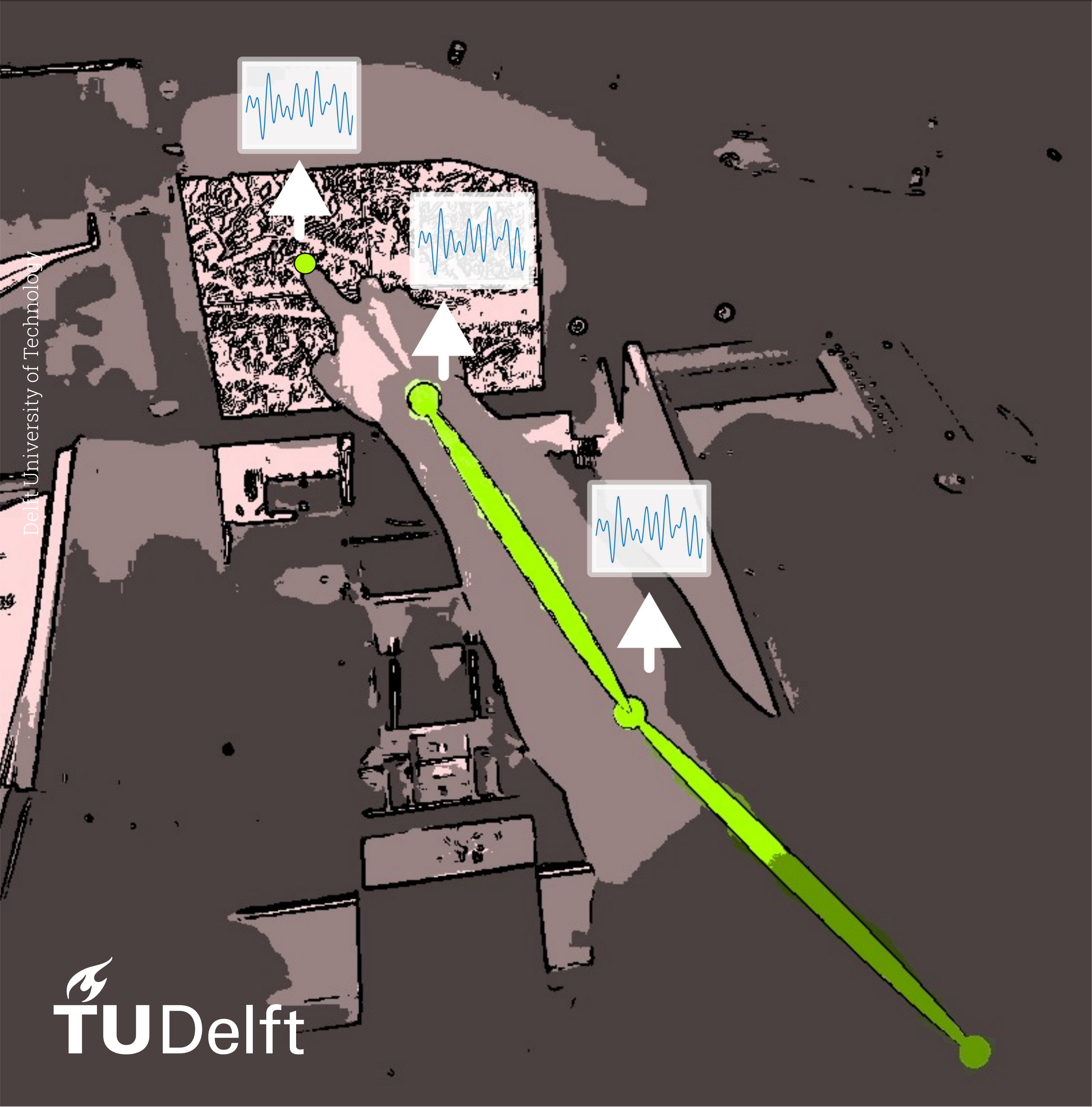
Towards Adaptive Biodynamic Feedthrough Modeling for Touchscreen Applications

MSc Thesis

Alexandru Iancu



Delft University of Technology

Towards Adaptive Biodynamic Feedthrough Modeling for Touchscreen Applications

by

Alexandru Iancu

to obtain the degree of Master of Science in Aerospace Engineering
at Delft University of Technology

Student number:	4837134		
Date:	April 24, 2024		
Thesis committee:	Dr. ir. D.M. Pool,	Delft University of Technology,	Supervisor
	Prof. dr. ir. M. Mulder,	Delft University of Technology,	Graduation chair
	Dr. ir. M.M. van Paassen,	Delft University of Technology,	Graduation member
	Dr. R. Merino Martinez,	Delft University of Technology,	Examiner

TUDelft

Preface

This report represents a compilation of the work and ideas that were developed in the past year while researching whether it is possible to use cameras in order to capture the biodynamic characteristics of humans in touchscreen operations. The MSc Thesis consists of three parts: A Scientific Paper (Part I) that summarizes the main findings of the experiments performed in the thesis phase, a Preliminary Report (Part II) that presents a literature study and computer-based simulations, and the MSc Thesis Appendices (Part III) that provide background information and additional results of the work performed in the MSc Thesis.

I would like to thank my amazing supervisor, Daan Pool. Through your great supervision, I was certainly able to get out of my comfort zone and try to take slightly more initiative in my thesis work. Moreover, your calm, friendly but also professional attitude is an inspiration for me. My sincerest appreciation goes toward your dedication to help and educate the hundreds of students that follow your classes and the ones that are supervised by you. Last but not the least, would also like to thank the other staff at C&S (Ferdinand, Andries, Harold and Olaf) for their support and help in this project.

After defending this thesis, my almost six years at TU Delft will come to an end. Back in 2018, when I started my first year, I certainly was overwhelmed by the fresh start that came with my move to The Netherlands. Slowly and steadily, new pages in ”my book of experiences” were being written through little failures, successes, hard work, reflections and friendships. I am and will be forever thankful to my parents for their support and for providing me the opportunity to get a good education far from home. Having the opportunity to study and live abroad has been one of the best lessons that I could have received during my youth. I am also grateful for my friends that I’ve made here in Delft through these years. You always bring a smile on my face and I am sure that we will remember the memories that we’ve made together during our student times for years to come.

Alexandru Iancu
Delft, April 2024

Contents

List of Figures

List of Tables

Nomenclature

Abbreviations

Abbreviation	Definition
AI	Artificial Intelligence
APC	Aircraft-Pilot Coupling
BDFT	Biodynamic Feedthrough
C	compliant neuromuscular setting
CE	Controlled Element
Cl	Close
CNN	Convolutional Neural Network
CNS	Central Nervous System
COM	Center Of Mass
CRT	Cathode Ray Tube
DOF	Degree Of Freedom
DLT	Direct Linear Transform
EFB	Electronic Flight Bag
El	Elbow
EMG	Electromyography
F	Finger fixed on scren
Fa	Far
FDFT	Force Disturbance Feedthrough
Fi	Finger
FMS	Flight Management System
FRF	Frequency Response Function
FT	Force Task
HMI	Human-Machine Interaction
HO	Human Operator
HREC	Human Research Ethics Commitee
IMU	Inertial Measurement Unit
LCD	Liquid Crystal Display
LTI	Linear Time-Invariant
M	Finger moving on the screen
MCDU	Multi-Function Control and Display Unit
MDFT	Motion Disturbance Feedthrough
N	Finger not touching the screen
NN	Neural Network
PAO	Pilot-Assisted Oscillations
PIO	Pilot-Induced Oscillations
PT	Position Task
RANSAC	Random Sample Consensus
RMS	Root-Mean-Square
ROM	Range-Of-Motion

Abbreviation	Definition
RR	Roll Ratcheting
RT	Relax Task
SRS	Simona Research Simulator
TLAM	Torque-less Activation Mode
Tr	Transition
Wr	Wrist

Symbol	Definition
Γ^2_{bdft}	coherence
ζ_{bdft}	damping ratio of the BDFT model
$\mu_{y_{T_{Fi}}} Y_{sc}$	mean of the windowed position of the finger expressed in Y_{sc}
$\overline{\mu}_{\vert\epsilon_{t-c}\vert}$	The mean of the absolute sample-wise difference between the signal observed at the screen and the signal extracted from the cameras
ω_{bdft}	frequency of the BDFT model
ω_d	disturbance frequencies
ω_{t_z}	target frequency in the Z_{sc} direction
ω_{t_x}	target frequency in the X_{sc} direction
$\sigma^2_{y_{T_f}}/\sigma^2_{y_T}$	percentage of the standard deviation of the feedthrough component from the total signal
$\sigma^2_{y_{T_r}}/\sigma^2_{y_T}$	percentage of the standard deviation of the remnant component from the total signal
ϕ_d	phases of the disturbance signal
ϕ_{t_z}	phase of the target signal in the Z_{sc} direction
ϕ_{t_x}	phase of the target signal in the X_{sc} direction
A_d	amplitudes of the disturbance signal
A_{t_z}	amplitude of the target signal in the Z_{sc} direction
A_{t_x}	amplitude of the target signal in the X_{sc} direction
a	slope that relates the finger's normal position from the screen to finger's vertical velocity
a_x	focal length expressed in the x pixel dimension
a_z	focal length expressed in the z pixel dimension
b	intercept that relates the finger's normal position from the screen to finger's vertical velocity
G_{bdft}	gain of the BDFT model
H_{bdft}	fitted BDFT frequency response function
$\hat{H}_{bdft}$	identified BDFT frequency response function
k	frequency index
n_d	number of times a sinusoid fits in the measurement time
R	rotation matrix between the two cameras
$RMS\, Z_{sc}$	Root-mean-square of a general signal expressed in Z_{sc}
S	stiff neuromuscular setting
r_{xx}	rotation element of the extrinsic matrix
T	translation vector between the two cameras
T_m	measurement time

Symbol	Definition
t_x	translation element of the extrinsic matrix
u_0	x coordinate of the principal point in the 2D sensor coordinate frame
v_0	x coordinate of the principal point in the 2D sensor coordinate frame
X_k	X coordinate in the 3D camera-fixed frame
x_s	x coordinate in the 2D sensor frame
X_{sc}	X coordinate in the 3D screen frame
X_w	X coordinate in the 3D world frame
Y_k	Y coordinate in the 3D camera-fixed frame
Y_{sc}	Y coordinate in the 3D screen frame
$y_{T_{El}}$	triangulated elbow location
$y_{T_{El_f}}$	feedthrough component of the triangulated elbow location
y_{Fi}	position of the finger joint obtained either using the cameras or the touchscreen
$y_{T_{Fi}}$	triangulated finger location
$y_{T_{Fi_f}}$	feedthrough component of the triangulated finger location
$y_{T_{Wr}}$	triangulated wrist location
$y_{T_{Wr_f}}$	feedthrough component of the triangulated wrist location
$y_{TSC_{Fi}}$	finger location obtained using the touchscreen
$y_{\dot{T}_{Fi}}\,Y_{sc}$	first derivative of the position of the finger expressed in Y_{sc}
$y_{\dot{T}_{Fi}}\,Z_{sc}$	first derivative of the position of the finger expressed in Z_{sc}
Y_w	Y coordinate in the 3D world frame
Z_k	Z coordinate in the 3D camera-fixed frame
z_s	z coordinate in the 2D sensor frame
Z_{sc}	Z coordinate in the 3D screen frame
Z_w	Z coordinate in the 3D world frame

Part I

Scientific Paper

Investigating the use of markerless pose estimators for capturing the adaptive biodynamic feedthrough characteristics in touchscreen operations

Alexandru Iancu*
Delft University of Technology, Delft The Netherlands

Given the expected increase of automation in the vehicles of the future, touchscreens are expected to be used in a wide variety of scenarios, including the ones that can become safety-critical. One disadvantage of the direct interaction approach that characterizes touchscreen operation is the feedthrough of accelerations through the human body that can induce unwanted activations, leading to unsafe situations. The current paper investigates the feasibility of applying a method that can track the biodynamic responses of the elbow, wrist and index finger using cameras in a stereo configuration and an open-source pose estimator (OpenPose). An experiment with six participants was performed to understand whether different neuromuscular settings, task instructions or degrees of arm extension induce an adaptive behavior of the biodynamic feedthrough to the recorded joints. The outcomes of the experiments showed that the finger is the body part that exhibits the most adaptive behavior in terms of feedthrough, being dependent on whether it is interacting with the screen with no velocity, moving on the screen or not touching the screen. Moreover, the difference between keeping the upper limb relaxed or stiff would need to be taken into account for the specified body part when the finger is not touching the screen, showing a decrease of 29% in the mean RMS of the vertical feedthrough component when the stiff condition is exhibited. The adaptive nature of the biodynamic feedthrough was also demonstrated concerning two different cases of arm extension (close and far from the body). This study proved the feasibility of using a linear mass-spring-damper model to the data gathered with the camera-based system to describe the feedthrough effects for all the body parts, neuromuscular settings and tasks, being able to explain on average at least 75% of the variance of the raw signals. The approach presented in this study can be further refined toward reaching real-time capabilities, which is a key step in accurately identifying the highly adaptive and subjective nature of the feedthrough of the accelerations.

I. Introduction

Touchscreens have become an invaluable part of society's day-to-day life. From smartphones, laptops, tablets and store checkouts, the interaction between humans and the environment that surrounds them has slowly transitioned to touch-based approaches. The conventional buttons have also been replaced by touchscreens in cars, with more than 97% of the road vehicles produced in 2023 having at least one touchscreen in their cabin [1]. The touchscreen transition did not only resort to road vehicles, the aviation domain being another example where important advancements were made towards new cockpit designs that use this technology. Nowadays, touchscreens are mainly used in the general aviation domain, examples being the avionics designed by Garmin [2] which use touch interaction for data entry and flight information retrieval [3]. Since the future vehicles are expected to become more and more autonomous, the role of humans will transition from actively controlling the vehicles to passively monitoring and commanding them through interfaces that most likely will rely strongly on touch technology.

The ability of the users to engage directly with the system is one of the touchscreens' key benefits [4]. Moreover, as Rouwhorst et al. [5] conclude, under some circumstances and with carefully designed user interfaces, touchscreens can decrease the workload and speed up tasks in the cockpit. From an operational point of view, the contents presented on the screens and their configurations can be rapidly exchanged by changing the interface with the user without the need to reconfigure any hardware [6]. On the other hand, the usability of touchscreen displays decreases when the user obscures parts of the display or when the dimensions of the targets are much smaller than the finger width [7]. Another possible

---

[*] MSc Student Control and Simulation Profile part of Control and Operations Section, Faculty of Aerospace Engineering Kluyverweg 1, 2629 HS, Delft, The Netherlands.

disadvantage is that a reduced interaction accuracy might be encountered in the use of touchscreens if tactile feedback is missing [8]. However, one of the most important issues with touchscreen interaction is represented by the accidental inputs that are indirectly caused by the unsteady conditions (air turbulence, rough parts of a road or sea) through which the vehicle travels [9].

Involuntary inputs that are caused by the transmission of disturbances through the human body are defined as biodynamic feedthrough (BDFT). Cockburn et al. [10] concluded that the induced disturbances negatively affected the error rates and the time to perform the task when operating a touchscreen. The first study that explicitly analyzed and quantified the influence of touchscreen BDFT from a cybernetic perspective was the one published by Mobertz et al. [11]. The fact that the influence of BDFT on the touchscreen input could be identified and modeled inspired Khoshnewiszadeh & Pool [12] to implement model-based cancellation techniques for mitigating biodynamic feedthrough in continuous dragging tasks. Although successful, it was pointed out that disturbances affect different people in different ways, based on their individual strategies for performing the task at hand. Similar conclusions were observed by Venrooij et al. [13], who analyzed the same idea of cancellation when using control inceptors. An additional insight provided by the same research group is that a relationship exists between the dynamics of the feedthrough of accelerations and how stiff the neuromuscular dynamics are [14]. Venrooij et al. [14] and Khoshnewiszadeh & Pool [12] suggested that for the signal cancellation to be effective, the model describing the BDFT would need to be adapted to the individual subject and to the task at hand.

Given the novelty of this topic, the research on BDFT modeling for touchscreen applications still has some unexplored areas. Previous studies have only focused on analysing the movement of the finger through signals that were obtained from the touchscreen. In this way, the BDFT analysis was limited only to tasks that required continuous interaction between the finger and the screen (dragging tasks). One example where recording only the signal on the screen is not sufficient is in pointing tasks, since most of the unwanted movement happens before the finger touches the screen. In this way, for analysing more realistic tasks, it is important not only to investigate the movement of the finger through an alternative source but also the movement of the other parts of the upper limb that would help to better understand how unwanted disturbances feed through the human body. Moreover, an explicit dependency of the BDFT on the neuromuscular settings was not explored in touchscreen operation. Lastly, the effects of BDFT on the whole limb when the nominal velocity of the finger on the screen is zero or when the finger is on the verge of touching the screen were not analysed.

Image-based markerless pose estimation methods that make use of artificial neural networks to automatically detect features on the human body [15–17] represent a rapidly growing research area in the era of Artificial Intelligence (AI) [18]. One particular advantage of these methods is that they are not intrusive, since no markers or devices are mounted on the human body. Moreover, since only normal cameras are required makes these approaches feasible to be used for tracking human motion in constrained spaces such as cockpits and cabins [19]. On the other hand, studies such as [19–21] proved that markerless methods are less reliable when compared to the less flexible and more expensive state-of-the-art marker-based systems. Nevertheless, the previous studies compared markerless and marker-based systems in highly dynamic situations such as walking, throwing or jumping while focusing on tracking the whole body. The existing literature did not investigate the accuracy of markerless pose estimation methods in the context of disturbance transmission through the human body, especially for the upper limb.

The current paper has two main goals. The first goal is to investigate whether camera-based pose reconstruction is feasible to be reliably used for capturing the biodynamic characteristics of humans while operating a touchscreen in turbulence. The second goal is to investigate the adaptability of BDFT to the task that is presented, to the stiffness of the upper limb and to the recorded joints of the right arm. To reach the goals, a human-in-the-loop experiment with six participants was performed in TU Delft's SIMONA Research Simulator (SRS) that simulated turbulence in the vertical direction. The experiment concerned the analysis of the upper limb movement of the subjects through camera-based markerless pose estimation methods (OpenPose [15]) and two cameras. In the first part of the experiment, the elbow, wrist and finger movement was investigated while varying the required neuromuscular settings (either compliant or stiff) that the subjects exhibited. In the same time, three different interaction methods with the screen were tested: finger touching the screen and moving with zero nominal velocity, finger touching the screen and moving circularly with a nominal velocity of 50 mm/s and finger not touching the screen and moving with zero nominal velocity. This part also investigated the feasibility of extracting a BDFT model using the information from the pose estimator and the cameras. The second part of the experiment focused on the effect of the degree of limb extension (from close to the body to fully extended) on the feedthrough susceptibility. This second part of the experiment had the goal of closing the gap toward more realistic target-reaching tasks.

In Section II, an overview of the steps that were taken to estimate the location of the landmarks of interest from the cameras and pose estimator will be given. The third section will present an overview of the experimental setup, conditions, independent and dependent variables of the study. The fourth section will first present the time domain analysis of the outcomes from the two parts of the experiment, followed by the results of the model fit and the comparison with the estimated frequency response functions from the first part of the experiment. Lastly, Sections V and VI will comment on the work and conclude it.

II. Method

Photogrammetry is the science that aims to reconstruct the geometrical characteristics of an environment, such as distances or 3D representations, based on the interpretation of 2D images that capture that environment [22]. In general, only analysing one perspective of the environment is not enough to reconstruct its contents. In this way, based on the analogy with human vision, at least two perspectives of the same scene are needed to estimate the depth of any visible object [22]. The method used in this study is based on the use of stereoscopic vision in order to reconstruct the 3D locations of the elbow, wrist and index finger landmarks identified with an image-based pose estimator. If the pixel locations of the landmarks are identified from multiple perspectives (two in this case), then it is possible to make an inference that uses pixels from the two-dimensional perspectives to estimate the 3D locations of the points of interest on the limb. This approach is considered promising since the movement of the limbs can be tracked without mounting any marker or device on the limbs, therefore making it feasible for future real-life applications. These principles have been previously tested in the literature, especially in the movement analysis/rehabilitation domains [16, 19, 21, 23–25], showing an increased potential of capturing human's biomechanics under various scenarios similar to the state-of-the-art marker-based systems.

The triangulation is performed by assuming a mathematical model that can describe the working principle of the cameras. In the field of photogrammetry, the cameras are most often idealised using the pinhole model [26]. For one camera, this model can describe how a point in the 3D space is projected on a 2D image. Equation (1) presents the pinhole camera model for one camera. This model maps a 3D point $[X_w, Y_w, Z_w]^T$ expressed in a world coordinate frame (w), to the pixel locations $[x_s, z_s]^T$ expressed in the sensor coordinate frame (s). The pinhole camera model uses coordinates that are defined in a projective space to facilitate linear operations (such as matrix multiplication). A vector $[x, y]^T \subset \mathbb{R}^2$ is equivalent with its counterpart expressed in a projective space $[x, y, 1]^T \subset \mathbb{R}^3$ which, in turn, is equivalent to any multiple m of the projective space vector $[mx, my, m]^T$. The interested reader is encouraged to consult MSc Thesis Appendices: Appendix C for a more in-depth explanation of the model. As can be observed in equation (1), the total camera matrix that maps from 3D to 2D is made of extrinsic parameters, which describe the pose of the camera in the world, and intrinsic parameters, which describe the internal characteristics of the camera such as the focal length, optical center and image dimensions. The terms r_{xx} in the extrinsic camera matrix describe a rotation matrix between the camera-fixed reference frame and the world reference frame. Similarly, the terms t_x represent the translation between the origin of the world reference frame and the camera's center of projection. The terms a_x and a_z in the intrinsic camera matrix represent the focal length expressed in the dimension of the pixels, and u_0, v_0 represent the location of the principal point (the position of the projection center expressed in the sensor frame) expressed in pixel dimensions [26]. The constant term m in this equation indicates that if one would like to extrapolate from 2D to 3D (i.e. the opposite transformation), it is not possible to reconstruct the 3D point since it will lie on a ray at an unknown distance from the camera. In this way, only one perspective cannot be used to reconstruct a 3D location and two cameras are needed.

$$m \begin{bmatrix} x_s \\ z_s \\ 1 \end{bmatrix}_s = \underbrace{\begin{bmatrix} a_x & 0 & u_0 \\ 0 & a_z & v_0 \\ 0 & 0 & 1 \end{bmatrix}}_{\text{Intrinsic parameters}} \underbrace{\begin{bmatrix} r_{00} & r_{01} & r_{02} & t_0 \\ r_{10} & r_{11} & r_{12} & t_1 \\ r_{20} & r_{21} & r_{22} & t_2 \end{bmatrix}}_{\text{Extrinsic parameters}} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}_w \tag{1}$$

To obtain the 3D locations of the features of interest, a number of steps were proposed that were based on previous work which focused on the use of OpenPose for the analysis of the whole-body motion [20, 24], as presented in Figure 1.

1. Step 1: Intrinsic and extrinsic camera calibration

The first step of the methodology is the intrinsic and extrinsic calibration of the two cameras. In other words, the parameters corresponding to equation (1) (including the distortion coefficients presented in MSc Thesis Appendices:

Appendix C) need to be estimated for both cameras that are being used in the experiment. The calibration method that was chosen for this experiment is the one developed by Zhang [27], which is implemented in MATLAB's *Computer Vision Toolbox (v2023b, Mathworks Inc., Natwick, USA)*[*]. The reason for choosing this method is that it is one of the most used methods for image processing/robotics applications and for estimating the camera parameters. This approach requires the use of a checkerboard pattern that is photographed at different locations and orientations in the image plane. For each frame that includes the checkerboard, a world reference frame is defined at the top left of the checkerboard pattern (Figure 1a, the reference frame has the subscript w and is presented at an offset with respect to its real location) and, by knowing the physical dimensions of the checkerboard's squares (in this case being 45 mm), observations in the world reference frame are created (since the checkerboard is planar, the observations in the world reference frame have only X_w and Y_w components). Moreover, edge detectors in the sensor (i.e. camera) planes (Figure 1a, subscript s_1 or s_2) create observations from two perspectives. By having observations in the world coordinate frame and in the sensor coordinate frame, inferences can be made regarding the intrinsic and extrinsic parameters of both cameras. First, the intrinsic parameters are estimated independently for each camera, by using sets of images with the checkerboard that span different locations and orientations. It needs to be mentioned that since the location and position of the checkerboard is different in each image, the extrinsic parameters also change per image; only the intrinsic ones remain constant for each of the cameras. Zhang's method is able to find a closed-form solution of the intrinsic matrix (per camera) given multiple observations of the checkerboard [27].

After the intrinsic matrix is obtained, it is also possible to obtain the extrinsics for each frame independently (although they do not seem helpful in this application). Once the intrinsic parameters were independently obtained for each of the cameras, it is of interest to obtain the general extrinsic matrices that relate both cameras (the rotation and translation between both cameras). Unlike the previous case where the estimations were performed independently, the rotation and translation between two cameras (R and t in Figure 1a) is obtained by observing the pose of the checkerboard simultaneously from the two perspectives. If the orientations and translations between the checkerboard and the first and second cameras can be obtained for each frame that contains the checkerboard then, since both cameras observe the same frames, it is possible to obtain an estimation of the rotation and translation between both cameras.

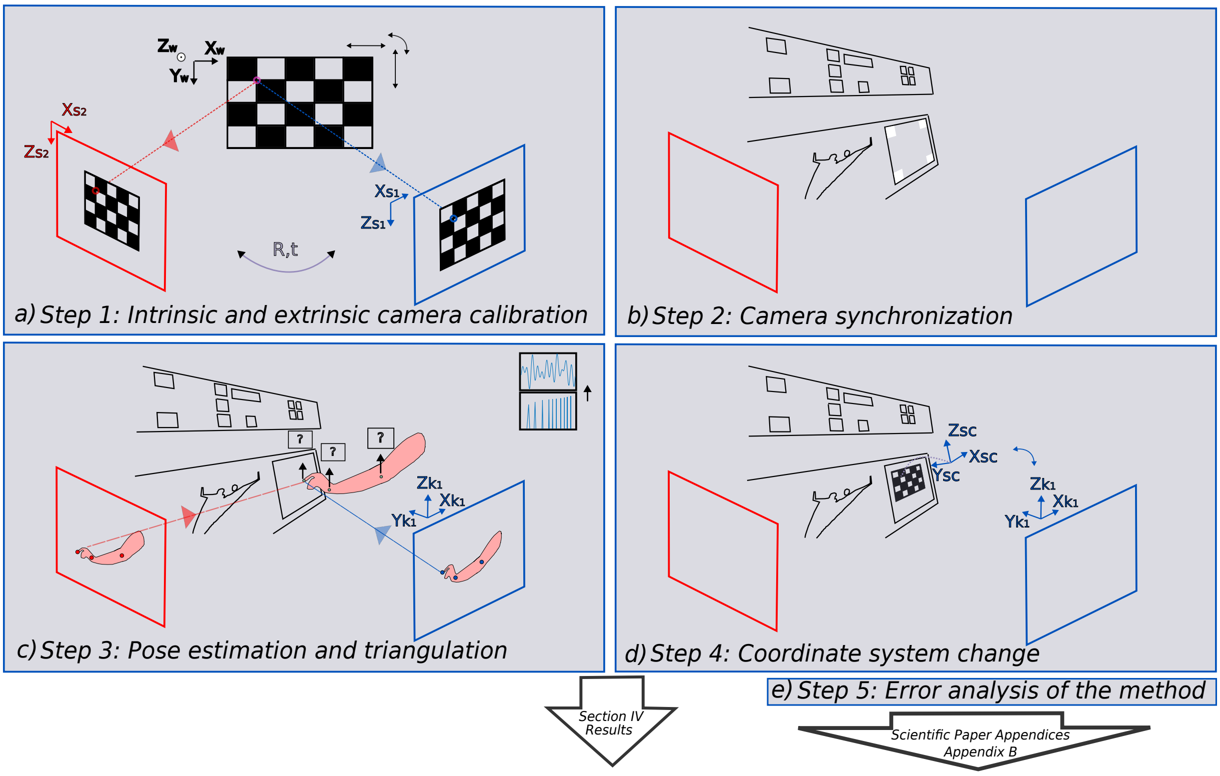


Fig. 1 **Overview of the steps that were performed in order to estimate the 3D locations of the limb landmarks.**

---

[*] https://nl.mathworks.com/help/vision/index.html?s_tid=CRUX_lftnav (Accessed on: 19.03.2024)

2. *Step 2: Camera synchronization*

Application of this approach requires continuous videos from both camera perspectives. Each frame of the videos is extracted creating a continuous list of frames that, at a first glance, do not possess any information regarding when a trial starts or ends. Before inputting the frames in the pose estimator, individual trials need to be isolated, and frames from both perspectives need to be synchronised. The synchronization of the frames is a crucial aspect since the photogrammetric approach assumes that the pixels identified in both perspectives correspond as much as possible. Inspired by [20] who used a light source to synchronize the cameras, this study used a visual cue in the form of four squares that was presented on the screen (Figure 1b) at a discrete time indicating the beginning and the end of each trial. In the post-processing stage, the synchronization was performed by manually selecting the common frames between two consecutive cues for both cameras (representing the start and the end of the trial). For each trial, it was verified that the number of frames between the starting cue and the ending cue was constant.

3. *Step 3: Pose estimation and triangulation*

After obtaining the camera matrices for both cameras and the frames of each trial were extracted, synchronized and undistorted, the pose estimation and triangulation step is performed. The OpenPose [†] software [15] was used to automatically obtain the pixel values of the points of interest (elbow, wrist, finger) from the recorded images. The reason for choosing this estimator is that the stereo applications implemented using OpenPose were investigated in previous validation experiments against the state-of-the-art marker-based estimation methods [19–21], proving its feasibility for estimating the body landmarks in 3D.

OpenPose is a markerless pose estimator that is based on a bottom-up approach [15] for estimating the body's key points. Bottom-up approaches first provide inferences of the discrete body landmarks, and then the inferences are assembled in a skeleton-type detection. These types of approaches can prove to be more efficient computationally when compared to other approaches (such as top-down methods). The architecture of OpenPose is based on Convolutional Neural Networks (CNNs) in combination with a pre-defined model (VGG-19 [28]) which is used in the early stages of the architecture to create an initial set of feature maps. Its architecture is based on two main principles: the creation of a mapping that can encode the orientation of the limbs (partial affinity fields as defined in [15]) and a mapping that can encode confidence maps of the body parts that describe a limb (two body parts can describe a limb). At inference time, the pose is created by using the information from the partial affinity fields and from the confidence maps that define the graph/tree-like structure of the skeleton.

Although OpenPose can record key points of multiple body parts, in this study only the information from the right elbow, wrist and index finger will be used. As observed in Figure 1c), if the pose estimator is applied for the synchronised images from both perspectives, the pixel locations corresponding to identical body parts can be triangulated to obtain the 3D location of the features of interest. The triangulation is performed using MATLAB's *Computer Vision Toolbox (v2023b, Mathworks Inc., Natwick, USA)*, which uses the Direct Linear Transform method (DLT) [26, p. 593] to estimate the 3D position of extracted pixel values from two or more perspectives.

4. *Step 4: Coordinate system change*

In order to make sure that the measurements between experiment sessions were repeatable, and an accurate comparison between the signals obtained from this measurement method and the ones recorded by the touchscreen could be made, the reference frame in which the triangulated points are expressed needed to be changed from a camera-fixed reference frame to a reference frame that is not changing between different experiment sessions. From the calibration step, matrices corresponding to the intrinsic parameters of both cameras and the rotation and translation information between the two cameras (R and t) were obtained. The triangulation was performed assuming that the perspective of the first camera was the center of the world. In other words, the extrinsic parameters of the first camera were set to unity, while the extrinsic parameters of the second one were set to the relative rotation and translation between the cameras. To be able to perform comparisons between experiment sessions (given the fact that the cameras were moved), the triangulated points would most effectively be expressed in a coordinate frame attached to the touchscreen (subscript sc in Figure 1d).

To perform the coordinate system change, a checkerboard pattern was presented on the screen, and frames from both perspectives were recorded. An edge detector was then used to identify the edges of the checkerboard. Four edges of the checkerboard were triangulated, defining the plane of orientation between the screen reference frame (denoted as

---

[†] https://github.com/CMU-Perceptual-Computing-Lab

sc) and the camera-fixed reference frame (denoted as k_1). It needs to be mentioned that the screen reference frame was fixed to the top left point that was identified on the checkerboard, and the camera fixed reference frame was fixed on the optical axis of the camera (not shown in Figure 1d). The approach of moving from the camera-fixed reference frame to the screen reference frame was to first subtract from the triangulated data of the four edges of the checkerboard the triangulated position of the top left checkerboard point. In this way, the triangulated points are still in the camera reference frame, but the top left point is the null vector. Additionally, by knowing the physical distances on the screen between the four points, the same four points were defined in the screen's reference frame, having coordinates only in the Z_{sc} and X_{sc} directions. Now, since the points of interest were obtained in both coordinate systems, Horn's algorithm can be applied [‡] to obtain the rotation matrix that relates the two sets of points. In this way, for the upper limb triangulated data, a translation was applied and then a rotation transformed the data in the screen reference frame. It needs to be noted that in this study, the "vertical" axis in which the data is expressed is in fact the Z_{sc} axis and not the vertical as defined by the gravity direction since the orientation of the screen with respect to the vertical direction is only 18°.

5. Step 5: Error analysis of the proposed method

Figure 2 (a, b, c, d) shows the possible sources of error for the chosen measurement system. Starting from the camera calibration, errors in the intrinsic and extrinsic calibrations might translate into erroneous camera parameters a). A factor that can contribute to erroneous intrinsics estimation is the checkerboard pattern [29]. Firstly, the checkerboard pattern might neither be perfectly manufactured nor perfectly planar. Secondly, the camera may imperfectly detect the edges of the checkerboard due to its limited resolution, edge detection algorithm, or motion blur. The estimation of the extrinsic parameters can also be affected by the previously mentioned checkerboard-related aspects, but also by the fact that synchronised pictures of the checkerboard need to be taken. Errors in the intrinsic and extrinsic parameters affect directly the triangulation since this step requires the use of camera parameters to estimate the 3D location of the pixels observed from both cameras.

In extreme scenarios, any movement of the cameras, b), will invalidate the camera calibration since the estimated rotation and translation matrices no longer have the same values as the ones estimated in the calibration step. As previously described, erroneous camera parameters translate into erroneous triangulation outcomes.

The imperfect synchronization of the cameras, c), can affect the triangulated outcomes through the lost correspondence between the pixel locations observed from the two perspectives. It is known that if two cameras observe the same scene, a correspondence between the pixels from the first perspective and the second one exists, through epipolar geometry [26]. Moreover, an accurate triangulation assumes that the points that are being triangulated correspond in both perspectives. In practice, if both cameras are imperfectly synchronized, one feature that is observed from one perspective might not align with the correct correspondence in the second image, introducing errors that can be quantified by the *reprojection error* [24]. The reprojection error is defined as the Euclidean distance in pixels between the identified point in the sensor frame and the projected point using the perspective projection approach (from 3D to 2D).

A pose estimator might interpret the same feature slightly differently in the two perspectives d). This aspect can be caused by the data that the pose estimator is trained on. This can create a mismatch between the corresponding pixels in the two perspectives and therefore will induce errors in the triangulation that can also be traced back using the reprojection error. A more detailed description and discussion of the method's accuracy will be given in the Scientific Paper Appendices: Appendix B.

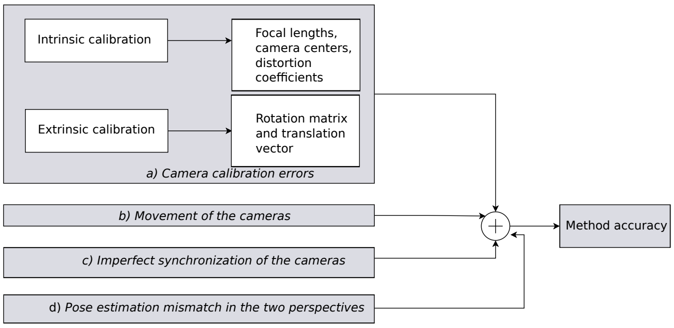


Fig. 2 Overview of the possible sources of error of the stereo pose estimation system.

[‡] https://nl.mathworks.com/matlabcentral/fileexchange/26186-absolute-orientation-horn-s-method (Accessed on: 19.03.2024)

III. Experiment

A. Apparatus

TU Delft's SIMONA Research Simulator was used for the experiment. The simulator is hydraulically actuated using six actuators in a hexapod configuration, allowing for a movement in six degrees-of-freedom [30]. In this research, movement was implemented only in the vertical (heave) direction, simulating vertical turbulence encountered in an aircraft. Turbulence in the heave direction can be considered one of the principal contributors to unwanted movements in aircraft and therefore attracted special attention in academic research [10, 11, 31]. Figure 3 presents an overview of the hardware setup in the simulator. For this experiment, a 15-inch Iiyama Pro-Lite TF1534MC-B1X touchscreen (TSC) was used (Figure 3a), having a width of 304 mm, a height of 228 mm and a resolution of 1024 px x 768 px. The touchscreen was positioned on the right-hand side of the simulator, mimicking the location of the primary flight display, and measuring the inputs of the users at 100 Hz. The study [32] found that the touchscreen latency is dependent on the input speed suggesting that for this study (nominal input speed between 0 mm/s and 50 mm/s) a latency of around 80 ms is to be expected.

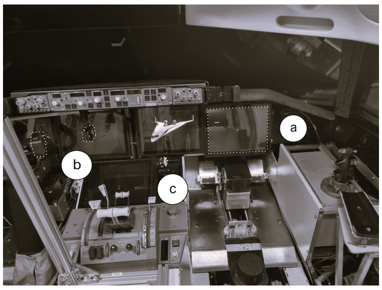


Fig. 3 The proposed set-up, consisting of the touchscreen (a), the two cameras positioned to the left of the subject (b) and the light (c).

Two GoPro (*model Hero 11 GoPro, Inc.*) cameras recorded the subjects at 30 Hz, with a shutter speed $1/120$ s and were firmly placed on the left side of the participants (Figure 3b). The positioning of the cameras was determined via preliminary tests, requiring that the elbow, wrist, finger and touchscreen could be observed from both perspectives while operating the touchscreen throughout the experiment procedure. MSc Thesis Appendices: Appendix C provides an overview of the relative distances and orientations between the two cameras (on average, the Euclidean distance between the two cameras is 40 cm). Lastly, additional light was introduced on the left-hand side of the participant to enhance the quality and contrast of the images of the upper limb in the low-light simulator environment (Figure 3c).

B. Experiment Design

1. Independent variables

The experiment was split into two main parts. The first part had the goal of identifying how the upper limb behaves in turbulent conditions when altering the neuromuscular setting and when performing a set of distinct tasks on the screen, while the second one had the goal of investigating the effect of the different levels of arm extension on the feedthrough observed at the finger level, in a simulated reaching task.

In the first part of the experiment, two independent variables were introduced. The first one (*NMS*) is the *neuromuscular setting*, with two levels: *Stiff* (*S*) and *Compliant* (*C*). The reason for introducing this independent variable was the absence of relevant literature that could explicitly relate the biodynamic feedthrough to the neuromuscular setting attained by the person operating the touchscreen. Moreover, the investigations performed in the field of biodynamic feedthrough for physical side-sticks ([14, 33, 34]) demonstrated a clear influence of the neuromuscular settings in the way the accelerations are fed-through the human body to the inceptor. Nevertheless, it is expected that the subjects will interpret subjectively what "stiff" and "compliant" means.

The second independent variable (*Task*) is the *task*, with three levels: *Finger fixed on the screen at one location* (F), *Finger moving on the screen* (M) and *Finger not touching the screen* (N). In the preliminary report of this MSc thesis [35], it was found that the most common interaction methods between the user and the touchscreen were based on either pointing actions (which implies that the interaction between the user and the screen is very brief) and dragging actions (concerning a longer interaction period) [36]. The author recognises that during the pointing actions, two separate states of the finger can be identified, namely a state where the finger is close to the screen, but not touching it (the *N* condition) and a state where the finger is on the screen (the *F* condition). To identify the particularities of each of the states, the pointing task action was split up into the two levels as presented above.

For each combination of NMS and $Task$, three points on the upper limb (*Body Part*) are measured, namely the elbow (El), wrist (Wr) and index finger (Fi) (Figure 4). To this date, no study which analysed the feedthrough of accelerations of such features on the upper limb in the context of touchscreen operation was found.

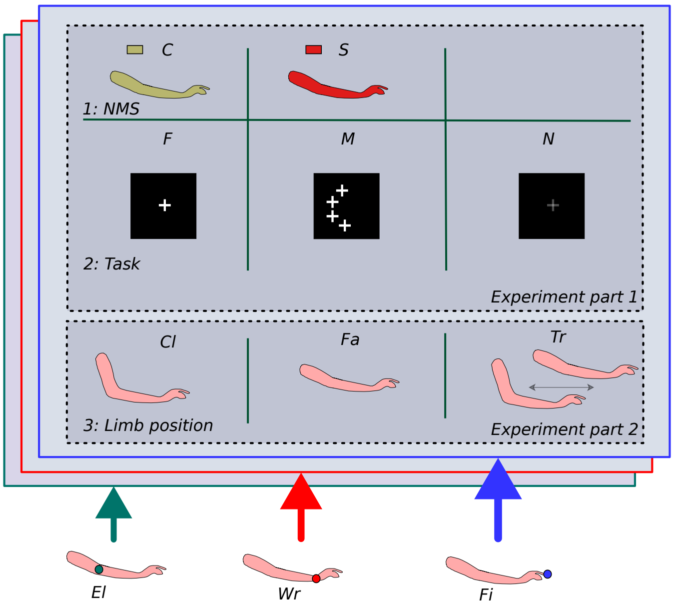


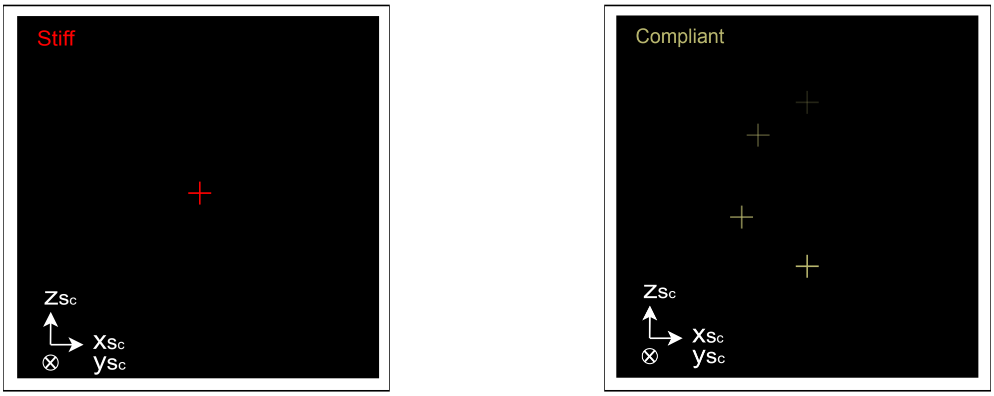
Fig. 4 Overview of the experiment conditions concerning the first and second parts of the experiment.

The second part of the experiment introduced a third independent variable (*Limb position*), namely the *position of the limb* having three levels, *Close* (Cl), *Far* (Fa) and *Transition* (Tr). The purpose of introducing this additional independent variable was to investigate the behavior of the limb in a simulated "reaching" task where it is expected that the turbulence will affect the limb more, the further (more extended) it is. It needs to be mentioned that in this scenario, the subject never touched the screen..

Through the presented independent variables, the goal was to cover most of the interaction methods that the user can face while operating a touchscreen. The analysis of each independent variable can represent a "*detail*" of the complex picture that describes the effects of turbulence in touchscreen operation.

C. Tasks, subjects and experimental procedure

1. Tasks



(a) Finger fixed on the screen + Stiff condition

(b) Finger moving on the screen + Compliant condition

Fig. 5 Overview of the possible display representations. For presentation purposes, the screen reference frame is shown at an offset with respect to its real location.

The design of the tasks was based on the independent variables that were included in the study. First, the interface that the user interacted with contained a colour-based text description (top left of the screen) indicating the required neuromuscular settings that should be attained (Figure 5). The description is relevant for the first part of the experiment since the second part required the subjects to be always compliant.

A cross was projected on the screen, representing the target that the subjects would need to follow with their finger. The cross-shape was used due to its minimal interference with the finger detection from the camera-based pose estimator. Depending on the level of *Task*, the cross either remained still on the screen (Figure 5a), or moved in a circular pattern (Figure 5b). Even though previous studies which analyzed BDFT in touchscreen operations ([11, 12, 37]) have proposed a more complex moving pattern, it was considered that a simpler one would encourage the subject to keep a constant finger velocity, time-invariant neuromuscular settings and would prevent fatigue throughout a trial, aspects that could help in avoiding any possible confounds in the analysis of the data.

It was considered that the combination of the previously mentioned representations (target not moving, target moving and stiff, compliant) would be enough to cover all the cases represented by the independent variables of the study. In this way, for *Task: F* the subjects were required to keep their finger in the region of the cross marker which was not moving (to ensure a consistent neutral steady-state position of the limb which the simulated turbulence would perturb). The condition *Task: M* required the subjects to follow the circular pattern described by the cross as well as possible. In contrast, *Task: N* required the subjects to raise their finger off the screen (in the negative Y_{sc} direction), trying to point to the indicated (static) cross presented on the screen (similar to *Task: F*).

As will be explained in Section III.C.3, the subjects did not need to interact with the screen during the second part of the experiment, and therefore what was present on the screen did not have an influence on the task evolution.

2. Subjects

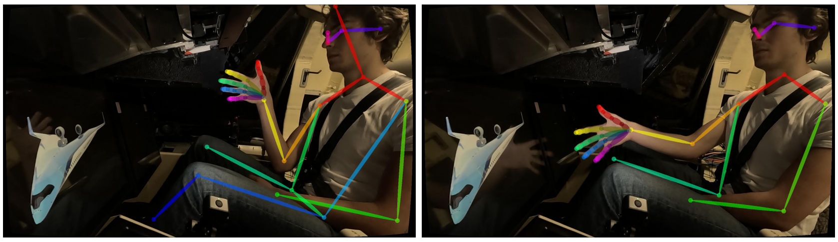
The participants were recruited from TU Delft's student population, consisting of 3 males and 3 females (6 participants). The mean age of the participants lies between 20 - 25 years, the dominant arm was right and all of them had extensive experience in touchscreen operation. The experiment was approved by TU Delft's Human Research Ethics Committee (HREC), *application number* 3648 and was performed according to its guidelines.

3. Experiment Description

The first part of the experiment consisted of the first two independent variables (*NMS*, *Task*), while simultaneously measuring the positions of the elbow, wrist and index finger, resulting in 2 (*NMS* levels) × 3 (*Task* levels) = 6 (conditions). For each participant, the order of conditions was randomized based on a Latin square design, while three repetitions were performed for each condition. The selected number of repetitions was based on the observation that the within-subject variability of a specific condition was low and on constraints regarding the time it would require to perform the data extraction and analysis. The second part of the experiment was performed immediately after the completion of the first part, consisting of one independent variable (*Limb position*) resulting in 3 conditions. The same randomization approach and number of repetitions were applied for this part of the experiment. Before the start of the measurements used for data analysis, the subjects performed a set of six training runs (*F*, *M*, *N*, *Cl*, *Fa*, *Tr*) for familiarisation purposes. Each trial had a length of 36 s, with a 30 s window of measurement and 3 s of fade-in and 3 s of fade-out. The total experiment duration (including the breaks) was 1.5 hr.

Before the participants were invited into the simulator, they were given a set of indications concerning the suggested approach toward the tasks. For the first part of the experiment, the subjects were kindly asked to arrange their seat such that they could reach the touchscreen comfortably with their right arm in an extended position (Figures 6a/6b). Moreover, a rather peculiar positioning of the hand was required, namely the operation of the touchscreen was instructed to be performed with the index finger while keeping the other fingers in the camera's sight. The reason for the previously mentioned aspects is the desire to maximize the identification performance of the pose estimator (in preliminary tests, difficulties were found to occur when not all the fingers were in the camera's sight with the chosen perspectives). Moreover, imposing a specific positioning of the hand implied that the subjects performed the task with a consistent pose, minimizing the confounds introduced by the possible differences in the reference limb positions that could have been attained by each subject [38].

(a) ***Task: F/M*** **(b)** ***Task: N***



(c) ***Limb position: Cl*** **(d)** ***Limb position: Fa***

Fig. 6 Overview of the suggested poses identified from the second perspective.

Part 1 of the experiment: For *Task: F*, the subjects were instructed to touch the screen without applying too much pressure, such that the finger could "glide" on the screen's surface under the influence of disturbances. In this case, the target on the screen represented more of a "reference" for the subject, around which he/she is disturbed. It is important to introduce such a reference to avoid the possibility of the subject actively changing the "trimmed" or "steady-state" position of the limb throughout a trial, affecting the BDFT identification results. The only instruction required for the *Task: M* was that the subjects would need to follow the target as accurately as possible. The condition *Task: N* can be considered similar to *Task: F* with the sole difference being in the fact that the finger did not touch the screen. The subjects were instructed to imagine that they point on the cross marker, but raise their finger away from the screen $2 - 3$ *cm* such that they would not touch the screen throughout the trial. The goal of the instruction was again to induce a steady-state position of the limb that is influenced as much as possible by the disturbance and, in the same time, to minimize the cognitive movement. Moreover, for each level of *Task*, the participants were required to keep either a *Compliant* neuromuscular setting or a *Stiff* one as consistent as possible throughout one trial. For the *Compliant* setting, the subjects were instructed to keep their limb relaxed while performing the trials. Similarly, for the *Stiff* setting, the subjects were instructed to comfortably activate their arm muscles throughout the task duration to resist the motion perturbations.

Part 2 of the experiment: For *Limb position: Cl*, the subjects were instructed to keep their arm relaxed while keeping their elbow and hand close to their body in a steady manner (Figure 6c) for the whole duration of the trial. Moreover, it was recommended that the elbow would be positioned slightly in front of the trunk and above the feet to provide a good perspective for the camera and to avoid any possible interference with any body parts during the application of the disturbance. Similar to *Task: N*, *Limb position: Fa* required the subject to keep their hand steady, and the arm fully extended (Figure 6d) with the only difference being that the palm was facing the cameras since no interaction with the screen was required. Lastly, for *Limb position: Tr*, the subjects were requested to transition between *Cl* and *Fa* continuously throughout the trial at their own discretion with pauses of $2 - 3$ *s* at both the *Cl* and *Fa* positions. It needs to be mentioned that for all the conditions of the second part of the experiment, the subjects were required to keep their arm relaxed.

D. Forcing functions

1. Disturbance forcing function

The disturbance signal was constructed from a sum of sinusoidal signals in order to create a quasi-random simulator movement. Quasi-random disturbances were used in earlier touchscreen BDFT experiments ([11, 12, 37]), due to the possibility of identifying the components in the touchscreen signal that were correlated with the disturbance. This is to facilitate the creation of Linear-Time-Invariant (LTI) models that could describe the feedthrough contribution in the touchscreen's signal. This study approached the design of the disturbance forcing functions characteristics in a similar manner to the one that was used in the previous studies. The biggest difference was in the adaptation of the measurement time, from $T_m = 81.92\ s$ to $T_m = 30\ s$ with implications that the number of times that a sinusoidal "fits" in the measurement time, n_d was adapted to create "peaks" in the frequency domain at the frequency values (ω_d) that matched previous research. A reduced measurement time was introduced due to the requirement that the subjects would need to keep their neuromuscular setting constant throughout the trials.

Disturbance					Target							
k	ω_d	n_d	A_d	ϕ_d	n_{t_z}	ω_{t_z}	A_{t_z}	ϕ_{t_z}	n_{t_x}	ω_{t_x}	A_{t_x}	ϕ_{t_x}
$-$	rad/s	$-$	m	rad	$-$	rad/s	mm	rad	$-$	rad/s	mm	rad
1	0.4189	2	$1.047 \cdot 10^{-1}$	3.3250	5	1.0472	47.7465	0	5	1.0472	47.7465	1.5707
2	0.8378	4	$8.076 \cdot 10^{-2}$	0.3014								
3	1.6755	8	$4.272 \cdot 10^{-2}$	6.0868								
4	2.9322	14	$1.941 \cdot 10^{-2}$	5.4549								
5	3.9794	19	$1.209 \cdot 10^{-2}$	5.4447								
6	5.4454	26	$7.543 \cdot 10^{-3}$	0.3780								
7	7.7493	37	$4.717 \cdot 10^{-3}$	3.9063								
8	10.4720	50	$3.423 \cdot 10^{-3}$	0.2617								
9	13.1947	63	$2.834 \cdot 10^{-3}$	1.4124								
10	17.3835	83	$2.405 \cdot 10^{-3}$	2.2162								

Table 1 Overview of the variables describing the disturbance and target signals.

The choice of the amplitudes (A_d) was based on a desired Root-Mean-Square (or Power) of the signal. Based on [37], the amplitudes were selected such that the acceleration of the simulator had an RMS of 0.75 m/s^2. Lastly, the phases (ϕ_d) were optimised by using a cresting technique as described in [39], in order to avoid simulator movements outside of the acceptable margins. A summary of the disturbance signal characteristics can be found in Table 1 (left side).

2. Target forcing functions

Consistent with the design of the disturbance forcing functions, the same studies have proposed a set of target forcing functions made of sum of sinusoidal signals [11]. The reason for introducing also sinusoidal forcing functions was to "isolate" the power of the screen input corresponding to the following task at discrete frequency values (peaks in the frequency domain), with the goal of obtaining an accurate biodynamic feedthrough identification. In the current study, forcing functions present in the target signal were designed only for *Task: M* condition where the number of sinusoidal signals decreased from three (previous research [11]) to one. The target forcing function being constructed from only one sinusoidal ensured that the subjects could keep a rather constant finger velocity and would limit the effects of fatigue during a trial while facilitating a time-invariant behavior from the subjects in terms of the required neuromuscular setting. The screen inputs two coordinates for the position of the target (vertical z and horizontal x). As it can be observed in Table 1 (right side), the z and x characteristics of the target signals differ only in the phase (a $\frac{\pi}{2}$ difference), therefore creating a target which moves in a circular pattern. The amplitude of the signal was scaled such that a nominal screen (and finger) velocity of 50 mm/s was induced. Lastly, to accurately estimate the BDFT estimate in the frequency domain, the target signal frequency was designed not to overlap with the frequencies of the disturbed signal.

E. Hypotheses

For each task, the subjects are required to attain a "compliant" and a "stiff" NMS setting. Intuitively, a stiffer condition of the arm rejects the applied disturbances in a more effective way, resulting in a smaller movement of the limb. With this in mind, the aim was to investigate whether a difference in the biodynamic behavior of a subject originates from a different way of approaching a task [33]. It is hypothesised that:

H-1: The feedthrough component for all three recorded joints (elbow, wrist, finger) is expected to be more pronounced for the compliant NMS setting compared to the stiff NMS setting for all tasks. (This hypothesis is linked to NMS)

Extensive research has been performed in the area of physical inceptors (i.e. sidesticks), relating BDFT and the NMS setting [34]. For lateral disturbances, it has been observed that by keeping the limb stiff, a better attenuation of the low-frequency accelerations ($\leq 9.5\ rad/s$) is present, while the tight coupling between the limb and the body can induce a higher transmission of higher frequency components ($> 9.5\ rad/s$) when compared to a compliant setting [34]. It needs to be mentioned that the outcomes from the stick-based studies can be fundamentally different from the case when the hand is not in contact with any inceptor since the BDFT dynamics computed in stick-based operations also include the dynamics of the stick [14]. This can affect the frequency range at which certain effects can take place and therefore the following hypothesis will not mention explicit frequency values. In this way, the following hypothesis can be posed:

H-2: The transfer-function representation of the BDFT of all three recorded joints is expected to have a higher gain in the low-frequency range, and a better attenuation of the higher frequencies in the compliant NMS compared to the stiff NMS for all tasks. (This hypothesis is linked to NMS)

During the experiment, the subjects were asked to perform three types of tasks: finger kept on the screen at a fixed location, finger moving on the screen according to a predefined trajectory, and finger away from the screen. Due to the effects of friction between the finger and the screen and due to the assumed differences in the ways the muscles of the upper limb are activated for each level of the task, it is expected that:

H-3: The behavior of the recorded joints under the influence of disturbances is different for all three tasks (F, M, N), with the case when the finger is away from the screen showing the largest variation in the behavior and when the finger is moving actively, the lowest. (This hypothesis is linked to Task)

In [12], it has been pointed out that in certain scenarios when the subjects kept their finger in one location, feedthrough was observed in low amounts since the subjects exerted large forces on the screen, or their finger just pivoted in a discrete location. In this study, it is expected that even in static situations, the finger will be affected by the disturbances, because the subjects were instructed to touch the screen gently to avoid pivoting, facilitating the observation of the feedthrough to the finger location. Moreover, based on the intuition that the subject might be required to use an overall stiffer neuromuscular setting when he/she is moving along the screen, compared to the case when the finger is not moving, it is expected that:

H-4: The finger moving on the screen increases the equivalent stiffness of the limb compared to the case when the finger is kept stationary. (This hypothesis is linked to the interaction between NMS and Task)

Intuitively, it is expected that when the arm is located closer to the body, it will be more directly coupled to it, and the disturbances will affect the finger position less, whereas a larger distance between the body and the finger will have a more negative effect (the finger will move more) due to different equivalent inertia properties of the upper limb, suggesting that:

H-5: The BDFT effects on the finger when the arm is close to the body replicate a "stiff" condition whereas the effects on the finger when the arm is far from the body replicate a "compliant" condition. (This hypothesis is linked to Limb position)

Besides the previously mentioned hypotheses, this study can also take several steps towards understanding the differences within or between subjects. Although not a primary goal (due to the relatively small sample size of 6 participants), [12] found a limited impact of the limb lengths and masses on the feedthrough of accelerations suggesting the hypothesis that the differences in BDFT effects between subjects could be less pronounced than the differences in BDFT within subjects (induced by the change in the neuromuscular settings).

F. Dependent Variables

For the experiment, the following dependent measures were selected:

- General measures:
  - The triangulated *time-traces* of the elbow ($y_{T_{El}}$), wrist ($y_{T_{Wr}}$) and index finger ($y_{T_{Fi}}$) landmarks, as identified by the pose estimator. Since disturbances were applied only in the heave direction, the analysis of the landmarks was made in the vertical (Z_{sc}) direction.
  - The *time-traces* of the finger touchscreen signal ($y_{TSC_{Fi}}$). This quantity can be used for verification purposes, where the signals provided by the screen and by the camera can be compared.
  - The *power spectral densities* (PSD) of the triangulated elbow, wrist and index finger landmarks that are expressed in the vertical direction. These measures are computed using the Fast Fourier Transform. For example, the cross-spectral density between the disturbance signal, f_d and the $y_{T_{Fi}}$ is denoted as $S_{f_d,y_{T_{Fi}}}$.
  - The estimation of the *biodynamic feedthrough* of the triangulated elbow, wrist and index finger landmarks that are expressed in the heave direction. The computation of the *biodynamic feedthrough* at the finger joint consisting of a disturbance *acceleration* f_d, a signal $y_{T_{Fi}}$ (for example) is defined for the disturbance frequencies ω_d as:

$$\hat{H}_{bdft}\ Finger(j\omega_d) = \frac{S_{f_d,y_{T_{Fi}}}(j\omega_d)}{S_{f_d,f_d}(j\omega_d)} \tag{2}$$

For each subject and condition described by *NMS*, *Task* and for every body part, the Fourier transforms of each of the three repetitions are averaged and then divided by the Fourier transform of the disturbance signal to obtain an averaged BDFT estimate.

- Quantitative measures used for data analysis:
  - The *Root-Mean-Square* of the feedthrough component of the time-domain positions of the limbs expressed in the heave direction. For all conditions described by the *NMS*, *Task* and for each recorded body part, the disturbance component of the signals in the time domain ($y_{T_{El_f}}$, $y_{T_{Wr_f}}$, $y_{T_{Fi_f}}$) is extracted by first interpolating the time-trace (from a 30 Hz sampling frequency to 100 Hz), Fourier transforming the signal, isolating only the frequencies at which the disturbance signal acts, and computing the root mean square of the signal that is transformed back to the time-domain. This approach follows the intuition from [34], which assumes that the total signal has a component coming from the disturbance, from the noise and optionally from the intentional movements. For each condition of the experiment, the average RMS of the three repetitions is used as a data point.
  - The *linear regression parameters and coefficient of determination* (a, b, R^2) describing the relationship between the moving average of the finger's normal distance from the screen ($\mu_{y_{T_{Fi}}} Y_{Sc}$) and the moving RMS of the first derivative of the finger's position in the vertical direction ($RMS\ \dot{y}_{T_{Fi}}\ Z_{sc}$). For each data point, i:

    $$(RMS\, \dot{y}_{T_{Fi}}\, Z_{sc})_i = a \cdot (\mu_{y_{T_{Fi}}} Y_{Sc})_i + b \tag{3}$$

    These parameters are estimated for each subject, by combining the data from each trial of the *Limb position: Tr* condition. This analysis has the goal of investigating whether a trend is present between the normal distance of the finger from the screen and the amount of feedthrough that occurs.
  - The *rate at which the finger moves toward or away from the screen* ($\dot{y}_{T_{Fi}}\ Y_{sc}$). This quantity is computed by taking the numerical derivative of the signal $y_{T_{Fi}}$ expressed in the direction that is normal to the screen surface, Y_{sc}. The role of this quantity is to provide a better understanding of the behavior of the subjects in the second part of the experiment, in the transitioning condition.
  - The *estimated coherences* (Γ^2_{bdft}) for each subject, each level of the *NMS*, *Task* and measured body part. The computation of the coherences follows from [37].
  - The *subject-level parameters of a fitted BDFT model*. This paper opted for a mass-spring-damper model (parameters: ω_{bdft}, ζ_{bdft}) in combination with a gain (parameter: G_{bdft}). A time-domain identification method [40] (initial values for G_{bdft}, ω_{bdft} and ζ_{bdft} are $[20, 6, 0.5]$, based on [12]) was applied to obtain the averaged BDFT parameters (from the three repetitions) corresponding to each subject and each level of *NMS*, *Task* and recorded body part. The current study is the first one to try to investigate the suitability of using LTI BDFT models for describing the feedthrough component at three locations (elbow, wrist, finger) on the limb using camera-based signals. The model has the following structure:

    $$H_{bdft}(j\omega) = G_{bdft} \frac{\omega^2_{bdft}}{(j\omega)^2 + 2\zeta_{bdft}\omega_{bdft} \cdot j\omega + \omega^2_{bdft}} \tag{4}$$

    The model structure presented in Equation (4) was chosen out of other linear models (such as a first-order lag or a second-order overdamped system) due to its ability to capture the variable BDFT behavior of the subjects. A comparison between the performance of multiple model structures is present in the MSc Thesis Appendices: Appendix G. Moreover, this model in combination with a delay was selected in other studies [12, 37] that investigated BDFT using touchscreen data. Since in Scientific Paper Appendices: Appendix B it is shown that this camera-based approach is able to capture the delay occurring in the touchscreen, the delay term included in the previous studies was omitted.
  - The *Variance Accounted For* (VAF) corresponding to each NMS, $Task$, and body part. The role of this measure is to quantify how well can a model explain the feedthrough component of each signal. The calculation of the VAF follows from [37], while the final value that is presented for each subject, task neuromuscular setting and body part is the averaged VAF of the three trials that are performed per condition.
  - The *distribution of the finger's signal variance across frequency* ($\sigma^2\ Fi$). This measure is obtained by computing the auto PSD using H_{bdft} and/or $\hat{H}_{bdft}$ at frequency points ω_d, and cumulatively integrating the power spectral density. The role of this measure is to give a better intuition about the frequencies that contribute the most to the variance of the time-domain signal.

IV. Results

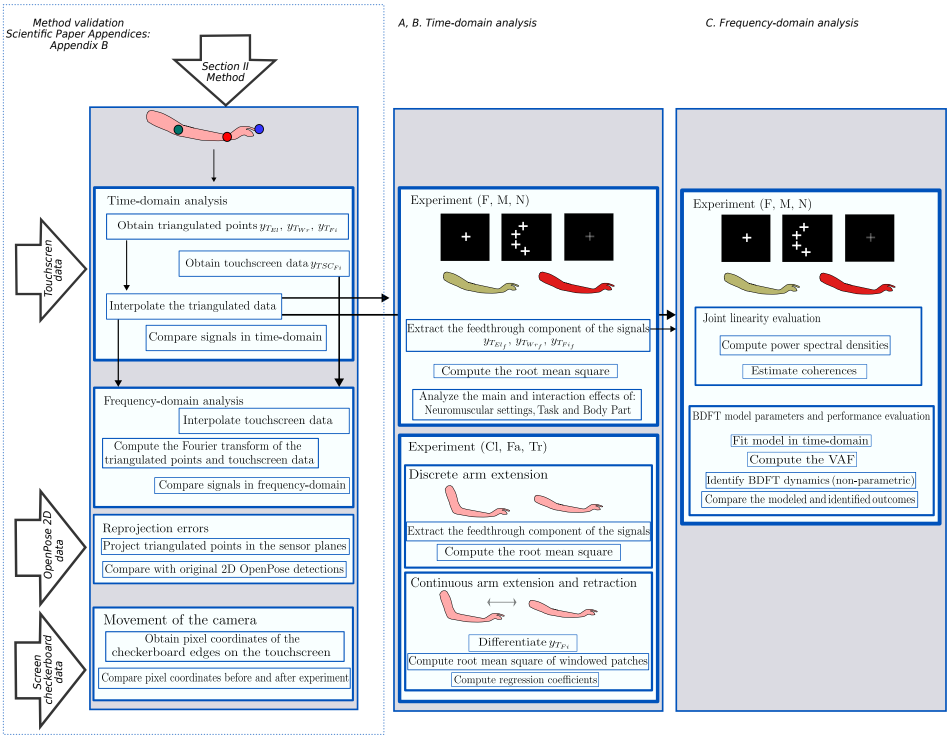


Fig. 7 Overview of the topics analysed in Section IV - Results.

Figure 7 presents a summary of the topics presented in Section IV - Results. The validation of the camera-based approach is presented in Scientific Paper Appendices: Appendix B. The Results section begins with the time-domain analysis of the first part of the experiment, investigating the effects of *NMS*, *Task* and of the recorded upper limb landmark on the RMS of the feedthrough component. The second part of the time-domain analysis focuses on the *Limb position* effects, covering separately the *Cl*, *Fa* conditions and the *Tr* condition. The frequency-domain analysis investigates the feasibility of fitting a second-order system to the feedthrough components of the recorded joints while showcasing different types of behaviors exhibited by the subjects.

A. Time domain analysis - Experiment (F, M, N)

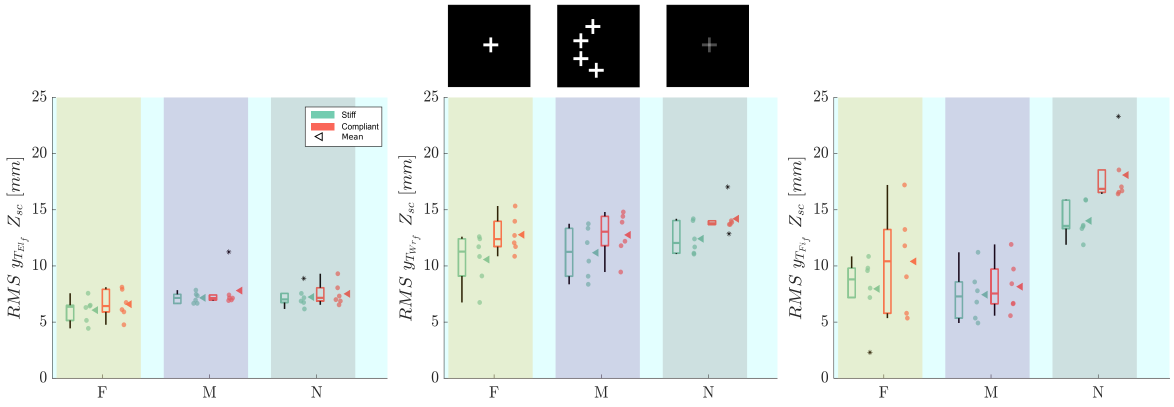


Fig. 8 Overview of the RMS values of the feedthrough component for the elbow, wrist and finger.

Dependent variable	NMS		Task		Body Part		NMS x Task		NMS x Body Part		Task x Body Part		NMS x Task x Body Part	
	$F(1)$	p	$F(2)$	p	$F(2)$	p	$F(2)$	p	$F(2)$	p	$F(4)$	p	$F(4)$	p
$RMS\ Z_{sc}$	7.845	**0.038**	15.056	**<0.01**	35.527	**<0.01**	2.682	0.651	5.480	**0.025**	39.199	**<0.01**	2.240	0.101

■ p < 0.05 (significant) □ p > 0.05 (not significant)

Table 2 Repeated measures ANOVA outcomes for the movement of the limbs for all the conditions described by *NMS*, *Task* and the measured body parts.

1. Effects of the NMS setting on the feedthrough for all different joints.

Figure 8 summarizes the amount of movement that is exhibited as a consequence of the vertical motion disturbance with the RMS of the time-domain feedthrough component of the limb's landmarks (elbow, wrist and index finger) expressed in the vertical direction (screen's Z_{sc} direction). The validity of using only the components that are correlated with the disturbance for the RMS calculation is demonstrated in Scientific Paper Appendices: Appendix A. The outcomes suggest that an important part of the limb's movement is due to the applied disturbances (up to 85% of the total signal variance can be represented by the feedthrough component).

Based on Figure 8, it can be observed that on average the stiff condition exhibits consistently a lower mean RMS compared to the compliant one. At the same time, on average, the difference between the two neuromuscular settings depends on the body part that is being examined and on the task that is performed (especially at the finger level). The smallest difference on average between the stiff and compliant scenarios was observed for the *N* task at the elbow joint (3% difference in the mean RMS between the stiff and compliant scenarios). In contrast, the biggest difference was observed for the *F* task at the finger location (30% difference in the mean RMS between the stiff and compliant scenarios) followed closely by the condition *N* at the finger location (29% difference in the mean RMS between the stiff and compliant scenarios).

Despite the relatively low number of samples (six per condition), a *Repeated-measures ANOVA* was used for the statistical analysis of the data (Table 2). *Kolmogorov-Smirnov* normality tests proved the normality for 16 out of 18 datasets (2 *NMS* × 3 *Task* × 3 *Body Part*). The main effect of the neuromuscular is consistent with the observation made that the mean RMS for the stiff setting is on average less than the mean RMS for the compliant one. A post-hoc analysis in the form of a *Paired-Samples T Test/Wilcoxon Test* was performed for all 18 conditions separately, showing only one significant result for the *N* task at the finger ($p = 0.03$). The significant difference between the stiff and compliant conditions, in this case, may be attributed to the fact that the finger does not make any contact with the screen while having a nominal finger velocity of zero mm/s and therefore it is possible to better control and alter the characteristics of the upper limb (i.e. reject the disturbances).

Table 2 suggests that an interaction is present between the effects of the neuromuscular settings and the examined arm joint. Figure 8 supports this interpretation showing that on average, the relative differences between the stiff and compliant scenarios increase with the analyzed joint (when starting from the elbow). One intuition of this interaction (while omitting pooling across all tasks) can be observed for the *N* task, where the mean of the compliant condition was 3% (elbow), 14% (wrist) and 29% (finger) higher than the stiff condition. This result indicates that for the condition where the subject is not touching the screen, the neuromuscular settings have a larger influence on the joint locations that are further along the kinematic chain (kinematic chain described by the shoulder, elbow, wrist, and finger parts). The task *F* shows a similar increasing pattern in the effect of the neuromuscular settings on the different body parts. A possible explanation for this behavior is the instruction to apply a low pressure on the screen in order to allow the finger to move under the influence of disturbances (and not to pivot around a static location). Nevertheless, given the large spread in the RMS values for the *C* neuromuscular setting, it is likely that the subjects interpreted the task in different ways, therefore pressing on the screen with different force amounts.

On the contrary, the mean difference between the neuromuscular settings for the finger was less than the mean difference for the wrist landmark (9% vs 14% increase between stiff and compliant settings) for the *M* task. One possible reason for this behavior is the fact that the finger is moving with a certain velocity along the screen while also pivoting on it which makes the influence of the neuromuscular setting at the finger level less visible.

At the finger, the same *M* task showed slightly different effects between neuromuscular settings compared to the tasks *F* and *N* (9% vs 30% vs 29% increase between the stiff and compliant cases). The fact that the subjects would need to already have a "baseline" activation of the muscles to perform the *M* task, makes the differences between the stiff and compliant conditions less evident. Despite this observation, no significant interaction between the NMS and the task was present, since the rest of the joints may balance out the overall effect of the interaction.

2. *Effects of different finger interactions with the screen on the feedthrough component*

The same results (Figure 8) are used to analyse the effect of the task on the biodynamic feedthrough component of the arm joints. For example, at the finger level, the most feedthrough occurred during *N*, with mean RMS values of 18.1 *mm* and 14.1 *mm* for the compliant and stiff conditions respectively. The smallest amount of movement was exhibited for *M* with a mean RMS of 8.2 *mm* and 7.4 *mm* for the same conditions. Additionally, by pooling the data across *NMS* and *Body Part*, the mean RMS values for *Task F*, *M* and *N* are 9.1 *mm*, 9.1 *mm* and 12.2 *mm*, respectively. This aspect, confirmed by the significant effect of *Task* in Table 2, indicates that by being in contact with the screen, the upper limb moves less on average (in certain parts) than the case when the upper limb is not touching the screen. This outcome is likely to be caused by the effects of pressure that the subjects apply on the screen (friction and pivoting that occur when the tasks are performed).

It can be observed that while the RMS of the finger decreased on average for *F* and *M* when compared to *N*, the same is not found for the wrist. This aspect is supported by the significant interaction between the task and body part from Table 2. By pooling the data across the levels of *NMS*, the mean RMS at the finger for *M* is approximately 15% lower than for the *F* condition. In contrast, for the wrist, the mean RMS for *M* is approximately 3% higher than for *F*. Additionally, the interaction effect is more pronounced for the wrist and finger in *M* and *N* conditions, where an increase of approximately 100% is observed for the pooled data at the finger and 11% for the wrist. The cause of this effect may be the elimination of contact with the screen that was used for *F* and *M* as a pivot for the finger.

3. *Effects of BDFT on different locations of the limb*

The last insight that can be seen from Figure 8 is how the effects of BDFT vary between the different joints. Pooling the data across all the levels of *NMS* and *Task*, shows that the elbow is moving on average less than the wrist (7.1 *mm* vs 12.3 *mm*) while the wrist moves slightly more than the finger (12.3 *mm* vs 11 *mm*). The fact that on average the finger moves less than the wrist, might be caused by the conditions *F* and *M* which decrease the average finger movement since they involve direct screen contact.

By re-analysing the interaction between the *Task* and *Body Part*, it can be observed that on average, the BDFT at the elbow and wrist is less influenced by the task compared to the finger. Another perspective of this interaction can suggest that for the conditions *F* and *M*, the wrist moves the most on average (mean RMS of 11.7 *mm* and 12 *mm*) while for the condition *N* the finger moves the most (mean RMS of 16.1 *mm*). This aspect reinforces the intuition that when the finger is touching the screen, the wrist is the most affected joint that is measured on the upper limb. On the other hand, when the finger is not touching the screen, the point furthest away from the shoulder moves the most as an effect of the overall inertia of the upper limb.

