# Biomimetic movement-based prosthesis control : dataset of natural movements and reference frame transformation for real-world settings

Bianca Lento

THÈSE PRÉSENTÉE POUR OBTENIR LE GRADE DE

**DOCTEUR**
**DE L'UNIVERSITÉ DE BORDEAUX**

École doctorale Sociétés, politique, santé publique
Sciences cognitives et ergonomie - Option sciences cognitives

---

# Contrôle biomimétique de prothèse à partir de mouvements naturels : base de données et transformation de référentiel pour une situation réelle.

---


**Présentée par :**
Bianca LENTO

**Réalisée à :**
l'**I**nstitut de **N**eurosciences **C**ognitives et **I**ntégratives d'**A**quitaine, Bordeaux

**Sous la direction de :**
Dr Aymar GOULLET DE RUGY


**Soutenue le :**
3 Octobre 2024

**Membres du jury :**

M. VERCHER Jean-Louis, Directeur de recherche émérite, ISM, CNRS, Aix-Marseille Université, Président

M. DOMINEY Peter Ford, Directeur de recherche, INSERM , Université Bourgogne Franche-Comté,Rapporteur

Mme AZEVEDO COSTE Christine, Directrice de recherche, INRIA, Université de Montpellier, Rapportrice

Mme MAURICE Pauline, Chargée de recherche, CNRS (Vandoeuvre-lès-Nancy), Examinatrice

M. DE RUGY Aymar, Directeur de recherche, INCIA, CNRS, Université de Bordeaux, Directeur de thèse

# Résumé

Les contrôles myoélectriques pour les prothèses transhumérales entraînent souvent un taux élevé d'abandon en raison de leurs performances insatisfaisantes. Inspirés des progrès réalisés dans les contrôles exploitant les mouvements résiduels, nous avons affiné une approche alternative utilisant un réseau de neurones artificiels entrainé sur les mouvements naturels de bras pour prédire la configuration des articulations distales en fonction du mouvement des articulations proximales et d'information sur l'objet à saisir. Des études antérieures ont montré que cette stratégie permet aux amputés de contrôler un avatar de prothèse dans un environnement de réalité virtuelle aussi bien qu'avec leur bras valide. Cependant, le déploiement de ce contrôle dans des scénarios réels requiert des développements supplémentaires. Il est nécessaire d'intégrer une caméra montée sur la tête et des algorithmes de vision par ordinateur pour estimer en temps réel la position et l'orientation de l'objet. Dans ce contexte, les informations sur l'objet ne seraient disponibles que dans un référentiel centré sur la tête de l'utilisateur, alors que notre contrôle repose sur l'objet exprimé dans un référentiel centré sur l'épaule.

Inspirés de la façon dont le cerveau exécute des transformations de coordonnées, nous avons développé et testé des solutions pour effectuer la transformation tête-épaule à partir des seules données d'orientation, disponibles en situation réelle. Pour développer ces algorithmes, nous avons constitué une base de données incluant la relation entre ces référentiels en demandant à vingt participants valides de saisir des objets dans diverses positions et orientations dans un environnement virtuel. Cette base de données comprenait les mouvements de la tête et du regard, ainsi que ceux du tronc, des épaules et des bras, capturant l'ensemble de la chaîne cinématique entre le but du mouvement et la main déplacée pour l'atteindre.

Ensuite, nous avons mis en œuvre deux méthodes pour obtenir la position de la tête dans le référentiel de l'épaule. La première consiste en un réseau de neurones artificiels entraîné hors ligne sur la base de données pour prédire cette position en fonction de la taille du participant et de l'orientation de sa tête et de son épaule. La seconde méthode s'inspire des processus d'intégration multisensorielle du cerveau et déduit la position de la tête dans le référentiel de l'épaule en comparant les données relatives à la main prothétique obtenues dans le référentiel de l'épaule par cinématique directe et simultanément dans le référentiel de la tête par la vision par ordinateur. Inspirés par les mécanismes neuronaux du codage de l'espace péripersonnel, nous avons adapté

en ligne les poids d'un réseau de neurones pour coder cette différence dans une carte spatiale.

Les résultats expérimentaux sur douze participants valides en réalité virtuelle ont démontré des erreurs persistantes avec la première méthode, qui n'a pas réussi à prendre en compte avec suffisamment de précision la spécificité de la morphologie de l'utilisateur, entraînant ainsi un contrôle inefficace de la prothèse. En revanche, la seconde méthode a réussi à coder efficacement la transition de la tête à l'épaule associée à différentes cibles dans l'espace.

L'efficacité de la seconde méthode a également été testée sur six amputés en réalité virtuelle, et une preuve de concept a été réalisée pour évaluer sa faisabilité en conditions réelles. Cette démonstration a été réalisée en contrôlant la plateforme robotique REACHY 2 en vision égocentrée, avec des marqueurs ArUco et un algorithme de vision artificielle pour détecter les objets à saisir et la main robotique. Les résultats suggèrent que, malgré les difficultés rencontrées dans la détection des objets, notre carte spatiale fonctionne efficacement dans des scénarios réels. Cette méthode pourrait également gérer des scénarios complexes, impliquant des déplacements de caméra ou des environnements perturbés.

***Mots-clefs :*** *Commande adaptative, changement de référentiel, contrôle de prothèse transhumérale, interaction homme-robot, réseau de neurones artificiels*

# Abstract

Myoelectric controls for transhumeral prostheses often lead to high rates of device abandonment due to their unsatisfactory performance. Grounded on advances in movement-based prosthesis control, we refined an alternative approach using an artificial neural network trained on natural arm movements to predict the configuration of distal joints based on proximal joint motion and movement goals. Previous studies have shown that this control strategy enabled individuals with transhumeral limb loss to control a prosthesis avatar in a virtual reality environment as well as with their valid arm. Yet, deploying this control system in real-world requires further development. A head-mounted camera and computer vision algorithms need to be integrated into the system for real-time object pose estimation. In this setup, object information might only be available in a head-centered reference frame, while our control relies on the object expressed in a shoulder reference frame.

Taking inspiration from how the brain executes coordinate transformations, we developed and tested solutions to perform the required head-to-shoulder transformation from orientation-only data, possibly available in real-life settings. To develop these algorithms, we gathered a dataset reflecting the relationship between these reference frames by involving twenty intact-limbs participants in picking and placing objects in various positions and orientations in a virtual environment. This dataset included head and gaze motion, along with movements of the trunk, shoulders, and arm joints, capturing the entire kinematic chain between the movement goal and the hand moved to reach it.

Following data collection, we implemented two methods to transform target information from the head to the shoulder reference frame. The first is an artificial neural network trained offline on the dataset to predict the head position in the shoulder referential given ongoing shoulder and head orientations and the participant height. The second method draws inspiration from multisensory integration in the brain. It derives the head position in the shoulder referential by comparing data about the prosthetic hand obtained in the shoulder referential through forward kinematics and simultaneously in the head referential through computer vision. Inspired by brain's mechanisms for peripersonal space coding, we encoded this difference in a spatial map by adapting the weights of a single-layer network of spatially tuned neurons online.

Experimental results on twelve intact-limbs participants controlling a prosthesis avatar in virtual reality demonstrated persistent errors with the first method, which failed to

adequately account for the specificity of the user's morphology, resulting in significant prediction errors and ineffective prosthesis control. In contrast, the second method elicited much better results and effectively encoded the transition from the head to the shoulder associated with different targets in space. Despite requiring an adaptation period, subsequent performances on already explored targets were comparable to the ideal scenario.

The effectiveness of the second method was also tested on six participants with transhumeral limb loss in virtual reality, and a physical proof of concept was implemented on a teleoperated robotic platform with simple computer vision to assess feasibility in real-life settings. One intact-limbs participant controlled the robotic platform REACHY 2 to grasp cylinders on a board. ArUco markers on the robot's end effector and cylinders coupled with a gaze-guided computer vision algorithm enabled precise object pose estimation. The results of this proof of concept suggest that despite challenges in object detection, our bio-inspired spatial map effectively operates in real-world scenarios. This method also shows promise for handling complex scenarios involving errors in position and orientation, such as moving a camera or operating in perturbed environments.

***Key words:*** *Adaptive control, reference frame transformation, transhumeral prosthesis control, human-robot interaction, artificial neural network*

# Acknowledgments

First of all, I would like to thank the members of the jury for agreeing to evaluate this thesis. In particular, I would like to thank Christine Azevedo Coste and Peter Ford Dominey for their roles as rapporteurs and Pauline Maurice and Jean-Louis Vercher for acting as examiners.

Words cannot express my gratitude to my supervisor, Aymar, for his support and mentoring throughout my research journey. Our paths crossed through a series of fortuitous coincidences and I am incredibly grateful for the opportunity to work under his guidance. These past three years have been challenging, yet incredibly rewarding. I've learned so much, from new programming languages to speaking French, and I owe much of my growth to him.

I would also like to acknowledge the French Direction Générale de l'Armement for funding my research. Without their support, none of this would have been possible. Following, I am deeply thankful to Dr. Rémy Klotz and the UGECAM Tour de Gassies for granting us the opportunity to conduct experiments with people with limb loss. Their support and cooperation were incredibly rewarding. Additionally, I wish to express my appreciation to all the participants who generously contributed their time and effort to our experiments over the past three years.

I extend my deepest gratitude to my colleagues Effie and Vincent. Effie, my 'thesis big sister', your support, advice, and mentorship have been indispensable throughout this journey. Your presence, both in and out of the office, has been a constant source of reassurance and inspiration. Vincent, I am truly thankful for your dedication and assistance at every stage of this project, through both the highs and the challenges.

My sincere thanks go out to all members of the Hybrid team. Daniel, your expertise and support have been invaluable assets throughout my research. Emilie, I am grateful for your continuous assistance, particularly during the experiments. I also want to thank Emilie and Etienne for giving me the incredible opportunity to participate in a parabolic flight. Furthermore, I would like to acknowledge Emeline and Lucas, the interns who contributed to my thesis, for their hard work and dedication. I am especially delighted to see Lucas continuing as a new team member.

I would like to extend my gratitude to Jean-René and Sandrine for welcoming me into their laboratory, and to all the members of the INCIA for their warm reception and assistance during my time here. Special thanks to the MOCOCO team for sharing the

office and creating enjoyable memories together.

My sincere thanks also go to everyone at Pollen Robotics for their collaboration and assistance with Reachy, and to Yannick Aoustin, Ganesh Gowrishankar, and Jean-Daniel Masson for their guidance, expertise, and invaluable advice throughout this journey.

I am profoundly grateful to my family for their support throughout my PhD journey. To my mother, whose constant presence and encouragement have been my guiding light. Despite not fully understanding my research, her belief in my abilities never wavered. To my father, whose advice and support have helped me during both the challenges and triumphs of this journey. To my brother, whose belief in my potential and encouragement, even from afar, has been a constant source of strength. And to all the women in my family, especially my grandmothers, whose remarkable strength and resilience have always inspired me.

To my friends, Andres, Sara, Chiara, Silvia, and Rosalia, I am deeply grateful. Despite the physical distance and only meeting a few times a year, your love and encouragement have always felt close, and your support has meant the world to me.

Last but certainly not least, I extend my heartfelt gratitude to Mostafa for his unwavering support throughout this journey. His encouragement and belief in me have been a steady source of reassurance and strength. Thank you for standing by me and supporting me through every step of this remarkable journey.

# Contents

**Résumé** . . . ii
**Abstract** . . . iv
**Acknowledgments** . . . vi
**Contents** . . . x
**List of Figures** . . . xii
**List of Tables** . . . xiii
**List of Abbreviations** . . . xiv
**Résumé étendu** . . . xvi
**Extended abstract** . . . xxii
1 **General introduction** . . . 2
1.1 Upper limb prostheses . . . 3
1.1.1 General information about upper limb amputation . . . 3
1.1.2 Types of upper limb prostheses . . . 4
1.1.3 Controlling upper limb prostheses . . . 5
1.2 Intuitive movement-based prosthesis control . . . 8
1.2.1 Control principle . . . 8
1.2.2 Adapting the control for real-world deployment . . . 10
2 **Dataset of natural arm and head movements** . . . 18
2.1 Materials and Methods . . . 20
2.1.1 Participants . . . 20
2.1.2 Experimental setup . . . 20
2.1.3 Virtual arm calibration . . . 21
2.1.4 Range of motion determination . . . 22
2.1.5 Task . . . 23
2.1.6 Return to Neutral Posture procedure . . . 24
2.1.7 Protocol . . . 25
2.1.8 Data reduction and metrics . . . 26
2.1.8.1 Filtering . . . 26

2.1.8.2 Metrics . . . 27
2.1.9 Statistical analysis . . . 27
2.2 Results . . . 27
2.2.1 Explored workspace . . . 28
2.3 DataPlayer . . . 29
2.4 Discussion . . . 31
3 **Two models for head to shoulder reference frame transformation** . . . 34
3.1 Armroot2Head Artificial Neural Network . . . 35
3.2 Space Map . . . 37
3.2.1 Tuning parameters . . . 40
3.2.2 Convergence and stability analysis . . . 41
3.3 Combination of Armroot2Head Artificial Neural Network and Space Map 44
4 **Experimental validation in a virtual environment** . . . 46
4.1 Materials and Methods . . . 47
4.1.1 Participants . . . 47
4.1.2 Methodological changes . . . 47
4.1.3 Protocol . . . 48
4.1.4 Data reduction and metrics . . . 49
4.1.4.1 Filtering . . . 49
4.1.4.2 Performances metrics . . . 49
4.1.4.3 Subjective metrics . . . 50
4.1.5 Statistical analysis . . . 50
4.2 Results . . . 51
4.3 Discussion . . . 54
5 **Proof of concept: from virtual reality to a robotic platform** . . . 58
5.1 Materials and Methods . . . 60
5.1.1 Robotic platform . . . 60
5.1.2 Experimental setup . . . 61
5.1.3 Protocol . . . 62
5.1.4 Metrics . . . 63
5.2 Results . . . 63
5.3 Discussion . . . 64
6 **Conclusions and perspectives** . . . 66
6.1 Conclusion . . . 67
6.2 Future directions . . . 70
**References** . . . 72

Appendix A **Publications and oral presentations** . . . . . 88
A.1 Publications . . . . . 88
A.2 Oral presentations . . . . . 88
Appendix B **Videos** . . . . . 90
Appendix C **Offline evaluations of PC ANN accuracy** . . . . . 91
Appendix D **Survey SUS** . . . . . 93
Appendix E **Survey rawProsTLX** . . . . . 94
Appendix F **Statistical analysis results of Exp1 and Exp2** . . . . . 96

# List of Figures

1.1 Upper limb amputation levels. . . . . 3
1.2 Different types of upper limb prostheses. . . . . 4
1.3 Overview of task and control strategy proposed by Mick et al., 2021. . . 9
1.4 Overview of the control strategy proposed by E. Segas et al., 2023. . . . 10
1.5 Problem definition. . . . . 11
1.6 Schematic representation of how goal-directed movements are planned in the central nervous system. . . . . 12
1.7 Neural network proposed by Pouget et al., 2002 for coordinate transformations using basis functions. . . . . 13
1.8 Anatomy of shoulder girdle and cervical spine. . . . . 15
1.9 Relationship between the head and the armroot reference frame. . . . . 16

2.1 Experimental setup in VR. . . . . 20
2.2 Final step of virtual arm calibration. . . . . 21
2.3 7-DOFs arm model. . . . . 22
2.4 Illustration of the task. . . . . 23
2.5 Return to Neutral Posture (RNP) procedure. . . . . 24
2.6 Dataset collection protocol. . . . . 25
2.7 Dataset analysis results. . . . . 28
2.8 Workspace explored in our dataset. . . . . 29
2.9 DataPlayer interface. . . . . 30
2.10 Examples of DataPlayer customization. . . . . 30

3.1 Training results for different input combinations of the A2H ANN. . . . . 36
3.2 A2H ANN offline training results. . . . . 37
3.3 Control scheme of A2H ANN combined with PC ANN. . . . . 37
3.4 Overview of task and co-adaptation process proposed by Couraud et al., 2018. . . . . 38
3.5 SMAP algorithm structure. . . . . 39
3.6 Control scheme of SMAP combined with PC ANN. . . . . 40
3.7 SMAP convergence analysis for angular errors of 60°, 88°, and 90°. . . 43
3.8 SMAP divergence curve. . . . . 44
3.9 Control scheme of SMAP and A2H combined with PC ANN. . . . . 45

4.1 Protocols of experiment 1 and 2. . . . . 48

4.2 Results of Exp1. . . . . 52
4.3 Results of Exp2. . . . . 53

5.1 Overview of the robotic platform REACHY 2. . . . . 60
5.2 Proof of concept. . . . . 61
5.3 Egocentric view of the robotic platform REACHY 2. . . . . 62
5.4 SMAP algorithm's loss. . . . . 64

C.1 Error in prediction of PC ANN. . . . . 92

# List of Tables

4.1 Exp2 participants' amputation description . . . . . 47

F.1 Exp1 Success rate - Pairwise comparisons using McNemar's test . . . . 96
F.2 Exp1 Movement time - Pairwise comparisons using Conover's all-pairs test . . . . . 96
F.3 Exp1 Shoulder spread volume - Pairwise comparisons using Conover's all-pairs test . . . . . 97
F.4 Exp1 SUS score - Pairwise comparisons using Conover's all-pairs test . 97
F.5 Exp1 rawProsTLX score - Pairwise comparisons using Conover's all-pairs test . . . . . 97
F.6 Exp2 Success rate - Pairwise comparisons using McNemar's test . . . . 97
F.7 Exp2 Movement time - Pairwise comparisons using Tukey's test . . . . . 97
F.8 Exp2 rawProsTLX score - Pairwise comparisons using Conover's all-pairs test . . . . . 98

# List of Abbreviations

**A2H** : Armroot2Head

**ADL(s)** : Activities of Daily Living

**ANN(s)** : Artificial Neural Network

**DOF(s)** : Degree(s) of Freedom

**EHI** : Edinburgh Handedness Inventory

**IMU(s)** : Inertial Measurement Units

**GNG** : Growing Neural Gas

**PC** : Proximo Contextual

**rawProsTLX** : Prosthesis Task Load indeX (first part)

**RNP** : Return to Neutral Posture

**SMAP** : Space Map

**SUS** : System Usability Scale

**VR** : Virtual Reality

# Résumé étendu

La perte de membres supérieurs affecte profondément la qualité de vie, réduisant l'autonomie individuelle dans les activités quotidiennes. Les prothèses sont des dispositifs artificiels conçus pour remplacer tout ou une partie des membres supérieurs, dans le but de restaurer les fonctionnalités essentielles et d'améliorer la qualité de vie de l'utilisateur. Les prothèses actives, en particulier, comportent une ou plusieurs articulations contrôlables, permettant aux utilisateurs de retrouver certaines fonctions motrices du membre manquant. Le schéma de contrôle prédominant pour les prothèses actives est le contrôle myoélectrique, qui repose sur des signaux recueillis à partir des muscles résiduels dans le moignon pour actionner les moteurs du dispositif.

Cette thèse se concentre sur les systèmes de contrôle des prothèses transhumérales, spécifiquement conçues pour les personnes à qui il manque une partie de leur membre au niveau de l'humérus. Les stratégies myoélectriques existantes pour contrôler ces dispositifs permettent principalement aux utilisateurs de manipuler séquentiellement la flexion/extension du coude, la pronation/supination du poignet et l'ouverture/fermeture de la main. Cependant, l'absence de contrôle sur les deux autres degrés de liberté (DDL) du poignet (flexion/extension et déviation latérale) limite la capacité d'orienter la main prothétique avec précision pour des tâches telles que la saisie d'objets. Par conséquent, les utilisateurs expriment la nécessité de dispositifs capables d'effectuer des actions de manière coordonnée avec tous les DDL manquants. Cette nécessité, associée à l'effort cognitif et physique considérable requis pour faire fonctionner les dispositifs existants, contribue significativement aux taux élevés d'abandon de prothèses.

Des alternatives prometteuses aux systèmes de contrôle myoélectrique sont les approches basées sur le mouvement, dans lesquelles le membre prothétique est contrôlé en fonction du mouvement du membre résiduel et de sa coordination naturelle avec les articulations manquantes. En général, ces méthodes apprennent la coordination entre les articulations proximales et distales grâce à des techniques de régression non linéaire, telles que les réseaux de neurones artificiels (RNA), les réseaux de fonctions à base radiale ou les régressions localement pondérées. Cependant, cartographier la relation entre le mouvement de deux articulations proximales et le mouvement de cinq articulations distales est intrinsèquement complexe. Pour relever ce défi, des informations contextuelles, telles que la position et l'orientation de l'objet à saisir, acquises grâce à des algorithmes de vision par ordinateur, ont été intégrées en tant que signaux d'entrée dans le système de contrôle. S'appuyant sur les avancées dans les con-

trôles de prothèses à base de mouvements naturels et de vision par ordinateur, mes collègues ont démontré qu'en ajoutant l'information sur l'objectif de mouvement aux informations de mouvements des articulations proximales, un RNA peut prédire avec précision la configuration des articulations distales nécessaire pour positionner et orienter la main de manière précise pour saisir l'objet ciblé. Le RNA, appelé Proximo Contextuel (PC), entraîné à partir des mouvements de bras de plusieurs participants, exploite la coordination naturelle du bras pour prédire la configuration des articulations distales. Cela se traduit par un contrôle plus naturel et intuitif du membre prothétique, ce qui a permis aux personnes ayant subi une amputation transhumérale de contrôler un avatar de prothèse et d'effectuer des tâches de prise et de placement d'objet dans un large espace dans un environnement de réalité virtuelle (RV) comme ils le feraient avec leur bras valide.

Malgré les résultats remarquables en RV, le déploiement du système de contrôle PC RNA dans des applications réelles présente des défis importants. Alors que les environnements de RV offrent un cadre idéal où toutes les informations nécessaires sont facilement accessibles, les scénarios du monde réel introduisent des complexités qui doivent être prises en compte. Par exemple, les utilisateurs peuvent porter une caméra montée sur la tête et des unités de mesure inertielle (Inertial Measurement Units – IMU en anglais) attachées à la caméra, au tronc et au membre résiduel. Dans ces conditions, les informations sur les objectifs de mouvement (c'est-à-dire l'une des entrées du système de contrôle) peuvent n'être disponibles que dans un référentiel centré sur la tête, grâce à la vision par ordinateur guidée par le regard. Cependant, le système de contrôle PC RNA repose également sur la position et l'orientation de la cible, exprimées dans un référentiel appelé armroot, dont l'origine se situe au niveau de l'articulation de l'épaule et dont l'orientation est déterminée par un capteur fixé sur le tronc de l'utilisateur. Par conséquent, les objectifs de mouvement, disponibles dans un référentiel centré sur la tête grâce à la vision par ordinateur guidée par le regard, doivent être transformés dans le référentiel armroot pour que le contrôle de la prothèse basée sur le PC RNA fonctionne efficacement. Bien que l'orientation de la tête et de l'armroot puisse être obtenues grâce aux IMU, les données de position ne sont pas facilement disponibles pour effectuer cette transformation.

L'objectif de cette thèse était d'adapter le contrôle PC RNA à des situations réelles, en abordant ce problème de transformation de référentiel. Plus précisément, l'objectif était de développer des méthodes pour récupérer la position de la tête dans le référentiel armroot.

Dans le domaine de la robotique, la résolution des problèmes de transformation des coordonnées, tels que la récupération de données sur une partie spécifique du corps, nécessite souvent l'utilisation de modèles cinématiques, qui décrivent analytiquement les relations entre les différents référentiels au sein de la chaîne. Cependant, la construction d'une chaîne cinématique de la tête à l'épaule représente un défi important en raison de l'anatomie complexe et des interactions dynamiques inhérentes à la colonne cervicale et à l'articulation de l'épaule. Par conséquent, les modèles cinématiques existants ont souvent recours à des simplifications au sein de la chaîne (par exemple, en omettant les capacités de translation au niveau de l'épaule), ce qui résulte en une

représentation partielle des mouvements complexes observés dans cette région.

Compte tenu de la complexité de la relation entre les deux référentiels, nous avons préféré aborder le problème en considérant le système comme une boîte noire et en apprenant son comportement entrée-sortie.

La première étape cruciale a consisté à collecter des données incluant la relation entre les référentiels de la tête et de l'armroot. Nous avons donc recueilli une base de données comprenant les performances de vingt participants valides effectuant des tâches de prise et de déplacement d'objets dans un environnement de RV. L'espace de travail dans lequel les participants ont exécuté la tâche a été maximisé grâce à une procédure systématique. Dans un premier temps, l'espace de travail le plus large possible a été déterminé par l'amplitude maximale des mouvements des participants, puis une carte auto-organisatrice a été utilisée pour mieux représenter l'espace réel couvert par les participants. De plus, un contrôle précis de la posture initiale du corps, à partir de laquelle les mouvements du bras ont débuté, a été réalisé en guidant les participants vers une posture neutre prédéterminée via des retours visuels du tronc et des épaules.

Parce qu'elle inclue les mouvements de la tête et du regard, ainsi que les mouvements du tronc, des épaules et des articulations des bras, notre base de données représente également une ressource précieuse dans divers domaines tels que le contrôle sensorimoteur, le contrôle des prothèses basé sur le mouvement, les interactions homme-robot, la robotique humanoïde, et surtout la vision par ordinateur guidée par le regard. En plus de la base de données, nous avons développé un projet Unity appelé DataPlayer, conçu pour rejouer et visualiser des données en 3D. L'objectif de ce DataPlayer était de créer un outil pour la recherche en vision par ordinateur, permettant de concevoir et de tester des algorithmes efficaces d'estimation de la position et l'orientation en 6D dérivés de la vision égocentrique avec les données du regard. Étant donné que la vraie information de l'objet est connue par conception dans la RV, et que l'environnement visuel peut être rejoué et manipulé à volonté, cette base de données peut être utilisée pour générer des données synthétiques cohérentes avec les mouvements physiologiques de la tête, du regard et du bras. Cette capacité est cruciale pour développer et tester des algorithmes d'estimation de la pose 6D dans des contextes réalistes.

Néanmoins, l'objectif premier de cet ensemble de données était de servir de base au développement et à la validation de deux méthodes pour résoudre le problème de la transformation de référentiel et déterminer la position de la tête dans le référentiel de armroot. Contrairement aux stratégies traditionnelles basées sur des modèles, qui impliquent la construction d'un modèle de la tête à l'armroot, nous avons abordé le problème en considérant le système comme une boîte noire et en apprenant le comportement entrée-sortie sans dépendre explicitement de modèles prédéfinis. Cette approche a permis de contourner les complexités associées à la construction de modèles et a offert une flexibilité et une adaptabilité dans la résolution du problème de transformation de référentiel.

La première approche, Armroot2Head (A2H), impliquait l'entraînement hors ligne d'un

RNA utilisant la base de données pour effectuer cette transformation. Le RNA prédit la position de la tête dans le référentiel de l'armroot en fonction des orientations de la tête et de l'épaule et de la taille du participant.

La deuxième approche, la carte spatiale (SMAP), code spatialement la position de la tête dans le référentiel de l'armroot dans une carte spatiale péripersonnelle bio-inspirée. Inspiré par les processus d'intégration multisensorielle du cerveau, l'algorithme SMAP calcule la position de la tête dans le référentiel de l'armroot en combinant des informations provenant de modalités sensorielles multiples (à savoir, la vision par ordinateur à partir d'une caméra montée sur la tête, combinée à la proprioception artificielle du cou et du bras à partir de capteurs placés sur la tête, le tronc et le bras). Plus précisément, le système peut calculer la position de la main prothétique dans le référentiel de l'armroot grâce à la cinématique directe tout en détectant simultanément la position de la main prothétique dans le référentiel de la tête à l'aide de la vision par ordinateur. Comme la valeur de la position de la tête par rapport à l'épaule dépend non seulement de la morphologie de l'utilisateur mais aussi de la posture spécifique de son bras et de son corps lorsque celui-ci atteint des cibles dans l'espace, nous avons opté pour la stocker dans une carte spatiale péripersonnelle. En s'inspirant de l'encodage de l'espace péripersonnel dans le cerveau et en tirant parti des idées issues de recherches antérieures menées par notre équipe sur les stratégies d'adaptation humaine, nous avons mis en œuvre un mécanisme d'encodage spatial utilisant un réseau de neurones d'une seule couche, disposés spatialement et accessible en fonction de la cible sur laquelle l'utilisateur fixe son regard et qu'il vise à atteindre et à saisir.

Les évaluations hors ligne de RNA A2H ont révélé une erreur de prédiction moyenne de 2 cm. En capitalisant sur la capacité du SMAP à apprendre des erreurs, nous avons mis en œuvre et testé une troisième solution, qui combine RNA A2H avec le SMAP, en utilisant ce dernier pour corriger les erreurs commises par le premier.

Nous avons mené deux expériences dans un environnement de RV impliquant des participants valides et amputées au niveau transhuméral pour comparer la performance de ces méthodes avec celle du scénario idéal PC, dans lequel le changement de référentiel de la tête à armroot a été exécuté avec une grande précision en utilisant toutes les informations disponibles dans le système de RV.

Les résultats expérimentaux sur douze participants valides ont démontré des erreurs persistantes avec le A2H RNA, entraînant des mouvements compensatoires de grande amplitude et des temps de mouvement plus longs. En revanche, la SMAP a présenté des résultats positifs. Bien que le processus d'apprentissage ait nécessité une période d'adaptation initiale, les performances sur des cibles déjà explorées ont été comparables au scénario idéal. Cela souligne l'efficacité de la deuxième approche dans la résolution du problème de transformation de référentiel et dans la détermination précise de la position de la tête dans le référentiel de l'armroot. La combinaison des deux méthodes a également donné des résultats favorables. Néanmoins, la deuxième approche seule s'est avérée suffisante pour la tâche, comme en témoigne l'expérience menée auprès de six participants amputés au niveau transhuméral. Ces derniers ont contrôlé efficacement le bras virtuel, atteignant un taux de réussite dépassant 95%,

tout en minimisant les mouvements compensatoires et en maintenant des temps de mouvement comparables au scénario idéal. Ils ont exprimé une grande satisfaction à l'égard du système de contrôle, manifestant un vif intérêt pour l'utilisation de ces méthodes de contrôle naturelles avec de vraies prothèses.

Nous avons également démontré l'efficacité du SMAP dans la détermination de la position de la tête dans le référentiel de l'armroot dans un scénario simple du monde réel, dans lequel les informations sur les objets et la main provenaient véritablement du référentiel de la tête grâce à la vision par ordinateur. En raison du manque de prothèses commercialement disponibles pour tester nos méthodes, nous avons réalisé une preuve de concept en téléopérant la plateforme robotique REACHY 2. En fixant une caméra Zed Mini sur la tête du robot et en la connectant au casque de réalité virtuelle, nous avons permis à l'utilisateur portant le casque de contrôler le robot depuis un point de vue égocentrique. Les cibles et la main du robot ont été détectés à l'aide de marqueurs ArUco et d'un algorithme de vision par ordinateur guidé par le regard. Un seul participant ayant une expérience préalable a contrôlé le robot pour effectuer des tâches de prise et de placement d'objet, alors que le SMAP a été utilisé pour calculer la position de la tête du robot dans son référentiel de l'armroot. Malgré les défis liés à l'absence de données visuelles, comme lorsque l'objet est hors champ de vision ou obstrué, la mémoire associée à notre codage spatial dans le SMAP s'est révélée robuste dans de telles situations, assurant un fonctionnement continu et fluide. Les résultats positifs de cette preuve de concept soulignent l'efficacité de l'algorithme, démontrant sa capacité à fonctionner efficacement dans des scénarios du monde réel.

Tout au long de cette thèse, nous avons utilisé l'algorithme SMAP pour calculer la position de la tête dans le référentiel de l'armroot et effectuer la transformation de référentiel de la tête à l'armroot. Les conclusions présentées ici contribuent non seulement à l'avancement des mécanismes de contrôle de prothèses, mais offrent également des perspectives précieuses sur les applications plus larges de l'algorithme SMAP. Celles-ci incluent divers scénarios impliquant des transformations de référentiel, une adaptation en ligne et des corrections d'erreurs tant en position qu'en orientation.

Le manuscrit est structuré comme suit :

Le **Chapitre 1** offre une vue d'ensemble de la perte du membre supérieur, des prothèses et des différentes stratégies de contrôle. Il présente les recherches antérieures de l'équipe sur le contrôle des prothèses et les questions centrales abordées dans cette thèse.

Le **Chapitre 2** détaille le protocole expérimental et l'approche méthodologique utilisés pour collecter la base de données des mouvements naturels de la tête et du bras, et en présente le contenu ainsi que les premières analyses.

Le **Chapitre 3** décrit en détail deux méthodes permettant d'effectuer la transformation tête-épaule nécessaire au déploiement de notre contrôle de prothèse basé sur le mouvement dans le monde réel.

Le **Chapitre 4** décrit deux expériences menées dans un environnement de RV avec des participants valides et des amputés transhuméraux pour tester les méthodes pro-

posées et les stratégies de contrôle qui y sont associées.

Le **Chapitre 5** présente une démonstration de faisabilité pour valider le travail de notre méthode dans un scénario réel en utilisant la plateforme robotique REACHY 2 et des algorithmes de vision par ordinateur.

Le **Chapitre 6** offre un résumé complet des conclusions et des résultats obtenus dans le cadre de cette thèse et présente les bases de futures recherches, en soulignant la pertinence du présent travail.

# Extended abstract

Upper limb loss profoundly affects the quality of life, diminishing individual autonomy in daily activities. Prostheses are artificial devices designed to replace all or part of the upper extremities, aiming to restore essential functionalities and significantly enhance overall user quality of life. Active prostheses, in particular, feature one or more controllable joints, enabling users to regain some motor functions of the missing limb. The predominant control scheme for active prostheses is myoelectric control, which relies on signals collected from residual muscles in the stump to drive the device's motors.

This thesis focuses on control systems for transhumeral prostheses, specifically designed for individuals missing part of their limb from the humerus. Existing myoelectric strategies for controlling these devices primarily focus on enabling users to sequentially manipulate elbow flexion/extension, wrist pronation/supination, and hand opening/closing. However, the lack of control over the remaining two degrees of freedom (DOFs) of the wrist (flexion/extension and lateral deviation) hinders the ability to orient the prosthetic hand accurately for tasks such as grasping objects. Consequently, users express the necessity for devices that can perform actions in a coordinated manner with all missing DOFs. This necessity, coupled with the considerable cognitive and physical exertion required to operate existing devices, significantly contributes to the high rates of prosthesis abandonment.

Promising alternatives to myoelectric control systems are movement-based approaches, whereby the prosthetic limb is controlled based on the motion of the residual limb and its natural coordination with the missing joints. Generally, these methods learn the coordination between proximal and distal joints through nonlinear regression techniques, such as artificial neural networks (ANNs), radial basis function networks, or locally weighted regressions. However, mapping the relationship between the movement of two proximal joints and the movement of five distal joints is inherently complex. To address this challenge, contextual information, such as the position and the orientation of the object to manipulate, acquired through computer vision algorithms, have been integrated as input signals in the control system. On this basis, my colleagues have demonstrated that integrating information regarding movement goals with proximal joint motion enables an ANN to accurately predict the distal joint configuration necessary for precise positioning and orientating of the hand to grasp the targeted object. The ANN, named Proximo Contextual (PC), trained on data from multiple participants performing the task, leverages natural arm coordination to predict the distal joint configuration.

This results in a more natural and intuitive control of the prosthetic limb, which has enabled individuals with transhumeral limb loss to control a prosthesis avatar and perform pick-and-place tasks in a wide range of locations in a VR environment as well as with their valid arm.

Despite the remarkable results in VR, deploying the PC ANN control system in real-world applications presents significant challenges. While VR environments offer an ideal setting where all necessary information is readily accessible, real-world scenarios introduce complexities that must be addressed. For instance, users may wear a head-mounted camera and Inertial Measurement Units (IMUs) attached to the camera, trunk, and residual limb. In such conditions, information about movement goals (i.e., one of the control system inputs) may only be available in a head-centered reference frame through gaze-guided computer vision. Yet, the PC ANN control system relies on target position and orientation expressed in a body reference frame known as armroot, with its origin located at the shoulder articulation and orientation of a sensor attached to the user's trunk. Therefore, movement goals available in a head-centered reference frame through gaze-guided computer vision need to be transformed into the armroot reference frame for the prosthesis control based on PC ANN to operate effectively. While orientation data can be obtained through IMUs, positional data may not be readily available to perform this transformation.

