Séminaire DataShape

Gait analysis with Inertial Measurement Units: from signal processing approaches to topological data analysis

Laurent Oudre
laurent.oudre@ens-paris-saclay.fr

June, 15th 2023

Contents

The Centre Borelli

The Centre Borelli

Fusion of two labs :

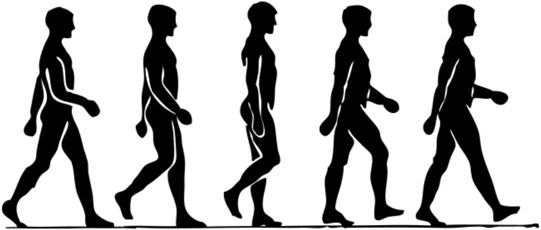
- The Centre de mathématiques et de leurs applications (CMLA) : applied mathematics for the study of complex phenomena and data
- The Cognition & Action Group (CognacG) : quantification and study of human and animal behavior

Main scientific questions

How to quantify the human behavior

- Adventure launched since 2012 : interdisciplinary collaboration between mathematicians, physicians, neuroscientists, engineers, biologists, etc...
- Implementation of measurement chains “pipelines”, platforms and intelligent tools but also of procedures for analysis, measurement and processing of data
- Creation of tools for diagnostic assistance, inter-individual comparison and longitudinal follow-up
- Integration into a clinical environment and interaction between algorithms and medical/neuroscience experts

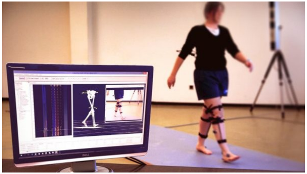
Study of locomotion



Why study the locomotion ?

- Most common dynamic human activity
- Reveals a large number of neurological, orthopedic, rheumatological disorders...
- Strong influence on daily life: risk of falling, frailty, mobility, dependence...

Study of locomotion



How to study locomotion?

- Historically : clinical examination by the physician, functional tests, clinical questionnaires

+	Ease of execution, clinician expertise
-	Lack of precision, difficult to compare two sessions

- Platforms for studying locomotion : instrumented mats, video/optical systems

+	Very precise, extraction of a large number of parameters, objective quantification
-	High cost, difficult to implement

General principles

★ Objective quantification of locomotion

→ Use of sensors and physiological measures

★ Longitudinal follow-up and inter-individual comparison

→ Need for a fixed protocol

★ Experimentation on the field

→ Lightweight sensors and fully automatic device for consultation and routine use

★ Clean and quality data

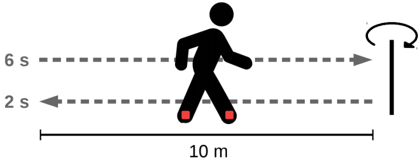
→ Control of the whole measurement chain, robust and reproducible algorithms

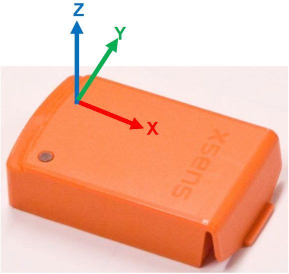
★ Willingness to capture clinician expertise

→ Clinical annotations and metadata

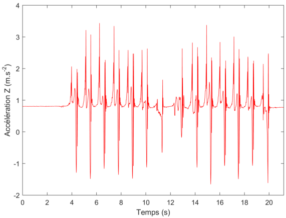
Protocol and sensors

- Comfort speed protocol: stop (6 sec), walk forward (10 m), turn around, return, stop
- Four wireless inertial units: left foot, right foot, lower back, head
- Nine signals per sensor : linear acceleration (3D), angular velocity (3D), magnetic field (3D)

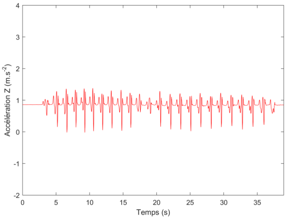




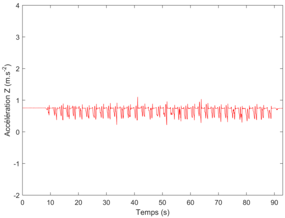
Signals



Healthy subject



Mild neurological disease



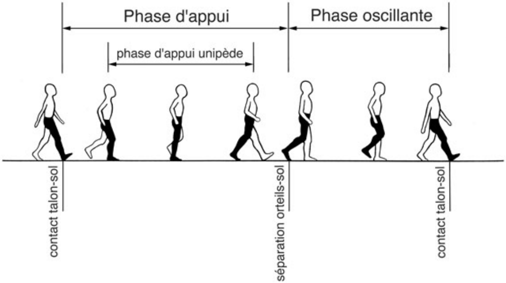
Severe neurological disease

Repetitive shapes

$\rightarrow$ How to detect the steps from accelerometer signals ?

