

POLITECNICO DI TORINO

Corso di Laurea in Ingegneria Biomedica

Tesi di Laurea

**Development of a Neural Interface
for Intuitive and Robust Control
of a Neuroorthosis in Daily Life
Using High-Density Surface
Electromyography**



Relatori

Prof. Luca Mesin

Prof. Alessandro Del Vecchio

Dominik Braun

Alberto GIAMMARIOLI

ANNO ACCADEMICO 2025-2026

Summary

Surface electromyography (sEMG) uses electrodes applied on the skin to capture the electrical activity of the muscles. One of its most common applications over the last decades has been myoelectric control: this electrical activity results in a signal that can be decoded and used as input for the control of an assisting device (e.g. prosthesis, orthosis, exoskeleton). The broader goal of the project to which this thesis belongs is the single-degree-of-freedom (1-DoF) myoelectric control of a hand orthosis for patients with spinal cord injury (SCI) causing impairment to their hands: capturing the residual activity of the forearm muscles, understanding whether the user wants to open or close the hand, and using that information to drive the orthosis actuation. The present thesis focuses on the decoding side of that pipeline: the design and experimental validation of an algorithm that, given the live forearm sEMG, reliably detects open–close intent while robustly suppressing false activations caused by non-target movements, all across a representative range of arm postures.

Despite many approaches in the literature reporting high accuracies, some of the persistent robustness challenges remain the limb-position effects — the alteration of the sEMG signal when the arm moves away from the calibration posture, causing instabilities and missed detections — and the false activations caused by non-target movements typical of Activities of Daily Living (ADL), such as wrist and arm movements. Most proposed solutions rely on inertial sensor fusion or large datasets for AI based frameworks; this work addresses both problems using only sEMG data, by adapting the Dynamic Time Warping algorithm with a runtime spatial correction (Spatial DTW) to a state machine framework, requiring a calibration of less than 15 minutes. Each subject recorded four templates per arm position for each of the two target gestures (closing and opening the hand) across the ADL positions. At inference time, the state machine uses the Spatial DTW to compute the distance between the real-time sEMG and the set of templates of the opposite class to the current state. If the distance goes below a per-subject threshold, then the change of state is triggered.

We validated the algorithm across ten most common ADL arm positions on 16 healthy subjects achieving a 98.4 % correct detection rate, 98.6 % rejection of pure wrist movements and 97.4 % rejection of false positives during ADL movements of

the arm (cohort F1 score = 98.8 %). As pilot studies, two subjects were retested one week later with the same calibration data showing minimal performance drift (F1= 99.4 % vs 99.8 % in their original session), while two SCI patients achieved 89.3 % correct detection rate (F1 score = 94.3 %). These results demonstrate that Spatial DTW can robustly control a state-machine, triggering state changes only in the presence of the target movement while ignoring everything else regardless of the arm position, suggesting Spatial DTW as a viable candidate for robust myoelectric control of hand assistive devices.

Acknowledgements

A mamma Barbara, papà Paolo e Viola,
che mi hanno sempre supportato, da vicino e da lontano

A nonna Lui, nonno Sergio, nonna Enza e nonno Tonino

Contents

List of Tables	7
List of Figures	8
List of Acronyms	14
1 Introduction	17
1.1 Neurophysiology and Electromyography	17
1.2 Muscle's Action Potential	19
1.3 Electrodes and Signal Acquisition	20
1.4 Spinal Cord Injury	23
1.5 Myoelectric Control	23
1.6 State of the Art - Literature Review	25
1.6.1 Dynamic Time Warping	27
1.6.2 MyoGestic Framework and CatBoost Classifier	28
1.7 Aim of the project	30
2 Materials and Methods	31
2.1 EMG Data Acquisition	31
2.1.1 Real-Time Acquisition	31
2.1.2 Filtering	32
2.2 Spatial Dynamic Time Warping Algorithm	34
2.2.1 Feature Extraction	34
2.2.2 Templates Extraction	35
2.2.3 Cosine DTW	36
2.2.4 Spatial Correction	38
2.2.5 The state-machine framework	40
2.3 Experimental Phase	43
2.3.1 Population	43
2.3.2 Laboratory Setup	43
2.3.3 Positions Overview	45
2.3.4 Templates Acquisition	45