The objective of this thesis was to adapt the PC ANN control to real-world settings, addressing this reference frame transformation issue. Specifically, the goal was to develop methods to retrieve the head position in the armroot reference frame.

In the realm of robotics, addressing coordinate transformation issues, such as retrieving data about a specific body part, frequently necessitates the utilization of kinematic models, which analytically describe the relationships between the different reference frames within the chain. However, constructing a kinematic chain from the head to the shoulder presents a significant challenge due to the intricate anatomy and dynamic interactions inherent in the cervical spine and shoulder girdle. Consequently, existing kinematic models often resort to simplifications within the chain (e.g., omitting translational capabilities at the shoulder level), resulting in a partial representation of the complex movements observed in this region.

Given the complexity of the relationship between the two frames, we opted to approach the problem as a black box by learning its input-output behavior.

The first crucial step involved gathering data that accurately reflects the relationship between the head and armroot reference frames. To this end, we collected a dataset comprising performances of twenty intact-limbs participants engaged in pick-and-place tasks within a VR setting. The workspace in which participants executed the task was maximized through a systematic procedure. Initially, the widest possible workspace was determined by the maximal range of motion of the participants. Then, a self-organizing map was employed to delineate the actual space covered by participants. Additionally, precise control over the initial body posture, from which arm movements were produced, was achieved by guiding participants to a predetermined neutral posture via visual feedback from the trunk and shoulders.

Including head and gaze motion, alongside movements of the trunk, shoulders, and arm joints, our dataset could represent a valuable resource across diverse domains such as sensorimotor control, movement-based prosthesis control, human-robot interactions, humanoid robotics, and foremost gaze-guided computer vision. In addition to the dataset, we developed a Unity project named DataPlayer, designed for replaying and visualizing data in 3D. The goal of this DataPlayer was to create a tool for computer vision research, enabling the design and testing of efficient 6D pose estimation algorithms derived from egocentric vision with gaze data. Since the ground truth of the object position and orientation is known by design in VR, and the visual environment can be replayed and manipulated at will, this dataset can be used to generate synthetic data consistent with physiological head, gaze, and arm movements. This capability is crucial for developing and testing 6D pose estimation algorithms in realistic contexts.

Nevertheless, the primary objective of this dataset was to serve as the basis for developing and validating two methods to solve the reference frame transformation issue and determine the head position within the armroot reference frame. In contrast to traditional model-based strategies, which entail constructing a model from the head to the armroot, we addressed the problem by treating the system as a black box and learning the input-output behavior without explicit reliance on predefined models. This approach bypassed the complexities associated with model construction and offered flexibility and adaptability in addressing the reference frame transformation issue.

The first approach, Armroot2Head (A2H), involved training offline an ANN using the dataset to do this transformation. The ANN predicts the head position in the armroot reference frame based on ongoing head and shoulder orientations and the participant's height. The second approach, Space Map (SMAP), spatially encoded the head position in the armroot reference frame within a bio-inspired peripersonal space map. Inspired by the brain's multisensory integration processes, the SMAP algorithm computes the head position in the armroot reference frame by combining information from multiple sensory modalities (i.e., computer vision from a head-mounted camera, combined with artificial neck and arm proprioception from sensors on the head, trunk and arm). Specifically, the system can compute the prosthetic hand position in the armroot referential through forward kinematics while concurrently detecting the prosthetic hand position in the head referential using computer vision. As the value of the head position is contingent not only upon the user's morphology but also on the specific posture of their arm and body while reaching targets in space, we opted for storing it in a peripersonal space map. Drawing inspiration from the encoding of peripersonal space in the brain and leveraging insights from previous research conducted by our team on human adaptation strategies, we implemented a spatial encoding mechanism using a single-layer network of spatially tuned neurons accessed based on the target the user gazes at and aims to reach and grasp.

Offline assessments of the A2H ANN revealed a mean prediction error of 2cm. Capitalizing on the SMAP's capacity to learn from errors, we implemented and tested a third solution, which combines the A2H ANN with the SMAP, employing the latter to correct errors made by the former.

We conducted two experiments within a VR environment involving participants with intact limbs and with transhumeral limb loss to compare the performance of these methods with that of the ideal scenario PC, wherein the change of reference frame from the head to the armroot was executed with high precision utilizing all available information within the VR system.

Experimental results on twelve intact-limbs participants demonstrated persistent errors with the first method, leading to extended compensatory movements and prolonged movement times. In contrast, the second approach exhibited positive outcomes. Although the learning process entailed an initial adaptation period, its subsequent performance on previously explored targets demonstrated comparability to the ideal scenario. This underscores the efficacy of the second approach in addressing the reference frame transformation issue and accurately determining the head position within the armroot reference frame. The combination of both methods also yielded favorable results. Nevertheless, the second approach alone proved sufficient for the task, as evidenced by the experiment involving six participants with transhumeral limb loss. They effectively controlled the virtual arm, achieving a success rate exceeding 95%, while minimizing compensatory movements and maintaining movement times comparable to those of natural movements. They reported high satisfaction with the control system, expressing a keen interest in utilizing such natural control methods with real prostheses.

We also demonstrated the effectiveness of the SMAP in determining head position in the armroot reference frame in a simple real-world scenario, in which the information about the objects and the hand were genuinely coming in the head referential from computer vision. Due to the lack of commercially available prostheses for testing our methods, we conducted a proof of concept teleoperating the robotic platform REACHY 2. By mounting a Zed Mini camera on the head of the robot and connecting it to the VR headset, we enabled the user wearing the headset to control the robot from an egocentric point of view. The targets and the robot's end effector were detected using ArUco markers and a gaze-guided computer vision algorithm. While a single participant with prior experience controlled the robot to perform pick-and-place tasks, the SMAP computed the position of the robot's head in its armroot reference frame. Despite challenges arising from missing visual data, such as when outside the field of vision or obstructed, the memory associated with our spatial coding in SMAP proved robust to such situations, ensuring continued smooth operation. The successful outcomes of this proof of concept underscored the algorithm's efficacy, showcasing its ability to function effectively in real-world scenarios.

Throughout this thesis, we have leveraged the SMAP algorithm to compute the head position within the armroot reference frame and perform the reference frame transformation from the head to the armroot. The findings presented here not only contribute to the advancement of prosthesis control mechanisms but also offer valuable insights into the broader applications of the SMAP algorithm. These include diverse scenarios involving reference frame transformations, online adaptation, and error corrections in both position and orientation.

The manuscript is structured as follows:

**Chapter 1** provides a comprehensive overview of upper limb loss, prosthetic devices, and various control strategies. It outlines the team's previous research in prosthesis control and the central issues addressed in this thesis.

**Chapter 2** details the experimental protocol and the methodological approach employed for collecting the dataset of natural head and arm movements.

**Chapter 3** extensively describes two methods to perform the head-to-shoulder transformation necessary for deploying our movement-based prosthesis control in real-world settings.

**Chapter 4** outlines two experiments conducted within a VR environment involving participants with intact limbs and with transhumeral limb loss to test the proposed methods and their associated control strategies.

**Chapter 5** presents a proof of concept to validate our method work in a real-life scenario using the robotic platform REACHY 2 and computer vision algorithms.

**Chapter 6** offers a comprehensive summary of the findings and results obtained in this thesis and lays the groundwork for future investigations, emphasizing the relevance of the present work.

# Chapter 1

# General introduction

This introductory chapter offers an overview of upper limb loss, prosthetics, and control strategies, laying the groundwork for this thesis. It explores the implications of upper limb loss and its physiological and societal impact, highlighting the crucial role of prosthetic technology in restoring autonomy. After examining the evolution of prosthetics and control strategies, the chapter outlines the team's previous research endeavors in prosthesis control and the central issues addressed in this thesis.

## 1.1 Upper limb prostheses

### 1.1.1 General information about upper limb amputation

Upper limb loss refers to the partial or complete absence of the upper limb. This condition profoundly impacts the quality of life for those affected by reducing their autonomy in daily activities. As illustrated in Figure 1.1, there are seven levels of upper limb amputation (Cordella et al., 2016): fingers or partial hand (transcarpal), wrist disarticulation, transradial amputation (below the elbow), elbow disarticulation, transhumeral amputation (above the elbow), shoulder disarticulation and forequarter amputation. The work proposed in this thesis addresses the population with limb loss at the transhumeral level (i.e., the absence of the limb from a portion of the humerus).

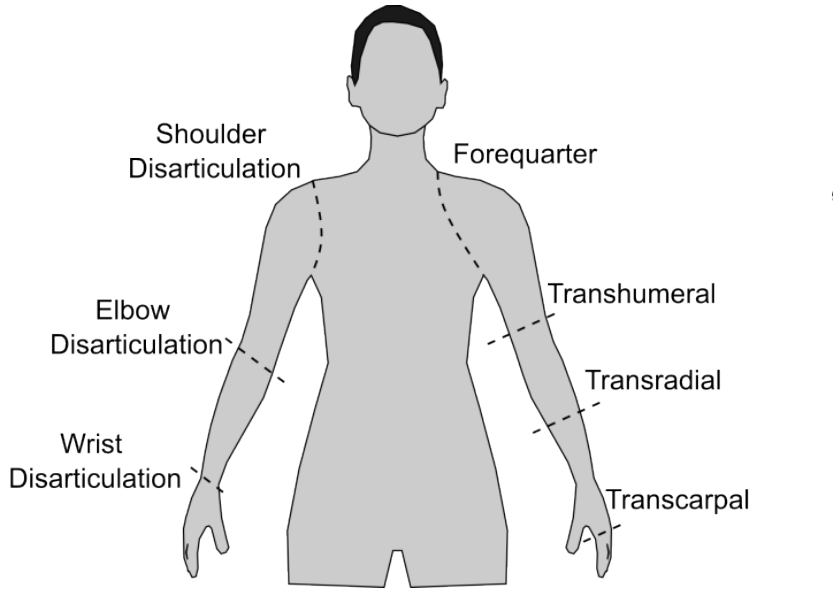



Figure 1.1: Upper limb amputation levels (Figure adapted from Figure I.1.1 in Ségas, 2023).

The population affected by upper limb loss can be categorized into two groups. The first group includes individuals suffering from agenesis, meaning that all or part of the limb did not form during embryonic development. The second group includes individuals who have undergone an amputation, a surgical procedure to remove all or part of a limb due to an accident, vascular diseases, or the presence of cancerous tumors.

In France, the Haute Autorité de Santé, 2010 documented that the number of individuals in the first category was 0.15 per 100,000 inhabitants in 2010. Regarding limb amputations, between 2011 and 2020, the "Scansanté.fr" database recorded 1616 upper limb amputations, with a majority occurring at the transhumeral level, averaging 76 amputations per year (Bruyant et al., 2023). While the overall number of people suffering from limb loss may seem low compared to the general population, the impact on the patient's quality of life and autonomy is significant (Gallagher et al., 2011; Jo-

hansen et al., 2018; Resnik et al., 2019). The higher the level of amputation, the more significant these impacts become.

### 1.1.2 Types of upper limb prostheses

Prostheses are artificial devices designed to replace all or part of the upper extremities. They play a crucial role in supporting individuals with upper limb loss by restoring essential functionalities and significantly enhancing their overall quality of life. According to the review by Ribeiro et al., 2019, upper limb prostheses can be classified into two categories based on the type of interaction with the user: passive prostheses, subdivided into aesthetic and functional, and active prostheses, subdivided into mechanical and electric (see Figure 1.2).

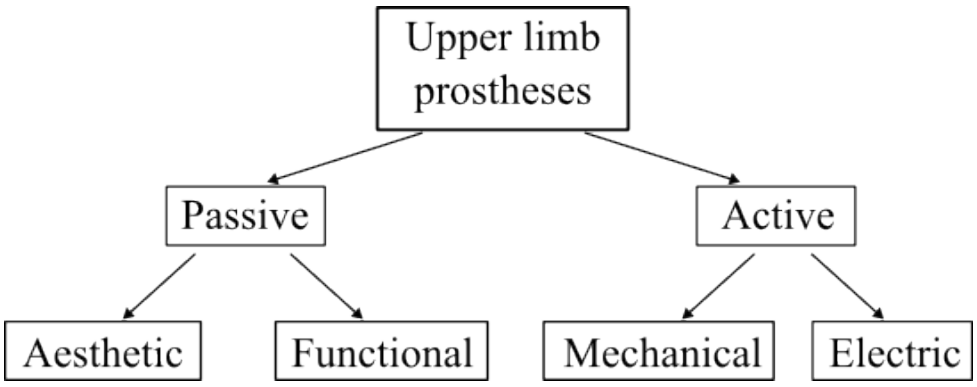



Figure 1.2: Different types of upper limb prostheses (Figure adapted from Ribeiro et al., 2019).

Passive prostheses are motionless and do not have active joints or degrees of freedom (DOFs). Aesthetic prostheses aim to complete the wearer's body image by closely mimicking the missing limb. Meanwhile, functional prostheses enable the performance of specific daily tasks by providing the ability to interface the prosthetic limb with tools, such as utensils and toothbrushes.

Active prostheses have one or more controllable joints, allowing the user to restore at least some motor functions of the missing limb. The activation of motors can result from either a mechanical or an electrical action. In prostheses with mechanical action, the motion of the prosthetic joints is triggered by the movement of another body part, usually the shoulder of the sound limb, through the transfer of force produced via cables. This system aims to re-establish a connection between the motion of a body part and an action in the prosthesis. However, since the mechanism of these prostheses relies on a harness, users are required to have sufficient strength and range of motion to manipulate the device effectively.

Prostheses with electrical action often employ motors controlled by buttons manipulated by the sound limb or motors controlled by signals collected from residual muscles in the stump.

### 1.1.3 Controlling upper limb prostheses

While cosmetics, electronic components, and computational efforts have made significant advancements, the control strategies currently employed in prosthetic devices have undergone relatively modest evolution since the 1960s, when myoelectric activity was used for the first time as a source of information regarding voluntary movement (Marinelli et al., 2023; Schmidl, 1965).

Electrically powered prostheses mainly rely on myoelectric activity recorded from surface electrodes placed on the skin over two antagonistic muscles of the residual limb (Farina and Sartori, 2016; Marinelli et al., 2023). These signals convey information about neuromuscular activity, offering insights into the intentions of individuals and, thus, a direct link between muscle contractions and prosthetic movements. Myoelectric control mechanisms can be classified into two main categories: "on-off" and proportional (Fougner et al., 2012). In the "on-off" control, signals exceeding a threshold trigger a constant-speed movement along one direction, while proportional control adjusts joint movement speed based on muscle contraction intensity. Nevertheless, both control types do not allow simultaneous control of more than one DOF. Joint movement is achieved sequentially by switching from one joint to another through co-contraction (simultaneous contraction of two antagonist muscles). For instance, in individuals with transhumeral limb loss, the control system relies on just two residual muscles, the biceps and triceps, to sequentially manage the complex movements of all four missing joints, along with the opening and closing of the hand. In such situations, myoelectric controls face a dimensionality challenge, as the control system must support a higher number of joints with two residual muscles. Moreover, since these muscles are not naturally engaged in controlling certain joints, this complicates the learning process and imposes a cognitive burden on the user. Additionally, issues like sweat, electrode displacement, or muscle fatigue may affect skin electrode signals, introducing inconsistencies between user commands and prosthesis responses (De Luca, 1997; Kyranou et al., 2018).

Pattern recognition approaches have emerged as solutions to overcome these challenges. These approaches employ machine learning techniques to interpret and categorize muscle signals, assigning them to specific movement classes (Hahne et al., 2014; Hudgins et al., 1993) and thus enabling simultaneous control of multiple DOFs (Hahne et al., 2018; Ortiz-Catalan et al., 2014). Regarding transhumeral limb loss, a promising strategy for achieving more intuitive prosthesis control involves integrating pattern recognition approaches with Targeted Muscle Reinnervation (TMR), a nerve-transfer surgery that redirects nerves from the amputated limb to nearby muscles (Mereu et al., 2021). By generating more natural control signals, TMR enables the simultaneous control of multiple DOFs, allowing actions such as flexing the elbow and closing the hand concurrently (Hargrove et al., 2017; Mereu et al., 2021).

Despite these promising advancements in upper limb prosthetics technology, the rates of device abandonment remain notably high (Salminger et al., 2022). According to a meta-analysis covering articles published between 1982 and 2007, the rejection rate for active prostheses in the adult population ranges between 20-30% (Biddiss and Chau,

2007) due to lack of functionality, comfort, inability to perform actions in a coordinated manner (simultaneous control of DOFs), and excessive visual attention requirements. User expectations include basic grasping actions, additional wrist movements, and the ability to control multiple joints simultaneously for the execution of Activities of Daily Living (ADLs), as reported by Cordella et al., 2016 and by Atkins et al., 1996 over three decades ago. Yet, most users rejecting the use of prostheses might reconsider their choice if affordable technological improvements were available (Biddiss and Chau, 2007).

To overcome the limitations of myoelectric control, researchers have explored alternative approaches whereby the prosthetic limb is controlled based on the motion of the residual limb and its natural coordination with the missing joints. In addition to being non-invasive and independent of the number of residual muscles, these approaches offer the advantage of restoring a connection between the movement of a body part and the prosthesis's interaction with the environment.

In 1998, D. Popovic and Popovic, 1998 demonstrated the feasibility of deriving elbow flexion/extension speed from shoulder speed using a multiplicative factor for horizontal pointing tasks. Successively, this factor was iteratively fine-tuned based on the spatial zone selected by the user (M. Popovic and Popovic, 2001). Similarly, Iftime et al., 2005 trained an Artificial Neural Network (ANN) to reconstruct existing synergies between the angular velocities of shoulder flexion/extension, elbow flexion/extension, and forearm pronation/supination. Their study demonstrated that while the ANN effectively captured the synergies for specific target movements, these relationships did not generalize across the entire workspace. This finding suggests that to achieve effective control, synergies involving a broader range of joint angles and movements must be exploited, highlighting the complexity and variability inherent in human arm control.

Building on this hypothesis, several research teams have proposed to enhance the quantity of pertinent information provided as input to the system by including more DOFs of the residual limb. The control scheme proposed by Kaliki et al. (Kaliki, Davoodi, and Loeb, 2008; Kaliki, Davoodi, and Loeb, 2008; Kaliki et al., 2013) integrated three ANNs to reconstruct elbow flexion/extension, wrist pronation/supination, and hand opening/closing from shoulder rotations and translations. Participants employed this control strategy to control an avatar arm in a virtual reality (VR) environment to reach an object placed on a table, grasp it, and bring it to the mouth. However, the observed task completion time was twice as long as the natural time, even after ten sessions of practice, and the proposed control scheme resulted in the non-natural use of shoulder translations.

Along the same line of thinking, the model proposed by Merad et al., 2016, reconstructing the speed of a prosthetic elbow using the residual limb angular velocities, enabled intact-limbs participants to use the prosthetic forearm attached to their upper arm to reach real targets in a 3D workspace. The same model, trained on data from intact-limbs participants, was then used by participants with limb loss to perform the same task (Merad et al., 2019, 2020). Despite the movement time being twice as long as expected for a natural movement, it remained faster than that achieved with conventional

myoelectric control. Furthermore, participants with limb loss reported that this control method was more intuitive than myoelectric control strategies.

In a different approach, Legrand et al., 2022 proposed a control system that utilizes compensatory movements executed by the user as an error signal to correct inaccuracies in the positioning of the prosthesis. This approach creates a coupling between body compensations and prosthetic movements to simultaneously control elbow flexion/extension and wrist pronation/supination. Four participants with transhumeral limb loss tested the control system while performing the Refined Rolyan Clothespin Test with an experimental prosthesis. The experiment demonstrated that the proposed method restores simultaneity between prosthetic joints while maintaining the level of performance of conventional myoelectric control without increasing compensatory motions or cognitive load.

Another interesting proposal by Gloumakov et al., 2019 involved encapsulating natural movement synergies observed during daily tasks into trajectory prototypes (i.e., a set of joint movements) and employing these trajectory prototypes to control a prosthesis (Gloumakov et al., 2020, 2022). In this method, users choose the prototype that best aligns with the task to control prosthetic joints movements. Experimental results demonstrated that task completion was faster using this simultaneous prosthesis control of the elbow and wrist than sequential myoelectric control. Despite exhibiting comparable compensatory movements, participants reported that the proposed method felt more natural and required less mental load than myoelectric control. However, the authors have not yet suggested an interface that is easy to use and can be integrated into a prosthesis for selecting among the proposed trajectory prototypes.

In recent years, there have been significant developments in the field of computer vision, especially applied to robotics. Notably, these advancements have focused on object pose estimation and identifying the optimal grip type necessary to reach and manipulate a given object (Du et al., 2019; Moullet et al., 2023). Since the 1980s, several studies have suggested that hand configurations for grasping an object or performing a task could be deduced from the properties of these objects or tasks through rule-based systems (Cutkosky, 1989) or ANNs (Iberall, 1988). On this basis, prosthesis control strategies have been devised by integrating residual limb motion with information about desired hand configuration, acquired through computer vision methods such as decision trees (Došen and Popović, 2011; Došen et al., 2010), finite-state machines (Markovic et al., 2014, 2015) and neural networks (DeGol et al., 2016; González-Díaz et al., 2019; Pérez De San Roman et al., 2017). In particular, the application of deep convolutional neural networks trained using egocentric videos and gaze data has exhibited remarkable efficacy in real-time object recognition across intricate real-world environments (González-Díaz et al., 2019; Pérez De San Roman et al., 2017).

Building on these advancements, my colleagues have developed a promising movement-based control system for transhumeral prostheses, which integrates information about movement goals with proximal joint motion, as explained in detail in the following section.

## 1.2 Intuitive movement-based prosthesis control

### 1.2.1 Control principle

The existing literature on movement-based approaches for transhumeral prostheses is limited to controlling one or two missing joints. Notably, most efforts have focused on reconstructing the elbow from the shoulder, aiming to place the end effector closer to the movement target, but leaving achievements of a proper orientation essentially unaddressed (Iftime et al., 2005; Merad et al., 2019, 2020; Merad et al., 2016; D. Popovic and Popovic, 1998; M. Popovic and Popovic, 2001). However, grasping an object requires positioning and orienting the hand in space appropriately. Without wrist mobility, the hand's orientation capabilities are severely restricted, necessitating compensatory movements (Olsen et al., 2019) that add stress to the body and lead to overuse issues in other joints (Mell et al., 2005; Spiers et al., 2018). Recognizing the importance of wrist angles, which are at least as crucial as the hand DOFs in executing ADLs (Kanitz et al., 2018; Montagnani et al., 2015b), several studies have attempted to reconstruct wrist pronation/supination, adding information to the control system, such as non-natural shoulder or trunk movements deliberately executed by the user (Kaliki et al., 2013; Legrand et al., 2022).

Grounded on the progress in movement-based approaches and computer vision techniques, Mick et al., 2021 recently demonstrated the feasibility of training an ANN to reconstruct all distal joints (i.e., elbow flexion/extension, wrist pronation/supination, wrist lateral deviation, and wrist flexion/extension) of an intact-limbs participant from proximal joint angles (i.e., shoulder flexion/extension, shoulder abduction/adduction, and humeral rotation) and information about the movement goal (i.e., the position and orientation of the target; see Figure 1.3). Ten intact-limbs participants were engaged in pick-and-place tasks in a VR environment, controlling the proximal joints of the virtual arm, while ANN predictions were employed to control the distal joints. The results of this experiment highlighted the effectiveness of the proposed control, with average movement times (1.22s) only slightly longer than those observed in naturally controlled movements (0.82s), along with comparable compensatory movements.

However, it is essential to acknowledge that the proposed control method has limitations that prevent its applicability to individuals with transhumeral limb loss. These limitations include the necessity for user-specific training data, which might be impracticable or time-consuming to collect, and the inability to measure the humeral rotation accurately using only a sensor placed on the upper arm due to soft tissue sliding.

To overcome these limitations, E. Segas et al., 2023 successfully adapted the control method to individuals with limb loss by designing a generic model, Proximo Contextual (PC), trained on natural arm movements gathered from multiple individuals (see Figure 1.4). Recognizing the challenges associated with reliable measurement of humeral rotation, this DOF was transferred as an output of the ANN, treating it as a predicted output rather than a necessary input for the control system. To enhance the precision and performance of the ANN, they increased the amount of training data by using the

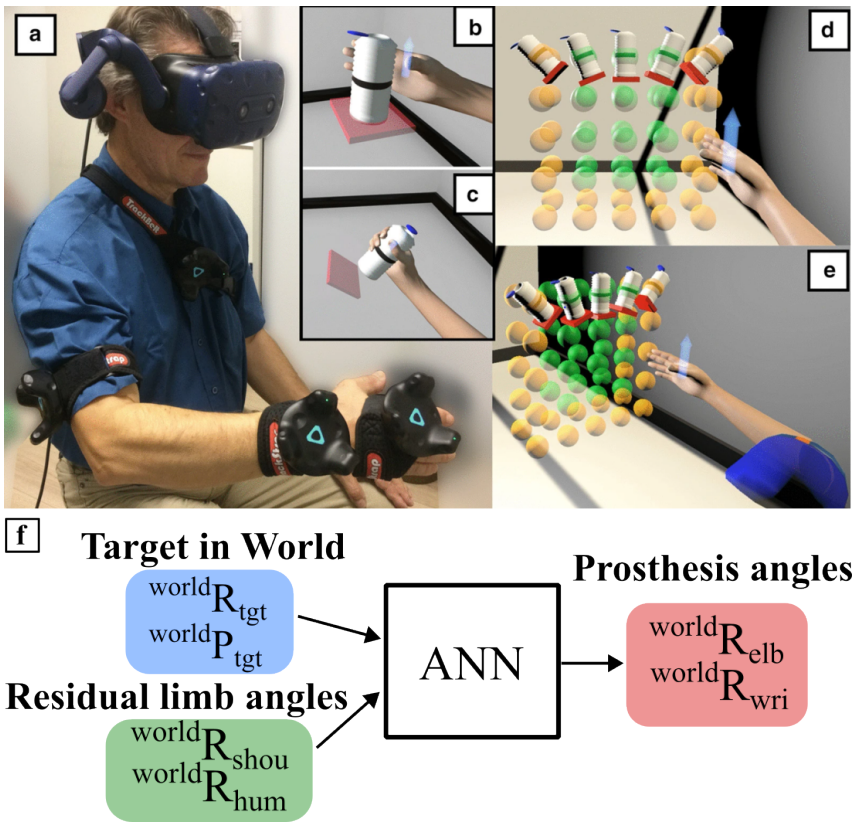



Figure 1.3: Overview of task and control strategy proposed by Mick et al., 2021. **a.** The participant's arm is reconstructed in the virtual environment using four HTC Vive sensors. **b-c.** The task is to control the virtual arm to pick a cylindrical object and place it on a platform. **d-e.** The targets are distributed over a grid of 5 X 5 X 2 (Width X Height X Depth) for the Acquisition phase (green and yellow spheres) and 3 X 4 X 2 for the Test phases (green spheres only). **f.** The ANN was trained to reconstruct the four distal joints from shoulder joint angles and information about the movement goal (i.e., the position and orientation of the target). ${}^{B}R_{A}$ and ${}^{B}P_{A}$ indicate the rotation and position of frame A expressed in the reference frame B (Figure adapted from Mick et al., 2021).

entire trajectories of the recorded natural arm movements instead of using only arm postures with the hand sufficiently close to the target. Furthermore, they expanded the applicable workspace of the control. To this end, they started by defining the widest possible workspace, determined by the maximal range of motion exhibited by participants, and then utilizing a self-organizing map (Fritzke, 1995) to represent the space covered by participants while generating natural arm movements within this space.

Critically, the training data was expressed in a body reference frame called armroot, which has its origin at the shoulder articulation and has the orientation of a sensor attached to the trunk of the operator. Using a common reference frame linked to the user's body allowed the data to be adapted to the user's morphology before training the subject-specific ANN. This adaptation was achieved by reconstructing the informa-

tion related to the movement goal (i.e., hand position and orientation) through forward kinematics using a chain based on the limb dimensions of the participant.

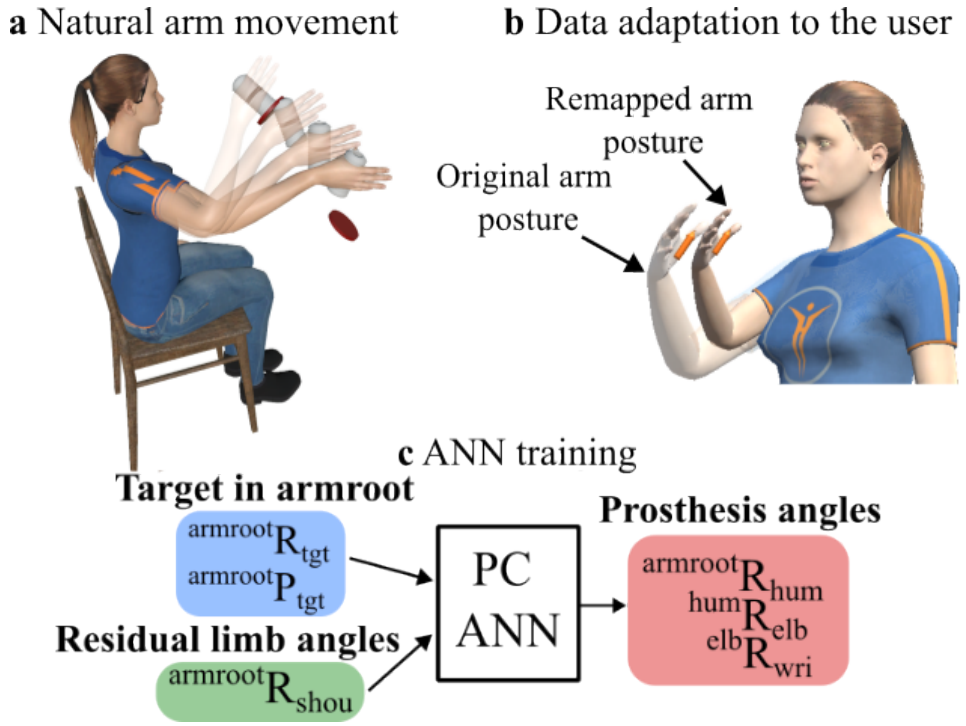



Figure 1.4: Overview of the control strategy proposed by E. Segas et al., 2023. **a.** Natural arm movements are recorded while participants pick and place a bottle in a wide workspace in VR. **b.** The data is adapted to the user's morphology by reconstructing the information related to the movement goal through forward kinematics using a chain based on the limb dimensions of the participant. **c.** PC ANN is trained to reconstruct distal joints (i.e., humeral rotation, elbow flexion/extension, wrist pronation/supination, wrist lateral deviation, and wrist flexion/extension) from proximal joint angles (i.e., shoulder flexion/extension and abduction/adduction) and information about the movement goal (i.e., the position and orientation of the target).

Overall, all those features enabled intact-limbs participants and participants with transhumeral limb loss to control a prosthesis avatar in VR and reach almost all targets (>99%) with the proposed prosthesis control with movement times similar to those observed in naturally controlled movements.

### 1.2.2 Adapting the control for real-world deployment

Despite the remarkable results achieved in VR, where the object position and orientation are perfectly known by design, deploying this control system in real-world settings requires the integration of a camera and computer vision algorithms into the system for real-time object pose detection and estimation. In this setup, information regarding the detected object might only be available in a head (or camera) centered reference

frame through gaze-guided computer vision (purple block in Figure 1.5). However, PC relies on the target position and orientation expressed in the armroot reference frame (blue block). Therefore, this information needs to be transformed from the head to the armroot reference frame.

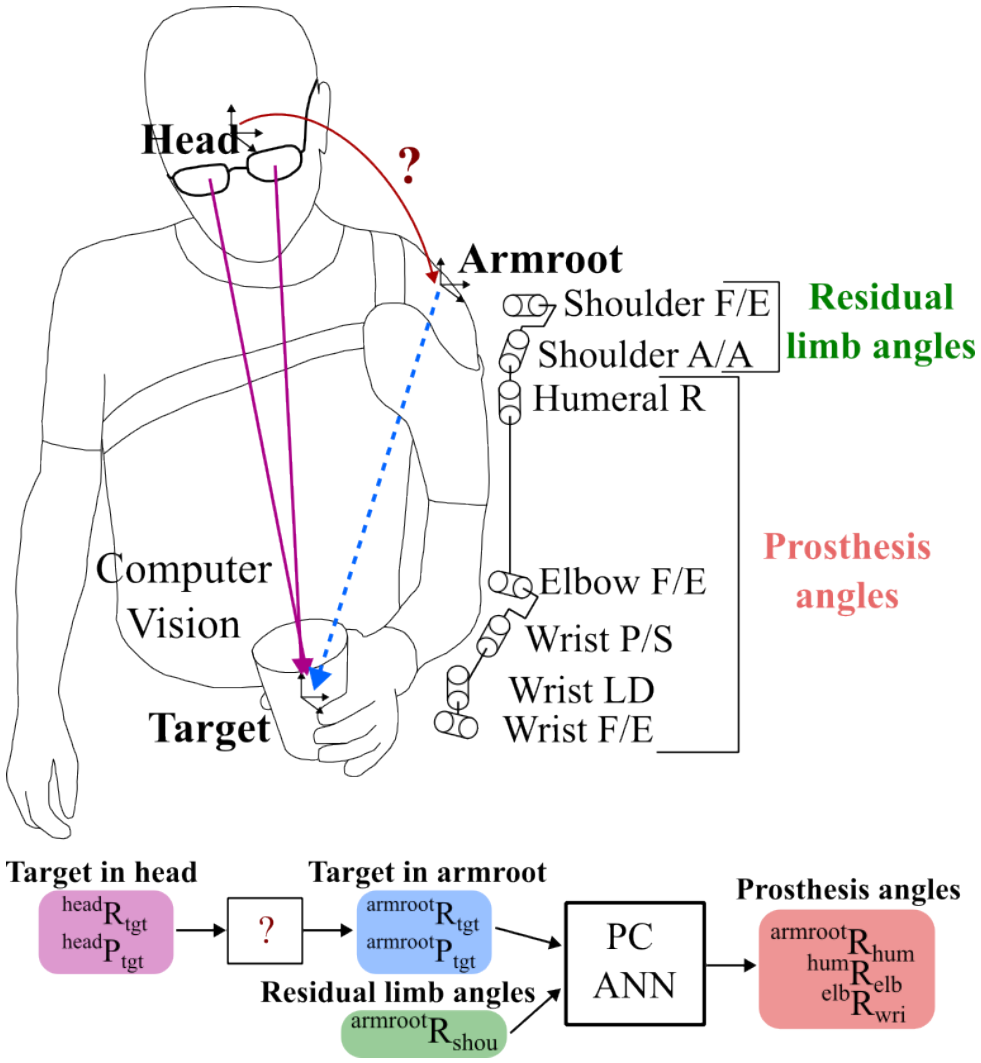



Figure 1.5: Problem definition. Computer vision algorithms retrieve information about the target in the head reference frame (purple block). This information requires transformation into the armroot reference frame (blue block) to be employed, along with the residual limb angles (green block), by the PC ANN to predict the angles to apply to the prosthesis (red block).

From a computational perspective, reference frame transformations require the knowledge of the position and orientation of both initial and final frames. While orientation data can be obtained from Inertial Measurement Units (IMUs) placed on the trunk and head, positional data might not be available in real-world settings.

Interestingly, ongoing theories (Batista et al., 1999; Blohm and Crawford, 2009; Pouget

et al., 2002; Salinas and Abbott, 2001; Snyder et al., 1998) suggest that reference frame transformations are computational mechanisms performed by the brain, especially when planning target-directed movements. Illustrated schematically in Figure 1.6, the brain acts as a sophisticated information processor, converting sensory inputs into motor commands to activate the muscles and coordinate the movement of different limb joints to achieve specific tasks (Holmes and Spence, 2004). Moreover, as the movement starts, the brain receives real-time sensory feedback, allowing for adjustments and corrections. This closed-loop system ensures the adaptability and accuracy of motor responses, especially in dynamic environments.