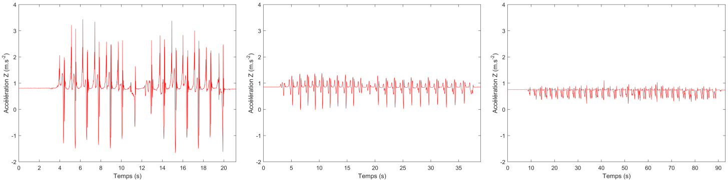
Contents

Step detection



- Locomotion is a structured activity composed of different phases (swing phase, stance phase)
- The basis *atom* for studying gait is the step
- How can we detect the start and end times of steps ?
- Useful to compute several features: number of steps, regularity, cadence, double stance times, etc.

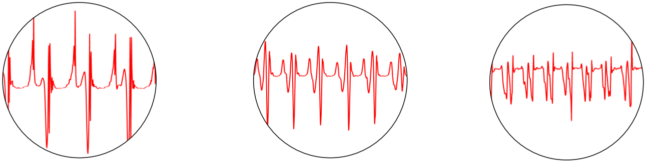
Easy problem ?



Sujet sain

Pathologie neurologique
peu sévère

Pathologie neurologique
sévère



Different durations, amplitudes and shapes

Scientific question

How to detect steps accurately and robustly on heterogeneous populations (healthy subjects, elderly, stroke, multiple sclerosis, Parkinson...) ?

Standard approaches are based on signal processing : filtering, peak detection...

$\rightarrow$ Tedious parameter tuning and most models are only adapted to one specific population (often young healthy subjects)

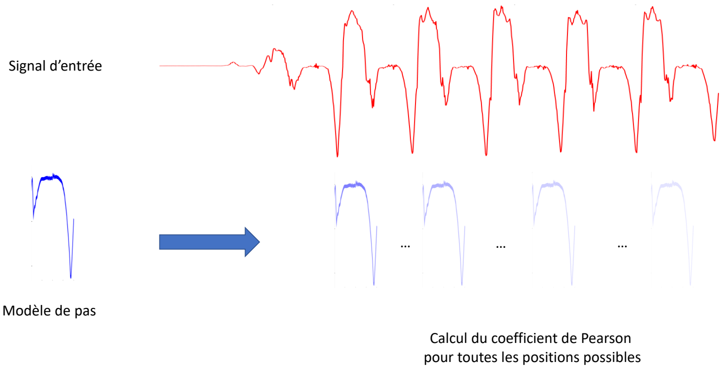
A few approaches use templates that are previously learned on the subject

$\rightarrow$ Difficult to deploy to general population and unable to adapt to possible changes of behavior (degenerative diseases, aging,...)

Proposed method

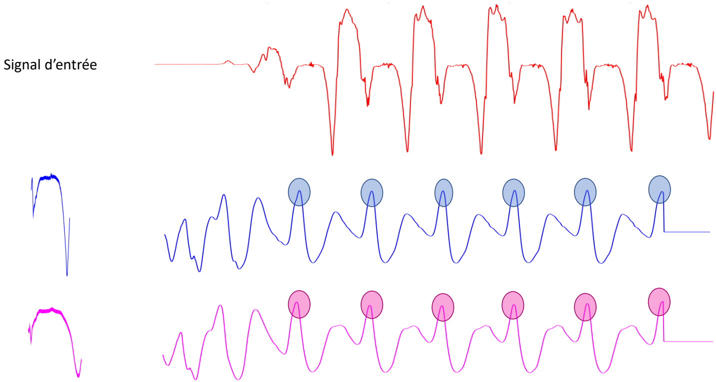
- Hypotheses
  - Robustness is crucial
    $\rightarrow$ One universal algorithm for all populations avoiding tuning parameters
  - Detection of all steps within the protocol
    $\rightarrow$ Must include U-turn, initiation and termination steps
- Proposition
  - Use pattern recognition techniques to avoid tuning parameters
  - Construction of a library of patterns composed of steps of various durations, amplitudes and shapes: we rely on several models and not just one
  - All models are tested simultaneously and the detection is done in a greedy way by choosing the best model for each detected step