B. Time domain analysis - Experiment (Cl, Fa, Tr)

1. *Understanding the effects of Close vs Far*

The simplified analysis presented in this section covers two cases, namely when the limb is close to the body (Figure 6c) and when the limb is far from the body (Figure 6d) and is trying to investigate whether the vertical feedthrough component of the finger changes in the two conditions. Possible scenarios that might require different extensions of the limb are when the touchscreen is not positioned far from the body, or when the user is reaching for the screen in a pointing task.

Figure 9 presents the RMS of the feedthrough component at the finger joint for the *Cl* and *Fa* conditions corresponding to the second experiment. The reliability of the results can be firstly demonstrated by comparing the mean RMS for the condition *Fa* (Part 2 of the experiment) and *NMS: C* + *Task: N* (Part 1 of the experiment). The mean RMS for the finger in the former condition is 20.8 *mm* while for the latter is 18.1 *mm*. One possible reason for the differences is that the *Fa* case enables a slightly relaxed approach (the palms were facing the cameras - Figure 6d) when compared to the less natural pose attained in the *N* condition (Figure 6b).

Comparing the *Cl* and *Fa* conditions from the second experiment, it can be observed that a difference between the two arm extensions exists (mean RMS of 12.5 *mm* vs 20.8 *mm*). The outcomes suggest that when the arm is close to the body, the influence of the disturbances is reduced. This confirms the interpretation that the closer the arm is to the body, the more rigid it is, therefore mitigating the effects of BDFT at the finger level.

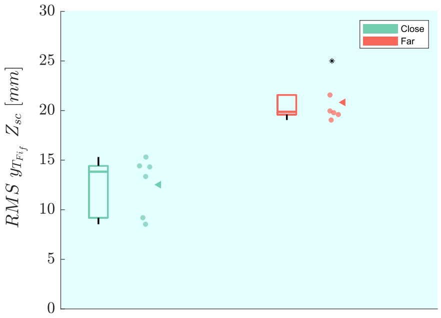


Fig. 9 Overview of the RMS values of the feedthrough component of the finger for the two different levels of arm extension (close to the body and far from the body).

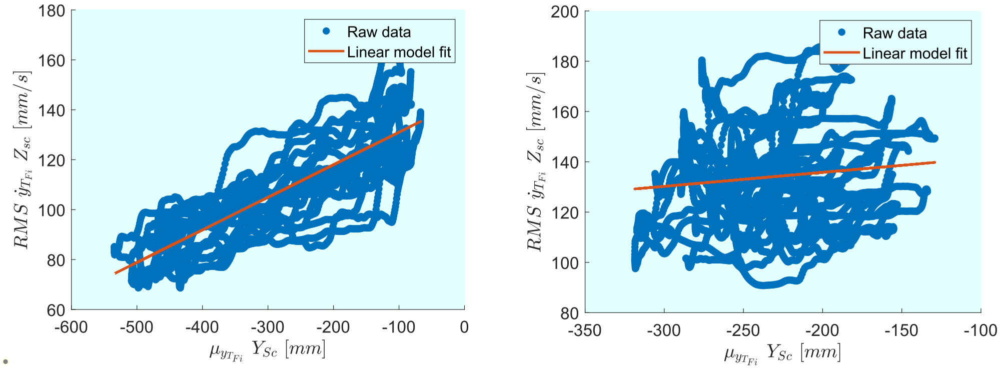
Table 3 Regression parameters for *Limb position: Tr*.

	Regression parameters		
$Subj.\#$	a	b	R^2
	$1/s$	mm/s	
S1	0.14	241.3	0.13
S2	0.15	182.4	0.18
S3	0.06	146.9	0.01
S4	0.13	144.1	0.72
S5	0.20	203.8	0.57
S6	0.12	134.6	0.45

2. Understanding the effects of Transition.

To analyse more realistic tasks, it was of interest to understand what happens with the finger landmark when the subjects slowly transition from a position when the limb is located close to the body (similar to *Cl*) to a position when the limb is fully extended (similar to *Fa*). In this way, the subjects were required to continuously transition between these positions.

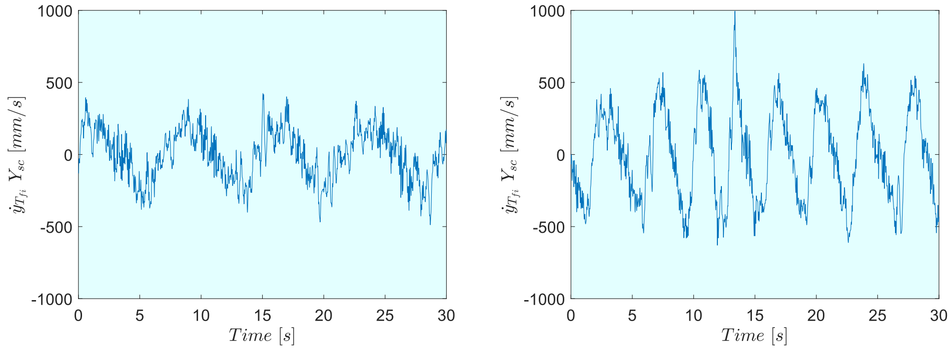
For the analysis of the *Tr* condition, only the finger joint is considered. To eliminate any trends in the data, the finger position (expressed in Z_{sc}) was numerically differentiated. A moving time window of three seconds was used to first obtain an average value of the finger position in the normal direction from the screen (Y_{sc}) to quantify the degree of arm extension and secondly to compute the RMS of the vertical finger velocity.



(a) Subject 4 (high R^2). **(b) Subject 3 (low R^2).**

Fig. 10 The relationship between the arm extension and RMS of the vertical finger velocity.

Table 3 presents the results of a linear model fit to the data of each subject. Weak linear fits were detected between the two variables for $S1$ - $S3$ (suggested by low R^2 values). On the other hand, for $S4 - S6$ possible signs of linear trends could be observed. An example of a linear trend and a trend that shows no correlation can be observed in Figure 10a and 10b. One possible reason for these results is presented in Figure 11, showing that people approached the task differently regarding the speed and also the frequency at which they moved toward or away from the screen (Y_{sc}) direction. The high *change in velocity* (essentially acceleration) showed in Figure 11b eliminates the linear relationship between arm extension and feedthrough since, in the window of 3 s (used for the computations), the arm changes its state (extends/retracts) more rapidly. Additionally, the amplification of noise in the differentiated signals could also affect the regression analysis.



(a) Subject 4 **(b) Subject 3**

Fig. 11 The evolution in time of the arm extension and retraction velocity.

C. Frequency domain analysis - Experiment (F, M, N)

1. Joint linearity evaluation

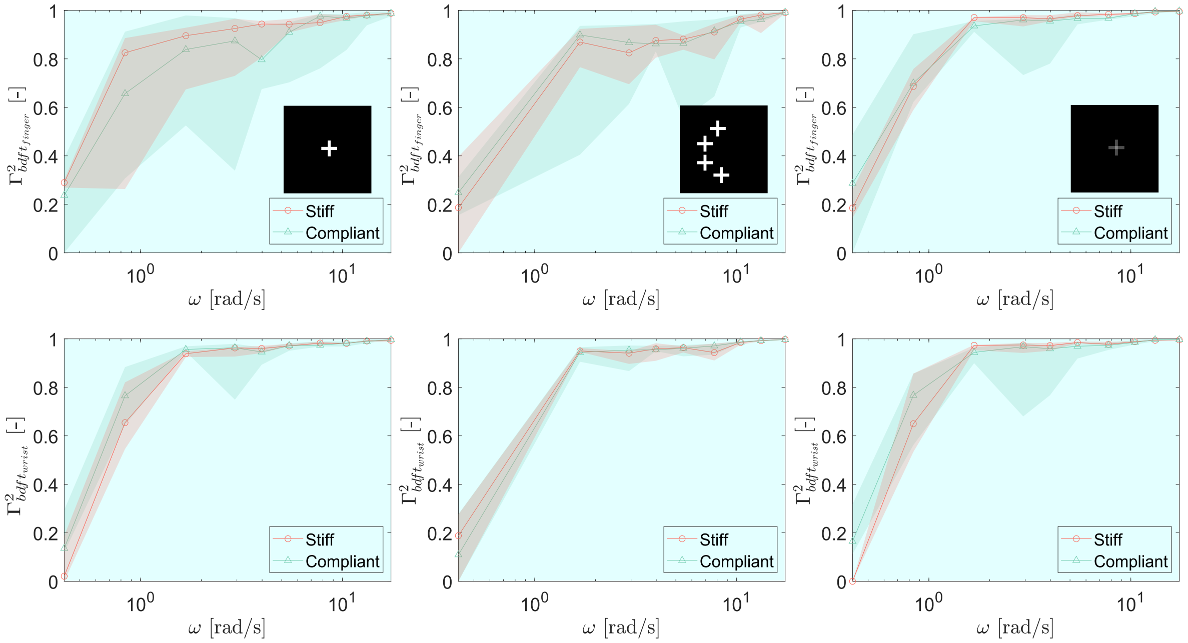


Fig. 12 Overview of the estimated coherence for the two levels of *NMS* and three levels of *Task* of the vertical finger and wrist positions. The medians of the datasets are represented by the continuous lines, while the 0.25 and 0.75 quantiles are represented by the shaded areas.

Figure 12 shows that the coherence values increase with frequency, and a higher coherence is observed on average for the conditions where the subjects are exhibiting a "stiffer" behavior (due to a better "coupling" between the disturbance and the upper limb). The cause for the increase in coherence with frequency might be attributed to the design of the forcing functions and to the rather short measurement time (the lowest frequency sinusoidal occurs only two times in the measurement period). Another point of view is that at low frequencies, the subjects might not respond purely to the induced disturbances, enabling the possibility of other intentional cognitive effects to take place.

An interesting outcome that can also be extracted from Figure 12 is the difference in coherence across the three tasks. When the finger is not touching the screen, a more linear response is measured compared to the case when the finger is touching the screen (either moving or not moving on the screen). The nonlinearities that occur can be caused by the combined effects of friction and pressure that the subject applies on the touchscreen. Lastly, it can be observed that on average, the coherence of the wrist for each task and neuromuscular setting is more consistent than the one of the finger confirming the fact that the BDFT at the wrist joint is not affected by the interaction between the finger and the screen.

2. *Model quality-of-fit*

In this paper, the VAF is used to quantify how much of the feedthrough in the signals is accounted for by the fitted mass-spring damper BDFT model. Figure 13 presents the VAF for all subjects and all levels of *NMS*, *Task* and each *Body Part*. Overall, it can be observed that, in all scenarios mean VAF values of at least 70% were obtained, confirming the feasibility of modeling the feedthrough in the upper limb using the proposed model. The study of [12] presented average VAF values of 90% in the condition that required the finger to move across the screen for the BDFT measured with the touchscreen. The lower performance of the model for the finger joint in *M* presented here can be caused by the elimination of two sinusoidal components in the target signal, the instruction to perform the tasks with the arm close to fully extended (Figure 6a, which was not the case in [12]) or by the elimination of the anti-static gloves that were used to reduce the non-linear friction effects in [12].

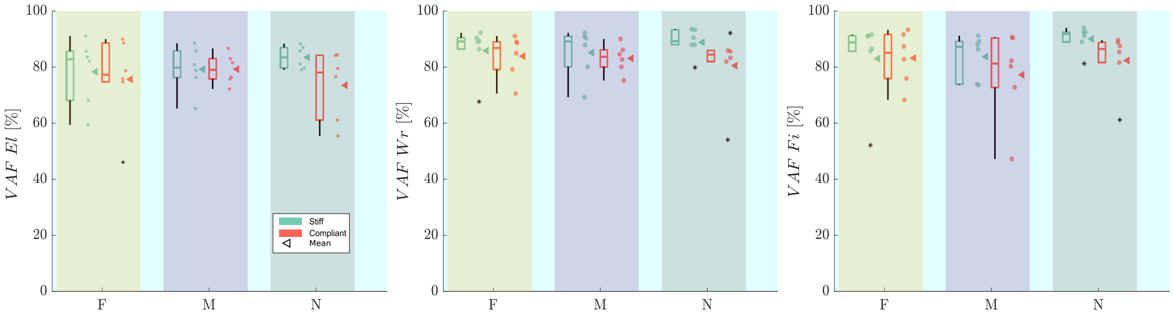
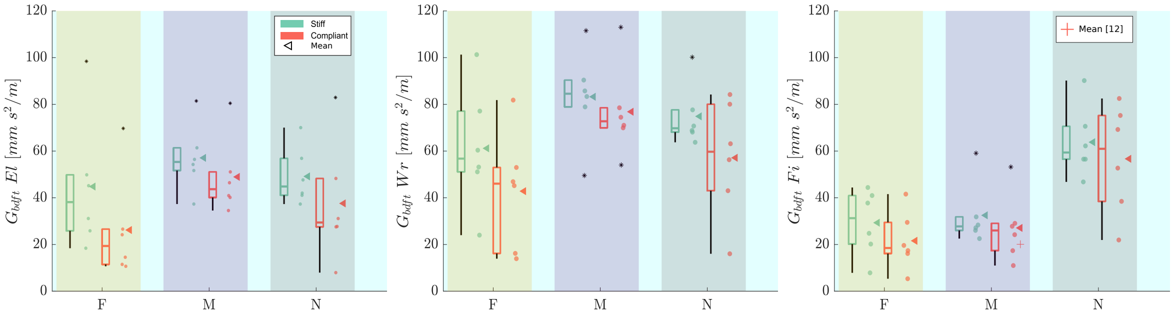


Fig. 13 Overview of the VAF for all measured conditions.

The trends in Figure 13 show that, on average, the VAF values for the stiff conditions are higher than for the compliant case. This result can be explained by the more consistent coherence values for the stiff condition (Figure 12). Additionally, at the finger level, the *N* condition showed slightly higher VAF values compared to the others. The reason for this behavior can also be attributed to how linear the subjects were while operating the tasks. Nevertheless, despite the fact that the interaction with the screen induces nonlinearities, it is still possible to model the feedthrough of the finger with a linear model with sufficient accuracy (e.g. for BDFT cancellation) using the approach proposed in this study.

3. *BDFT model parameters*



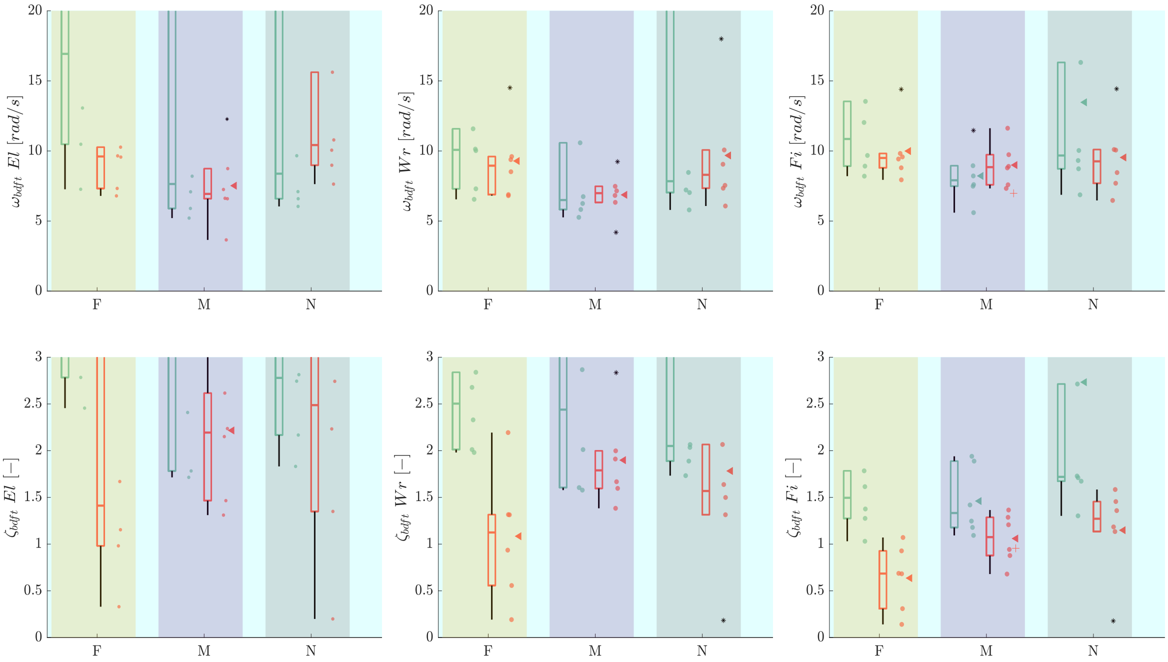


Fig. 14 Overview of the BDFT parameter for all measured conditions.

Figure 14 shows all BDFT parameters that were identified for each condition of Experiment 1. It can be observed that despite the reasonable VAF values obtained in Figure 13, in some cases unrealistic values of the parameters occur, especially for the natural frequency ω_{bdft} and damping ratio ζ_{bdft} of the elbow. One cause of this behavior can be that the model proposed in this study is not always compatible with elbow's response in the frequency range that the BDFT is measured. Another possible reason is the fact that the disturbances do not "excite" the elbow landmark significantly to facilitate identification. The highly unrealistic values tend to decrease when analyzing the wrist and the finger. It also needs to be pointed out that the erroneous values mainly occur for the *NMS: S* condition for which the effects of the BDFT are expected to occur in the higher frequency range. Since no clear conclusion can be taken from only observing the patterns in the data, the author has decided to provide examples that can explain the reason for obtaining these sets of parameters.

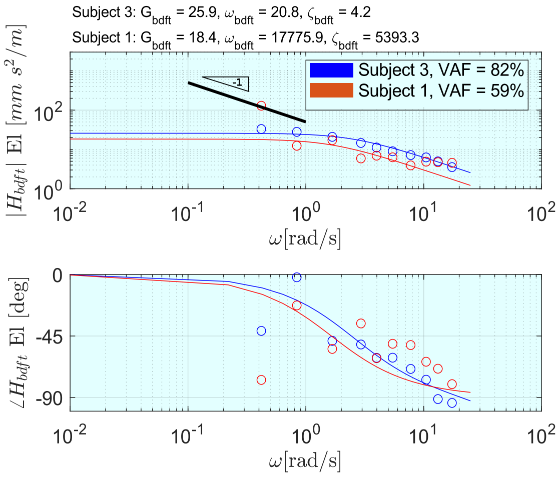


Fig. 15 Elbow BDFT estimation (circles) and model fit (full lines) for two subjects (*NMS: S*, *Task: F*).

Figure 15 presents two overdamped examples corresponding to two subjects that resulted in large estimated values of ω_{bdft} and ζ_{bdft}, but also in different performance estimates (quantified by the VAF). It can be observed that the response of some subjects allows for an accurate use of conventional model structures (such as the one presented in Section III.F) while for others does not. Large estimated values for the two pairs of parameters suggest that two poles on the real axis exist, where one of them has a large negative value (more negative than the largest frequency that is measured). This aspect implies that instead of fitting a second-order system, a first-order lag is describing the data in the frequency range analysed in this study (as can be observed by the phase of almost -90 deg of Subject 3). When analysing the rather poor fit of the model on the data from the first subject, it can be observed that at the higher frequencies, the magnitude of the transfer function does not show yet any decreasing pattern. This aspect can indicate that the complete behavior of the elbow, in this case, can not be determined given the rather limited frequency range at which the BDFT is recorded. The BDFT fit is affected by the limited frequency range at which the phenomena are measured and therefore the parameters can not be trusted 100%.

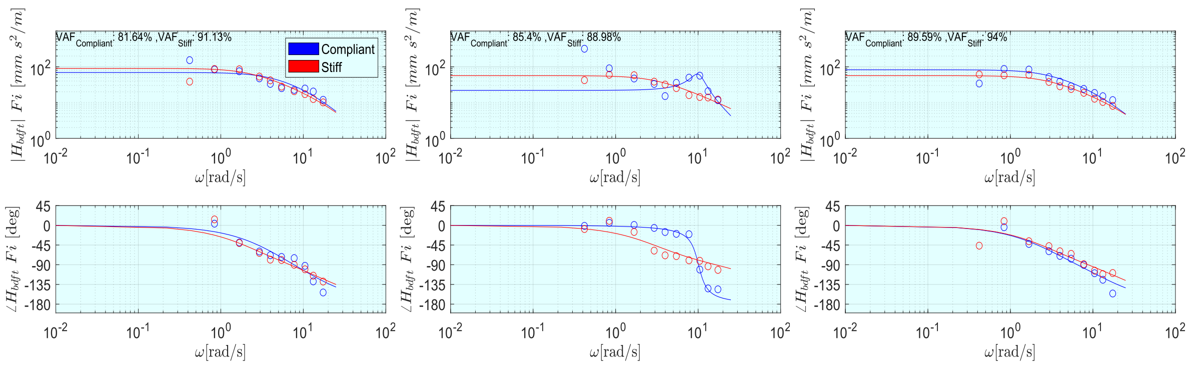


Fig. 16 Examples of finger frequency responses and model fits corresponding to *Task: N* for Subject 4 (left), Subject 5 (middle) and Subject 6 (right).

When analyzing the gain, G_{bdft} corresponding to all conditions from Figure 14 it can be observed that, on average, the stiff condition exhibits a higher gain compared to the compliant condition. This may seem counter-intuitive especially when taking into consideration the outcomes of Figure 8. In order to better understand the effect of the gain parameter, three individual responses of the subjects are shown in Figure 16. These responses correspond to the body part that exhibits the most adaptive behavior (finger) and the task that has not been investigated in any other study (finger not touching the screen).

Figure 16 (left), shows the largest differences between the compliant and stiff settings in the raw frequency response function (circular markers) toward the higher frequencies. It can be observed that at around 1-5 rad/s the compliant setting has a slightly lower magnitude than the stiff case, while at higher frequencies (10 rad/s) the compliant setting has a larger magnitude. The modeled BDFT of the compliant setting (blue) was not able to describe these complex effects, managing only to replicate the fact that at the higher frequencies, there is more BDFT than the stiff setting. In this scenario, the gain of the model for the stiff setting is larger than the gain of the model of the compliant one, while the mean RMS for the stiff condition is 15.9 mm and for the compliant one is 17.04 mm (also suggesting that the overall differences between NMS settings are not that high for this subject). At the same time, the break frequency of the compliant model was higher than the stiff one.

Figure 16 (middle) shows a considerable difference between the stiff and compliant settings, where the effects are larger compared to the previous case. At around 10 rad/s, an underdamped peak is present suggesting that by being compliant, more BDFT occurs around this frequency when compared to the stiff case. It is also interesting to observe that at the highest frequency (17 rad/s), both conditions seem to cross, but due to the limited frequency range, the behavior after the highest frequency can not be observed. The outcomes of the raw frequency response functions are exactly the opposite compared to the outcomes presented by [14], where it was shown that a stiffer condition exhibited higher BDFT compared to a compliant one around 10 rad/s.

Figure 16 (right) shows another type of response, where the gain of the fitted model corresponding to the compliant setting is indeed larger than the gain for the stiff setting. Unlike the previous two cases, the magnitudes of the BDFT for the stiff and compliant settings do not intersect. In this case, despite the apparent small differences between the stiff and compliant frequency responses, the difference in the RMS between the two conditions in time-domain is 5 mm. In the frequency-domain, the difference between the gain parameter of the two conditions is 26 $mm\ s^2/m$.

The goal of these three examples was to convey that the gain parameter does not solely define the amount of BDFT that occurs. It is expected that for a full understanding of the frequency-domain effects, all three model parameters would have to be analysed together while also making sure that the fit describes well the frequency response functions. Moreover, if model-based cancellation is desired and if the identified parameters are able to describe well the effects of the original signal (through a high VAF), the model can be used as a "black box" without seeking for a physical interpretation of the parameter values.

The author finds it important to also discuss what frequency values are the most important when fitting a model to the estimated frequency response functions. One example that supports this statement is the fit in Figure 16 (middle) where, despite the fact that only the last 5 frequency values are described well by the model in the compliant case, a VAF of 85.4% is obtained.

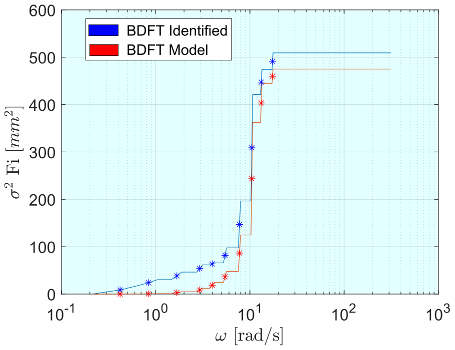


Fig. 17 Example distribution of the finger's signal variance across the frequency, corresponding to *NMS: C*, *Task: N*, finger and Subject 5.

Figure 17 presents the incremental distribution of the signal variance across the frequency spectrum corresponding to the modeled signal and the identified BDFT estimate for Subject 5. This figure shows the frequency values that contribute the most to the variance of the signal. It can be observed that the last five disturbance frequency values (stars) have the biggest contribution to the signal's variance; the increase is the largest there. In this way, if a model is capable of replicating well the high-frequency contents of a frequency response function, it would be able to describe most of the variance of the signal. This result is related to the outcomes of the coherence calculations which also emphasized that the behavior at the higher frequencies is more important for fitting a BDFT model (showing a higher coherence compared to the lower frequencies). Nevertheless, it is expected that this behavior would be highly related to the distribution of the power of the disturbance signal across the frequency spectrum.

V. Discussion

The current study proposed an experiment that evaluated the suitability of using pose estimators for analysing the biodynamic behavior of humans under different scenarios while operating a touchscreen in a turbulent environment. The turbulent environment was replicated by using TU Delft's SIMONA Research Simulator, with a multisine motion disturbance in the heave direction. Additionally, two GoPro cameras and the OpenPose software were used for pose estimation and triangulation. The first part of the experiment (Experiment 1) investigated the effects of the arm stiffness, type of finger interaction with the screen and body part on the observed feedthrough of accelerations. The second part of the experiment (Experiment 2) focused on investigating the effect of the arm extension on the feedthrough of accelerations. The goal of the experiments was to obtain a holistic image of how the upper limb is affected by BDFT in most of the scenarios that concern the user's interaction with the screen. By understanding where the upper limb is most influenced, future adaptive models will be able to counteract the negative effects of disturbances in the most sensitive scenarios. One of the most clear limitations of the study is the rather limited number of participants that took part in the experiment session. The decision on the number of participants was directly related to the amount of time would take to post-process the data. A more efficient data processing/analysis pipeline would allow for larger-scale tests, impacting the scientific value of the conclusions that can be obtained.

A. Discussion of the results

The comparison between the stiff and compliant neuromuscular settings showed a smaller RMS value occurring for the stiff condition at the wrist, elbow and finger and for all tasks (F, M, N). The less than 10% differences between the stiff and compliant conditions at the elbow joint might originate from the fact that either the elbow doesn't move significantly as a consequence of the disturbances, the upper limb muscles are not strong enough to be able to influence significantly the elbow movement or the markerless pose estimation system can not reliably capture the precise (small) movement of the elbow (MSc Thesis Appendices: Appendix H). Overall, it is expected that for a potential adaptive model that can take into consideration the neuromuscular settings, the elbow landmark could be ignored.

The effects of the changes in neuromuscular settings at the wrist were larger than for the elbow (between 14% and 20% increase in mean RMS for the compliant setting). Since the way the finger interacted with the screen did not alter significantly the differences in neuromuscular settings observed at the wrist, the biodynamic information from the wrist could be used in the future for investigating whether the differences in BDFT are a consequence of the anthropometric characteristics (limb length, mass) of the subject or from the task instruction (imposing different neuromuscular characteristics).

The finger exhibited the most different types of responses between the stiff and compliant cases when performing the different tasks (F, M, N). The clearest differences between these two conditions were in the scenario where the finger was away from the screen (*N*). It is assumed that since the friction/pressure effects from the interaction with the screen were non-existent, the subjects could control better their limb in order to show a difference between the stiff and compliant cases. The author certainly considers that neuromuscular-adaptive models would need to be researched for this case since more realistic tasks such as discrete pointing tasks [41] on a screen are mostly affected by the disturbances at the instance in time when the finger transitions from the condition of not touching the screen to the point where the screen is activated at the wrong location. The smallest differences were observed at the finger joint for the M condition (10% difference) due to the fact that the task required the subjects to actively move their finger on the screen. This finding can imply that for continuous (close to constant finger velocity) dragging tasks, neurmouscular-adaptive BDFT models would not be necessary. The results presented above suggest that *Hypothesis 1 which assumes that the feedthrough in the compliant setting is more pronounced for all tasks and upper limb locations, cannot be accepted.*

The outcomes of the first experiment showed that an interaction between the tasks (F, M, N) and the recorded arm joint occurred. By pooling across the compliant and stiff cases, the RMS values for the elbow and the wrist did not change significantly between the tasks (i.e. for the wrist the mean RMS of F was 11.68 mm, M was 11.97 mm, and N was 13.31 mm). This suggests that the movement of these body parts is not significantly affected by the way the finger interacts with the screen. On the other hand, the finger joint exhibited the most adaptive behavior (i.e. the RMS of F was 9.18 mm, M was 7.80 mm and N was 16 mm). For F and M the average RMS of the finger was less than the average RMS of the wrist indicating that the finger is pivoting on the screen when it is in contact with it, and the joint that is affected the most by the feedhtrough is the wrist (being at the middle of the kinematic chain described by the elbow and the finger). The implication of this outcome is that future applications that will track the upper limb, can investigate also the wrist for a complete understanding of the effects between this joint, the finger and the screen.

The finger is most affected by disturbances when it is not touching the screen. This implies that it is important to take into consideration the way the finger interacts with the screen when BDFT mitigation is desired, especially in the case when the finger is on the verge of touching the screen (*N*). Based on the guidelines specified in [42], the large feedthrough encountered at the finger in the *N* condition (average RMS of 18.1 mm in the compliant case), would imply that a touchscreen key of size 13 $mm \times 13$ mm would be missed, impacting the typing and pointing performance. Given the previously-mentioned considerations, *Hypothesis 3 which indicates that the behavior of the recorded joints under the influence of disturbances is different for all three considered tasks cannot be accepted since only the finger joint shows the most adaptive behavior.*

The current study analysed two different types of finger interactions with the screen (F, having a nominal finger velocity of zero mm/s and M having a nominal finger velocity of 50 mm/s). The outcomes from the time-domain analysis showed that the relative differences between the RMS in the stiff and compliant cases for the finger were larger for F than for M (30% increase and 9% increase in the mean RMS of the feedthrough at the finger between the stiff and compliant case). The author considers that F was performed with a high degree of subjectivity (given the large spread in the RMS values for the compliant condition) regarding how much pressure would be applied on the screen. Saying this, it is considered that the effect of finger velocity cannot be decoupled from the between-subject differences in the current experiment and therefore *Hypothesis 4 which indicates that the finger moving on the screen increases the equivalent stiffness of the limb compared to the case when the finger is kept stationary is rejected.* For a better understanding of the effect of finger speed on the BDFT, it is recommended that future studies will make use of pressure sensors for performing the task at predefined pressure levels.

This study also investigated the effects of the extension of the arm on the BDFT and whether these can be measured with pose estimators. For two discrete locations ($Close$ and Far from the body, Experiment 2), lower RMS values were obtained in the condition where the arm was close to the body compared to when the arm is retracted. It is expected that the lumped inertial properties of the upper limb better resist disturbances compared to the case when the arm is fully extended (RMS of the disturbance component at the finger of 12.5 mm vs 20.8 mm). Another perspective of these effects can be suggested by the larger moment arm between the root of the upper limb (shoulder) and the tip (finger) when the limb is extended, shifting the center of mass and altering the moment of inertia of the limb, being more susceptible to disturbances. These aspects *confirm Hypothesis 5, which states that more movement is to be expected at the finger level when the arm is extended compared to when the arm is retracted*, indicating that adaptive models could be introduced as a function of arm extension.

The analysis of the time-varying effects of BDFT during arm *Transition* (Experiment 2) did not result in any clear conclusions. For some subjects, no correlation was present (R^2 values < 0.2) between the mean distance from the screen and the finger velocity, while for others slight correlation was observed (R^2 values > 0.5). Since numerical differentiation is performed to obtain the vertical velocity of the finger, the amplification of the noise can be considered a limiting factor. Moreover, since the applied disturbance was a quasi-random signal, erroneous conclusions can be taken if, at a moment in time, a larger acceleration is exhibited in the period where the finger is away from the screen (where the smallest movement was expected). This analysis can be further improved by testing the effects of disturbances at different discrete positions away from the screen (instead of only two). In this way, by keeping the arm setting stationary for the duration of one trial, it is possible to clearly compare between the movement exhibited at different limb configurations, and frequency response functions could be estimated for each intermediate position.

When analysing the frequency-domain results, the computed coherences for the wrist were higher than for the finger due to nonlinearities induced by the screen friction. Moreover, the linearity of the dynamics of the joints increased with frequency since it is hypothesized that the response of the limb at the higher frequencies is mainly due to the involuntary effects of the disturbances. The rather poor linear behavior at low frequencies can be attributed, at least partly, to the design of the forcing function and limited measurement time (30 s vs 81.92 s [11, 12, 37]). Additional studies are needed to investigate what is the effect of the limited measurement time on the linearity of the system. The fact that the lower frequencies exhibit low coherence can have a negative impact on the reliability of the estimates of the frequency response functions at low frequencies.

The data from Experiment 1 was used for fitting a BDFT model, in the form of a linear second-order system in time-domain. The obtained VAF values showed for the first time that it is possible to use such simple linear systems for modeling, with an acceptable degree of accuracy, the feedthrough component of the limb movement extracted from camera-based pose estimators. This result indicates that using simple linear models, BDFT cancellation using model-based approaches is feasible for a wide range of tasks. The author considers it a remarkable result that the condition with the highest overall feedthrough in the time domain (*N* at the finger joint) could be modeled with such degree of accuracy (average VAF > 80%). Since a large part of the variance of the feedthrough can be explained by a linear model, then it is also expected that model-based cancellation in this realistic pointing scenario to be effective.

The results section also showed some extreme values for the estimated parameters ω_{bdft} and ζ_{bdft} especially for the elbow and wrist landmarks. The results suggest that either some parts of the limb can behave like a first-order lag, or that the frequency range of the disturbance signal was not large enough to capture the full BDFT dynamics of the limb. The possibility that the limited frequency range is not capturing the whole dynamics of the limb is also reinforced by Figure 16, where the BDFT estimates of the stiff and compliant cases tend to intersect at the highest frequency that is measured. It is recommended that future studies investigate the effects of BDFT on a larger frequency range, such as the one defined in [14], which covers frequencies from (0.94 $rad/\ s$ to 157 $rad/\ s$) using the SRS. The increase of the highest measured frequency has a clear impact on the frequency at which the signals are sampled.

The author could not find any relevant patterns in the estimated BDFT parameters for the finger landmark. The MSc Thesis Appendices: Appendix E shows that the averaging of the parameters (as used in this study) can produce erroneous outcomes if the parameter values identified for one of the runs are not consistent with the other runs. Additionally, in some cases, the choice of the initial parameter values influenced significantly the identified parameters (for example, Subject 1, *NMS: C*, *Task: N*). Future research can investigate the use of Genetic Algorithms for the definition of the initial values. One additional reason for the absence of patterns is the rather complex and subject-dependent behavior of the BDFT as observed from the raw frequency estimates in Figure 16. This indicates (as in [12]) that the BDFT parameters need to be adapted to the subject and to his/her subjective approach to the task. One possible option for identifying the time-varying parameters of such a model would be to research regressive BDFT model parameter estimations.

The behavior of the subjects especially for some of the examples given in Figure 16, could not be compared with existing literature. For example, one of the responses (of Subject 5) shows that for a very compliant approach, the BDFT at the higher frequencies (10 rad/s) shows an underdamped peak, where the BDFT for the compliant setting is more pronounced compared to the stiff one (which is exactly the opposite to what [14] found). Given the complexity and variability of the BDFT responses, *Hypothesis 2 which mentions that the transfer-function representation of the BDFT is expected to have a higher gain and a better attenuation of the higher frequencies in the compliant NMS compared to the stiff NMS is rejected.*

By analysing the complete identified and fitted BDFT frequency-response functions for one particular *Task* and *Body Part* (MSc Thesis Appendices: Appendix F), it is not exactly known whether the differences between persons come from the subjective interpretation of the NMS, limitations in attaining a certain neuromuscular setting, or from factors that are linked to the physical characteristics of the limb. The author suggests that in order to better decouple the subjective effects from the anthropometric effects, tests that better measure the quantities of interest such as the limb lengths, masses and muscle activations need to be made.

B. Discussion of experimental setup

The author strongly considers that the proposed approach is feasible to be used in future projects with the condition that several refinements will be made in certain parts of the setup. First, it should be emphasized that markerless pose estimation methods represent the best choice when taking into consideration the limited space that a cockpit or car cabin can provide. Moreover, other alternatives such as marker-based systems [43] or inertial measurement units [44] are more intrusive for the human, making them impractical to use in a real-life setting. The use of markerless systems is also facilitated by the fact that usually the movement of the limb is recorded while the user is seated in the vehicle of interest and the distance between the assumed cameras and the limb does not fluctuate significantly throughout the interaction with the vehicle.

The rather lengthy data analysis pipeline impacted the frame rate and the resolution at which the videos could be analysed. In this study, the videos were recorded at 30 fps and 1080p due to limitations in the number of frames (900 per trial) that would have needed to be analysed in the post-processing stages. The study performed by [20] suggests that the performance of pose estimations can increase with the fps and image quality, therefore the next step would be to investigate the performance of the pose estimation at higher fps (i.e. 60, 120) and with increased image quality. A higher frame rate and image quality can increase the precision of detecting the joints of interest. Additionally, recording the videos at a higher frame rate can provide a better alignment in time of the cameras if the current method of alignment is chosen.

In order to decrease the post-processing time of the data, automatic synchronization of the stereo system and having a close to real-time feed of the camera images on a computer is needed. This can be done by either feeding the GoPro cameras' recording directly to a computer or replacing the GoPro cameras with an off-the-shelf stereo camera (i.e. Leap Motion Controller [45]). Having a stereo camera will further simplify the setup since it can synchronise the videos from two perspectives without additional steps. Moreover, in this study, a rather naive approach was used to feed the data in the OpenPose software. The author suggests that if the OpenPose software is to be used in future projects, a better data pipeline shall be investigated by using OpenPose's Python implementation to feed the images/videos automatically to the pose estimator. The triangulation of the pixel values from two or more perspectives is based on a Singular Value Decomposition (SVD), and therefore it is considered to be computationally inexpensive to be used for close to real-time implementations.

The OpenPose software [15] provides the feasibility of performing the pose estimation in real-time (the authors report a runtime of 36 ms for an NVIDIA 1080Ti GPU). Since this study was testing the feasibility of using such a pose estimator, not all the features (including real-time inference) were investigated. The author finds it important to report several observations regarding the use of OpenPose as a pose estimator. Firstly, this pose estimator is prone to false-positive occurrences due to imperfect lighting conditions that can induce shades on the upper limb (MSc Thesis Appendices: Appendix H). A common false-positive detection was the inaccurate estimation of the wrist landmark due to the shade produced by the upper sleeve of the shirts that the subjects wore during the experiments. It was observed that a better model such as the one defined in [46] was able to reduce drastically the false positive occurrences. Moreover, given the method that OpenPose uses for inferring the pose detection (using tree/graph structures), both ends of a specific body part need to be seen in the image in order to obtain the pixel positions of either of the two landmarks. For example, in order to be able to obtain the wrist position, at least the elbow would need to be observed in the frame. Moreover, it was found that if the wrist/elbow position was wrongly identified, also the hand detection was not performed. This behavior is explained in [47], suggesting that the bounding box of the hand location (which is the first step in the key point estimation of the hand) is approximated using the elbow and wrist positions. Therefore, if the wrist/elbow positions are wrongly identified, no identification of the hand landmarks is available. It is therefore implied that for this method to be reliably applied, careful positioning of the cameras is required, making sure that the upper limb is continuously visible.

OpenPose was designed to efficiently detect the pose of multiple persons in an image. Its authors suggest in [15] that a possible trade-off can exist between speed and accuracy, mentioning the possibility of obtaining higher accuracy from other pose-estimation techniques (top-down approaches) at the cost of lower speed. Nevertheless, it is assumed that in the current applications, only one person would need to be tracked, OpenPose may not be the most optimal choice for the case of one-person estimation. Other pose estimators such as AlphaPose [17] or DeepLabCut [16] could also be viable solutions for future investigations. One possible advantage of DeepLabCut is the flexibility of using transfer learning for tracking key points defined by the researcher depending on their purpose. In this light, if only the finger and wrist joints are to be tracked, a hand pose estimator is worth investigating. Hence, methods/APIs used in Virtual Reality (VR) headsets could be viable alternatives to be used for hand tracking, such as those developed by Apple [§].

Lastly, the author considers that it is possible to improve the accuracy of the 3D estimation. In their book, [26] mention the possibility of using more optimal methods for the triangulation of the pixel values from two perspectives. Moreover, [19, 25] make use of filters such as a Kalman-smoother or Butterworth filter to suppress the jitter that was observed in the temporal OpenPose pose estimations. Additionally, to mitigate the inaccuracies caused by the different detection of the same feature from two perspectives, a Random Sample Consensus (RANSAC) approach can be used for the identified pixel locations by the pose estimator [25, 47].

VI. Conclusion

This study investigated the feasibility of capturing the adaptive biodynamic behavior of six subjects while operating a touchscreen in turbulence. In order to capture the movement of the upper limb (elbow, wrist and finger), a markerless pose estimation technique using the OpenPose estimator in combination with a stereo camera system was proposed. Comparison results against the data recorded from the touchscreen showed that mean errors of at most 1.6 mm and correlation coefficients of at least 0.94 could be obtained for the finger signals when using the proposed approach, proving the feasibility of using such a method for future applications. When analysing the time-domain data of the recorded joints, it was shown that the finger exhibited the most adaptive behavior to the imposed stiff and compliant neuromuscular settings and to the performed tasks. When comparing the differences between the stiff and compliant cases, the task M (finger moving on the screen) showed only a 9% increase in the mean RMS of the vertical finger position, whereas the condition N (finger away from the screen) showed an increase of 29%. This aspect indicates that if a potential neuromuscular-adaptive BDFT model would be applied to the finger, it would need to take into account the type of task that is performed. Moreover, when pooling the data across the two neuromuscular settings, the mean RMS of the finger decreased by 15% from F (finger fixed on the screen) to M and increased by 100% from N to M. These aspects show that the type of contact between the finger and the screen induces an adaptive behavior for the feedthrough at the finger joint. Moreover, another adaptive aspect that needs to be taken into account is the extension of the arm, where the condition of the arm being retracted showed a mean RMS of 12.5 mm whereas the condition of the arm being extended showed a mean RMS of 20.8 mm. A remarkable result of this study is that a simple linear mass-spring-damper system could be fitted for all conditions investigated, showing mean Variance Accounted For values of at least 75%. This result is favourable especially when analysing the task that was not analysed in previous studies and that exhibits the largest feedthrough, N, contributing towards the possible cancellation of the disturbance effects in the more realistic pointing tasks. With our findings, we provide a starting point toward future real-time implementations of adaptive BDFT models that are based on markerless pose estimation methods that can track the wrist and finger joints, to be used for facilitating a smarter and safer interaction with the vehicles.

References

[1] Bloomberg L.P., "Are Car Touch Screens Getting Out of Control?" , 2023. URL `https://www.bloomberg.com/news/features/2023-02-13/are-car-touch-screens-getting-out-of-control?embedded-checkout=true`, accessed: 20.03.2024.

[§] https://developer.apple.com/documentation/vision/detecting_hand_poses_with_vision

[2] Garmin Ltd., "Flight Decks and Displays," , 2024. URL `https://www.garmin.com/en-US/c/aviation/flight-decks-displays/`, accessed: 20.03.2024.

[3] Cirrus Design Corporation, "Cirrus Perspective Touch for the Vision Jet," , 2015. URL `https://cirrusaircraft.com/story/cirrus-aircraft-introduces-cirrus-perspective-touch-by-garmin-for-vision-sf50/`, accessed: 20.03.2024.

[4] Lee, Y.-L., "Comparison of the conventional point-based and a proposed finger probe-based touch screen interaction techniques in a target selection task," *International Journal of Industrial Ergonomics*, Vol. 40, No. 6, 2010, pp. 655–662.

[5] Rouwhorst, W., Verhoeven, R., Suijkerbuijk, M., Bos, T., Maij, A., Vermaat, M., and Arents, R., "Use of touch screen display applications for aircraft flight control," *2017 IEEE/AIAA 36th Digital Avionics Systems Conference (DASC)*, IEEE, 2017, pp. 1–10.

[6] Cockburn, A., Masson, D., Gutwin, C., Palanque, P., Goguey, A., Yung, M., Gris, C., and Trask, C., "Design and evaluation of braced touch for touchscreen input stabilisation," *International Journal of Human-Computer Studies*, Vol. 122, 2019, pp. 21–37.

[7] Albinsson, P.-A., and Zhai, S., "High precision touch screen interaction," *Proceedings of the SIGCHI conference on Human factors in computing systems*, 2003, pp. 105–112.

[8] van Grootheest, A., Pool, D. M., van Paassen, M., and Mulder, M., "Identification of Time-Varying Manual Control Adaptations with Recursive ARX Models," *2018 AIAA Modeling and Simulation Technologies Conference*, 2018, p. 0118. https://doi.org/10.2514/6.2018-0118.

[9] Degani, A., Palmer, E. A., and Bauersfeld, K. G., ""Soft" controls for hard displays: Still a challenge," *Proceedings of the Human Factors Society Annual Meeting*, Vol. 36, SAGE Publications Sage CA: Los Angeles, CA, 1992, pp. 52–56.

[10] Cockburn, A., Gutwin, C., Palanque, P., Deleris, Y., Trask, C., Coveney, A., Yung, M., and MacLean, K., "Turbulent touch: Touchscreen input for cockpit flight displays," *International Conference for Human-Computer Interaction (CHI 2017)*, 2017, pp. 6742–6753.

[11] Mobertz, X., Pool, D. M., van Paassen, M., and Mulder, M., "A cybernetic analysis of biodynamic effects in touchscreen operation in turbulence," *2018 AIAA Modeling and Simulation Technologies Conference*, 2018, p. 0115. https://doi.org/10.2514/6.2018-0115.

[12] Khoshnewiszadeh, A., and Pool, D. M., "Mitigation of Biodynamic Feedthrough for Touchscreens on the Flight Deck," *International Journal of Human-Computer Interaction*, Vol. 37, No. 7, 2021, pp. 680–692. https://doi.org/10.1080/10447318.2021.1890490.

[13] Venrooij, J., Mulder, M., Van Paassen, M. M., Abbink, D. A., Bülthoff, H. H., and Mulder, M., "Cancelling biodynamic feedthrough requires a subject and task dependent approach," *2011 IEEE International Conference on Systems, Man, and Cybernetics*, IEEE, 2011, pp. 1670–1675.

[14] Venrooij, J., Mulder, M., van Paassen, M. M., Mulder, M., and Abbink, D. A., "Relating biodynamic feedthrough to neuromuscular admittance," *2009 IEEE International Conference on Systems, Man and Cybernetics*, IEEE, 2009, pp. 1668–1673. https://doi.org/10.1109/ICSMC.2009.5346935.

[15] Cao, Z., Simon, T., Wei, S.-E., and Sheikh, Y., "Realtime multi-person 2d pose estimation using part affinity fields," *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 7291–7299.

[16] Mathis, A., Mamidanna, P., Cury, K. M., Abe, T., Murthy, V. N., Mathis, M. W., and Bethge, M., "DeepLabCut: markerless pose estimation of user-defined body parts with deep learning," *Nature neuroscience*, Vol. 21, No. 9, 2018, pp. 1281–1289.

[17] Fang, H.-S., Li, J., Tang, H., Xu, C., Zhu, H., Xiu, Y., Li, Y.-L., and Lu, C., "Alphapose: Whole-body regional multi-person pose estimation and tracking in real-time," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022.

[18] Liu, W., Bao, Q., Sun, Y., and Mei, T., "Recent advances of monocular 2d and 3d human pose estimation: a deep learning perspective," *ACM Computing Surveys*, Vol. 55, No. 4, 2022, pp. 1–41.

[19] Needham, L., Evans, M., Cosker, D. P., Wade, L., McGuigan, P. M., Bilzon, J. L., and Colyer, S. L., "The accuracy of several pose estimation methods for 3D joint centre localisation," *Scientific reports*, Vol. 11, No. 1, 2021, p. 20673.

[20] Nakano, N., Sakura, T., Ueda, K., Omura, L., Kimura, A., Iino, Y., Fukashiro, S., and Yoshioka, S., "Evaluation of 3D markerless motion capture accuracy using OpenPose with multiple video cameras," *Frontiers in sports and active living*, Vol. 2, 2020, p. 50.

[21] Zago, M., Luzzago, M., Marangoni, T., De Cecco, M., Tarabini, M., and Galli, M., "3D tracking of human motion using visual skeletonization and stereoscopic vision," *Frontiers in bioengineering and biotechnology*, Vol. 8, 2020, p. 181.

[22] Linder, W., *Digital photogrammetry: theory and applications*, Springer Science & Business Media, 2013.

[23] Arac, A., Zhao, P., Dobkin, B. H., Carmichael, S. T., and Golshani, P., "DeepBehavior: A deep learning toolbox for automated analysis of animal and human behavior imaging data," *Frontiers in systems neuroscience*, Vol. 13, 2019, p. 20.

[24] D'Antonio, E., Taborri, J., Mileti, I., Rossi, S., and Patané, F., "Validation of a 3D markerless system for gait analysis based on OpenPose and two RGB webcams," *IEEE Sensors Journal*, Vol. 21, No. 15, 2021, pp. 17064–17075. https://doi.org/10.1109/JSEN.2021.3081188.

[25] Needham, L., Evans, M., Cosker, D. P., and Colyer, S. L., "Can markerless pose estimation algorithms estimate 3D mass centre positions and velocities during linear sprinting activities?" *Sensors*, Vol. 21, No. 8, 2021, p. 2889.

[26] Hartley, R., and Zisserman, A., *Multiple view geometry in computer vision*, Cambridge university press, 2003.

[27] Zhang, Z., "A flexible new technique for camera calibration," *IEEE Transactions on pattern analysis and machine intelligence*, Vol. 22, No. 11, 2000, pp. 1330–1334.

[28] Simonyan, K., and Zisserman, A., "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[29] Huang, H., Liu, J., Liu, S., Jin, P., Wu, T., and Zhang, T., "Error analysis of a stereo-vision-based tube measurement system," *Measurement*, Vol. 157, 2020, p. 107659. https://doi.org/10.1016/j.measurement.2020.10765.

[30] Stroosma, O., Van Paassen, M., and Mulder, M., "Using the SIMONA research simulator for human-machine interaction research," *AIAA modeling and simulation technologies conference and exhibit*, 2003, p. 5525. https://doi.org/10.2514/6.2003-5525.

[31] Coutts, L. V., Plant, K. L., Smith, M., Bolton, L., Parnell, K. J., Arnold, J., and Stanton, N. A., "Future technology on the flight deck: assessing the use of touchscreens in vibration environments," *Ergonomics*, Vol. 62, No. 2, 2019, pp. 286–304. https://doi.org/10.1080/00140139.2018.1552013.

[32] Vrouwenvelder, S., Postema, F., and Pool, D. M., "Measuring the drag latency of touchscreen displays for human-in-the-loop simulator experiments," *AIAA Scitech 2021 Forum*, 2021, p. 0896. https://doi.org/10.2514/6.2021-0896.

[33] Venrooij, J., Abbink, D. A., Mulder, M., van Paassen, M. M., and Mulder, M., "Biodynamic feedthrough is task dependent," *2010 IEEE International Conference on Systems, Man and Cybernetics*, IEEE, 2010, pp. 2571–2578. https://doi.org/10.1109/ICSMC.2010.5641915.

[34] Venrooij, J., Abbink, D. A., Mulder, M., van Paassen, M. M., and Mulder, M., "A method to measure the relationship between biodynamic feedthrough and neuromuscular admittance," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, Vol. 41, No. 4, 2011, pp. 1158–1169. https://doi.org/10.1109/TSMCB.2011.2112347.

[35] Iancu, A., "Towards adaptive Biodynamic feedthrough modelling for touchscreen applications," *Preliminary Report, Unpublished*, TU Delft, 2023.

[36] Wynne, R. A., Parnell, K. J., Smith, M. A., Plant, K. L., and Stanton, N. A., "Can't touch this: Hammer time on touchscreen task performance variability under simulated turbulent flight conditions," *International Journal of Human–Computer Interaction*, Vol. 37, No. 7, 2021, pp. 666–679.

[37] Leto, G., "Effect of Turbulence Intensity and Variability on Model-Based Biodynamic Feedthrough Mitigation with Touchscreen Dragging Tasks," *MSc Thesis, Unpublished*, TU Delft, 2023. https://doi.org/http://resolver.tudelft.nl/uuid:287c7efb-d15d-4360-a785-c7441746b3f4.

[38] Masarati, P., Quaranta, G., and Jump, M., "Experimental and numerical helicopter pilot characterization for aeroelastic rotorcraft–pilot coupling analysis," *Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering*, Vol. 227, No. 1, 2013, pp. 125–141. https://doi.org/10.1177/0954410011427662.

[39] Pintelon, R., and Schoukens, J., "System Identification: A Frequency Domain Approach IEEE Press," *Piscataway, NJ*, 2001.

[40] Zaal, P. M., Pool, D. M., Chu, Q., Van Paassen, M. M., Mulder, M., and Mulder, J. A., "Modeling human multimodal perception and control using genetic maximum likelihood estimation," *Journal of Guidance, Control, and Dynamics*, Vol. 32, No. 4, 2009, pp. 1089–1099. https://doi.org/10.2514/1.42843.

[41] Dodd, S., Lancaster, J., Miranda, A., Grothe, S., DeMers, B., and Rogers, B., "Touch screens on the flight deck: The impact of touch target size, spacing, touch technology and turbulence on pilot performance," *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 58, SAGE Publications Sage CA: Los Angeles, CA, 2014, pp. 6–10.

[42] United States of America Department of Defense, *Department of Defense Design Criteria Standard: Human Engineering*, MIL-STD-1472G, January 2012.

[43] Figueiredo, P. R. P., Silva, P. L., Avelar, B. S., da Fonseca, S. T., Bootsma, R. J., and Mancini, M. C., "Upper limb performance and the structuring of joint movement in teenagers with cerebral palsy: the reciprocal role of task demands and action capabilities," *Experimental brain research*, Vol. 233, 2015, pp. 1155–1164. https://doi.org/10.1007/s00221-014-4195-3.

[44] Cutti, A. G., Giovanardi, A., Rocchi, L., Davalli, A., and Sacchetti, R., "Ambulatory measurement of shoulder and elbow kinematics through inertial and magnetic sensors," *Medical & biological engineering & computing*, Vol. 46, 2008, pp. 169–178. https://doi.org/10.1007/s11517-007-0296-5.

[45] Naglot, D., and Kulkarni, M., "Real time sign language recognition using the leap motion controller," *2016 international conference on inventive computation technologies (ICICT)*, Vol. 3, IEEE, 2016, pp. 1–5.

[46] Hidalgo, G., Raaj, Y., Idrees, H., Xiang, D., Joo, H., Simon, T., and Sheikh, Y., "Single-network whole-body pose estimation," *Proceedings of the IEEE/CVF international conference on computer vision*, 2019, pp. 6982–6991.

[47] Simon, T., Joo, H., Matthews, I., and Sheikh, Y., "Hand keypoint detection in single images using multiview bootstrapping," *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*, 2017, pp. 1145–1153.

[48] Johnson, P. W., and Blackstone, J. M., "Children and gender—differences in exposure and how anthropometric differences can be incorporated into the design of computer input devices," *SJWEH Supplements*, , No. 3, 2007, pp. 26–32.

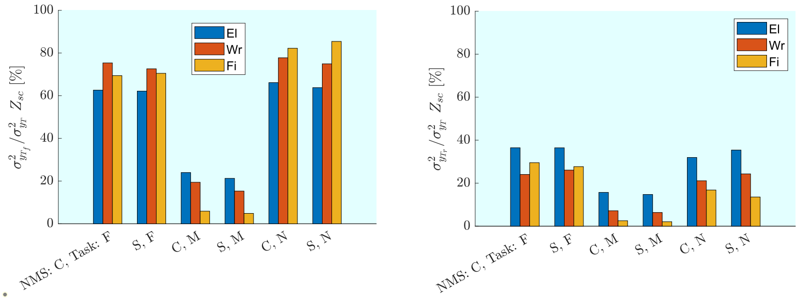
[49] Tung, J. Y., Lulic, T., Gonzalez, D. A., Tran, J., Dickerson, C. R., and Roy, E. A., "Evaluation of a portable markerless finger position capture device: accuracy of the Leap Motion controller in healthy adults," *Physiological measurement*, Vol. 36, No. 5, 2015, p. 1025.

[50] Pezzuti, E., and PP, P. V., "Accuracy in fingertip tracking using Leap Motion Controller for interactive virtual applications. Int," *Jour. On interactive design and manufacturing*, 2016, pp. 1–10.

Scientific Paper Appendices

A. Standard deviations of the disturbance and remnant signal components

The validity of the approach that is being used to extract the feedthrough component from the total signal is conditioned by the fact that the imposed acceleration "excites" enough the limb [11]. Figures 18a and 18b present the averaged (over all subjects) variance percentages of feedthrough and remanant from the total signal variance components present in the triangulated signals ($y_{T_{El}}$, $y_{T_{Wr}}$, $y_{T_{Fi}}$) expressed in the Z_{sc} direction.



(a) Overview of the percent amount of variance of the feedthrough component with respect to the total signal.

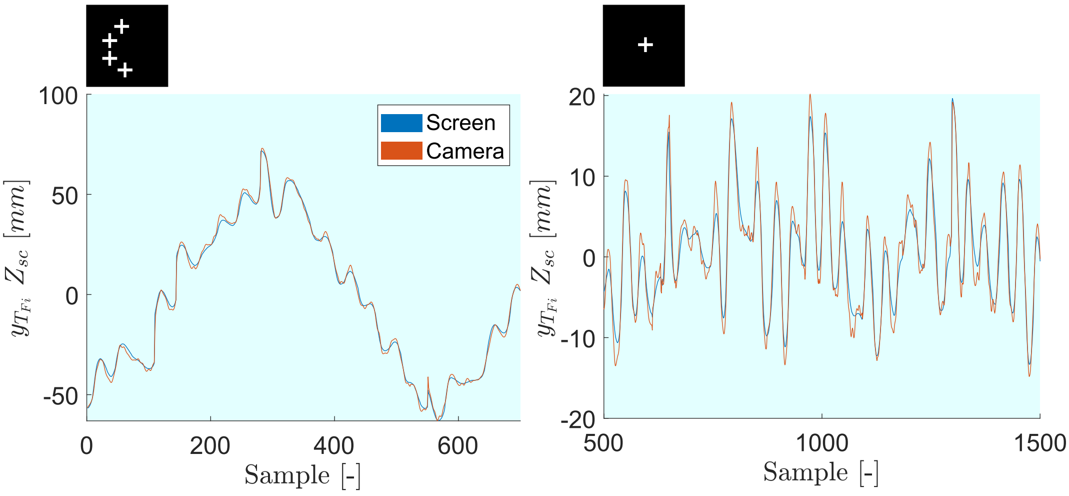
(b) Overview of the percent amount of variance of the remnant component with respect to the total signal.

Fig. 18 Overview of percentage feedthrough/remnant with respect to the total signal variance.

Figure 18 shows that on average, most of the variance of the total signal comes from the disturbance components. Moreover, for the *M* condition, a large part of the variance is due to the intentional movement (not shown). At least 4.8% of total signal variance is represented by the feedthrough component, indicating that it can be used reliably for BDFT analysis [11].

B. Method validation

1. Time-domain analysis



a) *Task: M*

b) *Task: F*

Fig. 19 Overview of the differences between the signals provided by the touchscreen and the camera for a particular subject.

The time-domain verification of the signals obtained using the cameras has the goal of identifying whether this new approach can replicate the effects of the applied disturbances oreviously measured from the touchscreen data [11, 12]. For verification, the signals measured by the touchscreen were considered a relevant baseline that can be used for comparison against the proposed setup. It needs to be mentioned that the verification is restricted only to the finger movement (since an alternative measurement of the wrist and elbow movement was not feasible in this study) and to the *Task: F* and *M*. The comparison between the screen and camera signals was done first by extracting the common touchscreen and camera signal parts where the subjects had no losses of signal from losing the contact with the screen. The time-domain accuracy was quantified by the mean and standard deviation of the absolute error between the two signal sources, and the Pearson correlation coefficient.

Figure 19 presents a comparison between the touchscreen data and the data obtained from the camera and pose estimator. Initially, it was observed that the signals from the camera were leading the signals measured on the touchscreen. The reason for the observed lag in the signals from the touchscreen is represented by the inherent delay between the actual time that the screen is pressed and the time the signal is recorded. The time-domain signals were aligned based on the criterion of maximum correlation. Based on the two plots, it can be observed that the camera is able to follow the general trends that are present in the touchscreen signal, both when the finger is moving on the screen (Figure 19a) and when the finger is stationary on the screen (Figure 19b). Clear differences between the screen and camera signals can be observed when the finger is stationary on the screen, where the tendency of the camera is to slightly overshoot the signal observed at the screen level. Moreover, in Figure 19b between the sample 500 and 1000, the signals obtained from the camera exhibit some high-frequency oscillations. These high-frequency oscillations can be caused either by the slight inaccuracies or jitter of the pose estimator (the finger can be detected in slightly different locations in consecutive frames) or by the filtering techniques used by the touchscreen. The overshoot of the camera signal can be caused by the inaccuracies of the pose estimator in detecting the tip of the finger (from both perspectives), the slightly inaccurate transformation to the screen coordinates or the movement of the cameras throughout the experiment.

Three new measures that numerically quantify the validity of the measurements are defined: $\overline{\rho}$ - The Pearson correlation coefficient between the signal observed at the screen and the one given by the cameras which is averaged over the three trials; $\overline{\mu}_{|\epsilon_{t-c}|}$ - The mean of the absolute sample-wise difference between the signal observed at the screen and the signal extracted from the cameras, averaged over the three trials; $\overline{\sigma}_{|\epsilon_{t-c}|}$ - The standard deviation of the absolute sample-wise difference between the signal observed at the screen and the signal extracted from the cameras, averaged over the three trials.

	NMS: C and *Task: F*			*NMS: S* and *Task: F*			*NMS: C* and *Task: M*			*NMS: S* and *Task: M*		
	$\overline{\rho}$	$\overline{\mu}_{\lvert\epsilon_{t-c}\rvert}$ *mm*	$\overline{\sigma}_{\lvert\epsilon_{t-c}\rvert}$ *mm*	$\overline{\rho}$	$\overline{\mu}_{\lvert\epsilon_{t-c}\rvert}$ *mm*	$\overline{\sigma}_{\lvert\epsilon_{t-c}\rvert}$ *mm*	$\overline{\rho}$	$\overline{\mu}_{\lvert\epsilon_{t-c}\rvert}$ *mm*	$\overline{\sigma}_{\lvert\epsilon_{t-c}\rvert}$ *mm*	$\overline{\rho}$	$\mu_{\lvert\epsilon_{t-c}\rvert}$ *mm*	$\sigma_{\lvert\epsilon_{t-c}\rvert}$ *mm*
S 1	0.9826	1.2708	0.9733	0.9482	1.0533	1.0057	0.9991	1.1484	0.9036	0.9995	0.9857	0.7388
S 2	0.9874	1.6221	1.3882	0.9862	1.3111	1.0066	0.9988	1.3463	1.0553	0.9990	1.3039	1.0273
S 3	0.9829	1.3159	1.2597	0.9810	1.1538	0.9643	0.9988	1.1822	1.1509	0.9989	1.2078	1.0660
S 4	0.9797	1.2932	1.1874	0.9809	1.3492	1.1108	0.9988	1.5033	1.1699	0.9990	1.3426	1.0524
S 5	0.9966	1.4352	1.1513	0.9857	1.3665	1.0046	0.9990	1.2615	0.9641	0.9991	1.3574	0.9857
S 6	0.9810	1.2257	1.0493	0.9752	1.6100	1.3146	0.9992	1.6051	1.1623	0.9993	1.3514	1.0522

Table 4 Summary of the correlation coefficients, mean and standard deviation of the absolute errors between the screen and the camera.

Table 4 presents a summary of the previously mentioned variables, indicating high correlation coefficients for all subjects and tasks (minimum 0.9482 for *S4*, *NMS: C*, *Task: F* and maximum 0.9995 for *S1*, *NMS: C*, *Task: M*). It needs to be observed that an "inflation" of the correlation coefficient is present in the case of *Task: M* due to the high amplitude of the target sinusoidal signal which dominates over the smaller amplitude disturbance components. On the other hand, the correlation coefficient for *Task: F* represents solely the effects of the disturbance, since the finger is not actively moving on the screen. When focusing on the mean of the absolute errors between the screen and camera signals, errors in the range of 1 - 2 *mm* are obtained.

2. *Frequency-domain analysis*

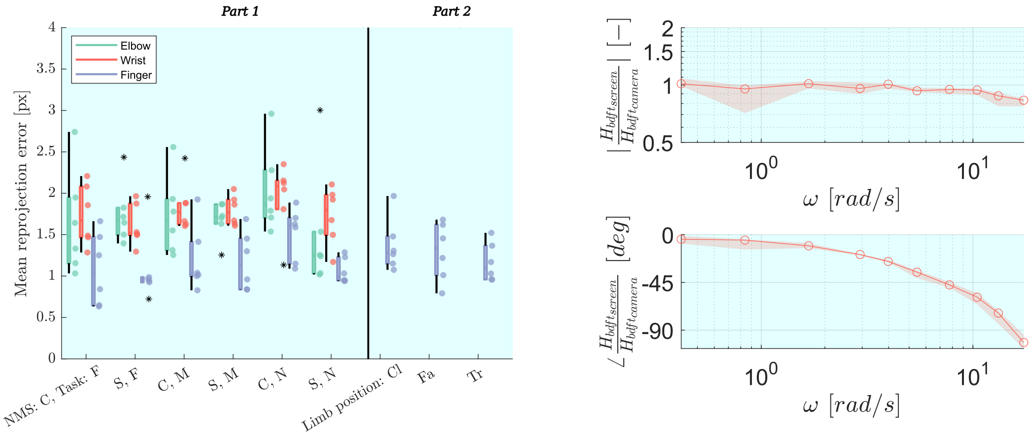
Another perspective regarding the analysis of the differences between the signals obtained from the touchscreen and cameras can be approached in the frequency domain. Of interest is to identify whether the frequency-dependent BDFT estimations are not different when using the screen signals and the signals from the cameras. With this in mind, the BDFT estimations using the two finger movement signal sources were compared and analysed.

Figure 20a presents the outcomes of the ratio between the BDFT computed using the touchscreen signals and the BDFT computed using the camera signals. The signals recorded using the touchscreen were interpolated using $splines$ in order to mitigate the effects of losses of contact between the screen and the finger. Ideally, the presented quantity represents the transfer function that directly relates the camera signals to the screen signals. Intuitively, for consistent measurements from both sources, it is expected that the magnitude ratio to be unity. It can be observed that for low frequencies, the median of the computed relationship almost reaches unity. For higher frequencies, the behavior seems to replicate the one of a low-pass filter. This outcome tends to be consistent with the high-frequency attenuation of the screen signals presented in the time-domain analysis. Nevertheless, currently, it is not known whether the differences arise from camera measurement limitations, or from a possible filter implemented in the screen's processing hardware.

A complementary perspective that also needs to be analysed is behavior of the ratio of the phases. This quantity describes the lead/lag between the signals recorded on the screen and the signals coming from the cameras. Since an estimation of the delay between the two signal sources was of interest, they were not aligned in time (as in the time-domain analysis section). For reference, an experimentally-computed delay of 80 ms corresponding to a finger movement of 50 mm/s is expected from [32]. A pure delay fit (not shown in the picture) of 100 ms was able instead to describe well the phase characteristics of Figure 15. The reason for this difference might be caused by the imperfect/inaccurate synchronization of the cameras that is based on a visual cue compared to the very rigorous laser-based computation of the delay presented in [32] and the additional "lag" that might be introduced by the screen filtering.

3. *Reprojection errors*

The investigation of the reprojection errors can provide additional insights regarding the strengths or weaknesses of the markerless pose estimation method. The reprojection error is obtained by projecting the triangulated point back on the image of the two cameras and then comparing it with the original pixel locations from which the triangulation was



(a) Summary of the mean reprojection errors of the elbow, wrist and finger for the conditions corresponding to the first and second part of the experiment.

(b) Frequency-domain verification. Picture depicting the outcomes of *NMS: S* and *Task: M*. Medians of the dataset are represented by the continuous lines, while the 0.25 and 0.75 quantiles are represented by the shaded area.

Fig. 20 Overview of the reprojection errors a), and frequency-domain bdft ratio between the camrea and touchscreen signals b).

performed. For perfectly-calibrated cameras, and for perfectly corresponding pixel values, the reprojection error should to be close to zero pixels.

The path of the reprojection error starts from the intrinsic calibration of the cameras. Firstly, each of the cameras are calibrated independently in order to obtain the intrinsic parameters (focal lengths, optical centres and distortion coefficients). The resulting estimated parameters lead to a reprojection error of approximately 0.27 px. Secondary to the intrinsic parameters, the cameras are calibrated also extrinsically in order to obtain the rotation and translation between the two cameras. An independent extrinsic calibration was performed for each subject, resulting in an average reprojection error for each subject of around 0.4 px (the baseline value from the intrinsic calibration increased). Saying this, according to [24], there is a high chance that reprojection errors that are higher than the indicated mean values are caused by mismatches in correspondences between pixels observed in the two images.

Figure 20b presents the mean reprojection error for each condition of parts one and two of the experiment. It can be observed that from all the recorded limb features, the finger possesses a lower reprojection error on average compared to the wrist and elbow (maximum 2 px). The reason for this behavior can be explained by the fact that the finger occupies a smaller area in the image when compared to larger body parts such as the wrist and the elbow. In this way, it is expected that there is more uncertainty in the pose estimator correctly detecting the exact correspondence of these features in both images. The presented reprojection errors can be considered acceptable due to the fact that most often, the reprojected points lie on the areas of the image corresponding to the finger, wrist and elbow.

Besides the slight mismatch in the detection of the features in both images, additional factors that can contribute to the reprojection error are the inaccurate synchronisation of the cameras and the movement of the cameras throughout the experiment. In this study, the two cameras were not synchronized electronically, but with a visual cue that it is presented on the screen. In this way, since the frame rate of both cameras is constant, the synchronisation is dependent on the time the recordings were started, allowing for a maximum offset that is dependent on the frame rate of the cameras (in this case, a maximum offset of 30 ms is expected). The movement speed of the limb, in combination with the possible offset in the synchronization of the cameras can also induce the effect of mismatch between identical features that are present in both images (even though the pose estimator manages to detect the features consistently in the two perspectives).

4. Movement of the camera

A factor that can also influence the accuracy of the camera-based measurements is represented by the movement of the cameras with respect to the position for which the extrinsic calibration took place. Section II explained that the intrinsic calibration has the role of estimating the rotation matrix and translation vector between the two cameras. These two quantities are pivotal for a correct triangulation of the pixel positions that are present in both perspectives. Due to the nature of the environment where the cameras are being used, a certain amount of movement is expected, especially in the vertical direction (since a disturbance in the heave direction is applied).

To investigate to what extent the cameras moved throughout an experiment session, frames of a known object were extracted before and after the experiment and further analysed. The selected object is a checkerboard pattern that was projected on the touchscreen. The checkerboard pattern represents a convenient representation that can be used for the evaluation since the pixel positions of its edges can be simply detected with pre-existing edge detectors. Since the screen and displayed checkerboard are firmly fixed in the cockpit, changes in the identified pixel values for consistent locations on the checkerboard before and after the experiment can be used to quantify the degree of movement of the cameras.

Figure 21 shows the checkerboard observed from the perspective of the first camera. This analysis was performed only for the first camera since the fixing method implied that it was more prone to movement than the second one. The analysis focused on investigating the change in pixel locations of two "extreme" edges (Figure 21 - Top left and Bottom right). For each situation (before and after the experiment), 50 frames were analysed and the average of the pixel locations corresponding to each of the two positions was computed, in order to avoid any effects linked to the stochastic nature of the edge detector. Lastly, a direct comparison between the average pixel values before and after the experiment was performed, by computing their absolute difference per direction.

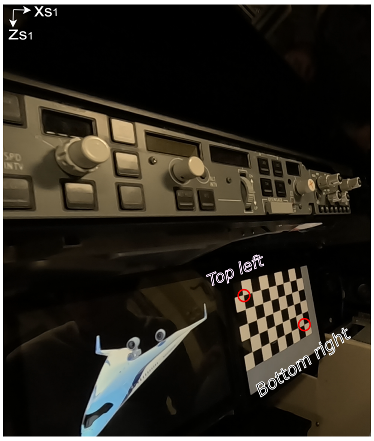


Fig. 21 Visual representation of the checkerboard observed from the first perspective.

	Δx_{s_1}		Δz_{s_1}	
	Top Left	Bottom Right	Top Left	Bottom Right
	px	*px*	*px*	*px*
S 1	0.2408	0.0308	0.6694	0.3318
S 3	0.1832	0.0304	0.6095	0.4249
S 4	0.0305	0.0752	0.3606	0.1407
S 5	0.1178	0.1768	0.7838	0.8048
S 6	0.0430	0.0318	0.3978	0.4549

Table 5 Summary of the change in detected pixel values of the checkerboard before and after the experiment sessions.

Table 5 presents the absolute difference between the detected pixels for the two locations on the checkerboard. At a first glance, it can be observed that bigger differences in the pixel values are present for the vertical (z_{s_1}) direction compared to the horizontal (x_{s_1}) direction (caused by the disturbance in the heave direction). Overall, it needs to be mentioned that the obtained differences seem acceptable, for all conditions differences of less than one pixel being achieved meaning that minimal camera movement existed between the start and the end of the experiment.

5. *Method accuracy discussion*

The validation steps compared the triangulated signals obtained using the camera setup and the pose estimator, with the signals recorded by the touchscreen in order to validate, for the first time, the accuracy of using camera-based markerless pose estimation techniques for the evaluation of the upper limb movement under turbulence. The overshoot of the camera-based signals and the high frequency oscillations could have been caused by the stochastic effects of the pose estimator where, depending on the position of the finger in the image, slight inconsistencies in the estimations of the fingertip point. The time-domain validation showed maximum mean absolute errors of approximately 1.6 *mm* and standard deviations of 1.3 *mm* for the tip of the finger. The obtained mean errors and standard deviations can be considered acceptable, taking into consideration the ambiguity of several millimetres/a couple of centimetres that can be present due to the physical dimensions of the tip of the finger, the study researched by [48] mentioning an average index finger width for an adult to be 20 *mm*. The literature [49, 50] evaluated the accuracy of a commercially available device that can be used for tracking the user's hands without markers (Leap Motion Controller) in discrete pointing tasks, reporting mean errors of $4 - 5$ *mm* [50] and RMS errors of 17.30 *mm* when compared to a marker-based system [49] (which seems sligtly worse than the current method). For a complete evaluation of the required accuracy, the touch-based target size guidelines need to be taken into consideration. For example, the guidelines provided by MIL-STD1472G (2012) [42], suggest that the optimal touchscreen key size should be 13 *mm* x 13 *mm* with a separation of 6 *mm* ¶, suggesting that the errors present in this study are most likely 10% - 20% of the required target size. A limitation of this analysis is that only the finger position was validated, while the wrist and elbow landmarks could not be evaluated by a separate measurement technique. For a more accurate validation procedure, the author suggests the comparison of the markerless system with the state-of-the-art Vicon/Qualisys marker-based systems [19].