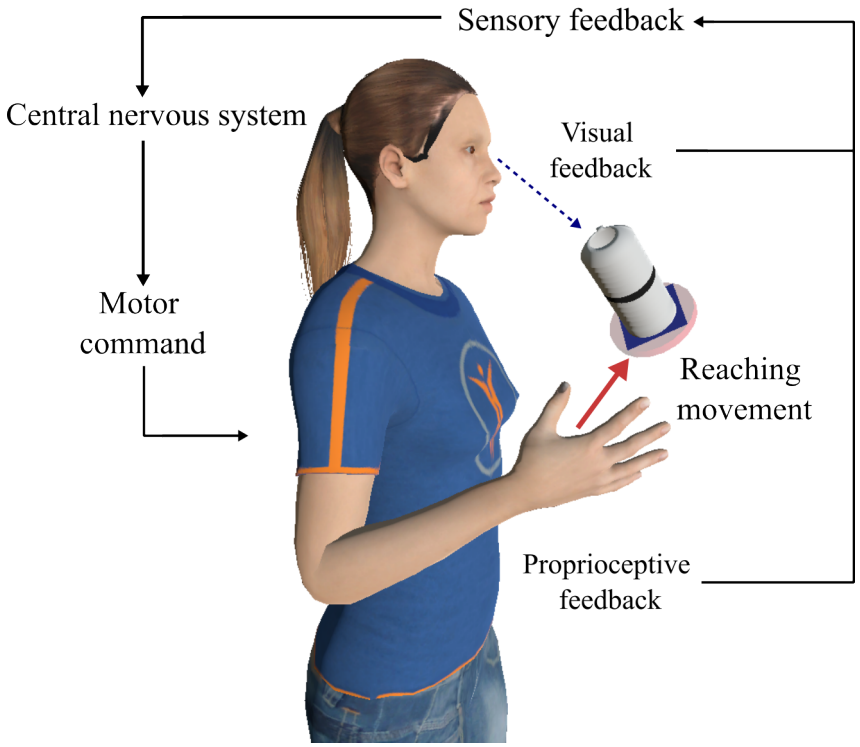



Figure 1.6: Schematic representation of how goal-directed movements are planned in the central nervous system. The brain converts sensory inputs into motor commands to coordinate the movement of different limb joints to achieve specific tasks.

Key brain regions involved in coordinate transformations include the lateral intraparietal area (LIP), primarily encoding data relevant to saccadic eye movements (Andersen et al., 1990), and the ventral intraparietal area (VIP), responsible for processing both visual and somatosensory information (Duhamel et al., 1998; Sereno and Huang, 2006). Within these areas, encoding appears to occur in distinct reference frames: eye-centered in LIP and head-centered in VIP (Duhamel et al., 1998; Mullette-Gillman et al., 2005). The neuroscience community widely acknowledges that these areas achieve coordinate transformations through mechanisms like gain fields and basis function representations, which enable stimuli to be represented simultaneously in multiple reference frames (Duhamel et al., 1998; Pouget et al., 2002; Snyder et al., 1998). Consequently, these areas are often modeled using neuron populations exhibiting gain fields

(Makin et al., 2013; Pouget et al., 2002; Pugach et al., 2019; Schillaci et al., 2014). The firing rate of these neurons depends on the position of the considered object relative to specific reference frames, such as arm-, head-, hand-, or eye-centered. Typically, these models comprise two input layers, each corresponding to a sensor modality and centered on their respective natural reference frames, containing units with bell-shaped activation functions.

Figure 1.7 illustrates the model proposed by Pouget et al., 2002 to emulate coordinate transformations from eye-centered to head-centered reference frame. The first input layer encodes the target in eye-centered coordinates, while the second encodes the eyes' position within the head reference frame. The intermediate layer combines the activity of the two receptive fields: each basis function unit computes the product of a pair of eye-centered and eye-position units. The output, the target in the head reference frame, is computed as the weighted sum of the activities of the basis function units. The network presented in this figure allows reference frame transformation in one direction (i.e., from the eyes to the head) but has been subsequently adapted to perform direct and inverse transformations.

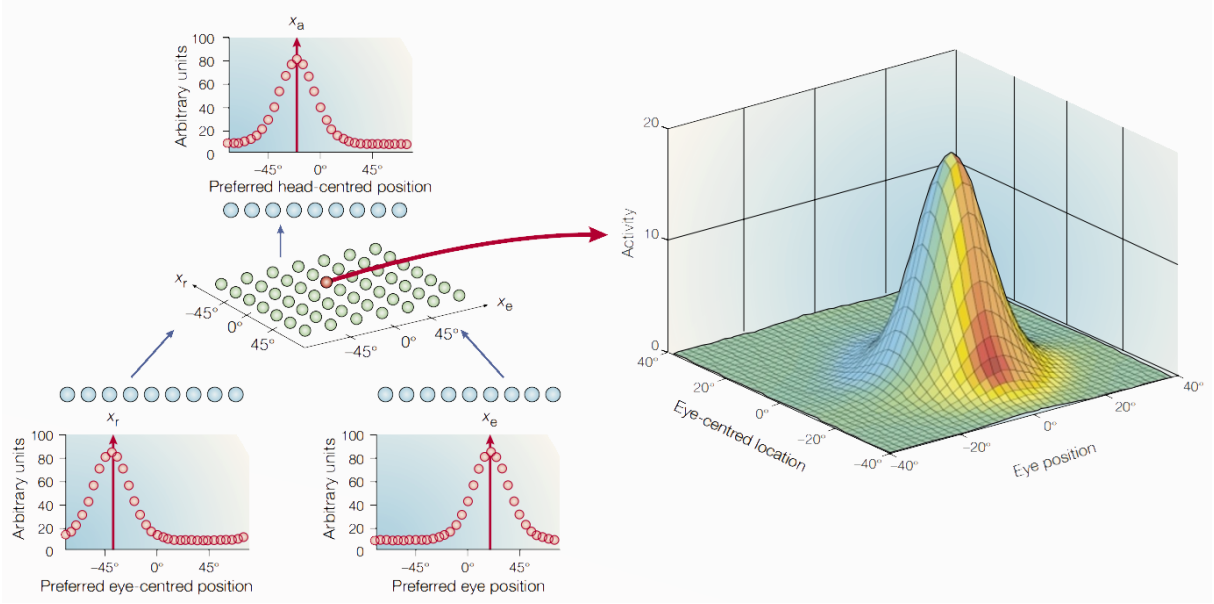



Figure 1.7: Neural network proposed by Pouget et al., 2002 for coordinate transformations using basis functions. The network enables the transformation of target information from eye-centered to head-centered coordinates.

In the realm of robotics, coordinate transformations are crucial for converting desired end-effector poses, specified in Cartesian or visual space, into motor commands issued in joint space, and vice versa. These transformations are often achieved through kinematic models, derived analytically and relying on prior knowledge about the robot's geometry. However, challenges arise when these models need to adapt to changes such as deformations of robotic parts or the integration of new tools. To address this problem, bio-inspired methods have been integrated into the control system, allowing

robots to autonomously learn the parameters of their kinematic model (Hoffmann et al., 2010; Nabeshima et al., 2006). Inspired by the brain's multisensory integration processes, these methods derive the pose of a body part, such as the end effector, from multiple sources like camera detection and forward kinematics. These poses are converted to a shared reference frame and compared. By using this difference as an error signal to iteratively adjust the parameters of their model, robots can progressively refine their understanding of their kinematics and adapt to changes in their configuration or environment. Commonly employed techniques include machine learning (Martinetz et al., 1990), self-calibration of the kinematic chain (D. J. Bennett et al., 1991; Gatla et al., 2007; Hollerbach and Wampler, 1996), or a hand/eye setup (Pachtrachai et al., 2021; Tsai and Lenz, 1988; Wang et al., 2021). All these approaches rely on a sampling period comprising various configurations, followed by an optimization procedure. However, rather than solely depending on predefined sampling periods, there is a growing interest in developing methods capable of learning online and adapting in cases of dynamic conditions (Hoffmann et al., 2010).

In particular, several approaches draw inspiration from the brain representation of our surroundings, the peripersonal space. This space is encoded in a multisensory, body-centered, and modular fashion (Farnè et al., 2005; Serino et al., 2015) and it continuously adapts in response to body movements, changes in the surroundings (Brozzoli et al., 2012; Di Pellegrino and Làdavas, 2015; Noel et al., 2015), and emotional states (Cléry et al., 2015). Furthermore, it can extend as a result of tool use (Farnè et al., 2005; Holmes and Spence, 2004; Maravita and Iriki, 2004; Maravita et al., 2002) or shrink due to traumatic amputation (Canzoneri et al., 2013). Building upon these characteristics, Nabeshima et al., 2006 employed an ANN to adapt the kinematic model of a humanoid robot in response to changes resulting from the integration of a new tool. The system learned to integrate visual and proprioceptive (joint angles) information through tactile interactions. Similarly, Hersch et al., 2008 developed a gradient descent algorithm enabling a 24-DOF humanoid robot (Hoap3) to learn the missing parameters of its kinematic chain (joint positions and orientations) by observing its body with a camera, and to adapt to changes, such as incorporating a stick as a body extension, within two to three minutes. Successively, the system was enhanced by employing optical flow to learn the neck-eyes kinematic chain and tested it on the iCub humanoid robot (Hersch, 2008). Sturm et al., 2009 took a further step by employing a flexible model based on Bayesian networks to enable a robotic manipulator arm to autonomously learn its kinematic structure and the geometrical relationships between its body parts through self-observation with a camera. Experiments conducted on both real and simulated robotic manipulators validated the effectiveness of this approach in addressing real-world challenges, including changes in the robot's body due to factors such as failure, repair, or material fatigue.

Within this context, a potential approach to obtain the head position within the armroot reference frame needed for transforming target information from the head to the armroot frame might involve constructing a kinematic chain connecting the two reference frames. Creating a model that accurately depicts the intricate relationship between the head and the shoulder positions requires a deep understanding of the complex

anatomy of these structures, shown in Figure 1.8.

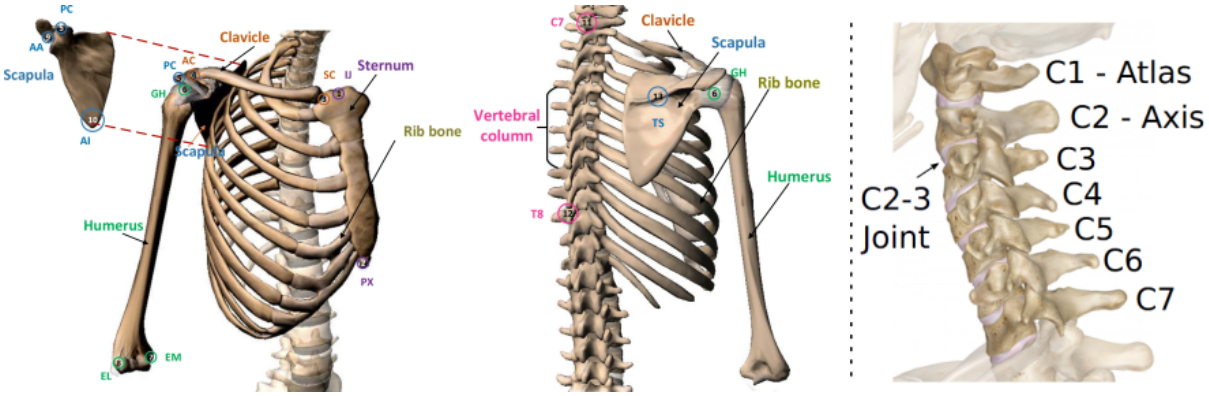



Figure 1.8: Anatomy of shoulder girdle and cervical spine (Figure adapted from Krishnan et al., 2019 and Barker et al., 2015).

The cervical spine, whose motion defines the head pose, has a complex musculoskeletal structure consisting of over twenty muscles and seven vertebrae, the first of which is the atlas (Standring, 2015). The vertebrae can be modeled as flexible springs, allowing flexion/extension and lateral bending motion, whereas the atlas bone enables head rotation and upper flexion/extension movements (translations). The range of movement between individual vertebrae may be small; however, the overall mobility of the cervical spine is extensive, as it depends on the cumulative effect of numerous small movements occurring across the seven joints. On the other hand, the shoulder girdle includes three synovial joints, the sternoclavicular, the acromioclavicular, and the glenohumeral joint, allowing translation in all directions (Standring, 2015). The motion of these joints depends on the scapulothoracic joint, a fictitious articulation responsible for approximately one-third of the shoulder's range of motion (Krishnan et al., 2019). Researchers often model this joint as either a fixed structure (Sah and Wang, 2009) or dynamic contact (Garner and Pandy, 1999; Maurel and Thalmann, 1999; Seth et al., 2016).

Capturing and modeling the movements of the head and the shoulder pose several challenges due to a combination of anatomical intricacies, functional complexities, mathematical representations, and limitations in measurement techniques (Krishnan et al., 2019; Rau et al., 2000). This results in the implementation of simplified kinematic models, which may lack a comprehensive representation of the diverse range of movements observed in reality. For instance, the most common representation of the kinematic chain between the head and the shoulder features 3 DOFs at the neck level, 3 DOFs at the trunk, and 3 DOFs at the shoulder, with only prismatic joints, enabling rotation but not translation. Furthermore, when attempting to implement such a model, for example, in a humanoid robot, additional simplifications are often necessary, such as employing a 2-DOF serial mechanism for the neck. Although this mechanism offers some mobility for the robot's head, its capabilities are usually limited, allowing only a partial scanning of the surrounding environment (Beira et al., 2006; Muñoz et al., 2024; Penčić et al., 2017).

The intricacy of the region between the head and the shoulder makes the interaction between these body parts inherently dynamic, characterized by continuous adaptations to accommodate various tasks and postures. The relative position between the two frames varies depending on the location of the object to be reached relative to the user's body, as depicted in Figure 1.9.

Figure 1.9: Relationship between the head and the armroot reference frame while grasping two different bottle poses.

Consider, for instance, the act of grasping an object situated above the user's head. In such instances, the individual instinctively raises their arm and shoulder girdle while adjusting their head position to secure a clear view of the object. Conversely, when the object lies below the level of the hip, the shoulder girdle lowers to facilitate the reaching. In both scenarios, the relative distance between the head and shoulder undergoes significant alterations, reflecting the human body's adaptability in response to varying tasks. Furthermore, the posture of the trunk and arms emerges as a pivotal factor influencing the positioning of the shoulder joint and overall upper body dynamics.

Similar situations arise in robotics, where complex systems integrate sensor modalities and flexible parts for which a kinematic model alone may not suffice to execute all the required reference frame transformations. Instead of implementing a system to learn these intricate models, such mappings can be acquired by treating the system

as a black box and learning its input-output behavior through a system identification process (De Barreto, 2003; Martinetz et al., 1990). For example, Chinellato et al., 2011, inspired by the encoding of the peripersonal space (Di Pellegrino and Làdavas, 2015; Serino, 2019) and sensorimotor transformations (Makin et al., 2013; Pouget et al., 2002), implemented a sensorimotor map with two radial basis function networks to perform bidirectional transformations in a humanoid robot between stereo visual information and oculomotor space, and between oculomotor and arm joint space, respectively. Through active exploration by gazing and reaching actions, the networks learned to execute proficiently direct and inverse transformations among different space representations and to adapt to varying conditions, including visual distortions or modified kinematic configurations. Likewise, Schillaci et al., 2014 proposed a biologically inspired model for coding internal representations of sensory maps (visual, auditory, and tactile) in the humanoid robot Aldebaran Nao. Inspired by the self-organizing properties of areas in the human brain, the authors employed a dynamic self-organizing map for online and continuous learning of static and dynamic data distributions.

In the same line of thought, Lallee and Dominey, 2013 introduced a novel approach to controlling a humanoid robot using an artificial neural network model. Unlike previously cited works, this model emulates the brain's multisensory integration processes and provides explanations and predictions for findings in neurophysiology and neuropsychology. By engaging in self-exploration, the model acquired a representation of the robot's body schema, integrating specific modalities such as arm proprioception, gaze proprioception, and vision, along with their multimodal interactions. Experimental results showcased that this method significantly enhances the robot's adaptability and functional performance, resulting in more efficient control.

In conclusion, a bio-inspired approach appears to be the most suitable for describing the head position in the armroot reference frame. Before embarking on the development of solutions to retrieve this information, it is crucial to gather data concerning the relation between these two reference frames. The subsequent chapter will delve into the experimental protocol and the methodological approach for collecting a dataset.

# Chapter 2

# Dataset of natural arm and head movements

Every object manipulation starts with a natural arm movement designed to reach and grasp it. Most of the time, this object is recognized based on visual information gathered during gaze-fixation. Understanding eye-hand coordination during reaching movements is a great challenge in human sensorimotor control, with huge applications in motor rehabilitation, humanoid robotics, human-robot interactions, and prosthesis control.

Regarding prosthesis control, we recently demonstrated that incorporating information about movement goals alongside proximal joint motion allows an ANN trained on natural arm movements to predict distal joints missing to people with transhumeral limb loss (Mick et al., 2021; E. Segas et al., 2023). Despite promising outcomes in a VR environment, implementing this control in real-world scenarios necessitates the integration of computer vision algorithms to retrieve movement goal information, which will be expressed in the head reference frame. Subsequently, this information must be transformed into the armroot reference frame to be effectively utilized by the control system. In this chapter, I introduce a dataset designed for collecting data about these two reference frames (head and armroot). This dataset will be used in Chapter 3 and 4 to develop and evaluate algorithms aimed at performing the necessary transformations.

Although several public datasets exist that include head and arm movements suitable for modeling the relationship between the head and the armroot reference frame (Averta et al., 2021; Huang et al., 2016; Mandery et al., 2015), the overall body postures from which arm movements were produced in those datasets were generally left free or poorly controlled. Moreover, none of those datasets were designed to maximize the workspace spanned by participants with their arms (i.e., the set of positions and orientations reachable depending on their morphologies and ranges of motion).

Furthermore, although 6D pose estimation is a very active area of computer vision (Labbé et al., 2020; Labbé et al., 2022; Liu, 2023; Nguyen et al., 2022) with applications far beyond prosthesis control (e.g., augmented reality, healthcare or industrial

robotics), precise determination of object pose from first person view/egocentric videos with gaze information acquired in natural context is still an open research issue (Rogez et al., 2015). Several public datasets are available for training and testing object 6D pose estimation models (Calli et al., 2015; C. Li et al., 2022), but none of those include realistic, egocentric visual information like gaze obtained in a functionally relevant context (i.e., with natural arm movements that require reaching and displacing objects).

The dataset outlined in this chapter was designed to overcome these aforementioned limitations. First, the initial body posture, from which arm movements were produced, was precisely controlled using visual feedback from the trunk and shoulders in a VR environment. Second, the workspace was maximized throughout a procedure that initially covers the widest possible workspace (i.e., set by the maximal range of motion of participants) and then subsequently uses a self-organizing map to represent the actual space covered by participants (i.e., producing natural arm movements within this space; E. Segas et al., 2023). Third, the dataset captured head and gaze motion alongside movements of the trunk, shoulders, and arm joints, providing comprehensive data on the entire kinematic chain between the object the participant is aiming and gazing at and the hand moved to reach it.

The work presented in this chapter was conducted with the assistance of a Master 1 student, Emeline Medan, who completed a two-month internship under my supervision.

The work presented in this chapter has been published (Lento, Segas, Leconte, Doat, Danion, Péteri, Benois-Pineau, and De Rugy, 2024), Scientific Data - Nature. The dataset and the codes for processing and analyzing the data are available as a Zenodo repository (Lento, Segas, Leconte, Doat, Danion, Péteri, Benois-Pineau, and de Rugy, 2024).

## 2.1 Materials and Methods

This section presents all the elements put in place to carry out the experiment presented in this chapter.

### 2.1.1 Participants

This study was conducted on twenty participants (six males and fourteen females) aged 19-44 (mean 25.1; SD 6.46), all with normal vision or vision corrected to normal. The Edinburgh Handedness Inventory (Oldfield, 1971; EHI) was administered to determine the participants handedness: they were all right-handed (mean EHI 81; SD 23). None of the participants suffered from any mental or motor disorders that might have affected their task performance. All participants provided informed consent, and the study received approval from the local ethics committee (CPP Est II: n°2019-A02890-57).

### 2.1.2 Experimental setup

The participant sat on a stool and wore the headset (ViveTM Pro, HTC Corporation) and five motion trackers (ViveTM Tracker, HTC Corporation), attached to the trunk, the main segments (upper arm, forearm, and hand) of the right arm and the upper arm of the left arm, using straps, as shown in Figure 2.1**a**.

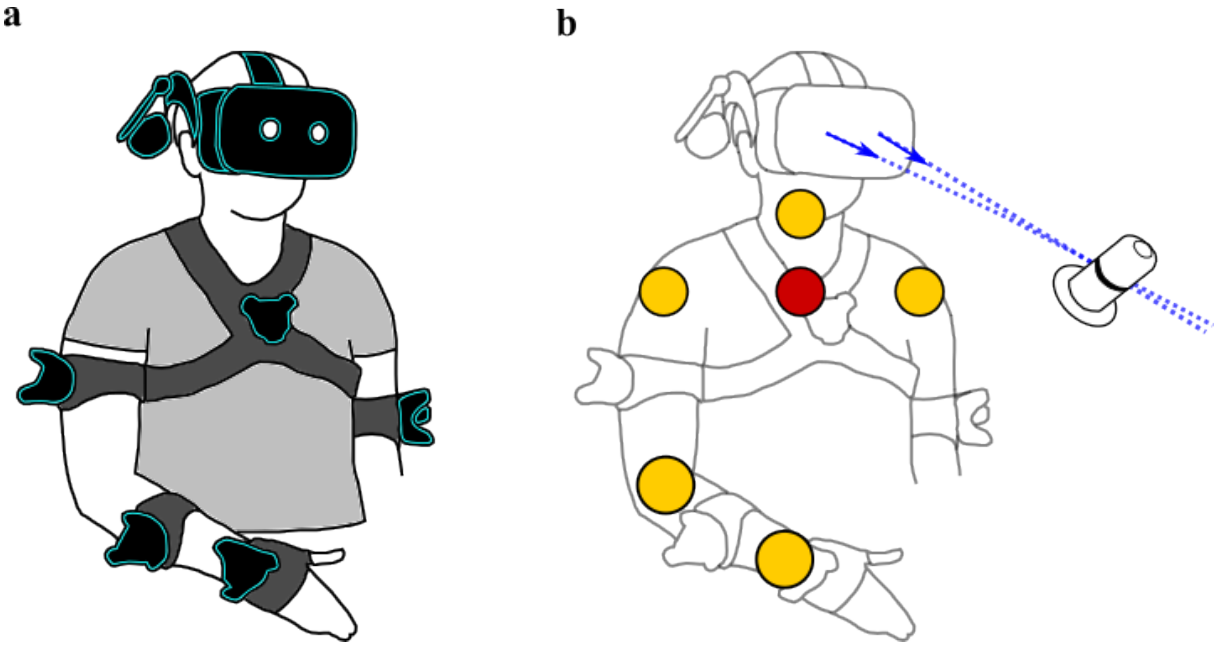



Figure 2.1: Experimental setup in VR. **a.** Participant wearing VR headset and trackers. **b.** Estimated joint centers' locations are represented by yellow and red spheres, the pupil axis by blue dashed lines, and the direction of the gaze vector of each eye by a blue arrow.

Each tracker and the headset provided measurements for their 3D position and orientation in relation to the reference frame of the virtual environment. Four cameras were positioned at the room corners to capture signals from the trackers and headset at a rate of 90 Hz (sampling rate), using SteamVR, developed by Valve Corporation, as middleware. The virtual scene's contents and interactions with the participant were managed by the Unity simulation engine (Unity Technologies). The virtual environment was scaled to match real-world dimensions: the ground plane was set at the same height as the actual floor and centered on the stool where the participant sat.

### 2.1.3 Virtual arm calibration

After the participant was equipped with the trackers and the headset, the software for eye tracking was calibrated throughout the method already included in SteamVR. Then, a calibration procedure was carried out to associate the virtual arm with the trackers and, therefore, with the participant's real arm. The virtual arm, created using the open-source software Make Human, consisted of three rigid segments (upper arm, forearm, hand) connected by spherical joints at shoulder, elbow, and wrist levels. Its thickness was adjusted according to the participant's height by applying a factor depending on its value.

The virtual arm calibration consisted of three steps: data recording, joint center estimation, and association of virtual and real arms. During the first step, the participant was instructed to perform slow movements using all the DOFs of both arms for 15s and then of the head for 10s. Based on these data, the method described in O'Brien et al., 2000 was employed to estimate the following joint centers' locations: left shoulder, neck, right shoulder, elbow, and wrist, shown in Figure 2.1**b** as yellow spheres. Additionally, the center of the trunk, represented by a red sphere in Figure 2.1**b**, was estimated as the orthogonal projection of the neck joint center onto the line connecting the right and left shoulders.

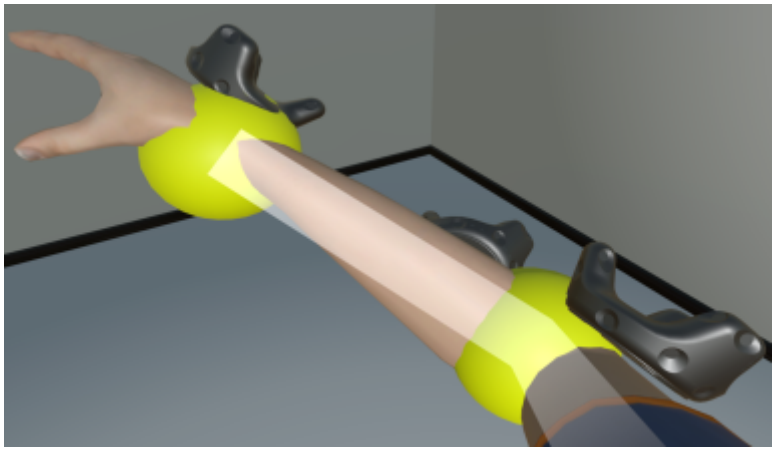
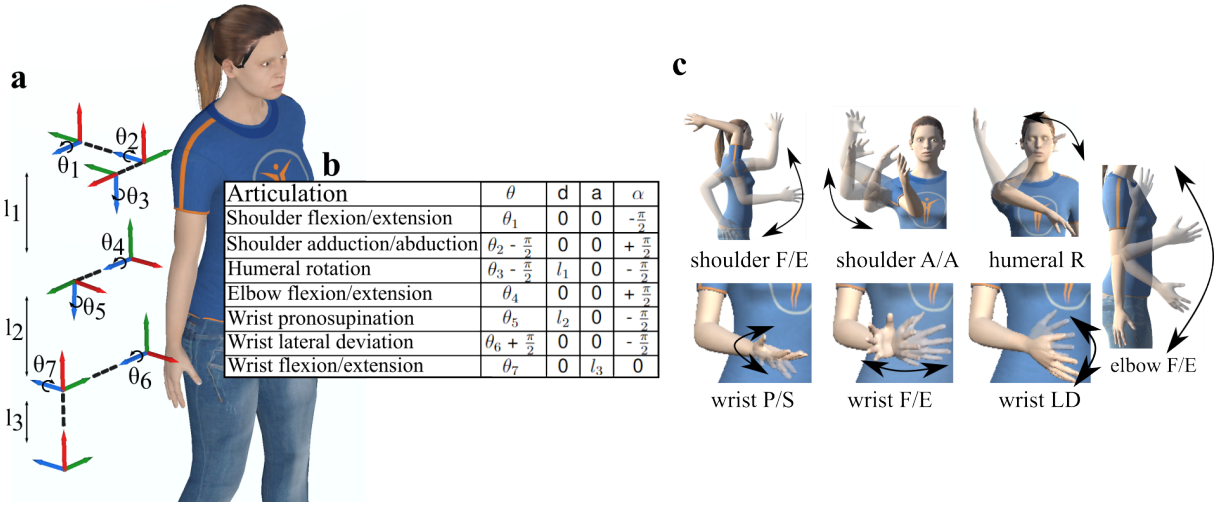


Figure 2.2: Final step of the virtual arm calibration. The participants overlay their real arm, represented in the virtual environment as yellow spheres connected by white segments, with the virtual arm (Figure II.1.4 in Ségas, 2023).

The joint centers were linked in position to the sensors. The shoulders, left and right, and elbow were associated with the corresponding upper arm sensors, while the wrist joint was associated with the forearm sensor. The shoulder of the virtual limb (right side) was positioned to coincide with the actual shoulder of the participant, and the virtual arm dimensions resized according to the participant ones, computed as the distance between estimated joint centers. In the last step, the virtual arm segments were linked in orientation to their respective sensor. The participants saw their real arm in the virtual environment as yellow spheres connected by white segments and had to overlay it with the virtual arm locked in a reference posture in order to bind them, as shown in Figure 2.2.

### 2.1.4 Range of motion determination

Following the calibration process, the headset was temporarily taken off to collect data with the purpose of calculating the range of motion for each DOF of the participant's arm (shoulder flexion/extension, shoulder abduction/adduction, humeral rotation, elbow flexion/extension, wrist pronation/supination, wrist lateral deviation, and wrist flexion/extension, see Figure 2.3).




| Articulation | $\theta$ | d | a | $\alpha$ |
|---|---|---|---|---|
| Shoulder flexion/extension | $\theta_1$ | 0 | 0 | $-\frac{\pi}{2}$ |
| Shoulder adduction/abduction | $\theta_2 - \frac{\pi}{2}$ | 0 | 0 | $+\frac{\pi}{2}$ |
| Humeral rotation | $\theta_3 - \frac{\pi}{2}$ | $l_1$ | 0 | $-\frac{\pi}{2}$ |
| Elbow flexion/extension | $\theta_4$ | 0 | 0 | $+\frac{\pi}{2}$ |
| Wrist pronosupination | $\theta_5$ | $l_2$ | 0 | $-\frac{\pi}{2}$ |
| Wrist lateral deviation | $\theta_6 + \frac{\pi}{2}$ | 0 | 0 | $-\frac{\pi}{2}$ |
| Wrist flexion/extension | $\theta_7$ | 0 | $l_3$ | 0 |

Figure 2.3: 7-DOFs arm model. **a.** Relative position of frames of reference using Denavit-Hartenberg convention (Denavit and Hartenberg, 1955). **b.** Denavit-Hartenberg parameters table. **c.** Elementary movements performed for each arm DOF during the range of motion assessment phase (Figure adapted from Ségas, 2023; reference frames assigned using the distal DH method).

Participants were asked to slowly perform a few repetitions of an elementary movement for each joint with maximal amplitude in each direction. For each movement, the experimenter performed a demonstration that the participants were required to mimic with their arm. This demonstration and movement recording lasted about 10s for each movement, and included at least four complete back-and-forth excursions. Then, the

joint angular limits were estimated as the extreme reached values (i.e., the minimum and maximum angles reached during each joint repetition).

### 2.1.5 Task

Data were collected while the participant was engaged in a pick-and-place task involving picking a bottle from one platform (red disk visible at the base of the bottle) and releasing it into another one.

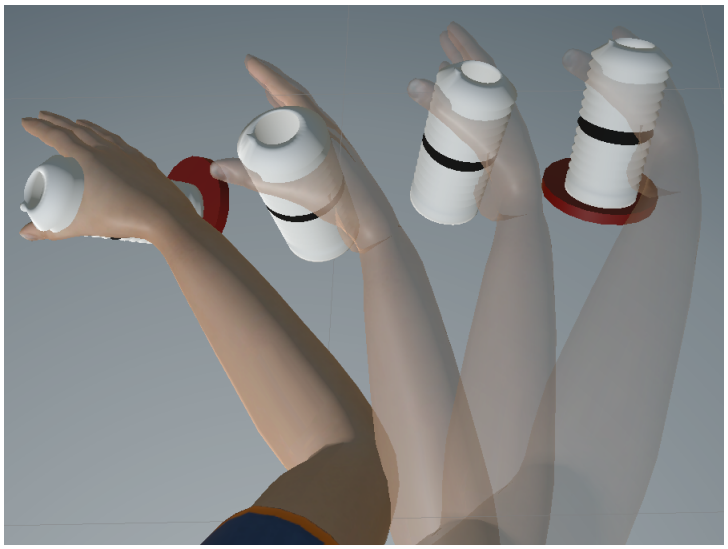


Figure 2.4: Illustration of the task. After the participants have picked a bottle from a previous platform, they will transport it toward the next platform, onto which they will release it.

The task was defined as the action of either picking up a bottle (target) from the platform or releasing it onto the next platform. The task did not involve directly the opening and closing of the hand. Instead, participants were required to place the virtual hand within a 2cm proximity of the bottle's center and match the virtual hand orientation with the bottle orientation within an angular tolerance of 5°. When these two criteria were met (i.e., virtual hand in the target zone), the bottle turned red. After 0.6s of continuous red coloring (validation time), the bottle was automatically grasped or released. A specific time was allocated for reaching each target (5s). If not achieved within that time, the task was considered failed, and a new target appeared, triggering a sound signal.

A shallow bottle was added on the target platform or in the virtual hand to help the participant fit the bottle positioning within the target zone.

Participants were instructed to perform the task at a natural pace. An illustrative video of an intact-limbs participant performing the task in the virtual environment is provided in Appendix B (Illustrative Video 1).

### 2.1.6 Return to Neutral Posture procedure

In most research focusing on a specific pattern of movement, it is a common procedure to have participants return to a neutral posture between trials (or movements) in order to maintain recording consistency and limit drifts or carry-over effects between trials (Light et al., 2002; Major et al., 2014; Montagnani et al., 2015a). Instead of merely relying on participants' compliance with instruction to return to a neutral posture, or upon the experimenter's observation of it, we capitalized on our VR set-up to design and apply a controlled procedure called Return to Neutral Posture (RNP). This was particularly suited for our purpose, which was to collect data about natural arm movements issued from a comfortable posture, including shoulder and trunk.

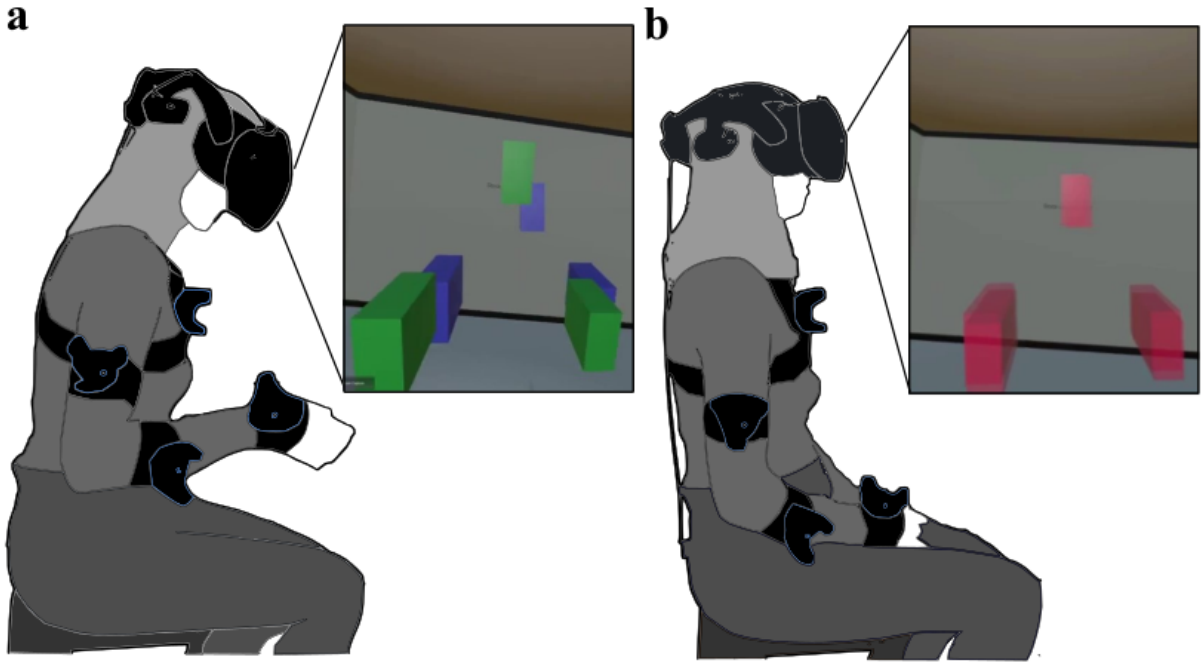



Figure 2.5: Return to Neutral Posture (RNP) procedure. **a.** After the participant completed a task, the RNP procedure started and the cubes appeared. The green cubes represented the neutral posture the participant had to return to, while the blue cubes represented the actual participant's shoulders and head positions. **b.** By overlapping the blue cubes with the green cubes within a given tolerance, all cubes turned red to indicate that the participant had successfully achieved the RNP.