Principle



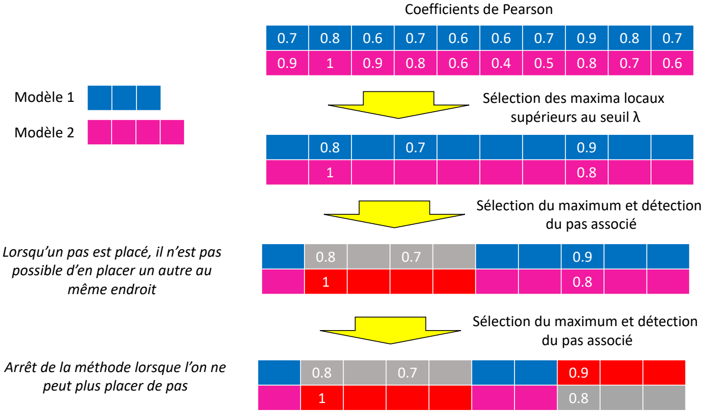
- Use of models extracted from real annotated steps
- Sliding computation of the Pearson correlation coefficients: independence with respect to the amplitude

Principle



- Several models of steps: different pathologies, gait styles, etc.
- Each model can provide a different step detection (see local maxima): which model should we choose ?

Principle



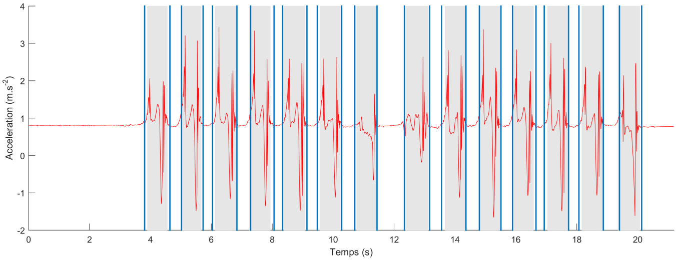
- Greedy procedure: each step is detected with the best available model
- Two steps cannot overlap

Results

	Proposed method		Pan-Tompkins	
Group	**Recall**	**Precision**	**Recall**	**Precision**
Healthy subject	**99.31 (1.75)**	**99.13 (1.86)**	99.14 (1.71)	97.09 (3.60)
Orthopedics pathologies	97.64 (2.73)	**98.20 (3.93)**	**98.78 (2.09)**	94.87 (5.09)
Neurological pathologies	**98.23 (3.42)**	**97.98 (3.33)**	96.80 (3.52)	95.49 (4.55)
Total	**98.34 (3.00)**	**98.30 (3.25)**	97.82 (3.07)	95.72 (4.56)

Results on 1020 recordings from different pathologies

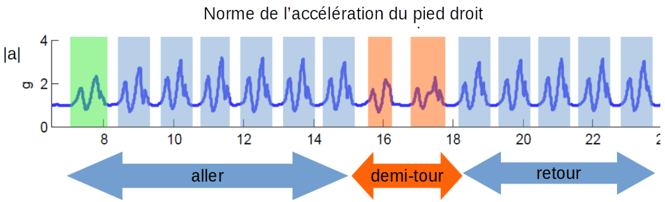
Results



- Precision and recall over ≥98% for all cohorts
- Robustness with respect to the pathology and to the type of step (U-turn, etc.)
- Possibility to adapt to another particular cohort by modifying the step library

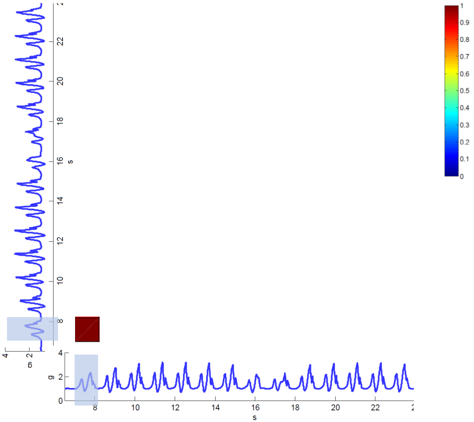
Contents

From features to visual feedbacks



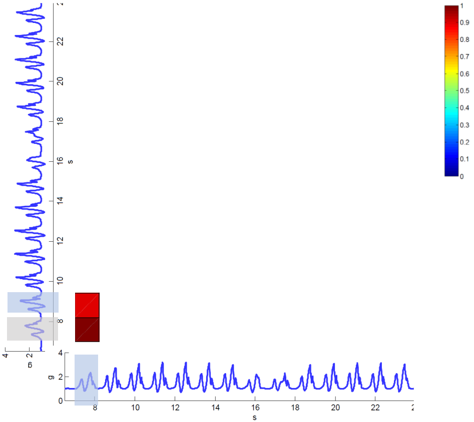
- Step detection allows to extract numerous useful features
- However, in most works, those are aggregated along the whole protocol, despite the fact that a great variability can occur
- How can we translate this variability into a condensed and ergonomic fashion ?

