

Apprentissage de Dictionnaires Convolutifs Personnalisés

Axel ROQUES^{1, 2, 3} Samuel GRUFFAZ^{1, 4} Kyurae KIM⁵ Alain O. DURMUS⁶ Laurent OUDRE¹

¹Centre Borelli, 4 avenue des Sciences, 91190 Gif-sur-Yvette, France

²Laboratoire GBCM, 292 rue Saint Martin, 75003 Paris, France

³Thales AVS, 1 rue du Général de Gaulle, 95520 Osny, France

⁴ENS Paris-Saclay, 4 avenue des Sciences, 91190 Gif-sur-Yvette, France

⁵University of Pennsylvania, Philadelphie, Pennsylvanie 19104, États-Unis

⁶Laboratoire CMAP, 91120 Palaiseau, France

Résumé – Les séries temporelles issues de la physiologie humaine se composent généralement de structures à la fois globales et locales : les premières sont partagées par l’ensemble d’une population, tandis que les secondes sont attribuables à la variabilité inter-individuelle. Par exemple, la mesure du cycle de marche montre à la fois des caractéristiques communes et des idiosyncrasies en raison des dispositions biomécaniques de chacun, ou bien de pathologies. Pour mieux représenter de telles données, nous proposons l’apprentissage de dictionnaires convolutifs personnalisés (PerCDL). PerCDL apprend conjointement un dictionnaire partagé par tous les signaux et un dictionnaire propre à chaque série temporelle, via la transformation spatio-temporelle personnalisée des structures communes. L’intérêt et l’efficacité de PerCDL sont démontrés sur des données synthétiques et des données de marche.

Abstract – Time series of human physiology tend to exhibit both global and local structures: the former are shared across a population, while the latter reflect inter-individual variability. For instance, measurement of the gait cycle present both common characteristics and idiosyncrasies due to biomechanical dispositions or pathologies. To better represent such data, this work introduces Personalized Convolutional Dictionary Learning (PerCDL). PerCDL jointly learns a dictionary shared by all signals and a signal-specific dictionary, through the spatio-temporal transformation of the common structures. The value and effectiveness of PerCDL are demonstrated on synthetic data and a human gait dataset.

1 Introduction

L’apprentissage de dictionnaire convolutifs (ou CDL pour *convolutional dictionary learning*; [2]) est une méthode d’apprentissage de représentation non supervisée qui décompose un ensemble de signaux en un dictionnaire, une collection de signaux de référence ou atomes, et un ensemble de coefficients qui indiquent la position et l’amplitude des atomes dans les données. Le CDL offre des capacités de représentation particulièrement intéressantes dans le cadre de séries temporelles issues de la physiologie humaine. Pour la marche par exemple, des cycles de marche réguliers et structurés peuvent être identifiés [6]. Toutefois, le CDL est incapable de traiter leur structure *locale-globale*. En effet, des structures communes (*globales*) peuvent exister au niveau de la population en raison de traits anatomiques similaires ou de circuits neuronaux communs. À l’échelle de l’individu, des variations personnelles (*locales*) autour de ces structures communes peuvent exister en raison d’idiosyncrasies, de différences de style perceptivo-moteur ou d’une pathologie [4]. Or, l’application naïve du CDL à l’ensemble d’une population (PopCDL) produira un dictionnaire qui ne tient pas compte des structures fines au niveau individuel, traitées comme du bruit. À l’inverse, appliquer le CDL à chaque individu indépendamment (IndCDL) exclut l’apprentissage des structures partagées et peut se révéler peu fiable en présence de bruit.

Pour répondre à cette problématique, nous proposons

l’apprentissage de dictionnaires convolutifs personnalisés (PerCDL), qui vise à identifier à la fois les structures globales et locales à partir d’une collection de séries temporelles. Son principe de fonctionnement est illustré en Figure 1. Brièvement, PerCDL apprend des atomes personnalisés à partir d’atomes communs partagés par la population. La personnalisation est assurée par la transformation spatio-temporelle des atomes communs à l’aide d’une déformation temporelle (DT).

Nos contributions sont les suivantes. 1) Nous proposons PerCDL, une nouvelle méthode pour l’apprentissage de représentations interprétables de séries temporelles qui met en évidence les écarts personnels par rapport aux structures communes (Section 2). 2) Nous présentons un méta-algorithme modulaire pour résoudre le problème d’optimisation associé (Section 3). 3) Nous évaluons PerCDL sur des données synthétiques et réelles (Section 4).