This procedure was implemented using the display of two sets of three cubes (see Figure 2.5), each cube corresponding either to a shoulder position (left or right) or to the head position, projected 45cm forward in front of the participant. The green cubes were fixed and represented the target posture the participant had to return to, while the position of the blue cubes was updated according to the actual participant's shoulders and head positions (see Illustrative Video 1 in Appendix B). The RNP was achieved by overlapping the blue cubes with the green ones within a given tolerance for a duration of 0.6s. The specific tolerances used were determined empirically to be tight enough

to ensure consistency in the given posture participants had to return to, while avoiding control difficulties that would be associated with too small tolerance. This was achieved with an angular tolerance of 5° for the head and shoulders, and a spatial tolerance of 2cm for the head and 3cm for the shoulders.

### 2.1.7 Protocol

The experimental protocol comprehended five phases: Familiarization, Initial acquisition, RNP familiarization, Test RNP after pauses, and Test RNP after target pairs.

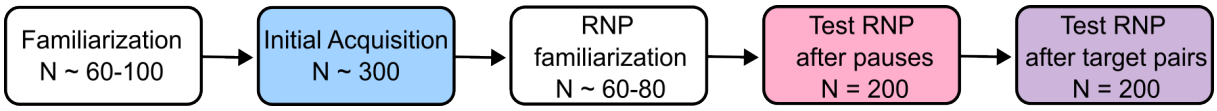



Figure 2.6: Dataset collection protocol.

Different target sets were generated, as explained below, for the different phases of the experiment. A target is defined as a position and orientation expressed in the initial shoulder reference frame to be reached by the participant's hand. All the targets were filtered according to the following criteria:

- Exclude targets pointing downward, defined by an angle between the target's axis and the vertical axis that would exceed 100°.
- Exclude targets too close to the participant's trunk, defined by a distance between the target's center and the participant's frontal plane that would not exceed a third of the participant's arm length.
- Exclude targets too close to the participant's legs, defined by a distance between the target's center and the horizontal plane passing through the participant's shoulder that would exceed two-thirds of the participant's arm length.

To generate the first target set, 7-DOFs arm angular configurations were drawn at random within the participant ranges of motion following a multivariate uniform probability distribution. Then, forward kinematics was used to transform these configurations into the target locations (i.e., positions and orientations), which were then filtered as indicated above (see E. Segas et al., 2023 for more details). This target set contained at least 300 targets, subsequently used in the Familiarization and Initial Acquisition phase.

The first phase aimed at familiarizing the participant with the virtual environment and the task. For this reason, the time to complete a task was extended to 10s. Once the participant seemed comfortable with performing the task, after approximately 60-100 targets, the experimenter could skip the remaining targets and proceed to the next phase.

The second phase, Initial Acquisition (IA), was designed to record data used to sample more accurately the reachable workspace, which is specific to each participant. This

was achieved using an unsupervised self-organizing map (more specifically, a Growing Neural Gas (GNG) algorithm; Fritzke, 1995), trained on all arm postures produced by each participant during the IA phase, to generate a new set of 200 targets that best represent those postures. This algorithm maps the space of the joint configurations executed by the participants during the IA phase by placing nodes (or neurons) and iteratively adjusting their positions based on the input data. By effectively learning the distribution of joint configurations, it identifies the neurons that best represent the arm configurations explored by the participant. These configurations are then transformed into targets through forward kinematics. This step was necessary to circumvent problems inherent to the fact that anatomical joints' limits are interdependent, which could lead to impossible arm configurations, in particular when maximal excursion at multiple joints is involved simultaneously. The resulting new set of 200 targets, which best represent the reachable workspace of each participant, was then used for all subsequent phases.

In the third phase, RNP familiarization, the participants freely chose their neutral posture to return to and got familiar with the RNP procedure, which was introduced every four targets. At any time, the participant could ask to reposition the cubes, defining that neutral posture, as the goal was to ensure the most natural and comfortable posture. As for the Familiarization phase, when the participant seemed comfortable with the neutral posture and the RNP procedure, the experimenter skipped the remaining targets and proceeded to the next phase (typically after 60-80 targets).

Two Test RNP phases were then conducted: one in which the RNP procedure occurred at the beginning of the phase and after each pause introduced every 50 targets (i.e., three pauses in a phase of 200 targets), and one in which the RNP procedure occurred after each single pick-and-place movement (2 targets). This last phase, Test RNP after target pairs (RNP ATP), was designed to prevent any change or drift in baseline posture that might occur from movement to movement. The preceding phase, Test RNP after pauses (RNP AP), aimed at testing whether a RNP procedure introduced only every 50 targets was sufficient to maintain a reliable neutral posture, as this would greatly simplify and accelerate future protocols using this pick-and-place task. Furthermore, to ensure that the arm movement between target pairs remained unaffected by the RNP procedure, the target set of the phase Test RNP ATP was adjusted by repeating the same target before and after the RNP procedure.

## 2.1.8 Data reduction and metrics

This subsection describes the filtering procedures applied to the data collected in the experiment, along with the metrics employed to assess and compare the performances.

### 2.1.8.1 Filtering

A filtering process was carried out to remove measurement errors associated with motion capture. As the sensors communicate via Bluetooth and the detection technique

involves infrared, some data may have been lost during recording. Two filters were employed: one for “freezing” behaviour, defined as a portion lasting at least 0.5s where a sensor’s measurement has not changed value, and one for “jumping” behaviour, triggered when a sensor’s measurement has jumped in position by at least 10cm between two samples. On average, this filtering process led to the exclusion of 2.6% of targets per participant and phase.

#### 2.1.8.2 Metrics

The collected data were analyzed and subjected to statistical treatment in terms of:

- success rate, defined as percentage of successful tasks achieved,
- movement time, defined as the time taken to reach and validate each target from the previous one,
- shoulder spread volume, defined as the volume of the ellipsoidal region containing 97% of the shoulder positions recorded while producing all movements of a phase, as a proxy of postural stability and body compensation within a phase (Mick et al., 2021).

Considering the high success rate obtained in all the phases (medians above 96%), the movement time and the shoulder spread volume were calculated on successful targets.

### 2.1.9 Statistical analysis

The statistical analysis was conducted using the R software with a significance level ($\alpha$) set at 0.05. A McNemar test was employed for evaluating success rate, whereas for movement time and shoulder spread volume, either a paired t-test or a Wilcoxon test was performed based on the normality test (Shapiro-Wilk) results.

## 2.2 Results

As illustrated in Figure 2.7, success rates were significantly lower and movement times significantly longer in the IA phase than in the two following RNP phases (IA vs RNP AP vs RNP ATP; n = 20; median success rates of 96.3% vs 98.2% vs 99.4%; Friedman test chi.sq = 24.08, $p < 0.001$; IA vs RNP AP, $p = 0.002$; IA vs RNP ATP, $p < 0.001$; median movement times of 1.61s vs 1.30s vs 1.22s; Friedman test chi.sq = 32.5, $p < 0.001$; IA vs RNP AP, $p < 0.001$; IA vs RNP ATP, $p < 0.001$). This is consistent with the view that some targets, drawn randomly within the range of motion determined individually at each joint of a participant, were difficult to reach, and that the set of targets established on the basis of movements actually produced in the IA phase, and used in the RNP phases, provides a better sample of the reachable space of each participant. The difficulty to reach some targets of the IA phase is further corroborated by four times difference in shoulder spread volume between this condition and the other

RNP phases (IA vs RNP AP vs RNP ATP; n = 20; median shoulder spread volume of 0.76 vs 0.17 vs 0.21; Friedman test chi.sq = 25.2, $p < 0.001$; IA vs RNP AP, $p = 0.001$; IA vs RNP ATP, $p < 0.001$).

When comparing the two RNP phases, none of the dependent variables exhibited a significant difference. Pick-and-place movements were performed with similarly high success rates (RNP AP vs RNP ATP; median success rates of 98.2% vs 99.4%; $p = 0.997$), reliable movement times (RNP AP vs RNP ATP; movement times of 1.30s vs 1.22s; $p = 0.366$) and led to similar shoulder spread volumes (RNP AP vs RNP ATP; median shoulder spread volume of 0.17 vs 0.21; $p = 1$), whether the return to a neutral posture procedure was introduced every 50 pick-and-place movements or between each pick-and-place movement. This lack of difference across RNP phases means that returning to a baseline posture only every 50 pick-and-place movements is sufficient to prevent postural changes or drifts, thereby allowing the release of postural constraints in future protocols using this task.

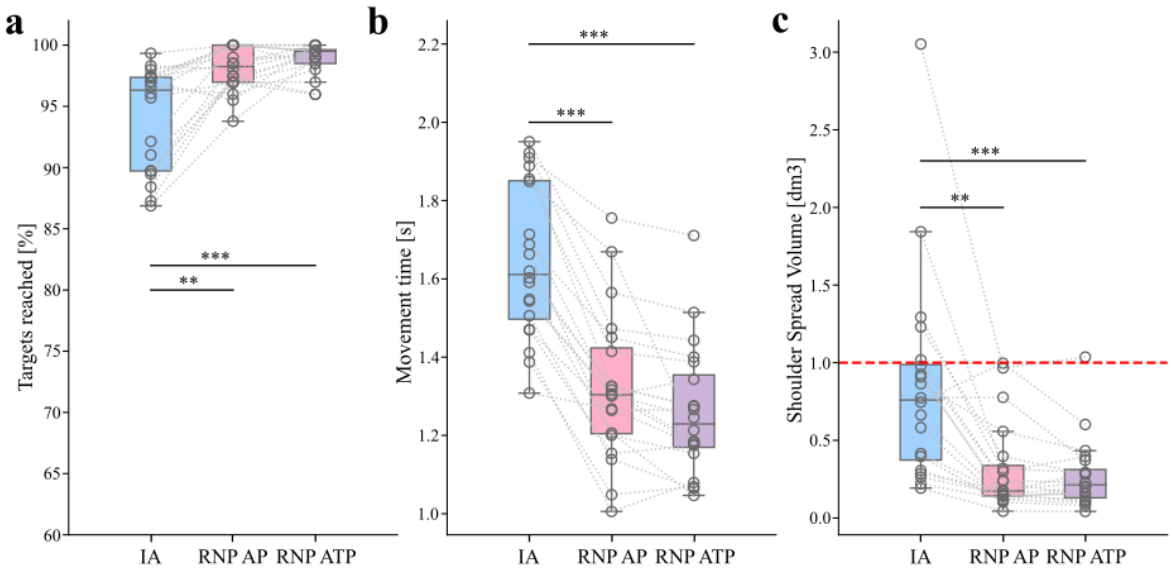



Figure 2.7: Dataset analysis results. Individual data are represented by hollow dots and dotted lines. Stars represent significant differences with ** for $p < 0.01$ and *** for $p < 0.001$. **a.** Success rate **b.** Movement time. **c.** Shoulder spread volume. The red line represents a volume of 1 $dm^3$ (=1 L).

### 2.2.1 Explored workspace

To provide an overview of the workspace effectively covered by our dataset, we computed from all arm configurations outputted by the GNG algorithm the corresponding targets (represented as red arrows in Figure 2.8) using forward kinematics with the limb dimensions of the avatar depicted in the figure. From the initial workspace, defined by the ranges of motion of the participants and represented by the targets selected within

those ranges of motion for the Initial Acquisition phase, the workspace defined by the GNG algorithm has shrunk by 13.74% (i.e., averaged reduction in the range of motion at each angular joint). Although this might appear to contradict our goal of maximizing the workspace covered by participants, experimental results confirmed that this was the price to pay to ensure high success rates with minimal postural adjustments, thereby keeping the focus on natural arm movements.

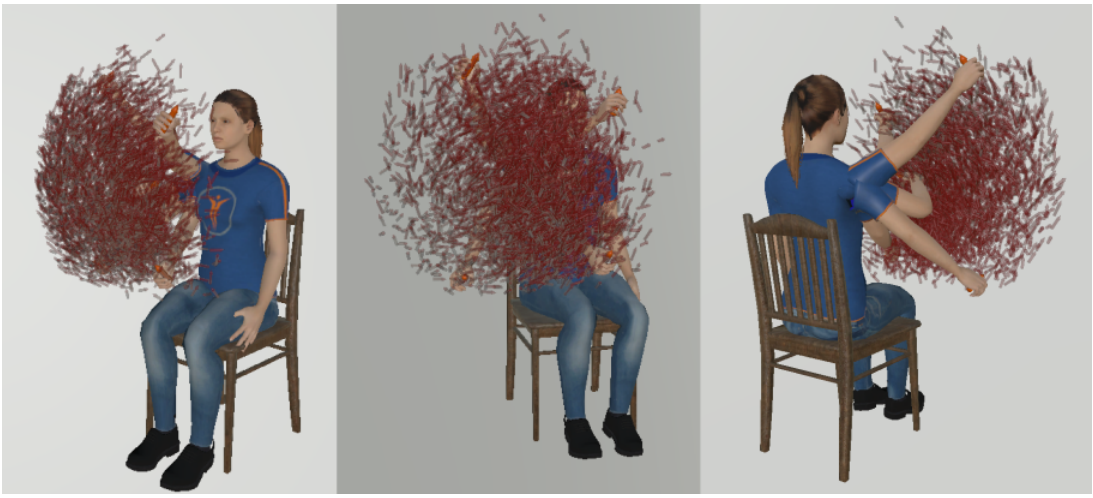


Figure 2.8: Workspace explored in our dataset. Red arrows indicate all targets generated by the GNG algorithm and used in subsequent experimental phases. Data from all participants are regrouped after proper scaling to the arm size of a single participant, and displayed with a few examples of arm configurations.

## 2.3 DataPlayer

*The following section outlines the implementation of a tool to replay the data collected in this experiment, led principally by my colleague Vincent Leconte.*

We developed a Unity project, named the DataPlayer, which enables the playback of recorded data from the experiment. By selecting a file from the dataset, the DataPlayer reproduces the arm movements executed by the participant during the pick-and-place task. Additionally, it displays gaze data, including the pupil axis and the point of gaze, computed using the method outlined in Bailakanavar et al., 2014. Figure 2.9 shows the interface of the DataPlayer. At the bottom, a control panel allows users to adjust replay speed, trigger pauses, or rewind as needed. In the upper-right corner, some text displays essential information about the data, such as the time, the target number, and if the tolerances for target validation are met (tgtRed). In the lower-left corner, a button allows the user to switch between two viewing modes: egocentric and free. In the egocentric view (Figure 2.9, left image), the perspective matches that of the participant during the experiment, while in the free mode (Figure 2.9, right image), the user can navigate around the scene, explore it from various angles and select a preferred view.

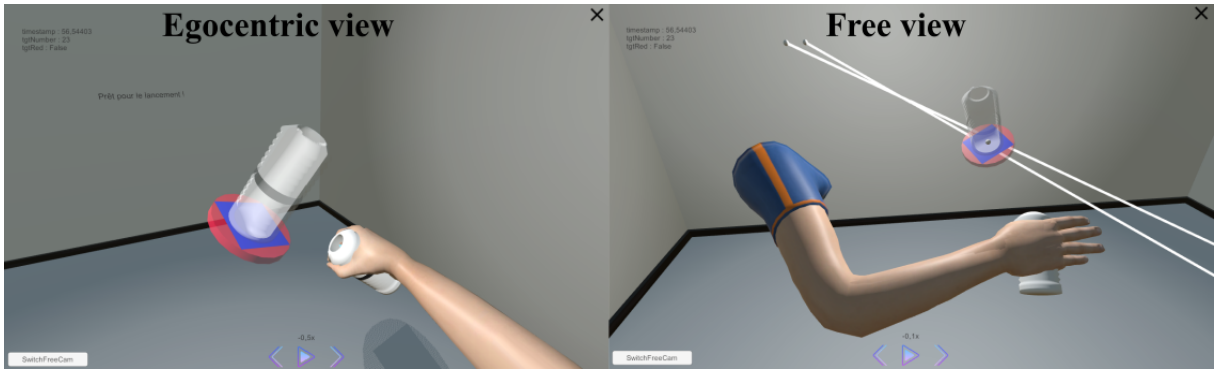



Figure 2.9: DataPlayer interface. The left image displays the DataPlayer showing an egocentric view from participant 8 reaching target 23 during the phase Test RNP AP. The right image shows the same data from a different viewpoint, selected using the free mode. The two white lines represent the pupil axis, and the white sphere indicates the point the participant is looking at.

The DataPlayer, being a Unity project, offers extensive customization options, granting users the freedom to tailor every aspect of the virtual environment according to their preferences and requirements. Users have the flexibility to modify various elements, including the shape, color, and dimensions of the target object. Furthermore, they can manipulate the surrounding environment by adding or removing objects, and adjust other parameters as needed (e.g., lighting conditions). Customization proves particularly advantageous in applications where the environment must reflect specific contexts. For instance, the ability to modify objects, attributes, and scenarios provides a fertile ground for computer vision scientists aiming to develop and test robust 6D pose estimation algorithms from egocentric vision with gaze (González-Díaz et al., 2019; Labbé et al., 2020; Labbé et al., 2022; Liu, 2023; Nguyen et al., 2022; Pérez De San Roman et al., 2017; Rogez et al., 2015). The project's malleability allows for simulating diverse real-world scenarios, transforming, for example, the environment into a kitchen setting, complete with culinary items, as shown in Figure 2.10 and in Illustrative Video 2 in Appendix B.

Figure 2.10: Examples of how the DataPlayer can be customized to combine real data about arm and gaze with various fictitious visual backgrounds and objects.

To facilitate this exploration and testing, we have integrated an option in the interface enabling the generation and recording of synthetic data based on the recorded data. This feature preserves the user's movements, gaze data, and target object poses while

allowing for modifications to all visual aspects of the environment. Consequently, researchers can leverage this capability to develop and evaluate computer vision algorithms for recognizing objects in complex environments, as the ground truth is known by design.

## 2.4 Discussion

In this chapter, I presented the experimental protocol designed to create a dataset that includes records of natural arm and head movements alongside visual and gaze information. Participants were actively involved in picking and placing objects in various positions and orientations in a virtual environment while ensuring a consistent seated posture.

While initially designed to gather natural arm and head movements and explore adaptation algorithms to perform reference frame transformations from orientation-only data, this dataset could represent a valuable resource across diverse domains such as sensorimotor control, movement-based prosthesis control, humanoid robotics, human-robot interactions, and gaze-guided computer vision.

This dataset can be used to revisit several issues lying at the core of human sensorimotor control. First, our 3D recordings of the entire kinematic chain (shoulder, elbow, wrist) are relevant to assessing how multiple joints are coordinated in a system exhibiting redundant DOFs. For instance, one may be interested in analyzing fluctuations in hand and joint kinematics across trials (as we never performed twice the same movement) and assessing to what extent a joint's deviation from its mean behavior can be compensated by adjustments in other joints (Latash et al., 2002; Scholz et al., 2000), so that the resulting hand position is minimally impacted. Additionally, the dataset can contribute to ongoing discussions about the straightness and smoothness of hand movements in 3D space (Atkeson and Hollerbach, 1985; Desmurget et al., 1995; Desmurget and Prablanc, 1997; Flash and Hogan, 1985; Hagiwara et al., 2020; Viviani and Flash, 1995).

As we recorded the maximal range of motion of each joint, this dataset can be used to investigate to what extent arm configurations can be accounted for by cost functions using postural comfort as a key variable (i.e., maximizing postures away from extreme joint angles; Cruse et al., 1993; Rossetti et al., 1994).

On a more applied side, the dataset offers a benchmark of natural arm movements with potential use in motor rehabilitation contexts (e.g., to generate assistance with upper limb exoskeleton for stroke rehabilitation; Nann et al., 2021) or for arms control in humanoid robotics (De Momi et al., 2016; Kim et al., 2006; Koskinopoulou and Trahanias, 2016). In the context of human-robot interactions, the dataset can be used either to feed methods aiming at reproducing natural arm movement on robotic platforms or to evaluate the human-likeness of these methods (see Gulletta et al., 2020 for a review of methods for human-like arm motion generation).

The incorporation of gaze recordings within the dataset can be leveraged to explore issues related to eye-head-arm coordination (Henriques et al., 2003; Pelz et al., 2001; Vercher et al., 1994) and 6D pose estimation from a first-person view or egocentric videos (González-Díaz et al., 2019; Labbé et al., 2020; Labbé et al., 2022; Liu, 2023; Nguyen et al., 2022; Pérez De San Roman et al., 2017; Rogez et al., 2015). Moreover, the DataPlayer allows users to replay and manipulate the visual environment as desired, as demonstrated in Illustrative Video 2 in Appendix B. This tool offers the opportunity to generate from the dataset new synthetic data that accurately reflects natural head, gaze, and arm movements. Since the ground truth of the object pose is known by design in VR, this synthetic dataset can be instrumental for the development and assessment of 6D pose estimation algorithms tailored for egocentric vision in real-world scenarios.

# Chapter 3

# Two models for head to shoulder reference frame transformation

As stated in Chapter 1.2, the prosthesis control previously developed by the team, PC ANN, relies on the position and orientation of the object to reach expressed in the armroot reference frame, whereas in real-world scenarios, it might only be available in a head-centered reference frame through gaze-guided computer vision. In this chapter, I present two methodologies for addressing the problem of performing reference frame transformation between the head and the armroot from incomplete, orientation-only data from the head and armroot, possibly retrieved with IMUs. Instead of using a kinematic model of the biomechanically complex region of the human spine and shoulder girdle, which would be hard to fine-tune for each participant (Krishnan et al., 2019), we took inspiration from how the brain performs coordinate transformation and represents peripersonal space (Chinellato et al., 2011; Hoffmann et al., 2010; Makin et al., 2013; Pouget et al., 2002; Salinas and Abbott, 2001; Serino, 2019; Snyder et al., 1998).

The work presented in this chapter was conducted with the assistance of Lucas Bardisbanian, who completed a six-month engineering internship under my supervision, developing the Armroot2Head ANN.

The work presented in this chapter is part of the following publication (Lento, Leconte, et al., 2024), IEEE Robotics and Automation Letters.

## 3.1 Armroot2Head Artificial Neural Network

Our first approach consisted in training offline an ANN, which we refer to as "Armroot2Head" (A2H), on the dataset collected in the previous experiment to predict the head position in the armroot reference frame. The A2H architecture consists of seven layers: an input linear layer, with different configurations tested as detailed later, two hidden layers, each with 128 units and rectified linear unit activation functions, three dropout layers with a dropout rate of 0.25 to prevent overfitting (one positioned after each layer), and an output linear layer.

Recognizing that the relation between the head and the armroot depends on the user's morphology as well as the particular head, trunk, and arm postures assumed during task execution, we explored diverse inputs for the ANN, ranging from head orientation to proximal joint angles (shoulder abduction/adduction and shoulder flexion/extension), participant height and segment dimensions between the shoulders, head, neck, and trunk. To identify the inputs that could effectively capture the intricate interdependencies within the data, we trained the ANN with various input combinations. The training procedure involved training the ANN on our dataset with the data of one participant excluded, and then testing the trained ANN on the data of the excluded participant. This iterative process was repeated for each participant and each combination of inputs. The outcomes of these offline training sessions are detailed in Figure 3.1, providing a visual representation of the performance of each ANN input configuration under evaluation and valuable insights into the impact of each input variable on the model's accuracy. As shown, removing proximal joint angles (SFE and SAA) increases the variability of the model's performance. In contrast, removing one segment dimension (N2H, T2N, or S2S) or all of them seems to improve the ANN's performance, suggesting that the segment dimensions contain low or redundant information, leading to overfitting. Therefore, we opted for the ANN with head orientation, participant height, and proximal joint angles as inputs, as this configuration demonstrated the best performance (refer to the ANN No Seg in the red rectangle in Figure 3.1).

Following the exclusion of segments between shoulders, head, trunk, and neck from the inputs of the A2H ANN, the model underwent retraining with additional data derived from a previous experiment conducted by the team (E. A. Segas et al., 2024). It is worth emphasizing that the supplementary data lacked recorded information about these segments, which prevented their inclusion in the previous evaluation. Figure 3.2a illustrates the outcomes of the A2H ANN offline training, presenting a scatter plot of actual data against predicted data for the x, y, and z components of the head-to-shoulder reference frame transformation. The coefficient of determination, denoted as $R^2$, offers insights into the model's fitting performance to the data. It is evident that the A2H ANN excels in determining positions along the y-axis but exhibits lesser accuracy along the other two axes. Furthermore, Figure 3.2b illustrates the distribution of prediction errors, showcasing a median prediction error of approximately 2cm. This value is lower compared to the previous evaluation (around 4cm), likely attributed to the increased volume of data.

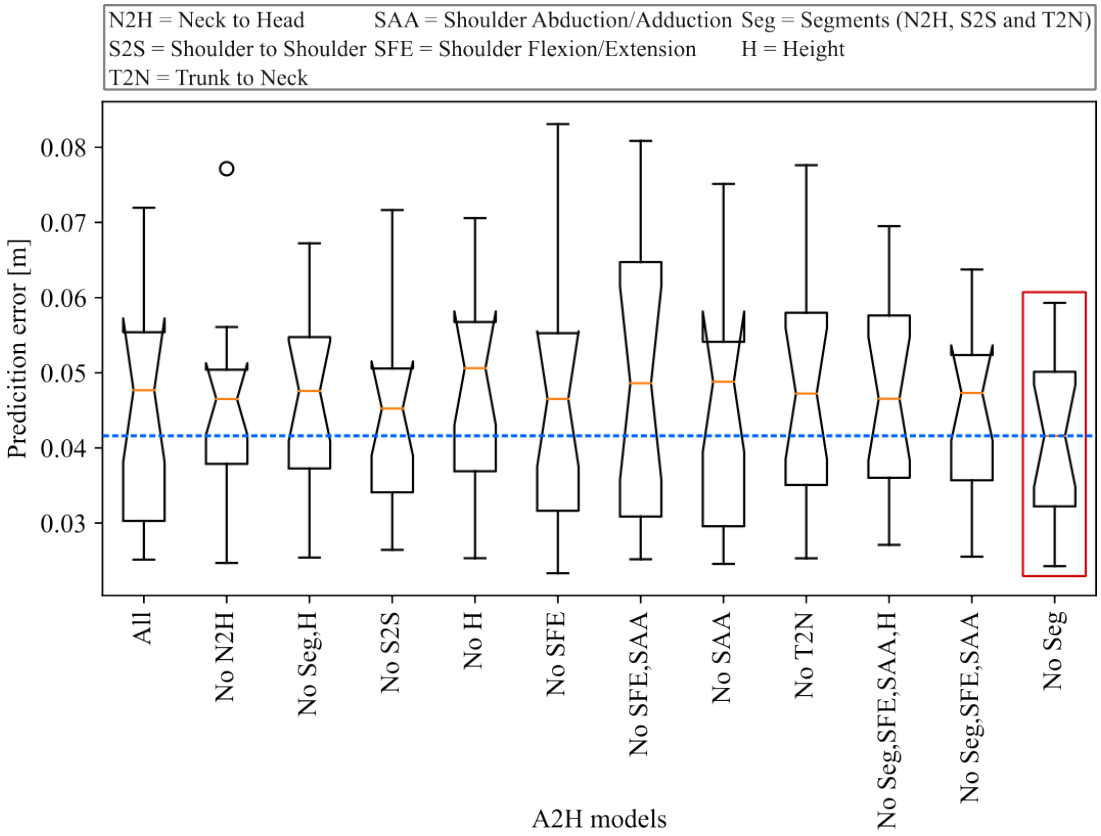



Figure 3.1: Training results for different input combinations of the A2H ANN. Each box plot represents the distribution of training accuracy across different input combinations, with whiskers indicating the range of accuracy values. The A2H ANN that exhibited the best performance corresponds to the input configuration labeled "No Seg," marked by a red rectangle. It demonstrated the lowest prediction error, as indicated by the blue dashed line.

Our first solution for deploying PC ANN in real-life settings was to combine it with the A2H ANN. Figure 3.3 illustrates the proposed control system, which integrates the PC ANN (red block) with the A2H ANN (yellow block). In such scenarios, we assume that the user will be wearing a head-mounted camera and that head and armroot orientations will be available from IMUs. Computer vision algorithms provide information about the rotation (R) and position (P) of the target in the head reference frame ($^{head}R_{tgt}$, $^{head}P_{tgt}$), which needs to be transformed into the armroot reference frame to be used by the PC ANN. First, the target orientation is transformed from the head to the armroot reference frame (rotation transform (RT) block). The same rotation is applied to the target position. Then, the A2H ANN predicts the head position in the armroot reference frame, which is applied to the resulting target position to complete the change of reference frame (translation transform (TT) block). The target information, now expressed in the armroot reference frame, is combined with the residual limb angles ($^{armroot}R_{shou}$) and used by the PC ANN to predict the angles of the prosthesis.

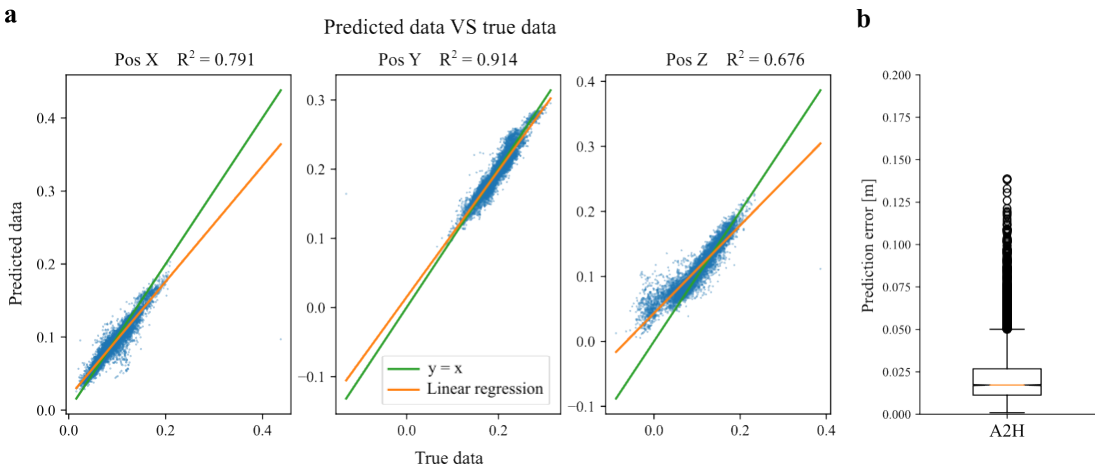



Figure 3.2: A2H ANN offline training results. **a.** Scatter plot comparing true data with predicted data. The orange line represents the linear regression of the data, and the green line represents the function y = x, indicating that the predicted data would be identical to the true data if the model were perfect. **b.** Box plot of median error in prediction of the ANN.

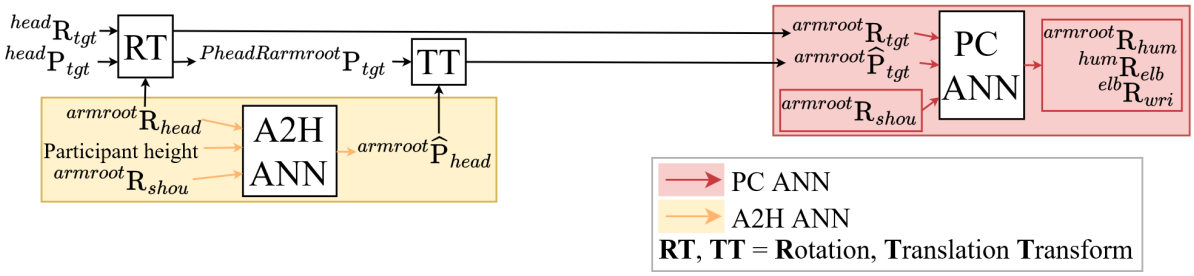



Figure 3.3: Control scheme of A2H ANN combined with PC ANN. The symbol $\widehat{P}$ stands for predicted.

Recognizing that errors in the A2H ANN predictions might lead to ineffective prosthesis control, causing participants to compensate by overusing their shoulder and trunk, we proposed and tested another approach, described in the next section.

## 3.2 Space Map

Prior research by the team already explored co-adaptation strategies in a simplified myoelectric context (Couraud et al., 2018), combined with a perturbation for which human adaptation is well characterized and modeled. As illustrated in Figure 3.4, participants were tasked with reaching targets displayed on a computer screen by controlling a cursor through myoelectric activity generated while producing isometric force at the wrist joint. They implemented a system that integrates two levels of adaptation—human

and machine—and tested it under both normal and perturbed conditions. The perturbed conditions involved constant directional error (visuomotor rotation), a scenario well-studied in human sensorimotor adaptation contexts (De Rugy and Carroll, 2010; Krakauer et al., 2000). The adaptation was achieved by dynamically adjusting the weights of the two models based on the difference between the input and output of the entire system.

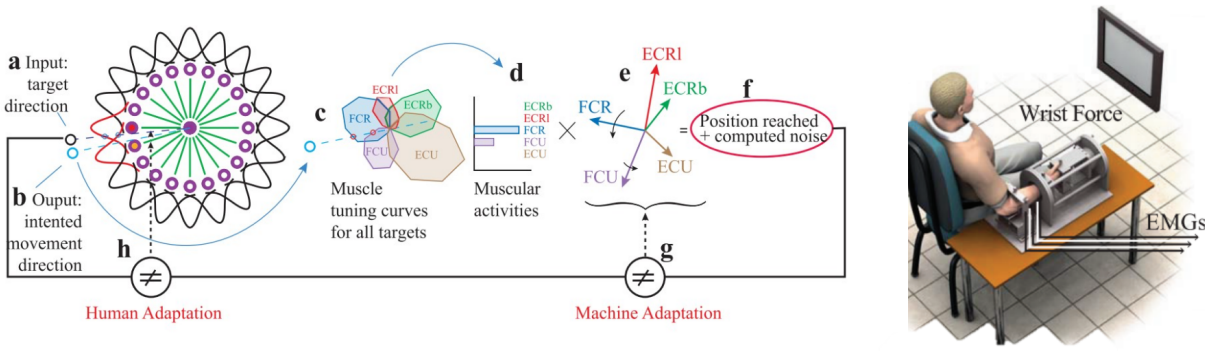



Figure 3.4: Overview of task and co-adaptation process proposed by Couraud et al., 2018. The adaptation was implemented in a network of directionally tuned neurons that takes the target direction as input (**a**) and produces the intended movement direction (**b**). Muscle tuning curves (**c**) were employed to extract the required muscular activities (**d**), which were then multiplied by the corresponding muscle-pulling vectors (**e**) to generate the resulting force vector. The force vector combined with noise defined the reached position on the target (**f**). This direction was compared to the input target direction (**a**) to estimate the angular error, which was used to adjust the direction of the pulling vectors of the virtual biomechanics (machine co-adaptation; **g**) as well as the weights of the network model that represents human adaptation (**h**; Figure adapted from Couraud et al., 2018).

Based on their model of human adaptation, realized as a network of directionally tuned neurons, I developed a similar algorithm with neurons distributed in a three-dimensional space and weights designed to encode multidimensional errors. The motivation behind this approach was to tackle the problem of head-to-shoulder reference frame transformation using a black box method while acknowledging that the relationship between the two frames is influenced by factors such as the pose of the target object and, more specifically, the posture of the arm, head, and upper body used to reach this object pose.