The locogram



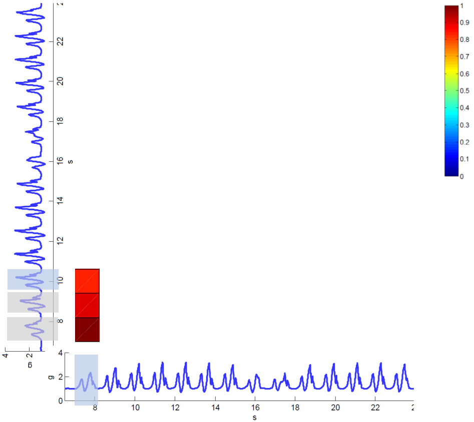
- Visualization of the whole gait exercise through a colored similarity matrix
- Use of the detected steps and comparison 2 by 2 through a correlation coefficient
- Each coefficient is associated to a color for easy visual inspection

The locogram



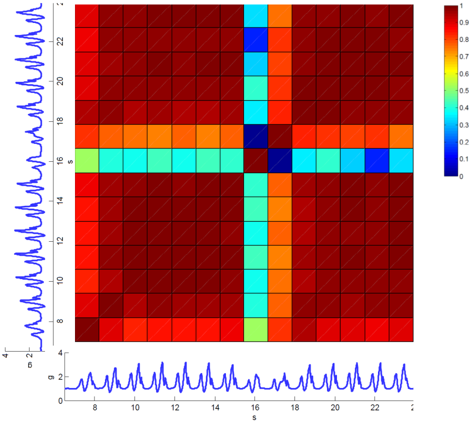
- Visualization of the whole gait exercise through a colored similarity matrix
- Use of the detected steps and comparison 2 by 2 through a correlation coefficient
- Each coefficient is associated to a color for easy visual inspection

The locogram



- Visualization of the whole gait exercise through a colored similarity matrix
- Use of the detected steps and comparison 2 by 2 through a correlation coefficient
- Each coefficient is associated to a color for easy visual inspection

The locogram



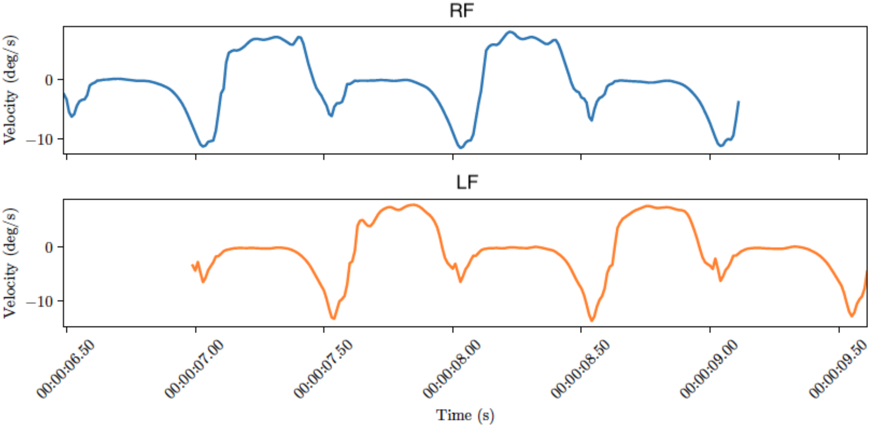
- Visualization of the whole gait exercise through a colored similarity matrix
- Use of the detected steps and comparison 2 by 2 through a correlation coefficient
- Each coefficient is associated to a color for easy visual inspection

Contents

Inter-individual comparison and longitudinal follow-up

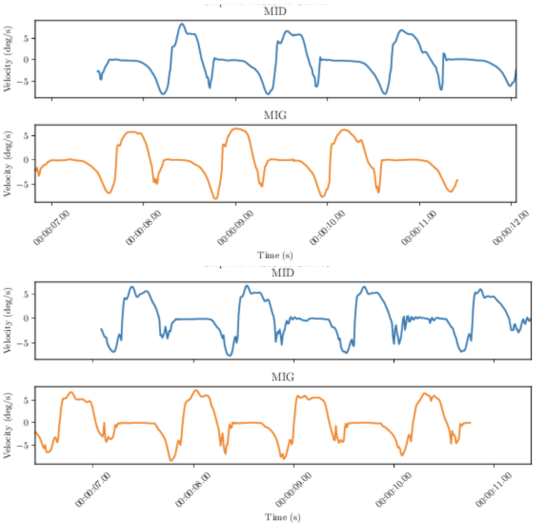
- Once we have extracted features and visual feedbacks, the main question is: how can we compare two gait signals ?
- Two aims: inter-individual comparison and longitudinal follow-up
- Larger question : could we do that from the raw data and without any processing step ?
- Could we somehow create a notion of *distance* between gait exercises ?
- Dataset : 22 MS patients, 10 healthy subjects. 2 trials at M0, 2 at M6 $\longrightarrow$ 16 signals per subject