[¶] https://www.esa-automation.com/wp-content/uploads/2017/10/04_Guidelines-for-designing-touch-screen-user-interfaces-.pdf

Part II

Preliminary Thesis Report

Note: This part has already been examined under the course AE4030

Literature Study

1

Introduction

The technological evolution in the past several decades has led to an increase in automation of the flight deck (Parasuraman & Manzey, 2010). If in the early times of aviation, the number of mechanical dials was in the hundreds, the introduction of the so-called ”glass cockpit” diminished the complexity of the cockpit and facilitated a reduction in workload leading to a decrease of pilots necessary to be present in the flight deck (Avsar et al., 2016b). Currently, most commercial aircraft make use of touchscreens only in very limited situations, such as data entry in the Electronic Flight Bags (EFB) (Rouwhorst et al., 2017). Nevertheless, touch screen technology for avionics operation is gradually being introduced in the cockpits of future business and civil transport aircraft.

Biodynamic Feedthrough (BDFT) can be defined as *”the transfer of accelerations through the human body [..] causing involuntary forces being applied to the control device”* (Venrooij et al., 2016b). In a less restricted formulation, it can be mentioned that disturbances caused by turbulence and wind gusts are ”transmitted” through the human body into involuntary inputs on a touchscreen. Cockburn et al. (2017) investigated touchscreen-specific tasks in a moving base simulator. It has been concluded that the induced disturbances affected negatively the error rates, the subjective workload, and the time to perform the task. The first study which explicitly analyzed and quantified the influence of BDFT from a cybernetic perspective in touchscreens was Mobertz et al. (2018). The fact that the influence of BDFT on the touchscreen input could be identified in the frequency domain inspired Khoshnewiszadeh and Pool (2021) to use it in a beneficial way. The study investigated the performance of model-based signal cancellation techniques for mitigating biodynamic feedthrough in dragging tasks. Moreover, it was pointed out that disturbances affect different people in different ways, based on their individual biodynamic properties and on their individual interpretation of the task which affects their neuromuscular settings. The same conclusion was observed in an earlier study researched by Venrooij et al. (2011a) which analyzed the same idea of model-based cancellation when using control inceptors. Moreover, an additional insight provided by the same research group is that a relationship exists between the dynamics of the feedthrough of accelerations and how stiff the neuromuscular system is (induced by the task that is being performed). Both studies suggested that in order for the signal cancellation to be effective, the model describing the BDFT would need to be adapted to the individual subject and to the task at hand due to inherent within and between-subject variability.

From the above-mentioned considerations, it can be concluded that BDFT is an ongoing issue related to the use of touchscreens in environments characterized by disturbances. Finding ways of reducing its impact can be an important factor in deciding whether the technology could be more widely used in safety-critical domains such as aerospace, automotive, or naval.

One step towards mitigating BDFT is to gain a better understanding of how the limbs of different individuals are affected by the task that needs to be performed, inducing different characteristics of the neuromuscular system. Having this understanding would allow the development of models which describe more accurately the adaptive character of BDFT. At the same time, adaptive models could be used for a more effective model-based cancellation.

The objective of this work is to investigate the feasibility of creating adaptive BDFT models for touchscreen operation by using motion analysis techniques. The studies researched by Mobertz et al. (2018) and Khoshnewiszadeh and Pool (2021) investigated effects of BDFT starting from the disturbance (model input) that lead to an unwanted effect in the input of the screen (model output). Saying this, all the effects between the input and the output were being lumped together in a form of a general mass-spring-damper system, making it hard to gain a physical interpretation regarding what actually happens with the limb in a disturbed environment. Moreover, the same mass-spring-damper system does not allow grouping the individual differences between and within subjects. This work tries to gain a better intuition regarding what happens between the input and the output of the ”BDFT system” by carefully analyzing the movement of the upper limb. The study will focus on investigating less-intrusive methods and set-ups for motion analysis such as markerless techniques based on image recognition or body-mounted accelerometers. To the best of the author’s knowledge, no study has used motion recording techniques to quantify and understand the effects of disturbances and their interaction with the different settings of the neuromuscular system in the study of BDFT in touchscreens and therefore tries to bring a new perspective to the research of this problem.

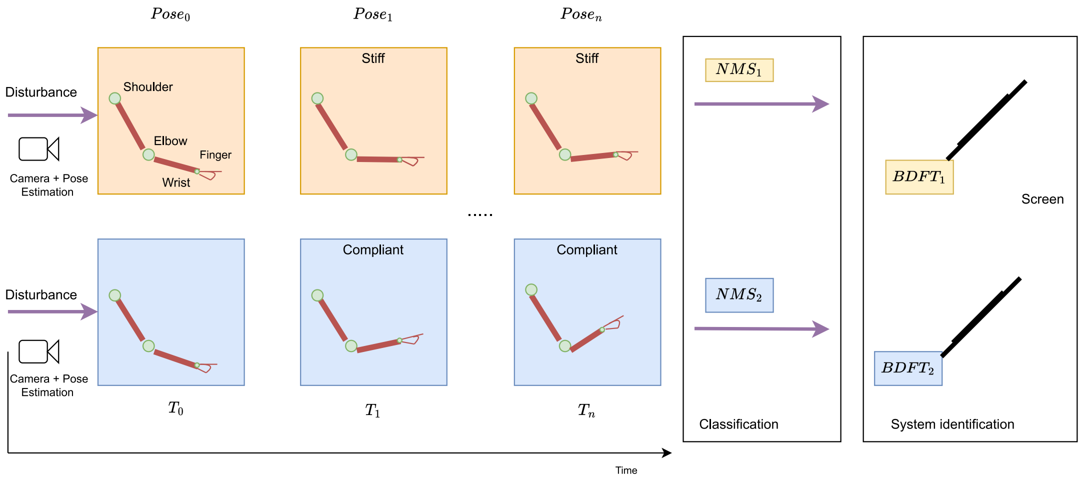


Figure 1.1: Graphical representation of the research goal.

Figure 1.1 presents a graphical representation of the research goal. As previously described, a relationship has been found between how stiff the neuromuscular system is and the resulting biodynamic feedthrough in stick-based systems. This research aims to investigate whether by capturing the motion of several features of the limb, it is possible to obtain any information regarding the variable nature of the stiffness of the neuromuscular system (NMS_1 and NMS_2) and how it induces a different (adaptive) behavior for the feedthrough of accelerations observed at the screen input ($BDFT_1$ and $BDFT_2$). In this way, the feasibility of constructing BDFT models which adapt to the subject’s task interpretation can be investigated.

The report will start with a brief investigation of the use of touchscreens in cockpits and in disturbed environments. Furthermore, an overview of different modeling techniques for BDFT will be given in the third chapter along with information regarding the most recent studies that covered the topic of feedthrough of accelerations through the human body in the aerospace domain. Chapter four will present the advantages and disadvantages of the most common motion recording techniques and will explain a series of applications and validation studies. Chapters five and six will present preliminary implementations linked to multibody systems and stereo reconstruction which can help in capturing the motion of limb features in the three-dimensional space. Lastly, chapter seven will present the research questions, the experiment design, and the hypotheses.

2

Touchscreens

This chapter aims to give the reader an overview of the advantages and disadvantages of using touchscreens as a user interface. Additionally, the types of tasks which are most suitable to be performed with touchscreens will be presented. Lastly, the performance of touchscreens in disturbed environments will be described in both subjective and objective manners.

2.1. Background information

Although the biggest technological advancements in aviation were made just one century ago, today's industry differs greatly in some areas, compared to how its pioneers envisioned it. As the aircraft could fly at faster speeds, at higher altitudes, and could reach farther destinations, more information was needed by the pilots in order to check the state of the aircraft, communicate, and to navigate within an airspace (Avsar et al., 2016b). The cockpits of the first commercial airplanes consisted of hundreds of mechanical gauges and instruments, while each one of them presented one type of information (i.e. airspeed, altitude). Given this, early airplanes required a crew ranging from three to five people to fly the aircraft safely (Avsar et al., 2016b).

The first milestone towards a more automated flight deck was the introduction of the Cathode Ray Tube (CRT) displays and of the Flight Management System (FMS), leading to the first generation of the so-called ”glass cockpit”. One of the first commercial aircraft produced with a ”glass cockpit” is the Airbus A310 (Sweet, 1995). The introduction of this type of cockpit facilitated new ways of displaying and integrating information allowing for a decrease in workload followed by a decrease in the number of people required to operate the aircraft to only two (Rudisill, 2000).

With a more automated flight deck, the role of the pilot gradually switched from ”controller” to ”supervisor” or, if one would like to follow the taxonomy proposed by Rasmussen (1983), from ”Skill-Based Behavior” to ”Knowledge-Based Behavior”. In a more extended discussion, Norman (1989) recognizes that despite the apparent advantages of automation, a crucial aspect that needs to be taken into account is the interaction of the human operator (HO) with automation. The previously-mentioned aspect needs to be taken into account especially when focusing on the limited feedback from the automation to the user. Moreover, the increased automation in the ”glass cockpits” has led to the introduction of the so-called ”automation paradox” (Wiener, 1989), indicating that the workload is reduced in the flight phases requiring a low workload and is increased in the flight phases requiring a high workload, due to difficulties in the interaction with the automation through the cockpit interfaces (Rudisill, 2000). Another interesting point recognized by Rudisill (2000) suggests that one needs to differentiate between manual and mental workload when analyzing the impact of automation. Automation, in this case, decreases the manual workload but can increase the mental workload, due to the fact that the pilots also need to manage the interfaces that they interact with (Rudisill, 2000).

Given the increased supervisory role of the pilots and the need for a more transparent interaction between the automation and the user, the following generations of the ”glass cockpit” have adapted to fulfill these new requirements. The more economical Liquid Crystal Displays (LCD) replaced the older CRTs, in some aircraft a new type of input device was introduced in the form of a touchpad (Abbott, 2017). Moreover, Electronic Flight Bags (EFB) possessing touchscreen technology have been introduced for a more compact way of accessing documents, checklists, and charts (Avsar et al., 2016b).

Touchscreens have become an integral part of people's day-to-day life and found their place in many industries. When focussing on the aeronautical domain, touchscreens were first introduced in the military sector (Lockheed Martin's F-35 (Fletcher & Huffman, 2010)). Moreover, advancements made in touchscreen technology have enabled it to slowly transition toward the highly regulated civil aviation domain. In the business jet sector, Gulfstream [1] has proposed a novel flight deck with 10 touchscreens that can be used for flight plan alterations, selection of the desired displayed contents, and overhead panel interaction (Watkins et al., 2018). Companies in the commercial aviation sector such as Boeing and Airbus follow certain standards which dictate the requirements for cockpit systems (Cockburn et al., 2017). One example of such a standard is ARINC661 (2016), which mentioned for the first time in 2016 the guidelines for the inclusion of touch inputs in cockpit display systems (Cockburn et al., 2017). Recently, “Boeing” (2016) (being the first one) and “Airbus” (2019) have announced the introduction of touchscreens in their newest long-haul products Boeing 777X and Airbus A350 XWB for data-entry tasks and chart and map navigation.

Looking back, it can be observed that the cockpits of commercial airliners have evolved from a very complex and workload-intensive environment with a large number of physical switches and buttons to a more automated and supervisory-focussed one. The switches are being slowly replaced by large screens with high information density content which, as presented, in the future cockpit designs will make more and more use of touch technology. One important question that needs to be answered is regarding to the advantages that touch technology would bring to the cockpit environment but also the drawbacks of this technology.

The ability of users to engage directly with the system they are interacting with is one of the touchscreens' key benefits (Lee, 2010). In other words, it is possible to interact with a system without the need for an intermediate element such as a touchpad or physical keyboard. Moreover, as Rouwhorst et al. (2017) concludes, under some circumstances, with carefully designed user interfaces, touchscreens can decrease the physical workload and speed up tasks in the cockpit. Touchscreens can be considered an intuitive interface to use since this type of interface has been present around the user in the form of smartphones or tablets for the last two decades (Watkins et al., 2018). From an operational point of view, the contents presented on the screens and the configurations can be rapidly exchanged by changing the interface with the user without the need of reconfiguring the cockpit's hardware (Cockburn et al., 2019). User's interaction with changing menus is accounted for by the interface design, without the need of including physical buttons for each type of command and menu available. On the other hand, the use of touchscreen displays might be influenced when the user's limb obscures parts of the display or when the dimensions of the targets are much smaller than the finger width (Albinsson & Zhai, 2003). Another possible disadvantage mentioned by van Zon et al. (2020) is that a reduced accuracy might be encountered in the use of touchscreens if tactile feedback is missing. Adding to this, without an additional source of feedback that can

---

[1] Gulfstream is an American business jet manufacturer.

provide information regarding the finger's location on the screen, the user would need to obtain this information from the visual system by alternatively focusing his/her eyesight from the task at hand to the touchscreen and vice-versa. In this way, it is possible that the apparent mental workload would increase and the user can be more prone to performing errors, due to multiple sources that would need attention from the user. Another possible disadvantage is represented by muscle fatigue which can increase in case the arm of the user needs to be held unsupported and in contact with the screen for a prolonged amount of time (Wang & Trasbot, 2011). Addressing the topic of this literature study, one of the most important disadvantages when discussing touchscreen interaction is represented by the accidental inputs (Degani et al., 1992) which are influenced by the environment through which the aircraft flies. Due to the direct input/output mapping of the touchscreen technology, a disturbance would lead to an accidental input. This drawback needs to be extensively researched before applying the technology to a wider scale in such a regulated, carefully designed, and safety-critical environment as the cockpit of an airplane.

This section presented information regarding the future of touchscreen technology which is slowly being introduced in the cockpits of the new airliners. It is still interesting to investigate whether this is the most suitable technology to be used in this case, and if not, in which scenarios can this technology perform better compared to other alternatives. The answer to these questions will try to be answered in the next section, which investigates the use of touchscreens as an alternative for being a user interface in cockpits.

2.2. Using touchscreens as user interface

One of the first scientific reports which compared the touch input with other devices for flight-deck applications is Jones (1990). The study analyzed three input methods: trackball, touchscreen, and a speech recognition system for discrete-pointing tasks such as subsystem display, warning or caution message clearing, and numerical entry. The experiments were performed in a static scenario using a cockpit from a fixed-base simulator. For each experimental task, objective measures such as task completion time and the number of errors were recorded while the questionnaires represented the subjective measures. The results have shown that the touchscreen possessed the lowest mean completion time across all tasks. The reason for this is the straightforward way of operating the system and performing the task. When focussing on the interaction between the pilot and the input concept, it can be mentioned that most of the pilots performed the best (the fastest) with the touchscreen concept. Nevertheless, not all the pilots performed the best with the touchscreens, the reasons being the acceptance/preference or the individual characteristics of the subjects. For the interaction between the input concept and the task, it was concluded that the touchscreen concept was well suited for tasks that involved ”on-off” operations such as selecting a subsystem. The results regarding the input error data have shown that the touchscreen concept resulted in the most recorded input errors, especially in the case of entering a reference altitude. It was hypothesized that the errors were due to the size and spacing between the available touch zones. Nevertheless, it is important to point out that subjectively, the preferred input method was dependent on the type of task being performed with the touchscreens representing the best solution for tasks involving *”on/off operations”* (Jones, 1990). Taking a step back, it needs to be mentioned that not a single concept performed the best for all the tasks and all the dependent measures. When comparing different input methods, the student considers the task design a very important factor since for some input methods the task design can provide ”shortcuts” which can alter the conclusions.

The above-mentioned study suggests two main points that the student considers important and that might be recurrently encountered throughout the report. Firstly, it has been observed that the touchscreens have been able to complete the tasks in the lowest average duration, but with many errors. It is expected that this phenomenon is not a coincidence, the literature describing it as the ”speed-accuracy tradeoff” (Norman, 1991). Saying this, the lower accuracy with a faster movement in combination with cluttered touch areas might be the reason for the increased error metric. Secondly, it was observed that no input method performed the best in all the tasks. This aspect is along the lines of the conclusion made by Rogers et al. (2005) mentioning that the selection of the input device is dependent on what specific task is required from the operator.

Stanton et al. (2013) recognized (based on Harris (2011)) that in the cockpit of an airplane, not only the cognitive aspects need to be taken into account when analyzing HO’s interaction with different input devices, but also the physical ones. Moreover, since the flight deck is an environment where multiple tasks need to be performed at the same time, a ”*holistic*” approach which contains both subjective and objective measures is recommended for the analysis in order to get a complete view on performance. Compared to the initially presented study, Stanton et al. (2013) analyzed four different input devices, namely: trackball, rotary controller, touchpad, and touchscreen. As previously mentioned, the study used objective measures such as interaction time and error rates and subjective measures such as workload (NASA-TLX questionnaire NASA (1986)), usability (System Usability Scale, Brooke et al. (1996)) and comfort (Cornell University Questionnaire for Ergonomic Comfort, Hedge et al. (1999)). A hypothesis of this study stated that the best performance over all the measures in discrete target acquisition tasks would be obtained by the touchscreens. The results showed that the touchscreens performed the best in terms of task times, whereas the rotary controller exhibited the least amount of errors. As previously mentioned, there is a possibility for a trade-off between the time the task is performed and the thoroughness of the task (amount of errors). Moreover, the touchscreens exhibited high evaluations for physical workload and effort, high total body discomfort but high evaluations for usability. For the attentive reader, these outcomes might contradict the statements of the first section, where it was mentioned that automation can decrease the physical workload. In this case, one reason for the increased physical workload and body discomfort is the necessity of the user to stay in contact with the screen for a prolonged duration (in the case of long tasks). Nevertheless, it needs to be mentioned that this outcome (physical workload) highly depends on how the touchscreen is positioned with respect to the users. The study concluded that even though the touchscreens have performed the best according to most of the measures, no input device performed consistently well across all the variables (in line with the previous study). When analyzing HO’s interactions in the cockpit, it has been recommended that a combination of interaction methods might represent the best solution. For example, touchscreens could be used for short tasks with high time pressure (tasks performed quickly while minimizing discomfort). Given this, it was highlighted that it is crucial to understand where and how the interface will be used in order to evaluate and understand its suitability.

The study researched by Eichinger and Kellerer (2014) tries to fill the body of knowledge not specifically accounted for by the previous two studies, that is analyzing different interaction methods in more realistic contexts. Even though previous studies recognized the importance of taking into account the context in which an interface is used when analyzing its suitability, none of them has explicitly accounted for additional informatory circumstances that the pilot can encounter. In a real scenario, the HO might need to perform several tasks at the same time, which can alter his/her performance when interacting with the elements in the cockpit. This study has introduced additional visual, cognitive, or motor tasks to be performed concurrently with the main discrete pointing task (possessing two levels of difficulty) when comparing touchscreen and trackball interaction. Even though the results of the study were in line with the previous ones (in terms of completion time and workload), it has been shown that different performance and workload outcomes were obtained for each additional task that was introduced. If the focus is put on performance, it can be mentioned that an additional task stimulating cognitive load has the least effect on this measure when compared to visual or motor tasks. On the other hand, if the focus is put on workload, an additional task stimulating motor load has the most detrimental effect. Moreover, when directly comparing the two methods, it was mentioned that additional motor tasks affect consistently in a negative way the performance of the trackball input. In the case of the touchscreen interaction, it was observed that in some of the cases, the performance was affected by the additional visual task. This behavior can be due to the need of the user to re-direct his/her eyesight towards the additional task. One possible limitation of this study is that it has not presented the additional tasks jointly, but separately. An interesting next step would be to combine the cognitive, visual, and motor demands in a single ”additional” task.

As part of the new generation of cockpits, manufacturers have envisioned a new concept for the Multi-Function Control and Display Unit (MCDU) using touch input (Wang et al., 2015). With this regard, the mentioned study analyzed the effects of the touch-key shape and size on MCDU usability. The subjects were faced with typical data-entry tasks (discrete screen inputs), being evaluated on the time-to-complete, errors, and also on a Likert scale which indicated the subjective usability. Concerning time, it was observed that the increase in key size facilitated a more rapid task completion. This is possibly due to the fact that a lower accuracy is needed for a larger key and therefore the speed of the task is increased. Moreover, as expected, an inverse trend was obtained for the error rate, where the larger key sizes had the smallest errors when compared to the smaller keys. Lastly, when comparing different key shapes (square and rectangle), it was observed that the square keys gave consistently lower error rates (with changing size) compared to the rectangle ones. The reason for this might be the decrease in the surface area of the square keys compared to the rectangle ones. Nevertheless, it is important to recognize that despite the advantages given by an increase in key size, this design choice might clutter the interface and might be susceptible to accidental touches to neighboring keys since the distance between two keys might decrease for a constant screen size (Wang et al., 2015).

Studies such as Thomas (2018), van Zon et al. (2020), or Xie et al. (2023) have analyzed the performance of touchscreens using Fitts' law. Thomas (2018) analyzed the performance of touchscreens using discrete pointing tasks in multiple directions (*ISO 9241-9* (2000)). It was observed that on average the throughput obtained by the touchscreen was the highest out of all input devices ($3.7 [bits/s]$). Moreover, it was observed that the touchscreen had a wide spread for the average error rate metric. This behavior was explained by the accuracy reduction when interacting with small targets. Xie et al. (2023) and van Zon et al. (2020) have included more realistic tasks such as heading selection, CDU data insertion, and touch-based navigation for the analysis of accuracy and throughput models. The conclusions of these studies confirm the findings of the previous ones concerning the direct interaction capabilities of touchscreens. Moreover, confirming what was previously mentioned, the size of the target is a relevant factor for the accuracy obtained in the touchscreen operation (smaller targets, poorer accuracy).

Most of the studies presented in this section analyzed and compared the use of touchscreens in static situations. Nevertheless, in a real-life scenario, airplanes/helicopters fly through disturbed environments, presumably affecting the HO's interaction with the elements inside the cockpit. Therefore, it is interesting to investigate, how disturbances affect the touchscreen operation, and which tasks are the most or the least affected in this case. These topics will be investigated in the next section.

2.3. Touchscreens in disturbed environments

As previously mentioned, the role of the HO has slowly transitioned from controller to supervisor. If, in the previous section, the accent was put mainly on airplane cockpits, it needs to be mentioned that a similar trend can be observed also in other scenarios such as automotive or maritime. All these application domains have in common one aspect namely, they are all characterized by disturbances (similar to when a car drives on a bumpy road). Saying this, given the increased interaction with advanced interfaces using touch technology, it is important to study how these interfaces perform in comparison with the classical interaction methods in disturbed environments and in which cases their use would be suitable.

One of the first studies which compared the use of multiple input devices such as touchscreen, mouse and trackball in vibration environments is Lin et al. (2010). The study focussed on the analysis of these devices using discrete pointing tasks (Fitts (1954)) under realistic filtered sea-wave vibrations with a Root-Mean-Square (RMS) acceleration of $0.22\,[m/s^2]$ and $0.34\,[m/s^2]$. The discrete Fitts' tasks concerned squares with different sizes, different distances between each other, and different angles with respect to each other (Figure 2.1).

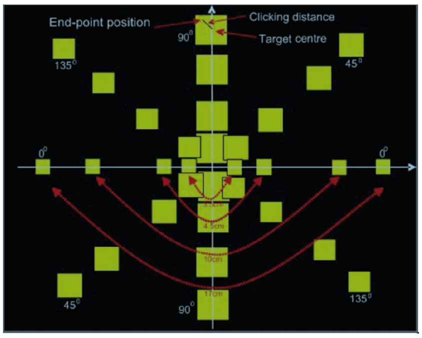


Figure 2.1: Discrete pointing task (Lin et al., 2010).

When analyzing movement time, it was observed that in general, it increased with increasing levels of vibration, with increasing distance between the targets, and with decreasing target size (Lin et al., 2010). Nevertheless, it was observed that touchscreens exhibited the fastest movement time under all vibration conditions due to their direct interaction capability. On the other hand, the error rate was significantly affected by the interaction effects between the pointing device and the vibration level. Based on this, it was shown that the error rate metric in the case of vibration for touchscreen interaction increased at a higher rate than for the other devices. Additionally, the error rate was significantly affected by the interaction between the pointing device and the target size. In line with the previous studies, the error rate decreases with increasing target size. One interesting measure that was analyzed in this study but not specifically reported in the others is the standard deviation of the endpoint location. It was reported that the standard deviation of the endpoint location for the touchscreens was the highest out of all other input methods, and it increased with the vibration level. The study suggested that a closer look at how the devices are operated is needed in order to understand their differences in performance under vibrational environments. It was reported that the trackball had better resistance to vibrations due to the forearm being rested on a hard surface, which attenuates part of the whole-body vibrations. On the other hand, the touchscreen was operated by moving the upper arm and the forearm (there is no mention of the wrist) which are being affected by the whole-body vibration in the case that no additional support is used. The understanding of which body parts are used to perform tasks using touchscreens can be used in the later stages of the report when several modeling approaches for the interaction between the arm and the disturbances will be investigated.

Compared to the previous study, the research performed by Dodd et al. (2014) introduced a more realistic task, namely a data-entry task similar to a waypoint entry task in an EFB. In the same time, the subject was instructed to keep the controlled aircraft in the simulated scenario at a steady velocity and altitude. The objective of the study was to investigate subjective workload, subjective fatigue, and the muscle activation of the subjects when faced with turbulence in touchscreen operation. As in the previous study, it was observed that the tasks took longer to complete in simulated turbulent conditions, and turbulence had a more negative impact on the smaller key sizes. When subjectively evaluating fatigue, it was reported that the smaller key sizes induced more fatigue throughout the arm. Electromyography (EMG) measurements of the forearm were used to quantify the muscle activation of the subjects, being observed that smaller key sizes and turbulence increased the muscle activity. Unfortunately, since the study does not report details regarding the turbulence signal (reporting only ”none” or ”moderate”) and/or RMS/peak acceleration values, it is difficult to interpret the realism of the scenario that was researched.

Taking a step back, besides the inherent response of the human body to external accelerations, human performance in disturbed environments can be affected by an impairment of human vision, proprioception, or vestibular reflexes (Dodd et al., 2014). Even though this report focuses on only one cause, namely vibration feedthrough, McLeod and Griffin (1989) propose two additional mechanisms which may influence the performance of the operator. One of them is ”visual impairment” (caused by the relative movement between the eyes and the display). Regarding the second one, ”neuromuscular interference”, it is hypothesized that *”such interference [between the vibration and neuromuscular processes] may reduce the signal-to-noise ratio between the intentional, task-related muscle activity and unintentional or random activity and may lead to perceptual confusion about the forces generated in the controlling limb”* (McLeod and Griffin (1989) p.p. 80). Saying this, it can be the case that when the subjects need to perform intentional movements (exhibited by the Central Nervous System (CNS)) it can sometimes be difficult to differentiate between the response inherently caused by the feedthrough of accelerations and the response caused by the other factors which also affect HO’s performance in disturbed environments.

The study performed by Avsar et al. (2016a) investigated the HO’s performance when operating touchscreens under conditions of increased G-force in the vertical direction. Figure 2.2 presents a simplified schematic regarding the biomechanics of the arm movement. This representation was used by the researchers in order to indicate the possibility of replicating prolonged g-forces in static conditions.

The same figure can be used to gain an intuition regarding possible solutions for modeling and understanding the biomechanics of the upper arm. In order to keep the arm fixed at a specific configuration (without any support), the muscles of the arm complex would need to create a torque (around a certain joint) that is used to counteract the weight of the limb. Moreover, two parameters in this figure can be subject-dependent, namely the moment arm (induced by different lengths of the limb) and the effective mass of de limbs (induced by the specific body type of the subject). With this in mind, it is expected that the biomechanics of each individual to be slightly different. A limitation of the figure presented by this study is that, depending on the situation, a torque would need to be applied at all the arm joints in order to reach the desired movement of the arm, since each segment has its own weight.

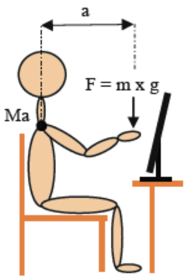


Figure 2.2: Biomechanics of touchscreen users, (Avsar et al., 2016a).

The researchers haven’t used Figure 2.2 to highlight possible differences in biomechanics due to individual differences in body characteristics, but to suggest a way of simulating increased g loads of the limb in static conditions. They used a weighted wristband in order to increase the moment that needed to be applied to the arm of the subject, and therefore simulating the effects of an increased g-force. The results for the discrete tapping tasks (Fitts (1954)) experiments showed an increase in movement time and a decrease in throughput with an increase in weight applied on the wrist. Moreover, one surprising result was an increase in aiming performance possibly due to a decrease in movement time. It needs to be recognized that the realism of this experimental set-up is limited, since in a real scenario, all the parts of the limb and body would exhibit an increase in weight due to an increase in g force.

Cockburn et al. (2017) is the first study that included an analysis of more realistic tasks when comparing the performance of the touchscreens with other input methods (such as trackballs) in simulated turbulent environments. The newly introduced tasks required a longer interaction time between the user's finger and the touchscreen and introduced two additional interaction possibilities, namely dragging and 'pinch and zoom'. For the slider tasks, large errors were recorded due to the difficulty of keeping prolonged contact with the touchscreen during the movement. Moreover, the prolonged contact with the screen increased the chances for the disturbances to negatively affect the task performance. Compared to the dragging tasks, the discrete selection tasks require rapid interaction with the touchscreen which decreases the time interval for the disturbance to negatively affect the task (Cockburn et al., 2017). In this study, two types of discrete target selection tasks were investigated, namely discrete acquisition of circular targets and numerical keypad entry tasks. The results of the study for the discrete pointing task showed that the trackball input exhibited fewer errors compared to the touchscreen input. It needs to be pointed out that, if a low accuracy for the task is required (i.e. keypad entry with wide keys), the touch input was on average the fastest in a disturbed setting. This study has shown that it can be the case that different tasks exhibit different performance and behavior characteristics when analyzing them in disturbed conditions. In general, the fast-discrete movement tasks show better suitability to this type of environment due to their limited interaction time with the screen. The study reported the motion consisting of *"non-periodic vertical displacements"* (Cockburn et al. (2017) p.p. 6747) and acknowledged the limitation of the motion's platform maximum displacement which limited the attainable vibration profiles. One question that still remains unanswered is how realistic are the introduced motion profiles compared to a real-life scenario.

Similarly, Alapetite et al. (2018) investigated tap, zoom, and drag and drop tasks in a disturbed environment. The study recognized some of the limitations of Stewart platforms not being able to produce long and sustained accelerations due to their limited motion profile. One of the peculiarities of this study is that the researchers also assessed the tasks in terms of the jerk (the first derivative of the acceleration) facilitated by their unusual experimental setting (using a rollercoaster to simulate a disturbed environment). In line with the previous study, the "dragging" task faced worse performance during turbulence when taking into account the time to complete the task and errors. Moreover, jerk affected the least the discrete action (tap), while the dragging task was slightly more sensitive to jerk compared to the zoom task.

Coutts et al. (2019) performed a study based on the same type of tasks as Cockburn et al. (2017), but introduced three touchscreen locations (overhead, center, and side) and a more complete disturbance profile including vertical, roll, and pitch movement. Confirming the previous studies, it was reported that the disturbance level affected negatively the error rate, task time, and accuracy for all the conditions. Moreover, the increase in key size for the discrete pointing task did not manage to eliminate completely the errors, the authors suggesting a need for confirmation when critical tasks are being performed. It is recommended that touchscreens shall be used for quick tasks, due to the acceptable accuracy of interacting using short presses. Slower types of interaction such as dragging are more prone to increase discomfort and require continuous interaction with the screen, which can affect the accuracy of the task in turbulent conditions. Consequently, quick actions using larger button sizes should be preferred over small ones or with a lengthy contact time.

Wynne et al. (2021) extended the work performed by Coutts et al. (2019), by investigating the performance of touchscreen gestures (swipe, tap, tap-and-hold, and drag) when performing the same type of task (number entry and pan). The drag gesture was the best in terms of completion time in a static scenario for the panning task when compared to the tapping gesture. Nevertheless, an increased level of turbulence had a negative impact on the dragging gesture, confirming the intuition from the previous studies. Moreover, it was pointed out that a more horizontal screen position exhibited less workload and discomfort ratings compared to a vertically oriented one. The reason for this is the possibility of keeping the hand and the wrist in a more neutral way compared to continuously extending the specific body parts. Single tap gestures were preferred over more complex ones such as swiping or sliding in turbulent scenarios when focusing on number entry tasks. Interestingly, this study mentions that for the dragging gesture, the increased contact time with the touchscreen might prevent the influence of disturbances. This fact disagrees with the previous two studies possibly because the previous studies analyzed tasks with different goals, each one using a different gesture while this one tries to investigate the best gesture for a common goal. In the end, it needs to be recognized that with different types of tasks, each can have different gestures that fit the specific task the best, while there is not one gesture that fits well for all the tasks.

2.4. Conclusions

The analysis of how humans interact with touchscreens in disturbed scenarios represents currently an interesting and value-adding topic given their potential to contribute toward a more intuitive Human-Machine Interaction (HMI) approach. Besides their intuitiveness, touchscreens were described to be a faster interaction method and they could save money due to their capabilities of efficiently displaying information. An area which has not been given great attention in the previously-mentioned studies is the analysis of the mechanisms regarding the decrease of performance in touchscreen operation, and the corresponding solutions and mitigation methods. It needs to be recognised that the decrease of performance in touchscreen operation has complex roots which can be tackled from multiple directions (purely cognitive, physical or an interaction between those two). In this report, the focus will be centered around the purely physical aspect of the interaction between the human, disturbance and the environment (touchscreen), that is the analysis of how the disturbances affect the human's body which in turn, cause involuntary inputs on the screen. It could be observed that most of the tasks that were analyzed in the presented studies were based on either pointing interaction methods, where the interaction between the finger of the user and the screen was very brief, and dragging interaction methods, which required a longer interaction time. Moreover, despite the fact that the dragging tasks were more affected by the disturbances, the discrete-pointing tasks also exhibited errors which could affect the performance of the tasks in safety-critical scenarios. In this way, the author would propose to further analyse two types of interaction methods, a discrete (fast) one and a continuous type of task, where the finger o the subject is kept in contact with a screen for an increased amount of time. In this way, the conclusions of the project can be applied and transferred to the most common touchscreen interaction methods.

3

Feedthrough of accelerations

This chapter aims to give an overview of the studies which investigated the effects and impact of the Biodynamic Feedthrough (BDFT) in the aerospace domain. Firstly, a general explanation regarding the effects of BDFT on HO's performance will be given. In addition, this chapter will present the modeling techniques that were proposed for predicting BDFT. It is important to mention that the focus will be on the methods the studies have used for constructing the models. Lastly, a short description of the studies that investigated the BDFT problem in the aerospace context will be given.

3.1. Effects of vibration on the performance of the Human Operator

The previous chapter showed that the performance of touchscreen use is affected in a negative way by the introduction of disturbances. However, most of the studies have investigated the effects of the parameters concerning the touchscreen's operation such as different types of tasks, interaction techniques, or screen positions (e.g. the studies presented by Cockburn et al. (2017) or Coutts et al. (2019)). In order to better understand how the performance of the touchscreen use is affected by vibrations, it is crucial to take a step back and understand what are the effects of several disturbance parameters such as frequency content or amplitude, on the performance of the HO when interacting with such a system.

It needs to be recognized that the body of knowledge concerning the effects of disturbance parameters on the performance of the HO when operating a touchscreen is limited. A field that received great attention in the past decades is the investigation of the effects of vibration on manual control performance using physical inceptors. Since this type of interaction method is currently found in every airplane, it is expected that some of the conclusions taken from these studies can be applicable to the newer interaction method.

McLeod and Griffin (1989) provide an exhaustive review of studies that investigated the effects of disturbances on the manual control performance with physical inceptors. The performance is analyzed in the context of changing the vibration variables (magnitude, frequency, axis), control system variables (controlled element dynamics, system dynamics), and the vibration exposure duration.

3.1.1. Effects of vibration frequency

The first investigated effects were linked to vibration in the vertical direction. When analyzing zero-order tasks and higher-order tasks under different vibration frequencies and equal acceleration magnitudes it was observed that the performance, or the feedthrough was affected in a non-monotonic fashion under varying frequencies. Lewis (1981) introduced vibrations between $2.5[Hz]$ and $12.5[Hz]$, specifying that the disruption in performance was the biggest at the frequency of $5[Hz]$ while at neighboring frequencies, the disruption was reduced (McLeod & Griffin, 1989). Similarly, when testing frequencies from $2[Hz]$ and $10[Hz]$, Allen et al. (1973) observed a maximum in control feedthrough at $4[Hz]$. It needs to be mentioned that the feedthrough of accelerations to the inceptor is not a straightforward phenomenon. McLeod and Griffin (1989) (based on Levison (1976)) mentioned that when analyzing the control feedthrough, interactions were present between the acceleration magnitude, control type, and vibration frequency. McLeod and Griffin (1989) concluded that the main frequency range in which the tasks are the most sensitive to the vertical vibration is from $3[Hz]$ to $8[Hz]$. Moreover, it has been pointed out that, in the same range of frequencies there is a maximum transmission of the vibration to the shoulders or head. Another very important aspect that will become recurrent in this report is the fact that the frequency that has the highest transmissibility depends on the biodynamic characteristics of the subjects (and therefore is subject-dependent) (McLeod & Griffin, 1989).

When discussing effects linked to vibrations in horizontal directions, McLeod and Griffin (1989) based on (Allen et al., 1973) reported a dependency of the feedthrough on the type of controlled element (CE) dynamics in the longitudinal direction. It was reported that the ”spring” control (more permissive) exhibited maximum feedthrough at $3[Hz]$ whereas the ”stiff” control at $1.3[Hz]$ and $4.5[Hz]$. In the lateral (y) direction, the maximum feedthrough was observed on average at the same frequencies for both types of CE’s, at $1.3[Hz]$ (with the ”spring” stick exhibiting more feedthrough at the lower frequencies (Allen et al., 1973)). As in the vertical direction, it was mentioned that the frequencies of maximum transmissibility are related to the frequencies of maximum transmission to the head and shoulders which are subject-dependent (McLeod & Griffin, 1989).

The effects presented in the previous two paragraphs were investigated using purely sinusoidal vibration profiles. The review study also presented a comparison between sinusoidal and non-sinusoidal vibration profiles. The study investigated by Lewis (1981) compared sinusoidal and random vibration representing amplitude-modulated sine-waves in zero-order tasks. It was reported that the random vibration produced less feedthrough compared to the sinusoidal vibration, although the reason for this aspect was not identified. When analyzing first-order tasks, no significant differences were present between the types of vibration. Nevertheless, the results suggested complex interactions between task difficulty, the type of vibration, and the duration of the task.

3.1.2. Effects of vibration axis

McLeod and Griffin (1989) mentions that the effects of the vibration axis on the feedthrough or performance depend on the axis that the task is performed. Usually, it is expected that the effects to be greatest if the task is performed on the same axis as the disturbance direction (i.e. roll task combined with lateral, y, vibration). Moreover, it has been also shown that in zero-order tasks, a disturbance in two simultaneous directions will produce a larger impact than any other disturbance in one single direction.

3.1.3. Effects of vibration magnitude

When analyzing the effects of vibration magnitude on the performance of zero-order tasks, the review has presented that both the minimum magnitude of vibration required to affect the tracking performance and the rate of increase in tracking error with increasing acceleration depend on the frequency of vibration. Lastly, it was recognized that a linear relationship exists between the magnitude of vibration and the tracking error. This relationship is dependent on the controlled device dynamics, the vibration frequency, and the system dynamics McLeod and Griffin (1989).

3.1.4. General effects

The study mentions additional effects that affect the feedthrough of vibrations to the control device. It was reported that the location of the control inceptor does not affect the feedthrough in a significant manner. On the other hand, when the upper limb is supported, the effect of the vibration can be diminished (reported under $10[Hz]$) in McLeod and Griffin (1989)).

There have also been reported effects of the mechanical characteristics of the manipulator and of the system dynamics on the vibration feedthrough. The dependence of the feedthrough on the manipulator characteristics might give the intuition that when being in contact with a surface/screen, the vibrations will feed through in a different manner compared to the hand not being in contact with any surface. This is an important aspect when modeling the limb when it is in contact with the screen.

Lastly, there have been mentioned possible effects of the workload and exposure duration on performance in disturbed environments. While an increase in workload can reduce the bandwidth of the tracking performance (which limits the frequency range at which the HO performs with small errors), the studies have not agreed on whether prolonged exposure to vibrations affects or not the performance. Nevertheless, it is expected that the increase in vibration exposure can affect cognitive processes such as (mental) fatigue or attention McLeod and Griffin (1989).

3.2. BDFT Modeling Techniques

3.2.1. Modeling techniques for physical manipulators

Early research performed in the field of feedthrough of accelerations through the human body focussed on extending the remarkable work performed by McRuer et al. (1967) regarding manual control.

One of the first studies which tried to model the effects of vibrational feedthrough to the control deflections is Allen et al. (1973). The study analyzed and partitioned the manual control behavior in three main elements: an element related to the HO's visual-motor response which is coherent with the target input, a portion linked to vibration input, and a remnant contribution that is not correlated either with the target input or vibration inputs. A novelty of the report is that it was one of the first reports that modeled the impact of vibrations in all three axes (fore-aft, lateral, vertical) separately.

Before the elaboration of the models, background information was given regarding the possible differences in the biodynamic response based on the direction of the disturbance. In the vertical direction, the response can be characterized by a complex mass-spring-damper system having the contribution from the following components: thorax-abdomen torso (resonant frequency range from $3[Hz]$ and $11[Hz]$), movements of the visceral organs (between $4[Hz]$ and $8[Hz]$), head-to-shoulder resonance (between $20[Hz]$ and $30[Hz]$). Moreover, it was mentioned that the overall head-to-seat impedance effects depend on the size of the subject and the posture, with a resonance interval between $4[Hz]$ and $6[Hz]$. In the horizontal plane, the biodynamic response is affected by the articulation of the hip joint and the spine which is being bent. Moreover, it was mentioned that lower resonance peaks (between $1[Hz]$ and $11[Hz]$

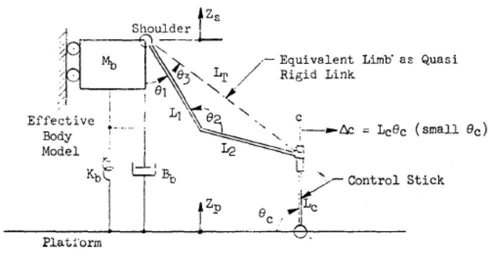


Figure 3.1: Feedthrough model for vertical motion (Allen et al., 1973).

and $3[Hz]$ and $11[Hz]$) are present compared to the vertical case.

The study made use of accelerometers which were positioned at key locations of the subject's body. The accelerometers were positioned on the shoulder, elbow, head and on the sternum. The reason for introducing these devices is to understand the transmissibility of vibrations but also to help in making several modeling decisions.

When focusing on the vertical (z axis) direction, the feedthrough data was recorded separately at increments of $1[Hz]$ starting from $2[Hz]$ up to $10[Hz]$. Moreover, in order to observe the feedthrough of the accelerations, the subjects had only to stabilize a pitch control task (without any target forcing function).

Figure 3.1 presents the proposed model which describes the feedthrough of platform accelerations in the vertical direction. Despite the complex whole-body and internal organ vibrational interactions (described by Von Gierke (1965) with multiple Degrees Of Freedom (DOF)), this study has opted for a simple 1 DOF mass-spring-damper model for the relative motion between the shoulder and the platform. The reason for this choice was the good agreement of the simple mass-spring-damper system with the experimental data for the sitting impedance.

When focusing on the upper limbs, it was assumed that the upper and the lower limb are represented by a rigid link with no equivalent mass. The rigid link assumption was backed up by the observation that the recorded accelerations of the shoulder and elbow were in phase across the frequency plateau. On the other hand, it was expected that inertia forces would be applied at the hand, due to the applied disturbance. It is hypothesized that the authors did opt for the massless rigid link, due to the fact that the rigid connection with the stick prohibits any inertial interactions from the vertical disturbance. In this way, it is assumed that the disturbances from the shoulder are directly fed through the stick. The resulting model describes the transfer function between the vertical acceleration and the stick deflection, being represented by a second-order system that is scaled by the (assumed) constant angles that the arm's subject makes with the vertical direction. This scaling term can be considered to represent changes in the model based on individual characteristics. Compared to the experimental results, this model successfully predicted the second-order nature of the feedthrough dynamics, giving a good approximation of the resonant peak. It needs to be mentioned that in the conditions of touchscreen operation, if the screen is not positioned horizontally, it is expected that inertia effects would be present throughout the arm due to it not being constrained vertically compared to the above-mentioned example.

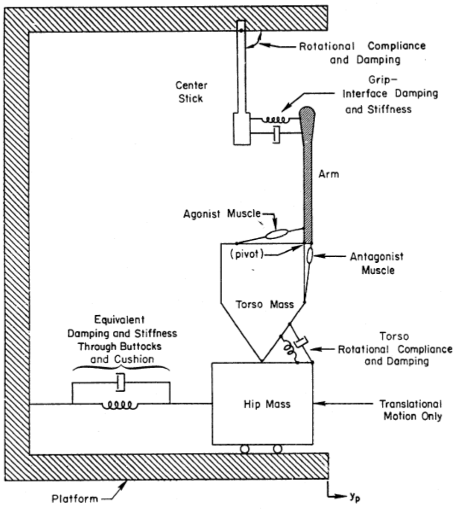


Figure 3.2: Feedthrough model for horizontal motion (Allen et al., 1973).

When analyzing the feedthrough of the horizontal motion, it needs to be mentioned that in this case the limb and body system performs both rotational and translational motion. Starting from the bottom, it is assumed that the hips perform only translational motion being connected to an equivalent mass-spring-damper system representing possible lower limb and seat interaction. The torso is assumed only to rotate with respect to the hip, being connected with an equivalent rotational spring and damper system. Lastly, the arm s assumed to move relative to the shoulder, by using a combination of agonist and antagonist muscle pairs. When discussing the equivalent effects of the closed loop neuromuscular system (NMS) on the upper limb's response, it can be considered that these effects can be described by an equivalent spring and damper system. It was observed that the lateral feedthrough could be explained mainly by the inertial effects present due to the arm's and stick's masses and inertias.

The work performed by Jex and Magdaleno (1978) continued the analysis of the feedthrough of disturbances in a manual control scenario using physical inceptors.

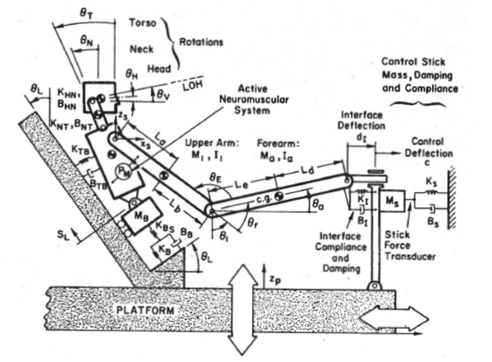


Figure 3.3: Biomechanical features used for describing the feedthrough (Jex & Magdaleno, 1978).

The model includes the torso, upper and lower arm links which are influenced by the neuromuscular system. Since the focus of the model is the analysis of the feedthrough of the accelerations to the stick, simplifications were made to the elements describing the lower body and the torso. The interaction of the model with the seats was modeled through a set of linear springs and dampers, as in Figure 3.3. It is interesting to observe the modeling choices made by this study, in comparison with the one made by Allen et al. (1973). Compared to the previous study, this model includes a neuromuscular component driving the limbs and, possibly the biggest difference is that in this case, the limbs include inertia parameters.

When taking a closer look at the neuromuscular system modeling adopted in Jex and Magdaleno (1978), several considerations can be made. It was proposed that only the lumped effects of the agonist/antagonist muscle pairs would be taken into account due to the complexity of modeling each individual muscle. The authors have chosen a linearized representation of the neuromuscular system consisting of tendon elements and a Hill-type muscle model (Hill, 1938). The linearized Hill-type muscle model consists of a force generator, and damping properties along with a series-elastic component representing the elasticity in the tendons and in the muscle and a parallel-elastic component representing the damping and spring characteristics of the inactive muscle (Van Paassen, 1994).

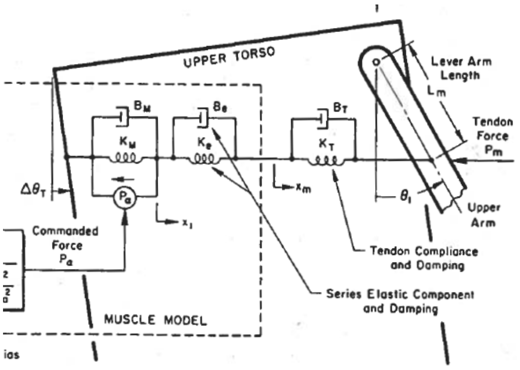


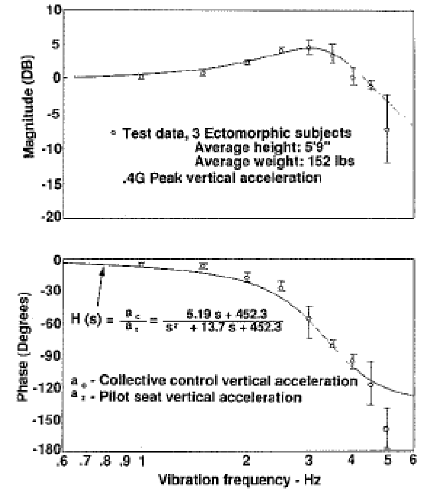
Figure 3.4: Neuromuscular system description (Jex & Magdaleno, 1978).

Based on Figure 3.4 and on the considerations made above, it is interesting to observe Jex & Magdaleno's choice of introducing both a series-elastic component and a tendon compliance and damping element separately. Moreover, it is interesting to understand why a linearized version of the Hill-type muscle model is used in this particular case. In Van Paassen (2021), it is mentioned that the force generated by the muscle depends on certain relationships, namely the force-velocity and the force-stretch/force-length relationships. It is important to recognize that these relationships are dependent on the activation levels of the muscle. Saying this, when the authors mention that a linearized neuromuscular system description is used, it means that these equivalent constants that describe the damping and elastic components correspond to a constant activation level of the muscle around which the linearization was made.

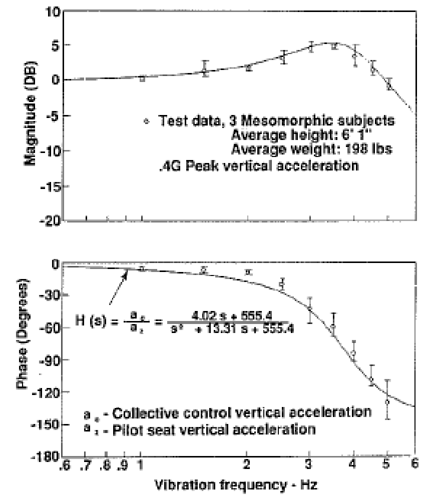
In a critical analysis, Venrooij et al. (2016b) pointed out a possible disadvantage of such complex models. It has been mentioned that the research community would rather apply simpler and more practical models, with a lower complexity which are elaborated to an intended purpose. Moreover, another possible disadvantage is a large number of parameters, which can be difficult to identify from data, as the study does not mention explicitly the approach used for parameter fitting.

The model developed by Mayo (1989) is one of the very few models to this date that tackles the problem of within-subject variability in the analysis of the biodynamic response of the HO. The study focussed on the acceleration transmissibility on the collective inceptor during helicopter's vertical motions. When analyzing possible factors linked to the HO which can affect the stability of the system, the study differentiates between two phenomena: Pilot-Induced Oscillations (PIO) and Pilot-Assisted Oscillations (PAO). Of more interest in this context are the PAOs since they result from the passive control activations due to the response of the body to vibrations. Mayo (1989) pointed out that when analyzing the feedthrough of vibrations to the physical inceptor, the response depends on the individual characteristics of the subjects. Some examples of such characteristics are: experience, body type, and neuromuscular system (Mayo, 1989). Taking a step back, these components have also been introduced in the previous two studies.

The goal of the study was to create a model which lumped the characteristics of the seat, limb, torso, and stick. In order to do this, accelerometers were positioned on the collective inceptor and on the pilot's seat. At the same time, a moving-base simulator was mimicking sinusoidal motions ranging from $1[Hz]$ and $5[Hz]$. In the trial phases of the experiment, it was observed that the acceleration feedthrough depended on the collective position. For example, it was mentioned that the shoulder accelerations were transmitted directly to the collective in the case of low angles of the stick. On the other hand, at higher angles, the neuromuscular system is contracted and it was hypothesized that some components of accelerations transmitted through the torso/shoulder are absorbed before interacting with the collective.



a: Ectomorphic Subjects



b: Mesomorphic Subjects

Figure 3.5: Resulting models, describing the feedthrough of accelerations to the collective input (Mayo, 1989).

In order to capture individual differences between the subjects, the authors proposed two different models depending on the anthropometric differences between the subjects Figure 3.5 (Mesomorphic and Ectomorphic body types). One interesting aspect highlighted by the authors is the slightly lower lumped resonance frequency when compared to the whole-body resonance frequency reported in other studies (such as Allen et al. (1973) in between $4[Hz]$ and $6[Hz]$). The reason for this phenomenon was explained by the fact that the obtained transfer functions also included the neuromuscular reflex which tries to counteract the disturbances. In this way, the subjects varied the stiffness of their limbs in order to account for the whole-body vibration effects. This fact can have great implications regarding the identification of the biodynamic characteristics of the arm possibly from images. With this, the identification problem becomes more complex since it can be difficult to identify and also when the time-varying changes in the apparent stiffness of the limb occur. One possibility is that these time changes are affected by the disturbance signal characteristics (a greater disturbance possibly induces a stiffer setting of the arm).

A qualitative observation based on the videotape measurements confirmed that a more contracted arm leads to a less sensitive stick motion to the feedthrough of accelerations. The limitation of this qualitative analysis is whether this phenomenon occurs along the whole frequency range or whether it occurs only at a certain range of frequencies. Moreover, it is still unclear how the researchers evaluated the amount of contraction of the neuromuscular system.

Biomechanical pilot models were also constructed for understanding and predicting aircraft-pilot couplings (APC). Koehler (1997) and Höhne (2000) focussed on extending the existing models for the analysis of roll ratcheting (RR). As mentioned in Höhne (2000), roll ratcheting represents a low amplitude and high-frequency aircraft-pilot coupling that occurs in the pitch axis. Usually, this behavior occurs in the frequency range between $1[Hz]$ and $3[Hz]$, due to inertial forces interacting with HO's limbs/body.

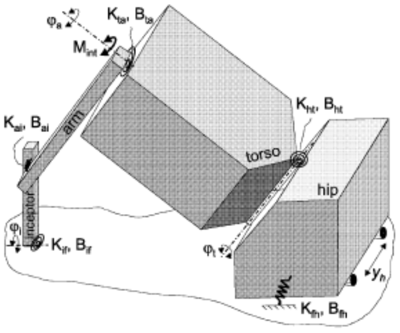


Figure 3.6: Biomechanical model used for predicting roll ratcheting (Koehler, 1997).

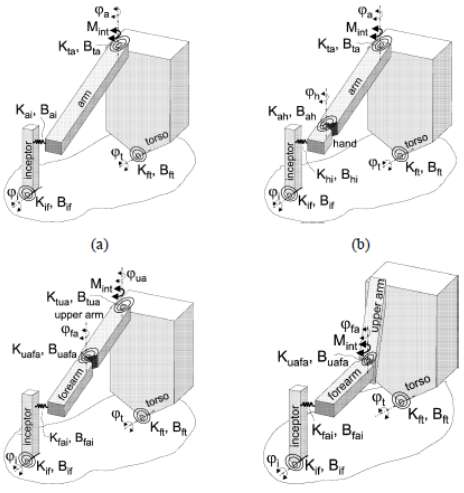
The model proposed by Koehler (1997) to predict the roll ratcheting behavior was based on the one from Allen et al. (1973), with slight modifications. The elements were interconnected using either linear or rotational mass-spring-damper systems that idealize the visco-elasticity of the body and of the muscular system. Each of the elements is assumed to have one DOF. The hip is assumed to be translating with respect to the airplane, the torso rotates with respect to the hip around an inclined axis, the arm rotates with respect to the torso and the inceptor rotates with respect to the aircraft. The model defined by Koehler (1997) has as its input the lateral acceleration of the HO, the acceleration in the roll direction, and the desired moment applied at the pilot's arm. The output is the stick control force.

In his study, Höhne (2000) recognized the shortcomings of the previous study, mentioning that the model developed by Koehler (1997) has no available information regarding the quality of the identified parameters and attained unrealistic values for stiffness and damping coefficients. Taking this into account, Höhne (2000) proposed a re-derivation of the equations of motion (EOM) and a re-identification of the parameters. The study considered approaching the problem as a multibody system. A possible way of solving such a problem is by determining its equations of motion using Lagrange's equation. One possible disadvantage of this formulation is that complex systems might contain lengthy equations for the velocities/accelerations of the center of mass of the components. Saying this, the study used a multibody simulation software, namely SIMPACK [1] to define the multibody model. The elastic and viscous parameters were obtained following a maximum-likelihood approach, and the standard deviations of the parameters were computed as a measure of reliability.

Since high standard deviations for the identified parameters were obtained when fitting the previously-defined model, the study proposed and tested additional models with different topologies. Four main groups of models have been defined as in Figure 3.7. Starting from the top-left to the bottom-right these models are defined as (Höhne, 2000): three-body models, wrist joint models, elbow joint models, and angle arm models. While performing the same fitting procedure as the original model, the authors made a couple of observations that can be used in the further stages of the report. It was observed that high standard deviations and

---

[1] https://www.3ds.com/products-services/simulia/products/simpack/



(a) (b)

(c) (d)

Figure 3.7: Additional biomechanical models for predicting roll ratcheting (Höhne, 2000).

correlations between the parameters were obtained with the models which consisted of two or more joints that had parallel rotation axes (i.e. Figure 3.7 c)). Moreover, if the experimental conditions show that the subjects do not move their upper limb with respect to the lower limb, the addition of an elbow joint in the model might seem redundant (since the whole upper limb can be considered one rigid body).

The takeaways of these two studies indicated the suitability of using a multibody approach for the biomechanical analysis of the human body in disturbed scenarios. Moreover, the importance of deciding the topology of the model based on the analysis of the ergonomics of the subject's movement while he/she is performing the task was highlighted. On the other hand, it needs to be mentioned that this parametrization consisted of spring and damping coefficients which were constant in time while it is known that the HO might change his/her equivalent spring and damping characteristics throughout the task. The study does not report any indication regarding this phenomenon or any possible solution to account for this.

Researchers from Politecnico di Milano (PoliMi) focussed on the research of BDFT in the context of rotorcraft. Mattaboni et al. (2008) and Mattaboni et al. (2009) are the first preliminary studies performed in this area that tried to model using a multibody approach the dependence of HO's biomechanical properties on the anthropometric characteristics and posture. The context of this research is based on the authors' insight that the transfer of accelerations is a highly variable phenomenon across, being dependent on the following factors: anthropometrics of the HO, cockpit's configuration, workload, and the ergonomics of performing the task. The main goal of the work of the two studies was to create an identification procedure that could be used for estimating the equivalent impedance for the upper limb's muscles.

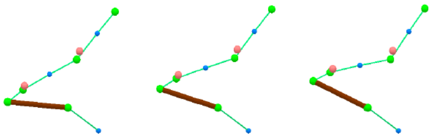


Figure 3.8: Multibody model describing HO's upper limb and collective (Mattaboni et al., 2008).

Figure 3.8 presents the defined multibody model used for the analysis of the interaction between the subject and the collective. The equivalent impedance of the muscles was defined with a viscoelastic constitutive law with parameters describing the coefficients of rotational springs and dampers. Since the scenario implied that the limbs kept a constant value for the collective deflection, the need for introducing more accurate muscle models which could include the actuation of the muscle was not deemed necessary. In this way, the simpler linear constitutive laws could create an equivalent mechanical model for the human body in this static-task scenario.

In order to identify the impedance characteristics of the upper arm, the motion of the limb had to be reconstructed. Since no optical techniques were available which could measure the motion due to the limitations of the moving-base simulator, the study used information from: accelerometers positioned on the seat of the subject, the collective rotation, and accelerometers positioned on the wrist and upper limb. It was observed that all the sources of information created an overdetermined system. The study proposed a least-squares problem to solve for the joint motions along with the corresponding velocities and accelerations. A power balance approach was used to obtain the parameters of the equivalent spring and damping coefficients by means of a least-squares formulation. The power-balance approach was tested using simulated measurements of accelerometers and collective positions. It was concluded that if simulated noisy measurements are to be used, the quality of the identification decreased considerably, providing unrealistic results for the parameters. This fact prohibited the application of this method to real measured data and questioned the robustness of the approach.

Critically analyzing the outcomes of the studies, it can be observed that both tried to obtain certain characteristics of the response of the body, based on external measurements. The multiple indirect sources of information caused an overdetermined system for the kinematics of the upper limb, and therefore a kineto-static problem had to be solved to obtain the motion of the model. It is not known how this approach would perform if direct measurements of the relevant points of the limb could be measured by using markerless motion techniques. Saying this, the present approach might represent an interesting solution that solves for the parameters of an assumed model for the constitutive laws of the joints by using movement information.

In a later study, Masarati et al. (2013a) extended the model defined in the previous studies, in order to model the dependence of the equivalent stiffness of the neuromuscular system on the muscular activation patterns.

The model possesses 7 degrees-of-freedom modeling all the 4 segments of the upper limb (humerus, radius, ulna, and hand) and 25 muscles. The numerous pairs of muscles were introduced such that the modeled 7 DOF system could be actuated. The study proposed a 3 DOF model for the shoulder complex, 1DOF for the humeroulnar joint, 1 DOF for the combination between the humeroradial and radioulnar joints, and 2 DOF for the wrist. Compared to the previous studies, in this case only the position of the end-effector of the limb is given while performing a maneuver at the collective level. In this conditions, the authors proposed an analysis based on multiple steps. Firstly, given the position on the end effector, an inverse kinematical analysis was be used to obtain the most plausible configuration of the whole upper limb (along with the velocities and accelerations of the points of interest). The kinematic analysis was

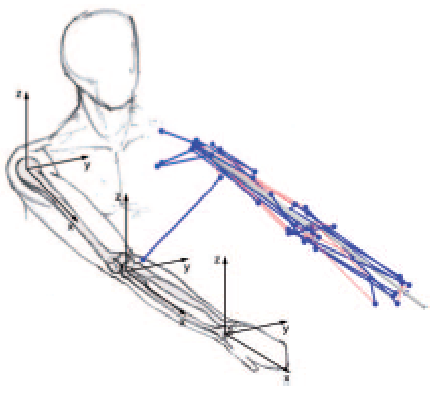


Figure 3.9: Biomechanical model of the upper limb including linking muscles (Masarati et al., 2013a).

performed based on the study researched by Pennestri et al. (2007) which introduced an optimization approach based on an ergonomic cost function that leads to realistic configurations for the upper arm. It needs to be mentioned that this step is performed due to the fact that only one position in the kinematic chain is known, the problem being kinematically underdetermined. It is hypothesized that if motion-capturing techniques would be applied, the previously-mentioned step would not be necessary. On the other hand, it is still unknown how the velocity and acceleration of the coordinates will be affected if a numerical differentiation scheme would be applied to noisy measurements.

After the coordinates of the upper limb were derived, along with their first and second derivatives, an inverse dynamic analysis is performed in order to define the joint torques needed to perform the specific task. The study goes one step further, by defining a set of 25 muscles along with their properties, including the reference length and insertion locations. It has been mentioned that the defined insertion locations were not defined in an anatomically correct manner, the goal being only to replicate a realistic torque configuration about the joints of the system. In a real-life scenario, this modeling approach can be considered very complex since it is hypothesized that there can be individual differences regarding the muscle characteristics of each person. (that are hardly measurable). The force induced by the muscles is assumed to be acting in a straight line between the corresponding insertion points. For defining the force applied by each muscle, Hill's model which describes the dependence of the muscle force on the elongation rate, peak force, elongation, and activation was defined. It has been recognized that computing the corresponding activation patterns which result in the obtained torque, is an undetermined problem since there can be multiple combinations of activations that lead to the same torque. The study has chosen an optimization approach that could minimize the square norm of the activations.

The resulting activation patterns represent the minimum activation patterns required to produce the required torques (Zanoni et al., 2021). In addition to this, the study recognized the importance of introducing the effects of the reflexive activity of the muscle model, since the reflexive activity is expected to occur in a perturbed environment, accounting for a considerable increase in stiffness, up to 120% (Masarati et al., 2013a). The reflexive activity was modeled by assuming a certain behavior for the perturbation activation patterns (being proportional to the perturbed muscle velocity and length). An additional activation contribution is present in terms of torque-less activation modes (TLAMs), which represent additional activations (from the baseline minimum activation patterns) that do not alter the computed toques. All these contributions add to the total activation of the muscles, which, in turn, were used to simulate time-varying changes in equivalent stiffness and damping properties of the neuromuscular system during the desired task.

The above-mentioned study researched a rather complex approach for simulating how the changes in muscle activation during the task affect the lumped stiffness and damping characteristics of a biomechanical system. The approach required the definition of pairs of muscles that were used to apply torques on the joints. The combination of inverse kinematics and dynamics showed that during a task, the biomechanical properties of the neuromuscular system change due to different mechanisms (intrinsic and reflexive) which are explained by changing the activation patterns of the muscles. Nevertheless, this modeling approach might deem complex since precise properties of the muscles need to be obtained (which can differ from person to person). On the other hand, this study reinforced the suitability of using a multibody model for biomechanical analysis.

In one of their recent studies, the researchers from PoliMi, extended the previous model of the upper limb with a multibody description of the spine (Zanoni et al., 2020). Inspired from (Kitazaki & Griffin, 1997), the authors have proposed a torso model consisting of 34 rigid bodies replicating the vertebrae and the visceral parts connected to them. It was assumed that each vertebra would be represented by a rigid body with rotational and translational DOF Figure 3.10. Each rigid body is assumed to be interconnected to its neighboring vertebras/visceral nodes by rotational and translational springs. Compared to the previous studies, the authors have decided not to include detailed modeling of the muscles, limiting the range of validity of the model.

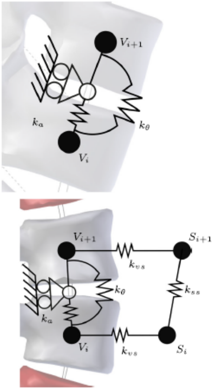


Figure 3.10: Overview of the connections between the vertebrae and visceral elements. (Zanoni et al., 2020).

A novelty of this study is that a parametrization of the model geometry, masses, and parameters was introduced. The goal of the parametrization was to produce the most plausible subject (not subject-specific) possessing a certain combination of age, body mass index, height, and gender. The position of the nodes of a particular parametrization represents a scaled version from a set of reference dimensions. The mass of the limbs is defined from a regression equation having as input the mass of the body. The intervertebral stiffness and damping coefficients are determined also as second-order polynomials as a function of the body mass index.

Similar to the previous study, the analysis was performed using a cascaded approach starting from the inverse kinematics, inverse dynamics, and computation of the muscular activations. In this scenario, the inverse kinematics was performed by enforcing the movement at the end effector of the hand and at the head. In this study, a more intuitive explanation was made regarding the torque-less activation modes, representing a way of modulating the impedance of the HO based on psychological characteristics (i.e. workload or fatigue). The authors mentioned that in general the results were in concordance with the experimental results of other studies when focusing on the seat-to-head transmissibility of accelerations. One clear discrepancy that the authors will investigate further, is the difference in the longitudinal response of the head to a vertical acceleration input.

One of the main advantages of using multibody approaches is the possibility of using physics-based principles in order to account for sources of variability between persons. In other words, by correctly defining the parameters of the models, the equations of motion can account for and describe changes in the human response due to changes in inertias, masses, lengths and configurations of the limbs. The complete model consists of 103 DOFs which might represent a slight disadvantage in terms of the computational effort that takes to perform the analysis. The authors mentioned that the dynamical problem was solved in the multibody solver MBDyn [2], but did not bring into the discussion any performance results. Moreover, since the generation of the nodes representing the spine is based on a ”general” model, it can be difficult to obtain and use spine configurations from a particular subject.

3.2.2. Modeling techniques for touchscreens

As previously described, the body of knowledge regarding modeling the BDFT in touchscreen operation is limited. Two studies from the same group from TU Delft were the only ones found by the author which focused on the cybernetic analysis and modeling of the BDFT in touchscreens. Mobertz et al. (2018) and Khoshnewiszadeh and Pool (2021) investigated the relationship between the disturbance signal applied on a 6 DOF simulator and the input observed on the touchscreen. In this way, a lumped-parameter model which could describe the contribution of the disturbance to the input of the screen was constructed. The model is presented in the following equation:

---

[2] http://www.mbdyn.org/

$$\hat{H}_{BDFT} = G_{BDFT}\frac{\omega_{BDFT}^{2}}{s^{2} + 2\zeta_{BDFT}\omega_{BDFT}s + \omega_{BDFT}}e^{-s\tau_{BDFT}} \tag{3.1}$$

Equation (3.1) shows that the proposed BDFT model contains a set of four parameters: G_{BDFT}, ω_{BDFT}, ζ_{BDFT}, τ_{BDFT} lumping the effects between the disturbance and the finger location on the screen (Khoshnewiszadeh & Pool, 2021). Khoshnewiszadeh and Pool (2021) mentions that the previously-mentioned model lumps the feedthrough contributions from the seat, spine, and upper limbs. One particular disadvantage of this approach is the inability to understand which body part has the biggest contribution towards the feedthrough of accelerations. Moreover, as it will be shown in a later section, it is difficult to understand and clearly point out the differences between and within the subjects only by analyzing/comparing the model parameters.

3.3. Biodynamic feedthrough studies

3.3.1. Biodynamic feedthrough studies with physical manipulators

As previously mentioned in the previous sections, BDFT is a complex phenomenon that can be affected by both internal factors (related to the human operator) but also by external ones (disturbance characteristics in terms of frequency, amplitude, etc.). Even though significant efforts were made in the 20th century to understand and characterize the implication of these factors in the manual control performance (i.e. Allen et al. (1973), Jex and Magdaleno (1978), Idan and Merhav (1990)), a commonly accepted framework that could be used to understand BDFT was still missing. In the 21st century, several research groups from PoliMi and TU Delft have made efforts to investigate further different aspects linked to BDFT in the use of physical inceptors.

A series of studies starting from Venrooij et al. (2009) further investigated the effects of BDFT on stick-based inceptors with the goal of creating a unified framework for analyzing this phenomenon. This initial study was one of the first ones which tried to relate neuromuscular admittance to BDFT. Neuromuscular admittance can be considered a measure that characterizes the arm dynamics in relationship with the control inceptor. A more intuitive explanation of the neuromuscular admittance was made by Van Paassen (2021) which describes it as how much the limb moves in response to the force that is exerted on it. It needs to be recognized that neuromuscular admittance is an adaptive quality, being dependent on *”the intrinsic and reflexive feedback, the level of co-contraction and task instruction”* (Venrooij et al. (2009) p.p. 1668). In order to enforce different settings of the neuromuscular system, the study proposed three disturbance rejection tasks that needed to be performed by the subjects. These tasks are: Position Task (PT) - requiring to resist the force perturbations (stiff neuromuscular setting); Force Task (FT) - requiring to minimize the force applied to the stick (compliant neuromuscular setting). It was observed that at low frequencies ($< 1.5[Hz]$), a low admittance is translated into a low BDFT, whereas for higher frequencies the effects of BDFT increased when the neuromuscular was in a stiff setting.

In other words, based on Figure 3.11 it can be mentioned that a stiff setting of the neuromuscular system can actually increase the feedthrough of disturbances to the control stick at higher frequencies, compared to a more compliant setting. Since it was hypothesized that during the experiment runs, the only adaptations were due to changing the neuromuscular admittance, it was assumed that a relationship between admittance and BDFT existed. Alternatively, the different tasks imposed different settings of the neuromuscular system (due to HO’s adaptation) which, in turn, induced different BDFT properties.

In other to better understand the mechanisms in the relationships between neuromuscular admittance (task interpretation) and BDFT, Venrooij et al. (2010a) split up the BDFT in two parts (Figure 3.12).

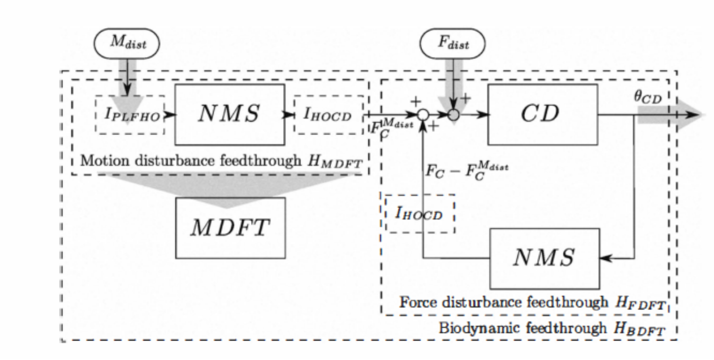


Figure 3.12: Overview of BDFT consisting of a motion disturbance part and force disturbance part (Venrooij et al., 2010a).

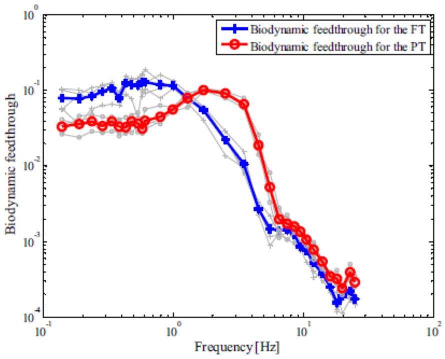


Figure 3.11: Overview of the biodynamic feedthrough measured during the PT and FT (Venrooij et al., 2009).

Figure 3.12 presents how BDFT is considered as a combination between Motion Disturbance Feedthrough (MDFT) and Force Disturbance Feedthrough (FDFT). MDFT represents the feedthrough of the motion disturbance on the forces on the inceptor, while FDFT represents the feedthrough of the resulting forces on the controlled element deflections. By exploring the Frequency Response Functions (FRF) of both components, a better understanding can be made regarding the behavior of Figure 3.11. It was observed that the FRF of the MDFT had the highest magnitude across almost all frequencies for the PT compared to the other tasks. This means that a stiff setting of the NMS transmits the contact forces the most. On the other hand, it was hypothesized that also due to the stiff setting these contact forces are not transferred back to deflections in the controlled element (being absorbed by the upper limb). In this way, the combination of a high MDFT and low FDFT, results in a low BDFT at low frequencies. Saying this, it is expected that the change in behavior at the higher frequencies to be due to changes in the behavior of the FDFT. The conclusion that the BDFT depends on the task that is being performed can lead to an additional point of interest.

A possibility for mitigating BDFT is by using a model-based approach that predicts the effects of BDFT and subtracts these effects from the total contribution on the manipulator (Venrooij et al., 2010b). For a good cancellation performance, it is, therefore, crucial to take into account the changes in the BDFT dynamics on the basis of the task at hand.

A continuation of the previous studies is the research performed by Venrooij et al. (2011b) in the field of BDFT and stick inceptors. The study proposed a frequency decomposition technique based on the assumption that the measured variables (the angle of the stick and the force applied on the stick) consist of contributions from the disturbance force applied on the stick (used for the identification of admittance), the motion disturbance, the cognitive action of the HO and a residual which accounts for other dependencies besides the ones mentioned. This assumption results in the possibility of decomposing the power of the total signal as a sum of the powers from each contribution variable that affects the signal. An example of such decomposition is Figure 3.13. Moreover, the frequency decomposition technique can deem helpful in identifying the relative contribution of each variable to the measurement quantity. This can be of interest when a frequency-based technique is used for identifying the BDFT dynamics. If the contribution from the motion disturbance to the output variable is not significant, then it can be expected that the power of the signal at the frequencies corresponding to the motion disturbance to be low, maybe so low that it can not be distinguished from the residual contributions. In this way, the confidence of the identified FRF is low since it can't be decided whether the contribution comes from the disturbance or is it just noise. One advantage of this technique is that the individual contributions can also be transformed into the time domain for quantitative analyses.