The algorithm, which we refer to as Space Map (SMAP), draws inspiration from the brain's multisensory integration processes (Makin et al., 2013; Pouget et al., 2002; Pugach et al., 2019; Salinas and Abbott, 2001; Snyder et al., 1998) and computes the head position in the armroot reference frame by combining information from multiple sensory modalities (i.e., computer vision from a head-mounted camera, combined with artificial neck and arm proprioception from orientation sensors on the head, trunk and arm). Specifically, the system can compute the prosthetic hand position in the armroot

reference frame through forward kinematics while concurrently detecting the prosthetic hand position in the head reference frame using computer vision. This hand position in the head reference frame can be aligned with the orientation of the armroot (i.e., multiplied by the orientation of the head in the armroot reference frame and thus expressed in a mixed reference frame which has position of the head and orientation of the armroot) and compared to the hand position in the armroot reference frame obtained through forward kinematics. The difference between the two represents the position of the head in the armroot reference frame. As this difference depends on both the user's morphology and the particular posture of their arm, head, and upper body while reaching targets in space, we chose to encode it in a spatial map (SMAP), the input of which is the target the user is reaching. Inspired by the brain's encoding of peripersonal space (Chinellato et al., 2011; Di Pellegrino and Làdavas, 2015; Hoffmann et al., 2010; Serino, 2019), this spatial encoding was achieved by adapting online the weights of a single-layer network of spatially tuned neurons, illustrated in Figure 3.5. Each neuron is characterized by a 3D position in the input space and a bell-shaped activation function centered on its position. When the target position is sent as input to the SMAP, each neuron fires according to its distance from this input, and the output is computed as the weighted sum of all firing rates.

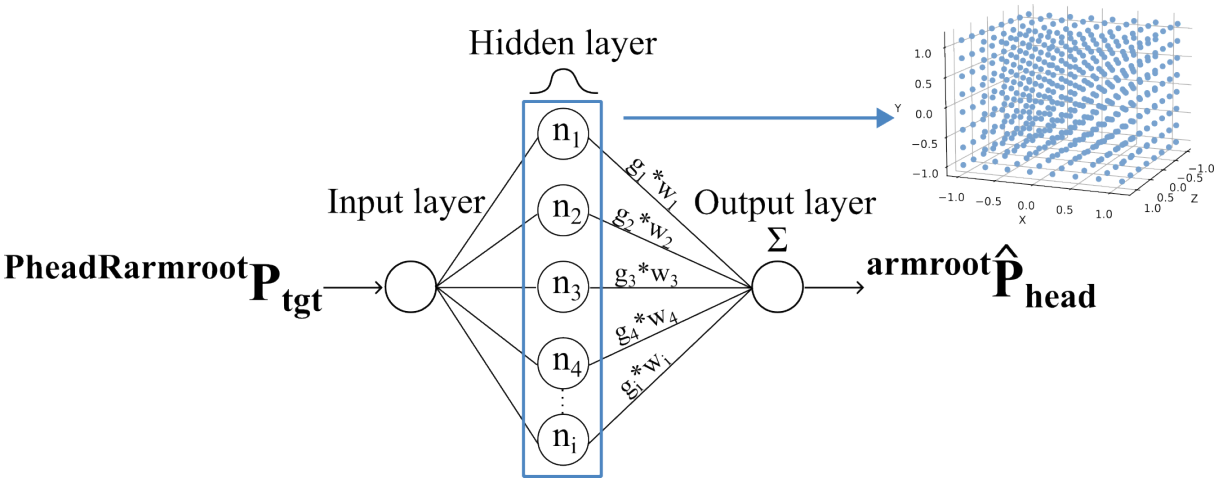



Figure 3.5: SMAP algorithm structure.

The weights of the SMAP are updated online according to the following formula:

$$w_i(t+1) = w_i(t) + \alpha \cdot g_i \cdot error$$

$$error = ({}^{armroot}P_{hand} - {}^{P_{head}R_{armroot}}\ P_{hand}) - \sum_{i}^{N} g_i \cdot w_i$$

in which $w_i$ is the weight of the *i-th* neuron, $\alpha$ is the learning rate, $g_i$ is the firing rate of the *i-th* neuron (Gaussian activation function), $N$ is the number of neurons, and $\sum_i^N g_i \cdot w_i(t)$ represents what the algorithm already learned.

Figure 3.6 illustrates the proposed control system, integrating the SMAP with PC ANN alongside the online adaptation process of the SMAP. Computer vision algorithms provide information about the target and the hand in the head reference frame. As in Figure 3.3, we transform the target orientation from the head to the armroot reference frame and apply the same rotation to the target and hand positions (rotation transform (RT) block). The SMAP takes as input the resulting target position (${}^{P_{head}R_{armroot}}P_{tgt}$) and computes the head position in the armroot reference frame. The predicted head position is then applied to the input to complete the change of reference frame (translation transform (TT) block). Simultaneously, real-time weight updates occur based on the difference between the hand position detected by computer vision and aligned with the orientation of the armroot (${}^{P_{head}R_{armroot}}P_{hand}$) and the hand position obtained through forward kinematics (${}^{armroot}P_{hand}$). The SMAP treats this difference as an error signal to adapt its weights and cancel the error.

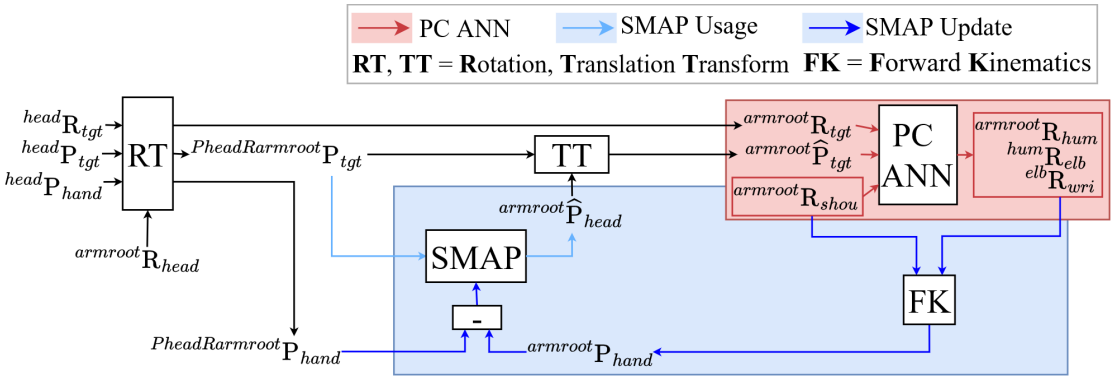



Figure 3.6: Control scheme of SMAP combined with PC ANN.

### 3.2.1 Tuning parameters

The algorithm's tuning process involved three parameters: the learning rate ($\alpha$), the width of the Gaussian function ($\Sigma$), and the distance between neurons ($n_d$). These parameters determine the algorithm's performance and efficiency.

The learning rate determines the magnitude of the error signal learned by the algorithm in each iteration of its weights update. A higher learning rate may result in faster convergence, but it might lead to unnatural behaviors of the virtual arm (e.g., movements in unintended directions). Conversely, a lower learning rate may lead to slower convergence but can offer more stable behaviors of the virtual arm. The Gaussian width determines the spatial generalization property of the algorithm. If its value is set too small, the algorithm becomes too specialized to the training data, resulting in reduced performance when presented with new, unseen data. On the other hand, if it is set too large, the algorithm may fail to capture essential patterns and variations in space. The distance between neurons directly influences the resolution of the algorithm. A smaller value results in higher precision, enabling more accurate corrections. However, it also

entails a higher number of neurons, leading to increased computational complexity and longer processing times.

In the tuning process, we explored different values for these three parameters of the algorithm. We replayed the data of one participant from the dataset presented in Chapter 2 while updating the weights of the SMAP as if it were operating in real-time. To ensure consistency, we repeated this process for two additional participants. Moreover, each participant's data was replayed twice as they reached the same set of targets a second time. This repetition provided insights into the algorithm's memory capacity, as explained in the next paragraph.

In the process of evaluating different sets of parameters, we took into account three constraints. The first constraint involved ensuring that the product of the learning rate and the sum of neuron firing rates remained equal to or less than 1 to prevent the algorithm from attempting to learn beyond its intended capacity (i.e., the magnitude of the error signal). The second constraint focused on the convergence rate, with the objective of ensuring that the participants could achieve the task successfully within a single trial. The last constraint pertained to the algorithm's capacity to memorize the required corrections. We aimed for the algorithm to memorize at least 50% of the corrections within a single trial. To evaluate this, we replayed the data twice and observed the convergence curve upon encountering the same target for the second time. This provided valuable insight into the extent to which the algorithm had assimilated the error. By incorporating these constraints, we aimed to enhance the robustness and generalization capability of our spatial encoding while mitigating issues such as overfitting and instability.

The tuning process automatically filtered out parameter sets that did not meet the three constraints described previously, retaining only sets that met the criteria and demonstrated superior performance. From these sets, we empirically chose the one that seemed to achieve the most natural and intuitive user experience. We set the learning rate at 0.004, while the Gaussian width and the distance between neurons were both determined to be 3cm, with this value representing the positional tolerance required for grasping.

### 3.2.2 Convergence and stability analysis

In addition to the previously presented tuning procedures, offline tests were conducted to evaluate the algorithm's strengths and limitations. These tests aimed to assess its ability to handle various error types and identify any inherent limits in learning capacity.

Findings from the tuning parameters process revealed that the algorithm effectively adapts its weights to absorb positional errors that are more complex than a constant offset. For instance, we tasked the algorithm with learning the head position in the armroot reference frame, whose value is not static but dynamically varies based on the specific head, body, and arm postures adopted when reaching various targets in space. Therefore, we opted to simulate a more complex scenario where the user wore a head-mounted camera whose pose could shift, introducing constant offsets in position and

orientation. Similar to the tuning process, we replayed the data from the dataset, while updating the weights of the SMAP as if it were working in real-time. Then, we examined the convergence of the SMAP, computed as the difference between the target position and orientation in the armroot reference frame corrected by the SMAP and its true value.

We conducted our testing within a simplified framework, introducing a constant error in orientation along one axis, which consequently induced a linear error in position along the other two axes. The objective was for the algorithm to learn this error. Hence, we did not take into account the reference frame transformation issue and assumed the head position in the armroot reference frame was known (we utilized the dataset from the VR experiment, where all data, including the recorded head position, was available). However, it is crucial to note that the principles and methodologies applied seamlessly extend to more intricate contexts involving multi-axial errors and, thus, more complex frameworks without any fundamental alterations.

The initiation point of our evaluation was to define a reachable target pose, an initial hand pose, and a constant angular error $\theta$ along the z-axis. This error, while being a constant offset in the orientation space, induces a linear error in position along the remaining two axes, which can be calculated using the following formula:

$$\begin{bmatrix} x_{det} \\ y_{det} \end{bmatrix} = \begin{bmatrix} cos(\theta) & -sin(\theta) \\ sin(\theta) & cos(\theta) \end{bmatrix} \begin{bmatrix} x_{true} \\ y_{true} \end{bmatrix} \tag{3.1}$$

in which $x_{det}, y_{det}$ are the coordinates of the target (or the hand) wrongly detected, $x_{true}, y_{true}$ are the coordinates of the true target (or hand) and $\theta$ is the error angle.

After computing the initially wrongly detected pose of the target and the hand, along with the corresponding error, this data was utilized to adapt the weights of the SMAP and correct the target pose. Assuming that the hand would precisely reach the corrected target pose, we iterated this process and evaluated the loss function as the difference between the actual and corrected target poses.

When analyzing the behavior of the algorithm during these offline tests, we encountered three distinct scenarios: convergence, infinite oscillation, and divergence, each exemplified in Figure 3.7. For each error angle, 60°, 88°, and 90°, three plots are presented in Figure 3.7: one depicting the algorithm's loss in position (left column), another showing the loss in orientation (middle column), and the third illustrating the trajectory of the corrected target throughout the adaptation process (right column). This trajectory begins from the initially wrongly detected position and should end at the final corrected position (i.e., actual target position), indicated by the green point. Furthermore, a color bar is included to represent the dynamics of the target position update, transitioning from yellow to dark purple.

Convergence occurred when the loss fell below the grasping constraint utilized during the VR experiments, indicating successful error correction. This can be observed in Figure 3.7 for an error angle of 60°, where both the position and orientation loss curves steadily decrease towards the constraint. Conversely, for an error angle of 88°, the loss

remained unchanged, resulting in a circular pattern in the position loss curve and failing to achieve convergence. This behavior suggests that the algorithm was stuck in an oscillatory loop without making progress toward minimizing the loss. Lastly, divergence occurred when the loss continued to increase over time, signifying the algorithm's failure to correct errors and optimize performance. This behavior is evident in Figure 3.7 for an error angle of 90°, where the loss curve consistently rises.

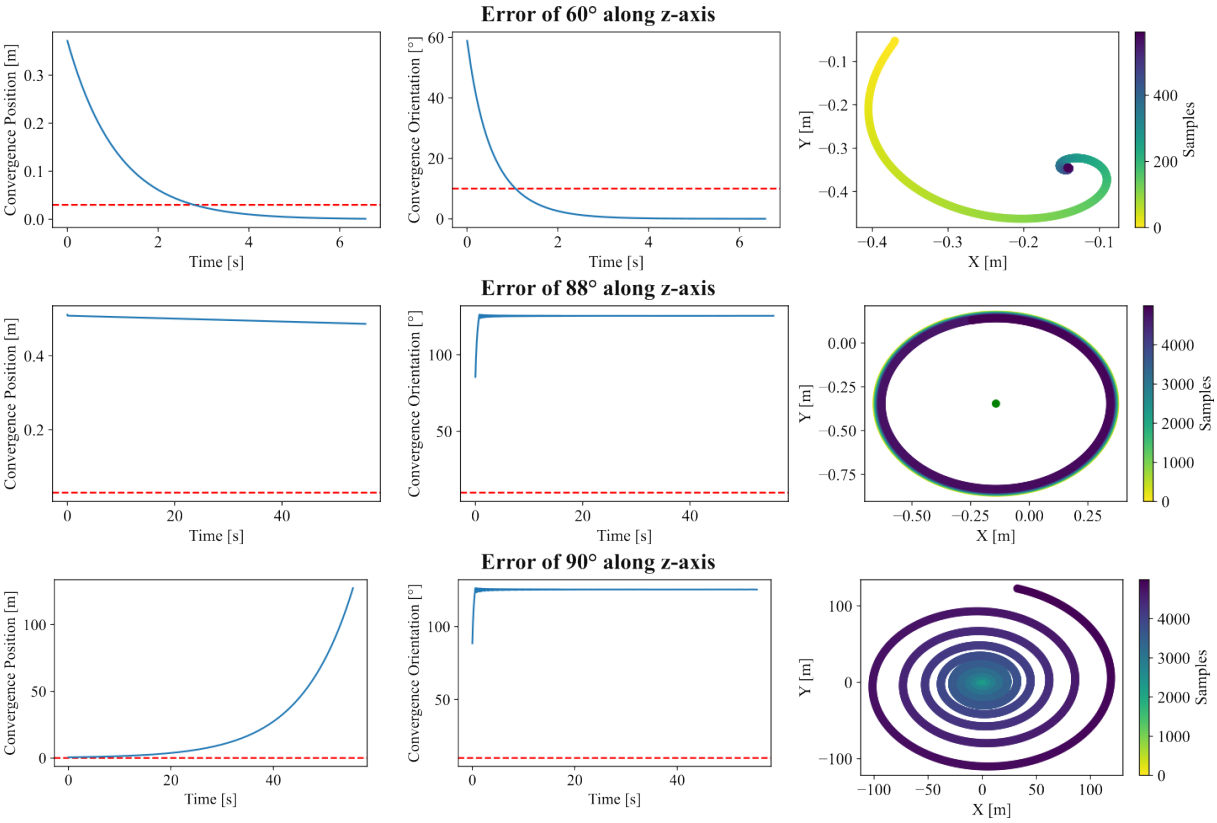



Figure 3.7: SMAP convergence analysis for angular errors of 60°, 88°, and 90°. Plots display the algorithm's loss in position and orientation, along with the trajectory of the corrected target from the initial wrongly detected position to the final corrected position (i.e., the actual target position, represented by the green point). The dashed red lines represent the tolerances on grasping used in the VR experiments (3cm and 10°).

The offline test results shown in Figure 3.7 revealed an inherent limitation in the algorithm's learning: it could not effectively correct a constant error in orientation larger than 88°. In such scenarios, the SMAP faces challenges in rectifying the errors, leading to instability in the adaptation process.

The convergence and stability of the algorithm hinge primarily on the learning rate, which determines the rate at which the model adapts its weights and assimilates the error. Therefore, we conducted additional offline tests, varying the learning rate values, to determine the angles at which the algorithm fails to learn the necessary corrections, resulting in divergence instead. This comprehensive analysis led to the creation of a

divergence curve, depicted in Figure 3.8, which offers valuable insights into the algorithm's learning limitations across various learning rates. The analysis demonstrates that as the learning rate approaches zero, the divergence angle converges to 90°, revealing that the algorithm can't correct orientation errors when the resulting angular error exceeds 90°.

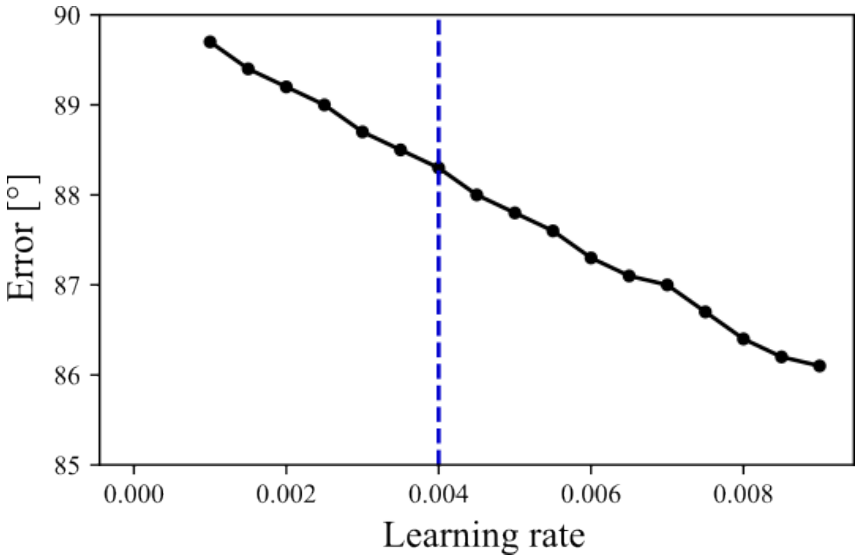



Figure 3.8: SMAP divergence curve. It shows the angular error at which the algorithm begins to diverge based on the learning rate. The circles denote data points acquired from the analysis. The blue dashed line represents the specific learning rate value chosen in the tuning process.

To summarize, the analysis of the convergence and stability of the algorithm revealed a significant limitation: with the parameters chosen through the tuning process, the SMAP cannot effectively correct orientation errors larger than 88°. Moreover, additional tests varying the learning rate value demonstrated inherent limitations in the algorithm's ability to correct orientation errors exceeding 90°, resulting in divergence rather than convergence. These findings highlight the limitations and boundaries of the algorithm's applicability in real-world scenarios.

## 3.3 Combination of Armroot2Head Artificial Neural Network and Space Map

Our first approach, A2H ANN, revealed a prediction error of approximately 2cm. Recognizing that this value might not align with our intended precision requirements and given the SMAP's ability to learn from errors, we implemented a third solution by combining the first two methods, thereby using the SMAP to correct errors that would persist in the A2H ANN predictions.

Figure 3.9 illustrates the control scheme combining the A2H ANN (yellow block) with the SMAP (blue block) and the PC ANN (red block). Computer vision algorithms provide information about the target and the hand in the head reference frame. First, we transform the target orientation from the head to the armroot reference frame and apply the same rotation to the target and hand positions (rotation transform (RT) block). The A2H ANN predicts the position of the head in the armroot reference frame, which is applied to the resulting target and hand position (first translation transform (TT) block). The SMAP takes as input the resulting target position and computes the correction to rectify the errors in the A2H prediction. This correction is applied to the SMAP input (second translation transform (TT) block). Simultaneously, real-time weight updates occur based on the difference between the resulting hand position translated by the A2H prediction and the hand position obtained through forward kinematics. The SMAP treats this difference as an error signal to adapt its weights and cancel the error.

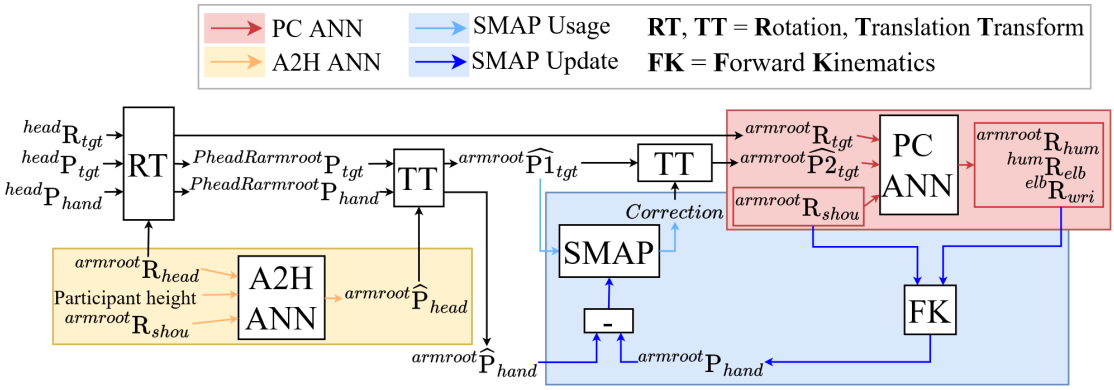



Figure 3.9: Control scheme of SMAP and A2H combined with PC ANN.

To assess the effectiveness and adaptability of the proposed methods and their associated control systems, we conducted two experiments in VR involving participants with intact limbs and those with transhumeral limb loss, as detailed in the following chapter.

# Chapter 4

# Experimental validation in a virtual environment

In the previous chapter, I presented three methods designed to retrieve the head position in the armroot reference frame and, thus, perform the required transformation from the head and armroot from reduced information that might be available in real-world settings. In this chapter, I describe two experiments conducted within a VR environment involving participants with intact limbs and with transhumeral limb loss. In this experimental setup, we simulated working with reduced information "as if" obtained from computer vision and compared the performances of the proposed methods and their associated control with those of the PC working in the ideal scenario, where target information is directly acquired and transformed in the armroot reference frame, as all information is readily available within the VR system.

The work presented in this chapter is part of the following publication (Lento, Leconte, et al., 2024), IEEE Robotics and Automation Letters.

## 4.1 Materials and Methods

This section outlines all the elements put in place to conduct the experiments presented in this chapter.

### 4.1.1 Participants

Exp1 was conducted on twelve intact-limbs participants (four males and eight females) aged 20-28 (mean 22.83; SD 1.92), whereas Exp2 was conducted on six participants presenting a transhumeral amputation (five males), aged 34-64 (mean 49; SD 11.5). Participants' handedness was assessed using the EHI (Oldfield, 1971; participants Exp1 mean 82; SD 17). The details regarding the amputation of each participant of Exp2 are outlined in Table 4.1.

Table 4.1: Exp2 participants' amputation description

| Id | Time since amputation | EHI pre-amputation | Amputated arm side |
|---|---|---|---|
| 1 | 12 years | Right-handed | Right |
| 2 | 14 years | Ambidextrous | Right |
| 3 | 2 years | Right-handed | Right |
| 4 | 37 years | Right-handed | Right |
| 5 | 2 years | Left-handed | Right |
| 6 | 25 years | Right-handed | Left |

All participants gave their informed consent in accordance with the Declaration of Helsinki and with the local ethics committee (CPP Est II: n°2019-A02890-57). None of the participants suffered from any mental or motor disorder that could interfere with the task performance (except limb difference in Exp2).

### 4.1.2 Methodological changes

The experimental setup and the core definition of the task remained unchanged from those of the dataset, presented in Chapter 2. The target set was generated using the dataset, specifically the data from the Initial Acquisition phases. From the joint configurations produced by the participants, we computed angular limits for each joint, setting the minimum value at the 5th percentile and the maximum at the 95th percentile. A generator algorithm drew at random 7-DOFs arm angular configurations following a multivariate uniform probability distribution. Then, the target positions and orientations were obtained through forward kinematics, using a chain based on the participant limb dimensions, and filtered according to the same criteria of the previous experiment, excluding targets too close to the participant's trunk or legs and targets pointing downward. This process was iterated until a target set of 100 targets was obtained. For Exp2, a supplementary filter was added, which considered the limits of the participant's residual limb.

Beyond the teleoperation of the virtual arm, the participants tested the PC ANN, the prosthesis control previously developed by E. Segas et al., 2023, and its combination with A2H ANN and SMAP (*cf* Protocol). While the proximal joints of the virtual arm continued mimicking the actual proximal joints of the participants, PC ANN predictions were employed to operate on the distal joints. We conducted an offline evaluation of the PC ANN's grasping accuracy and adjusted the tolerances to 3cm and 10° accordingly to test outcomes (see Appendix C).

Furthermore, when the grasping constraints were successfully met (i.e., virtual hand in the target zone), the participant was required to press a key on the keyboard using their opposite hand in Exp1 and their foot in Exp2 to grasp/release the bottle and complete the task.

At the onset of each phase, the reference frame in which the target set is expressed was readjusted based on the participant's initial posture to ensure natural shoulder movements.

In Exp2, the dimensions of the virtual arm on the side with limb difference were adjusted based on the participant's height and the size/length ratios of standard segments proposed in Herman, 2016.

### 4.1.3 Protocol

Experimental protocols are shown in Figure 4.1. The keyword "fam" stands for familiarization, meaning that these phases were designed to familiarize the participants with the environment and virtual arm controls. In these phases, the target set corresponded to the second half of the main target set. The other keywords represent the control used in that specific phase: "Nat" stands for natural, whereby all joints of the virtual arm follow the user's actual arm movements. "PC" stands for PC ANN, whereas "SMAP" and "A2H" stand for the use of the SMAP and A2H algorithms and their associated control systems (*cf* Chapter 3).

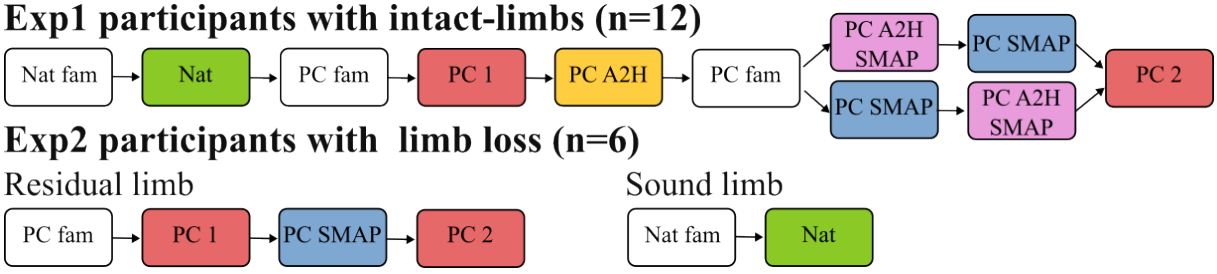



Figure 4.1: Protocols of experiment 1 and 2.

Exp 1 included nine phases and aimed to compare the performances of the three methods with that of the ideal case PC, in which the change of reference frame from the head to the armroot was performed with high precision from all the information available in the VR system. A first block (PC 1) was conducted early in both protocols

for participants to get used to this ideal control before being tested in our alternative controls, and a second block (PC 2) was conducted at the end of both protocols to ensure comparison with the ideal control once participants were fully accustomed to the device and controls.

In the phase PC A2H, the position of the head in the armroot referential was predicted by the A2H ANN. Despite offline tests revealing an average prediction error of approximately 2cm, we aimed to assess how this error might affect the control and the task performance. A “washout” phase, during which the participants reverted to using the ideal PC control again (PC fam), was then included to wash out possible compensations associated with this control before testing other controls.

In the phases PC SMAP and PC A2H SMAP, the head position was computed by the SMAP and the combination of the SMAP and A2H ANN, respectively. Recognizing that the learning process would take time and that the initial performances might not be optimal, we repeated the target set twice in these phases. We anticipated that the second attempt would yield results closer to the ideal case, as the algorithm would have assimilated the necessary corrections. Moreover, the order between PC SMAP and PC A2H SMAP was counterbalanced among participants.

Exp2 included six phases: four involving the residual limb followed by two involving the sound limb. Given the unsatisfactory results of A2H in Exp1, we chose to remove it from the protocol and compare only the SMAP with the ideal case PC.

### 4.1.4 Data reduction and metrics

This section explains the data filtering methods applied to the data recorded in the experiments described above and details the metrics used to evaluate and compare the performance of the proposed control strategies.

#### 4.1.4.1 Filtering

A filtering process was carried out to eliminate measurement errors associated with motion capture, thus excluding compromised targets from the analysis. Two filters were employed: one for "freezing" behavior, occurring when at least one sensor measure remained unchanged for a minimum of 0.5s, and another for "jumping" behavior, meaning when a sensor measure had a position discontinuity of at least 1cm between two consecutive samples. On average, this filtering process led to the exclusion of 2.52% targets per participant and phase in Exp1 and of 7.32% in Exp2.

#### 4.1.4.2 Performances metrics

Success rate, movement time, and shoulder spread volume were computed for each participant across the phases Nat, PC A2H, and PC 2. The phases PC SMAP and PC A2H SMAP were divided into two blocks (1B and 2B), with each block representing a repetition of the target set. Metrics were computed separately for each block. Given

the high success rate, with all medians exceeding 90%, movement time and shoulder spread volume were computed exclusively for successful targets across all phases.

#### 4.1.4.3 Subjective metrics

Qualitative feedback from participants about their experiences with each control strategy was assessed for a more comprehensive understanding of how the users perceive the control system, the mental effort required to use it, and their overall satisfaction with its usability. During the experiment, after testing for the first time each control (excluding the familiarisation phases), two surveys, the System Usability Scale (SUS; Gronier and Baudet, 2021) and the first part of the Prosthesis Task Load indeX (raw-ProsTLX; Parr et al., 2023), were administered to the participants to evaluate usability and cognitive load, respectively (see Appendix D and E).

The SUS assesses the usability of a system, defined as the user's ability to successfully perform the task for which the system was created (i.e., effectiveness) without overusing cognitive and physical resources (i.e., efficiency) and with a pleasant user experience (i.e., satisfaction). In this survey, the users assess their experience with the system by rating their agreement or disagreement with ten statements using a Likert scale (i.e., a 5-point scale ranging from 1: "Strongly Disagree" to 5: "Strongly Agree"). The total SUS score, ranging from 0 to 100, is calculated using the following formula:

$$SUS_{score} = \left( \sum_{I=1}^{5} (I_{odd} - 1) + \sum_{I=1}^{5} (5 - I_{even}) \right) * 2.5 \tag{4.1}$$

in which $I_{odd}$ and $I_{even}$ correspond to an odd or an even survey item, respectively. The score is then associated with the corresponding adjective of the scale proposed in Bangor, 2009. The higher the score, the better the participant perceives the system's usability.

In the first part of the ProsTLX, the participants rate eight categories from 0 to 20: mental demands, physical demands, visual attention, frustration, situational stress, time pressure, conscious processing, and uncertainty. The final score is the mean of the eight sub-ratings. A lower score indicates that participants required less cognitive effort to use this control strategy.

### 4.1.5 Statistical analysis

The statistical analysis was conducted using the R software with a significance level ($\alpha$) set at 0.05, applying a Bonferroni correction for multiple comparisons. Following the results of the normality test (Shapiro-Wilk) and the test for homogeneity of variances (Maulchy's Test), either a repeated measures ANOVA or a Friedman test was performed for movement time, shoulder spread volume, SUS score, and rawProsTLX score. If a significant difference was detected at this stage, post-hoc tests were carried out (Tukey's test or Conover's test, respectively). For success rate, a Cochran's Q test was performed, followed by a McNemar test if a significant difference was detected.

## 4.2 Results

Figure 4.2 illustrates the analysis results of Exp1. With PC A2H, the lower success rates, longer movement times, and increased shoulder spread volumes (PC vs PC A2H; n = 12; median success rates of 99% vs 91%; median movement times of 1.09s vs 1.64s; median shoulder spread volumes of 0.28$dm^3$ vs 1.41$dm^3$) all indicate that the error in the A2H predictions does affect the control.

In PC SMAP 1B, movement times and shoulder spread volumes are statistically higher than in PC due to the initial learning phase of the SMAP algorithm, during which it has to assimilate substantial errors (PC vs PC SMAP 1B; n = 12; median success rates of 99% vs 99%; median movement times of 1.09s vs 1.38s; median shoulder spread volumes of 0.28$dm^3$ vs 0.83$dm^3$). The longer movement times and increased shoulder spread volumes probably reflect the adaptation period required for the system to learn and memorize the necessary corrections. This is confirmed by the fact that the performances of PC SMAP 2B (the second repetition of the same target set) are drastically improved. The learning process required time, but once the corrections were memorized, its performances closely resembled those of PC (PC vs PC SMAP 2B; n = 12; median success rates of 99% vs 99%; median movement times of 1.09s vs 1.18s; median shoulder spread volumes of 0.28$dm^3$ vs 0.47$dm^3$).

The combination of A2H with SMAP exhibited performance similar to the ideal case PC. As the A2H ANN prediction error is relatively small, this control system has proved to be reliable and accurate in executing the reference frame transformation. Furthermore, as the SMAP has to learn small corrections, there is no statistical difference between the first and second repetitions of the target set (PC vs PC A2H SMAP 1B vs PC A2H SMAP 2B; n = 12; median success rates of 99% vs 99% vs 99%; median movement times of 1.09s vs 1.17s vs 1.14s; median shoulder spread volumes of 0.28$dm^3$ vs 0.31$dm^3$ vs 0.33$dm^3$).

The results of the subjective assessments (SUS and rawProsTLX scores) corroborate the objective metrics. In this context, the control Nat serves as the baseline, as the virtual arm mimicked the participant's actual arm movements. The only control strategy exhibiting a significant difference from the baseline is PC A2H. As the error in prediction did affect the control, participants necessitated additional physical and mental effort to achieve the task. Control strategies involving SMAP were evaluated with scores similar to the PC and Nat phases, suggesting that the algorithm does not significantly influence the virtual arm's behavior, thereby maintaining natural reaching movements (Nat vs PC vs PC A2H vs PC SMAP vs PC A2H SMAP; n = 12; median SUS scores of 92.5 vs 78.75 vs 55 vs 88 vs 83.75; median raxProsTLX scores of 3.19 vs 5.62 vs 6.78 vs 4.62 vs 4.5).