2 PerCDL

Notations. Notons $[k : l]$ l’intervalle d’entiers $\{k, \dots, l\} \subset \mathcal{P}(\mathbb{Z})$. Par extension $[l] = [1 : l]$. Pour tout vecteur $\mathbf{u} = (u_t)_{t \in [T]}$, sa norme l_2 est $\|\mathbf{u}\|_2 = \sqrt{\sum_{t=1}^T \|u_t\|^2}$ et sa norme l_0 $\|\mathbf{u}\|_0 = \sum_{t=1}^T 1_{\neq 0}(u_t)$. Nous écrirons $\mathcal{O}_P = \{O \in \mathbb{R}^{P \times P} : O^\top O = I_P\}$ et $C^1([0, 1])$ l’espace des fonctions continûment différentiables $g : [0, 1] \rightarrow [0, 1]$.

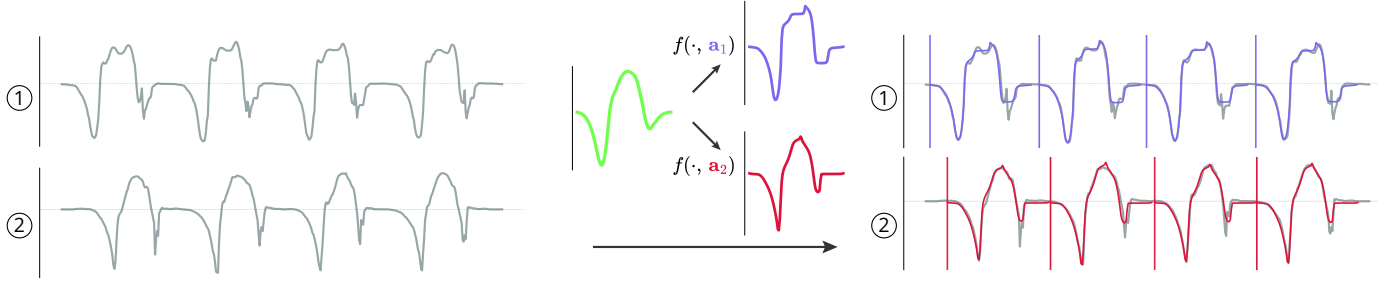


FIGURE 1 : Illustration de PerCDL. (Gauche) La vitesse angulaire du pied dans le plan sagittal de deux individus présentent des motifs structurés et répétés lors du cycle de marche (gris). (Milieu) PerCDL apprend un dictionnaire partagé (vert) et la variabilité spécifique aux participants en *personnalisant* l'atome commun par la fonction de transformation paramétrique f , selon des paramètres de personnalisation \mathbf{a}_1 et \mathbf{a}_2 . (Droite) PerCDL s'appuie sur la robustesse des structures communes identifiées et ses capacités de personnalisation pour obtenir une représentation fidèle des données.

2.1 Formulation du problème

Considérons un jeu de données $\mathbf{X} \in \mathbb{R}^{S \times N \times P}$ de S séries temporelles P -dimensionnelles $(\mathbf{x}^s)_{s \in [S]}$ de même taille N . Afin de tenir compte de la variabilité inter-individuelle, nous cherchons à apprendre des structures à l'échelle de la population et à l'échelle du signal. Soient un dictionnaire $\Phi = (\phi_k)_{k \in [K]} \in \mathbb{R}^{K \times L \times P}$ composé de K atomes communs de longueur L ($L < N$) et leurs poids et activations $\mathbf{Z} = (\mathbf{z}_k^s)_{s,k \in [S] \times [K]} \in \mathbb{R}^{S \times K \times (N-L+1)}$ associées. Chaque atome personnalisé $\hat{\phi}_k^s$ est dérivé d'un atome ϕ_k du dictionnaire commun Φ selon une fonction de transformation paramétrisée $f : (\mathbb{R}^P)^L \times \Theta \rightarrow (\mathbb{R}^P)^L$, où $\Theta \subset \mathbb{R}^M$ est l'ensemble des paramètres de personnalisation. La version personnalisée du k^{ieme} atome commun ϕ_k de la s^{ieme} série temporelle s'écrit, pour tout $(s, k) \in [S] \times [K]$, $\hat{\phi}_k^s = f(\phi_k, \mathbf{a}_k^s)$, avec $\mathbf{a}_k^s \in \Theta$ les paramètres de personnalisation.