2.3.5	Wrist Distractor Recording	47
2.3.6	Data for CatBoost	49
2.3.7	Model Creation	50
2.3.8	Testing Phase	51
2.3.9	Main Experiment	51
2.3.10	Wrist Movements False Positives Tests	53
2.3.11	Fingers Movements False Positives Test	55
2.3.12	The Mug-Holding Experiment	57
2.3.13	Tests on Patients	58
3	Results	59
3.1	Main Experiment Results	59
3.1.1	Latency	60
3.2	Robustness to Wrist Movements False Positives	62
3.3	Robustness to Finger Movements False Positives	64
3.4	Mug Holding Results	65
3.5	Temporal Stability	67
3.6	Patients Tests	69
3.7	Offline Results	69
3.7.1	CatBoost Comparison	71
3.7.2	Ablation Study	72
4	Discussion	85
4.1	Robustness Across Activities of Daily-Living	85
4.2	Comparison with States of the Art	87
4.3	Algorithm and Design Choices	88
4.4	Clinical Deployment	89
4.4.1	Limitations - Future Work	91
5	Conclusion	93

List of Tables

2.1	The ten arm positions used in the experimental phase, with the corresponding short name, ADL task each posture represents, and wrist orientation. Positions B, C, and E deliberately introduce a non-neutral wrist orientation to vary the forearm-rotation axis beyond what shoulder and elbow configuration alone provide.	46
2.2	Wrist calibration event sequence at position B (table reach). Each block consists of the six wrist distractor gestures interleaved with neutral rest periods. Between the two blocks an in-place pen→close transition is animated by the Control Hand. The order of the gestures is the same across subjects.	49
2.3	Pipeline of one round of the wrist-distractor test. Each round consists of four wrist-distractor blocks performed at two ADL positions, separated by in-place open↔close transitions (animated by the Control Hand) and one inter-position repositioning (performed freely by the subject). The block order is inverted between the two positions so that the subject's hand state is constant during the B→G repositioning.	55

List of Figures

1.1	Anatomical structure of a neuron	18
1.2	The action potential	19
1.3	Surface versus intramuscular EMG	21
1.4	Segmental organisation of the spinal cord	24
2.1	Sync-station containing four Muovi probes (in gray).	32
2.2	EMG bracelet with 32 electrodes (divided in two rows of 16).	32
2.3	Frequency response of the 4th-order Butterworth band-pass filter (20–500 Hz).	33
2.4	Frequency response of the notch filter bank (50 Hz and its harmonics up to 400 Hz).	33
2.5	Template extraction. The two one-second windows (red: closed template, blue: open template) are positioned automatically by the max-energy algorithm. The -200 ms shift makes sure the windows are placed before the beginning of the activation, capturing also the preparation for the movement. An important layer in the differentiation of the two templates is the intrinsic difference between this two activations, with different channels having higher and lower amplitudes, representing the different muscles of the fingers flexors and extensors.	36
2.6	Dynamic Time Warping alignment example	38
2.7	Saturating spatial correction function	39
2.8	The state-machine logic driving the final prediction. Whenever the distance between EMG and templates of the opposite class goes below the per-subject threshold, for five consecutive times, the change of state is triggered. In the real-time trials the algorithm was computing only the Spatial DTW distances to the opposite state; this figure also shows the distances with the same-state templates, drawn with dashed lines.	42