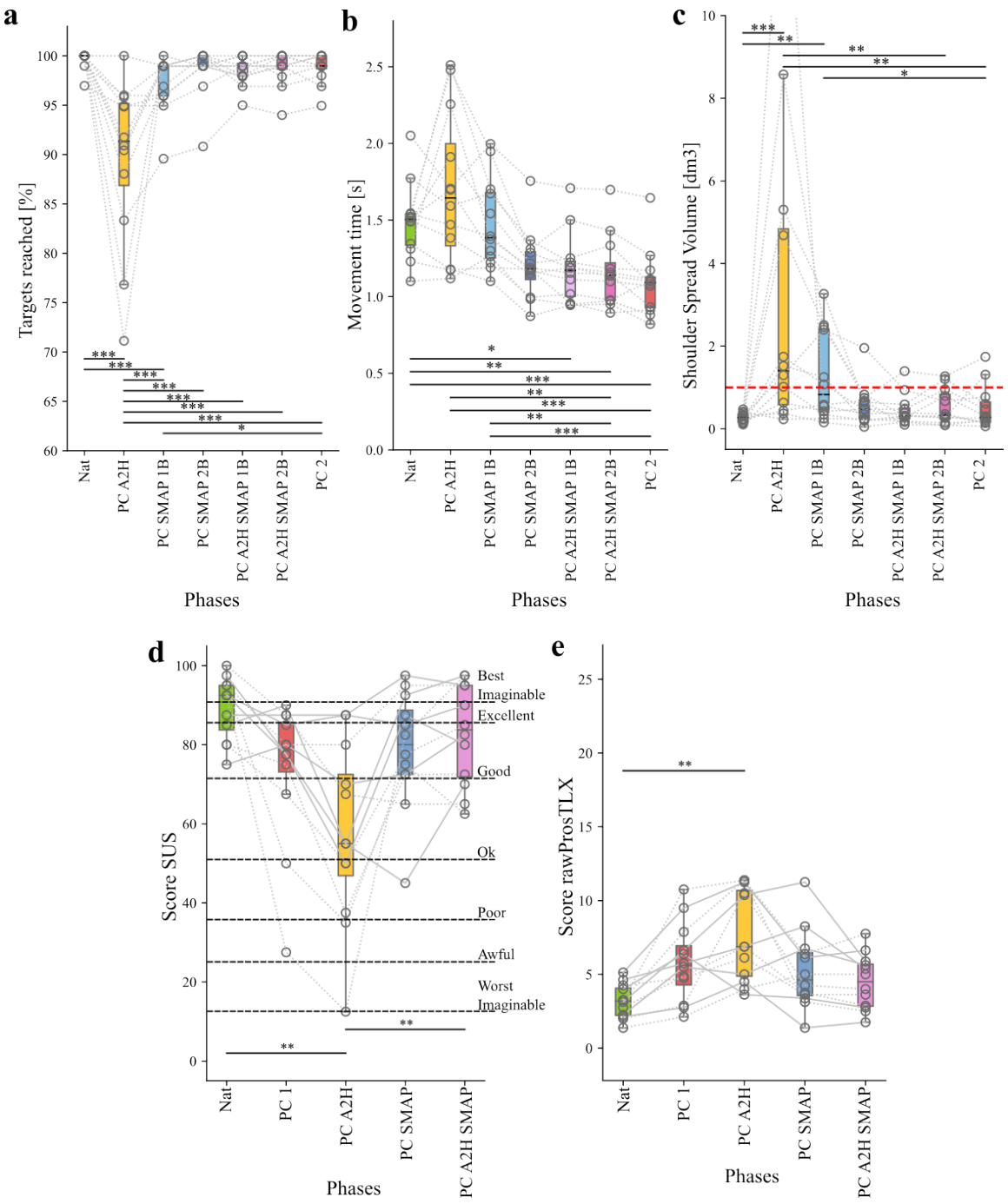



Figure 4.2: Results of objective (**a**-**c**) and subjective (**d**-**e**) metrics of Exp1 are presented for each phase. Individual data are represented by hollow dots and dotted lines. Stars represent significant differences between control strategies with * for $p < 0.05$, ** for $p < 0.01$ and *** for $p < 0.001$. **a.** Success rate (Cochran's Q test $Q = 304.45$, $p < 0.001$). **b.** Movement time (Friedman chi.sq = 56.57, $p < 0.001$). **c.** Shoulder spread volume (Friedman chi.sq = 36.85, $p < 0.001$). The dashed red line represents a volume of 1 $dm^3$(=1L). **d.** SUS score (Friedman chi.sq = 19.68, $p < 0.001$). **e.** rawProsTLX score (Friedman chi.sq = 18.02, $p < 0.01$). The detailed results of the statistical analysis are available in Table F.1-5 of Appendix F.

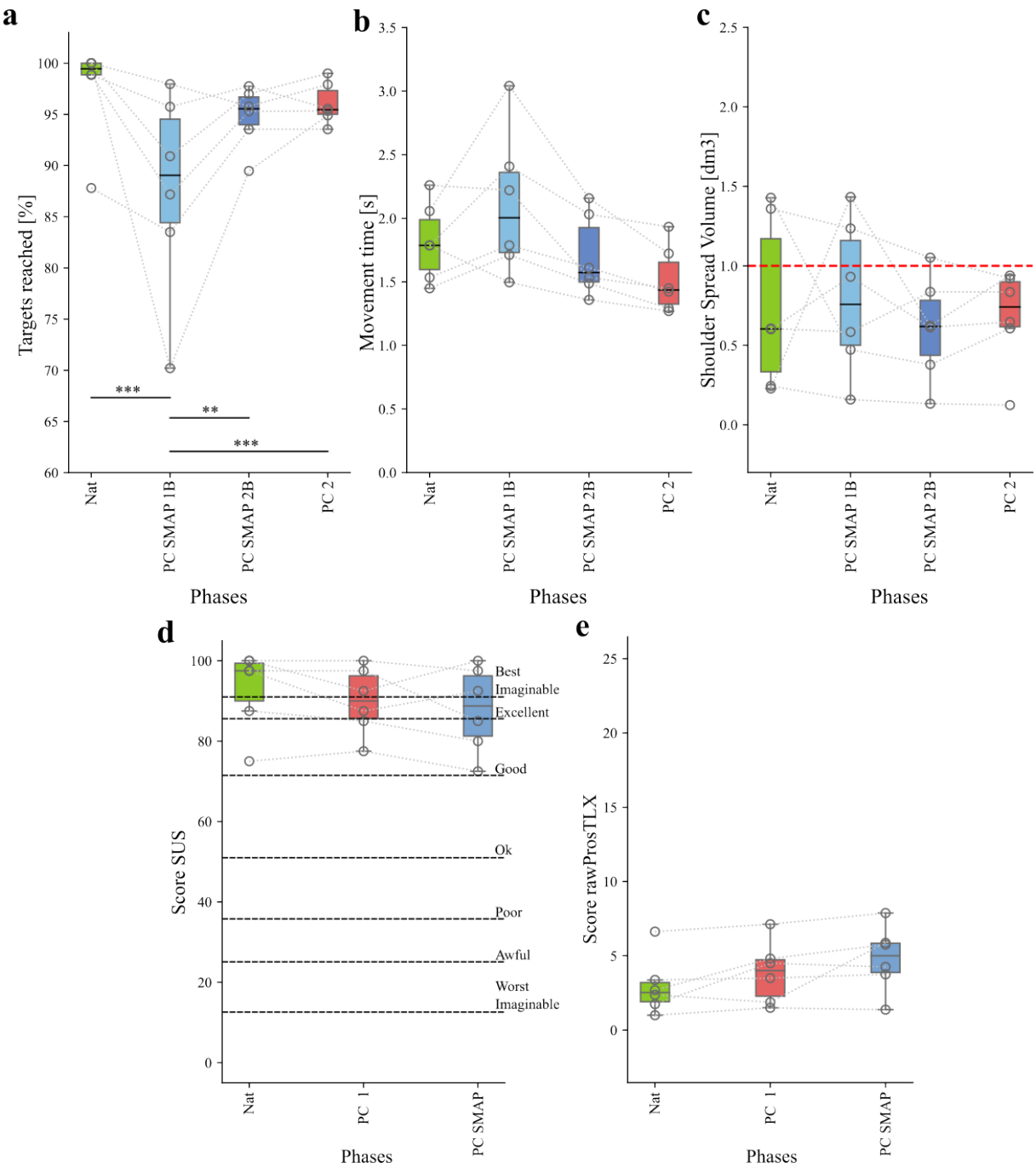



Figure 4.3: Results of objective (**a**-**c**) and subjective (**d**-**e**) metrics of Exp2 are presented for each phase. Individual data are represented by hollow dots and dotted lines. Stars represent significant differences between control strategies with ** for $p < 0.01$ and *** for $p < 0.001$. **a.** Success rate (Cochran's Q test Q = 45.57, p < 0.001). **b.** Movement time (RM ANOVA DFn = 3, DFd = 15, F = 6.15, p = 0.006, ges = 0.277). **c.** Shoulder spread volume (Friedman chi.sq = 1, p = 0.801). The dashed red line represents a volume of 1 $dm^3$(=1L). **d.** SUS score (Friedman chi.sq = 4.66, p = 0.097). **e.** rawProsTLX score (Friedman chi.sq = 7, p = 0.03). The detailed results of the statistical analysis are available in Table F.6-8 of Appendix F.

Figure 4.3 illustrates the analysis results of Exp2. Movement times tended to be longer, and shoulder spread volumes tended to be higher than in Exp1. Despite these observed differences, when evaluating the control strategies, PC SMAP 2B exhibited performances similar to PC across all metrics (PC vs PC SMAP 2B; n = 6; median success rates of 95% vs 95%; median movement times of 1.63s vs 1.77s; median shoulder spread volumes of 0.74$dm^3$ vs 0.62$dm^3$).

The longer movement times, increased shoulder spread volumes, and lower success rates in PC SMAP 1B (PC vs PC SMAP 1B; n = 6; median success rates of 95% vs 89%; median movement times of 1.63s vs 2.38s; median shoulder spread volumes of 0.74$dm^3$ vs 0.76$dm^3$) can be attributed to the initial adaptation period essential for the algorithm to learn the necessary corrections, a phenomenon also observed in the findings of Exp1.

The subjective evaluation carried out by participants with limb loss on the PC SMAP control yielded scores similar to those of the PC and Nat phases (Nat vs PC vs PC SMAP; n = 6; median SUS score of 97.5 vs 90 vs 88.75; median raxProsTLX score of 2.53 vs 4 vs 5). This concordance of scores effectively validates the conclusions drawn from the Exp1 results, indicating that the SMAP algorithm exerts minimal influence on the virtual arm's behavior. Consequently, the natural reaching movements remain largely unaffected by the integration of the SMAP control, solidifying its efficacy in preserving the same arm behavior observed in the PC phases.

Illustrative Video 3 in Appendix B illustrates a participant with transhumeral limb loss using PC ANN with SMAP. The orange and green virtual cylinders, not visible to the participant, represent the object transformed from the camera to the armroot reference frame without and with the correction applied by the SMAP. The green cylinder, therefore, acts as a dynamic representation of the convergence status of the SMAP. As the algorithm learns the proper corrections, its position is progressively adjusted toward the actual object position. In the second part of the video, leveraging its past learning, the algorithm applies the corrections it has already acquired, resulting in the green cylinder being already positioned closer to the actual object when a new object appears.

## 4.3 Discussion

In this chapter, I outlined two experiments conducted in VR involving participants with intact limbs and with transhumeral limb loss. These experiments were designed to evaluate the efficacy of the methods outlined in Chapter 3 in computing the head position in the armroot reference frame from minimal information that might be available in real-life settings. The performances of their respective control systems were compared with that of the ideal scenario, wherein all the information necessary to transform the target information from the head to the armroot frame is available to the system.

The first approach, A2H, proved ineffective in resolving this issue: the ANN, trained offline on a dataset of natural movements from multiple participants, failed to encapsulate sufficiently well the specificity of the user's morphology, resulting in a mean

prediction error of 2cm which was found to affect the prosthesis control. The second approach, SMAP, elicited much better results, as it effectively spatially encoded the transition from the head to the armroot associated with different targets in space. This bio-inspired spatial coding (Chinellato et al., 2011; Couraud et al., 2018; Hoffmann et al., 2010; Makin et al., 2013; Pouget et al., 2002; Salinas and Abbott, 2001; Serino, 2019; Snyder et al., 1998) enabled to capture online features specific to the user's morphology, as well as to the particular head, body, and arm postures employed when reaching various targets in space. Although the learning process required an adaptation period, subsequent performances on already explored targets were comparable to the ideal scenario, as observed from twelve participants with intact limbs testing in a condition of simulated amputation in Exp1. While the combination of both methods yields favorable results, the second approach alone suffices for the task, as shown on six participants with transhumeral limb loss tested in Exp2, eliminating the need for unnecessary complexity in the control system.

In addition, participants evaluated the proposed methods through two surveys assessing usability and cognitive load, respectively. The outcomes revealed a correlation between subjective and objective metrics. Notably, when employing the control system incorporating the A2H ANN, participants required additional physical and mental effort to complete the task. In contrast, the control systems integrating the SMAP had no discernible impact on virtual arm control, leading to natural reaching movements and lower cognitive load scores.

In this chapter, we employed the SMAP to perform an online "calibration" of the distance between the user's head and armroot reference frames (i.e., the head position in the armroot reference frame). Its value is determined by computing the difference between the hand pose acquired simultaneously in the armroot reference frame through forward kinematics (i.e., artificial proprioception) and in the head reference frame through computer vision (i.e., artificial vision). This difference is then stored in a map representing the peripersonal space of the user and is leveraged to translate target information from the head to the armroot reference frame, needed to deploy the control methodology proposed by E. Segas et al., 2023 in real-world scenarios.

Our approach of utilizing the hand pose to calculate the distance between the head and the armroot reference frame can be readily likened to self-observation methods employed for self-calibrating the parameters of the kinematic chain in robotics (Hersch et al., 2008; Hoffmann et al., 2010; Nabeshima et al., 2006; Sturm et al., 2009). Through self-observation, the robots learn the offsets in position and orientation between the different reference frames of their chain. Here, we refrain from defining a chain, as outlined in Chapter 1.2, due to the inherent complexity in modeling the intricate region between the head and the armroot. Furthermore, it is important to note that the offset learned by the algorithm is neither between two consecutive reference frames nor constant. Indeed, the head position in the armroot reference frame is influenced by the user's morphology and by the specific posture of the user's arm, head, and upper body produced while reaching targets in space. To account for this dynamic aspect, we chose to encode it within a map of the peripersonal space, accessed based on the target the user gazes at and endeavors to reach and grasp.

The algorithm holds the potential for handling more complex scenarios involving simultaneous errors and corrections in both position and orientation, as could occur when moving a camera on glasses or operating in perturbed environments. However, the results of these experiments alone do not demonstrate that the SMAP could be employed in real-life settings where the hand must be detected by computer vision. Therefore, we conducted a physical proof of concept in a simple real-world scenario, which will be presented in the next chapter.

# Chapter 5

# Proof of concept: from virtual reality to a robotic platform

While Exp1 and Exp2 assessed conditions simulating reduced information "as if" obtained from computer vision, they do not provide a basis for evaluating the feasibility of employing SMAP in a real-world setting. VR is an ideal environment where the system can consistently detect both the hand and the target and ensure flawless execution of commands for the virtual arm. However, the algorithm's functionality in conditions closer to real-world scenarios has yet to be established.

Due to the lack of commercially available prostheses in which it could be possible to test our methods, we conducted a proof of concept to demonstrate the possibility of transferring the SMAP from VR to a real-world environment while controlling the robotic platform REACHY 2. Designed by the company Pollen Robotics, this bio-inspired robot ranked 2nd at ANA Avatar XPRIZE, an international competition challenging teams to build a telerobotic avatar system to navigate and manipulate objects in a remote environment (Hauser et al., 2024).

Although determining the head position in the armroot referential is straightforward with a robot as its kinematic model is known, we aimed to verify that the SMAP is effective in this context whereby the hand position and orientation needed for the algorithm genuinely come in the head referential from computer vision. Selecting a robot as our experimental subject allows us to capitalize on the solvability of the problem but expose the algorithm to a scenario that mirrors the complexity of real-world applications. By focusing on retrieving the position of the robot's head in the armroot referential and learning its morphology, the involvement of multiple participants becomes unnecessary. Hence, we opted for a single participant with prior experience controlling the robotic platform. This approach allows us to concentrate on assessing the algorithm's applicability to the robotic system without unnecessary complexity introduced by involving multiple operators. It ensures that any observed outcomes are directly attributable to the algorithm's interaction with the environment, enhancing the clarity and validity of our evaluation.

The work presented in this chapter is part of the following publication (Lento, Leconte, et al., 2024), IEEE Robotics and Automation Letters.

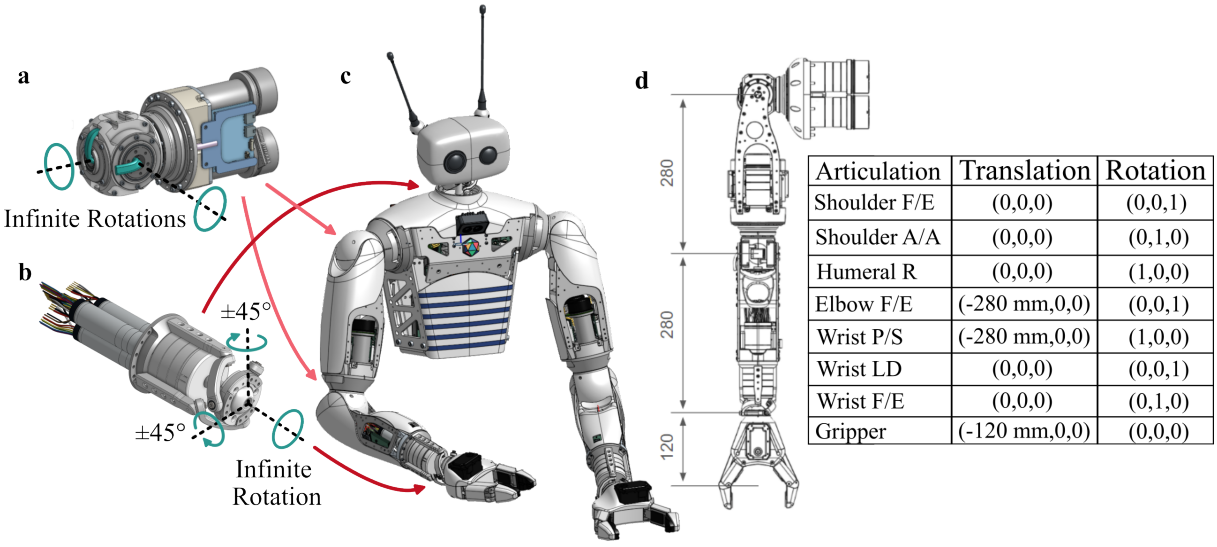
## 5.1 Materials and Methods

This section outlines all the elements put in place to conduct the proof of concept presented in this chapter.

### 5.1.1 Robotic platform

Developed by the company Pollen Robotics, the robotic platform REACHY 2 consists of a central body, a head, and two arms - right and left - with dimensions comparable to those of a human arm. Each arm possesses 7 DOFs corresponding to shoulder flexion/extension and adduction/abduction, humeral rotation, elbow flexion/extension, wrist pronation/supination, lateral deviation, and flexion/extension. It is terminated by an eighth DOF, the gripper or end effector, allowing it to grasp objects with a maximum span of 10cm.

The wrist and neck joints are actuated by a mechanism known as Orbita (Lapeyre et al., n.d.), illustrated in Figure 5.1b, which consists of three parallel brushless motors forming a perfect ball joint. It enables the robotic head to move in 3D with a high degree of mobility, mirroring the user's head movements and offering a more realistic egocentric view compared to traditional 2D (pan/tilt) robotic heads. The shoulder and elbow joints are actuated by a 2-DOF version of Orbita (Figure 5.1a), ensuring a tight correspondence between robotic and human arm movements.




| Articulation | Translation | Rotation |
|---|---|---|
| Shoulder F/E | (0,0,0) | (0,0,1) |
| Shoulder A/A | (0,0,0) | (0,1,0) |
| Humeral R | (0,0,0) | (1,0,0) |
| Elbow F/E | (-280 mm,0,0) | (0,0,1) |
| Wrist P/S | (-280 mm,0,0) | (1,0,0) |
| Wrist LD | (0,0,0) | (0,0,1) |
| Wrist F/E | (0,0,0) | (0,1,0) |
| Gripper | (-120 mm,0,0) | (0,0,0) |

Figure 5.1: Overview of the robotic platform REACHY 2. **a.** 2-DOF model of Orbita for shoulder and elbow joints. **b.** 3-DOF model of Orbita for wrist and neck joints. **c.** 3D model of REACHY 2. **d.** Geometric representation of the right arm with 7 DOFs and gripper, accompanied by a descriptive table. Figure adapted from the Pollen Robotics GitHub.

Illustrative Video 4 in Appendix B demonstrates the teleoperation of the head and right

arm of REACHY 2.

The communication with the robotic platform, enabling control over the actuators of the arms and the head, is facilitated with custom electronics and firmware written in Rust. The software stack is built around ROS 2 with an additional abstraction layer in the form of a Python SDK.

### 5.1.2 Experimental setup

The participant (intact-limbs, male, aged 26, EHI 100) sat on a stool beside the robot, which was fixed on a table. He wore the headset and four motion trackers (trunk, right upper arm, forearm, and hand). The experiment began with the same calibration procedures used in the VR experiments to link the participant's arm to a virtual reality arm, which would then be used to control the robotic arm. The task was to control the robot to reach and grasp a block of five polyethylene foam cylinders (with dimensions r = 1.8cm, h = 11cm) placed on a wooden board in various positions and orientations, as shown in Figure 5.2. The control mechanism employed was teleoperation, where the participant's arm and neck joint configurations were directly applied to manipulate the corresponding joints of the robot.

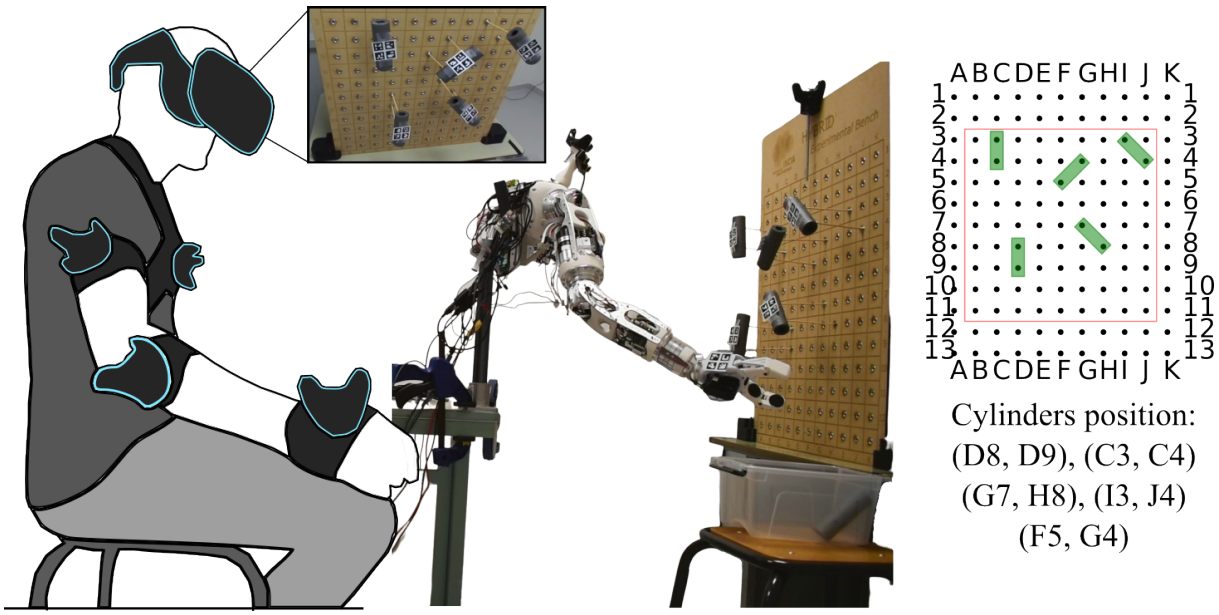



Figure 5.2: Proof of concept: experimental setup and placement of the five cylinders on the board.

We mounted a Zed Mini camera on the head of the robot and linked it to the VR headset. This setup allowed the participant wearing the headset to look through the Zed Mini camera and control the robot from an egocentric point of view.

ArUco boards, made of four ArUco markers each, were placed on the end effector (frontal and lateral sides) and the cylinders to detect their position and orientation. We

implemented a simple computer vision algorithm that used the point of gaze to discriminate which cylinder the participant was looking at and wanted to grasp. The choice behind opting for ArUco markers arises from the advantageous feature of enabling real-time and highly accurate object pose estimation.

### 5.1.3 Protocol

The participant teleoperated the robotic platform to reach and grasp the block of five cylinders. Visual feedback, presented as virtual cylinders, was provided to the participant, as shown in Figure 5.3. The red cylinder, representing the object detected by computer vision in the camera reference frame, was transformed in the armroot referential without and with the correction applied by the SMAP (the orange and green cylinders, respectively). The green cylinder, therefore, acted as a dynamic representation of the convergence status of the SMAP. As the algorithm iteratively refined its weights and learned the proper corrections to apply, this cylinder dynamically adapted its position to converge toward the actual location of the target object.

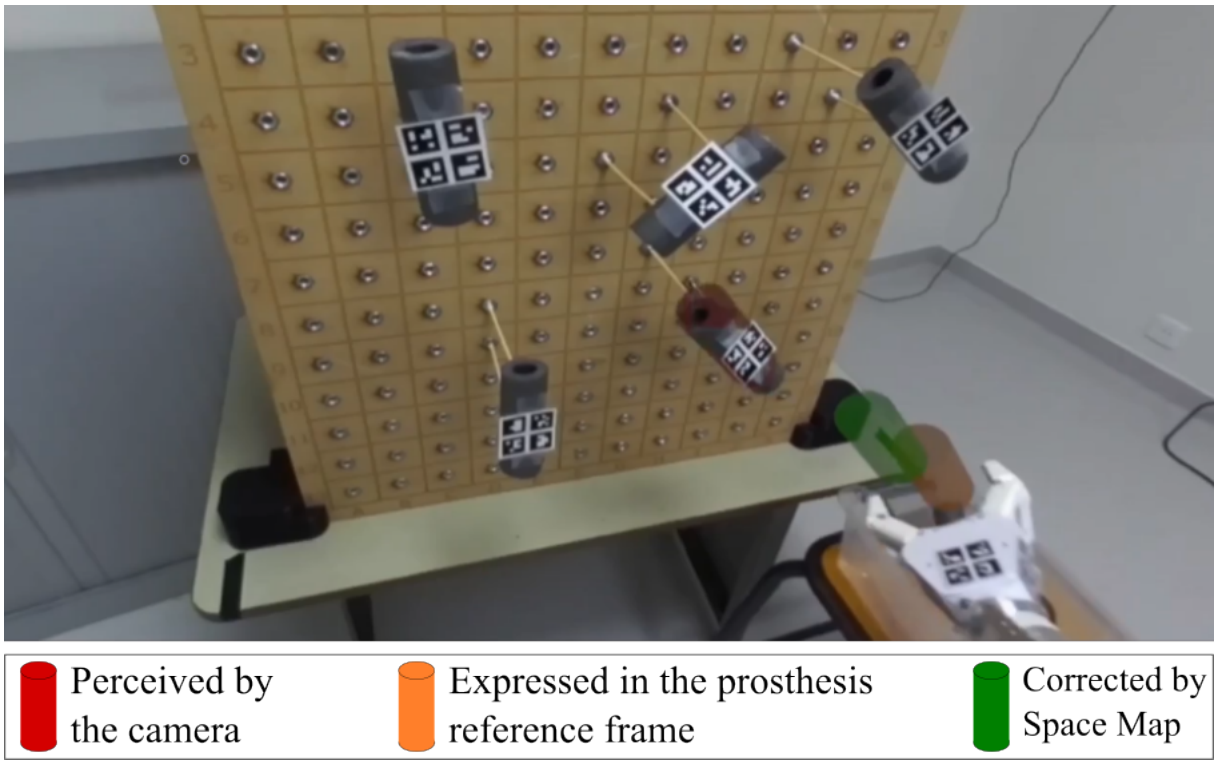



Figure 5.3: Egocentric view of the robotic platform REACHY 2. The participant controls the robotic platform REACHY 2 through teleoperation, operating from an egocentric point of view. The red virtual cylinder represents the object detected by computer vision in the camera reference frame, which is transformed in the armroot referential without and with the correction applied by the SMAP (the orange and green virtual cylinders, respectively).

The protocol included two phases: reaching at a natural pace (Nat Pace) and Grasping on Convergence (GoC), with the latter being recorded in Illustrative Video 5 in Appendix B. In Nat Pace, the participant performed the task at his natural pace. This phase aimed to determine the movement time required to naturally complete the task and assess whether the SMAP could assimilate the entire error within that timeframe, as observed in the VR experiments. For instance, we were uncertain about the algorithm's convergence time in this new scenario. Unlike in VR, where the system reliably receives data at 90Hz, in this setting, we were unsure about the rate at which the ArUco markers would be detected, particularly those positioned on the end effector. As a result, determining the frequency at which the algorithm's weights would be updated remained unclear, potentially leading to convergence times longer than anticipated. To address this uncertainty, we designed a second phase, GoC, in which the participant executed the block twice (1B and 2B), grasping the object only once the algorithm had converged (i.e. when he saw the virtual green cylinder on the object). By explicitly instructing the participant to grasp the targets only once the SMAP had convergence, we obtained the necessary data to address our uncertainty about convergence times.

### 5.1.4 Metrics

This section introduces the metrics used to assess the performances of the SMAP. No filtering process was applied for this study. The data were analyzed in terms of:

- movement time, calculated as the time elapsed between the beginning of the trial and the closing of the end effector on the target,
- algorithm's loss, defined as the difference between the actual value of the head position in the armroot referential and the output of the SMAP,
- algorithm's convergence time as the time elapsed between the beginning of the trial and the moment in which the loss of the algorithm was within the positional tolerance.

The phase GoC was divided into two blocks (1B and 2B), with each block representing a repetition of the target set. Metrics were computed separately for each block.

## 5.2 Results

Figure 5.4 illustrates the algorithm's loss computed for each target (cylinder to grasp) across different block repetitions. It demonstrates the SMAP's effective operation in real-world scenarios. Notably, it underwent an initial adaptation period, as exhibited in phase GoC 1B, which aligns with findings from the VR experiments. However, once the algorithm assimilated and memorized the necessary corrections, its performance improved drastically, as observed by much lower loss early on in phase GoC 2B (*cf* blue traces in Figure 5.4).

As previously discussed, we were uncertain whether the algorithm would converge in a

single trial, as observed in the VR experiments. Indeed, within the Nat Pace phase, the algorithm did not converge in a single trial, assimilating only 71% of the error (i.e., the robot's head position in the armroot reference frame) on average. The data analysis from phase GoC 1B shows that the algorithm required a median time of 5 seconds to converge and achieve a satisfactory solution. This is longer than the median task execution time of 2.75 seconds observed in Nat Pace, suggesting that the algorithm did not have enough time to fully converge. As expected, the algorithm's convergence times and the rate of the weights update were slower in this new scenario compared to VR experiments, primarily due to the utilization of the computer vision algorithm to detect the end effector. Fine-tuning the algorithm parameters could enhance its performance and ensure optimal adaptability to this application, thereby achieving convergence times closer to actual movement times.

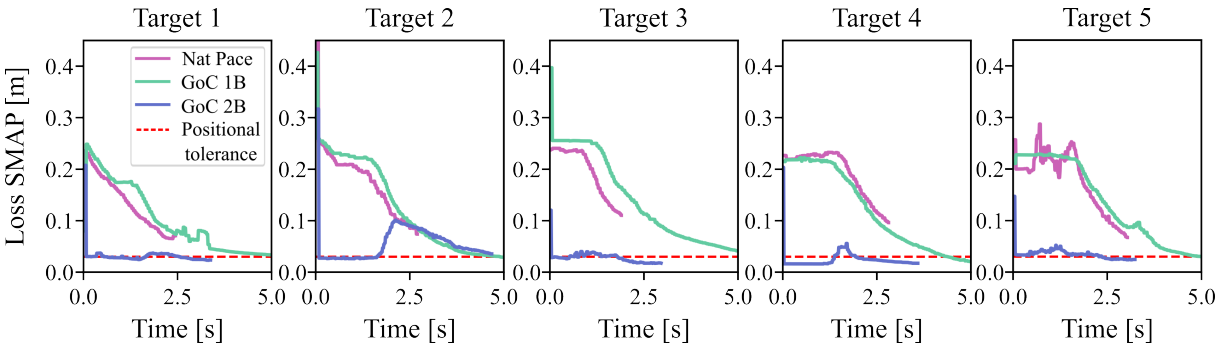



Figure 5.4: SMAP algorithm's loss. Each subplot represents a target, each color a different block repetition. The red dotted line represents the positional tolerance used in the VR experiments.

## 5.3 Discussion

In this chapter, I presented a proof of concept that allowed us to test the SMAP for deployment in practical, real-world applications. The experiment aimed to assess the algorithm's adaptability and feasibility in real-life settings in which the information about the target and the end effector was obtained through a camera and gaze-guided computer vision algorithm. The outcomes of this proof of concept highlighted the algorithm's effectiveness, showcasing its successful operation not only within a controlled VR environment but also in a simple real-world scenario.

VR is an ideal and simulated environment where users have full knowledge and control over the experimental conditions, such as the rate at which the system receives data. This consistency enabled us to conduct a tuning process for the algorithm parameters and meet specific requirements, such as achieving convergence times shorter than movement times. In our real-world implementation, detailed here, information regarding the target and the end effector, such as their position and orientation, was captured by a camera utilizing ArUco markers in conjunction with a gaze-guided computer vision

algorithm. Two challenges arose: firstly, uncertainty regarding the frequency at which the system received data, and secondly, potential difficulties in detecting the end effector when it was located outside the field of vision or obstructed from view. Indeed, experimental results revealed that the algorithm exhibited longer convergence times compared to the time required by the user to complete the task. Despite these potential complications, the algorithm incorporates a memory feature that has proven effective in addressing such situations. When reaching already explored targets, the algorithm leverages previously acquired knowledge, leading to rapid convergence. This memory feature contributes significantly to the algorithm's robustness, ensuring its ability to handle complex and dynamic scenarios.

# Chapter 6

# Conclusions and perspectives

This chapter revisits the objectives of this thesis, summarizes the findings and results obtained, and presents the conclusions drawn from the research. It lays the groundwork for future investigations, highlighting the relevance of the present work.

## 6.1 Conclusion

The objective of this thesis was to adapt the prosthesis control previously developed by my colleagues to real-world settings. Building on the progress in movement-based prosthesis controls (Iftime et al., 2005; Kaliki, Davoodi, and Loeb, 2008; Kaliki et al., 2013; Legrand et al., 2022; Merad et al., 2019, 2020; Merad et al., 2016; D. Popovic and Popovic, 1998; M. Popovic and Popovic, 2001) and computer vision algorithms (DeGol et al., 2016; Došen and Popović, 2011; González-Díaz et al., 2019; Markovic et al., 2015; Pérez De San Roman et al., 2017), my colleagues demonstrated that incorporating information about movement goals with proximal joint motion enables an ANN to precisely predict the distal joint configuration required for positioning and orienting the hand to grasp the targeted object (Mick et al., 2021; E. Segas et al., 2023). The ANN involved, named Proximo Contextual (PC ANN), trained on data from multiple participants performing the task, leverages natural arm coordination to predict the distal joint configuration. This results in a more natural and intuitive control of the prosthetic limb, enabling individuals with transhumeral limb loss to perform pick-and-place tasks in VR as well as with their natural arm (E. Segas et al., 2023).

While the PC ANN control system represents a significant advancement in control strategies for transhumeral prostheses, its implementation in real-world scenarios still presents challenges. VR provides an ideal environment for experimentation in which all necessary information is readily available. In real-world scenarios, however, users may wear a head-mounted camera and IMUs attached to the head, the trunk, and the residual limb. In such conditions, information about movement goals (i.e., one of the control system inputs) may only be available in a head-centered reference frame through gaze-guided computer vision. This data requires transformation into the armroot reference frame for deployment by PC ANN. Yet, limitations may arise in performing this transformation as we might have access only to orientation data through IMUs, lacking positional data.

The primary objective of this thesis was to address the reference frame transformation issue by devising methodologies to accurately determine the head's position within the armroot reference frame.

In the field of robotics, coordinate transformation problems, such as retrieving data about a specific body part, often involve the utilization of kinematic models. However, building a kinematic chain from the head to the shoulder poses a significant challenge due to the intricate anatomy and dynamic interactions of these body regions, as illustrated in Figure 1.8 (Krishnan et al., 2019; Standring, 2015). As a result, existing kinematic models frequently opt for simplifications within the chain leading to a partial representation of the movements inherent in this region. Even if we were to develop a precise kinematic model, challenges would arise in its implementation due to the necessity of fine-tuning the model parameters for each participant and accurately measuring the movements of these body parts. This would necessitate additional sensors, each meticulously calibrated to capture even the slightest motion. Given those difficulties associated with the complexity of the relationship between the two frames,

we preferred to approach the problem as a black box, learning its input-output behavior through a system identification process (Chinellato et al., 2011; De Barreto, 2003; Martinetz et al., 1990; Schillaci et al., 2014).