Soit $\mathbf{A} = (\mathbf{a}_k^s)_{s,k \in [S] \times [K]} \in \mathbb{R}^{S \times K \times M}$ la matrice des paramètres de personnalisation. Alors, le problème PerCDL est

$$\underset{\mathbf{Z} \geq 0, \Phi: \|\phi_k\|_2=1, \mathbf{A}}{\operatorname{argmin}} \sum_{s=1}^S \left\| \mathbf{x}^s - \sum_{k=1}^K \mathbf{z}_k^s * \hat{\phi}_k^s \right\|_2^2 + \lambda \|\mathbf{Z}\|_0, \quad (1)$$

où $*$ dénote l'opérateur de convolution multivariée tel que, pour deux séries temporelles, $\mathbf{u} = (u_i)_{i \in [N-L+1]} \in \mathbb{R}^{N-L+1}$ et $\mathbf{v} = (v_j)_{j \in [L]} \in (\mathbb{R}^P)^L$,

$$\mathbf{u} * \mathbf{v} = \left(\sum_{l=1}^L v_l u_{m-l+1} \right)_{m \in [N]} \in (\mathbb{R}^P)^N.$$

λ est un terme de pénalité, qui équilibre la parcimonie de la matrice d'activation et l'erreur de reconstruction.

2.2 Fonction de transformation

Le choix d'une fonction de transformation f appropriée est essentiel. Nous proposons ci-dessous une famille de fonctions de transformation bien adaptée (mais pas limitée) aux données physiologiques. Plus généralement, toute fonction de transformation *ad hoc* peut être utilisée.

Considérons un signal continu $s : [0, 1] \rightarrow \mathbb{R}^P$. Pour $t \in [0, 1]$, la version "déformée temporellement" de s est $s_{\text{warp}}(t) = s(\psi_{\mathbf{a}}(t))$, où $\psi_{\mathbf{a}} : [0, 1] \rightarrow [0, 1]$ est une fonction de DT paramétrisée par $\mathbf{a} \in \mathbb{R}^M$. Pour cette dernière, nous utilisons la méthode récemment introduite par CELLEDONI

et al. [1] que nous adaptons au cas d'atomes discrets. Soient deux hyperparamètres $D, W \in \mathbb{N}$ tels que $M = DW$. Alors, pour tout $\mathbf{a} = (a_w^d)_{d,w \in [D] \times [W]} \in \Theta$, la fonction de DT $\psi_{\mathbf{a}}$ utilisée s'écrit comme la composition de D déplacements :

$$\psi_{\mathbf{a}} = \psi_{D,\mathbf{a}} \circ \dots \circ \psi_{1,\mathbf{a}}, \quad (2)$$

avec, pour $d \in [D]$ et $t \in [0, 1]$,

$$\psi_{d,\mathbf{a}} = \operatorname{Id} + \sum_{w=1}^W a_w^d b_w, \quad b_w(t) = \frac{\sin(w\pi t)}{w\pi}. \quad (3)$$

La profondeur D et la largeur W contrôlent respectivement la *dilatation globale* et la *dilatation locale* de la transformation et déterminent donc toutes deux l'étendue de la variabilité autour des atomes communs (Figure 1).

3 Méta-algorithme

Un méta-algorithme (Algorithme 1) est proposé pour résoudre le problème d'optimisation en Équation (1).

Algorithme 1 : Méta-algorithme pour PerCDL.

Entrées \mathbf{X}, f

Sorties $\Phi, \mathbf{A}, \mathbf{Z}$

1 $\Phi, \mathbf{A}, \mathbf{Z} \leftarrow \text{initialisation}()$

2 **while** $i \leq n_{\text{init}}$ **do**

3 $\mathbf{Z} \leftarrow \forall s, \underset{\mathbf{Z}_s \geq 0}{\operatorname{argmin}} \left\| \mathbf{x}^s - \sum_{k=1}^K \mathbf{z}_k^s * \phi_k \right\|_2^2 + \lambda \|\mathbf{Z}\|_0$

4 $\Phi \leftarrow \underset{\Phi: \|\phi_k\|_2=1}{\operatorname{argmin}} \sum_{s=1}^S \left\| \mathbf{x}^s - \sum_{k=1}^K \mathbf{z}_k^s * \phi_k \right\|_2^2$

5 **end**

6 **while** $i \leq n_{\text{perso}}$ **do**

7 $\mathbf{A} \leftarrow \forall s, \underset{\mathbf{A}}{\operatorname{argmin}} \left\| \mathbf{x}^s - \sum_{k=1}^K \mathbf{z}_k^s * \hat{\phi}_k^s \right\|_2^2$