2.9	Laboratory setup: the subject was seated on the cube stool, with the laptop (on the right) showing the UI interface to control the experiment and plot the real-time EMG signal, and the supplementary monitor (left) which showed the Position Interface on the right side and the Virtual Hand Interface on the left side. On the desk, different sizes of the EMG-bracelets, and the reference electrode. . .	44
2.10	Virtual Hand Interface (VHI): the transparent control hand and the pink predicted hand, shown on the supplementary monitor.	44
2.11	The 10 ADL positions chosen for the experiment, which were shown to the subjects through the Position Interface.	46
2.12	The wrist position displayed in the Position Interface for wrist false positive calibration and the other wrist distractor experiment. . . .	48
2.13	Flowchart and setup of the main experiment. Left: subject setup, seated in front of the monitor displaying the Position Interface and the Virtual Hand Interface. Right: flowchart of the experiment, with the two rates being measured to extract the results.	54
2.14	The traditional Bavarian ceramic mug used in the mug-holding experiment.	58
3.1	Per-subject summary of the results of the main experiment. The middle column reports the number of transitions correctly detected out of all the events retained for the study, for each subject, pooling together all the positions.	61
3.2	Per-position summary of the results of the main experiment. The middle column reports the number of transitions correctly detected out of all the events retained for the study, for each position, pooling together events from different subjects.	62
3.3	Per-subject error breakdown. Each subject has two bars, one counting the errors due to missed events (red), one counting the instabilities (blue).	63
3.4	Per-position error breakdown. Each position has two bars, one counting the errors due to missed events (red), one counting the instabilities (blue).	63
3.5	Frequency error breakdown. Out of all the errors encountered in the main experiment (across all subjects), this pie diagram breaks down which were the most frequent errors.	64
3.6	Repositioning errors breakdown: how the repositioning errors were distributed across the subjects.	65
3.7	Repositioning errors pie diagram. Out of all the repositioning errors, this diagram shows which occurred during the OPEN and CLOSED states. The total number of events is 608.	66
3.8	Algorithm latency across subjects, computed only for correctly detected transitions.	66

3.9	Algorithm latency across positions. Values from all subjects are pooled per position to check whether detection was slower for certain postures.	67
3.10	Wrist distractors - per-gesture robustness for the OPEN state. . . .	67
3.11	Wrist distractors - per-gesture robustness for the CLOSED state. .	68
3.12	Wrist distractors - per-subjects robustness for the OPEN state. . . .	68
3.13	Wrist distractors - per-subjects robustness for the CLOSED state. .	68
3.14	Wrist distractors - False positives breakdown for the OPEN state. .	69
3.15	Wrist distractors - False positives breakdown for the CLOSED state.	70
3.16	Finger distractors - per-gesture robustness for the OPEN state. . .	70
3.17	Finger distractors - per-gesture robustness for the CLOSED state. .	71
3.18	Finger distractors - per-subjects robustness for the OPEN state. . .	71
3.19	Finger distractors - per-subjects robustness for the CLOSED state.	71
3.20	Finger distractors - False positives breakdown for the OPEN state. .	72
3.21	Finger distractors - False positives breakdown for the CLOSED state.	73
3.22	Mug experiment per subject breakdown. The table shows which activities did every subject attempt during the mug experiment. The tick means the subject did the activity with the prediction never mistakenly opening during the whole recording. The X means the algorithm mistakenly predicted an opening while the subject was doing that specific activity. If in a cell no tick is present it means the subject did not do that activity. On the last column, Passed* means the subject did not pass the experiment at first, but did pass it on a second trial.	74
3.23	Comparison of the performance in the second trial of the main experiment for S5 and S6, divided per position. The last row of each table reports the difference in false-positive rejection rate with respect to the first test. No repositioning errors occurred, matching the first experiment. The first event of the second trial of S5 happened when the subject wasn't ready therefore the event was discarded.	75
3.24	Comparison of the performance in the second trial of the wrist distractors experiment for S5 and S6, with the hand OPEN. The last row of each table reports the difference in false-positive rejection rate with respect to Fig. 3.10.	75
3.25	Comparison of the performance in the second trial of the wrist distractors experiment for S5 and S6, with the hand CLOSED. The last row of each table reports the difference in false-positive rejection rate with respect to Fig. 3.11.	76
3.26	Comparison of the performance in the second trial of the finger distractors experiment for S5 and S6, with the hand OPEN. The last row of each table reports the difference in false-positive rejection rate with respect to Fig. 3.16.	76