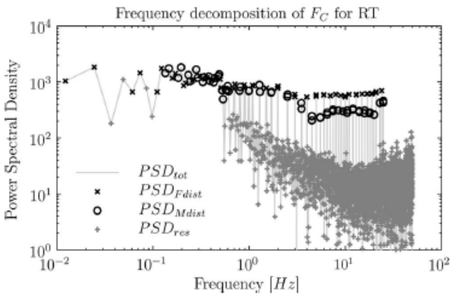


Figure 3.13: Overview of the frequency decomposition technique used for obtaining the relative contribution of each disturbance to a measured signal (Venrooij et al., 2011b).

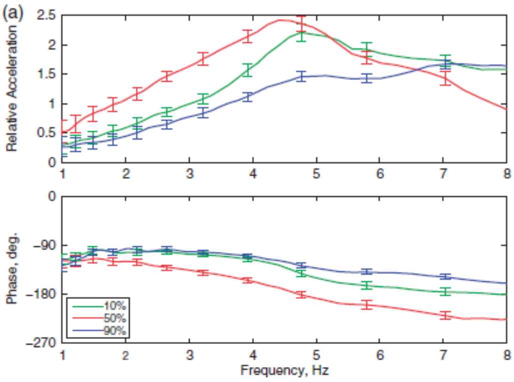
If the motion disturbance excites the measured variable enough, the identification of BDFT dynamics can be performed. In this way, the contribution of the motion disturbance to the measured signal can be estimated and used for the previously indicated model-based cancellation approach. Venrooij et al. (2011a) investigated the performance of different generality levels on the signal cancellation method. The different generality levels are represented through the way the estimated FRFs of BDFT are being used. The lowest level of generality is represented by the FRFs identified in the individual runs for each subject. One level of generality higher is represented by the FRF obtained by averaging all the estimated BDFT functions for one subject. The largest generality level is represented by averaging the previously defined subject-average models. In this way, the study estimated the percentage of BDFT that could be canceled using each type level of generality. It was observed that the global-level models performed worse compared to subject-average-level models in terms of how much percentage of the BDFT was canceled. This fact implies that if a model-based cancellation method is desired, the model would have to match the subject at hand. This aspect confirms the complexity of modeling the BDFT effects being subject-dependent and also task-dependent. One possible limitation of this study is that the results were obtained using optimal signal cancellation, or, in other words, the model was applied to the same runs that it was identified from. In this way, the validity of results outside the training domain are not known (for example testing on a person the model was not trained with). The reader should not forget that this study also tested the performance of model-based cancellation for the different tasks inducing different neuromuscular settings (tasks similar to the ones proposed in Venrooij et al. (2009)). It needs to be pointed out that worse cancellation results were obtained for models identified for tasks requiring different neuromuscular settings. This fact can suggest that within-subject variability might affect more BDFT behavior compared to the between-subject variability.

In line with the previous studies, Masarati et al. (2013b) recognized the importance of a BDFT model to account for variabilities due to the *" physiological dynamics of the pilot's neuromuscular system*" (Masarati et al. (2013b) p.p. 127), being dependent on anthropometric data, the cockpit layout and interfaces, the posture of the HO, the strategy of the pilot and his/her skills, and the cognitive activity of the HO (Masarati et al., 2013b). The study analyzed the transfer of disturbances in pilots with different physical characteristics and with different reference limb positions using body-mounted accelerometers. It has been observed that the differences between pilots appear especially in the accelerations of the elbow and the wrist.

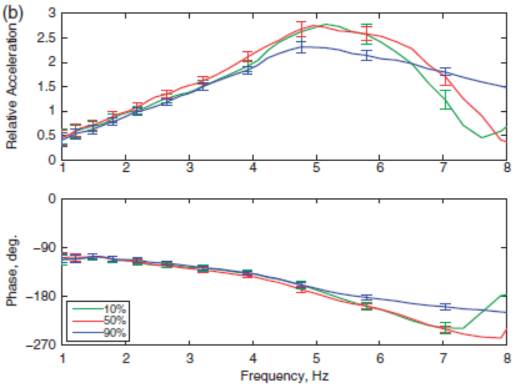
Moreover, it has been pointed out again that the muscular activation state can affect the BDFT even within a subject. For example, Figure 3.14 presents the feedthrough of accelerations to the collective for two different subjects and three different reference collective positions. The figure shows the within and between-subject characteristics of BDFT, implying that the FRFs of the two pilots are rather different and also depending on the reference limb position, the feedthrough is different within a subject (more clear results presented for Subject 1). The authors hypothesize that the differences might be due to the activation of the muscles rather than due to the different positions of the limbs.

In a similar field, Venrooij et al. (2013) emphasized that between and within-subject variability are important aspects to take into consideration when analyzing and modeling BDFT. The study investigated subjects with different somatotypes (between-subject variability) and tasks requiring changing the stiffness of the neuromuscular system (within-subject variability). The authors suggest that the results do not show a clear effect of the body type, whereas, the different neuromuscular settings influence the dynamics of BDFT.

In the review of the open issues regarding BDFT, Venrooij et al. (2016b) pointed out that until then, BDFT was identified only using linear-time-invariant situations. In order to analyze BDFT in more realistic settings, time-varying or adaptive techniques would need to be developed. In this way, the time-varying behavior of the HO could be accounted for in the model-based cancellation.



Pilot 1



Pilot 2

Figure 3.14: Overview of the differences in the feedthrough of accelerations of two subjects and three different values of the control inceptor location (Masarati et al., 2013b).

One of the first studies which investigated adaptive models in the context of model-based cancellation of BDFT in stick-based operation is Venrooij et al. (2016a). This study can be considered a continuation of the previous studies which investigated the change of BDFT dynamics and its effect on the cancellation performance when the task that the operator performed changed (force, position or relax tasks). A novelty of the study is that it proposed a two-step procedure for the evaluation of the cancellation performance: Firstly, two different models for the BDFT effects were identified using the ”classical” position and relax tasks for each subject; Secondly, a more realistic task was introduced (Highway-in-the-Sky) which, by varying some of the parameters of the outside visual cues, induced the same adaptive neuromuscular settings corresponding to a force and position task. In order to investigate the adaptive nature of the neuromuscular settings, the visual cues changed within a run at a particular time stamp. The cancelation performance was evaluated in the consistent conditions (PT/RT model applied to PT/RT segment of the second task) and in the incongruent conditions (PT/RT model applied to RT/PT segment of the second task). One important finding of the study was that for some subjects in the incongruent cancellation scenario, the performance of the cancellation was better than the performance of the consistent conditions (particularly in the RT condition). This aspect can give the indication that some subjects behaved differently in the identification runs compared to the runs of the validation session. For example, after examining the BDFT frequency response function for some of the subjects, it was observed that even though the second task imposed a relax task, the subjects behaved in such a way that their neuromuscular system was stiff (more similar to a position task), therefore inducing a better cancellation in the incongruent condition. This fact can lead to the conclusion that better cancellation results are obtained with a model that matches the current BDFT dynamics. Moreover, since it was shown that the BDFT dynamics depend on the neuromuscular settings, it can be concluded that for an ”accurate” cancellation, the models would need to be adapted to the current state of the neuromuscular settings.

One of the takeaways of the previously-mentioned study is the fact that even though the second task was designed to impose different neuromuscular settings, the way a particular subject approaches the task can vary from person to person. In other words, if one subject performs the task with a relaxed neuromuscular setting, one subject can perform the same task with a stiff setting. Moreover, it needs to be pointed out that in a real-life scenario, it is hard to clearly identify which type of task imposes which type of neuromuscular setting. In this way, the author considers that there is a need for an external perspective that can recognize the neuromuscular setting used by the subject independently from the task that is being performed.

3.3.2. Biodynamic feedthrough studies with touchscreens

One of the first studies which quantified the effects of BDFT in touchscreens was performed by (Mobertz et al., 2018). The subjects performed a continuous dragging task, with three different multisine perturbation directions (surge, sway, and heave) on two separate displays. The results demonstrated that the effects of feedthrough are the strongest when the disturbances act in the same direction as the inputs on the touchscreen. Moreover, by using the previously-mentioned frequency decomposition technique, the contribution of the motion disturbance signal on the touchscreen input could be evaluated for the different conditions. In this way, the feasibility of BDFT estimation using frequency response methods could be evaluated. For the cases where motion disturbance contribution on the touchscreen input was large enough, a mass-spring-damper model was fitted in the frequency domain which idealized the relationship between the motion disturbance and the feedthrough in the screen input (as described in the previous section).

The fact that the influence of BDFT on the touchscreen input could be identified in the frequency domain inspired Khoshnewiszadeh and Pool (2021) to use it in a beneficial way. The study investigated the performance of model-based signal cancellation techniques in dragging tasks which consisted of continuous multisine 2-D target following tasks and 2-D step inputs replicating a waypoint relocation. A mass-spring-damper model suggested by the previous study was fitted on the individual runs in order to model the effects of BDFT, and global-level models were constructed by averaging the parameters accordingly. Even though the cancellation performance was acceptable, it was concluded that models more tailored to particular individuals would allow a better cancellation performance in the case of model-based signal cancellation (similar conclusion as Venrooij et al. (2011a)). When discussing the individualization of the models, the study tried to understand the roots of the variation between and within the subject’s BDFT dynamics. Similar to previous research performed in inceptor-based systems, Khoshnewiszadeh and Pool (2021) agrees that there is a modest contribution from the anthropometric characteristics to the differences in BDFT behavior. On the other hand, he also hypothesizes as in other studies that differences might appear from different strategies that can be used to perform the task and differences in muscle activity. Moreover, it was also shown that not all the tasks were suitable for the model-based cancellation approach since sometimes an amplification of the BDFT was obtained by applying the model to the step task. One possible explanation for the degraded performance in the latter case is the possibility of nonlinear effects appearing due to the low velocity and static friction present at the finger interaction at the end of each step movement.

It is important to realize that the model used to characterize and cancel BDFT in touchscreens lumps together all the contributions starting from the motion disturbance signals and ending at the screen input. In this way, the impact of the disturbances on particular body parts can not be straightforwardly identified. In other words, it is difficult to know what effect is attributed to what part of the body, and what part of the body is most susceptible to disturbances. Moreover, as it was previously described, by using this kind of lumped-parameter models, one can only hypothesize the cause of the differences in the BDFT behavior between and within subjects. One particular research direction that could help towards the understanding of how certain body parts respond to disturbances would be to investigate the use of motion-capturing techniques that can track the evolution in time of some pre-defined features (for example on the upper limb). In this way, the author considers that a better understanding can be obtained regarding how the body behaves in a disturbed environment and the investigation of the existence of possible patterns which could provide information regarding within/between-subject variability can be performed.

3.4. Conclusions

The limited body of knowledge in modeling the BDFT effects in touchscreen operations made it an interesting topic for further research. Moreover, models which can capture either within-subject variability or between-subject variability have been scarcely analyzed in the previous literature. It is of interest to further explore methods that can possibly capture, quantify and adapt to these types of variabilities in touchscreen operation for the application of cancellation of the BDFT effects.

Efforts have been made in the 21st century to create a unified framework to describe BDFT. It was observed that a relationship between the feedthrough of accelerations and the neuromuscular setting existed. As a consequence, the hypothesized dependence of BDFT on the HO's interpretation of the task, and on the reference limb positions was demonstrated, being a source of within-subject variability. Along with this, the between-subject variability of the BDFT was also investigated when models with different generality levels were used for model-based cancellation of BDFT effects (Venrooij et al. (2011a) and Khoshnewiszadeh and Pool (2021)). In this way, the importance of capturing in a model both the variations within a subject and between subjects was brought to attention. It is important to mention that the between-subject variability can have two causes, one originating from different anthropometric characteristics (which was shown to have little impact), and the second one originating from the different approach of performing the task (different neuromuscular characteristics).

A clear dependence of BDFT on the settings of the neuromuscular system was demonstrated in stick-based operations. Moreover, it was shown that for a ”good” cancellation performance the BDFT models in inceptor-based systems would need to be adapted to the current setting of the neuromuscular system. It needs to be recognized that one way of detecting the setting of the neuromuscular system is by using motion-capturing approaches. Since previous studies which investigated the BDFT in stick-based operations focussed on two distinct neuromuscular settings, namely ”stiff” and ”compliant”, a value-adding first step toward the creation of adaptive BDFT models for touchscreen operation would be to investigate whether these two distinct settings can be identified by using data from a certain motion-capturing approach and related to the input signal observed at the screen level.

4

Motion recording techniques

This chapter aims to give an overview of the most widely known techniques that have been used to record the motion of the body or limb of the humans. The chapter will start with a short description of the marker-based motion-capturing techniques, which currently is the most widely used approach in this field. Continuing, the use of accelerometers for recording the motion of the limbs will be investigated. Lastly, the performance and applications of markerless motion recording techniques will be investigated and compared with the previous methods.

4.1. Background information

The goal of the motion-capturing techniques is to provide a kinematic analysis of the different body segments by using information from a multitude of landmarks positioned on or recorded from the limbs. In this way, an idealized configuration of the segments of the body can be obtained, along with the positioning of the individual landmarks in the physical space. Moreover, if the positioning of the landmarks is obtained, the relative locations between the bodily segments can be resolved along the measurement period.

Motion-capturing techniques have been used in a wide range of applications. Clinical applications have made the most use of these techniques in areas such as: postural analysis and control, stroke rehabilitation, and range of motion analysis for rehabilitation purposes (Yahya et al., 2019). Other applications such as movement analysis in sports, used for the analysis of movement ergonomics, and in movies used for realistic movement replication are also two important fields that made use of these techniques (Chen et al., 2020).

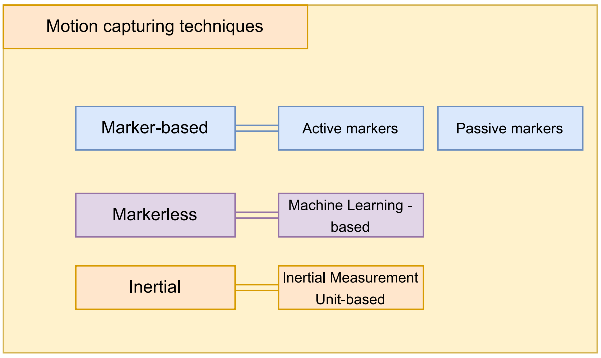


Figure 4.1: Overview of the most widely-used motion capturing techniques (Mugge, 2022).

Figure 4.1 presents a brief summary of the techniques that will be discussed in this report.

It needs to be mentioned that additional techniques which were not included in this representation exist and were not included since they were considered far from the scope of the current research topic. Saying this, the state-of-the-art in motion capturing (which is not included in the previous representation) is bi-planar videoradiography (fluoroscopy) (Wade et al., 2022). This method is based on bone movement analysis from continuous X-ray recordings. In this way, computer-aided software can be used for identifying the 3D centers of the joints and joint angles. This method has not been taken into consideration for further analysis since it can only capture one joint at a time (limited capture volume) and has potential drawbacks from the exposure to radiation (Wade et al., 2022).

4.1.1. Marker-based systems

The limitations of the previously-mentioned method have led to the marker-based techniques being the most commonly accepted method that has been used for capturing the motion. These systems make use of makers which are positioned on key locations of the human body. The two types of markers that can be used are passive (reflect infrared light) and active (emit light) (Cahill-Rowley & Rose, 2017). Special cameras (i.e infrared) depending on the marker type are used to record the position of the markers with a sub-millimeter accuracy (Buckley et al., 2019). In order to obtain the 3D location of the markers, triangulation methods are applied (i.e. Direct Linear Transform (DLT)) which combine the information (pixel locations) from multiple calibrated cameras. The combination of specially tailored cameras and markers allows for the precise identification of markers from the camera images. The 3D location of the markers is used to infer particular joint locations and joint angles. A possible disadvantage of the optoelectronic systems is represented by the need for careful positioning of the markers on the relevant parts of the limbs. In addition to this, the positioning of the markers is a time-intensive process and the repeatability of the measurements within the population study can be questionable (Wade et al., 2022). Moreover, optoelectronic systems are costly, and their use is restricted to special laboratory environments requiring multiple camera systems (D'Antonio et al., 2021) and specially trained people. Moreover, since the markers are applied on the skin, artefacts linked to the relative motion between the skin and the bones can be introduced which can create errors in the interpretation of the joint rotations (Schmidt et al., 1999). Lastly, the 3D position of the markers is facilitated by the fact that each marker is seen by at least two cameras (such that the triangulation procedure can be performed). In the case of movements that require large excursions, the occlusion of the markers can occur which can disrupt the measurement flow.

4.1.2. Inertial-baseds systems

A cheaper and less-restrictive alternative for capturing the motion is the use of Inertial Measurement Units (IMUs) (Palermo et al., 2014). The IMUs can indicate the linear acceleration of the device along with the angular velocity and, if also a magnetometer is introduced, the orientation with respect to the local magnetic field (Yahya et al., 2019). In this way, the positions and orientations of the analyzed limb segments can be obtained with respect to a non-moving reference frame. Usually, these can be obtained by combining the accelerometer measurements which can be integrated in time with kinematical models (Zhou & Hu, 2007). The principle of obtaining the orientation of the limbs with respect to a global frame is based on a series of 3D rotations and on the assumption that the IMUs are positioned on rigid segments represented by the limbs. In this way, an important step in obtaining an anatomically correct joint rotation, a sensor-to-segment calibration procedure needs to be performed where the orientation of the IMU with respect to a consistently-chosen anatomical axis is obtained (Roetenberg et al., 2009). It is important to realize that accelerometers do not provide perfect measurements of the quantities of interest. Most often, the obtained quantities are noisy and contain inherent biases which affect in turn the estimation of the desired measures. A common way to correct these problems is by using Kalman Filters (Roetenberg et al., 2007). A possible issue in some environments can be also represented by the inability of using the magnetometer due to ferromagnetic disturbances. In this condition, the reference yaw position can not be obtained reliably anymore. Studies such as Baldi et al. (2019) or Ligorio et al. (2020) have developed methods, and calibration techniques in order to overcome the limitation of not using the magnetometer information. One possible advantage of using accelerometers as a motion-capturing technique is the fact that they don't need specialized laboratories for performing the experiments. Moreover, the occlusion of the body parts does not need to be taken into account since this method does not rely on cameras. One possible disadvantage is the need for a calibration procedure, which can affect the outcomes of the measurements if not performed correctly Ligorio et al. (2020). Moreover, it is important to investigate the repeatability of the measurements within a subject and between the subjects (Öhberg et al., 2019). Lastly, it needs to be mentioned that most of the studies that have used this method did not assess the movement of the limbs in disturbed environments and therefore, the effects of disturbances on the estimated limb kinematics is currently not widely known.

4.1.3. Markerless systems

Recent advancements in computer vision and machine learning techniques facilitated the introduction of markerless motion-capture techniques. Markerless motion-capture techniques are able to detect body-specific landmarks from images or videos. Based on the dimensionality of the output features, the markerless can be categorized into 2D pose estimators and 3D pose estimators. Two of the most popular open-source 2D pose estimation algorithms are OpenPose (Cao et al., 2017) and DeepLabCut (Mathis et al., 2018). Moreover, based on the number of persons that can be identified, the 2D pose estimators can be further split up into single or multi-person pose estimators (Liu et al., 2022). Most of the techniques are based on Convolutional Neural Networks, making use of an encoder (high-level feature extractor) and decoder (estimation of the keypoint locations) (Liu et al., 2022). Again, two main approaches of implementing the decoder is by having a detection-based approach where heatmaps of the features are obtained (resulting in the features with the highest probability) or by having a regression-based approach (directly obtaining the target parameters). When discussing the problem of multi-person pose estimation, two main approaches can be identified: the top-down and bottom-up paradigms. Regarding the top-down approach, this framework identifies each person separately and then extracts the *”bounding box”* (Liu et al., 2022) features for each person. On the other hand, the bottom-up approach detects the features for all persons together, and then each feature is assigned to one person. Two of the most It needs to be mentioned that one disadvantage of the markerless-based approaches is represented by a decreased accuracy when compared to the other methods (especially the marker-based approaches). One possible reason for this is that these approaches depend on a set of manually labeled data which they are trained on. This, along with possibly limited scenarios in which the Neural Networks (NN) might be some of the aspects which lead to inaccuracies in these methods. Moreover, it is expected that their performance under occlusion situations and in poor lighting conditions to decrease. One advantage of these approaches is that they are non-intrusive and can be applied to videos/images captured by any camera available, making them a suitable approach to be investigated in non-laboratory specific environments or in space-constrained places.

The following sections will elaborate on the applications of the previously-mentioned methods in practical scenarios.

4.2. Applications of the motion capturing techniques

4.2.1. Applications of the marker-based systems

As described in the previous section, marker-based motion recording systems have mostly found applications in the clinical medicine domain. Their main goal is to provide a quantitative indication of the Range-Of-Motion (ROM) for different parts of the human body (i.e. gait analysis), with the goal of evaluating the potential implications of certain conditions or even to early diagnose some of these. On the other hand, another possible application sector is the sports domain (van der Kruk et al., 2018), where careful analysis of motion ergonomics can be made in order to enhance athlete performance. Both areas have in common the fact that most of the analyses are made by evaluating the joint angles (i.e. relative rotation between two adjacent body segments) instead of analyzing specifically the 3D location of each marker.

Researchers have tried to evaluate the accuracy and applicability of marker-based systems against the state-of-the-art, fluoroscopy approach. One such example is presented by Fiorentino et al. (2017) which compared the two approaches in the evaluation of the hip-complex ROM. The study investigated static poses and dynamic walking scenarios with the goal of analyzing the effects of soft tissue artifacts on the estimation of the pelvis, thigh, and hip joint angles. It was reported that the soft skin artifacts associated with the relative movement of the markers with respect to the skin were dependent on the subject and on the activity that was performed. Moreover, a general conclusion made by the study was that the apparent ROM was underestimated in the case of using the markers applied on the skin. Taking a step back, it might be the case that the joint angles might not be a useful quantity when analyzing the biodynamic response of the body under disturbed environments. Nevertheless, it can be expected that if the logistical drawbacks would not make this method an infeasible one, individual markers could be used for tracking in the 3D space particular limb landmarks. Saying this, the same study analysed also the position error between the two methods, reporting an error of at most $1[cm]$ in all directions.

One possible example of using marker-based systems for upper limb kinematic analysis is the research proposed by Figueiredo et al. (2015) on the topic of analysis of joint movement for subjects with cerebral palsy. It has been hypothesized that patients with cerebral palsy exhibit certain movement patterns that reflect a necessary accommodation of the subject to the task at hand due to limitations in the available movement space.

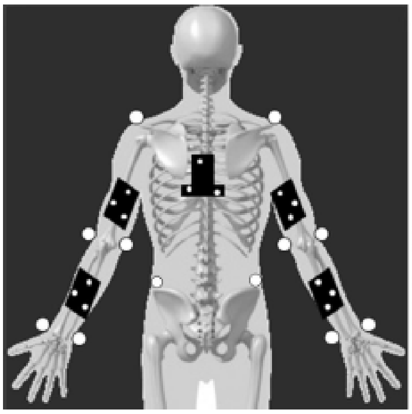


Figure 4.2: Overview of the set-up of the markers used for the kinematic analysis of the upper limbs (Figueiredo et al., 2015).

Figure 4.2 presents the upper body marker distribution that was used to track the upper limb kinematics. The three investigated segments namely the upper arm, forearm, and trunk were assumed to be rigid elements represented by the single white markers positioned on the distal and proximal anatomical landmarks. In turn, each defined segment's orientation and position were tracked by the cluster sets of three markers. The reason for using three-marker clusters is the fact that three points are needed to define a plane (orientation in space with respect to a reference system). The positioning of the markers and computation of the joint angles was defined in accordance with the guidelines proposed by the international society of biomechanics (Wu et al., 2005).

In a separate study, Rawal et al. (2018) investigated the upper limb ROM for driving applications. It was mentioned that the joint angles were obtained by combining the marker-based system with a rigid-body analysis and inverse kinematics. Moreover, it was recognized that the assumption for the rigidity of the upper limbs could be considered valid when the joint angles would need to be computed. The results of the articles are out of the scope of the current study, and therefore they will not be further investigated.

From the student's knowledge, very few studies analyzed the use of marker-based systems in the biomechanical analysis in disturbed environments. Jack et al. (2013) researched an experimental set-up consisting of a mechanical lever arm and reflective markers in a disturbed environment. As previously indicated, the study recognized the importance of the complex interaction effects of posture and muscle activation on the feedthrough of accelerations through the human body. The goal of the study was to investigate whether an accurate acceleration description could be obtained from the 3D tracking of a set of markers. The acceleration time traces were obtained by double differentiation of the position signals obtained from the camera tracking system. The study reported that the marker-based system could be used to feasibly detect translational vibrations up to $10[Hz]$ and rotational vibrations up to $3[Hz]$. Moreover, when compared to accelerometer measurements, the system was able to replicate the RMS linear and peak acceleration measurements with less than 5% difference suggesting that it could be a viable option for obtaining acceleration measures for seat-to-head transmissibility investigations. The authors recognized the sampling rate and the resolution of the cameras as two important parameters that can affect the quality of the results. The sampling rate is an important quantity since it dictates the frequency band that can be captured in the movement analysis (a low sampling frequency can ”lose” some information from the movement). Moreover, another aspect that needs careful consideration is the differentiation of noisy unfiltered position signals which can amplify the noise. Lastly, the six cameras used by the motion-capturing system might present a great disadvantage in terms of the mobility of the equipment and applicability to other scenarios such as motion platforms possessing greater dimensional constraints. Given this aspect, the study recommends the investigation of the feasibility of markerless systems for capturing motion.

4.2.2. Applications of the inertial-based systems

Most of the studies investigating the use of inertial-based systems for tracking motion focused on obtaining either the joint angles from devices positioned on adjacent segments, and/or the limb positions by combining the accelerometer, gyroscope, or magnetometer information with kinematic models. It is important to distinguish the two main applications of inertial-based systems that have been introduced in this report. If, in the previous sections accelerometers were used to understand how the disturbances are transmitted through the human body, this section describes the use of inertial systems for the kinematical analysis of the limbs.

One of the first studies which investigated the use of accelerometers in the estimation of the upper limb kinematics is the report presented by Zhou and Hu (2005). The goal of the research was to investigate methods of obtaining the joint kinematics of the upper limb with the help of one accelerometer mounted on the wrist and a kinematical chain model. The study used the intuition that the upper limb motion can be considered as a motion of rigid body segments in order to create a kinematic chain model which linked the shoulder to the wrist position. Since the only source of information was the wrist position (from integrating the accelerometer outputs), the kinematical model was used for an inverse kinematical analysis which was used to derive the elbow inertial position from a known wrist position. Due to the fact that the inertial sensors possess noise and biases, it can be expected that by integrating the raw values, the obtained wrist position would drift in time. In order to solve this problem, the researchers proposed an optimization approach based on the constraint that the length of the segment is constant. When comparing the method with a marker-based approach, deviations in the inertial coordinates were at most $3.5\,[cm]$. Part of the discrepancies were hypothesized to be present due to non-rigid behavior and tremors/vibrations of the upper limb.

The study performed by Cutti et al. (2008) proposed a methodology based on four accelerometers to investigate humerothoracic, elbow, and scapulothoracic kinematics in ambulatory conditions. Four accelerometers were positioned on the sternum, scapula, humerus and wrist in order to be able to record the relative rotation between the adjacent segments (segments assumed rigid). The methodology consisted in the positioning of the sensing units, a definition of a set of anatomically meaningful reference systems which are used for the joint angle computations, and their evaluation with respect to the set of axes defined by the sensing units. The definition of the anatomical axis system with respect to the sensor axis system was performed during a static trial, where the subject was instructed to stay still with the hands oriented downwards. In order to compute the joint angles, the axes of the anatomical reference frames were continuously updated based on the orientation of the sensing units and on the relation between the two axes, and individual angles were obtained by selecting a certain Euler-angle sequence. The comparison of the previously-mentioned approach with a marker-based system provided satisfactory results. It needs to be mentioned that similar to the marker-based systems, the approach of using accelerometers can also suffer from soft-tissue artifacts. Taking a step back, the relevance of obtaining the relative joint angles in the analysis of the feedthrough of disturbances is questionable. On the other hand, it is expected that the time traces of joint angles could possess some information regarding the stiffness of the neuromuscular system.

Similar to the marker-based systems, an approach based on IMUs requires careful positioning of the sensors on the subject's limbs. Moreover, it can be considered rather difficult to be able to replicate the same positioning across all the participants. In this way, it can be expected that the repeatability of the results would be affected which can introduce potential confounding factors. Saying this, Höglund et al. (2021) recognizes that a large between-study variability exists when it comes to the kinematical analysis of the upper limb, possibly being caused by a lack of standardization of sensor placements, procedures used for calibration, and biomechanical models. In this way, the study had the goal of investigating how upper limb kinematics are affected by different placement of IMUs. In order to investigate concurrently the different placement of IMUs, two pairs of IMUs were simultaneously mounted on the scapula, humerus, and lower arm. The results have shown that the between-sensor variability is lower for tasks performed on a single plane compared to more complex, compound tasks. Moreover, it was also observed that the different positioning of the sensors affected in a different way the kinematic results for the different segments of the upper limb. The sensor placement had a bigger influence on the kinematic behavior of the scapula (abduction and flexion) and of the elbow (prono-supination). One possible reason for these differences is the activation of the muscles which can disturb the positioning of the IMUs.

In order to be able to derive the relative rotations between two segments, a calibration procedure is needed for the measurement units. In this way, the relative rotation between the IMU axis and the defined anatomical axis is obtained. Ligorio et al. (2017) recognized that there are two main ways in which a calibration can be made. The anatomical method requires a static pose to be performed in a short time, and the sensor-to-segment calibration is made by imposing some assumptions about the position of the limbs in the static pose. A second calibration method is called the "functional" method, which requires the subject to perform planar motions in order to estimate the anatomical rotational axes. The functional method made use of an eigenvector analysis in order to estimate the elbow flexion axis, and the longitudinal axes of the lower arm and of the upper arm (again by assuming a kinematic chain model for the upper arm and lower arm segments). One of the main advantages of functional calibration methods is that they don't need an accurate placement of the sensors. Moreover, it has been reported that it can be possible to obtain more accurate elbow joint angles with functional calibration methods, compared to methods relying on an anatomical definition. Lastly, when compared to a marker-based system, the kinematical analysis provided by the sensors had a deviation of maximum $4[deg]$.

As previously presented, the IMUs consist of a set of accelerometers, gyroscopes, and magnetometers. The fact that the individual measurements from each sub-sensor can not provide directly the orientation of the IMU in inertial space and some of the quantities can be affected by noise, bias, and magnetic disturbances has inspired researchers to come up with novel solutions to solve for these shortcomings. One example is the study proposed by (Ligorio et al., 2016) which investigates a method consisting of an Extended Kalman Filter and the TRIAD method, in order to accurately obtain the orientation of the sensor. It needs to be recognized that under some circumstances, magnetic disturbances can be present in the environment in which the IMUs operate. Saying this, in those situations, the magnetometer information can not be reliably used anymore. Studies such as Truppa et al. (2021), Ligorio et al. (2020) or Truppa et al. (2022) investigated methods that have bypassed the use of the magnetometer for computing the orientation of the IMU using Kalman Filters or time propagation of fictitious axes based on initial measurements. One very important consideration was made in the study proposed by Truppa et al. (2022) referring to the fact that the Kalman Filter Algorithm was tested with either quasi-static or slow movements implying that the gravity vector was easily obtained. The study acknowledges that the algorithm's accuracy was not investigated in fast dynamic situations, which are certainly encountered in disturbed environments. Taking a step back, it has been observed that most of the studies that were investigated analyzed rather slow movements under static situations. In this way, it is difficult to hypothesize what are the effects of disturbances on the quality of the kinematical analysis when using IMUs.

4.2.3. Applications of the markerless systems

Both of the previously-described methods relied on either physical markers or inertial measurement units to be positioned on the limbs of the subjects. Moreover, both systems require a careful calibration procedure in order to make sure that anatomically meaningful quantities are obtained from the movement analysis. In recent years, researchers have focussed on alternative methods for identifying certain features of the human body without the need of mounting any device. One possible method for solving this problem is the use of pose estimation algorithms based on deep-learning techniques applied on pictures or videos.

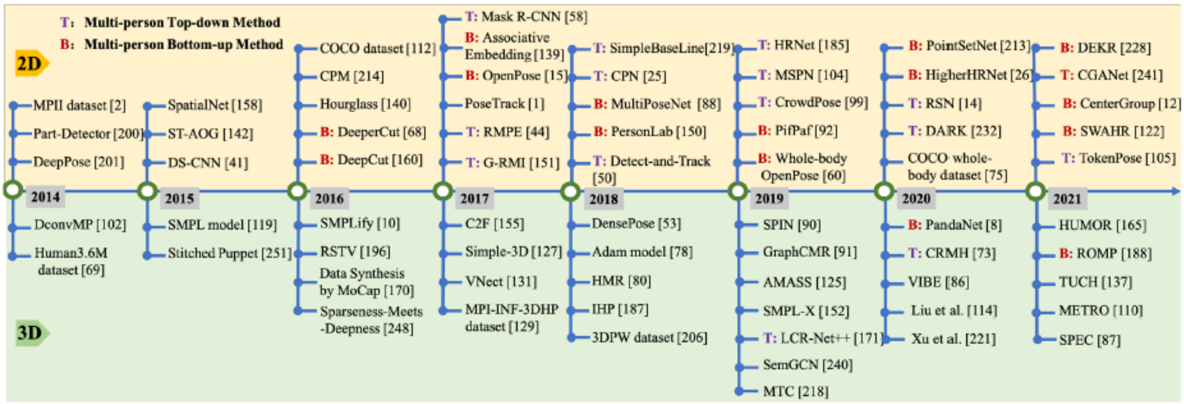


Figure 4.3: Timeline of the publication of the most well-known 2D and 3D pose estimation algorithms (Liu et al., 2022).

Figure 4.3 presents an overview of the published works and datasets linked to 2D and 3D pose estimation techniques. As can be observed, the body of knowledge in this field is large, and a thorough analysis of all its details can represent a literature study in itself. Saying this, the following section will only present a short description of some of the characteristics and applications of these methods.

Markerless methods can be split-up in two main categories based on the number of cameras that are being used, namely single-camera (monocular) and multi-camera systems (Wade et al., 2022). Furthermore, monocular systems can be used for both 2D and 3D detection of the features of the human body, whereas multi-camera systems are used for the 3D detection of these features. Two main approaches have been used for the representation of the human body in these approaches. The *"keypoint-based representation"* (Liu et al. (2022) p.p. 80:6) provides a simple overview of the keypoints of the body representing joint centers such as the elbow or shoulder. On the other hand, a richer description is provided by the *"model-based representation"*, where models which replicate the structural characteristics of the human body are exploited, such as volumetric models.

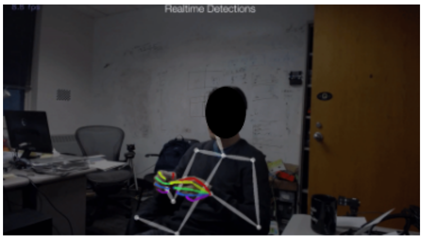


Figure 4.4: OpenPose features extracted from a human subject. [1]

When discussing monocular 2D keypoint approaches, OpenPose (Cao et al., 2017) can be considered one of the most intuitive and popular methods for the detection of the keypoints. The algorithm provides options for detecting up to 135 features ranging from different body parts to hand features. As pointed out by Wade et al. (2022) and observed in Figure 4.4, this algorithm can extract two features per each defined segment, defining the joint center locations in the proximal and distal areas. This aspect prohibits the evaluation of the 6DOF of the segment due to the fact that an additional point would have been needed to be defined in order to create a plane that can describes the rotation about a particular axis (Wade et al., 2022). One important aspect that needs to be recognized in the evaluation of features extracted from one camera is that the depth information is non-existent. In other words, the feature space is only present on one plane (the plane of the image). In this way, Wade et al. (2022) suggests that in order to obtain relevant biomechanical information from a pure 2D analysis, it needs to be assured that the camera plane is parallel to the plane of movement. If the camera plane is not parallel to the plane of movement, the joint angles could be affected by the projection error imposed by the monocular condition. Another example of depth loss is the example of Figure 4.4, where the left/right wrist and elbow keypoints might appear on the same plane on the image, but in reality there is a depth difference present.

An additional popular markerless motion recording algorithm is DeepLabCut developed by Mathis et al. (2018). One particular novelty exploited by the previously-mentioned methods is that it enables tracking user-annotated data or, in other words, features that are defined by the user by using transfer learning. Although not very prominent in the literature, several studies compared the planar 2D markerless estimation techniques with 3D marker-based solutions when using DeepLabCut's algorithm. Based on (Wade et al., 2022), several studies such as Drazan et al. (2021) (jumping) or Moro et al. (2020) (gait analysis of stroke survivors) focussed on the movement analysis of the lower-body limbs in the sagittal plane. The suitability of using markerless systems was enforced by the reported difference of $10 - 20[mm]$ compared to the marker-based systems when identifying the joint centers (Wade et al., 2022). On the other hand, the study performed by Stenum et al. (2021) which also investigated the human gait during walking reported spatial errors of 1 to $20[cm]$. Unfortunately, the authors have not

---

[1] https://github.com/CMU-Perceptual-Computing-Lab/openpose

reported any explanation why the differences were so big. One hypothesis offered by Wade et al. (2022) is the possibility of the estimation algorithm of wrongly labeling the limbs (i.e. left limb instead of right limb). In this way, the recommendation made by Wade et al. (2022) is that the positioning of the cameras shall be carefully analyzed such that the risk of occlusion is low. Nevertheless, this fact implies that an analysis of the expected movement profiles shall be made in order to eliminate potential factors which may influence the quality of the results. Lastly, given the accuracy evaluations of the previously-mentioned studies, it is expected that the overall performance of the system to be dependent on the type of movements that are being performed and the characteristics of the camera such as resolution and sampling rate.

It needs to be mentioned that monocular sets of images/videos can also be used for obtaining 3D keypoint configurations. The problem that focuses on transforming 2D information into 3D is a complex one, since information obtained from monocular systems induces depth ambiguity (Liu et al., 2022). In order to obtain depth information, methods that exploit prior geometric knowledge, statistical models for the limbs, and which use temporal information have been developed (Liu et al., 2022). One example of such a method is presented in the study researched by Chen et al. (2020), which uses neural networks which exploit both spatial and also temporal features to resolve the 3D pose estimation problem. When evaluating the methodology against a pre-labeled dataset with 3D features (thus not against any other motion analysis technique), the method performed with an accuracy of 30 to $40[mm]$. Another representative study that can be considered is the research performed by Rempe et al. (2021) which focuses on using already-labeled 2D data (i.e. from OpenPose) to estimate the motion in 3D, the shape of the body, the ground plane, and the contact points. The method is based on a generative and autoregressive model capturing the 3D motion dynamics. By training the autoencoder, a model of a distribution of pose transitions is obtained. The autoencoder is trained on a manually-labeled 3D dataset from motion capture experiments and is used as a prior for the motions at a specific point in time, helping in obtaining the depth information for the motion. One possible drawback of this method is the presence of rather large errors ranging from 26 to $35\,[cm]$ when comparing the 3D keypoints obtained from inputting 2D keypoints from OpenPose with manually-labeled 3D joint positions from a data-set. Besides the present ambiguity in depth information encountered by the monocular 3D markerless techniques (Liu et al., 2022), these methods were demonstrated to be slightly less accurate compared to their multi-camera counterparts (Wade et al. (2022), Desmarais et al. (2021)). Moreover, as Liu et al. (2022) points out, there is a lack of enough *in-the-wild* 3D datasets that are used to train these methods, meaning that their generality may be low (when compared to 2D datasets) or restricted to a certain camera set-up or perspective. Lastly, it was observed that validation studies of the monocular 3D methods in biomechanical applications are very scarce (according to the student's knowledge).

The intuition that a set of multiple images can provide much more information than a single one (Desmarais et al., 2021) inspired researchers to combine markerless systems with multiple calibrated cameras with the goal of obtaining a 3D representation of the features of interest. The principle behind this approach is to calibrate two or more cameras and in turn, obtain their relative location with respect to a global frame. In this way, if the same landmark is seen by two or more cameras, it can be triangulated, and its 3D location can be obtained in either a global frame or coincident with one of the camera frames. Both OpenPose and DeepLabCut provide information regarding using multiple cameras for triangulation. Usually, software is available online which can help with camera calibration and triangulation. For example Python/OpenCV (Bradski, 2000), MATLAB ("Computer Vision Toolbox Camera Calibration", 2023) or separate studies such as Theriault et al. (2014) (suggested by Mathis et al. (2018)) provide ready-to-use protocols or software for camera calibration.

The study performed by (Mathis et al., 2018) provides an intuition regarding how the labeled 2D data can be transformed in 3D. The study does not provide a detailed description of how the camera calibration was made, but describes that a set of 6 cameras (90 Frames Per Second and a resolution of 1080 pixels) were used in combination with the OpenCV library (which performed camera calibration and triangulation) in order to obtain a 3D description of a particular scene.

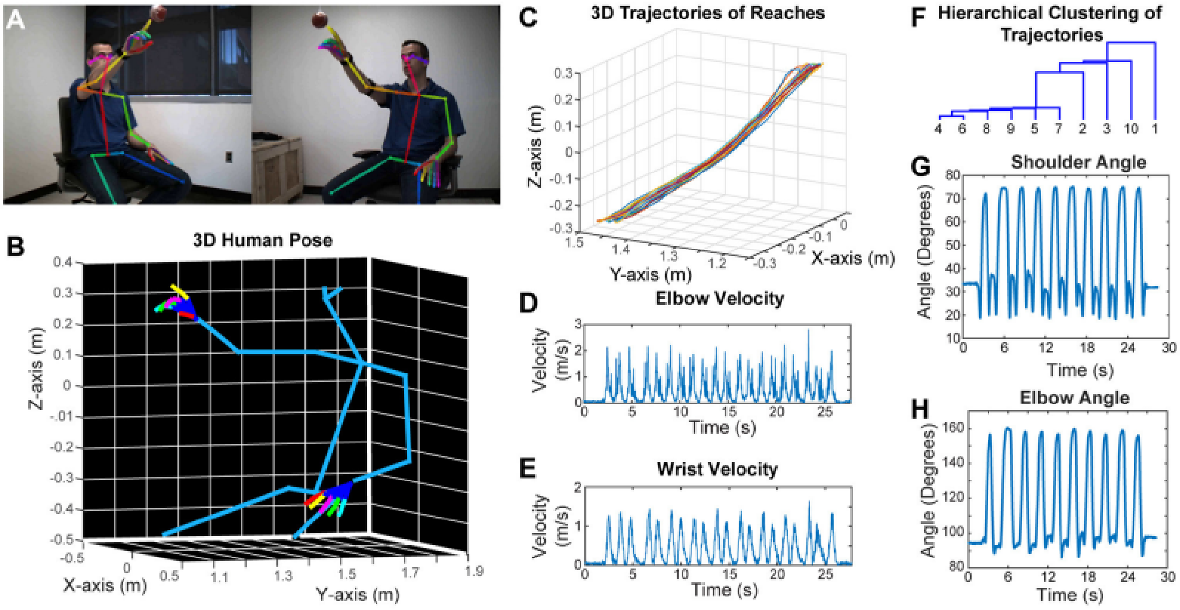


Figure 4.5: Oveview of the triangulation results for obtaining the 3D pose of a subject (Arac et al., 2019).

The study performed by Arac et al. (2019) (Figure 4.5) proposed a combination of OpenPose's 2D keypoint detection algorithm with a two-camera set-up (170 Frames Per Second) and MATLAB's Camera Calibration toolbox. In order to obtain the intrinsic and extrinsic parameters of the cameras, the researchers used a checkerboard pattern (10x 7 with 11.5 x $11.5[cm]$ squares) in combination with the previously-mentioned toolbox. Based on Figure 4.5, it could be observed that besides the 3D trajectories, the keypoint velocities were computed during a set of 10 reaches. The computed velocities need to be carefully analyzed due to the possibility of noise amplification in case of noisy position signals. One possible disadvantage of the study is that the approach was not validated using alternative motion capturing systems.

Nakano et al. (2020) performed a study where a multi-camera markerless system was compared to a marker-based system. The study included both slower movements like walking and faster tasks such as ball throwing. The 2D pose estimation with OpenPose was combined with five cameras which ran at either $120[Hz]$ or $30[Hz]$. The calibration procedure was performed using a calibration pole whose global coordinates were known. The synchronization of the videos from each camera was made with a bright light which represented the reference frame. The 2D to 3D extrapolation was made with the DLT method (Miller et al., 1980). The study reported a mean absolute position error between the time-series data (keypoints) recorded with the two methods to be less than $30[mm]$ in 80% of the trials. The presence of errors can be attributed to the not always correct manually-labeled training data used to train the markerless detection algorithm. There were also cases where the errors were slightly bigger, being attributed to OpenPose failing to detect the pose of the subject or attributing wrong keypoint positions. Other possible sources of error include the synchronization between the cameras or occlusion of some body parts. It is also important to mention the contribution of the camera's sampling rate and resolution to the accuracy of the results. In general, it can be expected that a camera with a higher sampling rate would lead to a more detailed capture of the motion. On the other hand, the accuracy results of the study which compared a camera with a better resolution and lower framerate with a lower resolution and higher framerate are inconclusive with a slight advantage over the camera with a higher framerate. The accuracy can be also affected by the speed of the task. It was observed that on average, the joint centers during slow movements were identified more accurately compared to the case of faster movements.

A similar study was performed by D'Antonio et al. (2021), which investigated the comparison between markerless triangulation based on OpenPose and an IMU system when analyzing the lower limb movement joint angles during walking/running. Besides the previously-mentioned comparison, the study also compared different sets of camera configurations.

The motions were recorded with two webcams which operated at 60 frames per second, being placed $1\,[m]$ apart. Moreover, the angle between the cameras was between 40 and $50\,[deg.]$, the different camera configurations preserving the relative positioning between each other. The calibration of the cameras was performed with the help of MATLAB's Stereo Camera Calibration app, using a checkerboard pattern (Figure 4.6) with 45 paired images. The study also proposed a manual validation of the calibration of the camera system based on epipolar geometry, by computing the average reprojection error of a triangulated feature during one trial. As in the previous study, the largest contribution to the errors was represented by the occlusion phenomenon which influenced the correspondence between two features in the two image planes. Moreover, the kinematical accuracy was affected by the positioning of the cameras, the activity that was performed, and also by the possible mismatch between OpenPose's predictions of the same real landmark on the separate picture planes. The study suggested an increased number of cameras would be suitable for better accuracy and better coverage of the body parts. The reason for this statement is that by using multiple cameras, the influence of one possible inaccurate detection of a feature by one camera can be counterbalanced (less weight) by the assumed accurate features detected by the rest of them. Although consistent angular trends were observed between the markerless system and the IMUs, the proposed method was not entirely suitable for an accurate kinematic analysis of the lower body flexion/extension during gait due to errors in joint angles larger than $5\,[deg]$. Another source of error that was not described by the authors might be represented by a mismatch between the anatomical axes defined by the accelerometers and the ones used by the markerless systems.

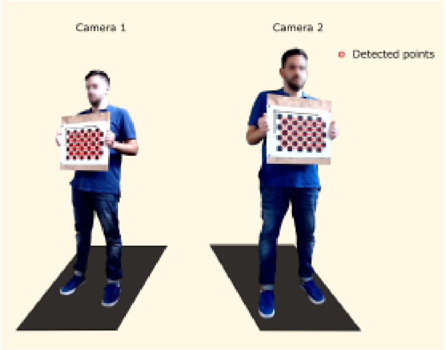


Figure 4.6: Overview of the checkerboard pattern used for calibration (D'Antonio et al., 2021).

In their review of markerless motion systems, Wade et al. (2022) included one of the few studies which compared three open-source markerless systems with the state-of-the-art marker-based system. Needham et al. (2021) compared markerless motion capturing systems such as OpenPose (Cao et al., 2017), DeepLabCut (Mathis et al., 2018) and AlphaPose (Fang et al., 2017) against a marker-based system with 15 cameras. The markerless systems consisted of 9 cameras which recorded data at $200\,[Hz]$. The results showed that OpenPose provided the closest results compared to the marker-based approach when using the stereo-view system. It was reported that feature differences of $16-34\,[mm]$ (walking), $23-48\,[mm]$ (running) and $14-36\,[mm]$ (jumping) were obtained.

The previously-mentioned study presents a series of current challenges that apply to markerless motion-capturing systems. Firstly, it is expected that the pose estimation algorithm can be affected by the clothing the subject uses. The *"apparent"* change in the shape of the limbs when using loose clothing might be a factor that affects the accuracy of the joint estimation. Moreover, as mentioned in the 3D monocular systems, the performance of the estimations depends highly on the datasets that the algorithms were trained on. Saying this, it can be expected that the performance of the estimations to drop in situations/movements that were not *"seen"* during the training. At the same time, not all datasets were specifically elaborated specifically for biomechanical applications. In other words, the annotated joint-center locations could not always be anatomically correct.

Other validation studies such as Lim (2019) or Kosourikhina et al. (2022) have investigated the accuracy of multi-camera 3D markerless systems. Lim (2019) used four cameras in the evaluation of joint angles during functional movements. When compared to a marker-based system, it was reported that the kinematic data under the form of joint angles obtained at the elbow level was not significantly different for a markerless system using OpenPose. Kosourikhina et al. (2022) reported a good agreement between two similar systems. The study recognized that the camera calibration step was slightly sensitive (but crucial), with iterations being needed in order to obtain accurate results for the stereo camera system. Another point of attention was the necessity of video synchronization between the cameras. Lastly, the report mentioned the possibility of obtaining *"excellent"* kinematic information by using markerless systems, with a relatively small accuracy tradeoff when compared to the use of the less flexible state-of-the-art systems.

Lastly, the study researched by Zago et al. (2020), focussed on analyzing the effects of the relative distance between the cameras, the movement direction of the subject, and the resolution of the cameras when comparing a multi-camera markerless system with a marker-based one. The focus was directed toward the analysis of the 3D gait parameters (evolution in time of the 3D landmarks on the lower limb) during walking. Intuitively, it was pointed out that if the subject covers a large distance in the working volume, several features can not be tracked accurately anymore due to the varying spatial resolution of the cameras. In other words, a feature can be observed "better" when it is positioned near a camera, compared to the case when it is positioned at a large distance with respect to the same camera. Moreover, another implication of this aspect is the fact that the pixel-to-distance conversion depends on the position of the subject with respect to the cameras Zago et al. (2020). In this way, it is hypothesized that not only it is difficult to track the desired features, but also the accuracy of the triangulation can be affected. One other interesting aspect described by this study concerns the increase in 3D reconstruction accuracy with the increase in relative camera distances (ranging from 1 $[m]$ to 1.8$[m]$). It is proposed that by increasing the relative distance between the cameras, the features are being observed from different perspectives (increased convergence) with the tradeoff that the allowable volume of observation is decreased. This fact tends to agree with the study researched by D'Antonio et al. (2021), which mentions the possibility of obtaining a negligible triangulation error if the cameras are positioned at 90 $[deg]$ relative to each other with the risk of not being able to observe the desired feature from both of the cameras.

The advantage of markerless motion capture approaches is that they are non-invasive, easily deployable (compared to the previous approaches), and relatively cheap. On the other hand, they are less accurate than the golden standard and suffer from problems regarding occlusion or false detections of the body features. However, this research will assume that only one hand of the subject would need to be tracked while he/she is seated (the distance from the subject to the camera does not fluctuate much) and the subject is required to perform rather slow tasks, representing mostly a static scenario unlike the previously-mentioned studies (cameras can be positioned close to the subject for more accurate joint detection). Taking this into consideration, it is assumed that the conditions imposed by the research will minimize the disadvantages portrayed by the markerless motion capture systems, making them a viable solution for further investigation. On the other hand, it is important to take into account that the performance of markerless systems in disturbed environments has not been analyzed before and represents an important aspect that needs to be investigated.

4.3. Conclusions

Advancements made in machine learning and image processing techniques helped with the introduction of markerless approaches for motion-capturing techniques. One of the most promising approaches from the author's perspective is represented by the 3D multi-camera approaches. These approaches make use of 2D markerless systems (i.e. OpenPose) combined with at least two calibrated cameras which can be used to triangulate the 3D position of the features. Special libraries in OpenCV or MATLAB have been constructed that can help with automatic camera calibrations using a checkerboard pattern. Several studies tried to validate these approaches, reporting a performance that is dependent on the speed of the performed action, camera parameters (resolution and sampling frequency), and camera positioning (minimizing occlusion, the distance to the user, and triangulation errors). Studies reported a rather acceptable accuracy of identifying body landmarks when compared to marker-based systems (mostly in the order of 10 to $40[mm]$. Even though one study has reported unacceptable joint-angle computations (for clinical use), the overall trends in the movement were equivalent to the ones obtained using accelerometers reinforcing the fact that even though in some cases the accuracy might not be sufficient, the markerless techniques can replicate trends observed in movement. It is considered that given the rather static scenario that the subject needs to perform (operating a touchscreen while seated), the fact that the upper limb that operates the screen would need to be tracked (minimizing occlusions), and the rather slow character of the movement, this approach can represent a valid option for the 3D analysis of motion under disturbed conditions.

5

Preliminary Biodynamic Model Analysis

The goal of this chapter is to give the reader an overview of the theoretical background that has been used for the preliminary analysis linked to the proposed biodynamic model of the upper arm. Moreover, an overview of the assumptions and modeling choices for idealizing the upper limb will be presented. The latter parts of the chapter will elaborate on the presentation of several scenarios that were tested in order to gain an intuition of the behavior of the multibody system in a disturbed environment when parameters linked to the geometry and the equivalent stiffness are altered.

5.1. Multibody dynamics background

The contents of the sections concerning multibody dynamics equations of motion and modeling approaches are tightly linked to the following sources: Moore (2022), Schwab (1998), and Ross (2020). The information from these sources will be arbitrarily combined in order to obtain a logical sequence of ideas.

5.1.1. Methods for obtaining the equations of motion

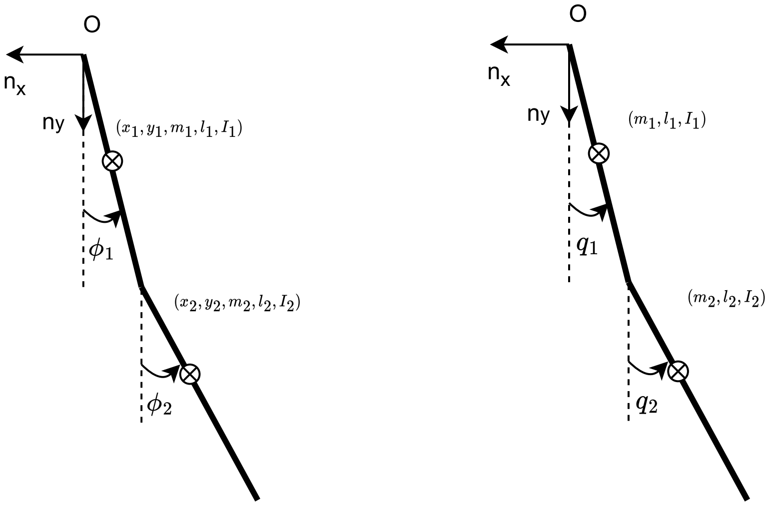


Figure 5.1: Representation of a multibody system with cartesian coordinates (Left) and with generalized coordinates (Right).

In order to better understand the differences between different methods that are being used to obtain the equations of motion from multibody systems, a use-case consisting of two bodies that can move in 2D has been proposed (Figure 5.1). Focussing on the left-most representation, each body has a particular mass $m\,[kg]$, a length $l\,[m]$, two coordinates that describe the position of the center of mass with respect to the inertial frame n_x/n_y (COM) $x\,[m]$ and $y\,[m]$, one angle which describe the orientation of the body with respect to the inertial frame $\phi\,[rad]$ and rotational inertia with respect to the (not presented) inertial z-axis $I\,[kg \cdot m^2]$. This classical multibody problem assumes that the first body is fixed with respect to the point O in the inertial coordinate frame, and is linked with the second body with a joint that enables a rotation with respect to the (not shown) inertial z-axis.

If the classical Newtonian approach is considered, each body would have 3 equations of motion linked to the change in linear momenta in the x and y axes and the change of angular momentum around the z axis resulting in 6 total equations. Since the system is constrained in two points, the total number of unknowns is 10 (the double derivatives of the x, y, θ quantities and the four reaction forces which keep the first body coincident to the point O and the second body linked to the first body). The other four equations are obtained with the help of the constraint equations which are differentiated twice with respect to time being also a function of the states of the bodies.

$$\begin{bmatrix} \boldsymbol{M} & \boldsymbol{A} \\ \boldsymbol{A^t} & \boldsymbol{0} \end{bmatrix} \cdot \begin{bmatrix} \ddot{\boldsymbol{x}} \\ \boldsymbol{F_r} \end{bmatrix} = \begin{bmatrix} \boldsymbol{F_{ext}} \\ \boldsymbol{c}(\boldsymbol{x}, \dot{\boldsymbol{x}}) \end{bmatrix} \tag{5.1}$$

Based on Schwab (1998), Equation (C.1) presents the matrix representation of the entire system of equations for all the bodies. The matrix $\boldsymbol{M}$ is a 6x6 diagonal matrix containing the masses and the inertia of the bodies. The matrix $\boldsymbol{A}$ is a 6x4 matrix which includes the contribution from the reaction forces. It is not a coincidence that the contribution from the reaction forces is exactly the transposed matrix/vector equation corresponding to the constraints. The vector $\boldsymbol{x}$ corresponds to $\boldsymbol{x} = [x_1, y_1, \theta_1, x_2, y_2, \theta_2]$. The vector $\boldsymbol{F_r}$ is represented by the reaction forces, $\boldsymbol{F_{ext}}$ by the external forces that are applied on the body (i.e. gravity) and the vector $\boldsymbol{c}$ is represented by the terms in the constraint equations which are not dependent on the double derivative of the states.

By using this approach, it can be observed that careful bookkeeping is needed in order to correctly define the constraint equations and reaction forces. Moreover, it can be observed that if the number of bodies or the number of constraints increases, the size of the matrices that need to be evaluated increases considerably which can lead to computational problems.

Schwab (1998) proposes an alternative formulation of the previously-mentioned equations of motion by using Lagrange multipliers. The method using Lagrange multipliers has the same dimensionality as the previous one, but it does not require an explicit formulation of the reaction forces, being directly accounted for by the multipliers applied to the constraint equations. In this way, the system does not solve for the reaction forces, but for the Lagrange multipliers.

The observation that the Euclidean description of the mass centers includes dependent and redundant coordinates (i.e. x depends on y through the length of the segments), suggested the introduction of an alternative parametrization method using generalized coordinates (Figure 5.1 - Right). The generalized coordinates can be considered a set of minimum coordinates which can describe the system, that in the same time minimize the required constraint equations. It can be observed that by defining the two quantities q_1 and q_2, the rotational movement is fully described without the need of enforcing any additional constraints. Moreover, the position of each body's COM can be uniquely described by the two generalized coordinates. The use of generalized coordinates would help in transitioning from Newton's equations to the definition of Lagrange's equations, an alternative possibility of solving the EOM.

Most of the multibody dynamical systems can include a series of constraints. According to Moore (2022) and Ross (2020), the two main types of constraints are holonomic and nonholonomic constraints. The holonomic constraints imply constraints on the configuration of the system, or in other words, on the position of the bodies (the above-defined example is a holonomic system since it is fixed in the inertial space and the two bodies are fixed with respect to each other). On the other hand, a nonholonomic system is characterized by constraints applied to the velocities of some points in the system (think about a shopping cart). in general, for a holonomic system, it is always possible (but not always trivial) to find a minimum set of independent generalized coordinates. For a nonholonomic system, the generalized speeds (for example the first derivative of the generalized coordinates) are not independent anymore. This aspect will introduce certain consequences which will be presented in the later stages of the section.

The derivation of Lagrange's equations is based on D'Alembert's principle, where the equations of motion are expressed along the directions of the available displacements (allowed by the constraint forces). It needs to be mentioned that the location of each of the bodies' COM can now be expressed in terms of the generalized coordinates. In this way, the admissible virtual displacements of each of the bodies can be defined as:

$$\delta \mathbf{r_j} = \sum_{i=1}^{n} \frac{\delta \mathbf{r_j}}{\delta q_i} \delta q_i, \qquad j = 1, 2, ..., N \tag{5.2}$$

Equation (C.2) presents the relationship between the virtual displacements δr of each of the N bodies with respect to the n (number of generalized coordinates) admissible virtual displacements of the generalized coordinates δq. Furthermore, Ross (2020) defines a new set of parameters as:

$$\delta \boldsymbol{\gamma_{ji}} = \frac{\delta \mathbf{r_j}}{\delta q_i} = \frac{\delta \mathbf{v_j}}{\delta \dot{q}_i}, \qquad j = 1, 2, ..., N; \;\; i = 1, 2, ..., n; \tag{5.3}$$

$$\delta \boldsymbol{\beta_{ji}} = \frac{\delta \boldsymbol{\phi_j}}{\delta q_i} = \frac{\delta \boldsymbol{\omega_j}}{\delta \dot{q}_i}, \qquad j = 1, 2, ..., N; \;\; i = 1, 2, ..., n; \tag{5.4}$$

It can be observed that for example, Equation (C.2) can be written in terms of Equation (C.3) where $\delta \boldsymbol{\gamma_{ji}}$ and $\delta \boldsymbol{\beta_{ji}}$ represent *"projection vectors"* Ross (2020). With this definition in mind, the generalized forces and moments are elaborated as:

$$Q_i = \sum_{j=1}^{N} (\mathbf{F_j} \cdot \boldsymbol{\gamma_{ji}} + \mathbf{M_j} \cdot \boldsymbol{\beta_{ji}}) \tag{5.5}$$

In the case of a holonomic system, it is possible to define the number of generalized coordinates to be equal to the number of degrees of freedom. In this case, the admissible virtual displacements of the generalized coordinates are independent of each other. This aspect leads to the elaboration of Lagrange's equations of a nonholonomic system:

$$\frac{d}{dt}\left(\frac{\delta L}{\delta \dot{q}_i}\right) - \frac{\delta L}{\delta q_i} = Q_i^{nc}, \;\; i = 1, 2, ..., n; \tag{5.6}$$

Equation (6.8) presents the Lagrange equations for a holonomic system in the case that n (the number of generalized coordinates) is equal to the number of DOF of the system. The term L represents the Lagrangian, which represents the difference between the total kinetic and potential energy of the system. The term Q_i^{nc} is represented by the generalized non-conservative forces and moments. One example of nonconservative force might be the force/moment induced by a damper, or an input force/moment.

A nonholonomic system implies that some of the generalized velocities become dependent on each other. This means also that the number of DOF of the system becomes less than the number of generalized coordinates. In this way, the admissible virtual displacements of the generalized coordinates are not independent on each other anymore. The same case if for a holonomic system, a minimum set of coordinates is not defined. In this case, all the constraints can be described by the following general form (Ross, 2020):

$$\sum_{i=1}^{N} a_{li}(q)\dot{q}_i + a_{lt}(q) = 0, \;\; l = 1, 2, ..., S; \tag{5.7}$$

The general form of constraints (Equation (5.7)) presents a set of S constraints which always has a term linearly related to the generalized velocity, $a_{li}(q)$ and a term which is not a coefficient of generalized velocities $a_{lt}(q)$. This results in the Lagrange equations form for a nonholonomic system:

$$\frac{d}{dt}\left(\frac{\delta L}{\delta \dot{q}_i}\right) - \frac{\delta L}{\delta q_i} = Q_i^{nc} + \sum_{l=1}^{S} \lambda_l a_{li}, \;\; i = 1, 2, ..., n; \tag{5.8}$$

One possible advantage of using the Lagrange formulation is the minimized number of coordinates that need to be solved for. If the example from Figure 5.1 is observed, only two unknowns are present in the form of $\ddot{q}_1$ and $\ddot{q}_2$. Moreover, the resulting system of equations obtained from the Lagrange equations is linear in the second derivative of the generalized coordinates vector, facilitating a rather straightforward computation. On the other hand, as it can be observed in Equation (5.8), additional unknowns in the form of Lagrange multipliers are introduced in the case of a nonholonomic system. This implies that the system needs to encompass new unknowns that need to be solved for.

An alternative method that can be proposed such that the use of Lagrange multipliers is not needed is the approach proposed by Kane and Levinson (1985). Both resources, Ross (2020) and Moore (2022) have different views on the nomenclature of the elements used for solving Kane's equations. Kane's equations are also an extension of D'Alembert's principle where new quantities such as *"generalized velocities"* and *"partial velocities"* are defined. For the rest of the report, the used nomenclature will be the one based on Moore (2022).

If, in the previous paragraphs, the generalized velocity was defined as the first derivative of the corresponding generalized coordinate, the more rigorous definition of the generalized velocity is given as (Moore, 2022):

$$\boldsymbol{u} = \boldsymbol{Y_k}(\boldsymbol{q}, t)\dot{\boldsymbol{q}} + \boldsymbol{z_k}(\boldsymbol{q}, t) \tag{5.9}$$

Based on Equation (5.9), the vector of generalized velocities **u** is dependent on a factor linearly related to the vector of the first derivative of the generalized coordinates $\dot{\boldsymbol{q}}$ and a term which is not related to the first derivative of the generalized coordinates. It needs to be observed that the number of generalized velocities is equal to the number of generalized coordinates and a valid option for the matrix Y_k is the identity matrix. In this case, the terms defined in Equations Equation (C.5) and Equation (6.8) are defined as *"partial velocities"* by Moore (2022). With these definitions, the number of independent generalized velocities for a nonholonomic system can be chosen to be n (the number of generalized coordinates) minus S (the number of nonholonomic constraints). In this way, the system of equations can be solved only in terms of the independent generalized speeds. The same approach is proposed also with a system of holonomic constraints (now with independent and dependent generalized coordinates) in the case where the number of selected generalized coordinates is not minimum (is not equal to the number of DOF). By following this approach to solving a constrained system, there is no need for additional unknowns attributed to the introduction of the constraints (previously defined as Lagrange multipliers). Due to space requirements, the final equations of motion will not be described in this report. Nevertheless, the final form of the equations of motion also solves for the unknown generalized velocities and the first derivative of the generalized velocities with the goal of obtaining the evolution of the states (generalized coordinates and generalized velocities) in time such as (Moore, 2022):

$$\begin{aligned} \boldsymbol{f_d}(\dot{\boldsymbol{u}}, \boldsymbol{u}, \boldsymbol{q}, t) &= 0 \\ \boldsymbol{f_k}(\dot{\boldsymbol{q}}, \boldsymbol{u}, \boldsymbol{q}, t) &= 0 \end{aligned} \tag{5.10}$$

Equation (5.10) presents the general form of equations that are attained by solving Kane's equations. The function $\boldsymbol{f_d}$ can be obtained by solving the dynamical problem (which includes forces and moments), and the function $\boldsymbol{f_k}$ can be obtained from the definition of the kinematical relationship between the first derivative of the generalized coordinates and the generalized speeds (Equation (5.9))

This section tried to give an overview regarding the possible methods of solving a multibody system. Starting from the Newton's definitions and reaching to Kane's equations, it can be observed that the number of unknowns that the system needs to be solved for decreases, and the need of additional terms which are due to the introduction of constraints is obsolete.

5.1.2. Problem set-up

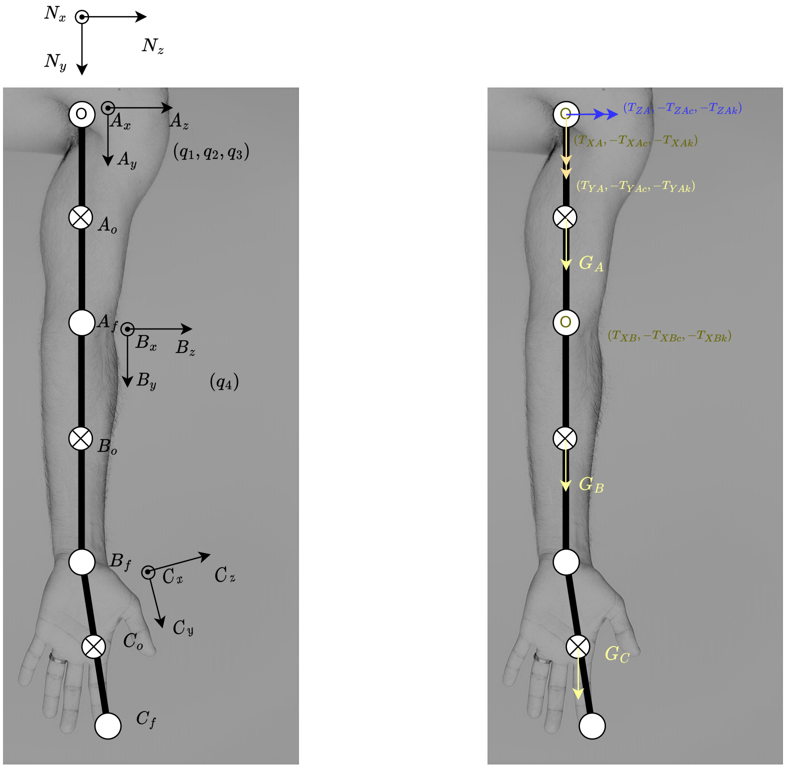


Figure 5.2: Upper arm coordinate systems (Left), an overview of the forces and moments (Right).

The proposed model used for the biomechanical analysis of the upper arm can be observed in Figure 5.2. It needs to be recognized that a model is a simplified representation of the real system which can replicate a bounded set of input-output relationships. As Van Paassen (1994) mentions, in order to obtain a simplified representation of the reality, the focus needs to be directed only towards a limited set of aspects that need to be modeled. In this case, the intention is to replicate the biodynamic behavior of the upper limb during the operation of a touchscreen. The constructed model will have two main uses: firstly, the goal would be to gain an understanding regarding the possible patterns of the response of each of the parts of the rigid body when subjected to disturbances; secondly, the goal would be to be able to observe the differences in the response in different locations on the simulated model, based on changes geometry, mass characteristics, and other parameters.

The implementations that will be presented in the following sections are based on Kane's equations and verified with the Lagrange approach. Moreover, the methodology used for implementing the multibody problem is inspired by the guidelines given by Moore (2022). The set-up of the problem is defined in $Python$ with the help of the symbolic computation tool $SimPy$, along with the mechanics toolbox $SimPy-Physics-Mechanics$.

5.1.3. Assumptions

The following assumptions can be introduced: Each segment has a fixed mass located as a point mass at its center of mass, the location of the center of mass remains constant during the movement, joints are considered hinge or ball and socket joints, the lengths and moments of inertia of the segments are considered constant during the movements (Mugge, 2022), the moments of inertia with respect to the local y axes are considered negligible.

Moreover, based on Figure 5.2, it is assumed that point O is coincident with the shoulder joint and can be prescribed a vertical movement pattern with respect to the inertial frame N. In other words, the point O does not have a translational degree of freedom, but, it can have an imposed (by the user) position, velocity, and acceleration signal with respect to the inertial reference frame.

Furthermore, it is assumed that the shoulder joint has 3 DOFs being able to rotate around each axis of its reference frame. The elbow joint allows for a rotation around the local x-axis. Lastly, the hand is assumed to have a fixed orientation with respect to the lower arm. These conventions were based on careful monitoring of the ergonomics of the upper limb when operating a touchscreen. It was observed that during the operation, the movement is induced mainly from the shoulder joint while the elbow just flexes/extends to accommodate for changes in distance with respect to the screen and the orientation of the hand is constant with respect to the lower arm.

5.1.4. Coordinate systems

The multibody kinematic chain consists of four main coordinate systems. Starting from the top, the first coordinate system, N is the inertial coordinate system. Every velocity/acceleration is computed with respect to this inertial coordinate system. The second coordinate system is the local coordinate system of the upper arm, A. This coordinate system is oriented with respect to the inertial reference system with a series of three consecutive rotations following the $'x, y, z'$ convention with the angles q_1, q_2, q_3 expressed in radians. Similarly, the local coordinate system of the lower arm, B follows a rotation with respect to the A coordinate system of $'q4'$ around the local $'x'$ axis. It needs to be mentioned that q_1, q_2, q_3, q_4 are the defined generalized coordinates for this system, resulting in a total of 4 DOF. The last reference frame is represented by the rotation of the local C axis with respect to the B axis. In this case, the rotation also follows the $'x, y, z'$ convention but the angles of rotation are constant (unlike the generalized coordinates that change during the simulation). In other words, the hand coordinate system keeps a fixed orientation with respect to the local B frame during a simulation run. Moreover, it needs to be mentioned that the coordinate frames in Figure 5.2 are in fact coincident with the root points of each segment (they were presented in the figure with a small offset for clarity reasons), such that they represent a set of principal axes (the off-diagonal inertia parameters is zero).

5.1.5. Forces

The modeled forces that are applied to the multibody system are introduced in Figure 5.2 (Right). As it can be observed, each of the bodies has a specific mass, inducing a force of gravity at each mass center.

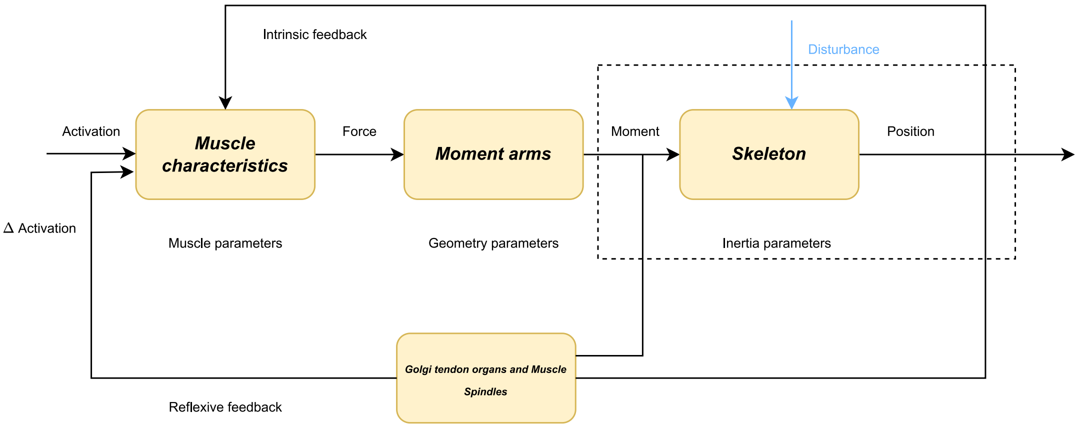


Figure 5.3: Block diagram describing neuromuscular system, adapted from Mugge (2022) and Lasschuit et al. (2008).

The additional moments that are applied on the kinematical chain model are based on Figure 5.3. The block diagram describes how the activation chain results in an applied force and, in turn, in a moment applied on the multibody system. Since the focus of this research is to describe the biodynamic characteristics of the upper limb, the focus of the model will be only represented by the relationship between the applied moments and the resulting position of the skeleton. It has been decided that an accurate representation of the muscles along with their very accurate moment arms is difficult to obtain and lies outside the scope of the analysis. Moreover, it can be observed that two main feedback loops are present in the loop, the first one is the intrinsic feedback which can be represented by the constant passive stiffness and damping of the muscle characteristics. On the other hand, the additional loop is the reflexive feedback which is usually present if disturbances are present, and can alter the activation levels of the muscles and, in turn, the lumped damping stiffness characteristics.