The first crucial step was to collect data that accurately reflects the relationship between the head and armroot reference frames. In Chapter 2, I presented the experimental protocol we designed to collect a dataset of twenty intact-limbs participants performing pick-and-place tasks in VR. The workspace in which participants executed the task was maximized through a systematic procedure. Initially, the widest possible workspace was determined by the maximal range of motion of the participants. Subsequently, a self-organizing map (Fritzke, 1995) was employed to delineate the actual space covered by participants, capturing the natural arm movements within this refined space (E. Segas et al., 2023). Concurrently, precise control over the initial body posture, from which arm movements originated, was achieved by guiding participants to a predetermined neutral posture via visual feedback from the trunk and shoulders within a VR environment.

As the dataset includes head and gaze motion, along with movements of the trunk, shoulders, and arm joints, it could represent a valuable resource across diverse domains such as sensorimotor control, movement-based prosthesis control, humanoid robotics, human-robot interactions, and foremost gaze-guided computer vision. The DataPlayer, the Unity project we developed to replay/visualize the data in 3D, enables the generation of new video data with various fictitious visual backgrounds and various objects. This last capability is particularly attractive to computer vision researchers aiming to design and test efficient 6D pose estimation algorithms from egocentric vision with gaze data.

The dataset served as the basis for developing and testing two methods to address the reference frame transformation issue and determine the head position within the armroot reference frame. Chapter 3 describes extensively these proposed methods. Unlike model-based strategies, which would necessitate constructing a model from the head to the armroot, our approaches were model-free, treating the system as a black box and learning the input-output relationship. The first method, Armroot2Head (A2H), involved training offline an ANN on the dataset to do this transformation. The ANN predicts the head position in the armroot reference frame based on ongoing shoulder and head orientations and the participant's height. The second method, Space Map (SMAP), spatially encoded the head position in the armroot reference frame within a bio-inspired peripersonal space map. This approach involves computing the head position in the armroot reference frame by comparing the hand positions captured simultaneously in both reference frames. Specifically, forward kinematics techniques determine the hand position in the armroot reference frame, while computer vision algorithms retrieve the hand position in the head reference frame. As the value of the head position depends not only on the user's morphology but also on the specific posture of their arm, head, and upper body while reaching targets in space, we opted for storing it in a peripersonal space map. Inspired by the encoding of peripersonal space in the brain (Chinellato et al., 2011; Di Pellegrino and Làdavas, 2015; Serino, 2019) as well as by previous research conducted by our team on human adaptation strategies

(Couraud et al., 2018), this spatial encoding was achieved through a single-layer network of spatially tuned neurons and accessed based on the target the user gazes at and aims to reach and grasp.

Chapter 4 outlines two experiments conducted to evaluate the efficacy of the proposed methods while intact-limbs participants and individuals with limb loss controlled a prosthesis avatar in VR. Experimental results on twelve intact-limbs participants demonstrated persistent errors with the first method, A2H ANN, leading to extended compensatory movements and prolonged movement times. In contrast, the second approach, SMAP, exhibited more positive outcomes. While the learning process required an initial adaptation period, its subsequent performance on already explored targets proved comparable to the ideal scenario, showcasing the effectiveness of this approach. Combining both methods also yielded favorable results; however, the second approach alone sufficed for the task, as shown on six participants with transhumeral limb loss, eliminating unnecessary complexity in the control system.

Individuals with transhumeral limb loss effectively controlled the virtual arm, achieving success rates exceeding 95% while minimizing compensatory movements and maintaining movement times comparable to the ideal scenario. Additionally, they reported high satisfaction with the control system, expressing a keen interest in utilizing such natural control methods with real prostheses.

We demonstrated the effectiveness of SMAP in determining head position in the armroot reference frame not only within a VR environment but also in a simple real-world scenario, whereby the hand position and orientation needed for the algorithm genuinely come in the head referential from computer vision. Due to the lack of commercially available prostheses for testing our methods, we conducted a proof of concept, described in Chapter 5, teleoperating the robotic platform REACHY 2 from an egocentric point of view. The targets and the robot's end effector were detected using ArUco markers and a gaze-guided computer vision algorithm. While a single participant with prior experience controlled the robot to perform pick-and-place tasks, the SMAP was tasked to compute the position of the robot's head in its armroot reference frame. The outcomes of this proof of concept highlighted the algorithm's effectiveness, showcasing its successful operation in a real-world scenario.

In this thesis, we employed the SMAP to compute the head position in the armroot reference frame and complete the reference frame transformation from the head to the armroot. The findings presented here not only contribute to the advancement of prosthesis control mechanisms but also offer insights into broader applications of the SMAP across various scenarios involving reference frame transformations, online adaptation, and error corrections. For instance, the algorithm can effectively handle reference frame transformations from a camera mounted on glasses to the armroot or address challenges such as camera misplacement or incorrect target detection, thereby allowing corrections in both position and orientation.

## 6.2 Future directions

In this thesis, we have made significant steps toward deploying the prosthesis control proposed by E. Segas et al., 2023 in real-world contexts. Our primary focus has been the development of an algorithm to determine the head's position within the armroot reference frame. This step is essential for converting target data from head-centric to armroot-centric coordinates. Looking ahead, two additional challenges emerge as we seek to deploy this control system in real-life contexts. First, we must integrate a robust computer vision algorithm into the system to retrieve information about the movement goal. Second, we need a suitable prosthetic device.

Grasping actions are typically preceded by visual fixations on the object of interest (Goodale, 1998), making eye-tracking a natural fit for our prosthesis control system (Corbett et al., 2012). However, achieving real-time estimation of both the position and orientation of objects that users are fixated on and intend to grasp is still an open research issue (Rogez et al., 2015). Experiments conducted by González-Díaz et al., 2019 with healthy subjects in real-life scenarios revealed that gaze information is noisy due to micro-saccades and distractors and that object recognition and localization require heavy preprocessing. To address this, recent computer vision models have incorporated two components: an algorithm for real-time object pose estimation, which primarily relies on a 3D object model (Ammirato et al., 2020; Calandre et al., 2021; Labbé et al., 2022; Y. Li et al., 2018; Liu, 2023; Nguyen et al., 2022; Xiang et al., 2017), and an object detector such as Yolov8 ("YOLOv8", 2022) or a filter based on gaze fixation coordinates (Bertinetto et al., 2016; González-Díaz et al., 2019; Mayer et al., 2022) to perform object recognition and optimize the object localization network.

To deploy our control system in a real-life scenario, the primary objective will be to estimate the 6-DOF pose of the object of interest from the RGB video stream relative to a reference frame attached to the glasses worn by the user. The dataset we collected in Chapter 2 serves as a valuable starting point. Notably, the DataPlayer tool developed for data replay features a functionality enabling the generation of synthetic data based on the recorded data. While preserving the user's movements, gaze data, and target object poses, this tool, being a Unity project, offers extensive customization options. Users can modify the shape and dimensions of the target object, alter the surrounding environment by adding or removing objects, and adjust other parameters as needed (e.g., lightness). This capability greatly facilitates the exploration and testing of computer vision algorithms, as the ground truth is inherently known by design. It allows for experimentation in diverse and complex scenarios, providing a reliable basis for evaluating algorithm performance and refining methodologies.

First, attention will be directed towards the development of the 6-DOF pose estimation algorithm under controlled conditions with known a priori information using the dataset and the DataPlayer. Next, an assessment of the proposed methodology will be conducted using the real-world corpus Grasping-InThe-Wild ("Grasping in the wild", 2017), which includes first-person videos and gaze maps captured in natural cluttered environments such as various objects in real kitchens. Finally, the method will be adapted

to address more challenging real-life scenarios where limited information is available about the object to be grasped.

Concurrently with the development of such computer vision models, deploying the PC ANN control system in real-life settings requires a prosthetic device with an actuated elbow and actuated wrist, and thus, five active joints. Replicating the complexity of the human wrist joint in artificial and prosthetic designs remains a challenge (Bajaj et al., 2019; Fan et al., 2022). As a result, commercially available wrist prostheses provide only a single active DOF, and, despite ongoing research efforts, devices featuring two active DOF, such as the ToMPAW modular arm (Kyberd et al., 2007), DARPA Modular Prosthetic Limb (“Modular Prosthetic Limb (MPL) v1”, 2022), DEKA "Luke" Arm (“Luke arm”, n.d.), and others (“Atom touch artificial arm”, n.d.; Boccardo et al., 2024; Demofonti et al., 2023; Montagnani et al., 2015a), have yet to be made commercially available for users.

In this thesis, due to the absence of commercially available transhumeral prostheses with five active joints, we conducted our experiments using VR and the robotic platform REACHY 2 as a physical proof of concept. One notable feature of this robot is its 3-DOF joint Orbita (Lapeyre et al., n.d.), which enables movement along all three directions at the neck and wrist level. Looking ahead, our project aims to advance toward the development of a prosthetic device equipped with Orbita at the wrist level for individuals with transradial limb loss. This prosthetic system will be coupled with a control mechanism similar to E. Segas et al., 2023, a camera and gaze-guided computer vision to detect objects of interest, along with the algorithm I developed, the Space Map, for transforming target information from the head to the shoulder reference frame.

Once these developments are completed, testing in real-world scenarios will become feasible. The final step will involve the integration of minimal sensors, such as IMUs positioned on the trunk, the residual limb, and the camera/glasses, into the system. Several studies on prosthesis control have utilized information from IMUs positioned on the residual limb to measure shoulder joint movements (D. A. Bennett and Goldfarb, 2018; Blana et al., 2016; Merad et al., 2019, 2020; Merad et al., 2016). Indeed, the accumulation of errors due to measurement noise and calibration issues can lead to drift in the signal integration process, ultimately reducing the reliability of measurements over time (Zhou and Hu, 2008). To address this challenge, numerous research teams are actively developing algorithms for drift reduction. These methods involve applying constraints, modeling sensor biases, or merging data from several sensors using complementary filters such as Kalman or particle filters (El-Gohary and McNames, 2015; Miller et al., 2004; Peppoloni et al., 2013; Rong Zhu and Zhaoying Zhou, 2004).

# References

Ammirato, P., Tremblay, J., Liu, M.-Y., Berg, A. C., & Fox, D. (2020). SymGAN: Orientation estimation without annotation for symmetric objects. *2020 IEEE Winter Conference on Applications of Computer Vision (WACV)*, 1657–1666. https://doi.org/10.1109/WACV45572.2020.9093450

Andersen, R., Bracewell, R., Barash, S., Gnadt, J., & Fogassi, L. (1990). Eye position effects on visual, memory, and saccade-related activity in areas LIP and 7a of macaque. *The Journal of Neuroscience*, *10*(4), 1176–1196. https://doi.org/10.1523/JNEUROSCI.10-04-01176.1990

Atkeson, C., & Hollerbach, J. (1985). Kinematic features of unrestrained vertical arm movements. *The Journal of Neuroscience*, *5*(9), 2318–2330. https://doi.org/10.1523/JNEUROSCI.05-09-02318.1985

Atkins, D. J., Heard, D. C. Y., & Donovan, W. H. (1996). Epidemiologic overview of individuals with upper-limb loss and their reported research priorities: *JPO Journal of Prosthetics and Orthotics*, *8*(1), 2–11. https://doi.org/10.1097/00008526-199600810-00003

Atom touch artificial arm. (n.d.). https://www.atomlimbs.com/

Averta, G., Barontini, F., Catrambone, V., Haddadin, S., Handjaras, G., Held, J. P. O., Hu, T., Jakubowitz, E., Kanzler, C. M., Kühn, J., Lambercy, O., Leo, A., Obermeier, A., Ricciardi, E., Schwarz, A., Valenza, G., Bicchi, A., & Bianchi, M. (2021). U-limb: A multi-modal, multi-center database on arm motion control in healthy and post-stroke conditions. *GigaScience*, *10*(6), giab043. https://doi.org/10.1093/gigascience/giab043

Bailakanavar, M., Liu, Y., Fish, J., & Zheng, Y. (2014). Automated modeling of random inclusion composites. *Engineering with Computers*, *30*(4), 609–625. https://doi.org/10.1007/s00366-012-0310-x

Bajaj, N. M., Spiers, A. J., & Dollar, A. M. (2019). State of the art in artificial wrists: A review of prosthetic and robotic wrist design. *IEEE Transactions on Robotics*, *35*(1), 261–277. https://doi.org/10.1109/TRO.2018.2865890

Bangor, A. (2009). Determining what individual SUS scores mean: Adding an adjective rating scale. *4*(3), 10.

Barker, S., Fuente, L. A., Hayatleh, K., Fellows, N., Steil, J. J., & Crook, N. T. (2015). Design of a biologically inspired humanoid neck. *2015 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, 25–30. https://doi.org/10.1109/ROBIO.2015.7407034

Batista, A. P., Buneo, C. A., Snyder, L. H., & Andersen, R. A. (1999). Reach plans in eye-centered coordinates. *Science*, *285*(5425), 257–260. https://doi.org/10.1126/science.285.5425.257

Beira, R., Lopes, M., Praca, M., Santos-Victor, J., Bernardino, A., Metta, G., Becchi, F., & Saltaren, R. (2006). Design of the robot-cub (iCub) head. *Proceedings 2006 IEEE International Conference on Robotics and Automation, 2006. ICRA 2006.*, 94–100. https://doi.org/10.1109/ROBOT.2006.1641167

Bennett, D. A., & Goldfarb, M. (2018). IMU-based wrist rotation control of a transradial myoelectric prosthesis. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *26*(2), 419–427. https://doi.org/10.1109/TNSRE.2017.2682642

Bennett, D. J., Geiger, D., & Hollerbach, J. M. (1991). Autonomous robot calibration for hand-eye coordination. *The International Journal of Robotics Research*, *10*(5), 550–559. https://doi.org/10.1177/027836499101000510

Bertinetto, L., Valmadre, J., Henriques, J. F., Vedaldi, A., & Torr, P. H. S. (2016). Fully-convolutional siamese networks for object tracking [Publisher: [object Object] Version Number: 3]. https://doi.org/10.48550/ARXIV.1606.09549

Biddiss, E. A., & Chau, T. T. (2007). Upper limb prosthesis use and abandonment: A survey of the last 25 years. *Prosthetics & Orthotics International*, *31*(3), 236–257. https://doi.org/10.1080/03093640600994581

Blana, D., Kyriacou, T., Lambrecht, J. M., & Chadwick, E. K. (2016). Feasibility of using combined EMG and kinematic signals for prosthesis control: A simulation study using a virtual reality environment. *Journal of Electromyography and Kinesiology*, *29*, 21–27. https://doi.org/10.1016/j.jelekin.2015.06.010

Blohm, G., & Crawford, J. D. (2009). Fields of gain in the brain. *Neuron*, *64*(5), 598–600. https://doi.org/10.1016/j.neuron.2009.11.022

Boccardo, N., Canepa, M., Stedman, S., Lombardi, L., Marinelli, A., Domenico, D. D., Galviati, R., Gruppioni, E., Michieli, L. D., & Laffranchi, M. (2024). Development of a 2-DoFs actuated wrist for enhancing the dexterity of myoelectric hands. *IEEE Transactions on Medical Robotics and Bionics*, *6*(1), 257–270. https://doi.org/10.1109/TMRB.2023.3336993

Brozzoli, C., Gentile, G., & Ehrsson, H. H. (2012). That's near my hand! parietal and premotor coding of hand-centered space contributes to localization and self-attribution of the hand. *The Journal of Neuroscience*, *32*(42), 14573–14582. https://doi.org/10.1523/JNEUROSCI.2660-12.2012

Bruyant, A., Guemann, M., & Malgoyre, A. (2023). Étude épidémiologique des amputations majeures des membres supérieur et inférieur en France. *Kinésithérapie, la Revue*, *23*(256), 3–12. https://doi.org/10.1016/j.kine.2022.05.014

Calandre, J., Peteri, R., Mascarilla, L., & Tremblais, B. (2021). Extraction and analysis of 3d kinematic parameters of table tennis ball from a single camera. *2020 25th International Conference on Pattern Recognition (ICPR)*, 9468–9475. https://doi.org/10.1109/ICPR48806.2021.9412391

Calli, B., Singh, A., Walsman, A., Srinivasa, S., Abbeel, P., & Dollar, A. M. (2015). The YCB object and model set: Towards common benchmarks for manipulation

research. *2015 International Conference on Advanced Robotics (ICAR)*, 510–517. https://doi.org/10.1109/ICAR.2015.7251504

Canzoneri, E., Marzolla, M., Amoresano, A., Verni, G., & Serino, A. (2013). Amputation and prosthesis implantation shape body and peripersonal space representations. *Scientific Reports*, *3*(1), 2844. https://doi.org/10.1038/srep02844

Chinellato, E., Antonelli, M., Grzyb, B. J., & Del Pobil, A. P. (2011). Implicit sensorimotor mapping of the peripersonal space by gazing and reaching. *IEEE Transactions on Autonomous Mental Development*, *3*(1), 43–53. https://doi.org/10.1109/TAMD.2011.2106781

Cléry, J., Guipponi, O., Wardak, C., & Ben Hamed, S. (2015). Neuronal bases of peripersonal and extrapersonal spaces, their plasticity and their dynamics: Knowns and unknowns. *Neuropsychologia*, *70*, 313–326. https://doi.org/10.1016/j.neuropsychologia.2014.10.022

Corbett, E. A., Perreault, E. J., & Körding, K. P. (2012). Decoding with limited neural data: A mixture of time-warped trajectory models for directional reaches. *Journal of Neural Engineering*, *9*(3), 036002. https://doi.org/10.1088/1741-2560/9/3/036002

Cordella, F., Ciancio, A. L., Sacchetti, R., Davalli, A., Cutti, A. G., Guglielmelli, E., & Zollo, L. (2016). Literature review on needs of upper limb prosthesis users. *Frontiers in Neuroscience*, *10*. https://doi.org/10.3389/fnins.2016.00209

Couraud, M., Cattaert, D., Paclet, F., Oudeyer, P. Y., & De Rugy, A. (2018). Model and experiments to optimize co-adaptation in a simplified myoelectric control system. *Journal of Neural Engineering*, *15*(2), 026006. https://doi.org/10.1088/1741-2552/aa87cf

Cruse, H., Brüwer, M., & Dean, J. (1993). Control of three- and four-joint arm movement: Strategies for a manipulator with redundant degrees of freedom. *Journal of Motor Behavior*, *25*(3), 131–139. https://doi.org/10.1080/00222895.1993.9942044

Cutkosky, M. (1989). On grasp choice, grasp models, and the design of hands for manufacturing tasks. *IEEE Transactions on Robotics and Automation*, *5*(3), 269–279. https://doi.org/10.1109/70.34763

De Barreto, G. A. (2003). Self-organizing feature maps for modeling and control of robotic manipulators. *Journal of Intelligent and Robotic Systems*, *36*(4), 407–450. https://doi.org/10.1023/A:1023641801514

De Luca, C. J. (1997). The use of surface electromyography in biomechanics [Publisher: Human Kinetics, Inc.]. *Journal of applied biomechanics*, *13*(2), 135–163.

De Momi, E., Kranendonk, L., Valenti, M., Enayati, N., & Ferrigno, G. (2016). A neural network-based approach for trajectory planning in robot–human handover tasks. *Frontiers in Robotics and AI*, *3*. https://doi.org/10.3389/frobt.2016.00034

De Rugy, A., & Carroll, T. J. (2010). Changes in muscle directional tuning parallel feedforward adaptation to a visuomotor rotation. *Experimental Brain Research*, *203*(4), 701–709. https://doi.org/10.1007/s00221-010-2280-9

DeGol, J., Akhtar, A., Manja, B., & Bretl, T. (2016). Automatic grasp selection using a camera in a hand prosthesis. *2016 38th Annual International Conference of*

*the IEEE Engineering in Medicine and Biology Society (EMBC)*, 431–434. https://doi.org/10.1109/EMBC.2016.7590732

Demofonti, A., Carpino, G., Tagliamonte, N. L., Baldini, G., Bramato, L., & Zollo, L. (2023). Design of a modular and compliant wrist module for upper limb prosthetics. *The Anatomical Record*, *306*(4), 764–776. https://doi.org/10.1002/ar.24911

Denavit, J., & Hartenberg, R. S. (1955). A kinematic notation for lower-pair mechanisms based on matrices. *Journal of Applied Mechanics*, *22*(2), 215–221. https://doi.org/10.1115/1.4011045

Desmurget, M., Prablanc, C., Rossetti, Y., Arzi, M., Paulignan, Y., Urquizar, C., & Mignot, J. C. (1995). Postural and synergic control for three-dimensional movements of reaching and grasping. *Journal of Neurophysiology*, *74*(2), 905–910. https://doi.org/10.1152/jn.1995.74.2.905

Desmurget, M., & Prablanc, C. (1997). Postural control of three-dimensional prehension movements. *Journal of Neurophysiology*, *77*(1), 452–464. https://doi.org/10.1152/jn.1997.77.1.452

Di Pellegrino, G., & Làdavas, E. (2015). Peripersonal space in the brain. *Neuropsychologia*, *66*, 126–133. https://doi.org/10.1016/j.neuropsychologia.2014.11.011

Došen, S., Cipriani, C., Kostić, M., Controzzi, M., Carrozza, M. C., & Popović, D. B. (2010). Cognitive vision system for control of dexterous prosthetic hands: Experimental evaluation. *Journal of NeuroEngineering and Rehabilitation*, *7*(1), 42. https://doi.org/10.1186/1743-0003-7-42

Došen, S., & Popović, D. B. (2011). Transradial prosthesis: Artificial vision for control of prehension: ARTIFICIAL VISION FOR CONTROL OF PREHENSION. *Artificial Organs*, *35*(1), 37–48. https://doi.org/10.1111/j.1525-1594.2010.01040.x

Du, G., Wang, K., Lian, S., & Zhao, K. (2019). Vision-based robotic grasping from object localization, object pose estimation to grasp estimation for parallel grippers: A review [Publisher: arXiv Version Number: 4]. https://doi.org/10.48550/ARXIV.1905.06658

Duhamel, J.-R., Colby, C. L., & Goldberg, M. E. (1998). Ventral intraparietal area of the macaque: Congruent visual and somatic response properties. *Journal of Neurophysiology*, *79*(1), 126–136. https://doi.org/10.1152/jn.1998.79.1.126

El-Gohary, M., & McNames, J. (2015). Human joint angle estimation with inertial sensors and validation with a robot arm. *IEEE Transactions on Biomedical Engineering*, *62*(7), 1759–1767. https://doi.org/10.1109/TBME.2015.2403368

Fan, H., Wei, G., & Ren, L. (2022). Prosthetic and robotic wrists comparing with the intelligently evolved human wrist: A review. *Robotica*, *40*(11), 4169–4191. https://doi.org/10.1017/S0263574722000856

Farina, D., & Sartori, M. (2016, April 22). Surface electromyography for MAN-machine interfacing in rehabilitation technologies. In R. Merletti & D. Farina (Eds.), *Surface electromyography : Physiology, engineering, and applications* (1st ed., pp. 540–560). Wiley. https://doi.org/10.1002/9781119082934.ch20

Farnè, A., Iriki, A., & Làdavas, E. (2005). Shaping multisensory action–space with tools: Evidence from patients with cross-modal extinction. *Neuropsychologia*, *43*(2), 238–248. https://doi.org/10.1016/j.neuropsychologia.2004.11.010

Flash, T., & Hogan, N. (1985). The coordination of arm movements: An experimentally confirmed mathematical model. *The Journal of Neuroscience*, *5*(7), 1688–1703. https://doi.org/10.1523/JNEUROSCI.05-07-01688.1985

Fougner, A., Stavdahl, Ø., Kyberd, P. J., Losier, Y. G., & Parker, P. A. (2012). Control of upper limb prostheses: Terminology and proportional myoelectric control—a review. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *20*(5), 663–677. https://doi.org/10.1109/TNSRE.2012.2196711

Fritzke, B. (1995). A growing neural gas network learns topologies. *Neural Information Processing Systems*, *7*.

Gallagher, P., O'Donovan, M.-A., Doyle, A., & Desmond, D. (2011). Environmental barriers, activity limitations and participation restrictions experienced by people with major limb amputation. *Prosthetics & Orthotics International*, *35*(3), 278–284. https://doi.org/10.1177/0309364611407108

Garner, B. A., & Pandy, M. G. (1999). A kinematic model of the upper limb based on the visible human project (VHP) image dataset. *Computer Methods in Biomechanics and Biomedical Engineering*, *2*(2), 107–124. https://doi.org/10.1080/10255849908907981

Gatla, C. S., Lumia, R., Wood, J., & Starr, G. (2007). An automated method to calibrate industrial robots using a virtual closed kinematic chain. *IEEE Transactions on Robotics*, *23*(6), 1105–1116. https://doi.org/10.1109/TRO.2007.909765

Gloumakov, Y., Bimbo, J., & Dollar, A. M. (2020). Trajectory control for 3 degree-of-freedom wrist prosthesis in virtual reality: A pilot study. *2020 8th IEEE RAS/EMBS International Conference for Biomedical Robotics and Biomechatronics (BioRob)*, 765–772. https://doi.org/10.1109/BioRob49111.2020.9224455

Gloumakov, Y., Bimbo, J., & Dollar, A. M. (2022). Trajectory control–an effective strategy for controlling multi-DOF upper limb prosthetic devices. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *30*, 420–430. https://doi.org/10.1109/TNSRE.2022.3151055

Gloumakov, Y., Spiers, A. J., & Dollar, A. M. (2019). A clustering approach to categorizing 7 degree-of-freedom arm motions during activities of daily living. *2019 International Conference on Robotics and Automation (ICRA)*, 7214–7220. https://doi.org/10.1109/ICRA.2019.8794421

González-Díaz, I., Benois-Pineau, J., Domenger, J.-P., Cattaert, D., & De Rugy, A. (2019). Perceptually-guided deep neural networks for ego-action prediction: Object grasping. *Pattern Recognition*, *88*, 223–235. https://doi.org/10.1016/j.patcog.2018.11.013

Goodale, M. A. (1998). Visuomotor control: Where does vision end and action begin? *Current Biology*, *8*(14), R489–R491. https://doi.org/10.1016/S0960-9822(98)70314-8

Grasping in the wild. (2017). https://www.nakala.fr/collection/11280/d5190c75

Gronier, G., & Baudet, A. (2021). Psychometric evaluation of the f-SUS: Creation and validation of the french version of the system usability scale. *International Journal of Human–Computer Interaction*, *37*(16), 1571–1582. https://doi.org/10.1080/10447318.2021.1898828

Gulletta, G., Erlhagen, W., & Bicho, E. (2020). Human-like arm motion generation: A review. *Robotics*, *9*(4), 102. https://doi.org/10.3390/robotics9040102

Hagiwara, T., Ganesh, G., Sugimoto, M., Inami, M., & Kitazaki, M. (2020). Individuals prioritize the reach straightness and hand jerk of a shared avatar over their own. *iScience*, *23*(12), 101732. https://doi.org/10.1016/j.isci.2020.101732

Hahne, J. M., Biebmann, F., Jiang, N., Rehbaum, H., Farina, D., Meinecke, F. C., Muller, K.-R., & Parra, L. C. (2014). Linear and nonlinear regression techniques for simultaneous and proportional myoelectric control. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *22*(2), 269–279. https://doi.org/10.1109/TNSRE.2014.2305520

Hahne, J. M., Schweisfurth, M. A., Koppe, M., & Farina, D. (2018). Simultaneous control of multiple functions of bionic hand prostheses: Performance and robustness in end users. *Science Robotics*, *3*(19), eaat3630. https://doi.org/10.1126/scirobotics.aat3630

Hargrove, L. J., Miller, L. A., Turner, K., & Kuiken, T. A. (2017). Myoelectric pattern recognition outperforms direct control for transhumeral amputees with targeted muscle reinnervation: A randomized clinical trial. *Scientific Reports*, *7*(1), 13840. https://doi.org/10.1038/s41598-017-14386-w

Hauser, K., Watson, E., Bae, J., Bankston, J., Behnke, S., Borgia, B., Catalano, M. G., Dafarra, S., van Erp, J. B. F., Ferris, T., Fishel, J., Hoffman, G., Ivaldi, S., Kanehiro, F., Kheddar, A., Lannuzel, G., Morie, J. F., Naughton, P., NGuyen, S., … Locke, D. (2024). Analysis and perspectives on the ANA avatar XPRIZE competition [Version Number: 1]. https://doi.org/10.48550/ARXIV.2401.05290

Haute Autorité de Santé, H. (2010). *Évaluation des prothèses externes de membre supérieur* (Guide usagers). Saint-Denis La Plaine. https://www.has-sante.fr/jcms/c_999782/fr/evaluation-des-protheses-externes-de-membre-superieur

Henriques, D. Y. P., Medendorp, W. P., Gielen, C. C. A. M., & Crawford, J. D. (2003). Geometric computations underlying eye-hand coordination: Orientations of the two eyes and the head. *Experimental Brain Research*, *152*(1), 70–78. https://doi.org/10.1007/s00221-003-1523-4

Herman, I. P. (2016). *Physics of the human body*. Springer International Publishing. https://doi.org/10.1007/978-3-319-23932-3

Hersch, M. (2008). Adaptive sensorimotor peripersonal space representation and motor learning for a humanoid robot [Publisher: [object Object]]. https://doi.org/10.5075/EPFL-THESIS-4289

Hersch, M., Sauser, E., & Billard, A. (2008). ONLINE LEARNING OF THE BODY SCHEMA. *International Journal of Humanoid Robotics*, *05*(2), 161–181. https://doi.org/10.1142/S0219843608001376

Hoffmann, M., Marques, H., Arieta, A., Sumioka, H., Lungarella, M., & Pfeifer, R. (2010). Body schema in robotics: A review. *IEEE Transactions on Autonomous Mental Development*, *2*(4), 304–324. https://doi.org/10.1109/TAMD.2010.2086454

Hollerbach, J. M., & Wampler, C. W. (1996). The calibration index and taxonomy for robot kinematic calibration methods. *The International Journal of Robotics Research*, *15*(6), 573–591. https://doi.org/10.1177/027836499601500604

Holmes, N. P., & Spence, C. (2004). The body schema and multisensory representation(s) of peripersonal space. *Cognitive Processing*, *5*(2), 94–105. https://doi.org/10.1007/s10339-004-0013-3

Huang, Y., Bianchi, M., Liarokapis, M., & Sun, Y. (2016). Recent data sets on object manipulation: A survey. *Big Data*, *4*(4), 197–216. https://doi.org/10.1089/big.2016.0042

Hudgins, B., Parker, P., & Scott, R. (1993). A new strategy for multifunction myoelectric control. *IEEE Transactions on Biomedical Engineering*, *40*(1), 82–94. https://doi.org/10.1109/10.204774

Iberall, T. (1988). A neural network for planning hand shapes in human prehension. *1988 American Control Conference*, 2288–2293. https://doi.org/10.23919/ACC.1988.4790106

Iftime, S., Egsgaard, L., & Popovic, M. (2005). Automatic determination of synergies by radial basis function artificial neural networks for the control of a neural prosthesis. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *13*(4), 482–489. https://doi.org/10.1109/TNSRE.2005.858458

Johansen, H., Bathen, T., Andersen, L. Ø., Rand-Hendriksen, S., & Østlie, K. (2018). Education and work participation among adults with congenital unilateral upper limb deficiency in norway: A cross-sectional study (B. A. Philip, Ed.). *PLOS ONE*, *13*(12), e0207846. https://doi.org/10.1371/journal.pone.0207846

Kaliki, R. R., Davoodi, R., & Loeb, G. E. (2013). Evaluation of a noninvasive command scheme for upper-limb prostheses in a virtual reality reach and grasp task. *IEEE Transactions on Biomedical Engineering*, *60*(3), 792–802. https://doi.org/10.1109/TBME.2012.2185494

Kaliki, R. R., Davoodi, R., & Loeb, G. (2008). Prediction of distal arm posture in 3-d space from shoulder movements for control of upper limb prostheses. *Proceedings of the IEEE*, *96*(7), 1217–1225. https://doi.org/10.1109/JPROC.2008.922591

Kaliki, R. R., Davoodi, R., & Loeb, G. E. (2008). PREDICTION OF ELBOW TRAJECTORY FROM SHOULDER ANGLES USING NEURAL NETWORKS. *International Journal of Computational Intelligence and Applications*, *07*(3), 333–349. https://doi.org/10.1142/S1469026808002296

Kanitz, G., Montagnani, F., Controzzi, M., & Cipriani, C. (2018). Compliant prosthetic wrists entail more natural use than stiff wrists during reaching, not (necessarily) during manipulation. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *26*(7), 1407–1413. https://doi.org/10.1109/TNSRE.2018.2847565

Kim, S., Kim, C., & Park, J. (2006). Human-like arm motion generation for humanoid robots using motion capture database. *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 3486–3491. https://doi.org/10.1109/IROS.2006.282591

Koskinopoulou, M., & Trahanias, P. (2016). A methodological framework for robotic reproduction of observed human actions: Formulation using latent space representation. *2016 IEEE-RAS 16th International Conference on Humanoid Robots (Humanoids)*, 565–572. https://doi.org/10.1109/HUMANOIDS.2016.7803331

Krakauer, J. W., Pine, Z. M., Ghilardi, M.-F., & Ghez, C. (2000). Learning of visuomotor transformations for vectorial planning of reaching trajectories. *The Journal of Neuroscience*, *20*(23), 8916–8924. https://doi.org/10.1523/JNEUROSCI.20-23-08916.2000

Krishnan, R., Björsell, N., Gutierrez-Farewik, E. M., & Smith, C. (2019). A survey of human shoulder functional kinematic representations. *Medical & Biological Engineering & Computing*, *57*(2), 339–367. https://doi.org/10.1007/s11517-018-1903-3

Kyberd, P. J., Poulton, A. S., Sandsjo, L., Jonsson, S., Jones, B., & Gow, D. (2007). The ToMPAW modular prosthesis: A platform for research in upper-limb prosthetics: *JPO Journal of Prosthetics and Orthotics*, *19*(1), 15–21. https://doi.org/10.1097/JPO.0b013e31802d46f8

Kyranou, I., Vijayakumar, S., & Erden, M. S. (2018). Causes of performance degradation in non-invasive electromyographic pattern recognition in upper limb prostheses. *Frontiers in Neurorobotics*, *12*, 58. https://doi.org/10.3389/fnbot.2018.00058

Labbé, Y., Carpentier, J., Aubry, M., & Sivic, J. (2020). CosyPose: Consistent multi-view multi-object 6d pose estimation. https://doi.org/10.48550/arXiv.2008.08465

Labbé, Y., Manuelli, L., Mousavian, A., Tyree, S., Birchfield, S., Tremblay, J., Carpentier, J., Aubry, M., Fox, D., & Sivic, J. (2022). MegaPose: 6d pose estimation of novel objects via render & compare [Publisher: arXiv Version Number: 1]. https://doi.org/10.48550/ARXIV.2212.06870

Lallee, S., & Dominey, P. F. (2013). Multi-modal convergence maps: From body schema and self-representation to mental imagery. *Adaptive Behavior*, *21*(4), 274–285. https://doi.org/10.1177/1059712313488423

Lapeyre, M., Rouanet, P., & Crampette, A. (n.d.). *Articulation having three degrees of freedom for a robot, and corresponding control method* (pat. No. 17/774,764).