3.27	Comparison of the performance in the second trial of the finger distractors experiment for S5 and S6, with the hand CLOSED. The last row of each table reports the difference in false-positive rejection rate with respect to Fig. 3.17.	77
3.28	Latency comparison. Changes in the algorithm latency distribution between the second trials of the main experiment for S5 and S6. . .	77
3.29	Patients results. The protocol for patients was a simplified version, without changes of position. Each patient kept following the ground truth until declared to be fatigued. Left: breakdown of the prediction of all the cues the patient followed. Right: algorithm latency distribution in every event	78
3.30	CatBoost comparison, main experiment Correct Detection Rate per arm position, against the Spatial DTW baseline.	78
3.31	CatBoost comparison, wrist movements rejection rate per gesture, hand OPEN, against the Spatial DTW baseline.	79
3.32	CatBoost comparison, wrist movements rejection rate per gesture, hand CLOSED, against the Spatial DTW baseline.	79
3.33	CatBoost comparison, finger movements rejection rate per gesture, hand OPEN, against the Spatial DTW baseline.	79
3.34	CatBoost comparison, finger movements rejection rate per gesture, hand CLOSED, against the Spatial DTW baseline.	80
3.35	Spatial-correction ablation, main experiment Correct Detection Rate per arm position, against the full Spatial DTW algorithm.	80
3.36	Spatial-correction ablation, wrist movements rejection rate per gesture, hand OPEN, against the full Spatial DTW algorithm.	80
3.37	Spatial-correction ablation, wrist movements rejection rate per gesture, hand CLOSED, against the full Spatial DTW algorithm. . . .	81
3.38	Spatial-correction ablation, finger movements rejection rate per gesture, hand OPEN, against the full Spatial DTW algorithm.	81
3.39	Spatial-correction ablation, finger movements rejection rate per gesture, hand CLOSED, against the full Spatial DTW algorithm. . . .	81
3.40	Fraction of events, per cued arm position, for which the cued position's own templates were absent from the three closest templates used in the distance aggregation, separated by correctly detected events (true positives) and errors (missed and unstable). Note that several positions have very few error events (e.g. position D: 1/1, position I: 0 errors), so the percentages for these positions should be interpreted with caution.	82
3.41	Cohort Correct Detection Rate as a function of the number of templates retained per position and class, from the full four down to one.	82

3.42	ABFGI protocol, main experiment Correct Detection Rate per arm position, using templates from only five arm positions (A, B, F, G, I), against the full ten-position Spatial DTW algorithm.	83
------	---	----

List of Acronyms

ADL Activities of Daily Living. [23](#), [34](#), [45](#), [60](#), [65](#), [85](#), [86](#)

BCI Brain–Computer Interface. [24](#)

CDR Correct Detection Rate. [52](#), [85](#), [89](#)

CNS Central Nervous System. [17](#), [18](#)

DoF Degree of Freedom. [24](#), [25](#), [29](#), [30](#), [49](#), [56](#), [91](#)

DTW Dynamic Time Warping. [27–31](#), [34](#), [36–38](#), [89](#), [90](#)

EMG Electromyography. [17](#), [20–29](#)

GT Ground Truth. [43](#), [45](#)

IIR Infinite Impulse Response. [33](#)

IMU Inertial Measurement Unit. [27](#)

LDA Linear Discriminant Analysis. [26](#), [28](#)

MAV Mean Absolute Value. [26](#)

MU Motor Unit. [20](#), [21](#)

MUAP Motor Unit Action Potential. [20](#), [21](#)