Gait signals: healthy subject



Angular velocity in Y-axis on right and left foot. Note that a segmentation algorithm has been used to only consider the *Walk* period

Gait signals: MS subject, M0/M6

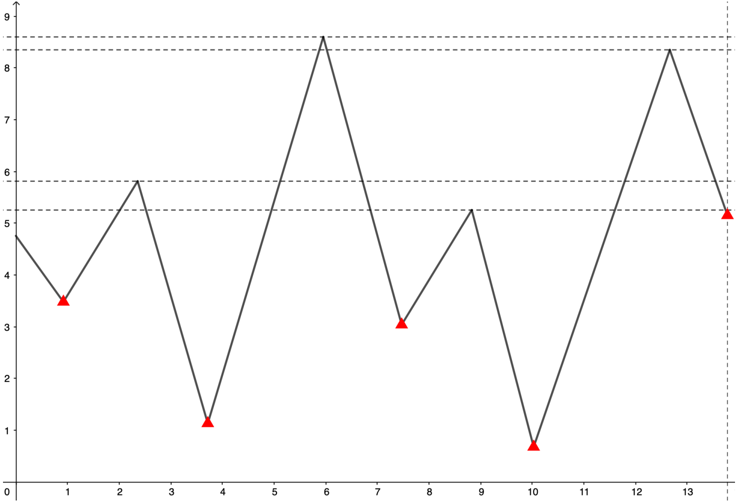


Top: M0. Bottom: M6.

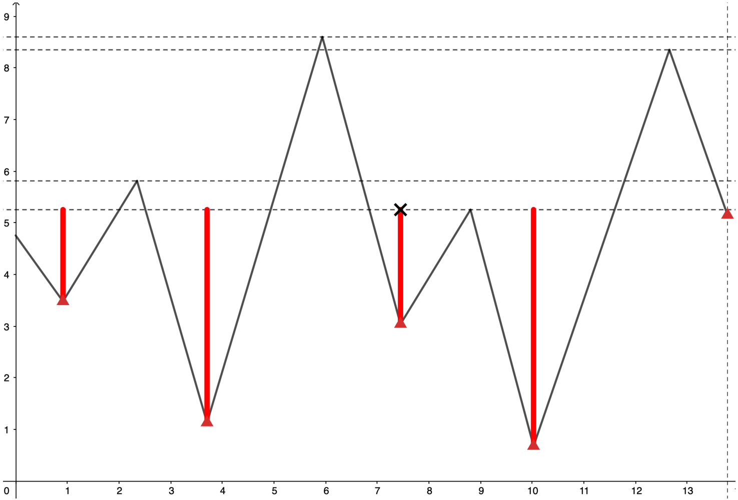
Topological data analysis (TDA)

- General idea: *persistence*: look at the data at different scales.
- Here: for $f : \mathbb{R} \to \mathbb{R}$ and threshold $\alpha \in \mathbb{R}$, study the evolution of sublevel sets $\{t,\ f(t) \leq \alpha\}$.
- Look at the evolution of connected components when α increases (*persistence*). The persistence barcode is the set of pairs (date of birth, date of death) of connected components.
- Concretely: construct a barcode by pairing local minima and maxima.

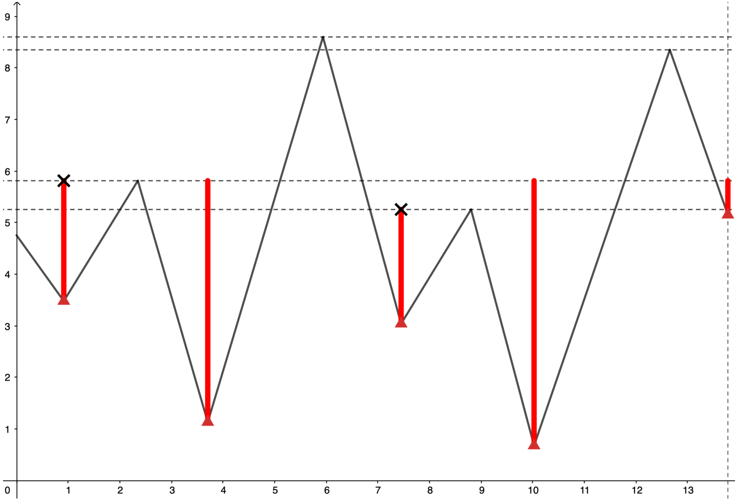
Construction of the persistence barcode from sublevel sets