Following the above-mentioned considerations, the current model contains three applied torques per each degree-of freedom (each joint direction that is not constrained). An active component is represented by the lumped contributions from all the muscles acting on a specific joint, resulting in the total muscle torque which has the goal to actuate the skeleton (T_{XA}, T_{YA}, T_{ZA}, T_{XB}). The rest of the torques represent the passive elements of the muscle components (kept constant) representing the lumped stiffness (T_{XAk}, T_{YAk}, T_{ZAk}, T_{XBk}) and damping characteristics of the muscles (T_{XAc}, T_{YAc}, T_{ZAc}, T_{XBc}). It needs to be mentioned that the reflexive contribution of the muscles or the time-varying characteristics of the muscles have not been modeled/analyzed in this stage of the work.

5.1.6. Parameters

Each segment of the elaborated multibody model of the upper arm has the following parameters: mass, length, inertia, equivalent spring, and damping coefficients. Usually, the geometrical parameters of the limbs such as masses, inertias, and locations of centers of mass are difficult to obtain in real scenarios. Some approaches make use of complex scanning techniques providing an overview of the limb density across different locations. Other methods have used interpolated/extrapolated data from cadavers in order to obtain the masses and center of masses for the limbs. One possible solution to the problem of estimating these parameters is to use regression equations based on quantities that are easily measurable such as height, and weight (Winter, 2009).

Table 5.1: Overview of the average anthropometric data based on Winter (2009).

	Segment Mass/ Total Body Mass [-]	Center of Mass/ Segment Length [-]	Radius of Gyration/ Segment Length [-]	Segment Length/ Height [m/kg]
Hand	0.006	0.506	0.297	0.108
Forearm	0.016	0.430	0.303	0.146
Upper arm	0.028	0.436	0.322	0.186

Table 5.1 presents a set of anthropometric characteristics, which have the goal of describing the average individual possessing a combination of height and mass. The Preliminary Results section will present how an alteration of several parameters of the model, thus replicating how either between-subject variability or within-subject variability, would affect the overall response of the upper-limb system when faced with disturbances.

5.2. Multibody simulations

This section presents a series of investigations centered around the variation of several parameters of the multibody model with the goal of obtaining an intuition regarding possible within or between-subject variability characteristics that can be encountered in the real-life scenario of acceleration feedthrough through the limbs. An overview of the selected cases is presented in Table 5.2.

Table 5.2: Overview of the analyzed cases of the multibody simulations

Case 1	Effect of limb length and mass.
Case 2	Effect of lumped stiffness and damping parameters.
Case 3	Effect of limb configuration.
Case 4	Investigation of the acceleration feedthrough along the limb.

Figure 5.4 presents a possible arm configuration that can be used for the analysis. For all the cases, an inverse dynamics problem is solved such that the active torques that are necessary to keep an initially defined configuration are applied to the bodies. Moreover, in this stage of the research, the interaction with the touchscreen has not been modeled. Nevertheless, some conclusions can be taken regarding possible differences between and within a subject when analyzing an arm under the subject of gravity. The set-up of the simulation would like to mimic the effect of disturbances on a *”trimmed”* condition of the upper limb (the limb configuration does not change within a simulation). Additionally, in a later section, a series of verification steps which make use of the computed torques under the *”trimmed”* condition will be presented. The role of the verification steps is to ensure the corectness of the chosen method that was used for solving the multibody equations of motion.

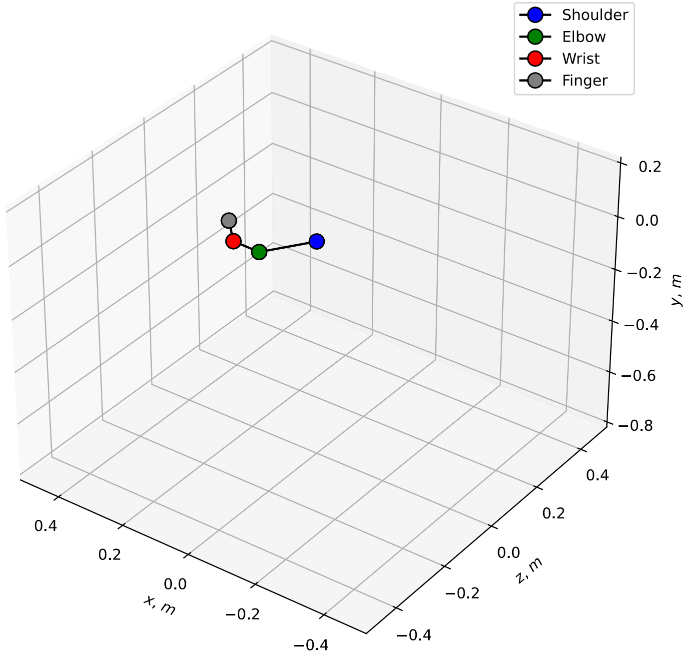


Figure 5.4: Overview of a possible arm configuration that is used in the simulations.

5.2.1. Effect of limb length

When analyzing the effect of the limb length, two sets of *"artificial"* limbs were created, as presented in Table 5.3. As mentioned in the previous chapters, a parametrization for the limb lengths, masses, mass centers, and moments of inertia was made based on two inputs, namely the height and mass of the *"simulated"* subject. The two sets of parameters were obtained based on two *"simulated"* subjects with the following characteristics: Subject 1 (Heavier/Longer), Height - $190\,[cm]$, Weight - $80\,[kg]$; Subject 2 (Lighter/Shorter), Height - $175\,[cm]$, weight - $65\,[kg]$;

Table 5.3: Overview of the independent variables for the Case 1

	$m_1\,[kg]$	$m_2\,[kg]$	$m_3\,[kg]$	$l_1\,[m]$	$l_2\,[m]$	$l_3\,[m]$
Heavier/Longer	2.24	1.28	0.48	0.35	0.28	0.21
Lighter/Shorter	1.82	1.04	0.39	0.33	0.26	0.19

Constant active torques were applied to each segment (solved using an inverse dynamics problem) such that a desired configuration of q_1, q_2, q_3, q_4 was kept in place, mimicking an upper limb that was hanging in the air at a particular position. The equivalent spring and damping coefficients of the available degrees-of-freedom were assumed to have constant values. A disturbance signal in the form of linear inertial acceleration (similar to the one used by Mobertz et al. (2018) and Khoshnewiszadeh and Pool (2021)) was applied at the shoulder level (point O) in the vertical (N_y) direction. The dependent variables are selected to represent the inertial acceleration at one location of the upper limb, namely points A_f which is expressed along the three directions of the inertial frame.

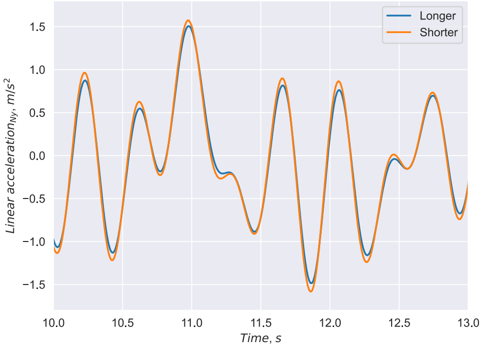


Figure 5.5: Overview of the time traces of the linear acceleration of the point A_f for the shorter and longer limbs expressed in the N_y direction.

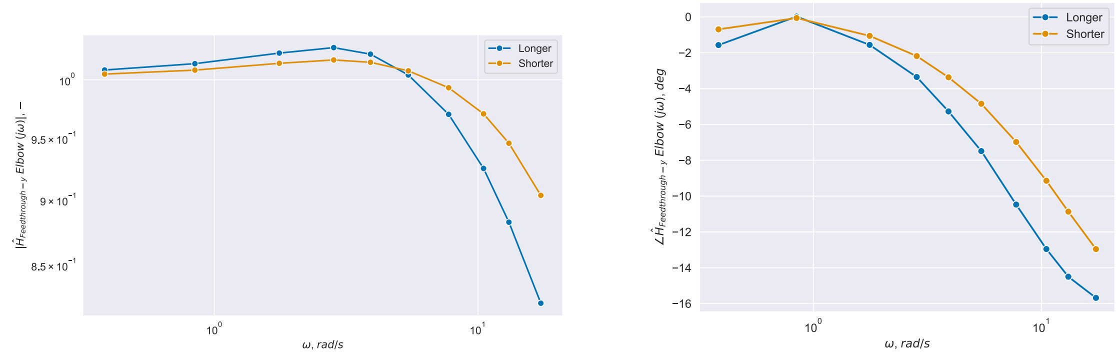


Figure 5.6: Non-parametric magnitude and phase representation for the estimated transfer function relating the acceleration of the point O to the acceleration of the point A_f in the vertical, N_y, direction.

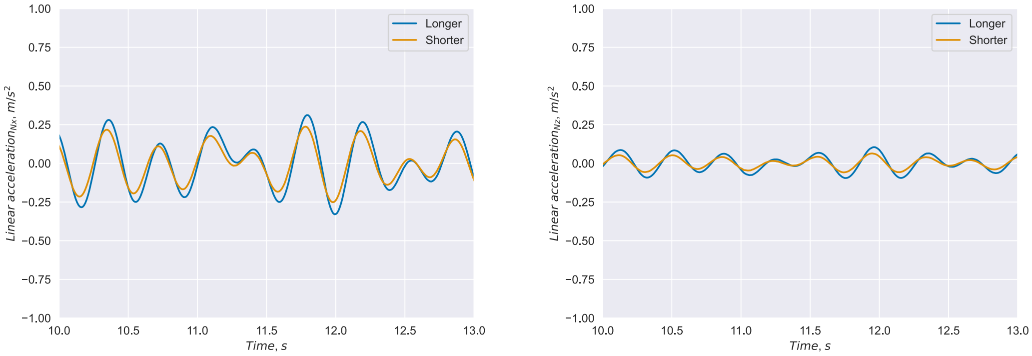


Figure 5.7: Overview of linear accelerations of point A_f for the shorter and longer limbs expressed in the N_x and N_z directions.

Takeaways - Limb length variation

Figure 5.10 presents a comparison between the inertial accelerations obtained at the point representing the elbow, A_f expressed in the vertical N_y direction. Moreover, an open-loop frequency response identification technique has been implemented in Figure 5.6 in order to gain more insight into the magnitude and phase of the transfer of accelerations between the shoulder point (O), and the point of interest (A_f).

It can be observed that due to the fact that the differences in the anthropometric characteristics are not very high, the responses do not exhibit significant changes. Nevertheless, a closer look at the time-domain signals reveals that the longer limbs exhibit a response with a slight attenuation of the high frequencies compared to the shorter limbs. Moreover, there is a slight lag present in the evolution in time of the long limbs. The transfer-function representation can suggest similar conclusions as observed in the time domain. The longer limbs attenuate more the frequencies around $10\,[Hz]$ compared to the shorter limbs. Moreover, the resonant peak is more pronounced and shifted to slightly lower frequencies for the long limbs. It needs to be mentioned that the simulations were performed by replicating a *"stiff"* setting for the neuromuscular system. This fact has the implication that the magnitude of the transfer function between the imposed acceleration and the computed acceleration at the target point does not deviate much from the value of one (representing the perfect transmission of accelerations), and the differences in phase are not significant. Moreover, in a real-life scenario, it is expected that the acceleration measurements (i.e. from an accelerometer) are not perfect and are contaminated with noise. This fact can make the already *"small"* differences seen in the *"analytical"* responses even harder to distinguish. Nevertheless, this analysis can provide a qualitative conclusion and hypothesis that the shorter the mass/weight of the limb, slightly higher frequencies will pass through it (in the vertical direction) providing that the stiffness characteristics are the same. Due to space constraints, the analysis focussed on the comparison of the accelerations obtained at the elbow location. A later section will present a comparison between the acceleration feedthrough at different locations of the arm within one "*subject*".

Since a 3D upper limb model has been proposed, a vertical acceleration would also impose accelerations in the other directions of the inertial reference frame. Figure 5.7 presents the residual accelerations of the elbow point, obtained from imposing a vertical acceleration at the shoulder. It can be observed that quantitatively, the accelerations obtained in the directions orthogonal to the direction of the imposed signal are smaller. A separate analysis (not presented) concluded that the transfer functions do not resemble the same "*low-pass*" behavior as in the previous case. Nevertheless, it is important to mention that these orthogonal accelerations can affect the interaction between the limb and the screen since the surface of the screen constrains the motion normal to the screen plane.

5.2.2. Effect of passive stiffness - Position analysis

The goal of analyzing the effects of changing stiffness is to observe how possible changes in neuromuscular settings within a participant would affect the behavior of the upper limb. In a real-life scenario, a motion-capturing system will be used to identify the time traces of relevant features on the upper limb. Saying this, before designing the real experiment, is interesting to investigate whether some information regarding the neuromuscular stiffness can be extracted from the resulting time traces of inertial positions of the simulated limb.

Table 5.4: Overview of the independent variables for the Case 2.

	$(k_{q1}\,[Nm/rad],\, c_{q1}\,[Nm/rad^2])$	$(k_{q2}\,[Nm/rad],\, c_{q2}\,[Nm/rad^2])$	$(k_{q3}\,[Nm/rad],\, c_{q3}\,[Nm/rad^2])$	$(k_{q4}\,[Nm/rad],\, c_{q4}\,[Nm/rad^2])$
Stiffer	(10, 10)	(10, 10)	(10, 10)	(10, 10)
Less stiff	(2, 3)	(5, 5)	(2, 3)	(1, 2)

Table 5.4 presents two sets of passive stiffness characteristics that are used for the analysis. The two sets of values would like to simulate two neuromuscular settings associated with a position task (stiff setting) and a force task (less stiff setting). Similar to the previous case, a disturbance in the form of vertical linear acceleration will be imposed at the shoulder level. On

the other hand, the lengths and masses of the segments are kept constant in order to observe the effects of changing just the *"real"* stiffness characteristics. The dependent variables of this scenario are the positions of the points A_f, B_f, and C_f expressed in the inertial frame and computed with respect to the shoulder position (since the shoulder position changes in the inertial frame due to the imposed acceleration). Since an acceleration disturbance is imposed in the N_y direction, the main effects are expected to be observed in the same direction.

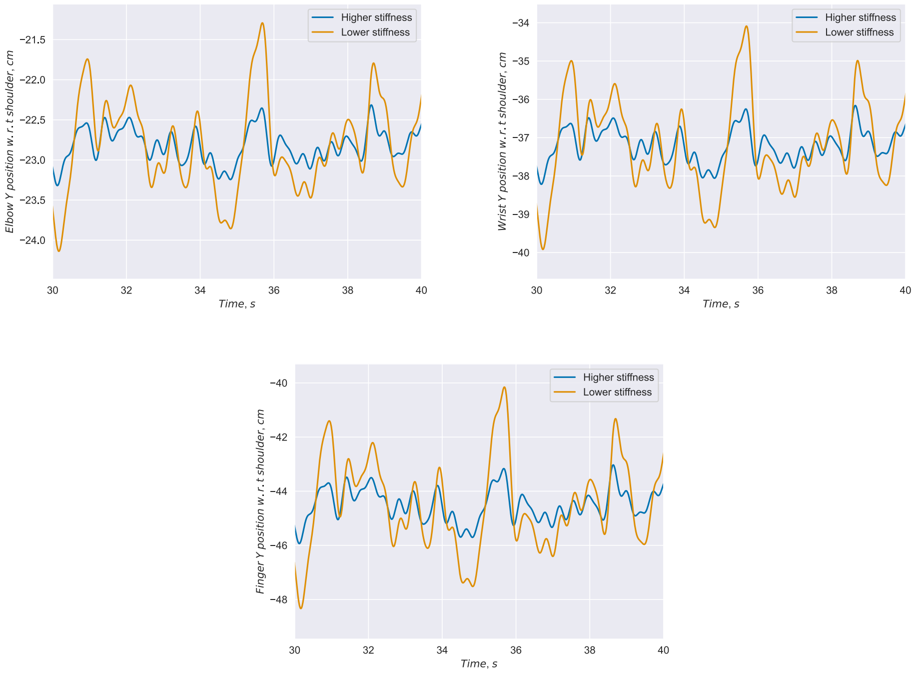


Figure 5.8: Overview of the positions of the points A_f, B_f, C_f expressed with respect to the point O along the y direction of the inertial frame.

Takeaways - Passive stiffness variation - Position analysis

Several conclusions regarding this scenario can be taken based on Figure 5.8. The reason for investigating the relative position of the landmarks with respect to the *"shoulder"* (point O), is based on the fact that in a real scenario, the camera will be fixed to the simulator, and therefore it will move with it, with respect to an inertial frame. Saying this, the camera is expected to observe only the relative movement between the limbs and not their absolute displacements (as if the camera would be fixed in an inertial space). As expected, the position of the landmarks in the y direction varies less with the stiffer system. Moreover, it can be observed that the difference in position between the stiffer system and the less stiff system increases with the location in the set of rigid bodies. In other words, the difference in the finger location is the biggest, whereas the difference in the elbow position is the lowest. This behavior can be explained by the set-up of the simulation where the arm is *"hanging"* against gravity without any end pivot point. It can be observed that the overall behavior across the analyzed points is the same due to the fact the simulations enforce linear visco-elastic properties at the joint locations. On the other hand, the overall differences in the behavior of the two systems encourage the author to ask the question whether *"some"* differences can be observed in a real-life scenario between the time-traces of several landmarks of the upper hand for two distinct neuromuscular settings (as described in the literature study part). Moreover, another interesting aspect to consider is to investigate also whether in a real-life scenario it is possible to classify/predict the *"correct"* neuromuscular setting by analyzing the outcomes of the motion capturing approach (the evolution in time of the recorded landmarks).

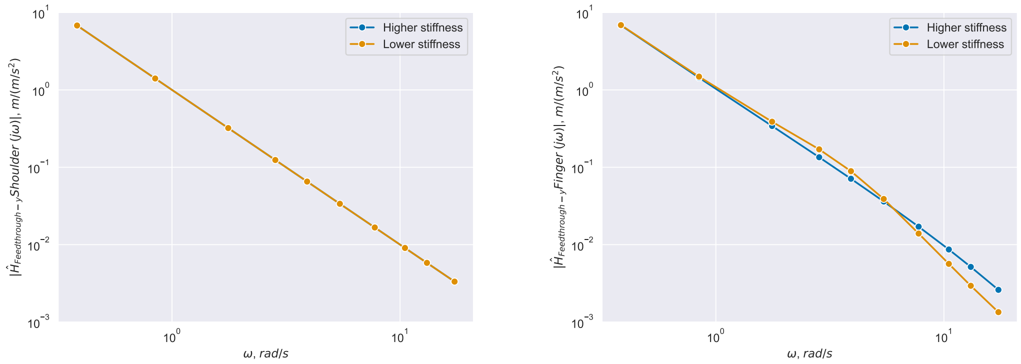


Figure 5.9: Non-parametric transfer function between the acceleration imposed at the shoulder and the inertial shoulder (Left) and finger (Right) positions.

A frequency-domain analysis was investigated in order to check whether any quantitative information can be obtained regarding the observed position time traces. Saying this, in a real-life scenario, a correlation between the movement of the arm (or of the recorded features on the upper limb) and the disturbance signal is expected. In this way, the simulated scenario estimated the transfer function which relates the acceleration imposed at the shoulder level and the positions of the shoulder and finger. The reader should remember that the transfer functions were estimated for the inertial position of the landmarks (and not for the relative positions). When focussing on Figure 5.9 (Left) it can be observed that the behavior for both systems is similar, replicating a double integrator. These outcomes shouldn't surprise anybody, since all the cases are based on an imposed vertical acceleration at the shoulder level. By integrating twice the inputted acceleration signal, the position signal is obtained. Figure 5.9 (Right) presents a slightly different behavior compared to the leftmost figure. The lower stiffness system shows an amplification area and an attenuation section (with respect to the higher stiffness case). The reason for this behavior can be related to the outcomes of the next section where it will be shown that the acceleration feedthrough in the lower stiffness setting possesses a resonant peak and a sharp attenuation of higher frequencies when compared to the higher stiffness setting. The previously-mentioned aspect in combination with the double-integrator behavior which transforms accelerations to positions, imposes the overall trend observed in the rightmost figure. It is important to mention that both differences in the time-domain and frequency-domain signals are dependent on the parameters of the system (lengths, masses, inertia). This example only portrayed a generic set of parameters, corresponding to the *Heavier/Longer* case. Moreover, this simulation was not considering the interaction with the touchscreen which can represent a pivot point in the system.

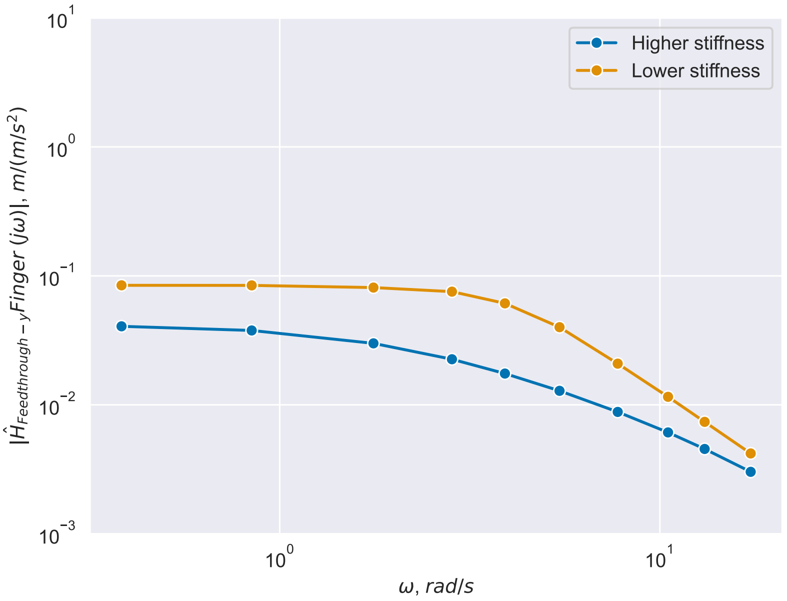


Figure 5.10: Non-parametric transfer function between the acceleration imposed at the shoulder and the position of the point C_f with respect to point O.

Figure 5.10 presents another perspective on the analysis of the landmark positions. This representation can be directly compared with the time-domain signals presented in Figure 5.8, representing the estimated transfer function which links the shoulder acceleration to the relative position of the point C_f with respect to the point O. It can be observed that the behavior of the estimated transfer function is different when compared to the estimated transfer function linked to the inertial positions (Figure 5.9). On one hand, it can be observed that the transfer function corresponding to the higher stiffness case attenuates more all the frequencies which were analyzed. This aspect can be confirmed by visually inspecting Figure 5.8 and observing that the blue signal is varying *”less”* when compared to the yellow one. An interesting behavior observed in Figure 5.8 (between 35 and 36 $[sec]$) and not present in the transfer-function representation is the *impression* that the lower stiffness signal attenuates more the higher frequency contents when compared to the higher stiffness signal. The author would have expected a trend in the frequency domain, similar to the one observed in Figure 5.9 (Right), where at some point the lower stiffness case would attenuate more the higher frequencies when compared to the higher stiffness scenario. The author hypothesizes that the expected trends are not observed due to the method that is used for computing the relative positioning between the landmarks which might cancel some of the parts of the frequency content (by subtracting the inertial position signal of the shoulder point, O from the inertial position signals of the landmarks). Nevertheless, this aspect needs to be further researched by the author in the next phase of the thesis.

5.2.3. Effect of passive stiffness - Acceleration analysis

Another interesting aspect that can be investigated when varying the passive stiffness of the system, is the feedthrough of linear accelerations through the limb. It needs to be mentioned that the same conditions were simulated as in the previous section, where only the equivalent spring and damping parameters (for all the degrees of freedom) are altered in order to simulate a stiffer or less stiff neuromuscular system.

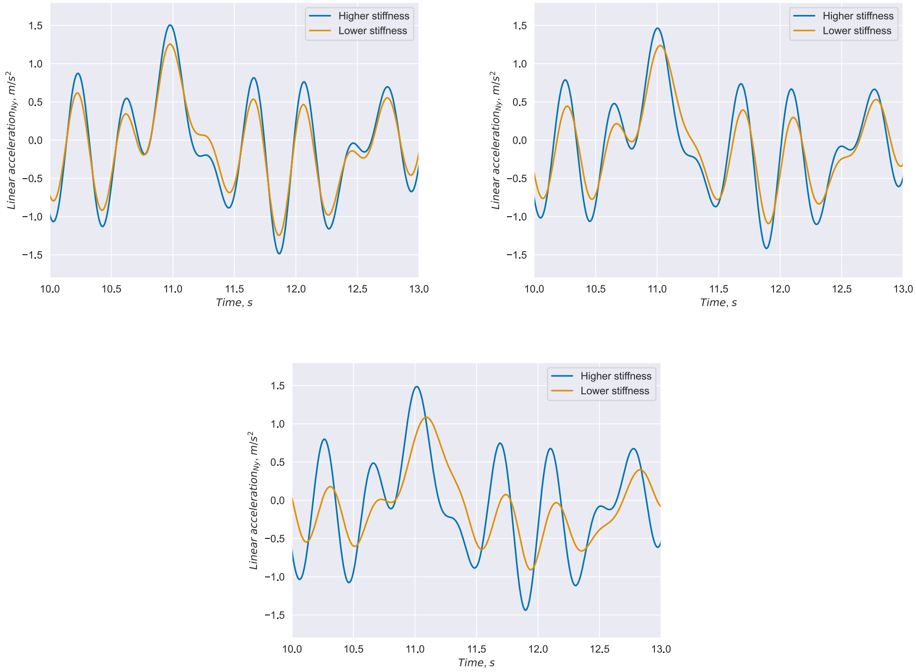


Figure 5.11: Overview of the inertial accelerations of the points A_f, B_f, C_o expressed along the y direction of the inertial frame.

Takeaways - Passive stiffness variation - Acceleration analysis

Figure 5.11 presents a comparison between the inertial acceleration of three points of the limb (A_f, B_f and C_o) with two simulated neuromuscular system characteristics. It can be observed that the effect of changing stiffness is more pronounced than the effect of different limb masses/lengths (Figure 5.6) when comparing the point A_f. Moreover, the effect is more pronounced as the points further down in the kinematic chain are analyzed. As a general trend, the acceleration corresponding to the lower stiffness system appears to be lagging behind the acceleration of the higher stiffness system. Moreover, the lower stiffness system sees a larger attenuation in the higher frequency contents of the signal.

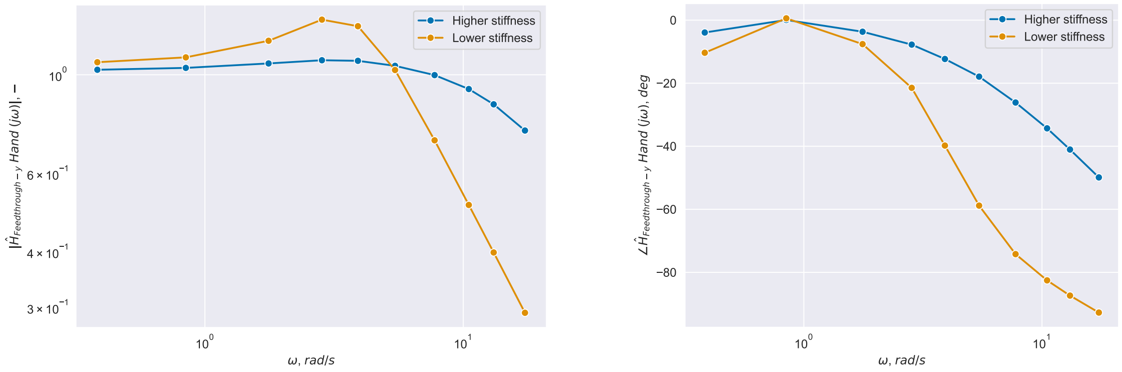


Figure 5.12: Non-parametric magnitude and phase representation for the estimated transfer function relating the acceleration of the point O to the acceleration of the point C_o in the vertical, N_y, direction.

For a better understanding of the time-domain signals, a non-parametric identification of the transfer function between the shoulder acceleration and the inertial linear acceleration of the point C_o expressed in the N_y direction has been performed in Figure 5.12. Unlike to the case where the effects of changing the limb lengths, a clearer difference in the phase can be observed between the stiffer and compliant systems. Additionally, it can be observed that the accelerations are feed-through more for the lower stiffness condition for frequencies close to $7[Hz]$ whereas for higher frequencies, the higher stiffness scenario exhibits a larger transmission of accelerations. A similar non-consistent behavior was also observed in Figure 3.11 where the BDFT for stick-based operations was identified in an experimental setting. The attentive reader could observe a difference in magnitude between the simulated BDFT transfer function (Figure 5.12) and the experimental BDFT transfer function (Figure 3.11). The differences can be explained by the fact that the BDFT transfer function identified in the experimental setting describes the transfer of the accelerations from the motion platform to the deflection $[rad]$ of a controlled device. In this way, the feedthrough of accelerations has a longer path, from the acceleration imposed to the moving base, to the feedthrough of the limbs, to the interaction between the limbs and the control device, and to the control device dynamics. On the other hand, the simulated scenario investigates only the relationship between the imposed acceleration of the shoulder, that feeds through to other locations on the limb. This non-consistent behavior between the BDFT and neuromuscular settings accentuates the complexity of the feedthrough phenomenon.

5.2.4. Effect of limb configuration

The goal of investigating the effects of different configurations of the limb is to observe how the computed linear accelerations are affected by different positions of the upper arm. From the presented literature, it is expected that an effect of limb position on the feedthrough of accelerations would be observed, given the fact that in a real-life scenario, different limb positions would impose different activations of the neuromuscular system. The two selected configurations can be seen in Figure 5.13.

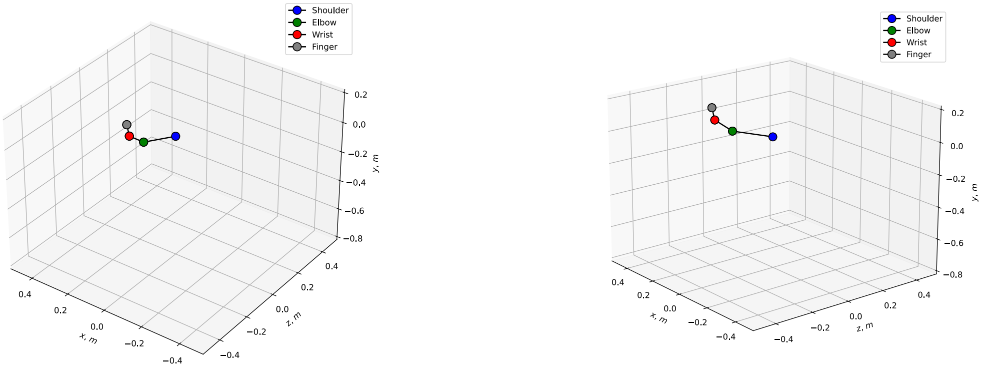


Figure 5.13: Overview of the first limb configuration (Left) and of the second limb configuration (Right).

As in the previous investigations, the above-shown configurations will be kept constant throughout one simulation, where a vertical acceleration will be imposed at the shoulder level, and with the constant stiffness characteristics corresponding to the *"stiffer"* scenario. The exact value of the configuration angles is given by Table 5.5.

Table 5.5: Overview of the independent variables for the Case 3

	$q_1[deg]$	$q_2[deg]$	$q_3[deg]$	$g_4[deg]$
Configuration 1	24	0	45	15
Configuration 2	24	0	75	15

The dependent variables in this scenario will be the inertial linear accelerations at the points A_f, B_f, and C_o expressed along the directions of the global coordinate system.

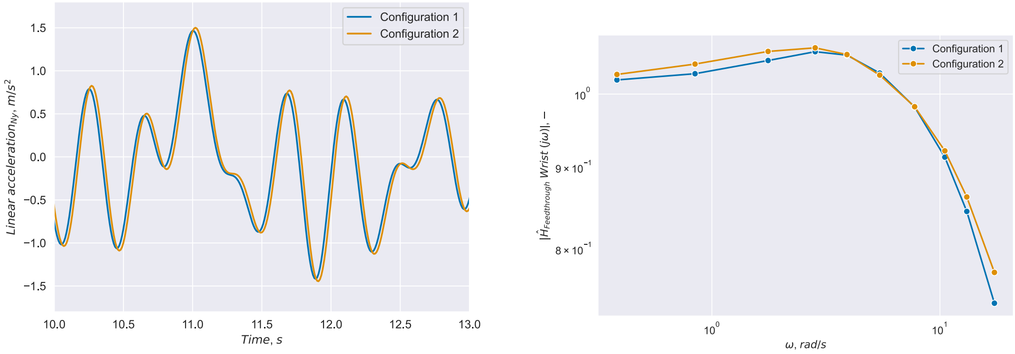


Figure 5.14: Overview of the time traces and non-parametric transfer function of the linear acceleration of point B_f for the two configurations expressed in the N_y direction.

Takeaways - Limb configuration

From Figure 5.14, it can be observed that the linear acceleration computed at the point B_f is slightly different when comparing the two configurations. The acceleration of the second configuration lags with respect to the first configuration. Moreover, based on the magnitude of the transfer function (the phase representation was left out due to space restrictions) representing the relationship between the vertical shoulder acceleration and elbow acceleration, it can be mentioned that the magnitudes are different in the areas before and after the corner frequency, with the first configuration having slightly attenuated components in those regions. Nevertheless, as in the previous section, the differences can be considered hardly significant if a real-life scenario would be analyzed. It is important to mention that the analysis was performed using similar passive stiffness characteristics. It can be expected that in reality, the second configuration would need a higher muscle activation to sustain the mass of the limbs, inducing a stiffer upper-limb system.

5.2.5. Effect of different locations on the limb

The last case of the analysis consists of investigating how the inertial linear acceleration imposed at the shoulder is fed through different parts of the limb. The simulated scenario corresponds to parameters from the Heavier/Longer, Less Stiff, and Configuration 1 scenario. Similarly, the vertical acceleration was imposed at the shoulder level. The dependent variables are represented by the computed linear accelerations at the points A_f, B_f, and C_o expressed in the vertical direction.

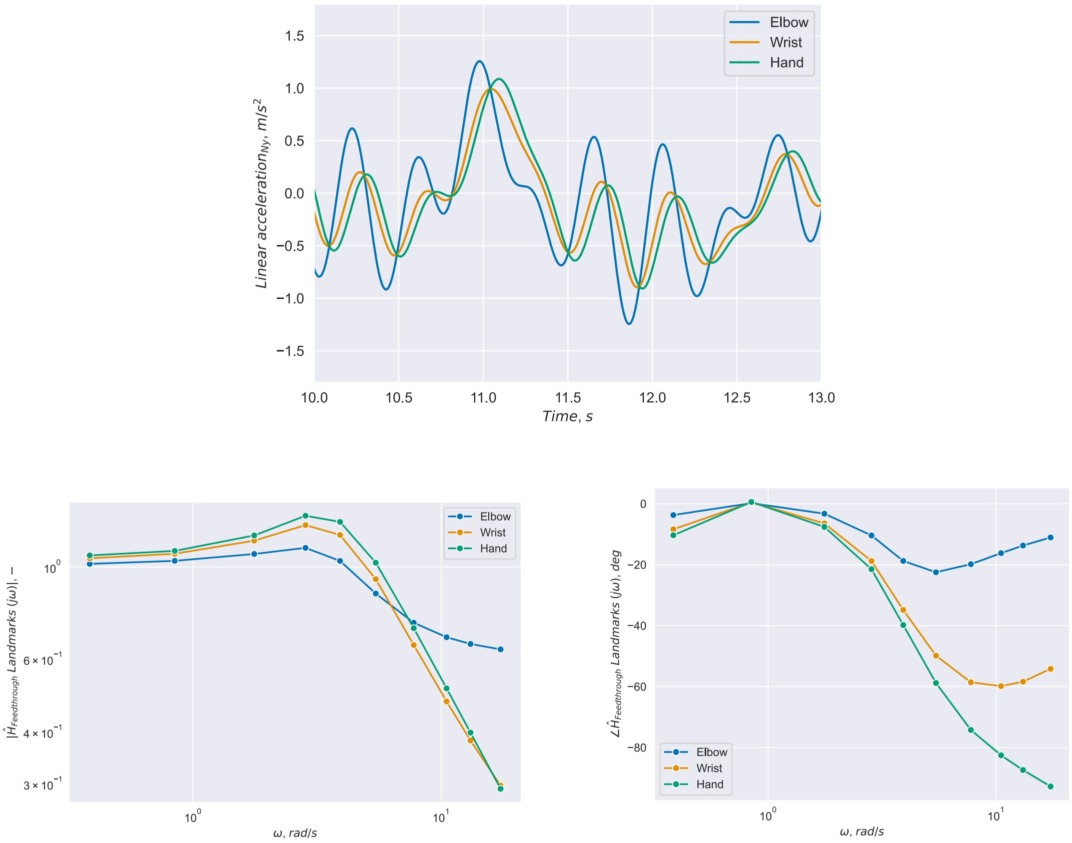


Figure 5.15: Overview of the time traces and non-parametric transfer function of the linear acceleration of points A_f, B_f, C_o expressed in the N_y direction.

Figure 5.15 presents the feedthrough of vertical inertial acceleration for the previously-mentioned points. It can be observed that the *"lag"* in the acceleration signal increases with the analyzed point which confirms the expectations based on physical phenomena. Moreover, the high frequencies are more filtered when looking at the wrist and hand. The reason for this phenomenon might be the fact that the wrist and hand are further in the kinematic chain compared to the elbow, and lumped inertia effects simulate a *"filter"* which is more effective at higher frequencies. The same conclusion can be observed from the non-parametric transfer function representation. As in previous cases, the frequencies below $7[Hz]$ are more fed through at the wrist and hand locations compared to the elbow, and higher frequencies see a higher attenuation compared to the same elbow landmark. Nevertheless, the overall behavior of the analysed landmarks is expected to change depending on alterations in geometry or passive stiffness/damping characteristics (the author presents here one used-case).

5.2.6. General conclusions of the analysis

The goal of this analysis was to discover and explain possible factors that have been reported in the literature which affect the BDFT behavior. It is widely known that the HO is one of the main sources in the BDFT variability (Venrooij et al., 2013). Saying this, the effects of between-subject variability (different body characteristics) and within-subject variability (different neuromuscular settings) need to be taken into consideration when analyzing the influence of BDFT on the human body. The simulations tend to agree with the point of view provided by Venrooij et al. (2013), which describes a modest influence of the body type on BDFT dynamics.

As it was observed in the simulated limbs, the effect of the acceleration transmission through limbs with different lengths and masses was limited. On the other hand, a larger effect was observed when changing the "*simulated*" neuromuscular settings. The same conclusions were also observed in the previously-mentioned study recognizing a non-negligible impact of the neuromuscular settings on the BDFT dynamics. Saying this, a possible research line would be to focus on the more evident within-subject variability, that is, to closely investigate methods that can detect the neuromuscular settings of a subject and which can relate these settings to the feedthrough of accelerations.

5.2.7. Shortcommings of the analysis

It needs to be mentioned that this analysis has several shortcomings regarding modeling.

Firstly, it has been mentioned that an acceleration signal was applied directly to the shoulder point. Based on the presented literature, it can be expected that the shoulder acceleration is different when compared to the acceleration of the moving base (through whole-body vibration properties). A simple approach that can be used to solve this *"inaccuracy"* is to link the shoulder with the moving-base through a lumped-parameter mass-spring-damper system with a resonant peak between 4 to $6\,[Hz]$ (the most common resonant peak location observed in the whole-body transmissibility). Nevertheless, the current analysis can explain the fact that several parameters such as linear/rotational inertia or limb lengths can affect the overall response of a multibody system replicating the upper limb.

Another shortcoming is that, unlike a real-life scenario, constant effective stiffness and damping parameters were used to simulate the neuromuscular system. Moreover, it is widely known that the neuromuscular system consists of a reflexive loop (active) and an intrinsic part. These simulations assumed only passive characteristics for the system.

Moreover, all the investigations were performed with the system fixed in a particular configuration (torques set such as the configuration was set constant without any disturbance). Similar to how a physical inceptor can affect the feedthrough characteristics of the upper limb, it can be expected that the contact with a hard surface such as a touchscreen, can affect the biodynamic properties of the upper limb. Efforts were made to simulate the interaction of the upper limb with the screen, but none of them succeeded at this stage. Firstly a collision model was proposed, but the next iterations included nonholonomic and finally holonomic constraints. The *"best"* results were obtained with the introduction of a holonomic constraint that fixed the endpoint of the kinematic chain to a plane (describing a screen). The difficulty of the approach comes from the fact that in the case of a holonomic system with more generalized coordinates than DOF, one of the generalized coordinates became dependent on the others (i.e. the elbow rotation). With one generalized coordinate becoming dependent on the others, also one of the torques became dependent on the rest of the torques, meaning that at each time step of the simulation a new dependent torque needed to be computed which increased the computational cost.

5.2.8. Verification procedures

Despite the fact that the above-mentioned results can be considered to be consistent with the physical interpretation of the phenomena, it is also important to ensure that the model outcomes and the code is correct. In order to do this, two different verification steps were proposed.

The goal of the first verification step is to identify whether the equations of motion and computed torques are correctly defined. A simplified scenario was selected, where no acceleration was imposed at point O, all the kinematic chain has a straight line configuration between O and A_f, and no stiffness and damping of the joints is defined. Additionally, a rotation of 90

$[deg]$ is defined with respect to the N_x direction. In this way, the kinematic chain is oriented horizontally, and the computation of the torques necessary to sustain gravity is trivial providing the case that the torques necessary for the *"trim"* condition need to be computed. The torques necessary to sustain gravity are computed via the inverse kinematical analysis and compared to their manual computation. The outcomes were similar in numerical value, indicating that a torque would need to be applied around the local B_x direction in order to counteract the effects of gravity of the segments B and C. Moreover, another torque would need to be applied around the local A_x direction in order to counteract the effects of gravity for the segments A, B, C.

The goal of the second verification step is to identify whether the evolution in time of the states is correct. In order to do this, an additional method that could solve the equations of motion was proposed (by using the Lagrange equations). A simplified scenario was defined with no acceleration at point O, and with no stiffness and damping parameters. An initially defined configuration was *"let free"* under the influence of the gravitational forces, and the evolution in time of the states q_1, q_2, q_3 and q_4 was recorded for both methods.

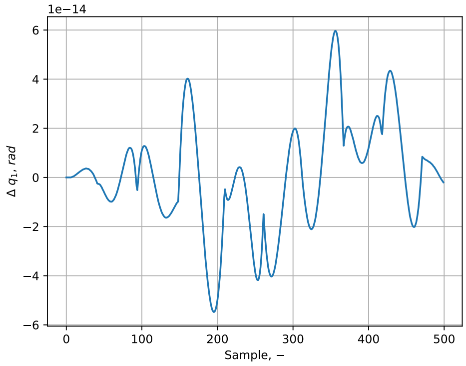


Figure 5.16: Overview of the difference between the state q_1 computed using Kane's and Lagrange's equations.

From Figure 5.16, it can be considered that the differences in state q_1 obtained using Kane's and Lagrange's equation are negligible. A possible explanation for the behavior of the graph is due to the influence of the numerical method that is used for solving the numerical differential equations. Differences of the same order of magnitude were observed for the rest of the states and the state derivatives, concluding to a successful verification procedure.

5.2.9. Additional implementations

An additional implementation was performed with the goal of defining a least-squares estimator for the torque components which describe the active and passive stiffness characteristics of the system. An initial simulation was performed with known active torque and passive stiffness characteristics while the states (generalized coordinates), derivatives, and double derivatives were stored. A separate program was then elaborated where new equations of motion were constructed and included four general torque variables (lumping the contribution of the active and passive components). By inputting all the stored information in Equation 5.10 (including the same geometric properties as the ones used in the initial simulation), the only four unknowns were the lumped torques at each time step. Since four dynamical equations were defined (four generalized coordinates), the unknown torques could be solved at each time step. Then, a least-squares estimator was defined as such:

$$\begin{bmatrix} 1 & q_1(0) & u_1(0) \\ 1 & q_1(1) & u_1(1) \\ 1 & q_1(2) & u_1(2) \\ .. & .. & .. \end{bmatrix} \begin{bmatrix} T_{01} \\ k_1 \\ c_1 \end{bmatrix} = \begin{bmatrix} T_{tot1}(0) \\ T_{tot1}(1) \\ T_{tot1}(2) \\ .. \end{bmatrix} \tag{5.11}$$

Where q_1 and u_1 are the stored generalized coordinates and velocities from the initial simulation, T_{01} is the unknown constant active torque for joint 1 around the direction corresponding to q1 and u1 (i.e x) and $k1$ and $c1$ are the unknown damping and spring coefficients of the joint 1 around the same direction. Lastly, T_{tot1} is the computed general torque around the same direction (computed from the new program). Under ideal conditions, it was observed that by solving the least squares equations, the estimated parameters were consistent with the parameters that were used to generate the data. This might represent an approach that can be used in combination with motion recording techniques. Using the 3D position information from the motion recording, the defined kinematical, dynamical model, and assumed limb lengths and masses (from regression equations), a set of estimated torques that create the recorded motion can be computed. With the computed torques, a certain model structure (like the one above) can be defined and parameters can be estimated. At this stage of this research, it is not known how noisy/imperfect the motion recording data is and in which way it will affect the inverse kinematics and dynamics approach. Moreover, since the motion recording techniques obtain the position of the limb features in a cartesian coordinate system, a methodology needs to be defined to transform them in the generalized coordinates space. Nevertheless, this method can be investigated in future stages of the research.

6

Preliminary Markerless Motion Capture Analysis

The goal of this chapter is to give the reader an overview of the theoretical background that has been used for the preliminary analysis linked to the camera model that is most widely used for triangulation applications. Secondly, a description linked to camera calibration and stereo reconstruction will be presented. Lastly, a series of validation steps are proposed and tested.

6.1. 3D reconstruction background

This section will arbitrarily use information from Stachniss (2021), Eisemann (2020), "Computer Vision Toolbox Camera Calibration" (2023), and "Camera Calibration and 3D reconstruction" (2023) in order to provide information regarding the camera model that is used for 3D reconstruction and the calibration techniques.

6.1.1. Camera model

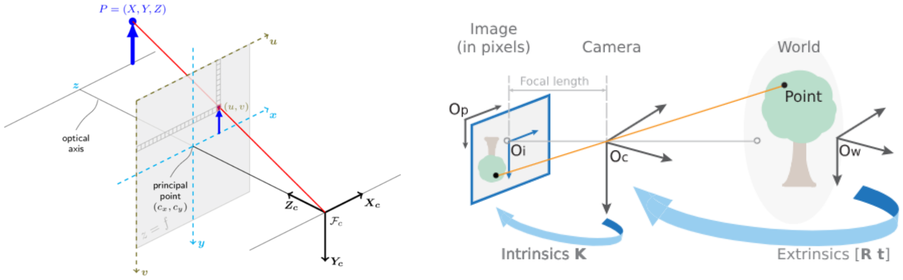


Figure 6.1: Two equivalent representations of the pinhole camera model "Camera Calibration and 3D reconstruction" (2023) (Left) and "Computer Vision Toolbox Camera Calibration" (2023) (Right).

One of the most often-used models used in photogrammetry applications is the pinhole camera model (Figure 6.1). The pinhole camera model idealizes how a point in the 3D world is projected on the 2D image plane. In other words, the pinhole camera model can provide a perspective projection of the 3D world. The camera model is described by a certain set of parameres such as: F_c/O_c camera center of projection, f - focal length (distance between the center of projection to the image plane), uv/Op plane (image plane), and a point P in the 3D world. Intuitively, it can be observed that a single camera can not solve for the 3D location of a point since a pixel location corresponds to a ray of 3D locations when taking into consideration a perspective approach. In this way, at least two cameras are needed to obtain the 3D location of a feature in the image plane.

In order to project a point in the 3D world on an image, three main steps need to be taken: Firstly, the world coordinate system that the point of interest is expressed in is related to the camera-fixed coordinat frame (3D to 3D transformation); Secondly, the ideal perspective projection is performed (3D to 2D transformation); Lastly, a metric-to-pixel transformation is performed (2D to 2D). Saying this, the general formula that lumps all the previously-mentioned steps is defined as:

$$m \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \boldsymbol{P} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}^w \tag{6.1}$$

Equation (6.1) presents the relationship between a point in the 3D world described in the X, Y, Z coordinates and the projected point u, v on the image plane. Since a perspective approach is used, both of the vectors are extended with one additional coordinate creating a set of homogenous coordinates. The homogenous coordinates can help with using one linear operation for both translation and rotation between reference frames. The constant m corresponds with the fact that homogenous coordinates are used for the perspective projection (the vector in homogenous coordinates of $[mu, mv, m]$ is equivalent to the vector $[um/m, vm/m, 1]$ in R^2. The matrix $\boldsymbol{P}$ is also called the camera matrix.

The first transformation expresses the world points ($\boldsymbol{X}^w$) in the camera-centred coordinate system ($\boldsymbol{X}^c$), suggested by Equation (6.7) (Eisemann, 2020).

$$\begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}^c = \begin{bmatrix} r_{00} & r_{01} & r_{02} & t_0 \\ r_{10} & r_{11} & r_{12} & t_1 \\ r_{20} & r_{21} & r_{22} & t_2 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}^w \tag{6.2}$$

The next step is represented by the (ideal) perspective projection where intermediate variables x' and y' are defined as:

$$x' = f\frac{X^c}{Z^c}$$
$$y' = f\frac{Y^c}{Z^c} \tag{6.3}$$

The quantities x' and y' are usually expressed in the units of the focal length f which is usually $[mm]$. The resulting image has pixel values, and therefore an additional transformation is needed to reach the final (u, v) values:

$$u = k_x x' + u_0$$
$$v = k_y y' + v_0$$
$$\tag{6.4}$$

By combining Equations (6.2), (C.7), (6.4), the resulting camera matrix is described by:

$$\boldsymbol{P} = \underbrace{\begin{bmatrix} a_x & 0 & u_0 \\ 0 & a_y & v_0 \\ 0 & 0 & 1 \end{bmatrix}}_{\text{Intrinsic parameters}} \underbrace{\begin{bmatrix} r_{00} & r_{01} & r_{02} & t_0 \\ r_{10} & r_{11} & r_{12} & t_1 \\ r_{20} & r_{21} & r_{22} & t_2 \end{bmatrix}}_{\text{Extrinsic parameters}} \tag{6.5}$$

The camera matrix presented in Equation (6.5) can be described by two matrices, an Intrinsic matrix that corresponds to the internal and specific parameters of the camera, namely focal length or sensor-to-pixel mapping, and an Extrinsic matrix which only describes the orientation of the camera with respect to a fixed world coordinate frame. Based on "Computer Vision Toolbox Camera Calibration" (2023), the Intrinsic matrix can be extended with other parameters in order to correct for distorted images (radial and tangential). It is expected that once the calibration was made, if the position of the camera is not changed and the focal length is not affected by automatic zoom or focus, the parameters for both of the matrices are constant. Moreover, if the camera is rotated with respect to the world frame, the intrinsic parameter matrix is expected to remain constant. Other possible sources of error such as changing the aperture due to different lighting conditions can be considered non-significant due to the fact that the aperture does not affect the focal length [1]. Saying this, in the description of "Computer Vision Toolbox Camera Calibration" (2023), it is recommended that the camera settings are set to "manual" in order to avoid any factors that can introduce time-varying changes in the intrinsic parameters.

6.1.2. Camera calibration

One of the most popular stereo calibration methods is the approach defined by Zhang (2000). This method is also implemented in the ready-to-use toolboxes for camera calibration in MATLAB and Python. This method makes use of a checkerboard pattern (Figure 6.2), in order to estimate the intrinsic parameters of the camera.

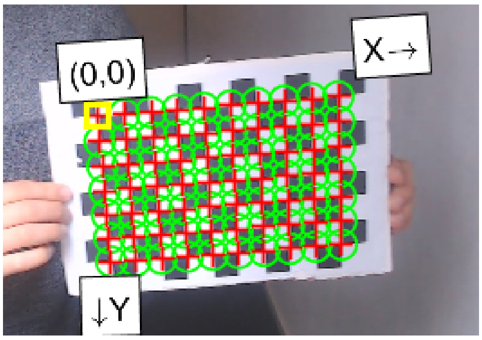


Figure 6.2: Camera calibration using a checkerboard pattern.

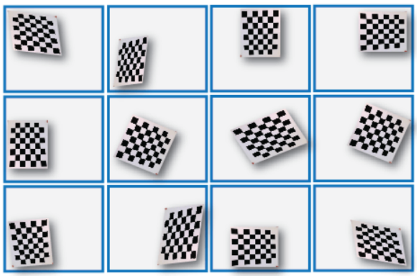


Figure 6.3: Multiple poses of the checkerboad presented in different images ("Computer Vision Toolbox Camera Calibration", 2023).

In order to estimate the camera parameters, a correspondence between the pixel locations and the world coordinate 3D point needs to be known. In this method, the fact that the checkerboard lies on a plane and the physical distances between the checkerboard squares is known simplifies the estimation problem. By defining a world coordinate frame woth the origin at one of the edges of the checkerboard, and by applying an edge detection algorithm, the 2D-3D pairs are found in a trivial manner. It needs to be mentioned that for each image of the checkerboard, the world coordinate system is changing, therefore this method is used only for estimating the intrinsic matrix for one camera. Since for every image, the global Z coordinate is zero, the total number of degrees-of-freedom of the combined matrix $\boldsymbol{P}$ decreases from 11 ($12 - 1$ since a projective space is assumed) to 8. In this way, each identified point of the checkerboard gives 2 equations (one for x-X correspondence and one for y-Y correspondence). Saying this, in order to obtain the reduced $\boldsymbol{P}$ matrix, at least four points need to be identified for each checkerboard. It needs to be mentioned that the reduced $\boldsymbol{P}$ matrix

---

[1] https://stackoverflow.com/questions/24299872/does-the-varying-of-aperture-influence-the-camera-calibration

changes from picture to picture (different positions of the checkerboard imply different global coordinate systems Figure 6.3). To derive the invariant part of the reduced matrix, a Singular Value Decomposition (SVD) problem is solved by combining multiple images with at least 3 views of the checkerboard (Stachniss, 2021). The exact mathematical details of the computation of the intrinsic matrix are considered to be outside the scope of the report and therefore, they will not be explained further.

It is highly reccomended that more than three images of the checkerboard pattern would need to be taken for a good-quality camera calibration (due to the inherent noise in the observations and for robustness purposes). One example is Figure 6.3, which presents the same checkerboard pattern in multiple images and poses. By displaying the checkerboard pattern in multiple poses, a higher robustness of the calibration is obtained. For a checkerboard photo collection obtained by using one camera, the intrinsic matrix is constant (assuming that the camera settings do not change), and the extrinsic matrix changes from picture to picture (due to changing global coordinate system). For one camera, Zhang's method estimates the (global) intrinsic matrix and the extrinsic matrix for each view. This is performed using the previously-mentioned SVD approach in combination with an optimization procedure which minimizes the squared error between the projected 3D edges of the checkerboard on the image plane and the ones identified in the image (by using for example Levenberg–Marquardt's algorithm (Huang et al., 2020)).

A quantitative measure which can help in evaluating the performance of the calibration is the so-called reprojection error. As presented in Figure 6.4, the reprojection error is defined as the difference in pixels between the edges identified on the checkerboard (green) and the projected (3D to 2D) edges of the checkerboard from the global coordinate frame (obtained by knowing the real dimensions of the squares on the board). Despite the fact that the goal is to obtain a reprojection error as close to 0 as possible, contributions from the checkerboard manufacturing, image processing errors and a poor image quality usually decrease the quality of the calibration (Huang et al., 2020).

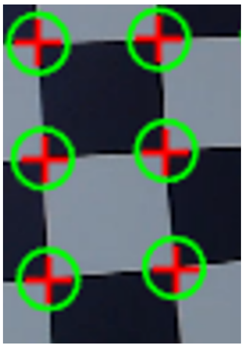


Figure 6.4: The reprojection error.

Besides the single-camera calibration feature, computer vision programming tools provide ready-to-use methods for stereo-camera calibration (two or more cameras). If synchronized pictures of the checkerboard are obtained for each of the cameras, by knowing the common world coordinate frame, the relative orientation between the two cameras can be estimated from a set of images. Usually, if a separate global reference frame is not defined, the triangulated location of the common feature that is observed from two cameras is expressed with respect to one of the cameras (which is set to be the global reference frame).

In this way, after the monocular camera calibration is performed separately for each camera, the intrinsic matrices are obtained. The extrinsics for one camera can be considered unity (center of the world) and for the other, the information obtained from the stereo camera calibration step can be used. By combining the intrinsic and extrinsic information of the cameras, each camera matrix $\boldsymbol{P}$ can be constructed and used for 3D reconstruction.

6.1.3. 3D reconstruction

The goal of the camera calibration was to obtain the complete camera matrices for two cameras. If, most commonly, these matrices are used separately to transform from a point in 3D space to a 2D one, by using the information from both cameras, it is possible to triangulate a common feature that is *”observed”* by both cameras in the 3D space.

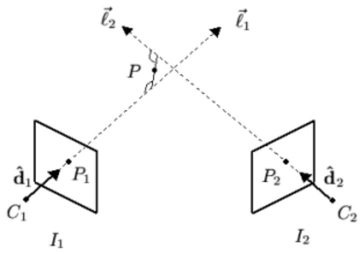


Figure 6.5: Triangulation of a common feature (P_1 and P_2) observed by two cameras (I_1 and I_2) (Slabaugh et al., 2001).

Figure 6.5 presents the intuition behind the triangulation. It might be the case that there are slight errors in the detection of the features P_1, P_2 and in the parameters of the camera matrices. In this way, the two lines (l_1 and l_2) will not intersect. One common method to solve for this issue is to solve for a minimization problem which finds the closest point of intersection between the two lines (here represented as P, not to be misunderstood with P, the camera matrix). One method that can be applied (which is also implemented in MATLAB's *triangulate* function) is the Direct Linear Transform.

By focussing on only one camera (Camera 1 with subscript 1), by denoting the 3D world point (expressed as a 4x1 column vector) $\boldsymbol{X_i}$, and by defining generic variables for the multiplication between the intrinsic and intrinsic matrix, the following (ideal) relationship can hold:

$$\begin{bmatrix} m_1u_1 \\ m_1v_1 \\ m_1 \end{bmatrix} = \begin{bmatrix} p_1^{11} & p_1^{12} & p_1^{13} & p_1^{14} \\ p_1^{21} & p_1^{22} & p_1^{23} & p_1^{24} \\ p_1^{31} & p_1^{32} & p_1^{33} & p_1^{34} \end{bmatrix} \boldsymbol{X_i} \tag{6.6}$$

By defining the three rows of the matrix as A_1^T, B_1^T, C_1^T:

$$\begin{bmatrix} m_1u_1 \\ m_1v_1 \\ m_1 \end{bmatrix} = \begin{bmatrix} — & A_1^T & — \\ — & B_1^T & — \\ — & C_1^T & — \end{bmatrix} \boldsymbol{X_i} \tag{6.7}$$

The resulting u_1 and v_1 pixels have the following form:

$$u_1 = \frac{A_1^T \boldsymbol{X_i}}{C_1^T \boldsymbol{X_i}}, \quad v_1 = \frac{B_1^T \boldsymbol{X_i}}{C_1^T \boldsymbol{X_i}}, ; \tag{6.8}$$

After further manipulations:

$$\begin{aligned} u_1 C_1^T \boldsymbol{X_i} - A_1^T \boldsymbol{X_i} = 0 \\ v_1 C_1^T \boldsymbol{X_i} - B_1^T \boldsymbol{X_i} = 0 \end{aligned} \tag{6.9}$$

From Equation (6.9), it can be observed that the system is underdetermined. In other words, there are three unknowns (the three coordinates of the unknown 3D point $\boldsymbol{X_i}$), and only two known quantities (the known pixel coordinates on the image, u_1 and v_1). The previously-mentioned fact proves why at least two cameras that see the same feature are needed in order to obtain the 3D location of the point. If a second camera (subscript 2) is introduced, then the following matrix equation can be defined:

$$\begin{bmatrix} u_1 C_1^T - A_1^T \\ v_1 C_1^T - B_1^T \\ u_2 C_2^T - A_2^T \\ v_2 C_2^T - B_2^T \end{bmatrix} \boldsymbol{X_i} = 0 \tag{6.10}$$

Equation (6.10) presents the scenario of obtaining the triangulated 3D point in ideal situations. In a real-life scenario, the observations from each camera will be noisy and the calibration of the cameras will not be perfect. This aspect imposes a non-zero value to the right of Equation (6.10). A solution to this problem is to use a least-squares method in order to solve for optimal intersection point. The solution for this least-squares approach is to use the Singular Value Decomposition of the transpose of the matrix which the point is multiplied with multiplied with the same matrix. Lastly, the eigenvector corresponding to the smallest eigenvalue of the resulting matrix can be considered the optimal point of intersection (the derivation of why is it this case has been omitted for brevity considerations).

6.2. Stereo reconstruction implementations

This section will present the approach that was used to perform the stereo 3D reconstruction using OpenPose and MATLAB's Camera Calibration toolbox. The approach will be presented in a step-by-step manner which will be concluded with several takeaways that can be used for future implementations.

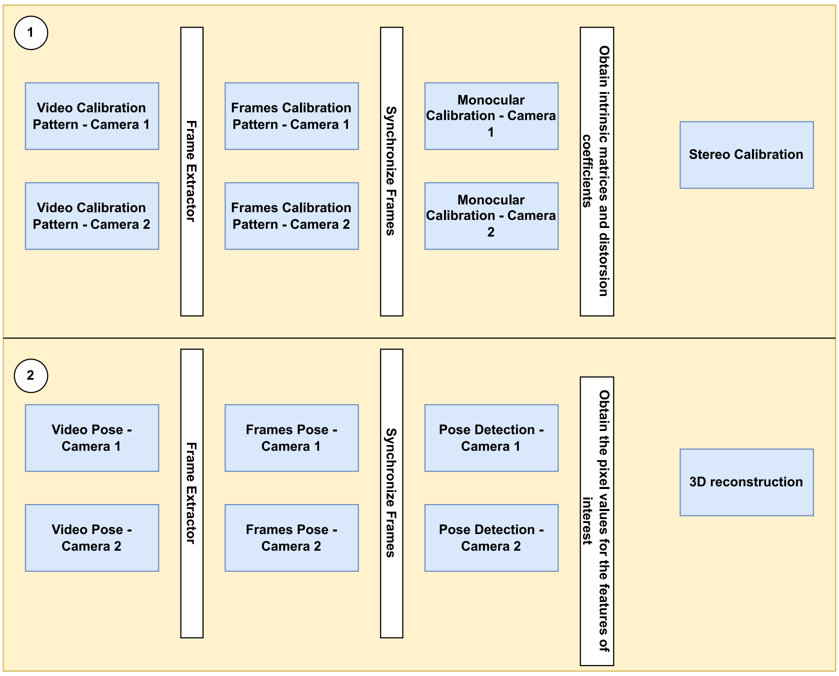


Figure 6.6: Overview of the stereo reconstruction process.

Figure 6.6 presents a short summary of how the author has decided to perform the stereo reconstruction. The methodology consists of two steps, the first one representing the stereo camera calibration (as presented in the previous section), while the second one represents the actual feature extraction and triangulation. In short, the first step consists of taking videos of the calibration pattern, extracting the frames from the videos, synchronizing the frames from the two cameras, performing the monocular camera calibration in order to obtain the intrinsic and distortion parameters for each camera and performing the stereo calibration for obtaining the relative rotation and translation between cameras. The second step also consists of video realization and frame extraction, moreover, the pose detection step is introduced in order to obtain the pixel coordinates for the features of interest, and lastly the 3D reconstruction uses the pixel coordinates in combination with the stereo calibration to solve for the 3D location of the landmarks. The following sections will present a preliminary experiment in which each of the above-mentioned steps is explained in more detail.

6.2.1. Experiment set-up

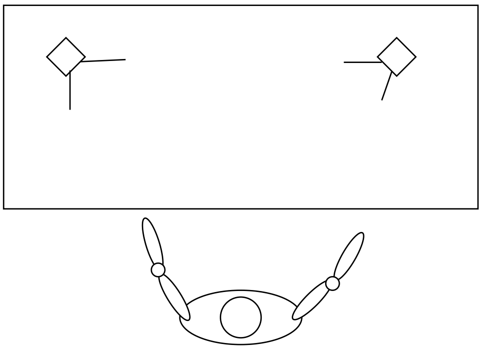


Figure 6.7: Setup used for investigating the feasibility of the stereo reconstruction method.

Figure 6.7 presents the setup of the cameras that has been used for evaluating the feasibility of the stereo reconstruction method. The angle of the cameras was chosen to be between 40 and $50\,[deg]$ as suggested by D'Antonio et al. (2021) in order to minimize the triangulation errors while being able to observe the features of the right upper limb from both cameras. Since OpenPose was used to extract the features from images, the goal for this analysis was at least to have the elbow, shoulder, and wrist features captured in the videos from both of the cameras.

6.2.2. Camera calibration

The stereo reconstruction step is dependent on the fact that both cameras need to be calibrated. This step extracts the intrinsic parameters and also estimates the relative rotation and translation of the cameras between each other. As mentioned in the previous sections, a calibration using a checkerboard pattern can be used and is supported by most of the toolboxes that support camera calibration. The used checkerboard pattern was described by 45x$45\,[mm]$ rectangles. During the camera calibration procedure, it is very important that the photos of the checkerboard pattern that are taken by both of the cameras are synchronized. Saying this, the user should ensure that both cameras operate at the same sampling rate. The solution for taking synchronized photos of the checkerboard pattern in different positions was to capture videos of the checkerboard which was moved in the image plane. The videos were synchronized by a light being turned on at a certain point in time or a *"clap"* was performed. After the videos were taken, the frames of the videos were extracted, and the frames until the event (light on or clap) were deleted for both cameras. This ensured that the frames from both cameras were aligned in time starting from the event. Based on the guidelines of using MATLAB's toolbox ("Computer Vision Toolbox Camera Calibration", 2023), synchronized frames were extracted such that the calibration pattern was captured at different orientations relative to the cameras. Moreover, the same source suggests that at least 10 to 20 images of the calibration pattern should be used and the angle of the pattern relative to the camera should be at most $45\,[deg]$. The paired frames are introduced in MATLAB's Stereo Camera Calibrator App (Figure 6.8).

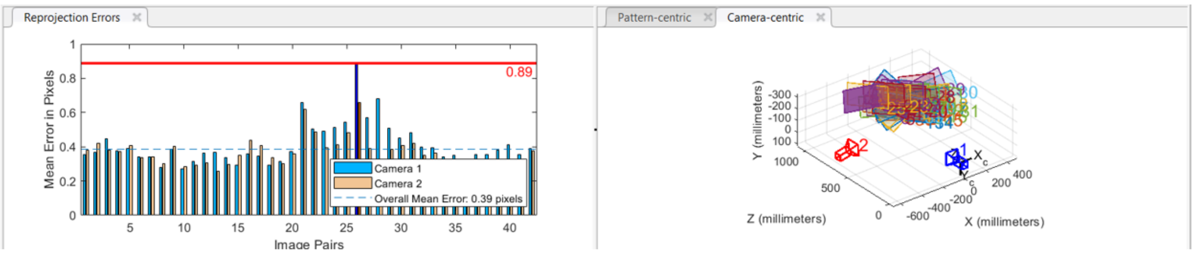


Figure 6.8: Outcomes of the camera calibration using MATLAB's Stereo Camera Calibrator App.

Figure 6.8 presents the outcomes of MATLAB's Stereo Camera Calibrator App. The box on the left shows the reprojection error for each pair of images. The reprojection error is the error in pixels between the point on the checkerboard which was identified by the edge detection algorithm and the reprojected point (on the image plane) using the camera matrix and the assumption that the global reference frame is at the top left of the pattern. Ideally, the literature mentions an acceptable reprojection error of maximum $1.5[px]$. The reprojection error can be considered a quantitative measure of the accuracy of the camera calibration. The box on the right shows the relative positioning between the cameras and the checkerboard in a camera-centric perspective.

6.2.3. Feature Identification and Stereo Reconstruction

After the cameras were calibrated, two new paired videos are taken which are inputted into the OpenPose feature extractor. The synchronization of the videos is performed in the same way as in the previous step. OpenPose can read a set of frames and output $.json$ files which indicate the pixel position of the identified features. It needs to be mentioned that during the preliminary tests, OpenPose could not be used to its highest accuracy setting due to limitations in computational power. OpenPose's user manual suggests using a computer fitted with a powerful GPU for optimal performance.

A custom MATLAB program was created to read the paired $.json$ files. With the paired pixel values corresponding to a common feature observed from two images, and with the parameters of the calibrated cameras, MATLAB's function $triangulate$ was used to compute the 3D reconstruction of the feature.

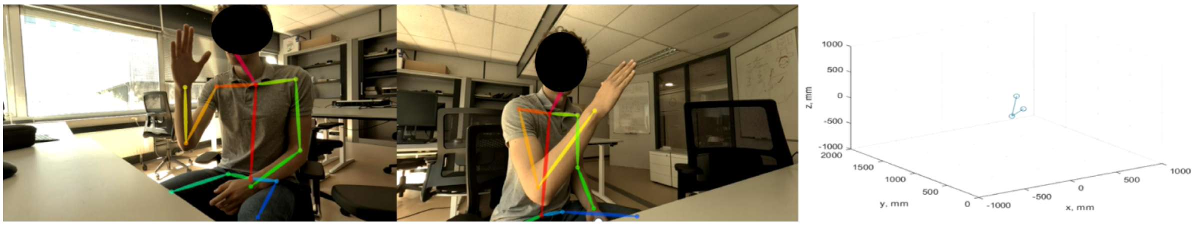


Figure 6.9: Outcomes of of the 3D stereo reconstruction.

Figure 6.9 presents one frame from a video for which the 3D reconstruction was applied. It can be observed that the global coordinate frame in which the 3D points are reconstructed is centered around one of the two cameras. In this case, the global Y coordinate is pointing away from the camera, the global Z coordinate is oriented vertically in the camera-fixed coordinate frame, and the last coordinate is orthogonal to the previous two, pointing to the right. Based on the comparison between the 3D reconstruction and the observed adjacent images, it can be mentioned that qualitatively, the pose of the upper arm is accurately described.

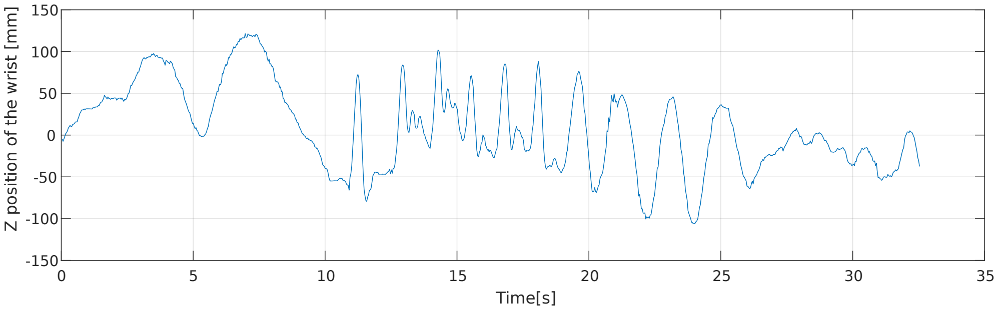


Figure 6.10: Triangulated Z position of the wrist landmark during a trial experiment.

Figure 6.10 presents one example of the Z position of the wrist landmark during a trial movement experiment. The trial was performed in static conditions and included functional movements of the arm in the image plane. Moreover, the author experimented with two neuromuscular settings (stiff and compliant) while performing movements that replicated the movement of the arm in disturbed conditions. The most prominent features can be observed in the time range from 10 to 18 [sec] (in the rest of the time, other functional movements were performed). The first, third, fourth, and fifth peaks replicated turbulent conditions with a compliant setting while the second and the third were performed with a stiff setting. It can be observed that the peaks obtained with a stiff setting end with a smaller displacement when compared to the other scenario, and their transients are more pronounced when compared to the other scenario. In other words, it can be considered that as expected from the simulated scenarios, the compliant setting attenuates more the highest frequencies. This fact can give the indication that there can be some fundamental differences that could be observed when analyzing the position signals when different settings of the neuromuscular system are induced in a disturbed environment. Lastly, it is also important to mention that it is not always the case that the Z position will coincide with the vertical position as perceived in reality. Since the coordinate frame is camera-fixed, it corresponds to the vertical plane in the camera-body frame. The following sections will provide a deeper understanding of the triangulation results.

Takeaways - Feature Identification and Stereo Reconstruction

Several takeaways can be noted from the different trials that were tested for the 3D reconstruction.

Initially, it was observed that the performance of the calibration (reprojection error) varied between different calibration trials. In some of the cases, the overall reprojection error for a calibration session was $11\,[px]$ which was considered unacceptable. After further research, it is hypothesized that the differences between the calibration sessions were due to changes in camera settings, and due to incorrect synchronization of the frames (it is important that the cameras record at the same framerates). As previously presented, the camera matrix consists of an intrinsic matrix describing the internal camera parameters and an extrinsic matrix that describes the positioning of the camera in the external world. New insights gained from MATLAB's guidelines mention that autofocus and changing zoom settings affect the focal length and thus the internal parameters of the cameras. Saying this, it is preferred that the camera settings are set to *"manual"* in order to avoid any changes in the camera parameters. Moreover, it is expected that besides the focal length, no other parameter from the outside world would affect the internal camera characteristics. It was mentioned that the aperture (affected by outside lighting) theoretically does not affect the camera parameters. It is expected that if the camera settings are set to manual and the intrinsic parameters are obtained, these parameters can be kept constant. If the relative positioning between the cameras also does not change, then it is hypothesized that once an initial calibration was performed, then additional calibrations are not needed.

Additionally, as presented by the literature, some artifacts were induced by OpenPose wrongly identifying the left hand to be the right hand due to similarities in the 2D perspective. Further research is needed to identify the optimal camera positioning that reduces these artifacts.

Lastly, the accuracy of the obtained 3D positions during the current trials was not validated. Other studies make use of marker-based camera systems in order to perform the validation procedure. A possible qualitative validation is to check similarities between the movement patterns in the videos with the movement patterns which are obtained in the 3D-reconstructed scene. The following section will give an overview of other several validation approaches which can prove/disprove the correctness of the method.

From a practical perspective, for each run, the videos that are taken need to be matched in time. A clear methodology of how the synchronization in time needs to be made along with a data analysis pipeline which starts from obtaining the videos to outputting the 3D locations of the desired features.

6.2.4. Validation procedures

According to Huang et al. (2020) the 3D reconstruction error of a point can have contributions from the error in the calibration (due to possible inaccuracies in the edge-detection algorithm or slight deviation between the frames) and from the error in matching the same feature in the two images (OpenPose might consider slightly different points to represent the joint centers based on the perspective of the cameras). Saying this, it is important to get a feel for the accuracy of the system and understand if its accuracy is sufficient for the current application. Unlike the qualitative validation which can only point out clear inaccuracies, it can be considered that a quantitative validation can provide a better understanding of the working principle and weaknesses of the current system.

The literature study chapters presented one of the most common methods for validating a markerless system, by using a more complex marker-based system. One possible disadvantage of this approach at this stage of the project is that the validation can not be done in the same environment as the experiment due to the physical constraints that a marker-based system is facing. Moreover, it can be considered that various calibration steps would need to be performed in order to get equivalent results from both of the systems. Two examples of such calibrations would imply ensuring a similar global reference frame for both systems and ensuring that the markers are placed in accordance to how the markerless system is able to detect a particular landmark. The author proposes that it is worth investigating less constrained and workload-intensive validation methods (with less accuracy) before taking into consideration a validation step using a marker-based system.