PI Position Interface. [44–47](#), [49](#), [51](#), [53](#), [56](#)

RMS Root Mean Square. [26](#), [29](#), [34](#), [35](#)

SCI Spinal Cord Injury. [23](#), [28](#), [30](#), [69](#), [85](#), [89](#)

sEMG surface Electromyography. [20](#), [22](#), [23](#), [31](#)

SVM Support Vector Machine. [26](#)

UI User Interface. [44](#)

VHI Virtual Hand Interface. [43](#), [47](#), [51](#), [52](#), [55–57](#), [60](#)

WL Waveform Length. [26](#)

Chapter 1

Introduction

This chapter provides the theoretical background necessary to contextualise the work presented in this thesis. It begins with the neurophysiological basis of electromyography and its use in myoelectric control, then reviews the state of the art in [Electromyography \(EMG\)](#)-based hand control with particular attention to the robustness challenges that motivated this work. The chapter concludes with a description of the MyoGestic framework in which the algorithm was developed and a statement of the specific aims of the thesis.

1.1 Neurophysiology and Electromyography

Every voluntary movement we perform physiologically originates in the [Central Nervous System \(CNS\)](#), in the motor cortex of the brain; here neural commands are generated and transmitted downward through the spinal cord and the peripheral nerves, until they reach the muscles we want to contract (1). These neural commands take the form of an electrical signal, which results from the electrical gradient caused by the ionic currents of sodium (Na^+) and potassium (K^+) between the inside and outside of each cell.

To fully understand the mechanism and where the electricity comes from, it is necessary to first understand what a neuron is. Neurons are the cells nerves are made of, and they are specialised in the long-distance transmission of those currents. Fig. 1.1 shows its anatomical details. At rest, the interior of a neuron is maintained at approximately -70 mV relative to the extracellular environment; this value is the resting membrane potential. This voltage is sustained by the selective permeability of the cell membrane to ions and by the continuous work of the sodium–potassium pump, which actively expels three Na^+ ions in exchange for two K^+ ions. This asymmetric transport maintains a higher sodium concentration outside the cell and a higher potassium concentration inside the cell. Since the membrane at rest is naturally more permeable to K^+ , potassium slowly leaks

outward due to its concentration gradient, leaving the interior progressively more negative. This charge stabilises at approximately -70 mV, the resting membrane potential. When a sufficiently strong excitatory input (in this case coming from the [CNS](#)) depolarises the membrane beyond a threshold, voltage-gated Na^+ channels (different from the sodium–potassium pump) open, allowing sodium to rush into the cell and driving the membrane potential toward positive values, reaching around $+40$ mV. Shortly after, Na^+ channels inactivate, while voltage-gated K^+ channels open, allowing potassium to rush out of the cell, and the membrane repolarises back toward -70 mV. Because K^+ channels are slower to close, the membrane briefly overshoots below -70 mV, producing a hyperpolarisation, before returning to rest. This whole mechanism is called the action potential, and it is the basis of the electrical signal propagated from the motor cortex of the brain to the muscles through the neurons. The temporal and spatial course of the action potential is shown in Fig. 1.2. Following an action potential, the portion of the membrane involved enters a refractory period during which it cannot be re-excited.

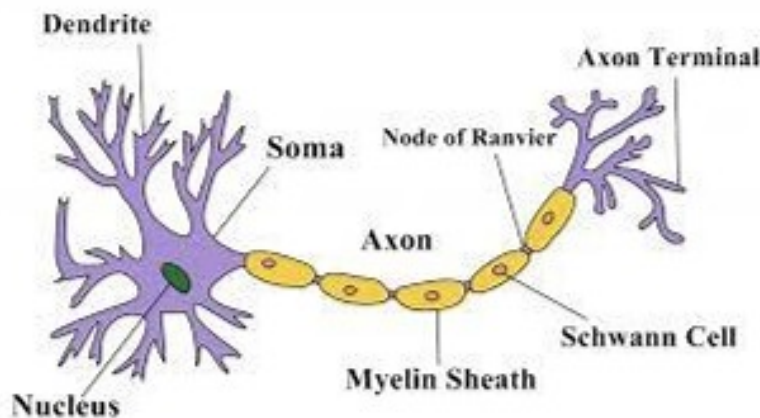


Figure 1.1. Anatomical structure of a motor neuron. The soma integrates incoming signals at the dendrites; the action potential propagates along the axon, whose myelin sheath (interrupted at the Nodes of Ranvier) enables saltatory conduction, until it reaches the axon terminals.

