Sujet de thèse : MALMOCAP-LS : MACHine Learning et MOtion CAPture pour l'analyse de la Langue des Signes

Modélisation et analyse statistique de données du mouvement pour la détection et la caractérisation de primitives, et la classification en langue des signes

Mots-clés : Machine Learning, Deep Learning, Capture de Mouvement, Analyse de trajectoires cinématiques, Analyse de Données Fonctionnelles, Groupes de Lie, Classification, Langue des Signes.

Contact : nicolas.brunel@ensiie.fr

Encadrants :

Académique : Nicolas BRUNEL, Laboratoire de Mathématiques et Modélisation d'Evry (UMR CNRS 8071), ENSIIE, Evry.

Entreprise : Rémi BRUN, MOCAPLAB.

Laboratoire : LaMME, Ecole Doctorale de Mathématiques Hadamard, Paris Saclay.

La capture du mouvement au service de la langue des signes

La thèse est une collaboration avec l'entreprise Mocaplab, une des rares sociétés Européennes spécialisée en Motion Capture. Parmi ces nombreux champs d'exploration de la motion capture, se trouve le domaine très spécifique de la langue des signes. Les sourds communiquent entre eux en utilisant les mouvements du corps, des mains, des doigts, du visage et des yeux. Ils arrivent à une capacité d'expression aussi riche et structurée que celle offerte par la parole.

L'objectif de cette thèse est d'identifier des primitives du mouvement des mains dans le contexte de la langue des signes, et de fournir des modèles et des outils de segmentation non-supervisée de séquence de langues de signes. Enfin, sur cette base, nous envisageons le développement de méthodes de classification de signaux de langue des signes.

Analyse de données fonctionnelles

Dans le contexte de la langue des signes, Il s'agit donc d'identifier le bon cadre mathématique et les propriétés indispensables dans le "signal mouvement". Il s'agira d'identifier des primitives, descripteurs et autres "traits cinétiques" impliqués dans la langue des signes, permettant d'imaginer en prime des solutions de segmentation et de classification et de reconnaissance. Pour cela, nous proposons d'utiliser des techniques issues de l'analyse des données fonctionnelles [Wang et al 2016]. Plus généralement, ces méthodes ont été étendues pour la modélisation de la variabilité et la prédiction de trajectoires multivariées, de formes, ou de trajectoires à valeurs dans des espaces éventuellement complexes. Dans le cas de courbes 1D, les problématiques classiques d'alignement, de variations en amplitude,...peuvent être définies de manière relativement univoques, mais lorsque l'on considère des courbes 3D, ces problèmes peuvent être appréhendés de différentes manières [Marron 2014].

Un point de départ de cette méthodologie consiste en l'estimation des trajectoires issues du motion capture et de la caractérisation de la géométrie et de la cinématique par des représentations fonctionnelles appropriées. Cette analyse simultanée est notamment possible dans la trajectoire d'une particule ponctuelle, en utilisant le cadre de Frénet-Serret et l'estimation du repère de Frénet [Brunel2014, Brunel2019, Park2019]. En premier point, nous proposons de développer des méthodes inférentielles permettant d'estimer le lien entre géométrie et vitesse curvilinéaire, à partir de données trajectorielles mesurées par Motion Capture. Puis dans un deuxième temps d'étendre ces techniques inférentielles à l'analyse des trajectoires d'un solide.

Recherche de primitives

A partir de ces estimateurs, nous souhaitons modéliser et décomposer la variabilité de ces trajectoires qui peuvent être à valeurs dans des groupes de Lie, tels que le Groupe Spécial Orthogonal ou le Groupe Spécial Euclidien. Ainsi l'exploitation de séquences annotées ou non de langues des signes et l'analyse de la diversité de ces mesures pourra permettre de traiter les différents points :

- explorer et estimer des modèles statistiques décrivant le lien entre vitesse curvilinéaire et trajectoire
- identifier les sources de variations les plus importantes dans la forme des trajectoires
- Identifier des principes d'invariance dans les mouvements de chaque main
- Modéliser les positions et vitesse relatives des deux mains dans le cas de la langue des signes.

Machine learning et Deep Learning

Enfin, dans un objectif de classification supervisée, nous envisageons d'utiliser des réseaux de neurones et le Deep Learning pour modéliser les données temporelles de mouvement (ces modèles ont montré un grand succès dans la modélisation de données temporelles et complexes, e.g. réseaux récurrents, convolutifs, orientés graphes...), ou encore de données à valeurs dans des groupes de Lie, [Li, et al 2018]. Nous bénéficions notamment de la possibilité de générer de nouvelles données adaptées grâce l'accès au plateau de MOCAP (8 jours) pour le test et l'amélioration des modèles de deep learning.

Références

[Stockoe2005] Stockoe Jr, William C. "Sign language structure: An outline of the visual communication systems of the American deaf." *Journal of deaf studies and deaf education* 10.1 (2005): 3-37.

[Wang et al 2016] Wang, Jane-Ling, Jeng-Min Chiou, and Hans-Georg Müller. "Functional data analysis." *Annual Review of Statistics and Its Application* 3 (2016): 257-295.

[Marron 2014] Marron, J. S., et al. "Statistics of time warpings and phase variations." *Electronic Journal of Statistics* 8.2 (2014): 1697-1702.

[Srivastava 2016] Srivastava, Anuj, and Eric P. Klassen. *Functional and shape data analysis*. Vol. 475. New York: Springer, 2016.

[Hicheur2005] Hicheur, Halim, et al. "Velocity and curvature in human locomotion along complex curved paths: a comparison with hand movements." *Experimental brain research* 162.2 (2005): 145-154.

[Viviani1995] Viviani, Paolo, and Tamar Flash. "Minimum-jerk, two-thirds power law, and isochrony: converging approaches to movement planning." *Journal of Experimental Psychology: Human Perception and Performance* 21.1 (1995): 32.

[Flash2019] Flash, Tamar, et al. "Motor compositionality and timing: combined geometrical and optimization approaches." *Biomechanics of Anthropomorphic Systems*. Springer, Cham, 2019. 155-184.

[Brunel2014] Brunel, Nicolas J-B., and Juhyun Park. "Removing phase variability to extract a mean shape for juggling trajectories." *Electronic Journal of Statistics* 8.2 (2014): 1848-1855.

[Park2019] Park, Juhyun, and Nicolas JB Brunel. "Mean curvature and mean shape for multivariate functional data under Frenet-Serret framework." *arXiv preprint arXiv:1910.12049* (2019).

[Brunel2019] Brunel, Nicolas J-B., and Juhyun Park. "The Frenet-Serret framework for aligning geometric curves." *International Conference on Geometric Science of Information*. Springer, Cham, 2019.

[Li, et al 2018] Li, Fanzhang, Li Zhang, and Zhao Zhang. *Lie group machine learning*. Walter de Gruyter GmbH & Co KG, 2018.