Additional validation approaches proposed by the author would consist of simple experiments where specially-designed movements are performed and the distances covered by the limb are compared with a known physical object (whose dimensions are known). Nevertheless, one last validation approach would be to use the same metric as the one used in the calibration step, namely the reprojection error of the triangulated point. The following section will focus on a set of experiments linked to the first approach that was proposed, namely comparing the limb movements with known distances in the real world.

The first validation experiment

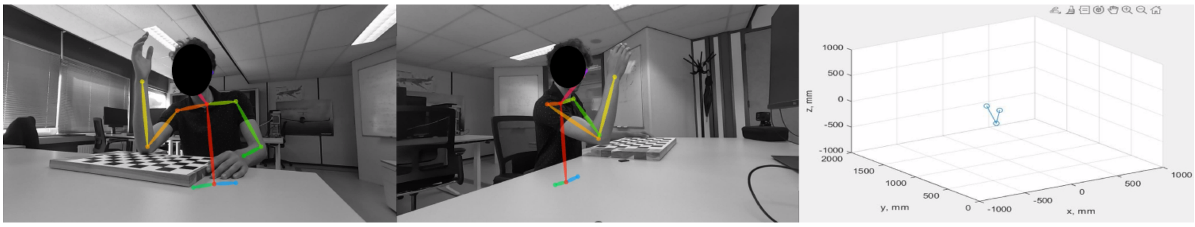


Figure 6.11: Overview of the first validation experiment.

Figure 6.11 presents a visual overview of the intuition behind the first validation experiment. The author used the intuition that the checkerboard pattern squares have a fixed and known geometry. It was proposed that the elbow would be always in contact with the table while it was moving in a straight line along one of the checkerboard's edges (the edge parallel to the edge of the table). In this way, by evaluating the time traces of the elbow landmark, a first-order approximation of the triangulation performance could be obtained.

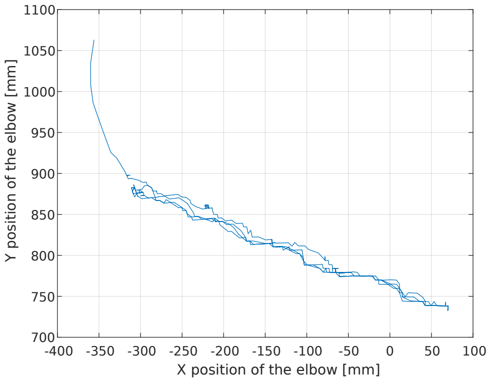


Figure 6.12: Triangulated position of the elbow in the XY plane of the camera-fixed global reference.

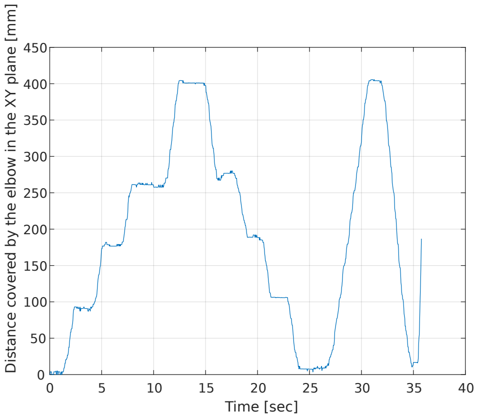


Figure 6.13: Euclidean distance in the XY plane of the elbow feature expressed with respect to the first frame of the recording.

Figure 6.12 presents the evolution of the (X,Y) pair of coordinates of the elbow during the validation trial. Since the movement of the elbow was performed across one of the checkerboard pattern's edges, the expectation would be to observe a linear trend between the two *"horizontal"* coordinates. The trends in the previously-mentioned figure replicate well a linear dependence between the X and Y coordinates. Since the raw data is used for the time traces, the obtained trends possess several imperfections (they are noisy). The cause of the imperfections is also represented by the noisy identification of 2D features along the trial (the pixel values of the identified elbow feature oscillate due to uncertainties in the landmark detection). It needs to be mentioned that the nonlinear part of the graph (top left) corresponds with the end of the trial when the subject took his elbow off the table.

In order to better quantify the accuracy of the method, another perspective on the problem would need to be presented, such as the one shown in Figure 6.13. The figure presents the Euclidean distance in the XY plane of the elbow feature, with respect to the initial frame of the video. The initial frame of the video coincides with the elbow position being at the leftmost edge of the checkerboard (seen from the camera perspective). As previously mentioned, the subject performed a movement such that the elbow was always in contact with the surface of the checkerboard while following one of its edges. From Figure 6.13, it can be observed that during the trial, the subject performed two excursions in a straight line to the other edge of the checkerboard and back. Moreover, in the first excursion, the subject stopped briefly once every two squares in length, besides the last stop which included three squares in length. Additionally, the figure shows that the length of one excursion is close to the value of 400 $[mm]$ which can be considered a good approximation of the value of 405 $[mm]$ (representing the length of the checkerboard with 9 (2+2+2+3) squares with one side having 45 $[mm]$). This result can be interpreted as a first-order approximation of the accuracy of the system, since the exact position of the elbow with respect to the checkerboard edges during the trial could have faced human errors, and additional errors could have been introduced by (an assumed) inaccurate detection of the elbow joint from the pose estimator.

An additional perspective that can be used to investigate the correctness of the approach is to use the information that the elbow was always in contact with a hard surface. In this way, it is expected that the vertical elbow position should not change during a trial.

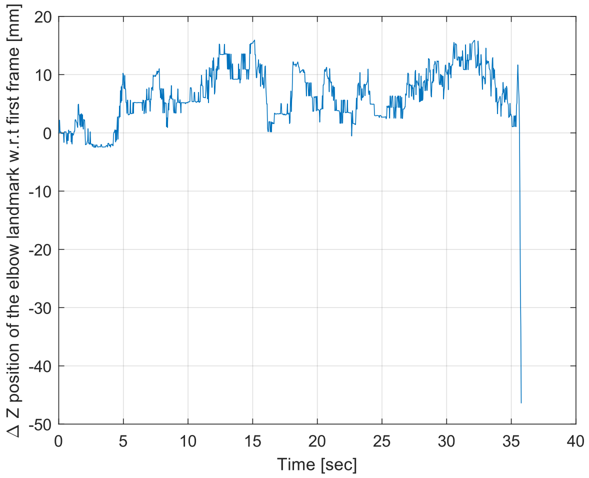


Figure 6.14: Difference in elbow's landmark Z position with respect to the initial frame.

Figure 6.14 presents the evolution in time of the elbow's triangulated landmark in the Z direction of the global reference frame during the same trial presented in the previous paragraph. It can be observed that the vertical position of the elbow appears to change slightly throughout the trial, with the biggest difference with respect to the initial frame of approximately 15 $[mm]$. Moreover, an interesting trend can be observed in the time frame between 25 and 35 $[sec]$ (which corresponds to the second excursion to the other edge of the pattern). The trend in the observed signal shows that when the elbow landmark was moving towards the right from the initial point (as seen from the camera's perspective) the apparent *"vertical"* position increased (with respect to the starting point) while when the landmark was moving back, (towards the left direction) the apparent vertical position decreased. The author hypothesizes that the reason for this behavior is the positioning of the camera which is not perfectly aligned parallel to the movement and more importantly exactly vertical (as perceived by the humans). This, in combination with the fact that the coordinate frame that the landmarks are expressed in is fixed to the camera, might be the cause for this observation. In order to further investigate this behavior, the author has proposed a second validation experiment.

The second validation experiment

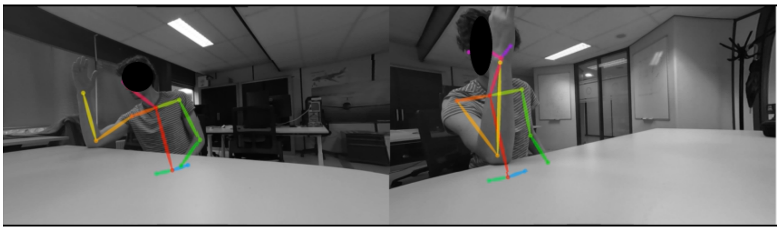


Figure 6.15: Overview of the second validation experiment.

Figure 6.15 presents the setup of the second validation experiment. In order to investigate whether the same linear trends can be observed in the Z component of the elbow landmark, an additional experiment was proposed where larger excursions of the arm were performed. In this scenario, the elbow was kept in contact with the edge of the table presented in the image, while it was moved to the right (as seen by the camera) and back to the left parallel to the edge.

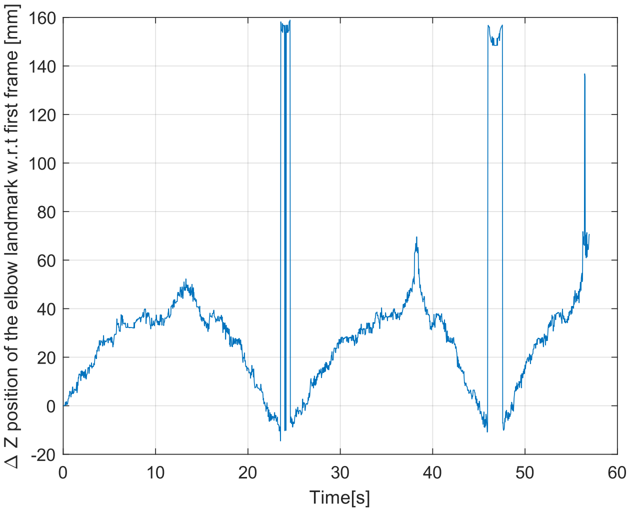
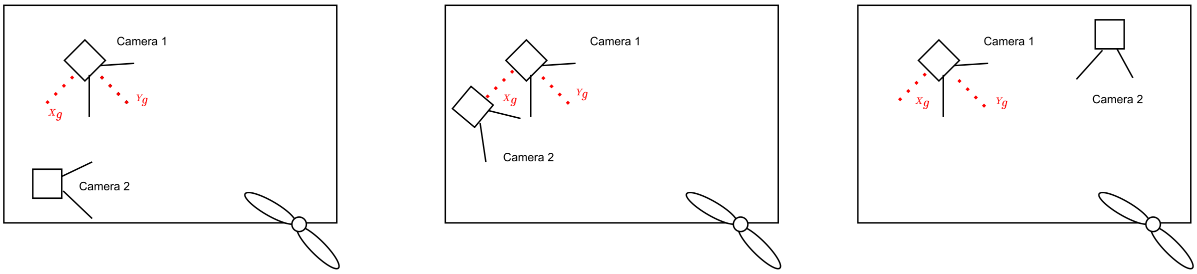


Figure 6.16: Difference in elbow's landmark Z position with respect to the initial frame (second validation session).

Figure 6.16 presents the outcomes from the second validation session. The dependent variable is similar to the one presented in Figure 6.14, namely the difference in the global Z position of the triangulated elbow's landmark with respect to the initial frame of the video. The time traces of the figure suggest that three arm excursions from right to left (as observed from the camera's perspective) and two excursions from left to right were performed. Moreover, since the excursions were larger (in the XY plane) compared to the previous scenario, a larger and clearer difference can be observed in the global Z component (up to 50 $[mm]$ when ignoring the signal losses) where in reality, if the camera is perfectly aligned vertically, there shouldn't be any differences (since the elbow was always tangent to the surface of the table). Again, the author hypothesizes that this increase and decrease in the dependent variable could have only one cause, namely that the camera-fixed coordinate system (at least in the vertical direction) is not aligned with the vertical direction as perceived in the real world. In other words, the camera is tilted to the left/right and up/down. One option to correct this inaccuracy would be to change the coordinate system the points are expressed in. One option would be to follow the intuition based on D'Antonio et al. (2021), namely to take one picture of the checkerboard pattern while it is perfectly oriented vertically and to obtain the rotation matrix between the camera and the pattern (by using the Camera Calibration Toolbox). In this way, all the points could be expressed with respect to the new coordinate frame (now with a better *"true"* vertical approximation) instead of having the points expressed in the camera frame. Further experiments that investigate the previously-mentioned approach can be performed in the thesis phase as a preparation for the main experiments.

The third validation experiment

The last validation experiment uses the intuition that the triangulated points are expressed with respect to the fixed coordinate system of one camera. In this way, the author proposes to alter the position of the second camera, while the position of the first one (and automatically of the global coordinate system) is kept constant. One important aspect is represented by the fact that the subject would need to keep his/her limb in a consistent position between the trials in order to be able to accurately compare the triangulation outcomes of each of the positions of the second camera. In this way, it is expected that by changing the position of the second camera (thus keeping the global coordinate system constant), and by keeping the limb in a consistent position, the triangulated landmarks shall be consistent between the trials.



Case 1 Case 2 Case 3

Figure 6.17: Setup used for the third validation experiment.

Figure 6.17 presents an overview of the cases that were investigated during the third validation trial. For this experiment, three main cases were selected, where the position of the second camera was altered. It needs to be mentioned that the global coordinate system (emphasized in red) is kept constant since the position of the first camera does not change. In order to keep the position of the features on the arm consistent between the trials, markings were defined on a horizontal surface which indicated the recommended position of the wrist and elbow (which are in contact with the horizontal surface of a table during one trial). When calibrating the cameras, the reprojection error for each of the three cases was 0.48, 0.49, and 0.57 $[px]$.

Figure 6.18: Overview of the third validation experiment observed from the perspective corresponding to Case 1.

Figure 6.18 presents an overview of the setup of the third validation experiment (camera positioning corresponding to Case 1). It needs to be mentioned that slight inaccuracies between the trials can be introduced by small differences in the subjective positioning of the elbow and wrist on the table.

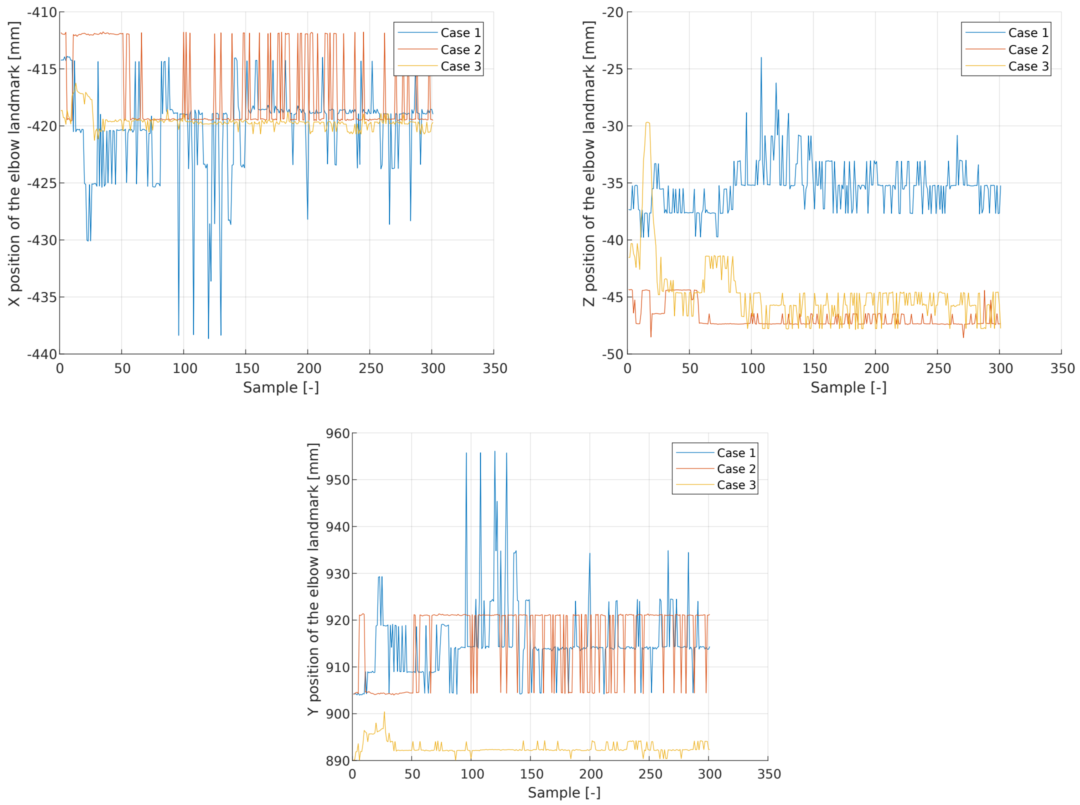
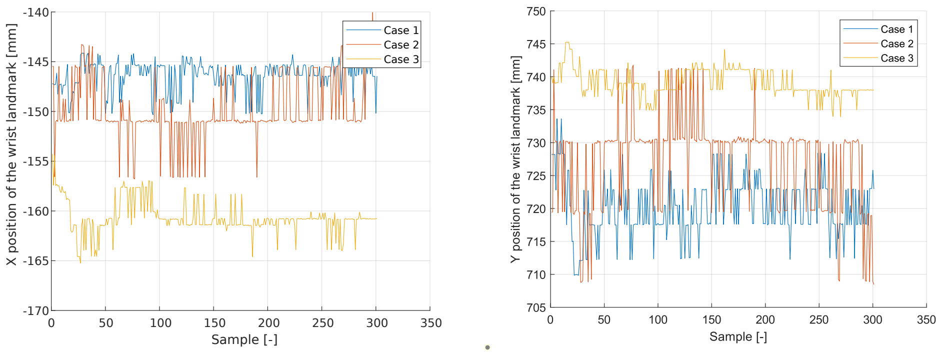


Figure 6.19: Comparison between the X, Y, and Z positions of the elbow across the three proposed cases.

Figure 6.19 presents the comparison between the X, Y, and Z positions of the elbow landmark across the three cases. From all representations, it can be observed that the maximum differences between the trials are in the order of magnitude of centimeters (more precisely, approximately 3 $[cm]$ when observing the Y position of the elbow and ignoring the noisy peaks). Nevertheless, it can be observed that the joint positions are noisy even though the limbs were not moving. From the figure, it can be observed that the highest uncertainty in the joint location (the signal with the highest spikes) is obtained when the camera is positioned directly in front of the subject (Case 1). In this way, due to the perspective that the camera observes the 3D world, the algorithm has difficulties in accurately describing the location of the limb features in the first scenario.



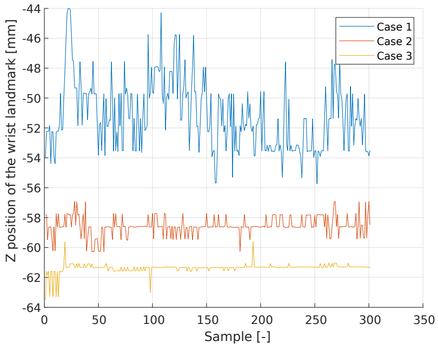


Figure 6.20: Comparison between the X, Y, and Z positions of the wrist across the three proposed cases.

The second feature which was kept in the same place across the trials was the wrist. Figure 6.20 presents the comparison between the X, Y, and Z positions of the wrist landmark across the three cases. As presented in the investigation which concerned the elbow, the signals can be considered consistent between each case with maximum differences of 2-3 $[cm]$. As in the previous analysis, the noise in the time series is due to the way the pose estimation algorithm detects the landmarks on the limb.

It can be observed that when analyzing both the elbow and wrist locations, small differences appear between the three cases. The author hypothesizes that the differences can be due to: 1) The algorithm detecting the landmarks slightly differently based on the location of the cameras, which affects in the end the triangulation outcome; 2) The subject slightly changing his pose between the trials. Nevertheless, the above-mentioned results can represent a first-order approximation and demonstration of the consistency between the trials.

Takeaways - Validation procedures

The goal of the validation experiments was to give a first-order approximation of the accuracy of the stereo system. The first two experiments showed that it is important to keep in mind that the outcomes of the triangulation are relative to a coordinate system fixed to one of the cameras. In this way, it is not always the case that the local vertical coordinate system fixed to one of the cameras is aligned to the vertical direction as perceived in a real-life scenario. To solve this *"problem"*, an additional coordinate transformation would need to be performed.

The last experiment validated the consistency of the triangulation approach, by varying the position of the second camera of the stereo system. Errors of at most 3 $[cm]$ were observed between the three test cases. The author hypothesizes that the origin of the errors stems from a combination of the subjective positioning of the limbs and a possible difference in landmark identification across the trials. An additional conclusion that can be taken is that not all the camera positions are suitable for feature identification. As observed from Figure 6.18 (right), slight inaccuracies in feature identification can appear when positioning a camera straight in front of the user due to the perspective from which the camera observes the 3D world.

One last outcome which was not presented in the validation section is the comparison between two triangulation methods. MATLAB's function *triangulate* was compared to the *DLT* method. When comparing the two methods, exactly the same results were obtained concluding that the *triangulate* function uses the *DLT* approach for triangulation.

6.2.5. Additional implementations

Even though the last sections focussed on tracking three individual features of the limb, namely the shoulder, elbow, and wrist, it is also interesting to explore the motion-tracking capabilities of the hand. The intuition provided by the author's supervisor, that by tracking the user's finger in a disturbed environment a prediction/correction of the intended activation location can be made, is certainly worth investigating for pointing tasks.

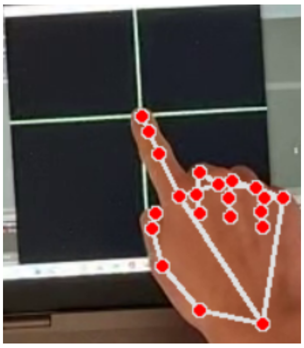


Figure 6.21: Overview of tracking the hand's features using OpenCV.

Figure 6.21 provides one example of extracting the features of the hand by using OpenCV's library. It needs to be mentioned that OpenPose also has the feature of extracting the features of the hands, but at the moment of writing the report, it was not tested due to the more advanced hardware requirements. Nevertheless, it is expected that during the thesis phase, OpenPose's hand-feature extractor will be used in order to have a consistent implementation with the feature tracker of the arm.

Once the finger position is tracked, the goal of the analysis would be to investigate whether it is possible to *"smooth out"* the effects of disturbances on the finger position in pointing tasks. Ahmad et al. (2018) suggests several methods for removing the perturbations of a noisy hand trajectory, by using Kalman filters (assuming a linear motion) or Sequential Monte Carlo particle filtering. Moreover, the same study also proposes a Bayesian predictor in which given the finger's location at a specific time point, the most probable selectable interface item is computed (without needing a physical activation on the screen). The author also proposes to investigate the intuition that the desired movements (e.g. reaching for the screen in pointing tasks) could be isolated/filtered in the frequency domain from the undesired movements, given the assumption that the arm responds linearly to the applied disturbances. In this way, it would be possible to *"filter out"* the unwanted movements during the reaching tasks and correct the location on the screen that the user pointed at. In the thesis phase, a clearer analysis of the possibilities for combining finger tracking and filtering techniques for obtaining a smooth trajectory and target selection in pointing tasks will be elaborated.

6.2.6. Influence of the camera positioning on the pose estimation

The last part of the analysis for the 3D reconstruction part concerns the investigation of the positioning of the cameras that can provide a successful pose estimation. In the author's opinion, a successful pose estimation implies that the features of interest are observed simultaneously from both cameras, and the pose recognition contains as few errors as possible while keeping in mind the constraints posed by the environment in which the cameras will operate. In order to investigate feasible camera positions, three cases were defined, while recognizing that some of them are similar to the ones presented in Figure 6.17. Additionally, the author will also comment on the setups presented in the earlier parts of the current section.

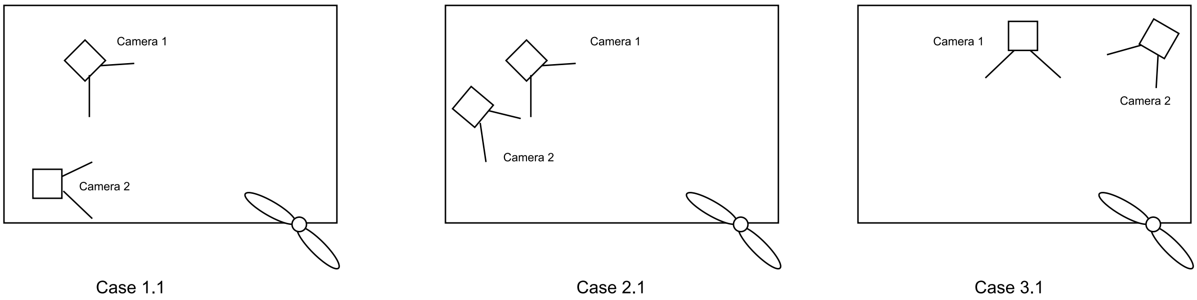


Figure 6.22: Setup used for investigating feasible camera positions.

Figure 6.22 presents the three cases that were initially proposed for the analysis. During the tests, the subject kept his arm horizontal and parallel to one of the table's edges.

Figure 6.23: Overview of the pose estimation corresponding to Case 1.1.

Figure 6.23 presents one frame corresponding to the setup attributed to Case 1.1. This case corresponds to one of the cameras facing directly the user. Despite the apparent *”complete”* feature detection, several aspects need to be pointed out. Firstly, it can be observed that the wrist detection for Camera 1 (left) was not accurately performed. While the reason for the inaccurate detection is not known, the author assumes that certain features on the subject's arm allowed for such a detection. Moreover, if the attention is directed to the second camera (right), it can be observed that due to the current perspective, several features of interest (such as the elbow or the wrist) can be occluded by the hand. If the hand would have been moved more toward the camera, it is assumed that it will cover most of the image, making the detection of other features of the arm infeasible. Nevertheless, this case emphasizes a possible trade-off that would need to be made between the feasibility of tracking the features on the arm and tracking the features on the hand.

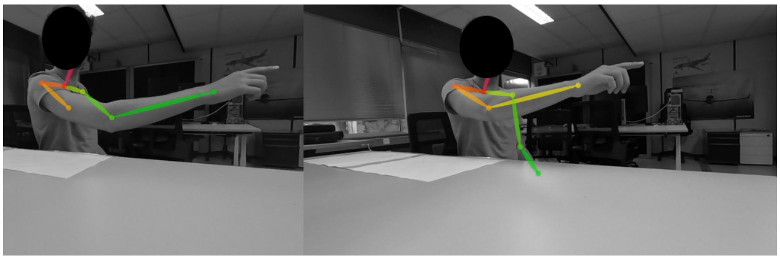


Figure 6.24: Overview of the pose estimation corresponding to Case 2.1.

Figure 6.24 presents one frame corresponding to the setup attributed to Case 2.1. The following case corresponds to the cameras being positioned on one side of the subject, observing the dominant hand. One particular problem with this positioning can be observed in Camera 1 (Left). In some frames, the algorithm combines the left shoulder landmark with the landmarks of the right arm. This confusion might be due to the fact that the left arm was not detected in the frame. Another reason for the confusion is the fact that the perspective of the first camera observes the line describing the two shoulders being almost perpendicular to the camera plane, and therefore it is assumed that the algorithm has difficulties in differentiating between the right and the left shoulder.

Figure 6.25: Overview of the pose estimation corresponding to Case 3.1.

Figure 6.25 presents one frame corresponding to the setup attributed to Case 3.1. This setup has one camera which is positioned to the back of the subject and one which is positioned parallel to the right limb. By observing only the selected frame, this positioning can be considered favorable for the stereo system, since it shows that all the features of the right upper limb are captured by both of the cameras. On the other hand, after further investigations, it was observed that for Camera 1 (Left), it can be the case that the wrist feature is occluded by the upper arm. Moreover, the location of the finger is hardly identifiable from the same perspective. Saying this, the difficulty of finding a positioning that both captures the features of the limb and hand needs to be emphasized.

Several other camera positionings can be considered valid to be used for the stereo system. One example of such positioning is represented by Figure 6.9, where both cameras are in front of the subject, facing him/her. In this way, the author assumes that it is possible to track not only the features of the upper limb but also the features of the hand. Currently, it is not known whether this positioning is permitted by the restricted space of the simulator.

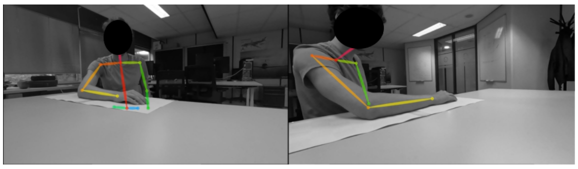


Figure 6.26: Overview of the pose estimation corresponding to Case 3.

One last camera positioning that can be considered feasible is the one corresponding to Case 3 of Figure 6.17, presented in Figure 6.26.

Takeaways - Camera positioning

The goal of the previous subsection was to investigate which camera positioning would suit the best the current application. It is expected that during the experiment phase, two types of tasks would be investigated, namely pointing tasks and prolonged interaction tasks. In this way, not only the features of the upper arm would need to be tracked but also certain features of the hand. One clear conclusion that can be taken is that it can be difficult to position the cameras such that they can record well both types of features. Another conclusion that can be taken is that the pose tracker is prone to erroneous outcomes such as inaccurate and incorrect labeling of some features. As a first-order approximation, the poses from the Figure 6.9 and Figure 6.26 can represent two feasible solutions for the positioning of the cameras. Nevertheless, it needs to be mentioned that multiple frames of a certain camera positioning would need to be analyzed in order to suggest better conclusions (compared to only one frame per positioning as presented in this subsection). The author recommends a further analysis of the positioning of the cameras in the constrained space of the simulator. It is hypothesized that by positioning the cameras higher than the subject and orienting them towards him/her, would be a way of capturing all the necessary information while minimizing the effect of the occlusions and of the false detections. Moreover, an environment with good lighting and contrast between the subject and the rest of the frame would ensure more reliable results from the pose estimator.

7

Research Question(s) and Experiment Description

The aim of this chapter is to propose the questions that will be answered during the thesis phase. First, the research objective is reiterated, followed by the elaboration of the research questions based on the main concepts identified in the research objective. Lastly, a description of the experiments will be given and the hypotheses will be presented.

7.1. Project goal and research questions

The proposed research objective is defined in the following way:

Investigate the feasibility of creating adaptive BDFT models for touchscreen operation by using motion analysis techniques.

Before continuing with the definition of the research questions, a brief recapitulation of the previous chapters can be performed:

- Based on **Chapter 2**, it was concluded that two of the most common tasks that are encountered in touchscreen operation are *discrete-pointing* tasks and *dragging* tasks;
- **Chapter 3** showed that one important aspect that influences the BDFT in stick-based operations is the adaptive nature of the NMS system. Saying this, for an accurate cancellation, the BDFT models would need to be adapted to the setting of the neuromuscular system that is imposed by the task or subjectively used by the HO. Moreover, it was concluded that the problem of within-subject variability concerning the BDFT effects in touchscreen operation was not yet tackled.
- Based on **Chapter 4**, it was reported that one feasible solution for recording limb features is to use markerless pose estimators. Moreover, by combining information from multiple cameras, it is feasible to obtain the 3D location of the desired features.

The literature study emphasized the limited body of knowledge present in the area of quantifying the effects of BDFT in touchscreen operation. Most of the presented models were developed focusing on the interaction of the HO with physical inceptors. As previously described, BDFT modeling and mitigation needs an adaptive approach in order to be effective due to inherent within and between-subject variability. From the author's knowledge, no study analyzed the individual arm movements of the subjects during touchscreen operation in disturbed environments. This study would like to investigate whether is feasible to use motion-capturing techniques to help model and predict the BDFT effects in an adaptive manner, by linking the movement of the limbs to the signals observed at the screen level and to the setting of the neuromuscular system selected by the subject.

The resulting research questions are the following:

a) *In which manner is the subject's arm movement affected by the introduction of disturbances when interacting with a touchscreen?*

- *What parts of the arm move relative to each other when disturbances are added?*
- *What parts of the arm are most affected by the disturbances?*
- *Does a correlation between the recorded movement of the limbs and the motion disturbance exist?*

b) *In which manner do different neuromuscular settings affect the recorded movement patterns of the limb when interacting with a touchscreen?*

- *Is it possible to capture the changes in neuromuscular settings with the chosen motion-recording technique?*
- *What landmarks of the arm are most affected by a change in the neuromuscular settings?*
- *How do the time-traces of the recorded landmarks change with a change in the neuromuscular settings?*
- *Does a change in the neuromuscular setting have an effect on the BDFT observed at the touchscreen input?*

c) *How do the recorded movement patterns of the arm in a disturbed environment change when the hand is not in contact with a touchscreen?*

d) *What is the approach that the HO takes when performing a discrete-pointing task on a touchscreen?*

7.2. Experiment design

The experiment design will give an overview of the current plan of the experiments along with a description of the hypotheses that will be tested. As a general rule, the previous section brought to attention the possibility of investigating two types of tasks, namely dragging tasks and discrete pointing tasks. The following section will be split up based on these two types of tasks.

7.2.1. Dragging task experiment

Goal of the experiment

The goal of the first experiment is to better understand the insights of how the upper limb is affected by the induced disturbances while the upper limb is in contact with the touchscreen and while inducing different settings of the neuromuscular system. Despite the fact that the title of the experiment includes the words *"dragging task"*, the task will be designed such that the subject would need to keep his/her hand continuously in contact with the screen at one particular location. The reason for this choice will be given in the next parts of the subsection.

Experiment and task description

Since the experiment requires the investigation of the arm's movement in a disturbed environment, TU Delft's SIMONA Research Simulator will be used as the main resource.

Figure 7.1: Outside representation of the SIMONA Research Simulator.

One possible representation of the simulator is given in Figure 7.1. It consists of a moving hexapod facilitating movement in 6 degrees-of-freedom, a cockpit mock-up where a variety of instruments (including the touchscreen) are mounted, and an outside-visual section.

The current study will focus on investigating the effects of BDFT when the disturbance signal is applied in one direction, namely in the vertical one. The signal that will be used to replicate a disturbed environment was selected to be the one used by Mobertz et al. (2018) and Khoshnewiszadeh and Pool (2021). The reason for selecting this signal is due to the possibility of extracting the frequency content from the signal recorded on the screen and from the time traces of the limb landmarks which is linearly related to the frequency content of the input signal (which is an acceleration signal). In this way, it would be possible to idealize the relationship between the imposed accelerations and the recorded signals in the form of transfer functions.

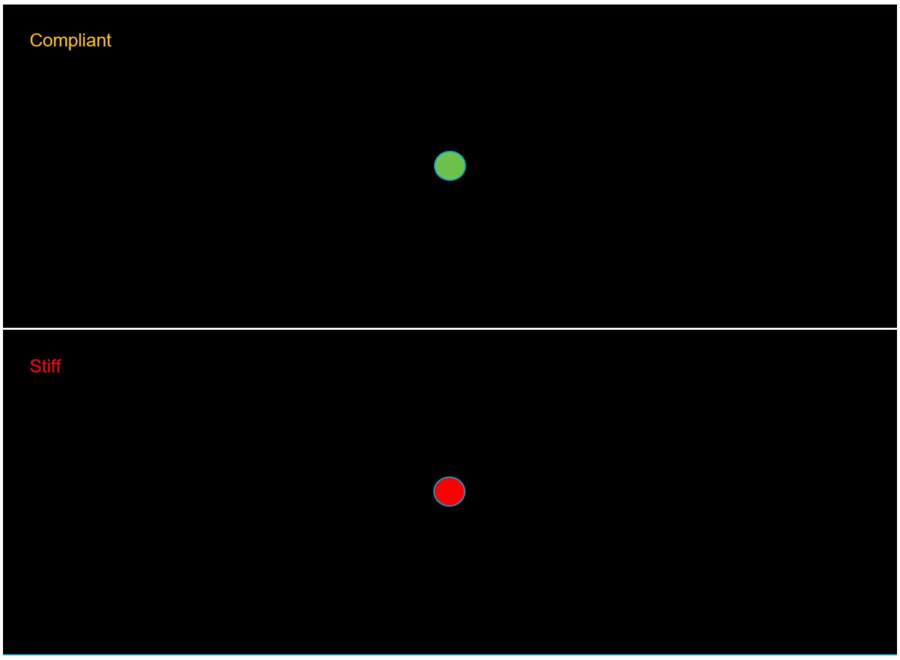


Figure 7.2: The variations of the task which has been proposed for the first experiment.

Figure 7.2 presents the two variations of the task which was proposed for the first experiment. The subject would need to keep his/her finger pointed at the red/green dot while encountering vertical disturbances. Moreover, two words can appear on the screen namely *”Stiff”* and *”Compliant”*. These words correspond to the two neuromuscular settings that the subject is suggested to use during the task. In this way, a comparison between the BDFT which is observed in a stiff setting can be made with the BDFT observed in a compliant setting. It needs to be pointed out that the two imposed neuromuscular settings are subjective and can vary from person to person. On the other hand, the same subjectiveness of the task (or neuromuscular setting) was also encountered in the experiments concerning stick-based operations (in position and force types of task). The reason for choosing a task that imposes a constant position of the finger throughout the trial is the hypothesis that it would facilitate a much easier context for keeping a constant neuromuscular setting. When taking into consideration the multisine task used by Mobertz et al. (2018) and Khoshnewiszadeh and Pool (2021), it can be the case that it would be much harder for the subjects to keep a constant neuromuscular setting due to the large excursions on the screen and to the task duration. When discussing the task duration, it can be expected that the subjects would find it difficult to keep their neuromuscular setting in the stiffer condition for large periods of time. In this way, a shorter experiment duration is suggested with the tradeoff that the BDFT dynamics at the low frequencies would not be able to be computed anymore. A practical aspect that was pointed out is the possibility of the subject to *”anchor”/”rest”* his/her finger on the required point on the screen, making the difference in BDFT between the two neuromuscular settings indistinguishable. While the previously-mentioned aspect can be considered valid, it is expected that during the trial, the subjects will wear an anti-friction glove, and to further reduce the confounding factors, the distance from the user to the screen can be adapted (to be interpreted as *increased*), in order to decrease the probability of the user to *”rest”* on the screen surface.

The first experiment will consist of two parts: a model training part and a validation part. In the first part, each participant will be required to keep one particular neuromuscular setting while continuously interacting with the screen for the whole trial duration. In the second part, the constructed models will be validated, where the neuromuscular setting will be altered throughout one trial (in order to investigate the robustness and the *”adaptability”* of the approach).

Measured variables

The first variable that will be measured is the signal observed at the screen level. As previously mentioned, the subjects will be required to keep their fingers on a discrete point, but due to the fact that disturbances will be induced, a variation of the signal around the target point will be expected.

The second variable that will be measured covers the X, Y, and Z positions (as defined by a simulator/camera fixed coordinate system) of the triangulated landmarks of the shoulder, elbow, and wrist.

Data analysis approach

The data analysis approach is closely linked to the *Experiment and task description*. The data from the first part of the experiment (model training) will be used to train machine-learning models for distinguishing between a stiff and compliant neuromuscular setting. The previously-mentioned machine learning models will be trained on the X, Y, and Z positions of the landmarks (providing that significant differences in the landmark positions will be observed between the two neuromuscular settings) and on the signals observed at the screen level for each subject. Moreover, by analyzing the signal observed at the screen level, and its correlation with the disturbance signal, two models/transfer functions for the biodynamic feedthrough (per subject) will be elaborated, one for the stiff setting and one for the compliant one (again, providing that significant differences will be observed).

The data from the second part will be used to validate the constructed models and to evaluate the signal-cancellation performance of the constructed transfer functions. Ideally, if the user would keep his/her neuromuscular setting very stiff during the validation trials, the models constructed for categorizing the landmarks and the signal observed on the screen would indicate a correct category. Ideally, the adaptiveness of the approach would be from the fact that if the correct category would be identified, the correct model/transfer function will be used for signal-based cancellation (the transfer function corresponding to the current neuromuscular setting). The first experiment will answer the first two research questions.

7.2.2. Discrete pointing task experiment

Goal of the experiment

The goal of the second experiment is to better understand the insights of how the upper limb is affected by the induced disturbances while it is pointing toward a discrete task on a touchscreen and while inducing different settings of the neuromuscular system.

Experiment and task description

The details regarding the Research Simulator, disturbance signal, and disturbance direction are similar to the ones presented in the previous experiment.

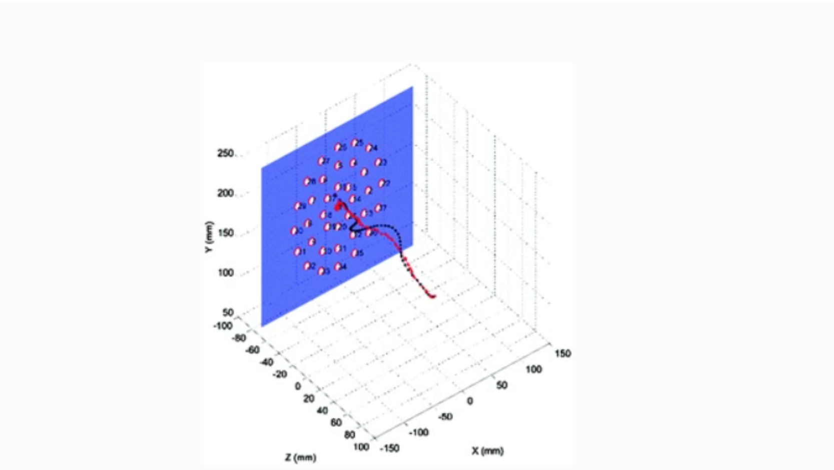


Figure 7.3: The intuition behind the task of the second experiment (Ahmad et al., 2018).

Figure 7.3 presents the intuition behind the task of the second experiment. Unlike the task of the first experiment, in this case, one dot appearing on the screen at a particular time would need to be pressed by the subject while keeping one of the two desired neuromuscular settings (stiff or compliant). The main idea behind the second experiment would be to investigate whether it is possible to *cancel* the contribution from the disturbance but this time of the triangulated finger landmark/trajectory (since the finger is interacting with the touchscreen dynamically) while moving toward the screen and interacting with it.

The second experiment will consist of two parts: the identification part and the validation part. In the first part of the experiment, the subject will be asked to keep his/her hand hanging *"in the air"* (static) while being disturbed and while keeping one of the two neuromuscular settings. The second part would consist of the actual discrete-pointing task where the user is required to point successively at one of the predefined locations on the touchscreen while keeping one of the two required neuromuscular settings.

Measured variables

The first variable that will be measured is the signal observed at the screen level. As previously mentioned, in this scenario, discrete tasks will be tested.

The second variable that will be measured covers the X, Y, and Z positions (as defined by a simulator/camera fixed coordinate system) of the triangulated landmarks of the finger.

Data analysis approach

When focussing on the first part of the experiment, the triangulated position of the finger landmark will be used in order to construct a model (to be interpreted as a transfer function) describing the transfer of accelerations to the finger position (A2P) with respect to a chosen coordinate system for each subject. The previously-mentioned aspect is linked with the hypothesis that the disturbed finger position acts linearly to the induced disturbances. It needs to be mentioned that the transfer function will be created in the direction that *"sees"* the largest contribution from the disturbance.

The second part of the experiment will use the intuition from the first part, and the hypothesis that the frequency content of the *"reach"* towards the screen will be different from the frequency content of the disturbance. In other words, it can be considered that the reaching task consists of a low-frequency movement (the reaching) in combination with movement correlated with the reaching action of the subject. In this way, it will be investigated whether by using the identified model, it is possible to cancel the disturbance contribution to the reaching task, and whether it is possible to correct the location of the screen the user pointed at. The second experiment will answer the last two research questions.

7.3. Hypotheses

- **H1:** The BDFT for the dragging task is expected to be more pronounced for the compliant neuromuscular setting compared to the stiff neuromuscular setting.
- **H2:** The triangulated landmarks for the dragging task are expected to vary more and to have a better attenuation of the higher frequencies in the compliant neuromuscular setting when compared to the stiff neuromuscular setting.
- **H3**: It is expected that the differences in BDFT or in limb position signals between subjects are less pronounced than the differences in BDFT when altering the neuromuscular settings.
- **H4:** It is expected that when the arm *"is left hanging"*, it will respond linearly to the induced disturbance, facilitating a transfer-function definition.
- **H5:** It is expected that the behavior of the upper limb landmarks will be different when being in contact with the screen compared to when the limb is *"left hanging"*.
- **H6**: It is expected that a positive cancellation performance of the signal in the pointing task will be obtained when using the model identified in a static limb scenario (describing the effects of the disturbance on the finger position) for consistent neuromuscular settings.

Reasoning:

- **H1 and H2:** As presented in the simulations of Chapter 5, when varying the stiffness of the multibody system, the position signals of the lower stiffness setting varied more (with respect to the shoulder landmark) when compared to the higher stiffness setting. Moreover, it is expected that a more compliant neuromuscular setting would create a weaker connection between the seat and the screen and therefore attenuating more the higher frequency content of the disturbance signal when compared to the stiffer neuromuscular setting.
- **H3:** The presented literature linked to stick-based operations has found a larger influence of the neuromuscular settings when analyzing BDFT in comparison to the anthropometrics of the subject.

- **H4:** It is expected that the movement of the arm will be correlated with the disturbance signal since the only forces that are acting on the arm are the inertial ones (except the forces that are necessary to keep the arm hanging against the gravity).
- **H5:** It is expected that the interaction forces between the finger and the screen will affect the whole arm since it is possible that the user will use the screen as a pivot point, minimizing the effects of the disturbances.
- **H6:** It is expected that while approaching the screen, the finger of the subject will be affected in the same (linear) way by the disturbances as it is when it is not moving and therefore making feasible the signal cancellation. In other words, it can be considered that the reaching task consists of a low-frequency movement (the reaching) in combination with movement correlated with the reaching action of the subject.

References

Abbott, K. H. (2017). Human factors engineering and flight deck design. *Digital avionics handbook*, 302–328. https://doi.org/10.1201/b17545

Ahmad, B. I., Langdon, P. M., & Godsill, S. J. (2018). Stabilising touch interactions in cockpits, aerospace, and vibrating environments. *Universal Access in Human-Computer Interaction. Methods, Technologies, and Users: 12th International Conference, UAHCI 2018, Held as Part of HCI International 2018, Las Vegas, NV, USA, July 15-20, 2018, Proceedings, Part I 12*, 133–145.

Airbus [Accessed: 2023-05-04]. (2019). https://www.airbus.com/en/newsroom/press-releases/2019-12-airbus-begins-deliveries-of-first-a350-xwbs-with-touchscreen

Alapetite, A., Møllenbach, E., Stockmarr, A., Minakata, K., et al. (2018). A rollercoaster to model touch interactions during turbulence. *Advances in Human-Computer Interaction*, *13*, 22–39.

Albinsson, P.-A., & Zhai, S. (2003). High precision touch screen interaction. *Conference on Human Factors in Computing Systems - Proceedings*, 105–112. https://doi.org/10.1145/642611.642631

Allen, R. W., Jex, H. R., & Magdaleno, R. E. (1973). *Manual control performance and dynamic response during sinusoidal vibration* (tech. rep.). Systems Technology Inc Hawth. https://apps.dtic.mil/sti/citations/AD0773844

Arac, A., Zhao, P., Dobkin, B. H., Carmichael, S. T., & Golshani, P. (2019). Deepbehavior: A deep learning toolbox for automated analysis of animal and human behavior imaging data. *Frontiers in systems neuroscience*, *13*, 20.

ARINC661. (2016). *Cockpit Display System Interfaces to User Systems ARINC Specification 661* (tech. rep.). AEEC. https://www.sae.org/standards/content/arinc661-6/

Avsar, H., Fischer, J. E., & Rodden, T. (2016a). Future flight decks: Impact of+ gz on touchscreen usability. *Proceedings of the International Conference on Human-Computer Interaction in Aerospace*, 1–8. https://doi.org/10.1145/2950112.2964592

Avsar, H., Fischer, J. E., & Rodden, T. (2016b). Mixed method approach in designing flight decks with touch screens: A framework. *2016 IEEE/AIAA 35th Digital Avionics Systems Conference (DASC)*, 1–10. https://doi.org/10.1109/DASC.2016.7778066

Baldi, T. L., Farina, F., Garulli, A., Giannitrapani, A., & Prattichizzo, D. (2019). Upper body pose estimation using wearable inertial sensors and multiplicative kalman filter. *IEEE Sensors Journal*, *20*(1), 492–500.

Boeing [Accessed: 2023-05-04]. (2016). https://www.boeing.com/777x/reveal/touchscreens-come-to-777x-flight-deck/

Bradski, G. (2000). The opencv library. *Dr. Dobb's Journal: Software Tools for the Professional Programmer*, *25*(11), 120–123.

Brooke, J., et al. (1996). Sus-a quick and dirty usability scale. *Usability evaluation in industry*, *189*(194), 4–7.

Buckley, C., Alcock, L., McArdle, R., Rehman, R. Z. U., Del Din, S., Mazzà, C., Yarnall, A. J., & Rochester, L. (2019). The role of movement analysis in diagnosing and monitoring neurodegenerative conditions: Insights from gait and postural control. *Brain sciences*, *9*(2), 34.

Cahill-Rowley, K., & Rose, J. (2017). Temporal–spatial reach parameters derived from inertial sensors: Comparison to 3d marker-based motion capture. *Journal of biomechanics*, *52*, 11–16.

Camera Calibration and 3D reconstruction [Accessed: 2023-06-01]. (2023). https://docs.opencv.org/2.4/modules/calib3d/doc/camera_calibration_and_3d_reconstruction.html

Cao, Z., Simon, T., Wei, S.-E., & Sheikh, Y. (2017). Realtime multi-person 2d pose estimation using part affinity fields. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 7291–7299.

Chen, Y., Tian, Y., & He, M. (2020). Monocular human pose estimation: A survey of deep learning-based methods. *Computer Vision and Image Understanding*, *192*, 102897.

Cockburn, A., Gutwin, C., Palanque, P., Deleris, Y., Trask, C., Coveney, A., Yung, M., & MacLean, K. (2017). Turbulent touch: Touchscreen input for cockpit flight displays. *International Conference for Human-Computer Interaction (CHI 2017)*, 6742–6753. https://doi.org/10.1145/3025453.3025584

Cockburn, A., Masson, D., Gutwin, C., Palanque, P., Goguey, A., Yung, M., Gris, C., & Trask, C. (2019). Design and evaluation of braced touch for touchscreen input stabilisation. *International Journal of Human-Computer Studies*, *122*, 21–37. https://doi.org/10.1016/j.ijhcs.2018.08.005

Computer Vision Toolbox Camera Calibration [Accessed: 2023-05-23]. (2023). https://nl.mathworks.com/help/vision/camera-calibration.html

Coutts, L. V., Plant, K. L., Smith, M., Bolton, L., Parnell, K. J., Arnold, J., & Stanton, N. A. (2019). Future technology on the flight deck: Assessing the use of touchscreens in vibration environments. *Ergonomics*, *62*(2), 286–304.

Cutti, A. G., Giovanardi, A., Rocchi, L., Davalli, A., & Sacchetti, R. (2008). Ambulatory measurement of shoulder and elbow kinematics through inertial and magnetic sensors. *Medical & biological engineering & computing*, *46*, 169–178.

D'Antonio, E., Taborri, J., Mileti, I., Rossi, S., & Patané, F. (2021). Validation of a 3d markerless system for gait analysis based on openpose and two rgb webcams. *IEEE Sensors Journal*, *21*(15), 17064–17075.

Degani, A., Palmer, E. A., & Bauersfeld, K. G. (1992). “Soft” controls for hard displays: Still a challenge. *Proceedings of the Human Factors Society Annual Meeting*, *36*(1), 52–56. https://doi.org/10.1177/154193129203600114

Desmarais, Y., Mottet, D., Slangen, P., & Montesinos, P. (2021). A review of 3d human pose estimation algorithms for markerless motion capture. *Computer Vision and Image Understanding*, *212*, 103275.

Dodd, S., Lancaster, J., Miranda, A., Grothe, S., DeMers, B., & Rogers, B. (2014). Touch screens on the flight deck: The impact of touch target size, spacing, touch technology and turbulence on pilot performance. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *58*(1), 6–10. https://doi.org/10.1177/1541931214581002

Drazan, J. F., Phillips, W. T., Seethapathi, N., Hullfish, T. J., & Baxter, J. R. (2021). Moving outside the lab: Markerless motion capture accurately quantifies sagittal plane kinematics during the vertical jump. *Journal of biomechanics*, *125*, 110547.

Eichinger, A., & Kellerer, J. (2014). Between laboratory and simulator: A cognitive approach to evaluating cockpit interfaces by manipulating informatory context. *Cognition, Technology & Work*, *16*, 417–427.

Eisemann, E. (2020). Visual Data Processing lecture notes.

Fang, H.-S., Xie, S., Tai, Y.-W., & Lu, C. (2017). Rmpe: Regional multi-person pose estimation. *Proceedings of the IEEE international conference on computer vision*, 2334–2343.

Figueiredo, P. R. P., Silva, P. L., Avelar, B. S., da Fonseca, S. T., Bootsma, R. J., & Mancini, M. C. (2015). Upper limb performance and the structuring of joint movement in teenagers with cerebral palsy: The reciprocal role of task demands and action capabilities. *Experimental Brain Research*, *233*, 1155–1164.

Fiorentino, N. M., Atkins, P. R., Kutschke, M. J., Goebel, J. M., Foreman, K. B., & Anderson, A. E. (2017). Soft tissue artifact causes significant errors in the calculation of joint angles and range of motion at the hip. *Gait & posture*, *55*, 184–190.

Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of experimental psychology*, *47*(6), 381. https://doi.org/10.1037/h0055392

Fletcher, M., & Huffman, D. (2010). Panoramic cockpit displays for tactical military cockpits. *Three-Dimensional Imaging, Visualization, and Display 2010 and Display Technologies and Applications for Defense, Security, and Avionics IV*, *7690*, 326–332. https://doi.org/10.1117/12.853032

Harris, D. (2011). *Human performance on the flight deck*. Ashgate Publishing, Ltd. https://doi.org/10.1201/9781315252988

Hedge, A., Morimoto, S., & Mccrobie, D. (1999). Effects of keyboard tray geometry on upper body posture and comfort. *Ergonomics*, *42*(10), 1333–1349. https://doi.org/10.1080/001401399184983

Hidalgo, G., Raaj, Y., Idrees, H., Xiang, D., Joo, H., Simon, T., & Sheikh, Y. (2019). Single-network whole-body pose estimation. *ICCV*.

Hill, A. V. (1938). The heat of shortening and the dynamic constants of muscle. *Proceedings of the Royal Society of London. Series B-Biological Sciences*, *126*(843), 136–195.

Höglund, G., Grip, H., & Öhberg, F. (2021). The importance of inertial measurement unit placement in assessing upper limb motion. *Medical Engineering & Physics*, *92*, 1–9.

Höhne, G. (2000). Computer aided development of biomechanical pilot models. *Aerospace science and technology*, *4*(1), 57–69.

Huang, H., Liu, J., Liu, S., Jin, P., Wu, T., & Zhang, T. (2020). Error analysis of a stereo-vision-based tube measurement system. *Measurement*, *157*, 107659.

Idan, M., & Merhav, S. (1990). Effects of biodynamic coupling on the human operator model. *Journal of Guidance, Control, and Dynamics*, *13*(4), 630–637.

ISO 9241-9 (tech. rep.). (2000). https://www.iso.org/standard/30030.html

Jack, R. J., Oliver, M., & Hayward, G. (2013). Using a high resolution motion capture system to determine 6-dof whole-body vibration accelerations. *Journal of Low Frequency Noise, Vibration and Active Control*, *32*(4), 273–284.

Jex, H. R., & Magdaleno, R. E. (1978). Biomechanical models for vibration feedthrough to hands and head for a semisupine pilot. *Aviation, space, and environmental medicine*, *49*(1 Pt. 2), 304–316.

Jones, D. (1990). Three input concepts crew interaction presented electronic for flight with information on a large-screen cockpit display (tech. report). https://ntrs.nasa.gov/citations/19900009078

Kane, T. R., & Levinson, D. A. (1985). *Dynamics, theory and applications*. McGraw Hill.

Khoshnewiszadeh, A., & Pool, D. M. (2021). Mitigation of biodynamic feedthrough for touch-screens on the flight deck. *International Journal of Human–Computer Interaction*, *37*(7), 680–692.

Kitazaki, S., & Griffin, M. J. (1997). A modal analysis of whole-body vertical vibration, using a finite element model of the human body. *Journal of Sound and Vibration*, *200*(1), 83–103.

Koehler, R. (1997). A new pilot model for roll ratcheting analysis. *DLR IB 11-97,32*.

Kosourikhina, V., Kavanagh, D., Richardson, M. J., & Kaplan, D. M. (2022). Validation of deep learning-based markerless 3d pose estimation. *Plos one*, *17*(10), e0276258.

Lasschuit, J., Lam, M., Mulder, M., van Paassen, R., & Abbink, D. (2008). Measuring and modeling neuromuscular system dynamics for haptic interface design. *AIAA modeling and simulation technologies conference and exhibit*, 6543.

Lee, Y.-L. (2010). Comparison of the conventional point-based and a proposed finger probe-based touch screen interaction techniques in a target selection task. *International Journal of Industrial Ergonomics*, *40*(6), 655–662. https://doi.org/10.1016/j.ergon.2010.09.001

Levison, W. H. (1976). *Biomechanical response and manual tracking performance in sinusoidal, sum-of-sines, and random vibration environments.* (tech. rep.). Bolt Beranek and Newman Inc., Cambridge, MA, United States.

Lewis, C. (1981). A comparison of the effects of sinusoidal and random whole-body vibration on manual tracking performance. *U.K. Informal Group Meeting on Human Response to Vibration, Heriot- Watt University, Edinburgh.*, [Unpublished manuscript].

Ligorio, G., Zanotto, D., Sabatini, A., & Agrawal, S. (2017). A novel functional calibration method for real-time elbow joint angles estimation with magnetic-inertial sensors. *Journal of biomechanics*, *54*, 106–110.

Ligorio, G., Bergamini, E., Pasciuto, I., Vannozzi, G., Cappozzo, A., & Sabatini, A. M. (2016). Assessing the performance of sensor fusion methods: Application to magnetic-inertial-based human body tracking. *Sensors*, *16*(2), 153.

Ligorio, G., Bergamini, E., Truppa, L., Guaitolini, M., Raggi, M., Mannini, A., Sabatini, A. M., Vannozzi, G., & Garofalo, P. (2020). A wearable magnetometer-free motion capture system: Innovative solutions for real-world applications. *IEEE Sensors Journal*, *20*(15), 8844–8857.

Lim, W. (2019). Markerless motion capture system based on webcams using openpose. *nternational Journal of Human Movement and Sports Sciences*, *10(5)*, 900–905.

Lin, C. J., Liu, C. N., Chao, C. J., & Chen, H. J. (2010). The performance of computer input devices in a vibration environment. *Ergonomics*, *53*(4), 478–490. https://doi.org/10.1080/00140130903528186

Liu, W., Bao, Q., Sun, Y., & Mei, T. (2022). Recent advances of monocular 2d and 3d human pose estimation: A deep learning perspective. *ACM Computing Surveys*, *55*(4), 1–41.

Masarati, P., Quaranta, G., & Zanoni, A. (2013a). Dependence of helicopter pilots' biodynamic feedthrough on upper limbs' muscular activation patterns. *Proceedings of the Institution of Mechanical Engineers, Part K: Journal of Multi-body Dynamics*, *227*(4), 344–362.

Masarati, P., Quaranta, G., & Jump, M. (2013b). Experimental and numerical helicopter pilot characterization for aeroelastic rotorcraft–pilot coupling analysis. *Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering*, *227*(1), 125–141.

Mathis, A., Mamidanna, P., Cury, K. M., Abe, T., Murthy, V. N., Mathis, M. W., & Bethge, M. (2018). Deeplabcut: Markerless pose estimation of user-defined body parts with deep learning. *Nature neuroscience*, *21*(9), 1281–1289.

Mattaboni, M., Fumagalli, A., Quaranta, G., & Masarati, P. (2009). Identification of the biomechanical behavior of a rotorcraft pilot arm. *XX AIDAA Congress, Milano, Italy*, 1–23.

Mattaboni, M., Fumagalli, A., Jump, M., Masarati, P., Quaranta, G., et al. (2008). Biomechanical pilot properties identification by inverse kinematics/inverse dynamics multibody analysis. *ICAS-International Council for the Aeronautical Sciences*, 14–19.

Mayo, J. R. (1989). The involuntary participation of a human pilot in a helicopter collective control loop. *15th European Rotorcraft Forum*, 81.1–81.12.

McLeod, R., & Griffin, M. (1989). Review of the effects of translational whole-body vibration on continuous manual control performance. *Journal of Sound and Vibration*, *133*(1), 55–115. https://doi.org/10.1016/0022-460X(89)90985-1

McRuer, D. T., Graham, D., & Krendel, E. S. (1967). Manual control of single-loop systems: Part i. *Journal of the Franklin Institute*, *283*(1), 1–29.

Miller, N. R., Shapiro, R., & McLaughlin, T. M. (1980). A technique for obtaining spatial kinematic parameters of segments of biomechanical systems from cinematographic data. *Journal of Biomechanics*, *13*(7), 535–547.

Mobertz, X., Pool, D. M., van Paassen, M., & Mulder, M. (2018). A cybernetic analysis of biodynamic effects in touchscreen operation in turbulence. *2018 AIAA Modeling and Simulation Technologies Conference*, 115–139.

Moore, J. K. (2022). Learn Multibody Dynamics. https://moorepants.github.io/learn-multibody-dynamics/

Moro, M., Marchesi, G., Odone, F., & Casadio, M. (2020). Markerless gait analysis in stroke survivors based on computer vision and deep learning: A pilot study. *Proceedings of the 35th annual ACM symposium on applied computing*, 2097–2104.

Mugge, W. (2022). *Neuromechanics and Motor Control lecture notes*, Delft University of Technology.

Nakano, N., Sakura, T., Ueda, K., Omura, L., Kimura, A., Iino, Y., Fukashiro, S., & Yoshioka, S. (2020). Evaluation of 3d markerless motion capture accuracy using openpose with multiple video cameras. *Frontiers in sports and active living*, *2*, 50.

NASA. (1986). Nasa task load index (nasa-tlx), version 1.0: Paper and pencil package.

Needham, L., Evans, M., Cosker, D. P., Wade, L., McGuigan, P. M., Bilzon, J. L., & Colyer, S. L. (2021). The accuracy of several pose estimation methods for 3d joint centre localisation. *Scientific reports*, *11*(1), 20673.

Norman, D. (1989). *The problem of automation: Inappropriate feedback and interaction not over automation (ics report 8904)* (tech. rep.). https://doi.org/10.1098/rstb.1990.0101

Norman, K. L. (1991). *The psychology of menu selection: Designing cognitive control at the human/computer interface*. Intellect Books. https://doi.org/10.5555/574380

Öhberg, F., Bäcklund, T., Sundström, N., & Grip, H. (2019). Portable sensors add reliable kinematic measures to the assessment of upper extremity function. *Sensors*, *19*(5), 1241.

Palermo, E., Rossi, S., Marini, F., Patanè, F., & Cappa, P. (2014). Experimental evaluation of accuracy and repeatability of a novel body-to-sensor calibration procedure for inertial sensor-based gait analysis. *Measurement*, *52*, 145–155.

Parasuraman, R., & Manzey, D. H. (2010). Complacency and bias in human use of automation: An attentional integration. *Human factors*, *52*(3), 381–410.

Pennestri, E., Stefanelli, R., Valentini, P., & Vita, L. (2007). Virtual musculo-skeletal model for the biomechanical analysis of the upper limb. *Journal of biomechanics*, *40*(6), 1350–1361.

Rasmussen, J. (1983). Skills, rules, and knowledge; signals, signs, and symbols, and other distinctions in human performance models. *IEEE Transactions on Systems, Man, and Cybernetics*, *SMC-13*(3), 257–266. https://doi.org/10.1109/TSMC.1983.6313160

Rawal, A., Chehata, A., Horberry, T., Shumack, M., Chen, C., & Bonato, L. (2018). Defining the upper extremity range of motion for safe automobile driving. *Clinical biomechanics*, *54*, 78–85.

Rempe, D., Birdal, T., Hertzmann, A., Yang, J., Sridhar, S., & Guibas, L. J. (2021). Humor: 3d human motion model for robust pose estimation. *Proceedings of the IEEE/CVF international conference on computer vision*, 11488–11499.

Roetenberg, D., Slycke, P. J., & Veltink, P. H. (2007). Ambulatory position and orientation tracking fusing magnetic and inertial sensing. *IEEE Transactions on Biomedical Engineering*, *54*(5), 883–890.

Roetenberg, D., Luinge, H., Slycke, P., et al. (2009). Xsens mvn: Full 6dof human motion tracking using miniature inertial sensors. *Xsens Motion Technologies BV, Tech. Rep*, *1*, 1–7.

Rogers, W. A., Fisk, A. D., McLaughlin, A. C., & Pak, R. (2005). Touch a screen or turn a knob: Choosing the best device for the job. *Human factors*, *47*(2), 271–288. https://doi.org/10.1518/0018720054679452

Ross, S. (2020). Analytical Dynamics - Lagrangian and 3D Rigid Body Dynamics Lecture Notes. https://drive.google.com/drive/folders/1-8ZGvfZTw7TxjSQmG3-JFShnvt6mCURk

Rouwhorst, W., Verhoeven, R., Suijkerbuijk, M., Bos, T., Maij, A., Vermaat, M., & Arents, R. (2017). Use of touch screen display applications for aircraft flight control. *2017 IEEE/AIAA 36th Digital Avionics Systems Conference (DASC)*, 1–10. https://doi.org/10.1109/DASC.2017.8102060

Rudisill, M. (2000). *Crew/automation interaction in space transportation systems: Lessons learned from the glass cockpit* (tech. rep.). NASA Langley Research Center Hampton, VA, United States. https://ntrs.nasa.gov/citations/20040086769

Schmidt, R., Disselhorst-Klug, C., Silny, J., & Rau, G. (1999). A marker-based measurement procedure for unconstrained wrist and elbow motions. *Journal of biomechanics*, *32*(6), 615–621.

Schwab, A. L. (1998). Multibody Dynamics B Lecture Notes. http://bicycle.tudelft.nl/schwab/Publications/LinSch02.pdf

Slabaugh, G., Schafer, R., & Livingston, M. (2001). Optimal ray intersection for computing 3d points from n-view correspondences. *Deliverable Report*, 1–11.

Stachniss, C. (2021). Photogrammetric Computer Vision lecture notes. http://www.ipb.uni-bonn.de/html/teaching/photo12-2021/2021-pho1-22-Zhang-calibration.pptx.pdf

Stanton, N. A., Harvey, C., Plant, K. L., & Bolton, L. (2013). To twist, roll, stroke or poke? a study of input devices for menu navigation in the cockpit. *Ergonomics*, *56*(4), 590–611. https://doi.org/10.1080/00140139.2012.751458

Stenum, J., Rossi, C., & Roemmich, R. T. (2021). Two-dimensional video-based analysis of human gait using pose estimation. *PLoS computational biology*, *17*(4), e1008935.

Sweet, W. (1995). The glass cockpit [flight deck automation]. *IEEE Spectrum*, *32*(9), 30–38. https://doi.org/10.1109/6.406460

Theriault, D. H., Fuller, N. W., Jackson, B. E., Bluhm, E., Evangelista, D., Wu, Z., Betke, M., & Hedrick, T. L. (2014). A protocol and calibration method for accurate multi-camera field videography. *Journal of Experimental Biology*, *217*(11), 1843–1848.

Thomas, P. R. (2018). Performance, characteristics, and error rates of cursor control devices for aircraft cockpit interaction. *International Journal of Human-Computer Studies*, *109*, 41–53.

Truppa, L., Bergamini, E., Garofalo, P., Vannozzi, G., Sabatini, A., & Mannini, A. (2022). Magnetic-free quaternion-based robust unscented kalman filter for upper limb kinematic analysis. *IEEE Sensors Journal*.

Truppa, L., Bergamini, E., Garofalo, P., Costantini, M., Fiorentino, C., Vannozzi, G., Sabatini, A. M., & Mannini, A. (2021). An innovative sensor fusion algorithm for motion tracking with on-line bias compensation: Application to joint angles estimation in yoga. *IEEE Sensors Journal*, *21*(19), 21285–21294.

van der Kruk, E., Schwab, A., van der Helm, F., & Veeger, H. (2018). Getting in shape: Reconstructing three-dimensional long-track speed skating kinematics by comparing several body pose reconstruction techniques. *Journal of biomechanics*, *69*, 103–112.

Van Paassen, M. M. (1994). Biophysics in aircraft control: A model of the neuromuscular system of the pilot's arm. *PhD thesis: Delft University of Technology*.

Van Paassen, M. M. (2021). Neuromuscular system lecture notes. *AE4316:Aerospace Human-Machine Systems Supplemental Material*.

van Zon, N. C., Borst, C., Pool, D. M., & van Paassen, M. M. (2020). Touchscreens for aircraft navigation tasks: Comparing accuracy and throughput of three flight deck interfaces using fitts' law. *Human factors*, *62*(6), 897–908. https://doi.org/10.1177/0018720819862146

Venrooij, J., Mulder, M., van Paassen, M. M., Mulder, M., & Abbink, D. A. (2009). Relating biodynamic feedthrough to neuromuscular admittance. *2009 IEEE International Conference on Systems, Man and Cybernetics*, 1668–1673.

Venrooij, J., Abbink, D. A., Mulder, M., van Paassen, M. M., & Mulder, M. (2010a). Biodynamic feedthrough is task dependent. *2010 IEEE International Conference on Systems, Man and Cybernetics*, 2571–2578.

Venrooij, J., Mulder, M., Van Paassen, M. M., Mulder, M., & Abbink, D. A. (2010b). A review of biodynamic feedthrough mitigation techniques. *IFAC Proceedings Volumes*, *43*(13), 316–321.

Venrooij, J., Mulder, M., Van Paassen, M. M., Abbink, D. A., Bülthoff, H. H., & Mulder, M. (2011a). Cancelling biodynamic feedthrough requires a subject and task dependent approach. *2011 IEEE International Conference on Systems, Man, and Cybernetics*, 1670–1675.

Venrooij, J., Abbink, D. A., Mulder, M., van Paassen, M. M., & Mulder, M. (2011b). A method to measure the relationship between biodynamic feedthrough and neuromuscular admittance. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, *41*(4), 1158–1169.

Venrooij, J., Pavel, M. D., Mulder, M., Van Der Helm, F. C., & Bülthoff, H. H. (2013). A practical biodynamic feedthrough model for helicopters. *CEAS Aeronautical Journal*, *4*, 421–432.

Venrooij, J., Mulder, M., Mulder, M., Abbink, D. A., Van Paassen, M. M., Van Der Helm, F. C., & Bülthoff, H. H. (2016a). Admittance-adaptive model-based approach to mitigate biodynamic feedthrough. *IEEE Transactions on Cybernetics*, *47*(12), 4169–4181.

Venrooij, J., Olivari, M., & Bülthoff, H. H. (2016b). Biodynamic feedthrough: Current status and open issues. *IFAC-PapersOnLine*, *49*(19), 120–125.

Von Gierke, H. E. (1965). Biodynamic response of the human body. *Bioastronautics and the Exploration of Space: Proceedings*, *385*.

Wade, L., Needham, L., McGuigan, P., & Bilzon, J. (2022). Applications and limitations of current markerless motion capture methods for clinical gait biomechanics. *PeerJ*, *10*, e12995.

Wang, L., Cao, Q., Chang, J., & Zhao, C. (2015). The effect of touch-key size and shape on the usability of flight deck mcdu. *ACHI 2015*, 234–238.

Wang, X., & Trasbot, J. (2011). Effects of target location, stature and hand grip type on in-vehicle reach discomfort. *Ergonomics*, *54*(5), 466–476. https://doi.org/10.1080/00140139.2011.564312

Watkins, C. B., Nilson, C., Taylor, S., Medin, K. B., Kuljanin, I., & Nguyen, H. B. (2018). Development of touchscreen displays for the gulfstream g500 and g600 symmetry™ flight deck. *2018 IEEE/AIAA 37th Digital Avionics Systems Conference (DASC)*, 1–10. https://doi.org/10.1109/DASC.2018.8569532

Wiener, E. L. (1989). *Human factors of advanced technology (glass cockpit) transport aircraft* (tech. rep.). NASA Ames Research Center Ames, CA, United States. https://ntrs.nasa.gov/citations/19890016609

Winter, D. A. (2009). *Biomechanics and motor control of human movement*. John Wiley & Sons.

Wu, G., Van der Helm, F. C., Veeger, H. D., Makhsous, M., Van Roy, P., Anglin, C., Nagels, J., Karduna, A. R., McQuade, K., Wang, X., et al. (2005). Isb recommendation on definitions of joint coordinate systems of various joints for the reporting of human joint motion—part ii: Shoulder, elbow, wrist and hand. *Journal of biomechanics*, *38*(5), 981–992.

Wynne, R. A., Parnell, K. J., Smith, M. A., Plant, K. L., & Stanton, N. A. (2021). Can't touch this: Hammer time on touchscreen task performance variability under simulated turbulent flight conditions. *International Journal of Human–Computer Interaction*, *37*(7), 666–679.

Xie, Y., Zhou, R., & Qu, J. (2023). Fitts' law on the flight deck: Evaluating touchscreens for aircraft tasks in actual flight scenarios. *Ergonomics*, *66*(4), 506–523.

Yahya, M., Shah, J. A., Kadir, K. A., Yusof, Z. M., Khan, S., & Warsi, A. (2019). Motion capture sensing techniques used in human upper limb motion: A review. *Sensor Review*, *39*(4), 504–511.

Zago, M., Luzzago, M., Marangoni, T., De Cecco, M., Tarabini, M., & Galli, M. (2020). 3d tracking of human motion using visual skeletonization and stereoscopic vision. *Frontiers in bioengineering and biotechnology*, *8*, 181.

Zanoni, A., Cocco, A., & Masarati, P. (2020). Multibody dynamics analysis of the human upper body for rotorcraft–pilot interaction. *Nonlinear dynamics*, *102*, 1517–1539.

Zanoni, A., Zago, M., Paolini, R., Quaranta, G., Galli, M., & Masarati, P. (2021). On task dependence of helicopter pilot biodynamic feedthrough and neuromuscular admittance: An experimental and numerical study. *IEEE Transactions on Human-Machine Systems*, *51*(5), 421–431.

Zhang, Z. (2000). A flexible new technique for camera calibration. *IEEE Transactions on pattern analysis and machine intelligence*, *22*(11), 1330–1334.

Zhou, H., & Hu, H. (2005). Inertial motion tracking of human arm movements in stroke rehabilitation. *IEEE International Conference Mechatronics and Automation, 2005*, *3*, 1306–1311.

Zhou, H., & Hu, H. (2007). Upper limb motion estimation from inertial measurements. *International Journal of Information Technology*, *13*(1), 1–14.

Part III

MSc Thesis Appendices

A

Experiment consent and briefing forms

The following appendix provides an overview of the two main documents that the subjects of the experiment received before performing it. Before the experiment session, the experimenter carefully explained the contents of the two documents and allowed the subjects to ask questions, and to refuse participation at any moment in time.

Experiment Briefing

"Towards adaptive biodynamic feedthrough modelling for touchscreen applications"

Foreword

Dear participant,

Thank you for being part of this research project! With your help, we can go one step further in the unknown territory of personalization of biodynamic feedthrough effects in touchscreen operation. The goal of this briefing is to give you the information necessary to perform the experiment safely and comfortably and to explain how your data will be used and processed further. If you still have additional questions regarding the experiment procedure or about the confidentiality of your data, please don't hesitate to contact the experimenter (the first name on the last page of this document).

Why are we doing this research?

Touchscreens have become an invaluable part of our everyday lives. From the possibility of rapidly typing messages to interacting with various widgets in the car or on the phone, it is undoubtedly that they have made our lives easier. With this in mind, the same technology is being slowly introduced in cockpits in order to facilitate a more intuitive interaction with the cockpit environment.

Figure 1: Cockpit with touchscreens.

Despite the fact that touchscreens provide a one-to-one mapping between the user input and the activation on the screen, this advantage is challenged by the fact that the outside disturbances of the airplane feed through the human body and negatively affect the performance when operating such a system. In this way, the probability of wrong activations is high, fact which is not acceptable in safety-critical domains such as the aerospace sector.