The ionic currents generated by an action potential act as input for the adjacent parts of the membrane, causing a cascade of other action potentials, meaning its propagation. The refractory periods ensure the potential propagates in only one direction. This mechanism is the reason how the action potential travel through the axon of the neurons. In myelinated motor neurons, this propagation does not occur continuously along the entire axon membrane; instead, the myelin sheath electrically insulates the axon, forcing the action potential to jump between the Nodes of Ranvier — the small gaps in the myelin where ion channels are concentrated (shown

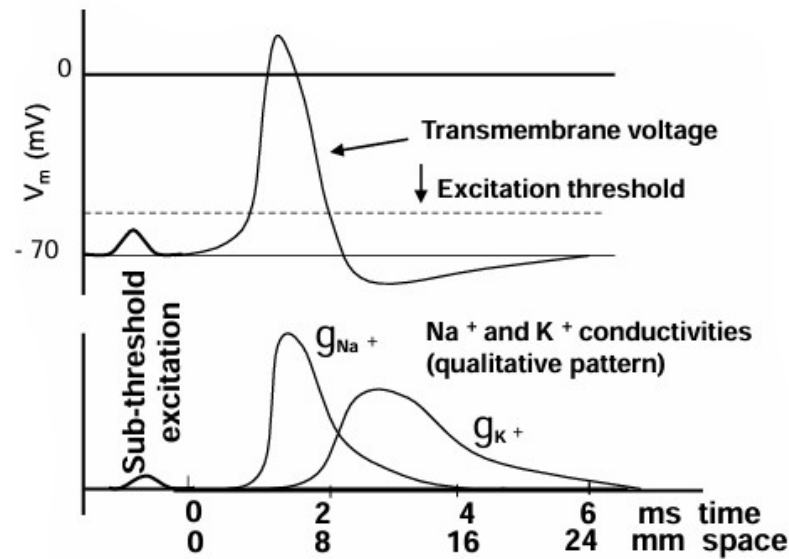


Figure 1.2. Time course of the transmembrane potential during an action potential (top) and the underlying Na^+ and K^+ conductances (bottom), from the Hodgkin-Huxley model (2). A sub-threshold excitation does not trigger firing; once the excitation threshold is crossed, the rapid rise in Na^+ conductance depolarises the membrane toward +40 mV, after which Na^+ inactivation and the slower K^+ conductance repolarise it, with a brief hyperpolarisation before rest.

in Fig. 1.1). This mechanism, called saltatory conduction, dramatically increases propagation speed while reducing the metabolic cost of signal transmission. In patients with multiple sclerosis, for example, the immune system attacks and degrades the myelin sheath (a process called demyelination) disrupting saltatory conduction and causing the characteristic neurological symptoms of the disease, from muscle weakness to loss of coordination (3). When the action potential reaches the axon terminal of a neuron, it is transmitted to the next neuron in the chain through the synapse: the action potential triggers the release of neurotransmitters from the pre-synaptic neuron, molecules which bind to the dendrites of the post-synaptic neuron, generating a new excitatory input in the following neuron and keeping the transmission alive.