8 $\mathbf{Z} \leftarrow \forall s, \underset{\mathbf{Z}_s \geq 0}{\operatorname{argmin}} \left\| \mathbf{x}^s - \sum_{k=1}^K \mathbf{z}_k^s * \hat{\phi}_k^s \right\|_2^2 + \lambda \|\mathbf{Z}\|_0$

9 $\Phi \leftarrow \underset{\Phi: \|\phi_k\|_2=1}{\operatorname{argmin}} \sum_{s=1}^S \left\| \mathbf{x}^s - \sum_{k=1}^K \mathbf{z}_k^s * \hat{\phi}_k^s \right\|_2^2$

10 **end**

Initialisation. Les activations \mathbf{Z} et les paramètres de personnalisation \mathbf{A} sont initialisés à zéro. Le CDL est sensible à

l'initialisation du dictionnaire [2], il est recommandé d'utiliser une estimation raisonnable des motifs attendus lorsque cela est possible.

Estimations initiales. Des premières estimations du dictionnaire commun et des activations associées sont obtenues (lignes 2-5). Nous employons une minimisation alternée entre une étape de codage parcimonieux (ligne 3) et une étape de mise à jour du dictionnaire (ligne 4). Tous les solveurs usuels peuvent être utilisés.

Personnalisation. Les paramètres de personnalisation \mathbf{A} sont appris (ligne 6-7), tout en affinant les activations \mathbf{Z} (ligne 8) et le dictionnaire commun Φ (ligne 9). Pour ce dernier, la transformation f peut introduire des non-linéarités qui rendent l'utilisation des techniques classiques réductrices. Dans le cas présent où f est linéaire, un solveur conventionnel peut être utilisé.

Complexité temporelle. Le coût supplémentaire de PerCDL par rapport au CDL classique est l'étape de mise à jour des paramètres, de complexité $\mathcal{O}(SKMNP T_{\text{grad}})$ avec T_{grad} le nombre de pas de gradient.

4 Applications

4.1 Données synthétiques

Les données synthétiques sont constituées de S séries temporelles de longueurs égales N , contenant chacune $r = 3$ répétitions de $K = 2$ motifs communs de longueur L , transformés à l'aide de paramètres de personnalisation aléatoires et distribués aléatoirement dans chaque série sans recouvrement. Le dictionnaire est initialisé avec l'atome personnalisé du premier signal.

Structures communes. La capacité de IndCDL, PopCDL et PerCDL à extraire les structures communes en fonction du nombre de signaux a été étudiée. Les données contenaient entre $S = 1$ et $S = 1024$ séries temporelles de longueur $N = 500$. Pour IndCDL, le barycentre euclidien des atomes de chaque signal a été calculé pour modéliser les structures partagées. La distance entre les vraies structures communes et les atomes obtenus est représentée en Figure 2.

La qualité des atomes communs identifiés par IndCDL et PopCDL se dégrade avec une augmentation du nombre de signaux. Cela suggère que les idiosyncrasies de chaque signal affectent négativement l'estimation des structures communes. Au contraire, PerCDL garantit une bonne identifiabilité des formes communes dans les données, qui s'améliore à mesure que la taille de l'ensemble de données augmente.

Robustesse au bruit. La robustesse des différentes méthodes a été évaluée sur des données synthétiques contaminées par du bruit impulsif. Tous les ensembles de données synthétiques testés contenaient $S = 32$ séries temporelles de longueur égale $N = 1000$, avec des atomes communs de longueur $L = 50$. Un quart des signaux, choisis aléatoirement, ont été corrompus : p pour cent de tous les échantillons ont été pollués par un pic positif ou négatif d'amplitude aléatoirement uniforme dans l'intervalle $[2, 2.5]$. Figure 3 présente les erreurs de reconstruction observées en fonction de la quantité de bruit.