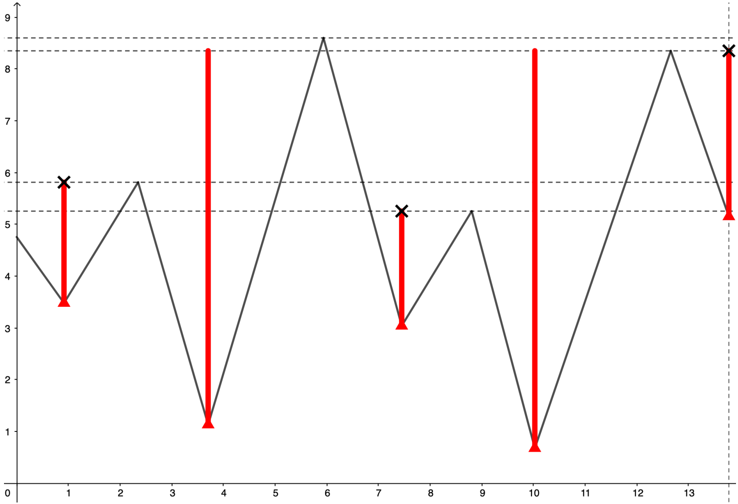
Construction of the persistence barcode from sublevel sets



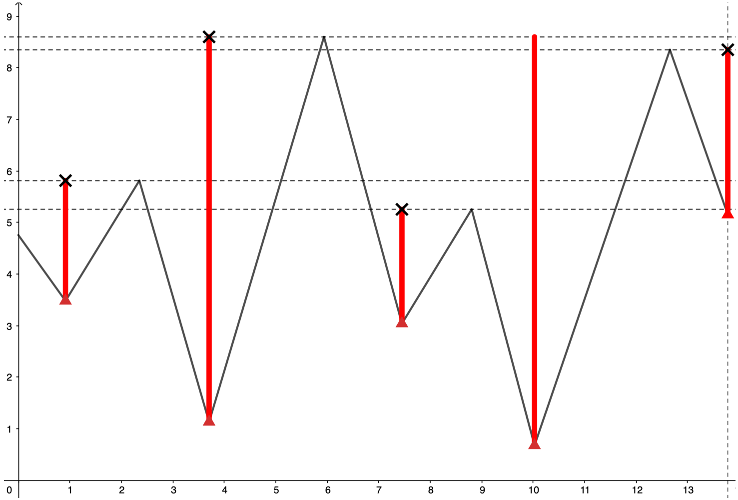
Construction of the persistence barcode from sublevel sets



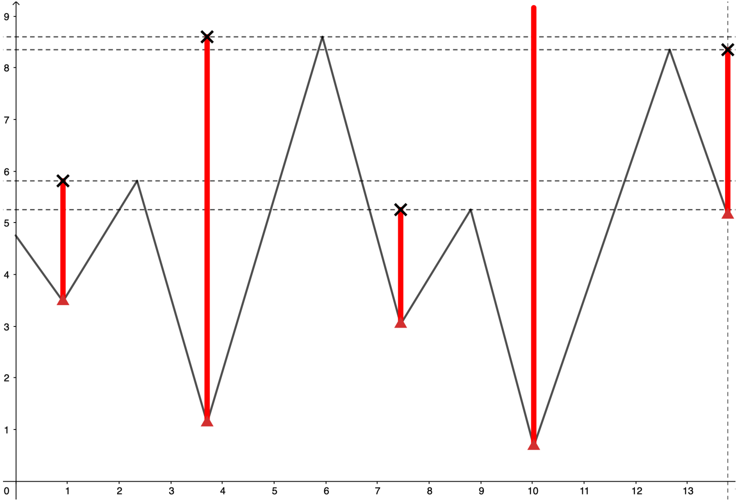
Construction of the persistence barcode from sublevel sets



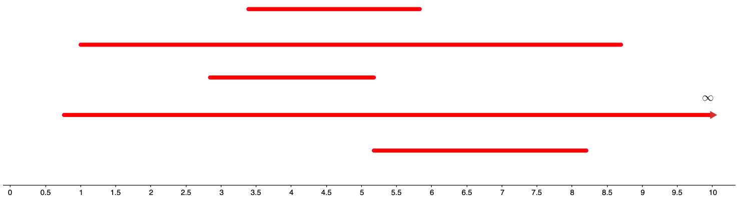
Construction of the persistence barcode from sublevel sets



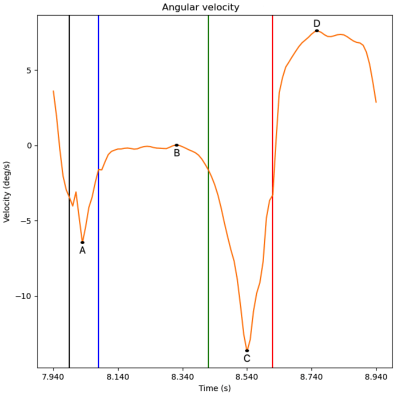
Construction of the persistence barcode from sublevel sets