1.2 Muscle’s Action Potential

At the end of the motor pathway, the action potential reaches the neuromuscular junction, the specialised synapse between the motor neuron and the muscle fibre. Here, the arriving action potential triggers the release of acetylcholine, a neurotransmitter which binds to the muscle fibre membrane and generates a new action

potential in the muscle fibre itself. The action potential in the muscle fibre does not only propagate electrically, it is also responsible for the mechanical contraction. The depolarisation of the fibre membrane causes the release of calcium ions (Ca^{2+}) into the cytoplasm, enabling the interaction between actin and myosin filaments that generates the mechanical force. At the interface between the nerves and the muscles, a single motor neuron innervates a number of muscle fibres. The complex formed by the motor neuron and the fibres it innervates is called a **Motor Unit (MU)**. A single muscle contains between 10 and 1000 motor units (1), each innervating between 50 and 1000 fibres. When the motor neuron fires (i.e. produces an action potential and transmits it to the muscle), all the fibres it innervates depolarise almost simultaneously, producing a characteristic electrical signal called a **Motor Unit Action Potential (MUAP)**. The brain modulates the force output produced by the muscles by controlling both the level of recruitment (i.e. the number and type of motor units recruited) and the firing rate, the frequency at which a given motor unit is activated.

1.3 Electrodes and Signal Acquisition

Measuring the electrical signal at the neuron level would require placing the acquisition source very close to the nerve, typically with an intramuscular needle electrode. Instead, since each motor neuron innervates many muscle fibres, the electrical activity recordable from the muscle surface is much more significant: a spatially amplified version of the central neural drive. Its amplitude, for example, reflects the degree of muscle activation and correlates with how much force is being produced (1). **surface Electromyography (sEMG)** therefore consists in placing a recording unit, called an electrode, on the skin above a muscle of interest, to record its electrical activity. An electrode is a transducer that converts the ionic currents flowing in the biological tissue into an electronic current measurable by the acquisition system. At the interface between skin and electrode, an electrochemical reaction occurs that generates a half-cell potential and a contact impedance, whose magnitudes depend on many factors, such as the electrode material, the skin preparation, and the contact area (1). This impedance introduces noise and distortion into the signal; to obtain a good-quality **EMG** signal it is therefore recommended to prepare the skin before placing the electrodes, typically by abrasion and cleaning, in order to reduce the contact impedance. Electrodes are commonly made of metal alloys; the most common materials are silver/silver chloride (Ag/AgCl), the standard for wet-gel electrodes, and gold or stainless steel for dry electrodes. The bracelet used in this work (HD20MM1602B, OT Bioelettronica, Torino, Italy) features gold-plated dry electrodes. When an electrode is placed on a muscle, it captures a weighted mixture of the electrical activity of all the motor units active within its pickup volume, with contributions decreasing with distance and filtered

by the tissue layers in between. This is analogous to placing a microphone in a room where many people are speaking simultaneously: the recorded signal is the superposition of all voices, with the closest ones to the microphone being dominant (the well-known “cocktail party” problem). In **EMG**, the different voices are the **MUAPs** of each motor unit. A single **MUAP** is essentially a spike, a short impulse, clearly visible in the intramuscular recording of Fig. 1.3 (blue); each motor unit fires at a rate that typically ranges from about 8 to 30 Hz (1), increasing with the level of force produced, and a normal contraction recruits a large number of motor units firing asynchronously. On the surface, these individual contributions overlap into the characteristic interference pattern shown in green in the same figure, which effectively records the combined electrical “sound” of the whole muscle.

