.
- [25] R. J. Kate, Using dynamic time warping distances as features for improved time series classification, *Data Mining and Knowledge Discovery* 30 (2) (2016) 283–312.
- [26] P. F. Lamb, M. Stöckl, On the use of continuous relative phase: Review of 555 current approaches and outline for a new standard, *Clinical Biomechanics* 29 (5) (2014) 484–493.
- [27] C.-W. Lee, S. G. Kim, S. S. Na, The effects of hippotherapy and a horse riding simulator on the balance of children with cerebral palsy, *Journal of physical therapy science* 26 (3) (2014) 423–425.
- 560 [28] S. Lloyd, Least squares quantization in PCM, *IEEE transactions on information theory* 28 (2) (1982) 129–137.
- [29] K. M. Menezes, F. Copetti, M. J. Wiest, C. M. Trevisan, A. F. Silveira, Effect of hippotherapy on the postural stability of patients with multiple sclerosis: a preliminary study, *Fisioterapia e Pesquisa* 20 (1) (2013) 43–49.
- 565 [30] A. G. Moraes, F. Copetti, V. R. Angelo, L. L. Chiavoloni, A. C. David, The effects of hippotherapy on postural balance and functional ability in children with cerebral palsy, *Journal of physical therapy science* 28 (8) (2016) 2220–2226.

- [31] G. Moriello, M. E. Terpstra, J. Earl, Outcomes following physical therapy
570 incorporating hippotherapy on neuromotor function and bladder control
in children with Down syndrome: A case series, *Physical & occupational
therapy in pediatrics* (2019) 1–14.
- [32] J. Nocedal, S. J. Wright, *Numerical Optimization*, Springer, New York,
NY, USA, second edn., 2006.
- 575 [33] G. W. Oehlert, *A first course in design and analysis of experiments*, 2010.
- [34] A. Olivier, F. Bonneau, J. Jevvrey, B. Isableu, et al., Rider-horse interac-
tion during transitions of paces: postural coordination of the rider accord-
ing to the level of expertise., 41ème journée de la recherche équine, jeudi
12 mars 2015, Paris, France (2015) 154–157.
- 580 [35] O. Oullier, B. G. Bardy, T. A. Stoffregen, R. J. Bootsma, Postural coor-
dination in looking and tracking tasks, *Human Movement Science* 21 (2)
(2002) 147–167.
- [36] J.-H. Park, J. S. H. You, Innovative robotic hippotherapy improves postural
muscle size and postural stability during the quiet stance and gait initiation
585 in a child with cerebral palsy: A single case study, *NeuroRehabilitation*
42 (2) (2018) 247–253.
- [37] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel,
M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, et al., Scikit-learn:
Machine learning in Python, *Journal of machine learning research* 12 (Oct)
590 (2011) 2825–2830.
- [38] P. J. Rousseeuw, Silhouettes: A graphical aid to the interpretation and
validation of cluster analysis, *Journal of Computational and Applied Math-
ematics* 20 (1987) 53 – 65.
- [39] B. Serrien, E. Hohenauer, R. Clijsen, W. Taube, J.-P. Baeyens, U. Küng,
595 Changes in balance coordination and transfer to an unlearned balance task

after slackline training: a self-organizing map analysis, *Experimental brain research* 235 (11) (2017) 3427–3436.

- [40] T. L. Shurtleff, J. W. Standeven, J. R. Engsborg, Changes in dynamic trunk/head stability and functional reach after hippotherapy, *Archives of physical medicine and rehabilitation* 90 (7) (2009) 1185–1195. 600
- [41] G. Van Rossum, F. L. Drake, *Python 3 Reference Manual*, CreateSpace, Scotts Valley, CA, ISBN 1441412697, 2009.
- [42] V. Vermöhlen, P. Schiller, S. Schickendantz, M. Drache, S. Hussack, A. Gerber-Grote, D. Pöhlau, Hippotherapy for patients with multiple sclerosis: A multicenter randomized controlled trial (MS-HIPPO), *Multiple Sclerosis Journal* 24 (10) (2018) 1375–1382. 605
- [43] R. Xu, I. Donald Wunsch, Survey of Clustering Algorithms, *IEEE Transactions on Neural Networks* 16 (3).
- [44] L. Zhuang, A. T. Walden, Sample mean versus sample Fréchet mean for combining complex Wishart matrices: A statistical study, *IEEE Transactions on Signal Processing* 65 (17) (2017) 4551–4561. 610

Appendix

Algorithm 1: Head Cycle detection

Input: Horse data matrix $MatHorse$, Head data matrix $MatHead$, for one Subject P

Output: Horse cycle matrix $CycleHorse$, Head cycle matrix $CycleHead$.

```
; /* PeakDetect:find the indices  $Loc$  at which the peaks  $Peak$ 
occur. */
; /* The peaks are separated by more than the minimum peak
distance  $d$ . */
1 Function PeakDetect( $MatHorse, d$ ):
2   return  $Peak, Loc$  ;
; /* cycledetect: divides  $MatHead$  into adjacent
sub-matrices of different lengths. */
3 Function CycleDetect( $Loc, MatHead$ ):
4   for  $i \leftarrow 2$  to  $length(Loc)$  do
5      $Inter(i) = Loc(i) - Loc(i - 1)$ 
6    $Inter(1) = loc(1)$ 
; /* mat2cell : Convert array to cell array */
7    $CycleHead = mat2cell(MatHead, Inter)$ 
8   return  $CycleHead$  ;
```

Algorithm 2: K-means Cycle Clustering

Input: data matrix X

Output: Clusters C_1, \dots, C_k

1 Let C_1, \dots, C_k be the initial cluster centers.

2 **repeat**

3 For each x_i in X , assign it to the closest cluster C_k of center C_k ;
 /* by euclidean distance or by Soft-DTW distance */

4 For each cluster C_k , update its center C_k by averaging all cycles x_i
 that have been assigned to it.

5 **until** *convergence*;

6 **return** C_1, \dots, C_k ;