Construction of the persistence barcode from sublevel sets

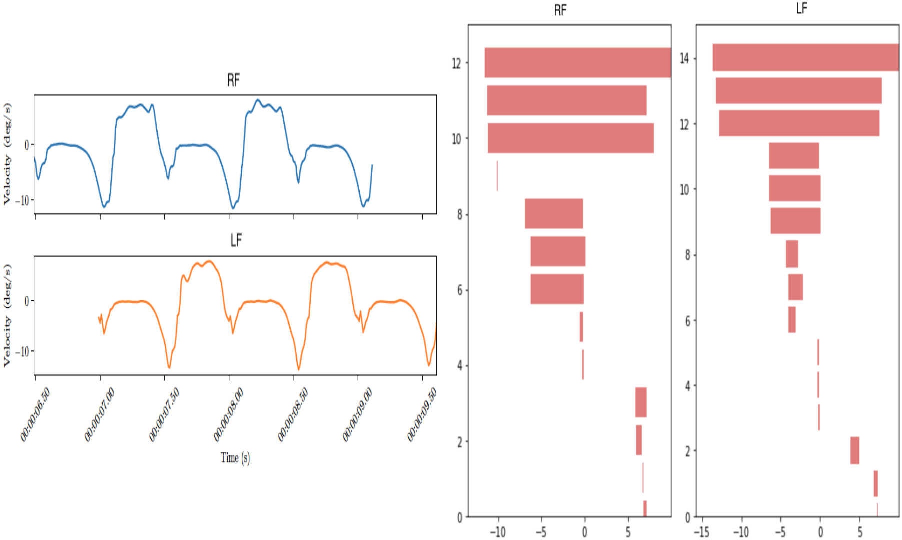


Interpretation (1/2): main bars from each step

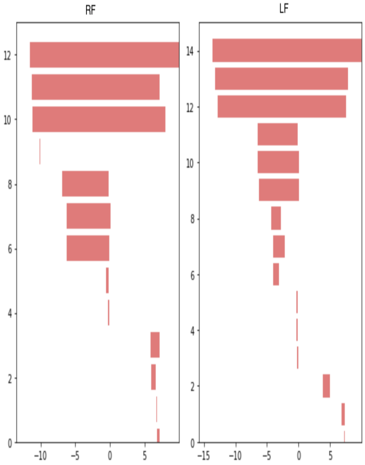


From left to right: Heel Strike (black), Foot Flat (blue), Heel Off (green) and Toe Off (red).

Interpretation (2/2): long bars count the steps



The bottleneck distance

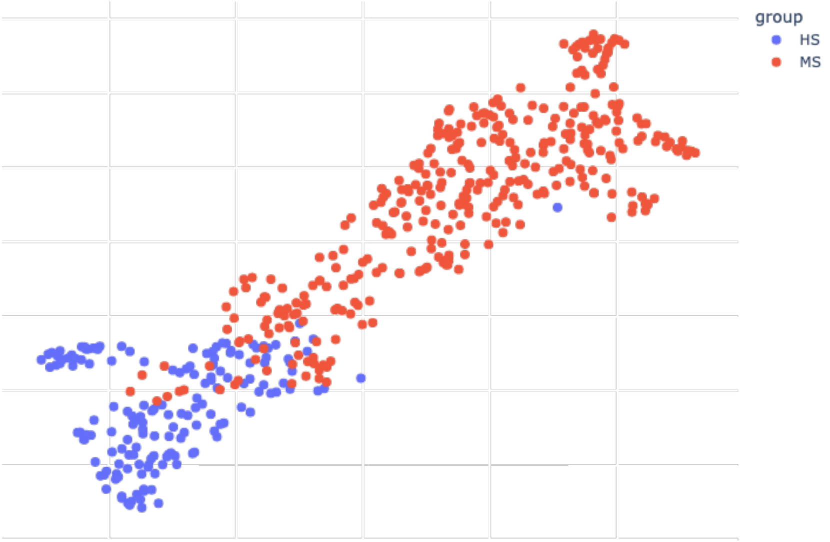


- A distance between barcodes can be defined: make pairs of bars and take the highest start/end difference. Keep the optimal pairing.
- Different number of steps $\Rightarrow$ high bottleneck distance.
- For a signal with k steps, we remove the k most persistent bars to focus on oscillations.