Alors que les performances de toutes les méthodes se dégradent avec l'augmentation du bruit, PerCDL s'avère plus robuste que IndCDL et PopCDL. Ce contraste entre IndCDL et PerCDL ne peut pas s'expliquer par une différence dans le

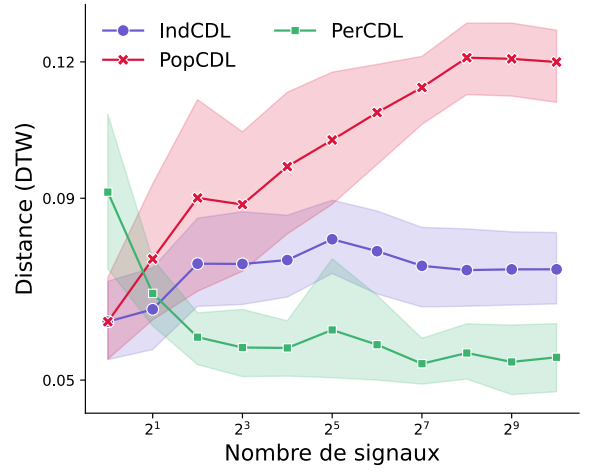


FIGURE 2 : Distance moyenne entre les atomes communs identifiés et les structures réelles en fonction du nombre de signaux. Les régions en transparence représentent l'intervalle de confiance à 95%.

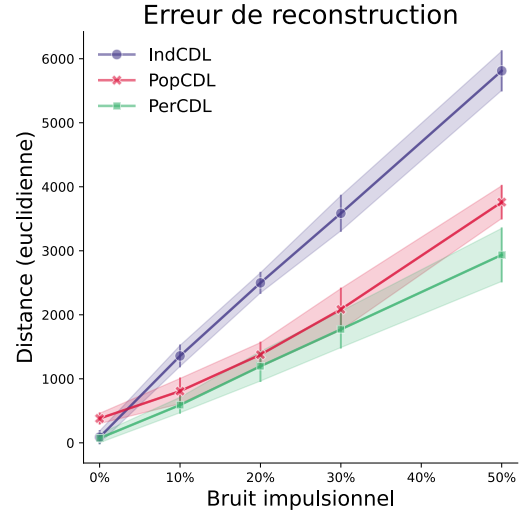


FIGURE 3 : Erreur de reconstruction en présence de bruit impulsif. Les barres d'erreur représentent deux déviations standard.

nombre de paramètres, car les deux méthodes en avaient un nombre similaire ($W * D \approx L$ dans cette expérience). Dès lors, nos résultats indiquent que les informations sur les structures partagées par la population, combinées aux possibilités de représentation supérieures offertes par la personnalisation, offrent à PerCDL une meilleure résistance face au bruit et aux artefacts dans les données.

4.2 Données de marche

PerCDL a été appliqué à un jeu de données sur la marche [3]. Trois groupes de participants (Sain, Orthopédique ou Neurologique selon l'état de santé), de taille égale $n = 50$, ont porté une centrale inertielle sur la face dorsale de leur pied pour enregistrer sa vitesse angulaire dans le plan sagittal lors d'un rapide aller-retour. Un seul atome commun ($K = 1$) a été recherché, de taille $L = 100$, afin de correspondre à la longueur du cycle de marche. Le dictionnaire a été initialisé

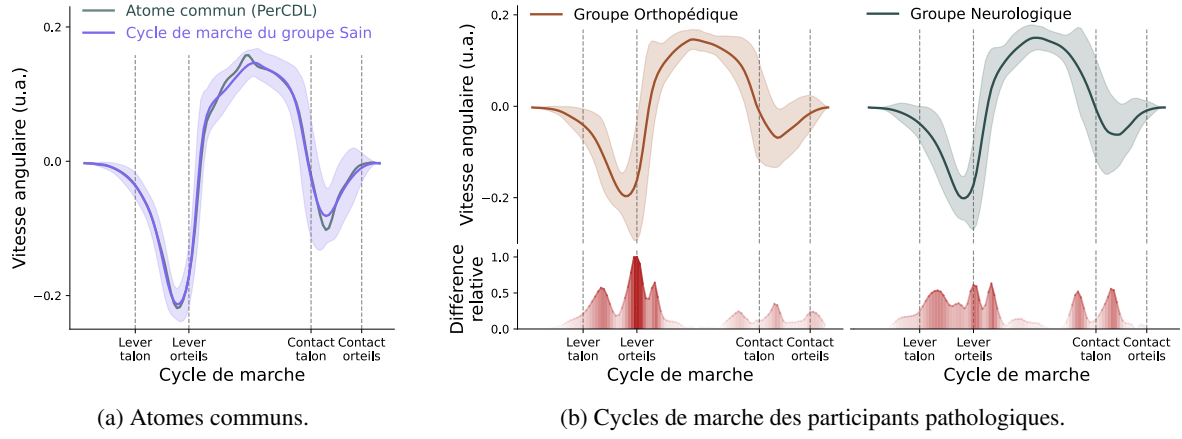


FIGURE 4 : Application de PerCDL aux données de marche. (a) Atomes communs identifiés par PerCDL (gris) et cycle de marche moyen du groupe Sain (bleu). (b) Moyenne des atomes personnalisés appris par PerCDL dans les groupes Orthopédique (marron) et Neurologique (vert). Sur le graphe du dessous, les courbes rouges représentent la variabilité normalisée relative au cours du cycle de marche du groupe avec le groupe Sain. Les régions en transparence représentent deux déviations standard.