One possible solution that was proposed by our department is to create models that can quantify the unwanted contribution of the disturbance to the input of the screen. In this way, in ideal conditions, it would be possible to cancel the unwanted inputs on the screen, keeping only the voluntary components. *Previous research focussed on only analyzing the relationship between the disturbance (input) and the signal observed at the screen level (output), and concluded that for a "good" cancellation performance, the models would need to be adapted to the biomechanical and/or neuromuscular characteristics of the user (one example being the stiffness of the arm).*

The current research will try to better understand what happens between the external disturbance and the touchscreen input, by analyzing the biomechanical characteristics of the limbs of the subjects using markerless camera-based pose estimation techniques. In this way, this study will tell us whether is it indeed feasible to use these techniques in order to capture some of the biomechanical characteristics of the human under several scenarios (by varying the stiffness of the arm, and the task), when operating a touchscreen in turbulence.

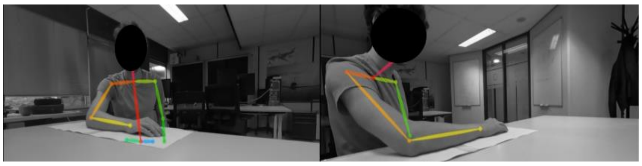


Figure 2: Markerless pose estimation technique.

What will the research look like?

This research will investigate whether is it possible to capture biomechanical information of the arm using camera-based pose estimation techniques while operating a touchscreen in turbulence-mimicking scenarios. In this way, this section will provide a step-by-step guide to how the experiment will take place.

Step 0: SIMONA Research Simulator

In order to successfully investigate the biodynamic characteristics of the arm under turbulent conditions, SIMONA Research Simulator will be used. SIMONA is a moving-base simulator that is able to move in space with 6 degrees of freedom. During the experiment, we will only focus on *the heave direction* (vertically up and down). It needs to be mentioned that in a real-life scenario, the turbulence can be considered unpredictable, but in our case, we will try to mimic turbulence with a signal constructed from a sum of sinusoids, which can help us in the post-processing of the data. Nevertheless, this signal will feel like real moderate turbulence to you!

Figure 3: Outside view of the SIMONA Research Simulator.

Step 1: Finding your way through the simulator

As shown in *Figure 3*, the initial position of the simulator is “down”, being linked to the platform with an aerial bridge. After the initial briefing, you will be kindly asked by the experimenter to enter the simulator *via the aerial bridge* and take your seat.

After entering the simulator, you are asked to take a seat in the rightmost seat (“Seat R” presented in *Figure 4*). Near the seat you will find two cameras, which are used to capture the movement of your arm when operating the touchscreen under turbulence – you can safely ignore them. Moreover, in front of you, you will find the touchscreen on which a series of tasks will be presented (which will be explained shortly).

Initially, the seat will be positioned in the farthest position from the screen. With the help of the lever on the left side of the seat, you will be able to vary the distance between you and the screen. *Please adjust your seat making sure that you are not seated too close to the screen. You should be able to reach the screen with the arm extended, without moving your back forward from the backrest.*

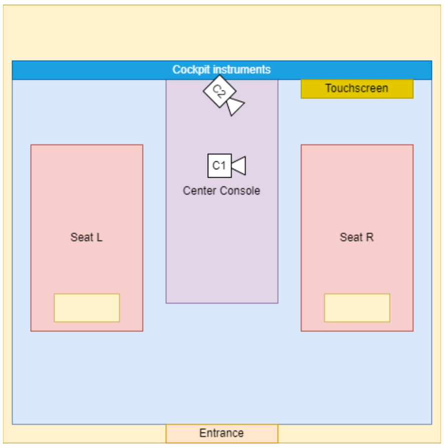


Figure 4: Top view of the simulator.

After the seating position is chosen, you will observe a circular 5-point harness mount. Before continuing, please make sure that all the seatbelts are securely fastened.

Once everything is secured, we will close the door and retract the moving bridge. It is important to keep the seatbelt on for the whole period that the bridge is disconnected. The experimenter will announce when you are safe to take the seatbelt off at the end of the experiment or before the scheduled/unscheduled brakes. Lastly, you will communicate with the experimenter via the ambient microphones which are positioned throughout the cockpit.

Step 2: Task(s) description

The selected tasks have the goal of replicating the most common ways the user is operating a touchscreen.

Step 2.1: Task 1 – Finger kept on the screen in one location

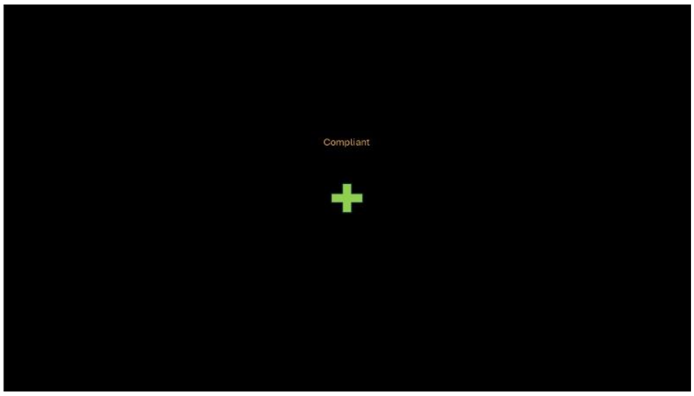

Compliant

Figure 5: Task 1 version 1.

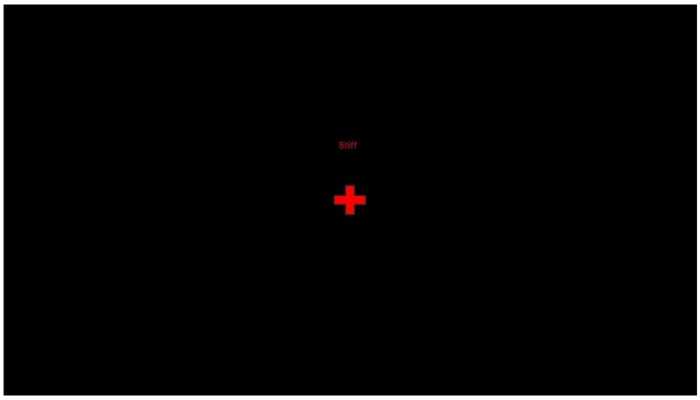

Stiff

Figure 6: Task 1 version 2.

During the first task, the subject is required to keep his/her finger in contact with the screen on the green/red cross. The colour of the crosses is related to the neuromuscular setting (arm stiffness) that the subject is requested to attain *continuously* during the specific trial. For the whole duration of the trial, the subject will continuously see one cross of one colour which is not moving, and he/she is required to keep the arm relaxed for the green cross or stiff for the red cross while in the same time pointing with the finger *continuously* at the dot. With the red cross, the subject is kindly asked to make his/her hand as stiff as comfortably possible (by activating the necessary muscles of the arm).

Important considerations:

1) *In order to ensure that the pose estimation software is able to correctly identify the finger landmarks, the subject is asked to operate the touchscreen with the index finger while keeping his/her palm open and fingers in the sight of the camera (the palm shall be pointing away from the cameras), as described in Figures 7 and 8.*

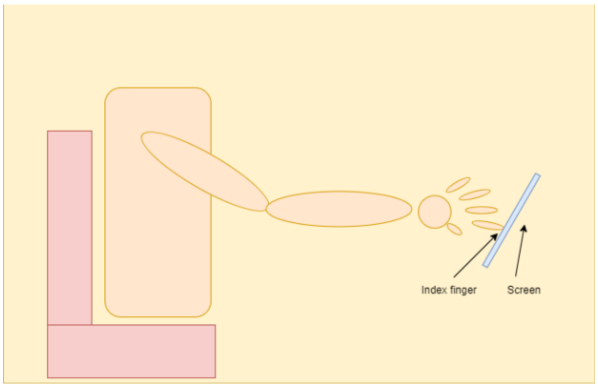


Figure 7: Requested position of the hand version 1 - Task 1.

Figure 8: Requested position of the hand version 2 – Task 1.

2) *Since this task requires that the finger is in contact with the screen continuously, some instructions need to be pointed out in order to avoid potential confounding factors linked to the interaction between the screen and the finger. The subject is asked not to rest his/her finger on the screen (to put a lot of pressure) and to pivot around a potential anchor on the screen. The goal is to let the finger "slide" smoothly along the screen, movement which is dictated by the feedthrough of the accelerations through the hand. Saying this, the experimenter recommends that an ideal approach for this task is to put as little pressure as possible on the screen while ensuring that the finger is in contact with it.*

Step 2.2: Task 2 – Finger on the screen at multiple locations (dragging task)

Unlike the first task, which required the finger to be positioned in one discrete position on the screen, the second task requires the subject to follow a continuous target on the screen, while trying to exhibit the required neuromuscular settings.

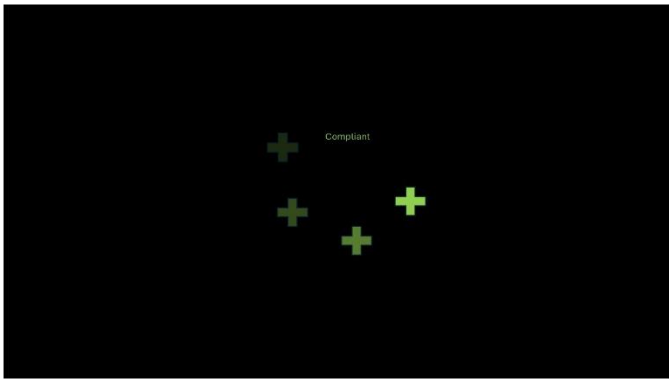


Figure 9: Task 2 version 1.

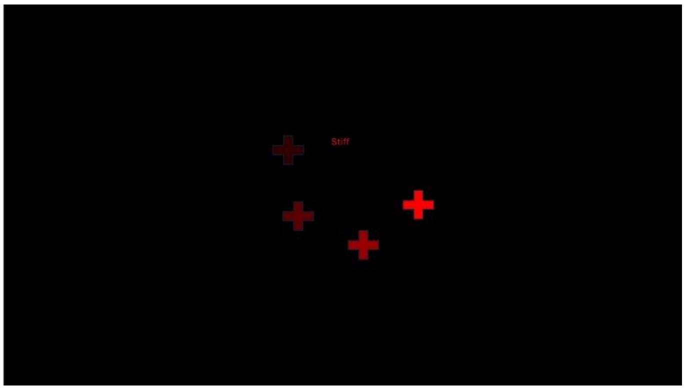


Figure 10: Task 2 version 2.

The green/red cross will move across the screen and shall be followed with the finger by the subject. The subject will only see the fully coloured target cross (i.e. Figure 9 and 10), the purpose of the transparent crosses is to express the movement of the cross along the screen.

Important considerations: The same considerations that were mentioned for the first task apply also for this task.

Step 2.3: Task 3 – Finger kept off the screen at one location

The third task is similar to the first one, with the crucial difference that the subject is asked to keep his/her finger off the screen during the duration of the trial. On the screen, the same two versions of the first task will be presented (target not moving), while the subject will keep the finger at a distance of 4-5 [cm] away from the screen, pointing towards the shown point, but not touching the screen. The goal of this task is to investigate how the response of the arm changes when the hand is not in contact with the screen.

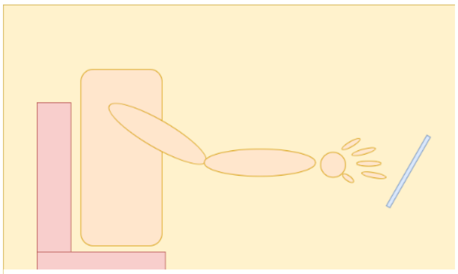


Figure 11: Requested position of the hand - Task 3

Important considerations:

1) *The subject is kindly asked to keep his/her finger pointed at the screen as if he/she is pointing at the imaginary cross. The reason for this is that on average the arm ideally shall stay in one steady position throughout one trial.*

Step 2.3: Task 4 – Simulating a reaching task

The last task which is tested is inspired from the natural motion of the user to reach for the screen (i.e. in discrete selection tasks).

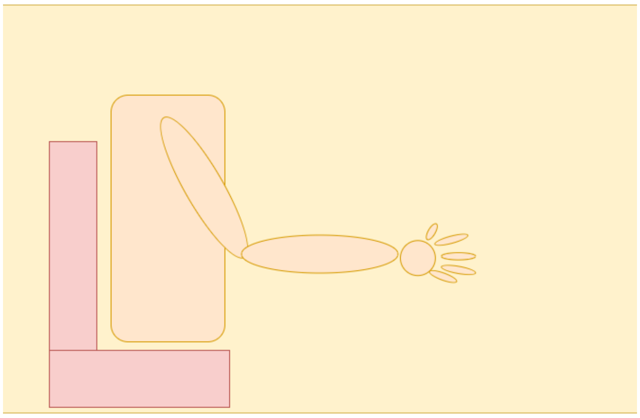


Figure 9: Task 4 position 1.

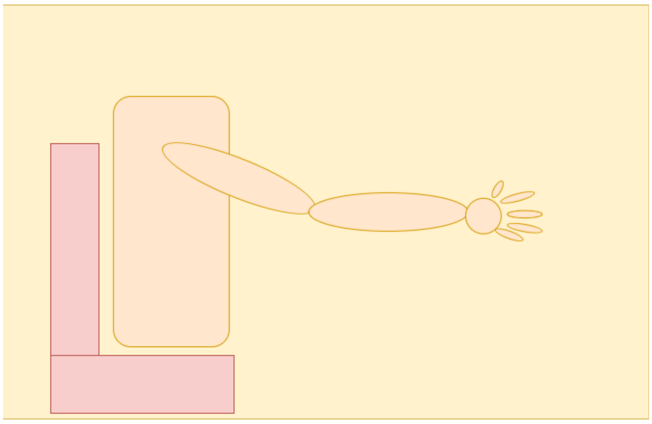


Figure 10: Task 4 position 2.

The task proposes two positions that need to be attained by the subject. Position 1 requires the subject to keep the lower arm close to the torso with the elbow slightly in front of the torso and the hand again open but with the thumb pointing up, and the palm facing the cameras. For the second position, the arm needs to be extended toward the screen (and left hanging), similar to the third task. During this task, the subject shall keep his/her arm *relaxed* (i.e. only one neuromuscular setting).

Task 4 will have three different versions: version 1 which requires the subject to keep position 1 fixed throughout one trial, version 2 which requires the subject to keep position 2 fixed throughout one trial and version 3 which requires the subject to transition between position 1 and position 2. The transitioning trials can be interpreted as keeping position 1 for 2/3 [s], transitioning to position 2, keeping position 2 for 2/3 [s], going back to position 1 and repeating until the end of the trial. The transition is performed as interpreted by the subject, without having specific requirements.

Step 2.4: Task summary

The experiment will consist of the following tasks (presented in an arbitrary order):

Task 1 version 1 – 3 runs

Task 1 version 2 – 3 runs

Task 2 version 1 – 3 runs

Task 2 version 2 – 3 runs

-- Scheduled break 1--

Task 3 version 1 – 3 runs

Task 3 version 2 – 3 runs

Task 4 version 1 – 3 runs

Task 4 version 2 – 3 runs

Task 4 version 3 (transitions) – 3 runs

Expected experiment duration (including brakes): 1 [hr.]

Step 2.5: Experiment progression summary

Once the aerial bridge is disconnected and the simulator is brought to its neutral position, the experimenter will explain to the subject the contents of the next task and will begin with the trials. Before each trial, the experimenter will ask the subject whether he/she is ready and will start with a countdown announcing the start of the trial. At any discomfort, including signs of nausea/motion sickness, the subject is encouraged to announce the experimenter, and the experiment will be stopped. Moreover, at any desired time, the subject can ask for a period of break.

How will your data look like, and how do we make sure that it remains confidential?

Three categories of data will be saved for this experiment.

1) Touchscreen-related data – 2D time traces representing the coordinates of the screen activations.
2) Videos that are being inputted in a pose estimation software (OpenPose) that extracts the pixel values of your *shoulder, elbow, wrist, and* hand landmarks.

   We understand that this represents very sensible information for the subject, and therefore, here we will elaborate on the steps that are taken to ensure that no privacy violation is performed.

   During the duration of the experiment, you will be filmed continuously. After the experiment ends, the videos will be segmented into video frames on the experimenter's computer, transferred to a computer located in SIMONA's building, and then deleted from the experimenter's computer. Once the data arrives on the TU Delft's computer, it will be sent through the pose estimator which extracts the pixel locations of the landmarks and anonymizes the subject.

Figure 11: The normal output of the pose estimator.

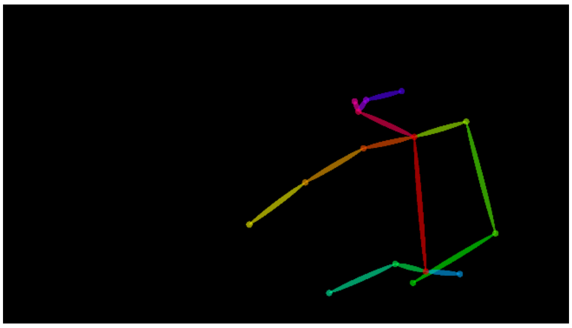


Figure 12: Anonymized output from the pose estimator.

   It needs to be mentioned that the rendered pictures of the pose can be only used to investigate the quality of the data, while the most crucial information that can be obtained from the pose estimator is represented by the pixel locations of the previously-mentioned landmarks (which cannot indicate any form of personal information).

   After the required pixel locations and the anonymized outputs are obtained, the original pictures are deleted, and the only information that is left on the computer is represented by the previously mentioned two types of data.

3) Personal information such as: Name, email address, age, and gender.

With this in mind, we will take all the necessary steps to ensure that your personal information remains confidential. The previously-mentioned data will be uploaded in a safe storage environment from TU Delft whose access is only allowed for the researchers that contribute to this experiment. To ensure confidentiality, your name will be attributed to a non-representative participant number. In this way, the name and participant number connection will be known only by the researchers. Lastly, the informed consent form that you will need to sign will be kept in a secure location in order to avoid the disclosure of any personal information.

The findings of this investigation will possibly be reported in upcoming scientific journals. Any publications (MSc thesis report, scientific papers, reports) about the study will never include your name and pictures.

Contact information researcher:	Contact information research supervisor:

Informed Consent Form

"Towards adaptive biodynamic feedthrough modelling for touchscreen applications"

Please tick the appropriate boxes	**Yes**	**No**
Taking part in the study		
I have read and understood the study information or it has been read to me. I have been able to ask questions about the study and my questions have been answered to my satisfaction.	○	○
I consent voluntarily to be a participant in this study and understand that I can withdraw from the study at any time, without having to give a reason.	○	○
I understand that taking part in the study involves the recording of my arm movements when providing control inputs on a touchscreen while being subjected to simulated aircraft turbulence in the SIMONA Research Simulator.	○	○
I understand that taking part in the study involves the video recording of my movements throughout the experiment session. Moreover, through the experiment briefing I was made aware that the obtained videos and extracted frames are anonymized after processing and under no circumstances any personal identification can be derived from the anonymized video frames.	○	○
I am aware that some basic identification information (name, email address, age, etc.) will be collected before the experiment, but will not be shared in identifiable form beyond the research team.	○	○
Risks associated with participating in the study		
I understand that taking part in the study involves the risk of physical discomfort such as mild motion sickness and that I should always report such discomfort to the experimenter.	○	○
Safety		
I confirm that the researcher has provided me with detailed safety and operational instructions for the hardware used in the experiment. Furthermore, I confirm that I have understood the researcher's instructions for guaranteeing that the experiment will be performed in line with current TU Delft COVID-19 guidelines and that this experiment shall at all times follow these guidelines.	○	○
I understand that this research study has been reviewed and approved by the TU Delft Human Research Ethics Committee (HREC). I am aware that I can report any problems regarding my participation in the experiment to the researchers using the contact information below.	○	○
Use of the information in the study		
I understand that information collected (excluding information that can identify me) will be used for scientific reports and publications.	○	○

I understand that personal information collected about me that can identify me, such as name and email address, will not be shared beyond the study team, and that my confidentiality as a participant in this study will remain secure.	○	○

Future use and reuse of the information by others

I give permission for the anonymised arm movement data that I provide to be archived on the 4TU.ResearchData repository so it can be used for future research and learning.	○	○

Signatures

________________________	__________________	________
Name of participant:	Signature	Date

I have accurately read out the information sheet to the potential participant and, to the best of my ability, ensured that the participant understands to what they are freely consenting.

________________________	__________________	________
Researcher name:	Signature	Date

Contact information researcher:	Contact information research supervisor:

Participant demographic information

"Towards adaptive biodynamic feedthrough modelling for touchscreen applications"

Age range:

○ 18 – 19
○ 20 – 24
○ 25 – 29
○ 30 – 34
○ 35 – 39
○ 40 – 44
○ 45 – 49
○ 50 – 55
○ 55+

Gender: __________

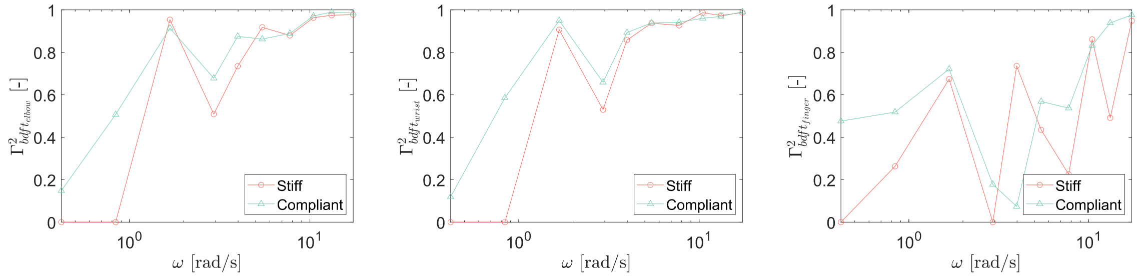
Handedness: __________

Participant number: __________
(filled out by the researcher)

B

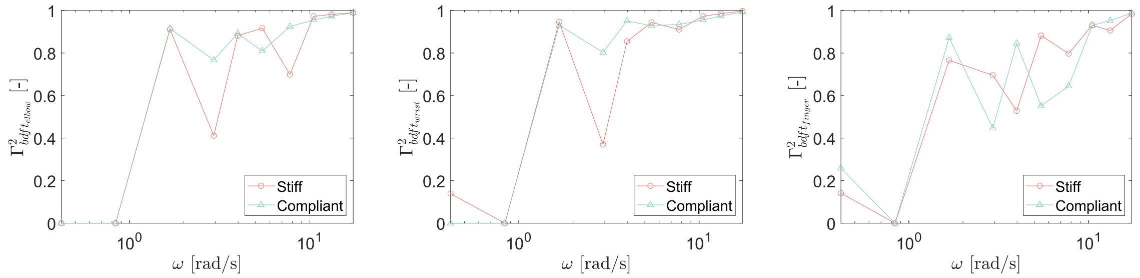
Coherence calculations

The following appendix proxides an overview of the coherence estimates for each subject, body-part and combination of *NMS* and *Task* independent variables. The information captured by the estimated coherences can be crucial for understanding which frequency points in the nonparametric estimates of the BDFT can be considered ”reliable” or not. Moreover, the computed quantity will provide the reader an insight regarding the linearity of the behavior corresponding to the three recorded body-parts in the context of the selected tasks and neuromuscular conditions.



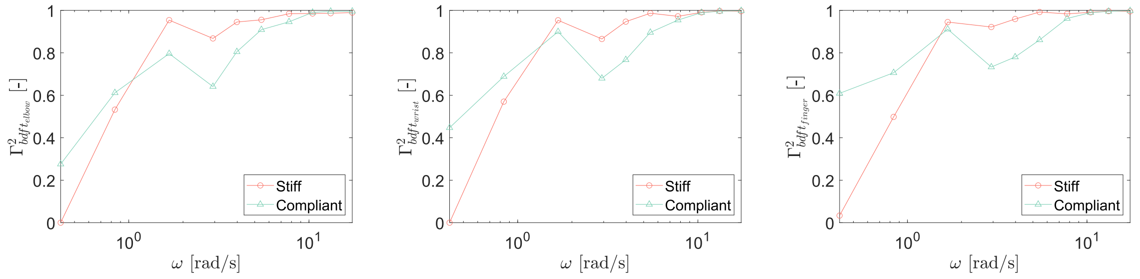
(a) Coherence of the elbow landmark **(b)** Coherence of the wrist landmark **(c)** Coherence of the finger landmark

Figure B.1: Overview of the estimated coherence calculation for Subject 1, Task: F.



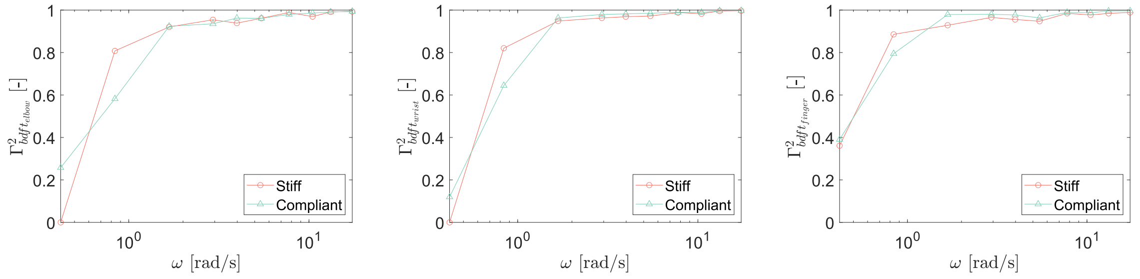
(a) Coherence of the elbow landmark **(b)** Coherence of the wrist landmark **(c)** Coherence of the finger landmark

Figure B.2: Overview of the estimated coherence calculation for Subject 1, Task: M.



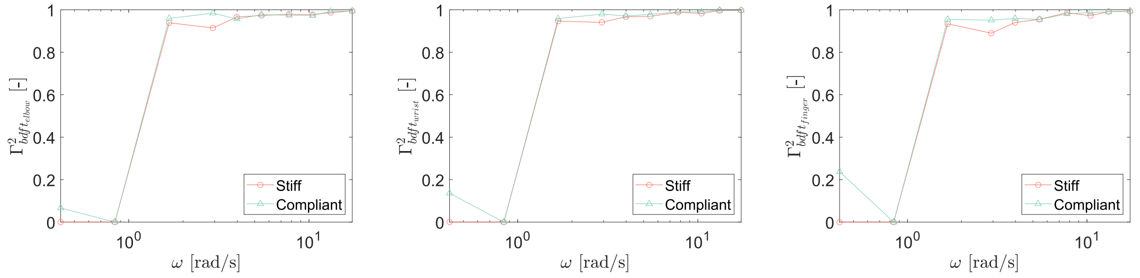
(a) Coherence of the elbow landmark **(b)** Coherence of the wrist landmark **(c)** Coherence of the finger landmark

Figure B.3: Overview of the estimated coherence calculation for Subject 1, Task: N.



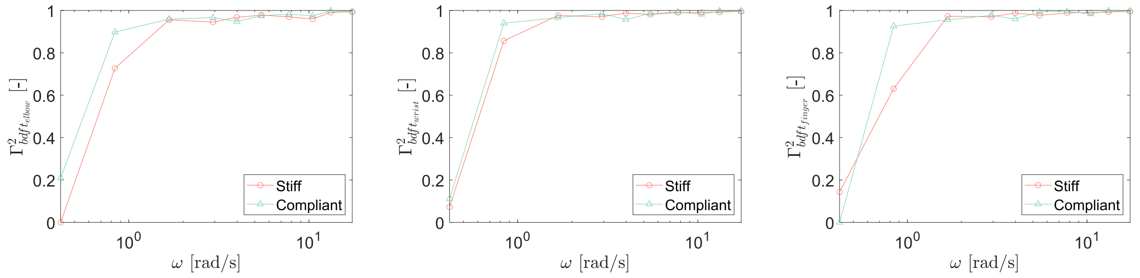
(a) Coherence of the elbow landmark **(b)** Coherence of the wrist landmark **(c)** Coherence of the finger landmark

Figure B.4: Overview of the estimated coherence calculation for Subject 2, Task: F.



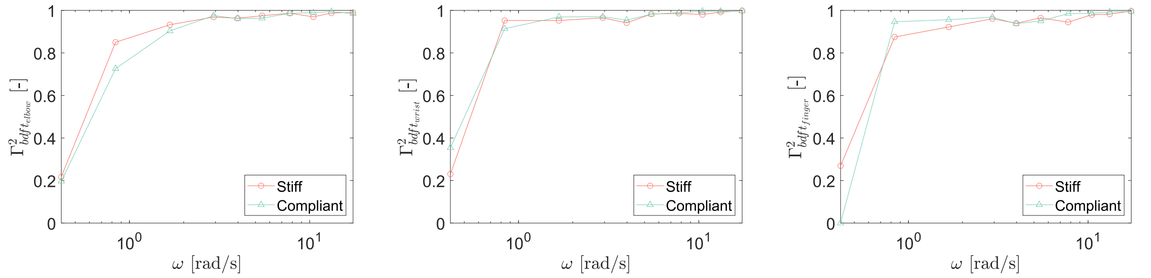
(a) Coherence of the elbow landmark **(b)** Coherence of the wrist landmark **(c)** Coherence of the finger landmark

Figure B.5: Overview of the estimated coherence calculation for Subject 2, Task: M.



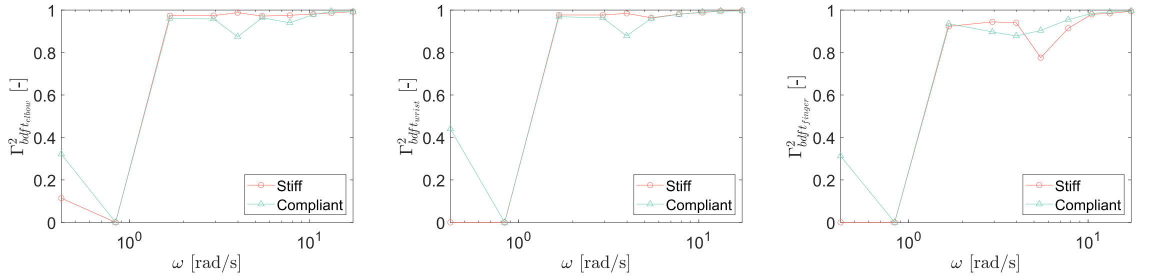
(a) Coherence of the elbow landmark **(b)** Coherence of the wrist landmark **(c)** Coherence of the finger landmark

Figure B.6: Overview of the estimated coherence calculation for Subject 2, Task: N.



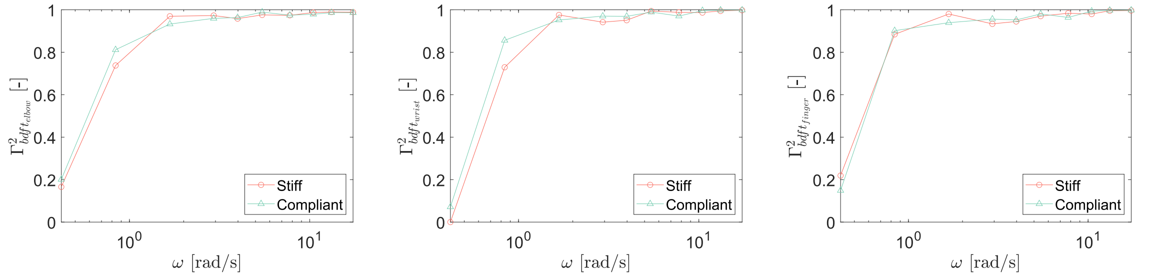
(a) Coherence of the elbow landmark **(b)** Coherence of the wrist landmark **(c)** Coherence of the finger landmark

Figure B.7: Overview of the estimated coherence calculation for Subject 3, Task: F.



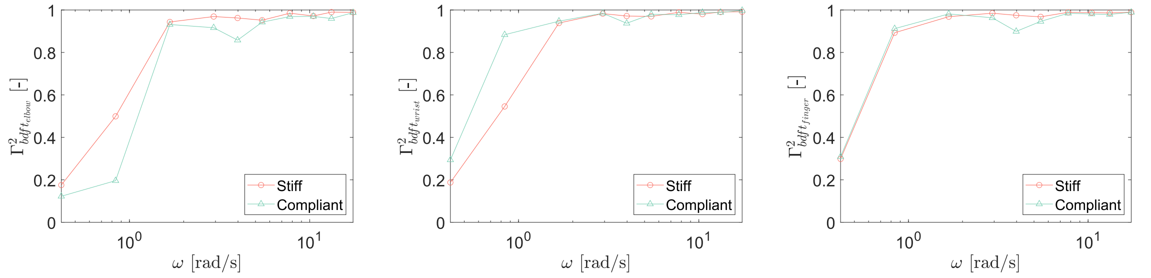
(a) Coherence of the elbow landmark **(b)** Coherence of the wrist landmark **(c)** Coherence of the finger landmark

Figure B.8: Overview of the estimated coherence calculation for Subject 3, Task: M.



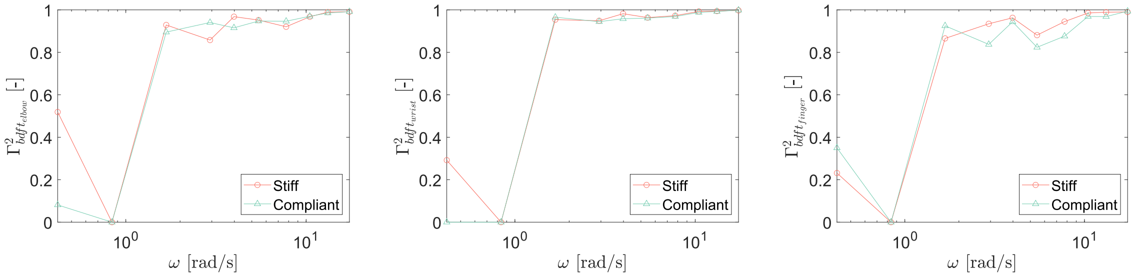
(a) Coherence of the elbow landmark **(b)** Coherence of the wrist landmark **(c)** Coherence of the finger landmark

Figure B.9: Overview of the estimated coherence calculation for Subject 3, Task: N.



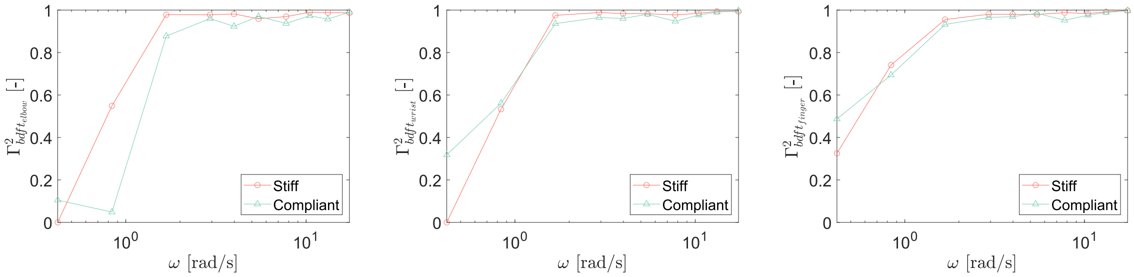
(a) Coherence of the elbow landmark **(b)** Coherence of the wrist landmark **(c)** Coherence of the finger landmark

Figure B.10: Overview of the estimated coherence calculation for Subject 4, Task: F.



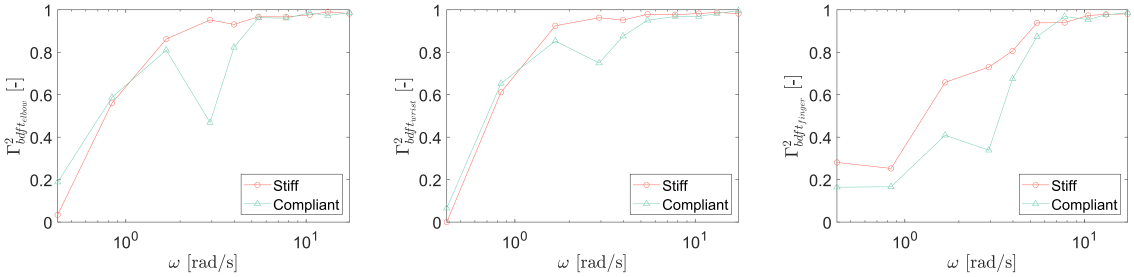
(a) Coherence of the elbow landmark **(b)** Coherence of the wrist landmark **(c)** Coherence of the finger landmark

Figure B.11: Overview of the estimated coherence calculation for Subject 4, Task: M.



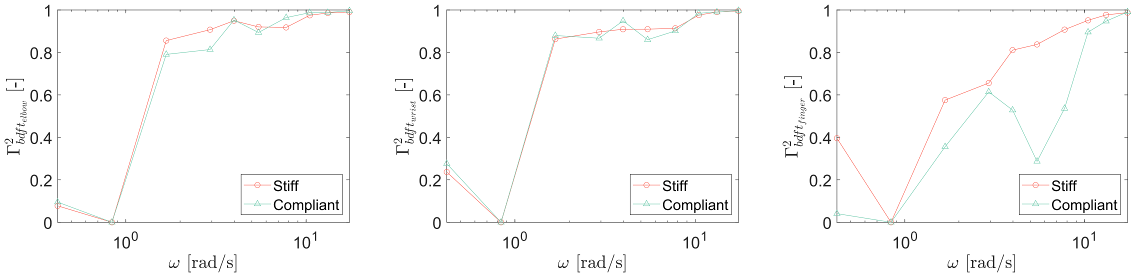
(a) Coherence of the elbow landmark **(b)** Coherence of the wrist landmark **(c)** Coherence of the finger landmark

Figure B.12: Overview of the estimated coherence calculation for Subject 4, Task: N.



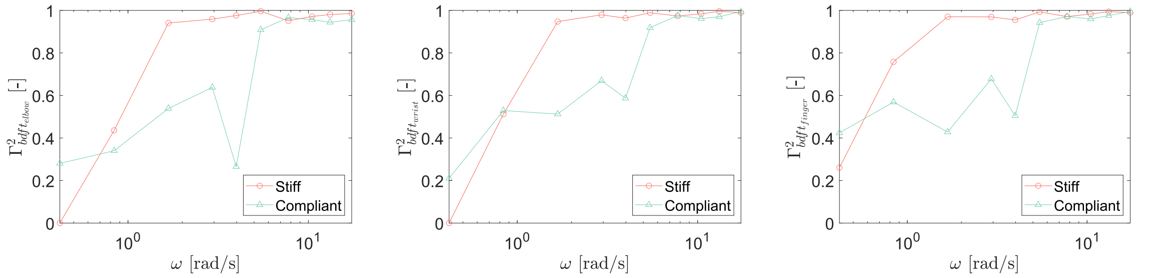
(a) Coherence of the elbow landmark **(b)** Coherence of the wrist landmark **(c)** Coherence of the finger landmark

Figure B.13: Overview of the estimated coherence calculation for Subject 5, Task: F.



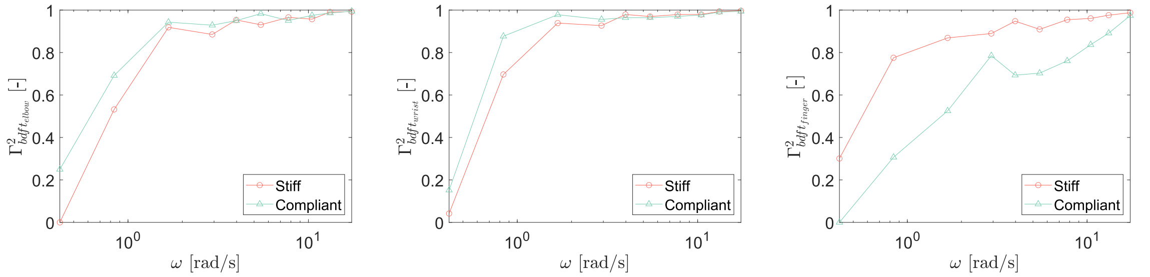
(a) Coherence of the elbow landmark **(b)** Coherence of the wrist landmark **(c)** Coherence of the finger landmark

Figure B.14: Overview of the estimated coherence calculation for Subject 5, Task: M.



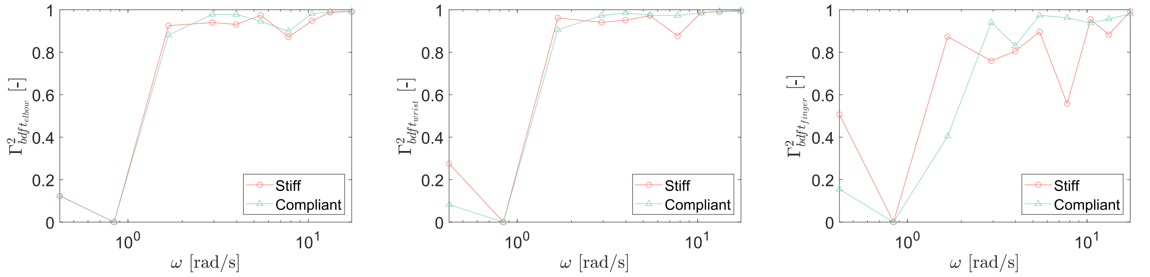
(a) Coherence of the elbow landmark **(b)** Coherence of the wrist landmark **(c)** Coherence of the finger landmark

Figure B.15: Overview of the estimated coherence calculation for Subject 5, Task: N.



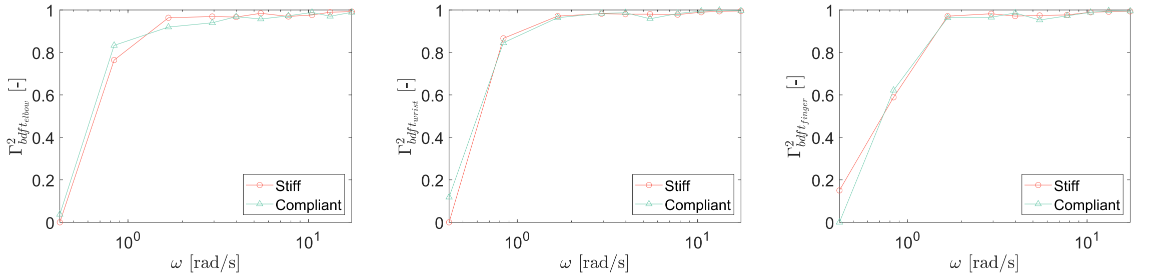
(a) Coherence of the elbow landmark **(b)** Coherence of the wrist landmark **(c)** Coherence of the finger landmark

Figure B.16: Overview of the estimated coherence calculation for Subject 6, Task: F.



(a) Coherence of the elbow landmark **(b)** Coherence of the wrist landmark **(c)** Coherence of the finger landmark

Figure B.17: Overview of the estimated coherence calculation for Subject 6, Task: M.



(a) Coherence of the elbow landmark **(b)** Coherence of the wrist landmark **(c)** Coherence of the finger landmark

Figure B.18: Overview of the estimated coherence calculation for Subject 6, Task: N.

B.1. Discussion of the outcomes corresponding to Appendix B

The figures presented above showed that the linearity of the recorded landmarks depends on: the landmark that is being analysed, the task, the neuromuscular settings and the subject that was analysed.

When focusing on the recorded landmark, it can be observed that for some subjects, the wrist landmark behaves more linearly compared to the finger landmark in the cases where the finger is interacting with the screen (e.g. Figure B.14b and Figure B.14c). This behavior can be caused by nonlinear effects such as finger slip or friction with the screen surface.

Another aspect that is influenced by the screen interaction is the change in coherence between the cases where the subject is interacting with the screen and when it is not. An example of such change is observed in Figure B.18b and Figure B.18c where the condition of the finger not touching the screen exhibits a more linear behavior as expected.

In some scenarios, a stiffer approach toward a task induces a more linear behavior compared to the case when the subject is more relaxed. This aspect can be explained by the fact that being stiff creates a stronger dynamical coupling between the disturbances and the movement of the limb (e.g. Figure B.15c).

Lastly, it can be observed that the linearity of the body parts depends on the subject that is being analysed. For example, the response exhibited in Figure B.15c is different than the response presented in Figure B.18c, especially for the compliant condition. These differences can be caused by different interpretations regarding how ”compliant” one should be. It was observed that the subject whose compliant response corresponds to Figure B.15c, was able to relax the most (having the highest RMS value of the finger landmark), fact that reduced the linearity of the system in the low-frequency range.

C

Pinhole camera model

The following appendix provides an overview regarding the pinhole camera model, estimated intrisic parameters for both cameras, and the extrinsic parameters for each of the six experiment sessions.

C.1. Perspective projection insights

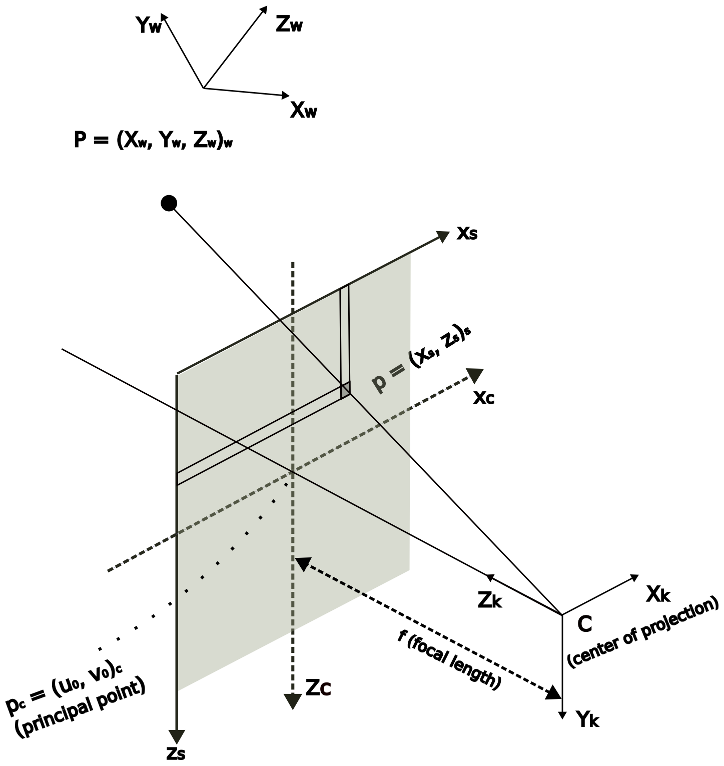


Figure C.1: Overview of the pinhole camera model.

The contents of this section (which are adapted from [1] and [2]) have the goal of explaining the steps that construct the pinhole camera model. Based on Figure C.1, the pinhole camera

---

[1] https://www.ipb.uni-bonn.de/html/teaching/photo12-2021/2021-pho1-20-camera-params.pptx.pdf

[2] https://docs.opencv.org/4.x/d9/d0c/group__calib3d.html

model consists of three different coordinate frames: The world coordinate system (subscript w), the camera-fixed coordinate system (subscript k), the image coordinate system (subscript c) and the sensor coordinate system (subscript s).

In this way, the previously mentioned coordinate systems are linked to the generation of the camera model through a series of coordinate transformations: A 3D to 3D transformation which links the world coordinate system to the camera-based coordinate system, a 3D to 2D transformation (described also as ideal perspective projection) which transforms a point from the camera-based coordinate system to a point in the image coordinate system and a 2D to 2D transformation which transforms a point from the image coordinate system to a pixel value in the sensor coordinate system.

The 3D to 3D transformation is represented by a matrix multiplication that relates the two coordinate systems through a translation matrix and a rotation vector, as follows:

$$\begin{bmatrix} X_k \\ Y_k \\ Z_k \\ 1 \end{bmatrix}_k = \begin{bmatrix} r_{00} & r_{01} & r_{02} & t_0 \\ r_{10} & r_{11} & r_{12} & t_1 \\ r_{20} & r_{21} & r_{22} & t_2 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}_w \tag{C.1}$$

After point P was transformed to the camera-fixed coordinate frame, the ideal perspective projection is performed to bring point P into the image coordinate system, using the transformed 3D points in the camera-fixed reference frame (subscript *s*) and the focal length (*f*):

$$x_c = \frac{X_k}{Z_k} f \quad and \quad z_c = \frac{Y_k}{Z_k} f \tag{C.2}$$

It needs to be mentioned that the units in which the values of Equation C.2 are expressed are the units of the focal length which can be metric (ex. mm). To express these quantities in pixels, a translation (by u_0 and v_0) and scaling (m_x and m_z) by the pixel dimensions needs to be performed, in the following manner:

$$x_s = m_x \cdot x_c + u_0 \quad z_s = m_z \cdot z_c + v_0 \tag{C.3}$$

In this way, by combining Equation C.2 and Equation C.3:

$$x_s = m_x \frac{X_k}{Z_k} f + u_0 \quad z_s = m_z \frac{Y_k}{Z_k} f + v_0 \tag{C.4}$$

Summarizing Equation C.4 in a matrix format:

$$\begin{bmatrix} m x_s \\ m z_s \\ m \end{bmatrix}_s = \begin{bmatrix} m_x f & 0 & u_0 & 0 \\ 0 & m_z f & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X_k \\ Y_k \\ Z_k \\ 1 \end{bmatrix}_k \tag{C.5}$$

And combining Equation C.1 and Equation C.5, the following formula expresses the relationship between a given 3D point in space and a projected point in 2D:

$$m \begin{bmatrix} x_s \\ z_s \\ 1 \end{bmatrix}_s = \underbrace{\begin{bmatrix} a_x & 0 & u_0 \\ 0 & a_z & v_0 \\ 0 & 0 & 1 \end{bmatrix}}_{\text{Intrinsic parameters}} \underbrace{\begin{bmatrix} r_{00} & r_{01} & r_{02} & t_0 \\ r_{10} & r_{11} & r_{12} & t_1 \\ r_{20} & r_{21} & r_{22} & t_2 \end{bmatrix}}_{\text{Extrinsic parameters}} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}_w \tag{C.6}$$

All the formulas that were presented above made use of the projective-space representation of a vector. A vector $[x, y]^T \subset \mathbf{R}^2$ is equivalent with its projective-space counterpart $[x, y, 1]$ $\subset \mathbf{R}^3$ which, in turn, is equivalent to any multiple m of the projective-space vector $[mx, my, m]$.

In reality, cameras possess distortions. The pinhole camera model does not account for the nonlinearities induced by them, and therefore a correction needs to be made in order to either transform an undistorted image into a distorted one or vice-versa. The distorted (x_{s_d}, y_{s_d}) pixels have the following formula:

$$\begin{aligned} x_{s_d} &= x_s + (x_s - u_0)[k_1(x^2+z^2] + k_2(x^2+z^2)^2 + k_3(x^2+z^2)^3] \\ z_{s_d} &= z_s + (z_s - v_0)[k_1(x^2+z^2) + k_2(x^2+z^2)^2 + k_3(x^2+z^2)^3] \end{aligned} \tag{C.7}$$

Equation C.7 presents how the distorted pixel values (x_{s_d}, z_{s_d}) are modeled as a function of the undistorted pixel values (x_s, z_s), the image centers $(u_0.v_0)$, normalized image coordinates $(x = \frac{X_k}{Z_k}, z = \frac{Y_k}{Z_k})$, and the distorsion coefficients $(k_1.k_2, k_3)$. All these parameters are estimated using Zhang's method (Zhang, 2000).

C.2. Intrinsic parameters

As previously mentioned, the intrinsic parameters represent one part of the pinhole camera model. This section introduces the estimated intrinsic parameters for each of the two cameras (K_{C_1}, K_{C_2}) and the two pairs of distorsion coefficients (D_{C_1}, D_{C_2}).

C.2.1. Camera 1

$$K_{C_1} = \begin{bmatrix} 916.39 & 0 & 961.77 \\ 0 & 914.04 & 538.99 \\ 0 & 0 & 1 \end{bmatrix}$$

$$D_{C_1} = (0.0192, -0.0611, 0.0533)$$

Mean reprojection error - Camera 1: 0.27 $[px]$

C.2.2. Camera 2

$$K_{C_2} = \begin{bmatrix} 917.62 & 0 & 960.97 \\ 0 & 915.35 & 536.86 \\ 0 & 0 & 1 \end{bmatrix}$$

$$D_{C_2} = (0.0206, -0.0606, 0.0499)$$

Mean reprojection error - Camera 2: 0.29 $[px]$

C.3. Extrinsic parameters

This section introduces the estimated extrinsic parameters for every extrinsic calibration session. The extrinsic parameters are represented by a translation (T_S) and a rotation (R_S).

C.3.1. Subject 1

$$R_{S_1} = \begin{bmatrix} 0.892 & 0.024 & -0.452 \\ -0.019 & 0.999 & 0.016 \\ 0.452 & -0.006 & 0.892 \end{bmatrix} \quad T_{S_1} = \begin{bmatrix} 336.088 \\ -162.388 \\ -21.329 \end{bmatrix}$$

Mean reprojection error - Subject 1: 0.40 $[px]$

C.3.2. Subject 2

$$R_{S_2} = \begin{bmatrix} 0.890 & 0.016 & -0.455 \\ -0.029 & 0.999 & -0.023 \\ 0.454 & 0.033 & 0.890 \end{bmatrix} \qquad T_{S_2} = \begin{bmatrix} 336.562 \\ -161.573 \\ -26.277 \end{bmatrix}$$

Mean reprojection error - Subject 2: 0.41 $[px]$

C.3.3. Subject 3

$$R_{S_3} = \begin{bmatrix} 0.904 & 0.026 & -0.428 \\ -0.026 & 0.999 & -0.006 \\ 0.428 & 0.005 & 0.904 \end{bmatrix} \qquad T_{S_3} = \begin{bmatrix} 332.922 \\ -152.299 \\ -6.826 \end{bmatrix}$$

Mean reprojection error - Subject 3: 0.41 $[px]$

C.3.4. Subject 4

$$R_{S_4} = \begin{bmatrix} 0.897 & 0.038 & -0.440 \\ -0.022 & 0.999 & 0.042 \\ 0.442 & -0.028 & 0.897 \end{bmatrix} \qquad T_{S_4} = \begin{bmatrix} 330.687 \\ -161.494 \\ -7.525 \end{bmatrix}$$

Mean reprojection error - Subject 4: 0.40 $[px]$

C.3.5. Subject 5

$$R_{S_5} = \begin{bmatrix} 0.893 & -0.006 & -0.451 \\ -0.037 & 0.998 & 0.060 \\ 0.449 & -0.070 & 0.891 \end{bmatrix} \qquad T_{S_5} = \begin{bmatrix} 334.724 \\ -160.992 \\ 1.753 \end{bmatrix}$$

Mean reprojection error - Subject 5: 0.39 $[px]$

C.3.6. Subject 6

$$R_{S_6} = \begin{bmatrix} 0.891 & -0.001 & -0.454 \\ 0.018 & 0.999 & 0.033 \\ 0.453 & -0.037 & 0.891 \end{bmatrix} \qquad T_{S_6} = \begin{bmatrix} 333.552 \\ -161.181 \\ -1.852 \end{bmatrix}$$

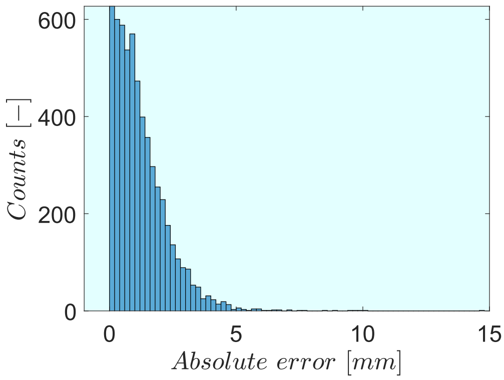
Mean reprojection error - Subject 6: 0.41 $[px]$

D

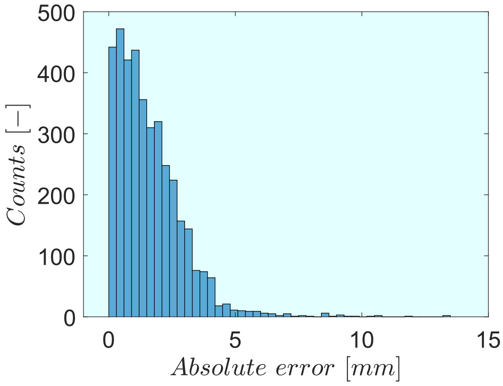
Validation histograms

The following appendix provides an overview of the complete histograms that describe the absolute errors between the signals obtained from the touchscreens and the signals obtained from the stereo system.

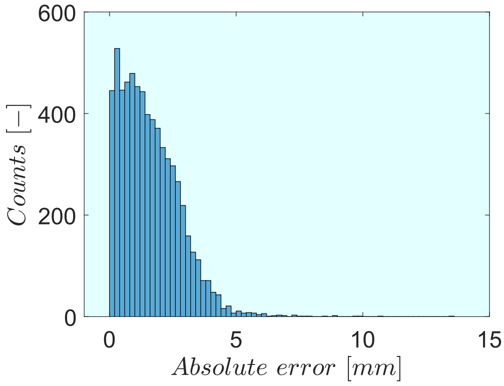
This visualisation was performed by combining the data from the three repetitions and all subjects in the four different conditions that concerned the interaction between the finger and the screen. Moreover, this visualisation complements the information provided in the scientific paper, since it provides a more complete understanding of the error distribution.



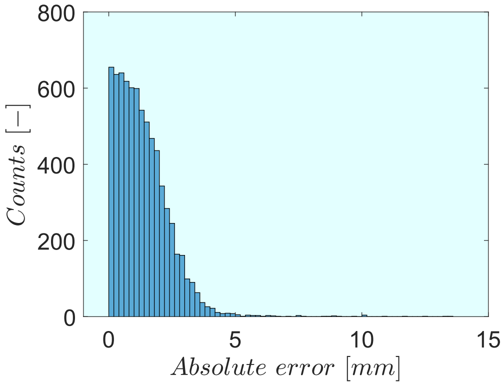
(a) Histogram of absolute errors for NMS: C and Task: F



(b) Histogram of absolute errors for NMS: S and Task: F



(c) Histogram of absolute errors for NMS: C and Task: M



(d) Histogram of absolute errors for NMS: S and Task: M

E

Second-order system parameter estimations

The following appendix provides an overview of the run-specific and averaged parameters of the time-domain identification method using a mass-spring-damper model. The goal of this appendix is to give the reader the understanding that in some cases, the averaging of the parameters (the ones reported in the scientific article) might induce erroneous results due to the spread in the values of the parameters identified using this method of identification.

Table E.1: Overview of the run-level and averaged parameters for Subject 1, NMS: Compliant, Task: F

	G_{bdft}			ω_{bdft}			ζ_{bdft}		
	Elbow	Wrist	Finger	Elbow	Wrist	Finger	Elbow	Wrist	Finger
Run 1	18.68	14.65	5.01	3397.83	16.85	16.15	601.91	0.71	0.36
Run 2	7.48	11.96	4.49	13.86	13.67	14.48	0.48	0.45	0.23
Run 3	8.29	15.18	6.54	13.57	13.02	12.56	0.61	0.52	0.33
Average	11.48	13.93	5.35	1141.75	14.51	14.40	201.00	0.56	0.31

Table E.2: Overview of the run-level and averaged parameters for Subject 1, NMS: Stiff, Task: F

	G_{bdft}			ω_{bdft}			ζ_{bdft}		
	Elbow	Wrist	Finger	Elbow	Wrist	Finger	Elbow	Wrist	Finger
Run 1	6.65	14.54	4.48	16.99	15.07	11.18	0.95	1.18	0.81
Run 2	27.65	33.50	11.93	50063.47	125685310.27	686.52	15443.64	28746786.93	183.72
Run 3	20.94	23.98	7.20	3247.25	23.17	12.54	735.23	3.74	2.17
Average	18.41	24.01	7.87	17775.90	41895116.17	236.75	5393.28	9582263.95	62.24

Table E.3: Overview of the run-level and averaged parameters for Subject 1, NMS: Compliant, Task: M

	G_{bdft}			ω_{bdft}			ζ_{bdft}		
	Elbow	Wrist	Finger	Elbow	Wrist	Finger	Elbow	Wrist	Finger
Run 1	40.91	70.04	28.71	15.98	10.96	11.11	4.46	3.21	2.23
Run 2	43.98	76.73	15.70	8.37	7.70	10.46	2.76	2.58	0.99
Run 3	35.56	66.34	7.70	12.47	9.05	13.30	3.33	2.71	0.40
Average	40.15	71.03	17.37	12.27	9.24	11.62	3.52	2.83	1.21

Table E.4: Overview of the run-level and averaged parameters for Subject 1, NMS: Stiff, Task: M

	G_{bdft}			ω_{bdft}			ζ_{bdft}		
	Elbow	Wrist	Finger	Elbow	Wrist	Finger	Elbow	Wrist	Finger
Run 1	72.46	96.99	19.92	162.62	183.78	11.42	83.11	97.57	2.06
Run 2	87.03	109.80	36.43	8.00	6.01	4.97	4.80	3.19	1.65
Run 3	84.84	127.81	28.71	247.37	2133.55	6.41	187.76	1721.41	2.11
Average	81.45	111.53	28.35	139.33	774.45	7.60	91.89	607.39	1.94

Table E.5: Overview of the run-level and averaged parameters for Subject 1, NMS: Compliant, Task: N

	G_{bdft}			ω_{bdft}			ζ_{bdft}		
	Elbow	Wrist	Finger	Elbow	Wrist	Finger	Elbow	Wrist	Finger
Run 1	52.78	106.97	85.01	417.08	26.06	15.48	152.43	11.48	3.89
Run 2	25.35	11.63	15.93	1688.60	14.12	14.02	394.52	0.29	0.26
Run 3	4.54	10.48	14.50	13.94	13.82	13.80	0.19	0.22	0.21
Average	27.56	43.03	38.48	706.54	18.00	14.43	182.38	4.00	1.46

Table E.6: Overview of the run-level and averaged parameters for Subject 1, NMS: Stiff, Task: N

	G_{bdft}			ω_{bdft}			ζ_{bdft}		
	Elbow	Wrist	Finger	Elbow	Wrist	Finger	Elbow	Wrist	Finger
Run 1	46.34	90.21	61.71	853.07	27.86	17.08	289.65	10.92	3.33
Run 2	36.10	75.36	82.51	314.62	624.85	60.82	87.65	195.09	17.03
Run 3	29.50	40.85	42.44	11.79	10.42	10.78	2.50	1.82	1.45
Average	37.31	68.80	62.22	393.16	221.05	29.56	126.60	69.28	7.27

Table E.7: Overview of the run-level and averaged parameters for Subject 2, NMS: Compliant, Task: F

	G_{bdft}			ω_{bdft}			ζ_{bdft}		
	Elbow	Wrist	Finger	Elbow	Wrist	Finger	Elbow	Wrist	Finger
Run 1	21.97	36.63	40.04	10.74	10.07	9.99	1.42	1.06	1.01
Run 2	29.85	53.85	42.50	7.97	7.80	8.37	1.76	1.33	1.01
Run 3	20.74	50.22	42.17	12.11	10.29	9.92	1.83	1.56	1.19
Average	24.19	46.90	41.57	10.27	9.38	9.43	1.67	1.32	1.07

Table E.8: Overview of the run-level and averaged parameters for Subject 2, NMS: Stiff, Task: F

	G_{bdft}			ω_{bdft}			ζ_{bdft}		
	Elbow	Wrist	Finger	Elbow	Wrist	Finger	Elbow	Wrist	Finger
Run 1	32.44	67.16	45.40	11.11	8.91	8.96	2.45	2.00	1.57
Run 2	32.80	59.67	43.37	16.91	11.96	11.64	3.77	2.33	1.93
Run 3	28.14	54.28	34.13	11.18	9.20	8.44	2.13	1.71	1.35
Average	31.13	60.37	40.96	13.07	10.03	9.68	2.78	2.01	1.61

Table E.9: Overview of the run-level and averaged parameters for Subject 2, NMS: Compliant, Task: M

		G_{bdft}			ω_{bdft}			ζ_{bdft}	
	Elbow	Wrist	Finger	Elbow	Wrist	Finger	Elbow	Wrist	Finger
Run 1	41.30	75.71	55.48	7.99	7.06	6.86	1.90	1.55	1.27
Run 2	35.24	62.45	44.79	11.32	9.29	9.28	2.47	1.74	1.42
Run 3	46.26	97.45	59.28	6.92	6.08	5.85	2.08	1.72	1.18
Average	40.93	78.53	53.19	8.74	7.48	7.33	2.15	1.67	1.29

Table E.10: Overview of the run-level and averaged parameters for Subject 2, NMS: Stiff, Task: M

		G_{bdft}			ω_{bdft}			ζ_{bdft}	
	Elbow	Wrist	Finger	Elbow	Wrist	Finger	Elbow	Wrist	Finger
Run 1	38.36	81.32	60.65	6.87	6.23	6.21	1.69	1.60	1.32
Run 2	36.45	78.14	60.12	8.31	7.04	5.79	2.12	1.76	1.31
Run 3	37.16	77.29	56.44	6.08	5.49	4.79	1.54	1.38	1.11
Average	37.32	78.92	59.07	7.09	6.25	5.60	1.78	1.58	1.25

Table E.11: Overview of the run-level and averaged parameters for Subject 2, NMS: Compliant, Task: N

		G_{bdft}			ω_{bdft}			ζ_{bdft}	
	Elbow	Wrist	Finger	Elbow	Wrist	Finger	Elbow	Wrist	Finger
Run 1	38.13	61.42	60.03	11.91	9.63	10.18	2.88	1.99	1.46
Run 2	63.37	107.35	95.29	8.11	6.74	7.58	3.23	2.37	1.83
Run 3	43.32	83.85	70.40	6.93	6.27	7.61	2.11	1.84	1.46
Average	48.27	84.21	75.24	8.98	7.55	8.46	2.74	2.07	1.58

Table E.12: Overview of the run-level and averaged parameters for Subject 2, NMS: Stiff, Task: N

		G_{bdft}			ω_{bdft}			ζ_{bdft}	
	Elbow	Wrist	Finger	Elbow	Wrist	Finger	Elbow	Wrist	Finger
Run 1	42.08	68.57	66.03	9.99	8.55	9.39	3.05	2.11	1.66
Run 2	39.50	68.19	69.82	9.44	8.40	9.28	2.61	2.04	1.71
Run 3	44.42	67.57	75.97	9.55	8.46	9.29	2.77	1.95	1.77
Average	42.00	68.11	70.60	9.66	8.47	9.32	2.81	2.03	1.71

Table E.13: Overview of the run-level and averaged parameters for Subject 3, NMS: Compliant, Task: F

		G_{bdft}			ω_{bdft}			ζ_{bdft}	
	Elbow	Wrist	Finger	Elbow	Wrist	Finger	Elbow	Wrist	Finger
Run 1	24.43	47.70	24.70	6.81	6.89	9.08	0.88	0.88	0.62
Run 2	19.32	36.29	28.33	7.88	7.99	9.20	0.76	0.70	0.63
Run 3	36.03	74.96	35.39	5.70	5.78	8.13	1.31	1.23	0.82
Average	26.59	52.98	29.47	6.80	6.89	8.80	0.98	0.94	0.69

Table E.14: Overview of the run-level and averaged parameters for Subject 3, NMS: Stiff, Task: F

	G_{bdft}			ω_{bdft}			ζ_{bdft}		
	Elbow	Wrist	Finger	Elbow	Wrist	Finger	Elbow	Wrist	Finger
Run 1	26.39	45.60	25.78	13.93	10.31	12.27	2.76	1.77	1.31
Run 2	33.85	52.25	23.29	14.64	9.35	10.86	4.54	2.00	1.57
Run 3	17.37	55.47	25.29	33.81	15.08	17.50	5.33	3.21	2.47
Average	25.87	51.11	24.79	20.79	11.58	13.54	4.21	2.33	1.78

Table E.15: Overview of the run-level and averaged parameters for Subject 3, NMS: Compliant, Task: M

	G_{bdft}			ω_{bdft}			ζ_{bdft}		
	Elbow	Wrist	Finger	Elbow	Wrist	Finger	Elbow	Wrist	Finger
Run 1	27.03	63.29	26.30	6.84	6.21	8.55	1.21	1.30	0.78
Run 2	34.83	61.29	29.91	6.50	6.64	8.79	1.37	1.21	0.88
Run 3	41.77	85.25	27.19	6.44	6.15	9.03	1.81	1.64	0.97
Average	34.54	69.94	27.80	6.59	6.33	8.79	1.47	1.38	0.88

Table E.16: Overview of the run-level and averaged parameters for Subject 3, NMS: Stiff, Task: M

	G_{bdft}			ω_{bdft}			ζ_{bdft}		
	Elbow	Wrist	Finger	Elbow	Wrist	Finger	Elbow	Wrist	Finger
Run 1	59.34	88.85	26.20	5.41	5.42	8.48	2.03	1.64	1.15
Run 2	57.63	99.29	30.20	4.67	4.91	8.85	1.67	1.61	1.29
Run 3	38.03	83.02	25.04	5.56	5.47	9.52	1.44	1.56	1.09
Average	51.67	90.39	27.15	5.21	5.27	8.95	1.71	1.60	1.18

Table E.17: Overview of the run-level and averaged parameters for Subject 3, NMS: Compliant, Task: N

	G_{bdft}			ω_{bdft}			ζ_{bdft}		
	Elbow	Wrist	Finger	Elbow	Wrist	Finger	Elbow	Wrist	Finger
Run 1	27.69	59.73	57.60	7.61	7.27	8.21	1.30	1.24	1.02
Run 2	35.57	53.03	45.53	16.76	11.83	12.89	3.87	1.87	1.26
Run 3	30.08	56.11	55.00	8.00	8.06	9.09	1.52	1.39	1.12
Average	31.11	56.29	52.71	10.79	9.05	10.06	2.23	1.50	1.13

Table E.18: Overview of the run-level and averaged parameters for Subject 3, NMS: Stiff, Task: N

	G_{bdft}			ω_{bdft}			ζ_{bdft}		
	Elbow	Wrist	Finger	Elbow	Wrist	Finger	Elbow	Wrist	Finger
Run 1	35.81	76.28	48.13	7.39	7.50	10.83	1.84	1.86	1.50
Run 2	47.88	80.54	47.11	6.75	6.62	9.76	1.92	1.77	1.28
Run 3	39.57	76.11	45.16	7.16	6.97	9.52	1.73	1.57	1.13
Average	41.08	77.64	46.80	7.10	7.03	10.03	1.83	1.73	1.30

Table E.19: Overview of the run-level and averaged parameters for Subject 4, NMS: Compliant, Task: F

		G_{bdft}			ω_{bdft}			ζ_{bdft}	
	Elbow	Wrist	Finger	Elbow	Wrist	Finger	Elbow	Wrist	Finger
Run 1	12.12	52.74	23.18	10.77	8.66	10.06	0.92	1.65	0.85
Run 2	17.13	49.89	21.00	7.31	6.78	8.21	1.31	1.27	0.72
Run 3	14.39	32.93	14.66	10.62	10.11	11.20	1.23	1.02	0.48
Average	14.55	45.19	19.61	9.56	8.52	9.82	1.15	1.31	0.68

Table E.20: Overview of the run-level and averaged parameters for Subject 4, NMS: Stiff, Task: F

		G_{bdft}			ω_{bdft}			ζ_{bdft}	
	Elbow	Wrist	Finger	Elbow	Wrist	Finger	Elbow	Wrist	Finger
Run 1	47.94	81.12	40.37	6.72	6.36	8.58	2.39	1.96	1.30
Run 2	52.86	94.20	38.92	7.80	6.59	8.15	3.06	2.41	1.34
Run 3	34.72	56.03	33.96	7.29	6.70	7.88	1.92	1.57	1.19
Average	45.17	77.12	37.75	7.27	6.55	8.20	2.46	1.98	1.28

Table E.21: Overview of the run-level and averaged parameters for Subject 4, NMS: Compliant, Task: M

		G_{bdft}			ω_{bdft}			ζ_{bdft}	
	Elbow	Wrist	Finger	Elbow	Wrist	Finger	Elbow	Wrist	Finger
Run 1	52.75	90.34	21.30	7.12	6.77	9.63	3.24	2.60	0.98
Run 2	52.40	73.17	29.28	6.02	6.42	7.96	1.97	1.67	1.05
Run 3	48.10	59.92	22.04	6.73	7.28	9.12	2.64	1.73	0.80
Average	51.08	74.48	24.21	6.62	6.83	8.90	2.62	2.00	0.94

Table E.22: Overview of the run-level and averaged parameters for Subject 4, NMS: Stiff, Task: M

		G_{bdft}			ω_{bdft}			ζ_{bdft}	
	Elbow	Wrist	Finger	Elbow	Wrist	Finger	Elbow	Wrist	Finger
Run 1	51.12	72.37	31.12	4.82	5.12	6.74	1.80	1.51	0.93
Run 2	55.99	85.38	30.96	7.25	6.83	8.43	3.02	2.35	1.15
Run 3	55.69	92.30	33.44	5.63	5.53	7.29	2.41	2.17	1.20
Average	54.26	83.35	31.84	5.90	5.82	7.49	2.41	2.01	1.09

Table E.23: Overview of the run-level and averaged parameters for Subject 4, NMS: Compliant, Task: N

		G_{bdft}			ω_{bdft}			ζ_{bdft}	
	Elbow	Wrist	Finger	Elbow	Wrist	Finger	Elbow	Wrist	Finger
Run 1	20.09	49.14	50.10	9.19	8.56	9.10	1.28	1.29	1.06
Run 2	36.24	88.52	102.67	5.50	5.35	5.53	1.38	1.50	1.45
Run 3	26.83	51.76	54.93	8.22	8.12	8.47	1.38	1.16	1.04
Average	27.72	63.14	69.23	7.64	7.34	7.70	1.35	1.31	1.18

Table E.24: Overview of the run-level and averaged parameters for Subject 4, NMS: Stiff, Task: N

		G_{bdft}			ω_{bdft}			ζ_{bdft}	
	Elbow	Wrist	Finger	Elbow	Wrist	Finger	Elbow	Wrist	Finger
Run 1	65.53	106.50	109.63	5.96	5.47	6.14	2.37	1.93	1.77
Run 2	45.97	84.93	73.53	7.26	7.05	8.13	2.29	1.93	1.60
Run 3	59.13	109.08	87.37	4.91	4.86	6.35	1.84	1.80	1.65
Average	56.88	100.17	90.18	6.04	5.79	6.87	2.17	1.89	1.67

Table E.25: Overview of the run-level and averaged parameters for Subject 5, NMS: Compliant, Task: F

		G_{bdft}			ω_{bdft}			ζ_{bdft}	
	Elbow	Wrist	Finger	Elbow	Wrist	Finger	Elbow	Wrist	Finger
Run 1	8.88	16.57	16.64	9.84	9.73	9.65	0.26	0.21	0.17
Run 2	7.30	16.22	17.32	9.17	9.23	9.23	0.14	0.14	0.10
Run 3	15.85	15.81	14.45	9.94	9.84	9.85	0.58	0.22	0.15
Average	10.68	16.20	16.14	9.65	9.60	9.58	0.33	0.19	0.14

Table E.26: Overview of the run-level and averaged parameters for Subject 5, NMS: Stiff, Task: F

		G_{bdft}			ω_{bdft}			ζ_{bdft}	
	Elbow	Wrist	Finger	Elbow	Wrist	Finger	Elbow	Wrist	Finger
Run 1	55.80	39.82	15.48	10.75	9.60	11.62	4.69	1.98	0.72
Run 2	50.49	53.64	31.53	94310.34	11.42	11.32	41463.81	3.22	1.69
Run 3	43.20	66.02	13.55	10.20	9.39	13.15	3.26	2.82	0.67
Average	49.83	53.16	20.19	31443.76	10.14	12.03	13823.92	2.68	1.03

Table E.27: Overview of the run-level and averaged parameters for Subject 5, NMS: Compliant, Task: M

		G_{bdft}			ω_{bdft}			ζ_{bdft}	
	Elbow	Wrist	Finger	Elbow	Wrist	Finger	Elbow	Wrist	Finger
Run 1	35.80	43.31	3.61	8.99	8.72	12.39	2.40	2.09	0.28
Run 2	56.67	64.60	12.76	6.00	6.02	7.66	2.30	1.99	1.00
Run 3	46.85	54.16	16.60	6.76	6.73	9.16	2.01	1.65	0.76
Average	46.44	54.03	10.99	7.25	7.15	9.74	2.24	1.91	0.68

Table E.28: Overview of the run-level and averaged parameters for Subject 5, NMS: Stiff, Task: M

		G_{bdft}			ω_{bdft}			ζ_{bdft}	
	Elbow	Wrist	Finger	Elbow	Wrist	Finger	Elbow	Wrist	Finger
Run 1	47.76	44.01	17.15	70023313.47	12.71	11.53	24979493.42	3.26	1.17
Run 2	67.88	50.75	26.97	8.32	8.56	9.80	4.82	2.45	1.56
Run 3	53.64	53.85	23.47	12.50	10.48	13.07	4.70	2.89	1.53
Average	56.43	49.53	22.53	23341111.43	10.58	11.47	8326500.98	2.87	1.42

Table E.29: Overview of the run-level and averaged parameters for Subject 5, NMS: Compliant, Task: N

		G_{bdft}			ω_{bdft}			ζ_{bdft}	
	Elbow	Wrist	Finger	Elbow	Wrist	Finger	Elbow	Wrist	Finger
Run 1	8.08	15.09	20.38	10.47	10.45	10.49	0.21	0.20	0.19
Run 2	8.40	16.21	22.38	9.94	9.93	9.96	0.20	0.17	0.17
Run 3	7.40	16.80	23.04	9.77	9.84	9.86	0.18	0.17	0.17
Average	7.96	16.03	21.94	10.06	10.07	10.10	0.20	0.18	0.18

Table E.30: Overview of the run-level and averaged parameters for Subject 5, NMS: Stiff, Task: N

		G_{bdft}			ω_{bdft}			ζ_{bdft}	
	Elbow	Wrist	Finger	Elbow	Wrist	Finger	Elbow	Wrist	Finger
Run 1	71.26	68.73	53.59	40327991.76	84.25	28.28	19293699.95	23.12	4.69
Run 2	27.61	53.07	54.32	12.87	11.77	12.29	2.92	2.34	1.97
Run 3	44.10	69.46	61.71	6.83	7.20	8.39	1.86	1.73	1.48
Average	47.65	63.75	56.54	13442670.49	34.40	16.32	6431234.91	9.06	2.71

Table E.31: Overview of the run-level and averaged parameters for Subject 6, NMS: Compliant, Task: F

		G_{bdft}			ω_{bdft}			ζ_{bdft}	
	Elbow	Wrist	Finger	Elbow	Wrist	Finger	Elbow	Wrist	Finger
Run 1	67.31	82.51	7.16	6.62	6.25	9.67	2.96	2.03	0.52
Run 2	54.91	61.36	27.45	9.00	7.93	7.36	3.66	2.10	1.31
Run 3	86.89	101.60	17.54	6.36	6.25	6.80	3.33	2.45	0.95
Average	69.70	81.82	17.39	7.33	6.81	7.94	3.32	2.19	0.93

Table E.32: Overview of the run-level and averaged parameters for Subject 6, NMS: Stiff, Task: F

		G_{bdft}			ω_{bdft}			ζ_{bdft}	
	Elbow	Wrist	Finger	Elbow	Wrist	Finger	Elbow	Wrist	Finger
Run 1	51.18	76.25	26.58	19.71	9.69	11.17	7.89	3.21	1.19
Run 2	124.62	90.28	47.28	5.25	6.13	7.70	5.80	2.10	1.22
Run 3	119.48	137.33	59.35	6.48	6.08	7.89	5.01	3.21	1.71
Average	98.43	101.29	44.40	10.48	7.30	8.92	6.23	2.84	1.38

Table E.33: Overview of the run-level and averaged parameters for Subject 6, NMS: Compliant, Task: M

		G_{bdft}			ω_{bdft}			ζ_{bdft}	
	Elbow	Wrist	Finger	Elbow	Wrist	Finger	Elbow	Wrist	Finger
Run 1	75.90	112.50	41.39	3.51	4.38	7.56	1.10	1.62	1.85
Run 2	97.58	120.65	20.08	3.93	4.38	8.91	1.76	1.87	1.13
Run 3	67.92	105.86	25.70	3.53	3.81	6.29	1.08	1.31	1.12
Average	80.47	113.00	29.06	3.66	4.19	7.59	1.31	1.60	1.36

Table E.34: Overview of the run-level and averaged parameters for Subject 6, NMS: Stiff, Task: M

	G_{bdft}			ω_{bdft}			ζ_{bdft}		
	Elbow	Wrist	Finger	Elbow	Wrist	Finger	Elbow	Wrist	Finger
Run 1	76.69	107.05	37.02	7.95	7.13	10.00	4.45	4.01	2.80
Run 2	52.18	68.41	20.02	9.67	7.02	7.85	4.06	2.80	1.41
Run 3	55.30	81.80	21.08	6.98	6.08	6.80	2.49	2.77	1.46
Average	61.39	85.75	26.04	8.20	6.75	8.22	3.67	3.19	1.89

Table E.35: Overview of the run-level and averaged parameters for Subject 6, NMS: Compliant, Task: N

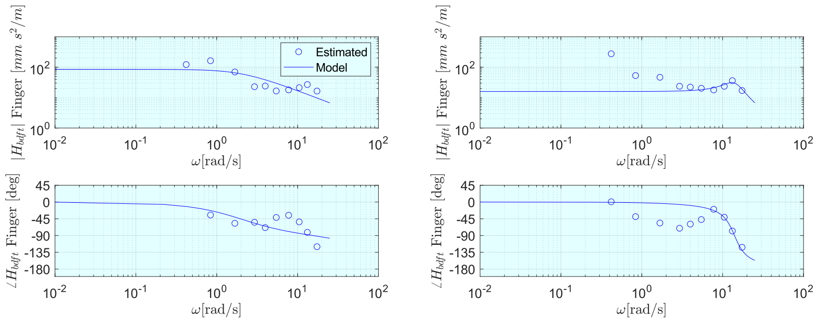
	G_{bdft}			ω_{bdft}			ζ_{bdft}		
	Elbow	Wrist	Finger	Elbow	Wrist	Finger	Elbow	Wrist	Finger
Run 1	108.85	71.44	76.37	35.61	7.56	7.82	22.71	1.82	1.45
Run 2	59.07	77.65	83.67	6.56	5.77	6.12	2.97	1.63	1.37
Run 3	80.89	91.09	87.47	4.69	4.90	5.48	2.25	1.47	1.26
Average	82.93	80.06	82.51	15.62	6.07	6.47	9.31	1.64	1.36

Table E.36: Overview of the run-level and averaged parameters for Subject 6, NMS: Stiff, Task: N

	G_{bdft}			ω_{bdft}			ζ_{bdft}		
	Elbow	Wrist	Finger	Elbow	Wrist	Finger	Elbow	Wrist	Finger
Run 1	75.11	76.69	63.31	6.23	6.99	8.53	2.75	2.16	2.00
Run 2	71.46	72.77	58.42	6.29	7.08	8.85	2.67	2.10	1.85
Run 3	63.54	62.72	47.85	7.26	7.59	8.81	2.81	1.93	1.33
Average	70.03	70.73	56.53	6.59	7.22	8.73	2.74	2.07	1.73

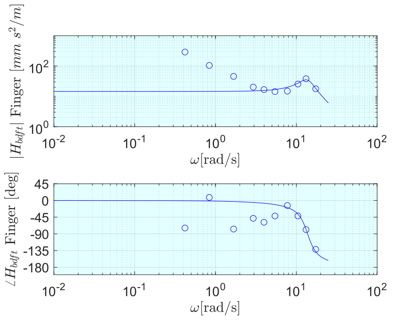
E.1. Discussion of the outcomes corresponding to Appendix E

The final values that are being plotted in the scientific paper are the run-averaged BDFT parameters. It was observed that in several cases, the parameter averaging might induce erroneous conclusions regarding the modeling outcomes and the biodynamic behavior of certain subject. One example is Table E.5, where the damping ratio for the finger landmark corresponding to the first trial is ten times higher compared to the value of the last two trials. Figure E.1 shows the raw BDFT estimates along with the response using the run-level models. It can be observed that for the last two trials, the model is capturing well the underdamped peak, whereas the response of the first model is not. The rather poor fit corresponding to the first figure is attributed to the initial values that are inputted in the maximum-likelihood estimator.