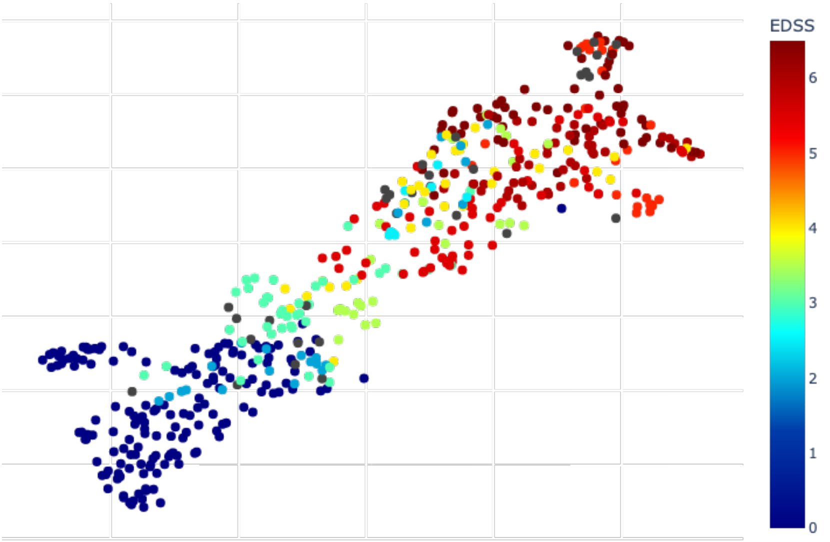
From signals to 2D points

- **UMAP** algorithm. Input: distance matrix. Output: 2D point cloud (1 signal $\longleftrightarrow$ 1 barcode $\longleftrightarrow$ 1 point).
- The Euclidean distance on the point cloud respects the structure induced by the bottleneck distance.
- Partition of the point cloud into groups (healthy/MS, M0/M6, etc.)
- Features to study the separability and density of the groups (silhouette scores, mean squared distance, diameter).

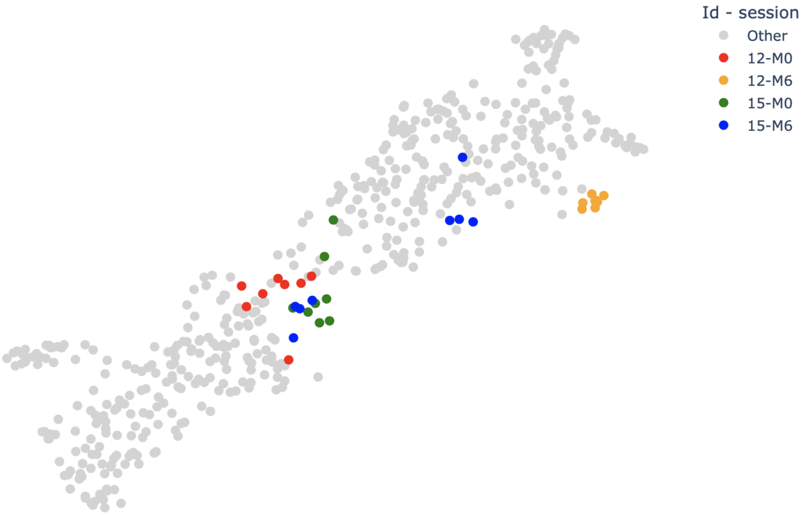
HS VS MS



EDSS groups



Focus on two patients



Contents

Perspectives

- More data : clinical trial with several neurological pathologies + free-living environment on longer periods
- Improvements of step detection techniques using elastic distances instead of Pearson coefficients
- Elaborate on the locogram to create distances between symbolic time series
- Use TDA to automatically extract patterns and anomalies from gait data

References

- Step detection

- L. Oudre, R. Barrois-Müller, T. Moreau, C. Truong, A. Vienne-Jumeau, D. Ricard, N. Vayatis, and P.-P. Vidal. Template-Based Step Detection with Inertial Measurement Units. Sensors, 18(11):4033, 2018
- T. Dot, F. Quijoux, L. Oudre, A. Vienne-Jumeau, A. Moreau, P.-P. Vidal, and D.Ricard. Non-Linear Template-Based Approach for the Study of Locomotion. Sensors, 20(7):1939 , 2020
- R. Barrois-Müller, R. Dadashi, C. de Waele, T. Moreau, L. Oudre, D. Ricard, C. Truong, N. Vayatis, P. P. Vidal, and A. Yelnik. Method for characterising a gait. U.S. Patent No. 11,229,380. Washington, DC: U.S. Patent and Trademark Office, 2022

- Locogram

- R. Barrois-Müller, L. Oudre, D. Ricard, and P.-P. Vidal. Locogram software: tool for analysing gait exercises. U.S. Patent No. 11,253,172. Washington, DC: U.S. Patent and Trademark Office, 2022

- TDA

- A. Bois, B. Tervil, A. Moreau, A. Vienne-Jumeau, D. Ricard, and L. Oudre. A topological data analysis-based method for gait signals with an application to the study of multiple sclerosis. PloS one, 17(5):e0268475, 2022

Thank you for your attention