avec le cycle de marche extrait d'un participant sain choisi au hasard. Les hyperparamètres de personnalisation ont été fixés à $D = 4$, $W = 15$ après un bref processus d'ajustement.

L'une des premières tâches en analyse de la marche consiste à identifier les cycles de marche. Puisque le début précis d'un cycle est une notion floue, un seuil de tolérance fixé à $\tau_{\text{tol}} = 100$ échantillons a été fixé pour permettre la comparaison avec d'autres méthodes. Le Tableau 1 montre que PerCDL obtient des scores compétitifs avec l'état de l'art [5].

TABLE 1 : Segmentation des cycles de marche.

	Voisard et al.	PerCDL
Sensitivity	0.959	0.988
FPR	0.004	0.033

PerCDL permet avant-tout la représentation, l'interprétation et l'analyse des données. Figure 4(a) révèle que l'atome commun identifié (gris) est similaire à la moyenne des atomes personnalisés de la population saine. Ainsi, la structure partagée identifiée par PerCDL a une signification intrinsèque et sert de base cohérente à la personnalisation. Cette observation est prometteuse car elle pourrait faciliter la surveillance d'écarts fins par rapport à une "ligne de base saine", qui peuvent indiquer le début d'un processus pathologique [4].

La Figure 4(b) illustre les atomes personnalisés obtenus chez les différents groupes. Bien que tous présentent des cycles de marche similaires, des différences inter-groupes perceptibles peuvent être identifiées à différentes phases du cycle de marche. La représentation de la différence relative normalisée de la variabilité locale entre chaque groupe pathologique et la population saine met ce point en évidence (courbes rouges). Nos résultats suggèrent que le groupe Orthopédique se distingue d'un cycle de marche sain principalement au début de la phase oscillante (indiqué par "Lever des orteils"), tandis que le groupe Neurologique diffère principalement à la fin de la phase d'appui (entre "Lever du talon" et "Lever des orteils") et juste avant la fin de la phase oscillante ("Contact du talon").

5 Conclusion

Nous avons introduit l'apprentissage par dictionnaire personnalisé convolutif (PerCDL), un cadre mathématique pour l'apprentissage de représentations interprétables de séries temporelles, qui capture à la fois les structures communes et leurs variations locales. Un méta-algorithme a été proposé et ses performances ont été démontrées empiriquement sur des données synthétiques et réelles. Le cadre proposé est par nature général et pourrait être étendu à un large panel d'applications, e.g., en modifiant la fonction de transformation f . La principale limite de la méthode proposée est celle du coût de calcul associé à l'étape IPU. Pour résoudre ce problème, une variante de PerCDL en apprentissage fédéré est en cours de développement.

Références

- [1] Elena CELLEDONI et al. "Deep neural networks on diffeomorphism groups for optimal shape reparametrization". In : *BIT Numerical Mathematics* 63.4 (2023). Publisher : Springer, p. 50.
- [2] Cristina GARCIA-CARDONA et Brendt WOHLBERG. "Convolutional dictionary learning : A comparative review and new algorithms". In : *IEEE Transactions on Computational Imaging* 4.3 (2018). Publisher : IEEE, p. 366-381.
- [3] Charles TRUONG et al. "A Data Set for the Study of Human Locomotion with Inertial Measurements Units". In : *Image Processing On Line* 9 (2019), p. 381-390.
- [4] Pierre-Paul VIDAL et Francesco LACQUANITI. "Perceptual-motor styles". In : *Experimental Brain Research* 239.5 (2021). Publisher : Springer, p. 1359-1380.
- [5] Cyril VOISARD et al. "Automatic gait events detection with inertial measurement units : healthy subjects and moderate to severe impaired patients". In : *Journal of NeuroEngineering and Rehabilitation* 21.1 (2024), p. 104.
- [6] Michael W WHITTLE. *Gait analysis : an introduction*. Butterworth-Heinemann, 2014.