Compromise clustering for quaternion time-series and its application for gait analysis in Multiple Sclerosis

Pierre Drouin[1,2], Aymeric Stamm[1], Laurent Chevreuil[2] Vincent Graillot[2] David Laplaud[3] Lise Bellanger[1]

[1]Laboratoire de Mathématiques Jean Leray, Faculté des Sciences et Techniques, Nantes, France
[2]Uman E-Health Solution, UmanIT, Nantes, France
[3]Centre d'investigation Clinique, équipe Neurologie, Centre Hospitalier Universitaire, Nantes, France
E-mail for correspondence: `pdrouin@umanit.fr`

Abstract: Recent approaches in gait analysis involve the use of quaternion time-series representing the body segments rotation and/or orientation in 3D space during the walk. Following the proposition of Motl(2017) in, we hereby propose a method to analyze walking data measured on patients with Multiple Sclerosis (MS). The subject matter is to find groups of multiple sclerosis patients with similar walking deficiencies. Expanded Disability Status Scale is currently used to assess overall disability in MS. This information may be taken into account when searching for the grouping structure to provide clinically relevant partition. There are two ways of incorporating external information: either through constraints or through a compromise. The former usually forces some observations to belong to the same cluster or some cluster to be structured in a particular fashion (Dinler(2016)) while the latter only uses the external information to guide the search of the grouping structure. A recent clustering approach known as `periodclust` (Bellanger(2020)) builds on the principal of hierarchical agglomerative clustering and accounts for external information in the way similarity between observations is measured. We propose its generalisation for quaternion time-series and present results of its application on the walking data of patients with MS.

Key words: Time series; Compromise Clustering; Quaternion; Human Gait

Motl RW, Cohen JA, Benedict R, Phillips G, LaRocca N, Hudson LD, Rudick R (2017). Validity of the timed 25-foot walk as an ambulatory performance outcome measure for multiple sclerosis. Multiple Sclerosis Journal, Vol. 23(5) 704–710.

Dinler D and Tural MK (2016). Analysis of Longitudinal Data. Unsupervised learning algorithms.

Bellanger L, Coulomb A, Husi P (2020). Models for Discrete Longitudinal Data. Data Analysis, and Rationality in a Complex World.