Latash, M. L., Scholz, J. P., & Schoner, G. (2002). Motor control strategies revealed in the structure of motor variability: *Exercise and Sport Sciences Reviews*, *30*(1), 26–31. https://doi.org/10.1097/00003677-200201000-00006

Legrand, M., Marchand, C., Richer, F., Touillet, A., Martinet, N., Paysant, J., Morel, G., & Jarrasse, N. (2022). Simultaneous control of 2dof upper-limb prosthesis with body compensations-based control: A multiple cases study. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *30*, 1745–1754. https://doi.org/10.1109/TNSRE.2022.3186266

Lento, B., Leconte, V., Bardisbanian, L., Doat, E., Segas, E., & De Rugy, A. (2024). Bioinspired head-to-shoulder reference frame transformation for movement-based arm prosthesis control. *IEEE Robotics and Automation Letters*, *9*(9), 7875–7882. https://doi.org/10.1109/LRA.2024.3433751

Lento, B., Segas, E., Leconte, V., Doat, E., Danion, F., Péteri, R., Benois-Pineau, J., & De Rugy, A. (2024). 3d-ARM-gaze: A public dataset of 3d arm reaching movements with gaze information in virtual reality. *Scientific Data*, *11*(1), 951. https://doi.org/10.1038/s41597-024-03765-4

Lento, B., Segas, E., Leconte, V., Doat, E., Danion, F., Péteri, R., Benois-Pineau, J., & de Rugy, A. (2024, January 25). Data and code for 3d-ARM-gaze: A public

dataset of 3d arm reaching movements with gaze information in virtual reality. https://doi.org/10.5281/ZENODO.10567366

Li, C., Sun, S., Song, X., Song, H., Akhtar, N., & Mian, A. S. (2022). Simultaneous multiple object detection and pose estimation using 3d model infusion with monocular vision [Publisher: arXiv Version Number: 2]. https://doi.org/10.48550/ARXIV.2211.11188

Li, Y., Wang, G., Ji, X., Xiang, Y., & Fox, D. (2018). DeepIM: Deep iterative matching for 6d pose estimation [Publisher: [object Object] Version Number: 4]. https://doi.org/10.48550/ARXIV.1804.00175

Light, C. M., Chappell, P. H., & Kyberd, P. J. (2002). Establishing a standardized clinical assessment tool of pathologic and prosthetic hand function: Normative data, reliability, and validity. *Archives of Physical Medicine and Rehabilitation*, *83*(6), 776–783. https://doi.org/10.1053/apmr.2002.32737

Liu, Y. (2023). Gen6d: Generalizable model-free 6-DoF object pose estimation from RGB images. https://doi.org/10.48550/arXiv.2204.10776

Luke arm. (n.d.). https://mobiusbionics.com/luke-arm/

Major, M. J., Stine, R. L., Heckathorne, C. W., Fatone, S., & Gard, S. A. (2014). Comparison of range-of-motion and variability in upper body movements between transradial prosthesis users and able-bodied controls when executing goal-oriented tasks. *Journal of NeuroEngineering and Rehabilitation*, *11*(1), 132. https://doi.org/10.1186/1743-0003-11-132

Makin, J. G., Fellows, M. R., & Sabes, P. N. (2013). Learning multisensory integration and coordinate transformation via density estimation (G. Blohm, Ed.). *PLoS Computational Biology*, *9*(4), e1003035. https://doi.org/10.1371/journal.pcbi.1003035

Mandery, C., Terlemez, O., Do, M., Vahrenkamp, N., & Asfour, T. (2015). The KIT whole-body human motion database. *2015 International Conference on Advanced Robotics (ICAR)*, 329–336. https://doi.org/10.1109/ICAR.2015.7251476

Maravita, A., & Iriki, A. (2004). Tools for the body (schema). *Trends in Cognitive Sciences*, *8*(2), 79–86. https://doi.org/10.1016/j.tics.2003.12.008

Maravita, A., Spence, C., Kennett, S., & Driver, J. (2002). Tool-use changes multimodal spatial interactions between vision and touch in normal humans. *Cognition*, *83*(2), B25–B34. https://doi.org/10.1016/S0010-0277(02)00003-3

Marinelli, A., Boccardo, N., Tessari, F., Di Domenico, D., Caserta, G., Canepa, M., Gini, G., Barresi, G., Laffranchi, M., De Michieli, L., & Semprini, M. (2023). Active upper limb prostheses: A review on current state and upcoming breakthroughs. *Progress in Biomedical Engineering*, *5*(1), 012001. https://doi.org/10.1088/2516-1091/acac57

Markovic, M., Dosen, S., Cipriani, C., Popovic, D., & Farina, D. (2014). Stereovision and augmented reality for closed-loop control of grasping in hand prostheses. *Journal of Neural Engineering*, *11*(4), 046001. https://doi.org/10.1088/1741-2560/11/4/046001

Markovic, M., Dosen, S., Popovic, D., Graimann, B., & Farina, D. (2015). Sensor fusion and computer vision for context-aware control of a multi degree-of-freedom

prosthesis. *Journal of Neural Engineering*, *12*(6), 066022. https://doi.org/10.1088/1741-2560/12/6/066022

Martinetz, T., Ritter, H., & Schulten, K. (1990). Three-dimensional neural net for learning visuomotor coordination of a robot arm. *IEEE Transactions on Neural Networks*, *1*(1), 131–136. https://doi.org/10.1109/72.80212

Maurel, W., & Thalmann, D. (1999). A case study on human upper limb modelling for dynamic simulation. *Computer Methods in Biomechanics and Biomedical Engineering*, *2*(1), 65–82. https://doi.org/10.1080/10255849908907979

Mayer, C., Danelljan, M., Bhat, G., Paul, M., Paudel, D. P., Yu, F., & Van Gool, L. (2022). Transforming model prediction for tracking. *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 8721–8730. https://doi.org/10.1109/CVPR52688.2022.00853

Mell, A. G., Childress, B. L., & Hughes, R. E. (2005). The effect of wearing a wrist splint on shoulder kinematics during object manipulation. *Archives of Physical Medicine and Rehabilitation*, *86*(8), 1661–1664. https://doi.org/10.1016/j.apmr.2005.02.008

Merad, M., De Montalivet, E., Legrand, M., Mastinu, E., Ortiz-Catalan, M., Touillet, A., Martinet, N., Paysant, J., Roby-Brami, A., & Jarrasse, N. (2020). Assessment of an automatic prosthetic elbow control strategy using residual limb motion for transhumeral amputated individuals with socket or osseointegrated prostheses. *IEEE Transactions on Medical Robotics and Bionics*, *2*(1), 38–49. https://doi.org/10.1109/TMRB.2020.2970065

Merad, M., De Montalivet, E., Legrand, M., Touillet, A., Martinet, N., Paysant, J., Roby-Brami, A., & Jarrassé, N. (2019). Improving the control of prostheses in arm amputees with approaches based on motor coordination. *Computer Methods in Biomechanics and Biomedical Engineering*, *22*, S445–S447. https://doi.org/10.1080/10255842.2020.1714976

Merad, M., De Montalivet, E., Roby-Brami, A., & Jarrasse, N. (2016). Intuitive prosthetic control using upper limb inter-joint coordinations and IMU-based shoulder angles measurement: A pilot study. *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 5677–5682. https://doi.org/10.1109/IROS.2016.7759835

Mereu, F., Leone, F., Gentile, C., Cordella, F., Gruppioni, E., & Zollo, L. (2021). Control strategies and performance assessment of upper-limb TMR prostheses: A review. *Sensors*, *21*(6), 1953. https://doi.org/10.3390/s21061953

Mick, S., Segas, E., Dure, L., Halgand, C., Benois-Pineau, J., Loeb, G. E., Cattaert, D., & De Rugy, A. (2021). Shoulder kinematics plus contextual target information enable control of multiple distal joints of a simulated prosthetic arm and hand. *Journal of NeuroEngineering and Rehabilitation*, *18*(1), 3. https://doi.org/10.1186/s12984-020-00793-0

Miller, N., Jenkins, O., Kallmann, M., & Mataric, M. (2004). Motion capture from inertial sensing for untethered humanoid teleoperation. *4th IEEE/RAS International Conference on Humanoid Robots, 2004.*, *2*, 547–565. https://doi.org/10.1109/ICHR.2004.1442670

Modular prosthetic limb (MPL) v1. (2022). https://www.jhuapl.edu/prosthetics/program

Montagnani, F., Controzzi, M., & Cipriani, C. (2015a). Is it finger or wrist dexterity that is missing in current hand prostheses? *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *23*(4), 600–609. https://doi.org/10.1109/TNSRE.2015.2398112

Montagnani, F., Controzzi, M., & Cipriani, C. (2015b). Exploiting arm posture synergies in activities of daily living to control the wrist rotation in upper limb prostheses: A feasibility study. *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2462–2465. https://doi.org/10.1109/EMBC.2015.7318892

Moullet, E., Bailly, F., Carpentier, J., & Coste, C. A. (2023, June 27). Vision-based interface for grasping intention detection and grip selection : Towards intuitive upper-limb assistive devices. Retrieved June 3, 2024, from http://arxiv.org/abs/2306.15280

Mullette-Gillman, O. A., Cohen, Y. E., & Groh, J. M. (2005). Eye-centered, head-centered, and complex coding of visual and auditory targets in the intraparietal sulcus. *Journal of Neurophysiology*, *94*(4), 2331–2352. https://doi.org/10.1152/jn.00021.2005

Muñoz, J., De Santos-Rico, R., Mena, L., & Monje, C. A. (2024). Humanoid head camera stabilization using a soft robotic neck and a robust fractional order controller. *Biomimetics*, *9*(4), 219. https://doi.org/10.3390/biomimetics9040219

Nabeshima, C., Kuniyoshi, Y., & Lungarella, M. (2006). Adaptive body schema for robotic tool-use. *Advanced Robotics*, *20*(10), 1105–1126. https://doi.org/10.1163/156855306778522550

Nann, M., Cordella, F., Trigili, E., Lauretti, C., Bravi, M., Miccinilli, S., Catalan, J. M., Badesa, F. J., Crea, S., Bressi, F., Garcia-Aracil, N., Vitiello, N., Zollo, L., & Soekadar, S. R. (2021). Restoring activities of daily living using an EEG/EOG-controlled semiautonomous and mobile whole-arm exoskeleton in chronic stroke. *IEEE Systems Journal*, *15*(2), 2314–2321. https://doi.org/10.1109/JSYST.2020.3021485

Nguyen, V., Du, Y., Xiao, Y., Ramamonjisoa, M., & Lepetit, V. (2022). PIZZA: A powerful image-only zero-shot zero-CAD approach to 6 DoF tracking. https://doi.org/10.48550/arXiv.2209.07589

Noel, J.-P., Grivaz, P., Marmaroli, P., Lissek, H., Blanke, O., & Serino, A. (2015). Full body action remapping of peripersonal space: The case of walking. *Neuropsychologia*, *70*, 375–384. https://doi.org/10.1016/j.neuropsychologia.2014.08.030

O'Brien, J., Bodenheimer, R., Brostow, G., & Hodgins, J. (2000). Automatic joint parameter estimation from magnetic motion capture data [Artwork Size: 8 pages, 1.17 MB ISBN: 9780969533894 Medium: application/pdf Publisher: Canadian Human-Computer Communications Society]. *Proceedings of Graphics Interface 2000*, *Montréal*, 8 pages, 1.17 MB. https://doi.org/10.20380/GI2000.09

Oldfield, R. (1971). The assessment and analysis of handedness: The edinburgh inventory. *Neuropsychologia*, *9*(1), 97–113. https://doi.org/10.1016/0028-3932(71)90067-4.

Olsen, N. R., George, J. A., Brinton, M. R., Paskett, M. D., Kluger, D. T., Tully, T. N., Duncan, C. C., & Clark, G. A. (2019, October 17). An adaptable prosthetic wrist reduces subjective workload. https://doi.org/10.1101/808634

Ortiz-Catalan, M., Håkansson, B., & Brånemark, R. (2014). An osseointegrated human-machine gateway for long-term sensory feedback and motor control of artificial limbs. *Science Translational Medicine*, *6*(257). https://doi.org/10.1126/scitranslmed.3008933

Pachtrachai, K., Vasconcelos, F., Edwards, P., & Stoyanov, D. (2021). Learning to calibrate - estimating the hand-eye transformation without calibration objects. *IEEE Robotics and Automation Letters*, *6*(4), 7309–7316. https://doi.org/10.1109/LRA.2021.3098942

Parr, J. V. V., Galpin, A., Uiga, L., Marshall, B., Wright, D. J., Franklin, Z. C., & Wood, G. (2023). A tool for measuring mental workload during prosthesis use: The prosthesis task load index (PROS-TLX) (J. Selvaraj, Ed.). *PLOS ONE*, *18*(5), e0285382. https://doi.org/10.1371/journal.pone.0285382

Pelz, J., Hayhoe, M., & Loeber, R. (2001). The coordination of eye, head, and hand movements in a natural task. *Experimental Brain Research*, *139*(3), 266–277. https://doi.org/10.1007/s002210100745

Penčić, M., Čavić, M., Savić, S., Rackov, M., Borovac, B., & Lu, Z. (2017). Assistive humanoid robot MARKO: Development of the neck mechanism (I. Bondrea, C. Simion, & M. Inţă, Eds.). *MATEC Web of Conferences*, *121*, 08005. https://doi.org/10.1051/matecconf/201712108005

Peppoloni, L., Filippeschi, A., Ruffaldi, E., & Avizzano, C. A. (2013). A novel 7 degrees of freedom model for upper limb kinematic reconstruction based on wearable sensors. *2013 IEEE 11th International Symposium on Intelligent Systems and Informatics (SISY)*, 105–110. https://doi.org/10.1109/SISY.2013.6662551

Pérez De San Roman, P., Benois-Pineau, J., Domenger, J.-P., Paclet, F., Cataert, D., & De Rugy, A. (2017). Saliency driven object recognition in egocentric videos with deep CNN: Toward application in assistance to neuroprostheses. *Computer Vision and Image Understanding*, *164*, 82–91. https://doi.org/10.1016/j.cviu.2017.03.001

Popovic, D., & Popovic, M. (1998). Tuning of a nonanalytical hierarchical control system for reaching with FES. *IEEE Transactions on Biomedical Engineering*, *45*(2), 203–212. https://doi.org/10.1109/10.661268

Popovic, M., & Popovic, D. (2001). Cloning biological synergies improves control of elbow neuroprostheses. *IEEE Engineering in Medicine and Biology Magazine*, *20*(1), 74–81. https://doi.org/10.1109/51.897830

Pouget, A., Deneve, S., & Duhamel, J.-R. (2002). A computational perspective on the neural basis of multisensory spatial representations. *Nature Reviews Neuroscience*, *3*(9), 741–747. https://doi.org/10.1038/nrn914

Pugach, G., Pitti, A., Tolochko, O., & Gaussier, P. (2019). Brain-inspired coding of robot body schema through visuo-motor integration of touched events. *Frontiers in Neurorobotics*, *13*, 5. https://doi.org/10.3389/fnbot.2019.00005

Rau, G., Disselhorst-Klug, C., & Schmidt, R. (2000). Movement biomechanics goes upwards: From the leg to the arm. *Journal of Biomechanics*, *33*(10), 1207–1216. https://doi.org/10.1016/S0021-9290(00)00062-2

Resnik, L., Ekerholm, S., Borgia, M., & Clark, M. A. (2019). A national study of veterans with major upper limb amputation: Survey methods, participants, and summary findings (A. Eshraghi, Ed.). *PLOS ONE*, *14*(3), e0213578. https://doi.org/10.1371/journal.pone.0213578

Ribeiro, J., Mota, F., Cavalcante, T., Nogueira, I., Gondim, V., Albuquerque, V., & Alexandria, A. (2019). Analysis of man-machine interfaces in upper-limb prosthesis: A review. *Robotics*, *8*(1), 16. https://doi.org/10.3390/robotics8010016

Rogez, G., Supancic, J. S., & Ramanan, D. (2015). First-person pose recognition using egocentric workspaces. *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 4325–4333. https://doi.org/10.1109/CVPR.2015.7299061

Rong Zhu & Zhaoying Zhou. (2004). A real-time articulated human motion tracking using tri-axis inertial/magnetic sensors package. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *12*(2), 295–302. https://doi.org/10.1109/TNSRE.2004.827825

Rossetti, Y., Meckler, C., & Prablanc, C. (1994). Is there an optimal arm posture? deterioration of finger localization precision and comfort sensation in extreme arm-joint postures. *Experimental Brain Research*, *99*(1). https://doi.org/10.1007/BF00241417

Sah, S., & Wang, X. (2009). Determination of geometric constraints between the ribcage and scapula in the shoulder complex: A cadaver study. *Computer Methods in Biomechanics and Biomedical Engineering*, *12*, 223–224. https://doi.org/10.1080/10255840903093979

Salinas, E., & Abbott, L. (2001). Chapter 11 coordinate transformations in the visual system: How to generate gain fields and what to compute with them. In *Progress in brain research* (pp. 175–190, Vol. 130). Elsevier. https://doi.org/10.1016/S0079-6123(01)30012-2

Salminger, S., Stino, H., Pichler, L. H., Gstoettner, C., Sturma, A., Mayer, J. A., Szivak, M., & Aszmann, O. C. (2022). Current rates of prosthetic usage in upper-limb amputees – have innovations had an impact on device acceptance? *Disability and Rehabilitation*, *44*(14), 3708–3713. https://doi.org/10.1080/09638288.2020.1866684

Schillaci, G., Hafner, V. V., & Lara, B. (2014). Online learning of visuo-motor coordination in a humanoid robot. a biologically inspired model. *4th International Conference on Development and Learning and on Epigenetic Robotics*, 130–136. https://doi.org/10.1109/DEVLRN.2014.6982967

Schmidl, H. (1965). The inail myoelectric b/e prosthesis. *Orthotics and Prosthetics*, *19*(4), 298–303.

Scholz, J. P., Schöner, G., & Latash, M. L. (2000). Identifying the control structure of multijoint coordination during pistol shooting. *Experimental Brain Research*, *135*(3), 382–404. https://doi.org/10.1007/s002210000540

Segas, E., Mick, S., Leconte, V., Dubois, O., Klotz, R., Cattaert, D., & De Rugy, A. (2023). Intuitive movement-based prosthesis control enables arm amputees to

reach naturally in virtual reality. *eLife*, *12*, RP87317. https://doi.org/10.7554/eLife.87317.3

Segas, E. A., Leconte, V., Doat, E., Cattaert, D., & Rugy, A. D. (2024, April 23). Movement-based prosthesis control with angular trajectory is getting closer to natural arm coordination. https://doi.org/10.22541/au.171386556.60813345/v1

Ségas, E. (2023). *Contrôle biomimétique de prothèses à partir des mouvements résiduels et d'informations contextuelles* [Doctoral dissertation]. http://www.theses.fr/2023BORD0075/document

Sereno, M. I., & Huang, R.-S. (2006). A human parietal face area contains aligned head-centered visual and tactile maps. *Nature Neuroscience*, *9*(10), 1337–1343. https://doi.org/10.1038/nn1777

Serino, A. (2019). Peripersonal space (PPS) as a multisensory interface between the individual and the environment, defining the space of the self. *Neuroscience & Biobehavioral Reviews*, *99*, 138–159. https://doi.org/10.1016/j.neubiorev.2019.01.016

Serino, A., Noel, J.-P., Galli, G., Canzoneri, E., Marmaroli, P., Lissek, H., & Blanke, O. (2015). Body part-centered and full body-centered peripersonal space representations. *Scientific Reports*, *5*(1), 18603. https://doi.org/10.1038/srep18603

Seth, A., Matias, R., Veloso, A. P., & Delp, S. L. (2016). A biomechanical model of the scapulothoracic joint to accurately capture scapular kinematics during shoulder movements (L. Ren, Ed.). *PLOS ONE*, *11*(1), e0141028. https://doi.org/10.1371/journal.pone.0141028

Snyder, L. H., Grieve, K. L., Brotchie, P., & Andersen, R. A. (1998). Separate body- and world-referenced representations of visual space in parietal cortex. *Nature*, *394*(6696), 887–891. https://doi.org/10.1038/29777

Spiers, A. J., Gloumakov, Y., & Dollar, A. M. (2018). Examining the impact of wrist mobility on reaching motion compensation across a discretely sampled workspace. *2018 7th IEEE International Conference on Biomedical Robotics and Biomechatronics (Biorob)*, 819–826. https://doi.org/10.1109/BIOROB.2018.8487871

Standring, S. (2015). *Gray's anatomy e-book: Gray's anatomy e-book*. Elsevier Health Sciences. https://books.google.fr/books?id=b7FVCgAAQBAJ

Sturm, J., Plagemann, C., & Burgard, W. (2009). Body schema learning for robotic manipulators from visual self-perception. *Journal of Physiology-Paris*, *103*(3), 220–231. https://doi.org/10.1016/j.jphysparis.2009.08.005

Tsai, R., & Lenz, R. (1988). Real time versatile robotics hand/eye calibration using 3d machine vision. *Proceedings. 1988 IEEE International Conference on Robotics and Automation*, 554–561. https://doi.org/10.1109/ROBOT.1988.12110

Vercher, J., Magenes, G., Prablanc, C., & Gauthier, G. (1994). Eye-head-hand coordination in pointing at visual targets: Spatial and temporal analysis. *Experimental Brain Research*, *99*(3). https://doi.org/10.1007/BF00228987

Viviani, P., & Flash, T. (1995). Minimum-jerk, two-thirds power law, and isochrony: Converging approaches to movement planning. *Journal of Experimental Psychology: Human Perception and Performance*, *21*(1), 32–53. https://doi.org/10.1037/0096-1523.21.1.32

Wang, G., Quan, W., Li, Y., Fang, S., Chen, H., & Xi, N. (2021). Fast and accurate 3d eye-to-hand calibration for large-scale scene based on HALCON. *2021 IEEE 11th Annual International Conference on CYBER Technology in Automation, Control, and Intelligent Systems (CYBER)*, 230–234. https://doi.org/10.1109/CYBER53097.2021.9588160

Xiang, Y., Schmidt, T., Narayanan, V., & Fox, D. (2017). PoseCNN: A convolutional neural network for 6d object pose estimation in cluttered scenes [Publisher: [object Object] Version Number: 3]. https://doi.org/10.48550/ARXIV.1711.00199

YOLOv8. (2022). https://docs.ultralytics.com

Zhou, H., & Hu, H. (2008). Human motion tracking for rehabilitation—a survey. *Biomedical Signal Processing and Control*, *3*(1), 1–18. https://doi.org/10.1016/j.bspc.2007.09.001

# Appendix A

# Publications and oral presentations

## A.1 Publications

Lento, B., Segas, E., Leconte, V., Doat, E., Danion, F., Péteri, R., Benois-Pineau, J., & de Rugy, A. (2024, January 25). Data and code for 3D-ARM-gaze: A public dataset of 3d arm reaching movements with gaze information in virtual reality. https://doi.org/10.5281/ZENODO.10567366

Lento, B., Segas, E., Leconte, V., Doat, E., Danion, F., Péteri, R., Benois-Pineau, J., & De Rugy, A. 3d-ARM-gaze: A public dataset of 3d arm reaching movements with gaze information in virtual reality. Sci Data 11, 951 (2024). https://doi.org/10.1038/s41597-024-03765-4

Lento, B., Leconte, V., Bardisbanian, L., Doat, E., Segas, E., & De Rugy, A. (2024). Bioinspired head-to-shoulder reference frame transformation for movementbased arm prosthesis control. IEEE Robotics and Automation Letters, 9(9),7875-7882. https://doi.org/10.1109/LRA.2024.3433751

Leconte, V., Segas, E., Lento, B., et al. (2024, August 01) First-person teleoperation of humanoid robot Reachy by valid and arm-amputated participants with a novel movement-based prosthesis control. TechRxiv. https://doi.org/10.36227/techrxiv.172253934.48844352/v1

## A.2 Oral presentations

Poster presentation at ISEK (International Society of Electrophysiology and Kinesiology) Congress, Quebec City, 22-25 June 2022: “Intuitive prosthesis control based on residual (stump) motion tested with valid subjects and amputees on the robotic platform REACHY”

Poster presentation at École du GT8: Apprentissage et neurosciences pour la robotique (GT8 school: Learning and neuroscience for robotics), Moliets-et-Ma, 17-20 Oc-

tober 2022: “Intuitive prosthesis control based on residual (stump) motion tested with valid subjects and amputees on the robotic platform REACHY”

Poster presentation at Journée des Jeunes Chercheurs en Robotique 2023, Moliets-et-Ma, 16 October 2023: “Biomimetic prosthesis control from natural movements: personalization of a generic model using co-adaptation”

Poster presentation at Journées Nationales de la Recherche en Robotique (JNRR), Moliets-et-Ma, 17-20 October 2023: “Biomimetic prosthesis control from natural movements: personalization of a generic model using co-adaptation”

Participation at the Forum Innovation Défense 2023 as an exhibitor, Paris, 23-25 November 2023

Participation in the competition Ma Thése en 180s 2024 (semifinals and finals of University of Bordeaux): https://www.youtube.com/watch?v=HQx8WFD4bAI

Oral presentation type MT180s at the “Journée de l’École doctorale Sociétés, politique, santé publique 2024” (Doctoral School Day), 10 April 2024

Oral presentation at INCIA Data Blitz Talk, 12 June 2024: “Bioinspired head-to-shoulder reference frame transformation for movement-based arm prosthesis control”

Oral presentation at 5th International NeuroErgonomics Conference (NEC’24), Bordeaux, 8-12 July 2024: “Bioinspired adaptive spatial map for prosthesis control”

# Appendix B

# Videos

**Illustrative Video 1** : https://youtu.be/4ZRYN6ljeCQ

A participant performing the RNP procedure and the pick-and-place task.

**Illustrative Video 2** : https://www.youtube.com/watch?v=RZhN5IR34_c

Examples of new video data generated with the DataPlayer by combining real data about arm and gaze with various fictitious visual backgrounds and objects.

**Illustrative Video 3** : https://youtu.be/IsEasynQHkk

A participant with transhumeral limb loss using our Proximo Contextual prosthesis control with Space Map.

**Illustrative Video 4** : https://youtu.be/7GQemjYP5rI

REACHY 2 teleoperated from an egocentric point of view.

**Illustrative Video 5** : https://youtu.be/SQNKJrkQBXE

Space Map adaptation for head-shoulder transformation on REACHY 2 teleoperated from egocentric vision.

# Appendix C

# Offline evaluations of PC ANN accuracy

Before carrying out the experiments detailed in Chapter 4, we conducted offline evaluations to assess the ability of the PC ANN to position and orient accurately the virtual hand. The objective was to determine the most suitable grasping constraints for using this control system. During these tests, we utilized data from the Initial Acquisition phase of the experiment described in Chapter 2. We extracted joint configurations of participants when their virtual hand was within the target zone. These configurations, coupled with the segment dimensions of a designated test participant, were employed in forward kinematics calculations to determine the positions and orientations of the virtual hand in the armroot reference frame. Subsequently, the PC ANN, trained on the segment dimensions of the designated test participant, received as input the obtained virtual hand positions and orientations, along with the proximal joint angles, and predicted the distal joint angles. These newly predicted joint configurations, consisting of the actual proximal joints and the predicted distal joints, were employed in forward kinematics calculations to determine the predicted virtual hand positions and orientations. The prediction error was computed as the difference between the actual and predicted virtual hand positions and orientations.

Figure C.1 illustrates the outcomes from the offline evaluations of the PC ANN accuracy. The graph displays the distribution of errors, showcasing the percentages falling below each distinct value. Vertical lines correspond to specific percentiles (50, 90, and 99%), providing insights into the network's performance. Concerning position errors, it is worth noting that 99% of the values are below 5.7cm. A substantial portion, 90%, of these errors falls below the 3cm threshold. Furthermore, half of the errors are below 1.3cm, highlighting a high level of precision in positioning the hand. Concerning orientation errors, 99% of the values are below 18.9°. The network demonstrates enhanced accuracy in orientation, with 90% of errors falling below 8.2° and 50% below 3.6°.

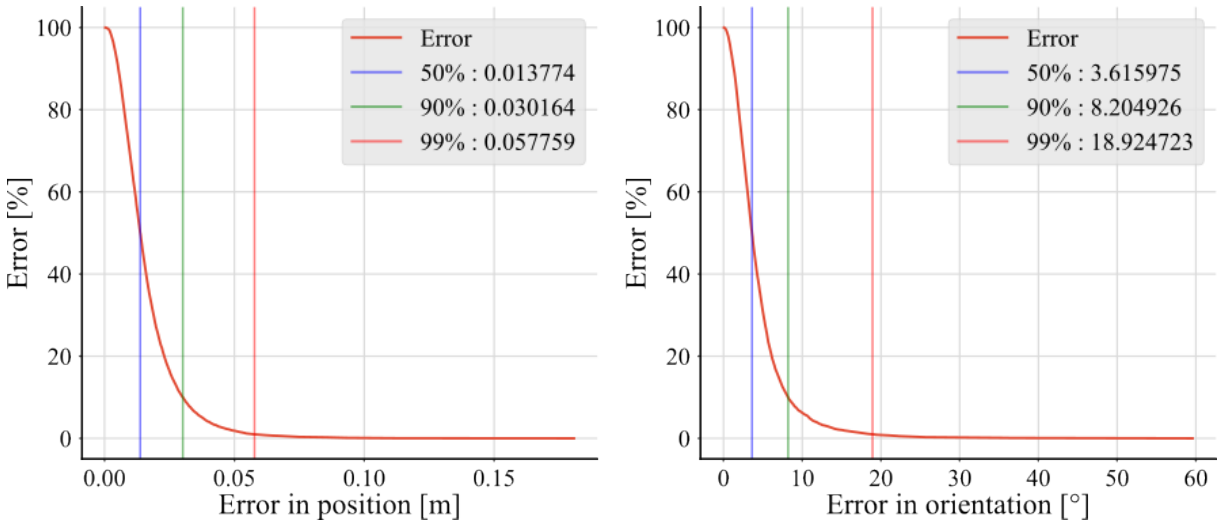



Figure C.1: Error in prediction of PC ANN.

# Appendix D

# Survey SUS

Évaluez chacune de ces affirmations en indiquant à l'aide de l'échelle à quel point vous êtes en accord avec elles.

| Item du F-SUS | 1 (Pas du tout d'accord) | 2 | 3 | 4 | 5 (Tout à fait d'accord) |
|---|---|---|---|---|---|
| 1. Je voudrais utiliser ce contrôle de bras fréquemment. | | | | | |
| 2. Ce contrôle de bras est inutilement complexe. | | | | | |
| 3. Ce contrôle de bras est facile à utiliser. | | | | | |
| 4. J'aurais besoin du soutien d'un technicien pour être capable d'utiliser ce contrôle de bras. | | | | | |
| 5. Les différentes fonctionnalités de ce contrôle de bras sont bien intégrées. | | | | | |
| 6. Il y a trop d'incohérences dans ce contrôle de bras. | | | | | |
| 7. La plupart des gens apprendront à utiliser ce contrôle de bras très rapidement. | | | | | |
| 8. Ce contrôle de bras est très lourd à utiliser. | | | | | |
| 9. Je me suis senti·e très en confiance en utilisant ce contrôle de bras. | | | | | |
| 10. J'ai eu besoin d'apprendre beaucoup de choses avant de pouvoir utiliser ce contrôle de bras. | | | | | |

# Appendix E

# Survey rawProsTLX

Les 8 échelles de mesure suivantes ont été imaginées pour évaluer votre expérience durant la tâche. Veuillez évaluer la procédure en indiquant le chiffre qui correspond le plus à votre expérience dans chacune des 8 échelles présentées. Les échelles vont de « Faible » à gauche à « Élevé » à droite. Veuillez écouter les descriptions attentivement.

**Charge mentale**

À quel point était-il mentalement fatigant d'utiliser ce contrôle de bras durant cette tâche ?

| | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 Faible | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 Élevé |

**Demande physique**

À quel point était-il physiquement fatigant d'utiliser ce contrôle de bras durant cette tâche ?

| | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 Faible | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 Élevé |

**Attention visuelle**

À quel point deviez-vous regarder le bras virtuel alors que vous bougiez durant cette tâche ?

| | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 Faible | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 Élevé |

**Traitement conscient**

À quel point deviez-vous penser à votre façon de bouger durant cette tâche ?

| | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 Faible | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 Élevé |

**Frustration**

À quel point étiez-vous inquiet·iète, découragé·e, irrité·e. ou contrarié·e durant cette tâche ?

| | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 Faible | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 Élevé |

**Stress situationnel**

À quel point vous sentiez-vous anxieux·se ou stressé·e durant cette tâche ?

| | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 Faible | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 Élevé |

**Pression temporelle**

À quel point vous sentiez-vous pressé.e ou précipité.e durant cette tâche ?

| | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 Faible | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 Élevé |

**Incertitude**

À quel point ce contrôle de bras était-il imprédictible durant la réalisation de cette tâche ?

| | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 Faible | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 Élevé |

# Appendix F

# Statistical analysis results of Exp1 and Exp2

Table F.1: Exp1 Success rate - Pairwise comparisons using McNemar's test

| | Nat | PC 2 | PC A2H | PC A2H SMAP 1B | PC A2H SMAP 2B | PC SMAP 1B |
|---|---|---|---|---|---|---|
| PC 2 | 9.56e-01 | - | - | - | - | - |
| PC A2H | 1.01e-19 | 6.20e-17 | - | - | - | - |
| PC A2H SMAP 1B | 2.44e-01 | 1 | 5.71e-16 | - | - | - |
| PC A2H SMAP 2B | 6.09e-01 | 1 | 4.72e-17 | 1 | - | - |
| PC SMAP 1B | 4.64e-04 | 3.57e-02 | 1.26e-11 | 3.32e-01 | 8.17e-02 | - |
| PC SMAP 2B | 1.60e-01 | 1 | 5.80e-17 | 1 | 1 | 1.96e-01 |

Table F.2: Exp1 Movement time - Pairwise comparisons using Conover's all-pairs test

| | Nat | PC 2 | PC A2H | PC A2H SMAP 1B | PC A2H SMAP 2B | PC SMAP 1B |
|---|---|---|---|---|---|---|
| PC 2 | 3.6e-05 | - | - | - | - | - |
| PC A2H | 1 | 7.3e-05 | - | - | - | - |
| PC A2H SMAP 1B | 0.04123 | 0.98119 | 0.07114 | - | - | - |
| PC A2H SMAP 2B | 0.00207 | 1 | 0.00389 | 1 | - | - |
| PC SMAP 1B | 1 | 0.00015 | 1 | 0.12034 | 0.00720 | - |
| PC SMAP 2B | 0.09276 | 0.51295 | 0.15532 | 1 | 1 | 0.25464 |

Table F.3: Exp1 Shoulder spread volume - Pairwise comparisons using Conover's all-pairs test

| | Nat | PC 2 | PC A2H | PC A2H SMAP 1B | PC A2H SMAP 2B | PC SMAP 1B |
|---|---|---|---|---|---|---|
| PC 2 | 1 | - | - | - | - | - |
| PC A2H | 0.00056 | 0.00530 | - | - | - | - |
| PC A2H SMAP 1B | 1 | 1 | 0.05429 | - | - | - |
| PC A2H SMAP 2B | 1 | 1 | 0.00720 | 1 | - | - |
| PC SMAP 1B | 0.00530 | 0.04123 | 1 | 0.32340 | 0.05429 | - |
| PC SMAP 2B | 1 | 1 | 0.15532 | 1 | 1 | 0.79509 |

Table F.4: Exp1 SUS score - Pairwise comparisons using Conover's all-pairs test

| | Nat | PC | PC A2H | PC A2H SMAP |
|---|---|---|---|---|
| PC | 0.5006 | - | - | - |
| PC A2H | 0.004 | 0.7559 | - | - |
| PC A2H SMAP | 1 | 0.8625 | 0.0087 | - |
| PC SMAP | 0.9813 | 1 | 0.3753 | 1 |

Table F.5: Exp1 rawProsTLX score - Pairwise comparisons using Conover's all-pairs test

| | Nat | PC | PC A2H | PC A2H SMAP |
|---|---|---|---|---|
| PC | 0.1046 | - | - | - |
| PC A2H | 0.0041 | 1 | - | - |
| PC A2H SMAP | 1 | 1 | 0.0752 | - |
| PC SMAP | 0.4759 | 1 | 0.8145 | 1 |

Table F.6: Exp2 Success rate - Pairwise comparisons using McNemar's test

| | Nat | PC 2 | PC SMAP 1B |
|---|---|---|---|
| PC 2 | 1 | - | - |
| PC SMAP 1B | 1.53e-06 | 1.84e-05 | - |
| PC SMAP 2B | 1.10e-01 | 1 | 1.97e-03 |

Table F.7: Exp2 Movement time - Pairwise comparisons using Tukey's test

| | Nat | PC 2 | PC SMAP 1B |
|---|---|---|---|
| PC 2 | 0.54738 | - | - |
| PC SMAP 1B | 0.54341 | 0.06188 | - |
| PC SMAP 2B | 0.95269 | 0.84383 | 0.27089 |

Table F.8: Exp2 rawProsTLX score - Pairwise comparisons using Conover's all-pairs test

| | Nat | PC |
|---|---|---|
| PC | 0.34 | - |
| PC SMAP | 0.08 | 1 |