(a) Raw and modelled BDFT for Subject 1, NMS: Compliant, Task: N, Trial: 1

(b) Raw and modelled BDFT for Subject 1, NMS: Compliant, Task: N, Trial: 2



(c) Raw and modelled BDFT for Subject 1, NMS: Compliant, Task: N, Trial: 3

Figure E.1: Raw frequency-response-functions and modeled functions using the set of the initial values corresponding to the scientific paper.

When the set of initial parameter values is changed, it was possible to attain a better fit to the data corresponding to the first trial.

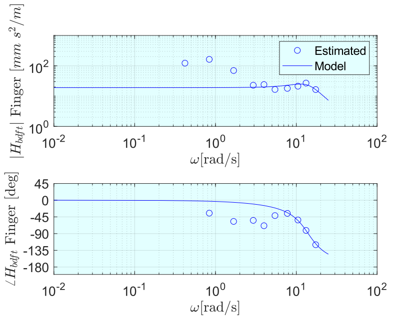
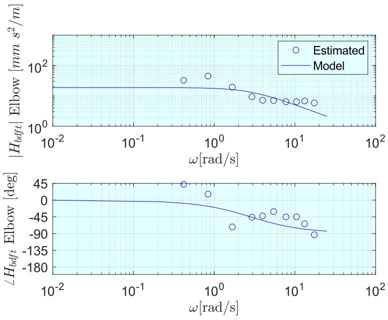


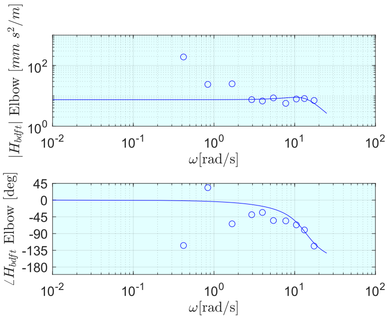
Figure E.2: Raw and modelled BDFT for Subject 1, NMS: Compliant, Task: N, Trial: 1 with the updated initial values.

E.2. Overview of the unrealistic values of the second-order system

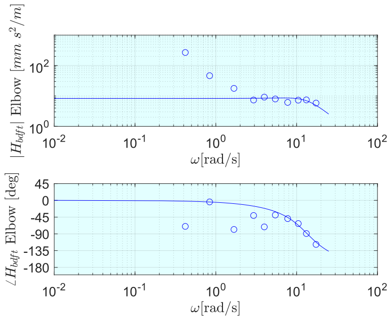
The highly unrealistic values for the elbow, wrist corresponding mainly to Subject 1 and Subject 5 can be attributed to the dynamic effects not being totally visible in the range of frequencies that are being measured, as presented in the figure below. In this way, fitting a second-order system might introduce artefacts in the parameter estimation process.



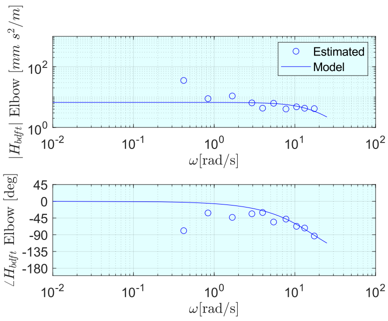
(a) Raw and modelled BDFT for Subject 1, NMS: Compliant, Task: F, Trial: 1



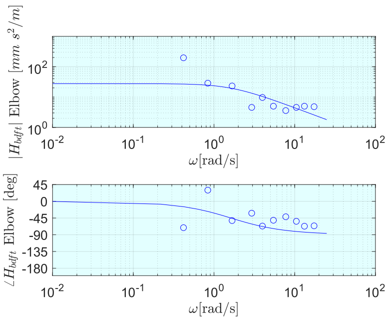
(b) Raw and modelled BDFT for Subject 1, NMS: Compliant, Task: F, Trial: 2



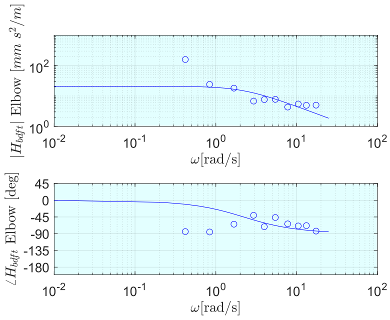
(c) Raw and modelled BDFT for Subject 1, NMS: Compliant, Task: F, Trial: 3



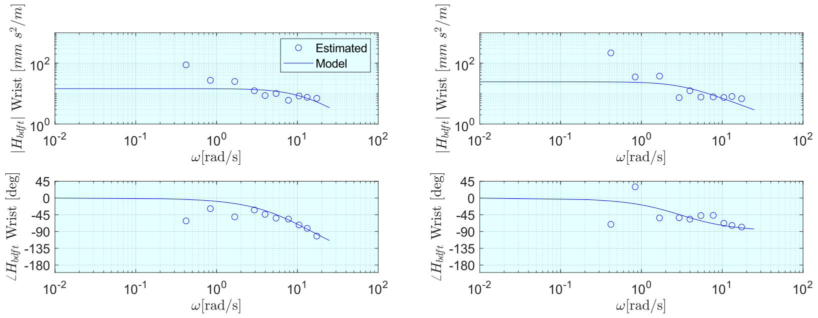
(a) Raw and modelled BDFT for Subject 1, NMS: Stiff, Task: F, Trial: 1



(b) Raw and modelled BDFT for Subject 1, NMS: Stiff, Task: F, Trial: 2

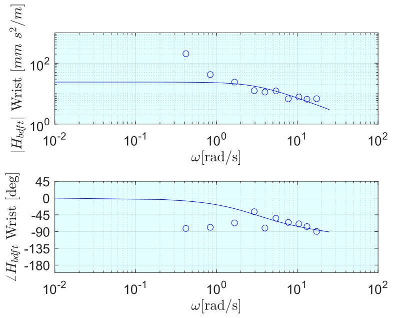


(c) Raw and modelled BDFT for Subject 1, NMS: Stiff, Task: F, Trial: 3

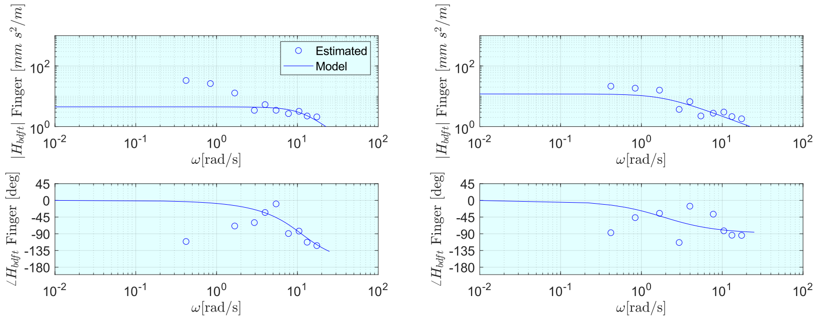


(a) Raw and modelled BDFT for Subject 1, NMS: Stiff, Task: F, Trial: 1

(b) Raw and modelled BDFT for Subject 1, NMS: Stiff, Task: F, Trial: 2

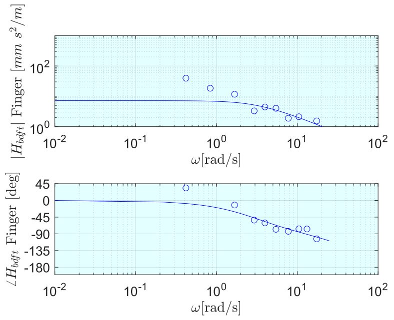


(c) Raw and modelled BDFT for Subject 1, NMS: Stiff, Task: F, Trial: 3

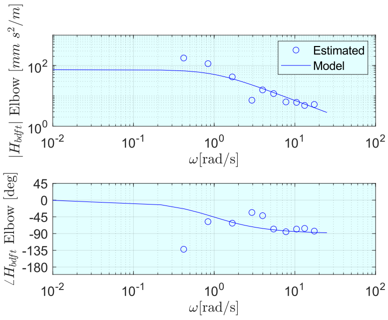


(a) Raw and modelled BDFT for Subject 1, NMS: Stiff, Task: F, Trial: 1

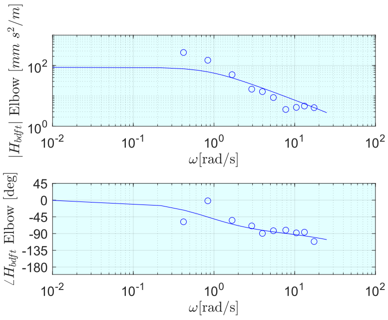
(b) Raw and modelled BDFT for Subject 1, NMS: Stiff, Task: F, Trial: 2



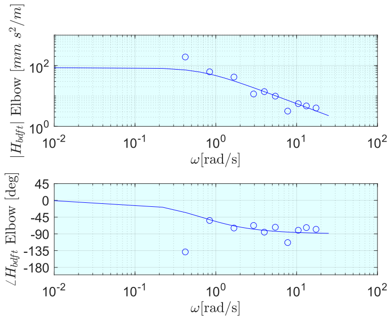
(c) Raw and modelled BDFT for Subject 1, NMS: Stiff, Task: F, Trial: 3



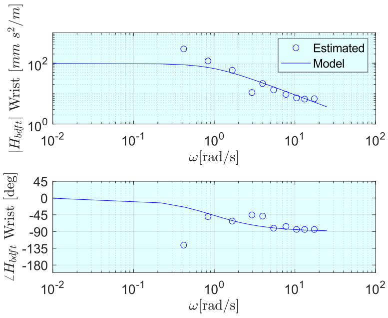
(a) Raw and modelled BDFT for Subject 1, NMS: Stiff, Task: M, Trial: 1



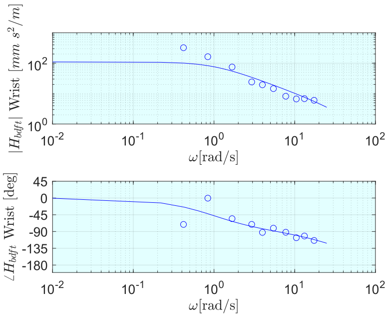
(b) Raw and modelled BDFT for Subject 1, NMS: Stiff, Task: M, Trial: 2



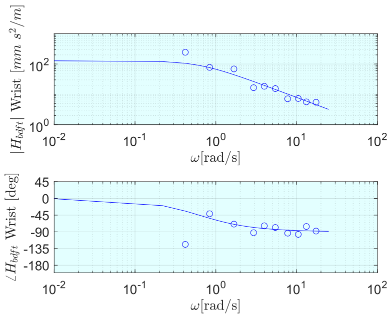
(c) Raw and modelled BDFT for Subject 1, NMS: Stiff, Task: M, Trial: 3



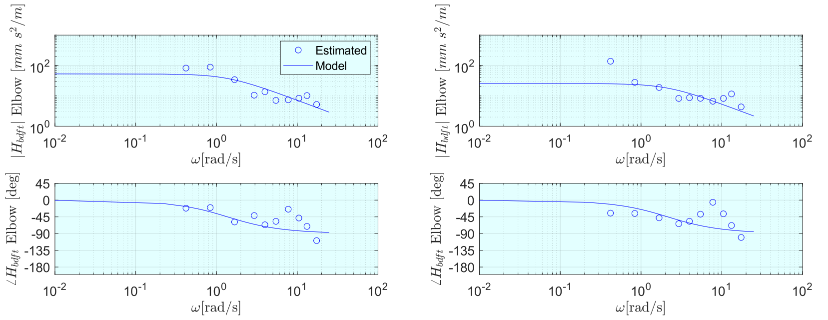
(a) Raw and modelled BDFT for Subject 1, NMS: Stiff, Task: M, Trial: 1



(b) Raw and modelled BDFT for Subject 1, NMS: Stiff, Task: M, Trial: 2

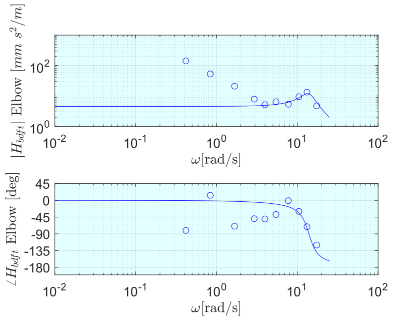


(c) Raw and modelled BDFT for Subject 1, NMS: Stiff, Task: M, Trial: 3

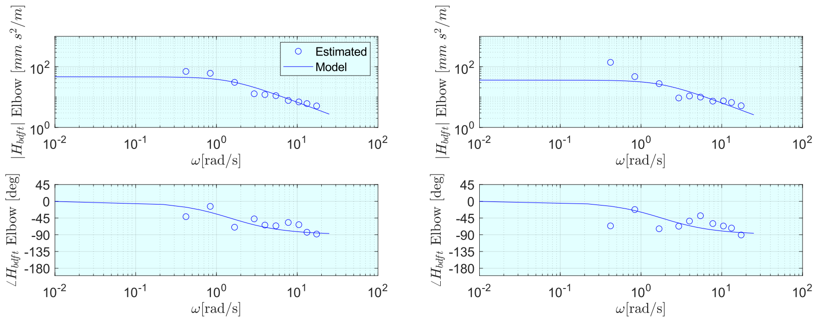


(a) Raw and modelled BDFT for Subject 1, NMS: Compliant, Task: N, Trial: 1

(b) Raw and modelled BDFT for Subject 1, NMS: Compliant, Task: N, Trial: 2

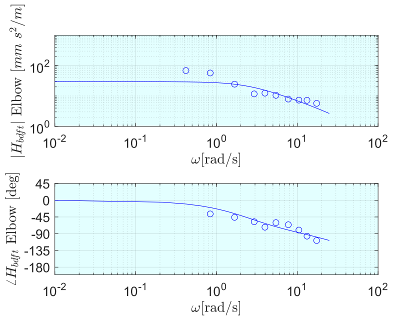


(c) Raw and modelled BDFT for Subject 1, NMS: Compliant, Task: N, Trial: 3

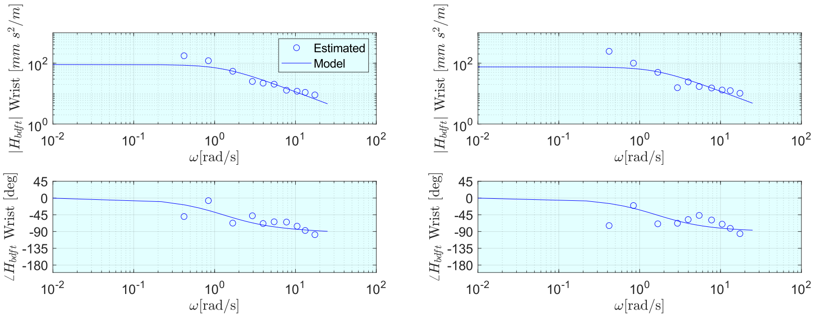


(a) Raw and modelled BDFT for Subject 1, NMS: Stiff, Task: N, Trial: 1

(b) Raw and modelled BDFT for Subject 1, NMS: Stiff, Task: N, Trial: 2

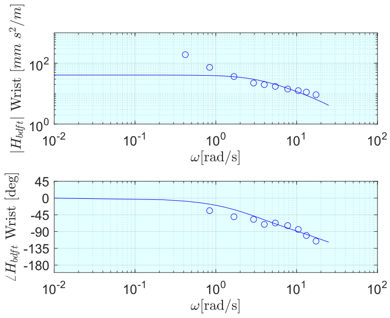


(c) Raw and modelled BDFT for Subject 1, NMS: Stiff, Task: N, Trial: 3

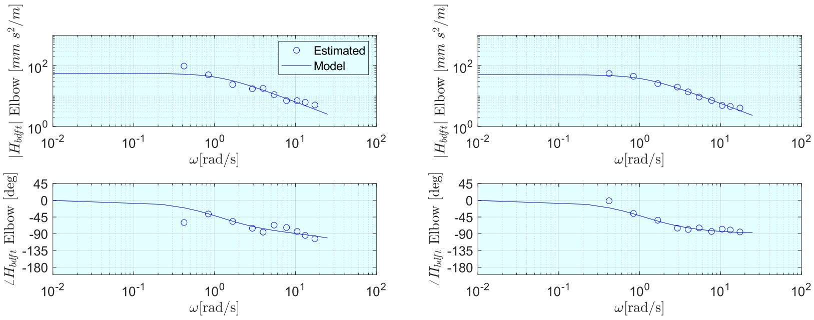


(a) Raw and modelled BDFT for Subject 1, NMS: Stiff, Task: N, Trial: 1

(b) Raw and modelled BDFT for Subject 1, NMS: Stiff, Task: N, Trial: 2

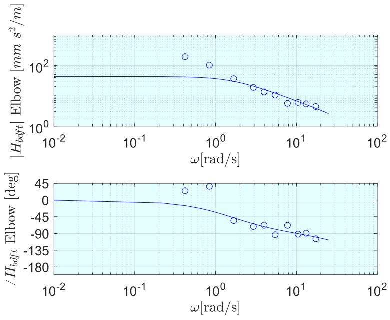


(c) Raw and modelled BDFT for Subject 1, NMS: Stiff, Task: N, Trial: 3

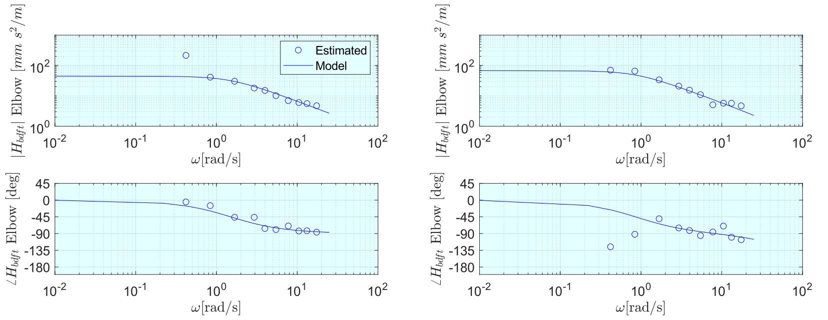


(a) Raw and modelled BDFT for Subject 5, NMS: Stiff, Task: F, Trial: 1

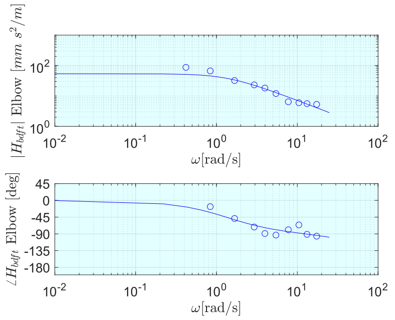
(b) Raw and modelled BDFT for Subject 5, NMS: Stiff, Task: F, Trial: 2



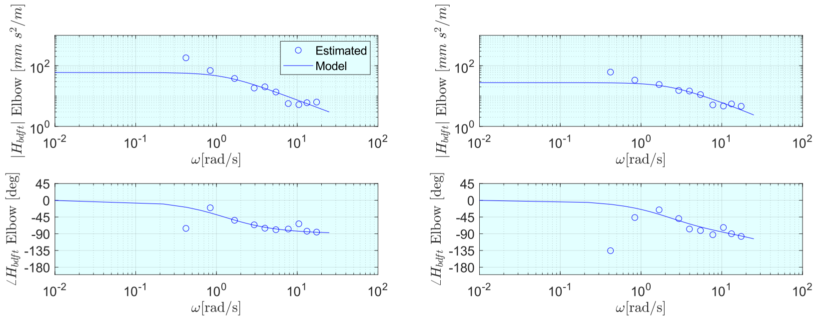
(c) Raw and modelled BDFT for Subject 5, NMS: Stiff, Task: F, Trial: 3



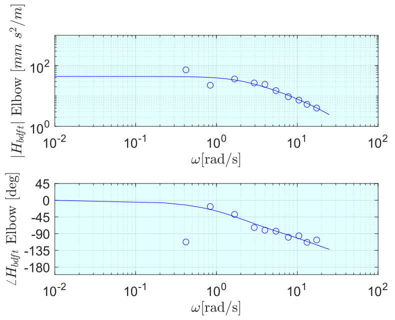
(a) Raw and modelled BDFT for Subject 5, NMS: Stiff, Task: M, Trial: 1 **(b)** Raw and modelled BDFT for Subject 5, NMS: Stiff, Task: M, Trial: 2



(c) Raw and modelled BDFT for Subject 5, NMS: Stiff, Task: M, Trial: 3



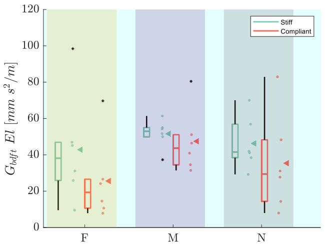
(a) Raw and modelled BDFT for Subject 5, NMS: Stiff, Task: N, Trial: 1 **(b)** Raw and modelled BDFT for Subject 5, NMS: Stiff, Task: N, Trial: 2



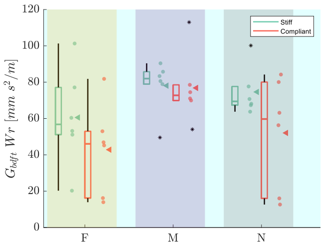
(c) Raw and modelled BDFT for Subject 5, NMS: Stiff, Task: N, Trial: 3

E.3. Parameter identificatio results using a new set of initial values

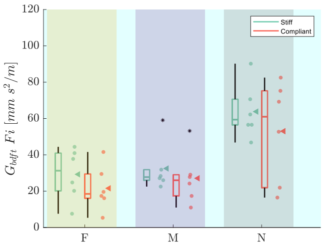
The following figures show the model parameters obtained using the initial parameter values for G_{bdft}, ω_{bdft} and ζ_{bdft}, [10, 14, 0.2] instead of [20, 6, 0.5] as used in the scientific paper.



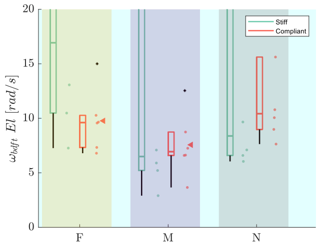
(a) Parameter identifications for the elbow.



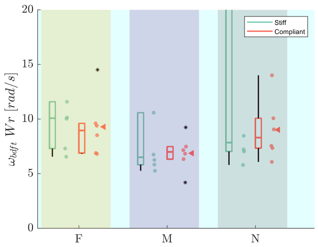
(b) Parameter identifications for the wrist.



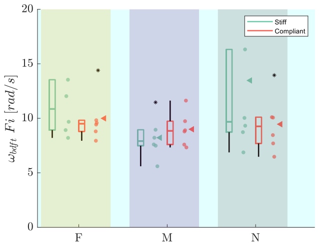
(c) Parameter identifications for the finger.



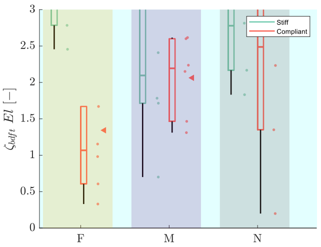
(a) Parameter identifications for the elbow.



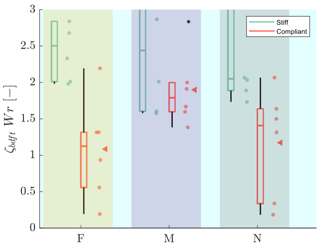
(b) Parameter identifications for the wrist.



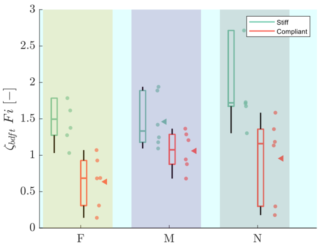
(c) Parameter identifications for the finger.



(a) Parameter identifications for the elbow.



(b) Parameter identifications for the wrist.

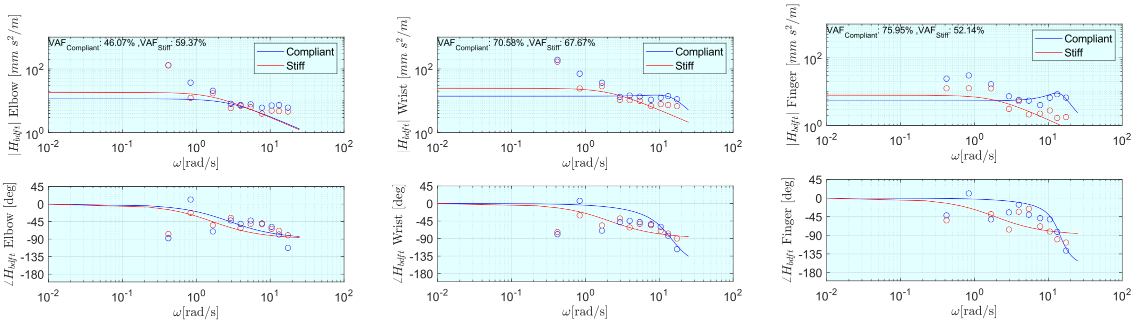


(c) Parameter identifications for the finger.

F

Raw and modeled BDFT frequency response functions

The following appendix provides an overview of the raw and modeled BDFT frequency response functions for all the subjects, tasks, neuromuscular settings and body parts.

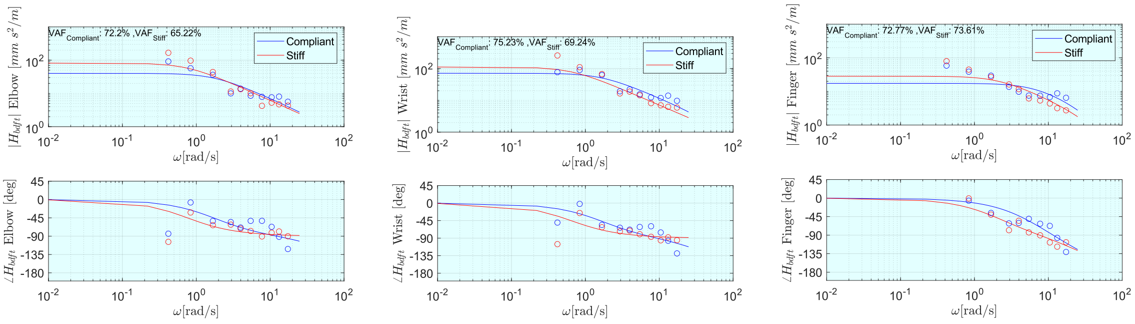


(a) Raw and modelled frequency response functions - elbow landmark*.

(b) Raw and modelled frequency response functions - wrist landmark.

(c) Raw and modelled frequency response functions - finger landmark*.

Figure F.1: Frequency response functions for Subject 1, NMS: Stiff and Compliant, Task: F

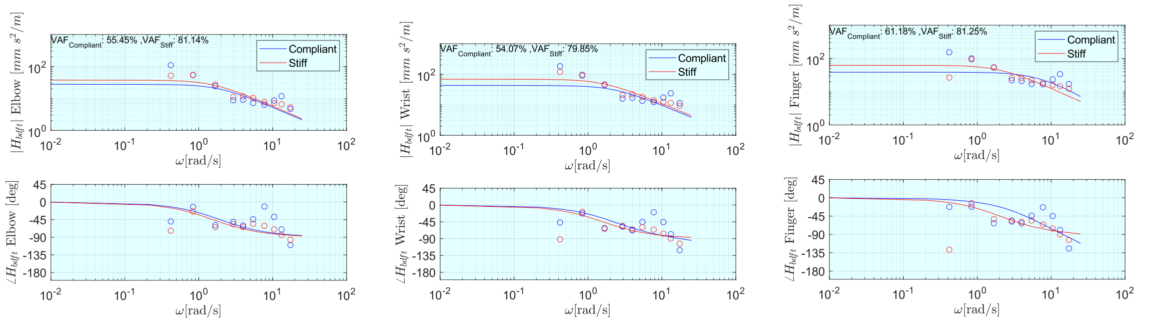


(a) Raw and modelled frequency response functions - elbow landmark*.

(b) Raw and modelled frequency response functions - wrist landmark*.

(c) Raw and modelled frequency response functions - finger landmark.

Figure F.2: Frequency response functions for Subject 1, NMS: Stiff and Compliant, Task: M

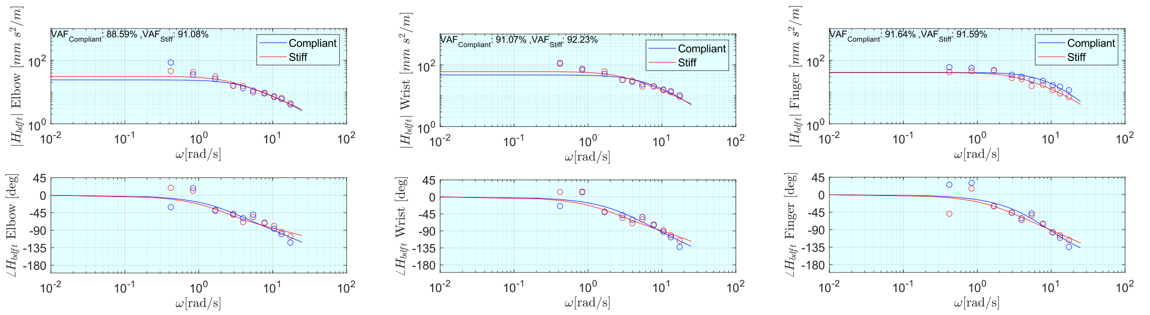


(a) Raw and modelled frequency response functions - elbow landmark*.

(b) Raw and modelled frequency response functions - wrist landmark*.

(c) Raw and modelled frequency response functions - finger landmark.

Figure F.3: Frequency response functions for Subject 1, NMS: Stiff and Compliant, Task: N

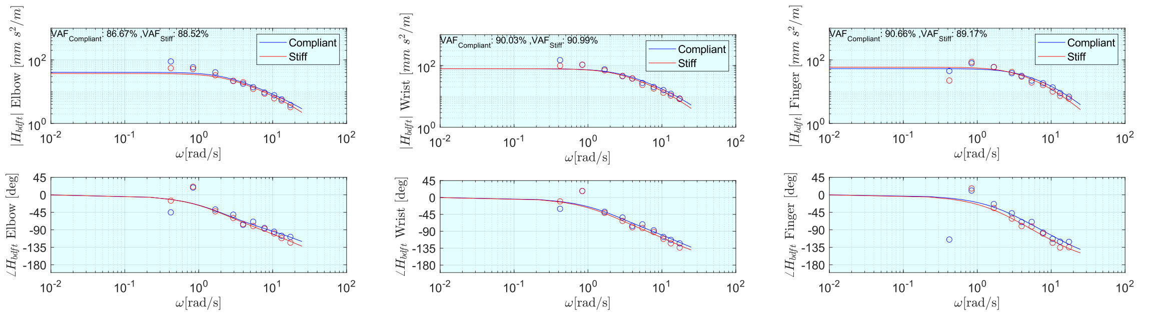


(a) Raw and modelled frequency response functions - elbow landmark.

(b) Raw and modelled frequency response functions - wrist landmark.

(c) Raw and modelled frequency response functions - finger landmark.

Figure F.4: Frequency response functions for Subject 2, NMS: Stiff and Compliant, Task: F

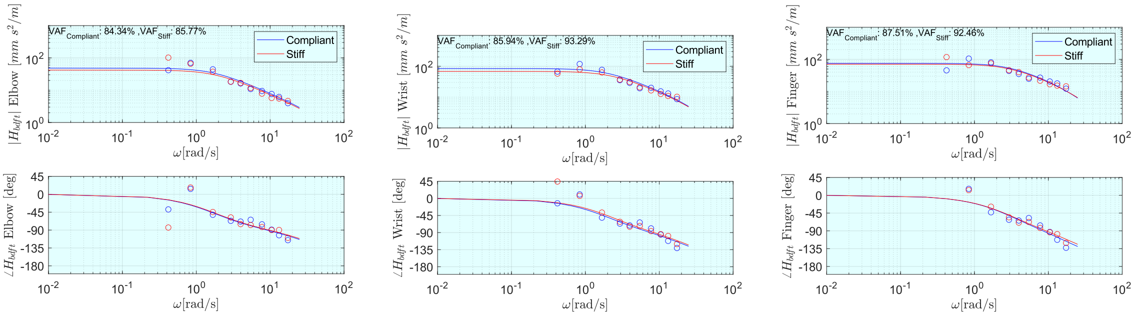


(a) Raw and modelled frequency response functions - elbow landmark.

(b) Raw and modelled frequency response functions - wrist landmark.

(c) Raw and modelled frequency response functions - finger landmark.

Figure F.5: Frequency response functions for Subject 2, NMS: Stiff and Compliant, Task: M

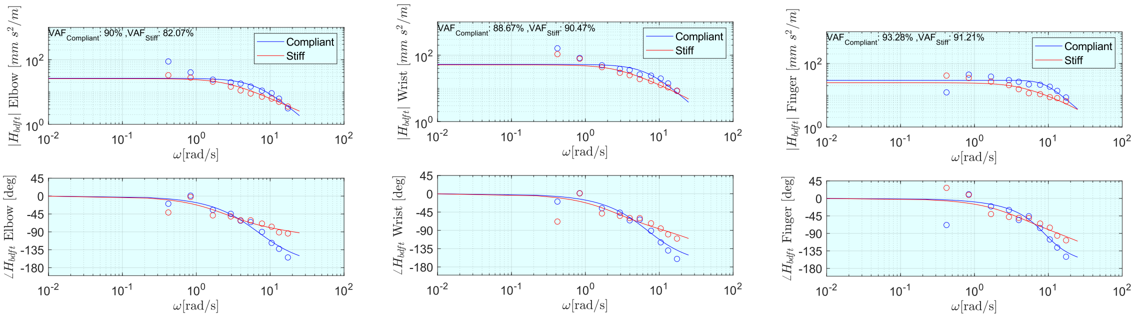


(a) Raw and modelled frequency response functions - elbow landmark.

(b) Raw and modelled frequency response functions - wrist landmark.

(c) Raw and modelled frequency response functions - finger landmark.

Figure F.6: Frequency response functions for Subject 2, NMS: Stiff and Compliant, Task: N

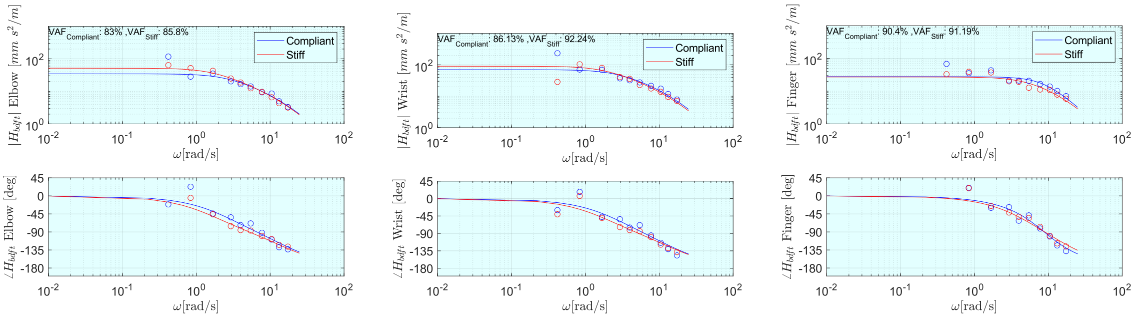


(a) Raw and modelled frequency response functions - elbow landmark.

(b) Raw and modelled frequency response functions - wrist landmark.

(c) Raw and modelled frequency response functions - finger landmark.

Figure F.7: Frequency response functions for Subject 3, NMS: Stiff and Compliant, Task: F

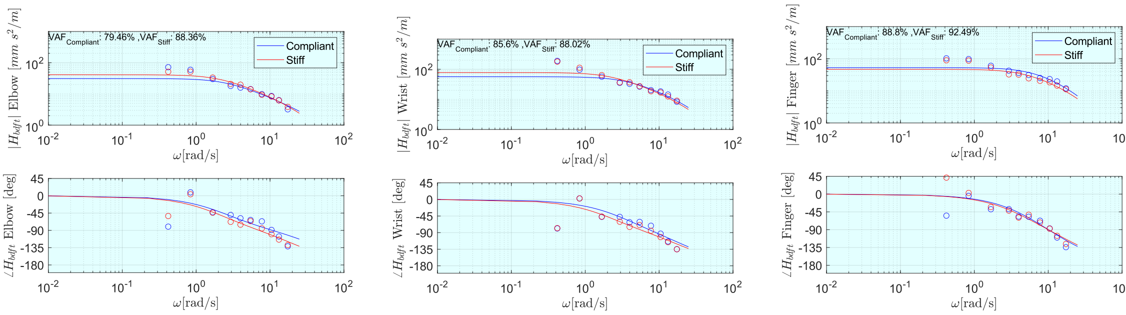


(a) Raw and modelled frequency response functions - elbow landmark.

(b) Raw and modelled frequency response functions - wrist landmark.

(c) Raw and modelled frequency response functions - finger landmark.

Figure F.8: Frequency response functions for Subject 3, NMS: Stiff and Compliant, Task: M

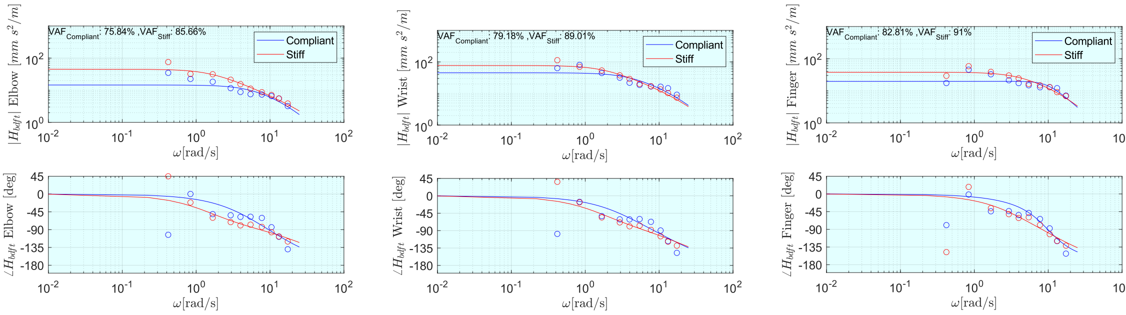


(a) Raw and modelled frequency response functions - elbow landmark.

(b) Raw and modelled frequency response functions - wrist landmark.

(c) Raw and modelled frequency response functions - finger landmark.

Figure F.9: Frequency response functions for Subject 3, NMS: Stiff and Compliant, Task: N

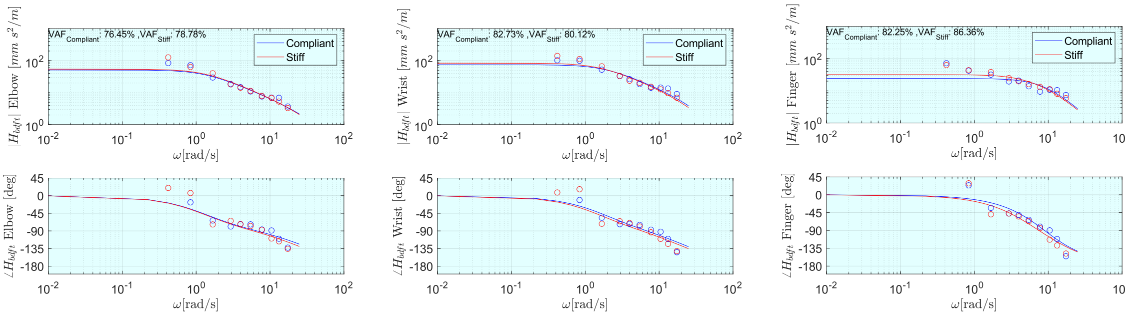


(a) Raw and modelled frequency response functions - elbow landmark.

(b) Raw and modelled frequency response functions - wrist landmark.

(c) Raw and modelled frequency response functions - finger landmark.

Figure F.10: Frequency response functions for Subject 4, NMS: Stiff and Compliant, Task: F

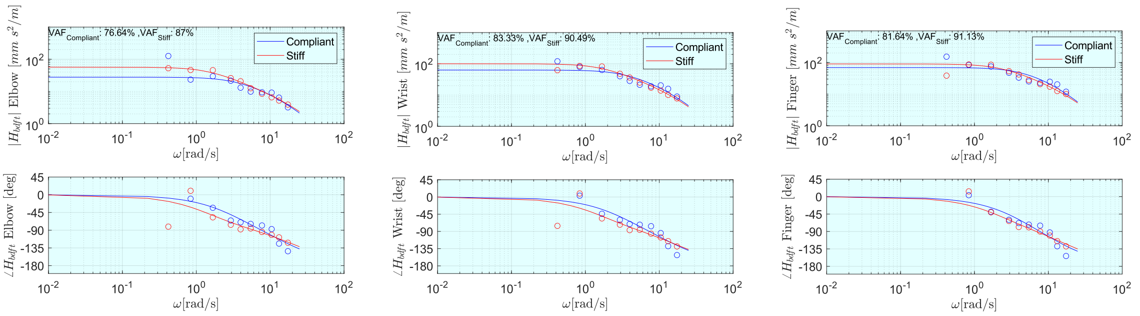


(a) Raw and modelled frequency response functions - elbow landmark.

(b) Raw and modelled frequency response functions - wrist landmark.

(c) Raw and modelled frequency response functions - finger landmark.

Figure F.11: Frequency response functions for Subject 4, NMS: Stiff and Compliant, Task: M

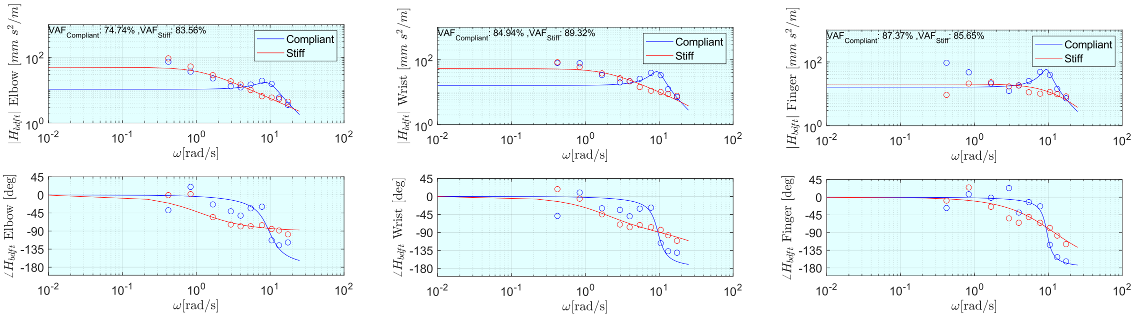


(a) Raw and modelled frequency response functions - elbow landmark.

(b) Raw and modelled frequency response functions - wrist landmark.

(c) Raw and modelled frequency response functions - finger landmark.

Figure F.12: Frequency response functions for Subject 4, NMS: Stiff and Compliant, Task: N

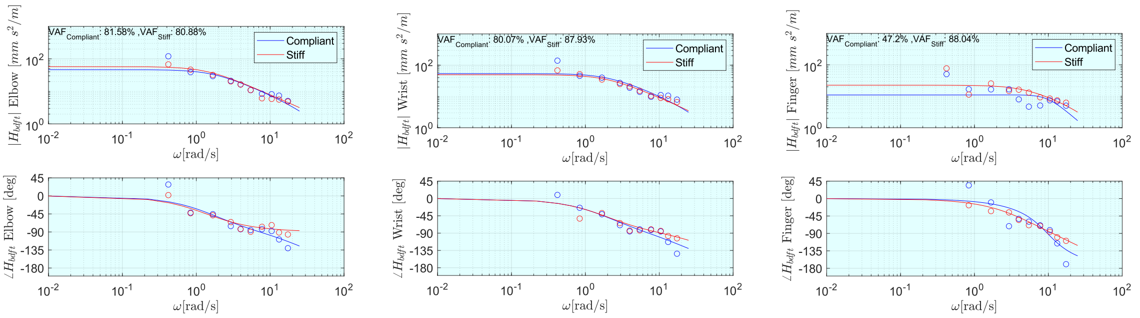


(a) Raw and modelled frequency response functions - elbow landmark*.

(b) Raw and modelled frequency response functions - wrist landmark.

(c) Raw and modelled frequency response functions - finger landmark.

Figure F.13: Frequency response functions for Subject 5, NMS: Stiff and Compliant, Task: F

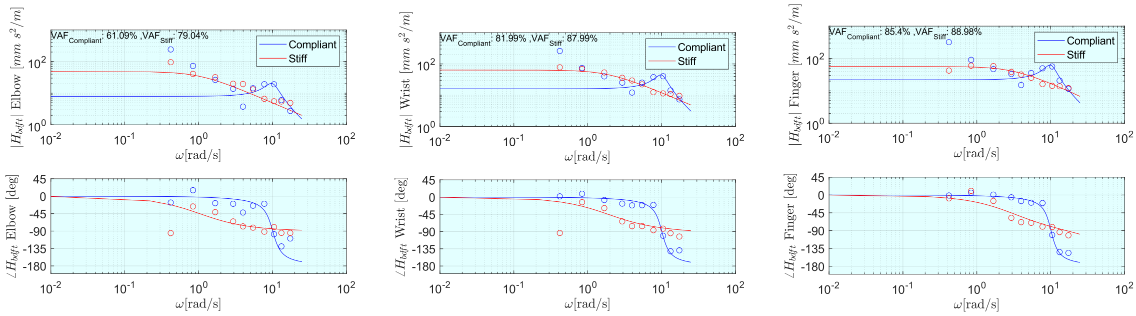


(a) Raw and modelled frequency response functions - elbow landmark*.

(b) Raw and modelled frequency response functions - wrist landmark.

(c) Raw and modelled frequency response functions - finger landmark.

Figure F.14: Frequency response functions for Subject 5, NMS: Stiff and Compliant, Task: M

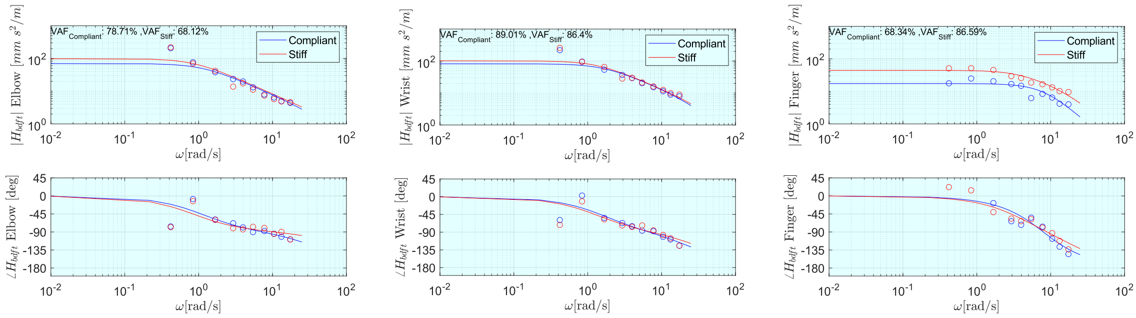


(a) Raw and modelled frequency response functions - elbow landmark*.

(b) Raw and modelled frequency response functions - wrist landmark*.

(c) Raw and modelled frequency response functions - finger landmark*.

Figure F.15: Frequency response functions for Subject 5, NMS: Stiff and Compliant, Task: N

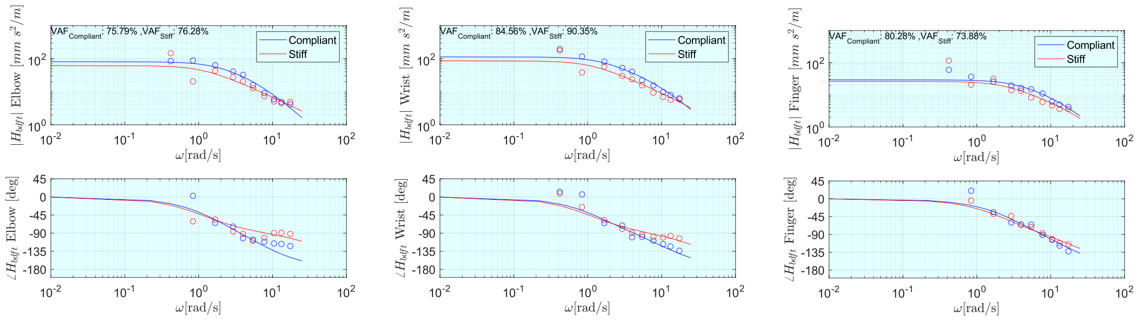


(a) Raw and modelled frequency response functions - elbow landmark.

(b) Raw and modelled frequency response functions - wrist landmark.

(c) Raw and modelled frequency response functions - finger landmark.

Figure F.16: Frequency response functions for Subject 6, NMS: Stiff and Compliant, Task: F

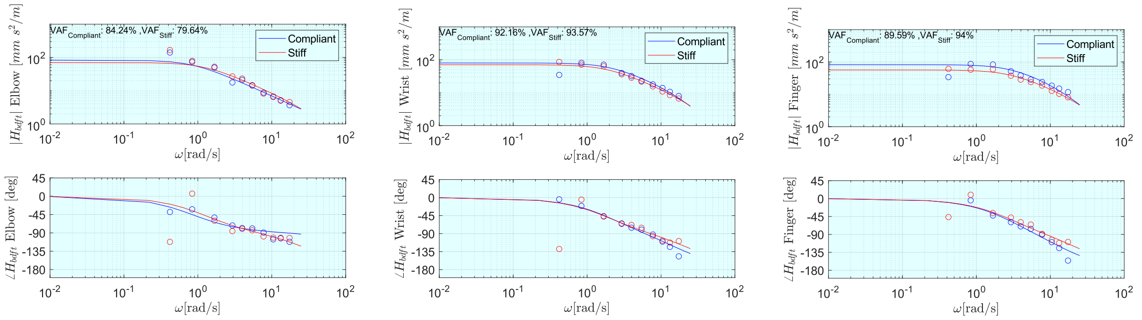


(a) Raw and modelled frequency response functions - elbow landmark.

(b) Raw and modelled frequency response functions - wrist landmark.

(c) Raw and modelled frequency response functions - finger landmark.

Figure F.17: Frequency response functions for Subject 6, NMS: Stiff and Compliant, Task: M



(a) Raw and modelled frequency response functions - elbow landmark.

(b) Raw and modelled frequency response functions - wrist landmark.

(c) Raw and modelled frequency response functions - finger landmark.

Figure F.18: Frequency response functions for Subject 6, NMS: Stiff and Compliant, Task: N

G

BDFT model choice analysis

The following appendix provides a comparison between different feasible BDFT model structures that could be used to describe the BDFT effects. The goal of this section is to reason the choice of the model that is used in the scientific article:

G.1. Models used

In order to describe the BDFT effects, the following models were chosen:

”First order”

$$Model = G_{bdft}\frac{\omega_{bdft}}{(j\omega + \omega_{bdft})} \tag{G.1}$$

”First order + delay”

$$Model = G_{bdft}\frac{\omega_{bdft}}{(j\omega + \omega_{bdft})}e^{-j\omega\tau_{bdft}} \tag{G.2}$$

”Second order overdamped”

$$Model = G_{bdft}\frac{\omega_{bdft}^2}{(j\omega)^2 + 2\cdot\omega_{bdft}\cdot j\omega + \omega_{bdft}^2} \tag{G.3}$$

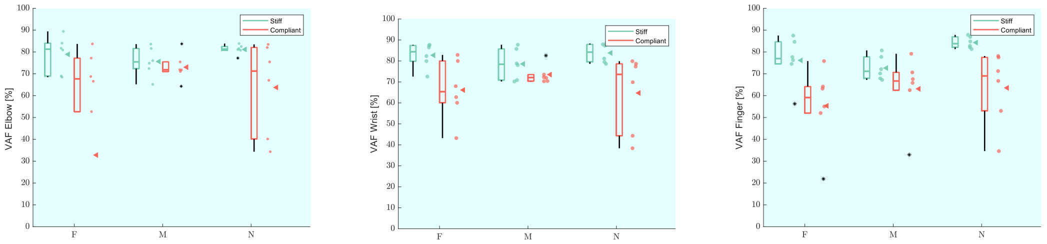
”Second order + zeta”

$$Model = G_{bdft}\frac{\omega_{bdft}^2}{(j\omega)^2 + 2\cdot\omega_{bdft}\cdot\zeta_{bdft}\cdot j\omega + \omega_{bdft}^2} \tag{G.4}$$

G.2. VAF calculations

For each of the model structures, the VAF was computed:

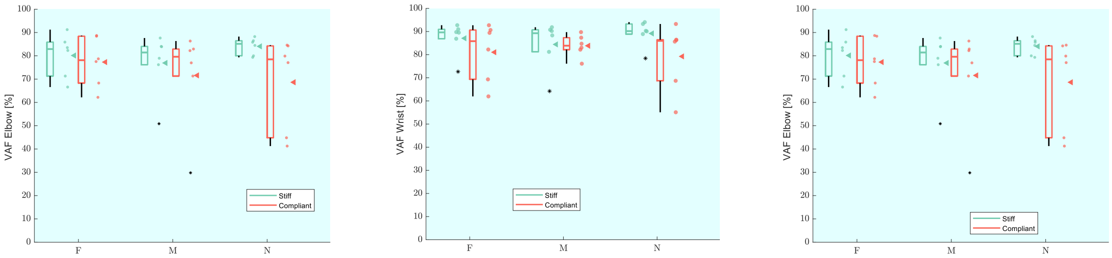
First order



(a) VAF of the elbow landmark. **(b)** VAF of the wrist landmark. **(c)** VAF of the finger landmark.

Figure G.1: VAF of the "first order" model.

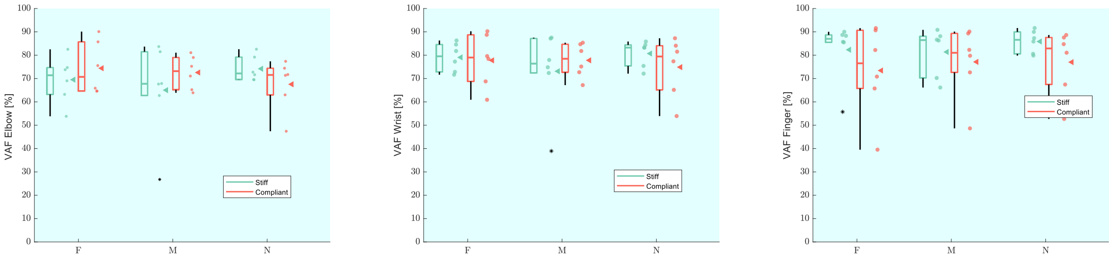
First order + delay



(a) VAF of the elbow landmark. **(b)** VAF of the wrist landmark. **(c)** VAF of the finger landmark

Figure G.2: VAF of the "first order + delay" model.

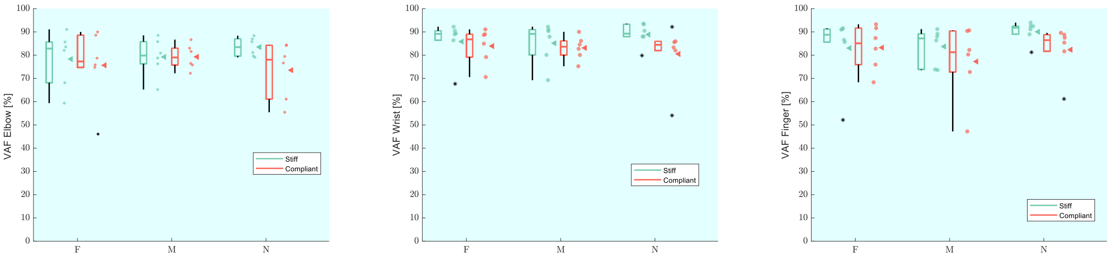
Second order



(a) VAF of the elbow landmark. **(b)** VAF of the wrist landmark. **(c)** VAF of the finger landmark

Figure G.3: VAF of the "second model overdamped" model.

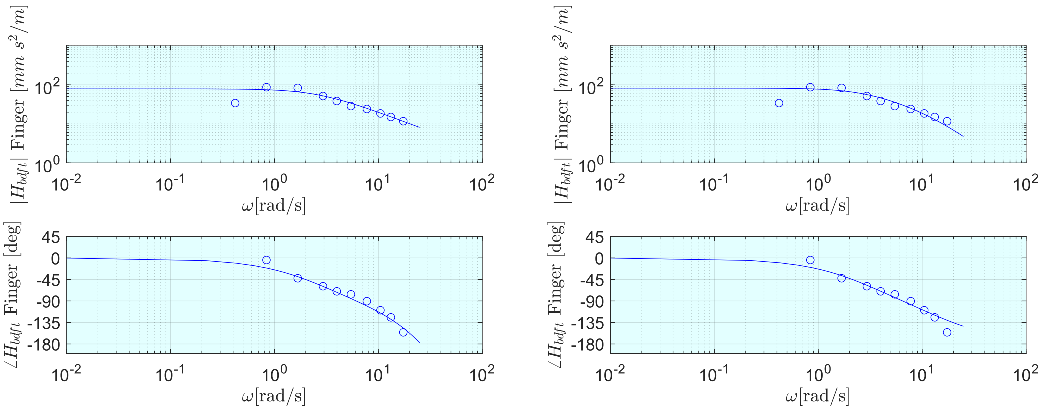
Second order + zeta



(a) VAF of the elbow landmark. **(b)** VAF of the wrist landmark. **(c)** VAF of the finger landmark

Figure G.4: VAF of the "second model + zeta" model.

The VAF results indicate that the two model structures that explain the best the variance of the signals are "first order + delay" and "second order + zeta". There are some scenarios where the "first order + delay" behaves slightly better than the "second order + zeta", one example being:

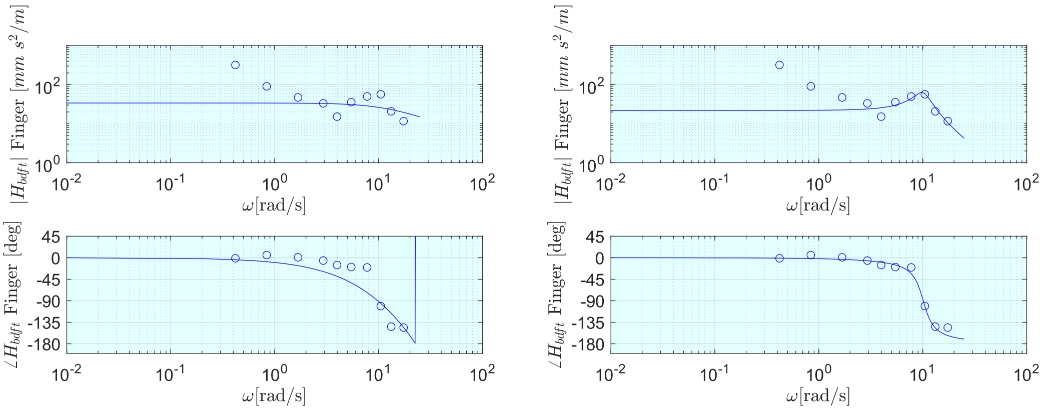


(a) Raw and modelled BDFT for Subject 6, NMS: Compliant, Task: N, VAF: 92.5%

(b) Raw and modelled BDFT for Subject 6, NMS: Compliant, Task: N, VAF: 89.6%

Figure G.5: Comparison between "first order + delay" and "second order + zeta" models.

On the contrary, there are also some scenarios where the "first order + delay" behaves worse better than the "second order + zeta", one example being:



(a) Raw and modelled BDFT for Subject 5, NMS: Compliant, Task: N, VAF: 92.5%

(b) Raw and modelled BDFT for Subject 5, NMS: Compliant, Task: N, VAF: 89.6%

Figure G.6: Comparison between ”first order + delay” and ”second order + zeta” models.

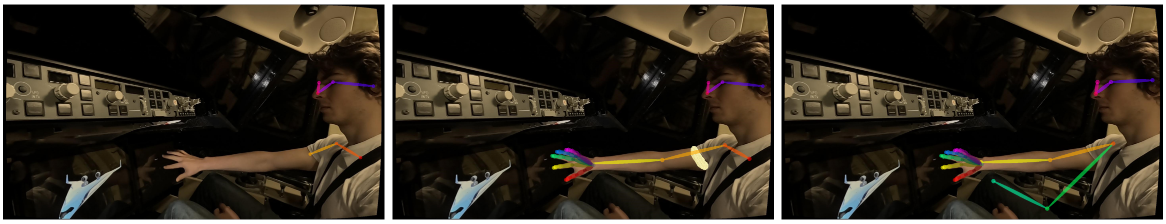
It was considered that the model ”second order + zeta” was able to describe the underdamped behavior of certain subjects, providing more robust modelling capabilities for the different behaviors that the subjects were exhibiting. In this way, the previously mentioned model structure was chosen in the scientific article.

H

Observations and limitations of the stereo markerless tracking

The following appendix provides an overview of the limitations of the approach selected in this study, based on the OpenPose markerless pose estimator and two stereo cameras.

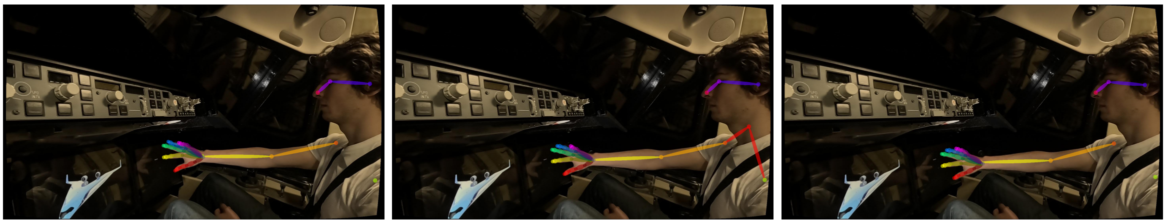
H.1. OpenPose failure modes and solutions



(a) Pose estimation failure, model - Cao et al., 2017. **(b)** Pose estimation failure mitigation, model - Cao et al., 2017. **(c)** Pose estimation failure mitigation, model - Hidalgo et al., 2019.

Figure H.1: OpenPose failure mode and mitigation procedures.

Figure H.1a presents one of the most common failure modes that were encountered during the experiment session. The uneven lighting that was observed on the upper limb, accentuated by the shade that was produced in the upper part of the limb due to the short-sleeved shirts that were worn by the subjects increased the possibility of artefacts from the OpenPose model proposed by Cao et al., 2017. The artefact consisted in a false-positive detection of the elbow, followed by an absence of the detection of the wrist and hand features. If the model from Cao et al., 2017 is to be used, a solution to mitigate the problem is to alter the original image, by introducing colors that eliminate the shading effects as in Figure H.1b. The option that was chosen in the current study is to mitigate the effects of the false-positive detections by using an updated model, the one proposed by Hidalgo et al., 2019 which, as it can be observed in Figure H.1c, is able to eliminate the false-positive detections without altering the original image.



(a) Elbow estimation at the beginning of the movement, model - Hidalgo et al., 2019.

(b) Elbow estimation in the second frame of the movement, model - Hidalgo et al., 2019.

(c) Elbow estimation in the second frame of the movement, model - Hidalgo et al., 2019.

Figure H.2: Possible sources of elbow measurement inaccuracies.

Figure H.2 presents three consecutive frames describing a ”downward” motion of the upper limb. It might be the case that in certain scenarios, the elbow detection and automatically the triangulated landmark induce unreliable results. This aspect can be observed especially when the elbow does not move significantly between two consecutive time-frames. It can be observed that despite the fact the elbow in Figure H.2b moved slightly lower than in Figure H.2a, the detected landmark did not follow exactly the downward trend. This phenomenon can be caused by the positioning of the camera with respect to the elbow, keeping in mind that the elbow is a larger body keypoint when compared to the wrist and the finger.

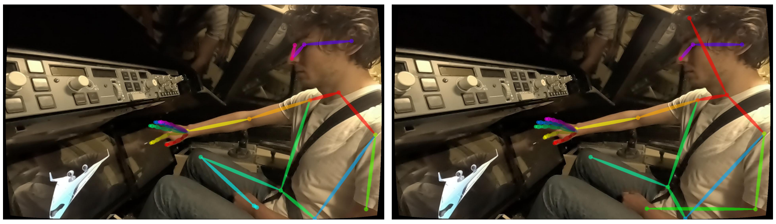


Figure H.3: The low lighting impact on the pose estimation.

Figure H.3 shows how low lighting affects the estimation of the hand landmarks. Preliminary tests were performed in order to analyse whether the simulator’s lighting good enough in order to ensure an accurate detection of the landmarks of interest. The outcomes based on the previously shown picture suggested the addition of an additional lighting source for the official experiment session.

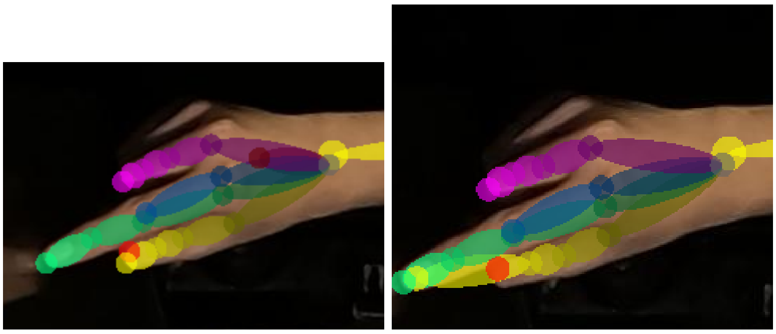


Figure H.4: Inaccuracies in the hand feature detection.

The subjects were instructed to keep their palm open during the experiment session, in order to avoid any possible artefacts that might occur in the pose estimation. Figure H.4

presents two of the possible artefacts if only the index finger would be shown to the camera. The reason for these effects is the probably limited training data of the hand feature estimator.

H.2. Effect of the synchronization on the reprojection error

As presented in the scientific paper, the synchronization of the cameras was performed by using a visual cue presented on the screen. Due to slightly different activation times of the cameras, it is highly probable that the two cameras are not perfectly aligned in time.

(a) Inaccurate camera synchronization - perspective 1.

(b) Inaccurate camera synchronization - perspective 2.

(c) Accurate camera synchronization - perspective 1.

(d) Accurate camera synchronization - perspective 2.

Figure H.5: Examples of slightly inaccurate and accurate camera synchronizations.

Figure H.5 presents two examples, showing a case where the synchronization frame between the two cameras is less accurate (the cameras ”see” slightly different points in time) and a case where the two perspectives capture the same time instance. It is expected that if two cameras record the scene at a small offset, a larger reprojection error of the triangulated points will be observed compared to the case where the cameras are perfectly in sync.

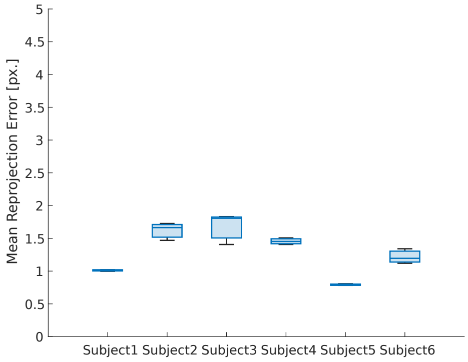


Figure H.6: Overview of the mean reprojection errors for Limb position: Fa.

To emphasize the effect of synchronization on the accuracy of results, the subject-level reprojection errors were analysed for the independent variable Limb position: Fa (one of the conditions that induces the largest speed of the hand). Figure H.6 shows that for subjects one and five, there is a considerably lower reprojection error compared to the others. When taking a closer look at the synchronizations for the two subjects, it was observed that the synchronizations were similar to the Figure H.5c and Figure H.5d indicating an accurate synchronization. Nevertheless the differences between the accurate and inaccurate synchronizations seem to be small in this scenario.