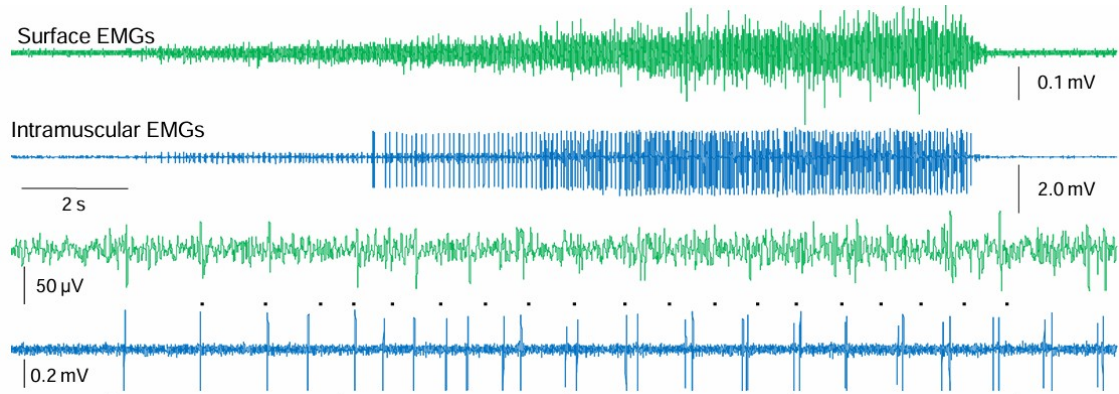


Figure 1.3. Comparison of surface (green) and intramuscular (blue) **EMG** recorded from the same muscle, shown at two temporal scales. At the coarser scale (top), the surface signal appears as the smooth interference pattern resulting from the superposition of many motor units, while the intramuscular signal is more selective. At the finer scale (bottom), the discrete discharges of individual motor units become visible in the intramuscular trace (dots mark successive firings of the same **MU**), each corresponding to a **MUAP**. Note the difference in amplitude scale between the surface (μV –0.1 mV range) and intramuscular (0.2–2 mV range) recordings.

The **EMG** signal is a voltage, and a voltage in electrical terms is never an absolute value but a difference of potential with respect to a reference. Every **EMG** recording therefore requires at least one active electrode and one reference. The reference can be internal to the acquisition device (e.g. the ground), or it can be an electrode as well, generally placed on an electrically neutral site such as a bony prominence with little muscle activity underneath. When one or more electrodes all measure the **EMG** signal with respect to a single common reference, the configuration is called monopolar. This configuration preserves the maximum spatial

information from each electrode independently, but is more susceptible to interferences that appear identically across all channels, called common-mode interferences. An alternative is the bipolar configuration, where instead of measuring each electrode against a common reference, the voltage difference is computed directly between two adjacent active electrodes. Since both electrodes are close to each other, they share most of the common-mode interference, which is cancelled by the subtraction. The bipolar configuration also acts as a spatial filter, improving the selectivity of the recording toward the portion of muscle directly underneath the electrode pair while reducing the contribution of distant sources. This property is known as crosstalk rejection, crosstalk being the unwanted presence in the **EMG** signal of activity from muscles further from the electrodes and not under study.

The raw **sEMG** signal at the surface is in the range of microvolts to millivolts, and is therefore highly susceptible to interference from external sources: above all the power line at 50 Hz, but also motion artefacts and electromagnetic interference. Electrodes are commonly paired to a differential amplifier, as part of the acquisition system, whose aim is to amplify the physiological signal (the difference between the active electrode and the reference, referred to as differential-mode) without amplifying the common-mode. The ability of the amplifier to reject common-mode signals is quantified by the common-mode rejection ratio (CMRR), which should be as high as possible for reliable **sEMG** recording (5). Once amplified, the analogue signal is converted to a digital representation by an analogue-to-digital converter (ADC). Once digitised, the signal can be inspected as a discrete time series representing the voltage recorded at each electrode over time. This is the raw **sEMG** signal, ready for further processing and analysis. The physiological **EMG** frequency band lies approximately between 10 Hz and 500 Hz. Therefore, to capture the whole activity without incurring aliasing, an acquisition system with a sampling frequency of at least 1000 Hz is required, as dictated by the Nyquist–Shannon sampling theorem (4). The probe used in this work (Muovi, OT Bioelettronica) samples at 2000 Hz. Despite the built-in common-mode rejection described above, it is common practice to apply filtering to the recorded **EMG**. To remove information outside the physiological band, band-pass filtering is usually applied to discard components below 10 Hz (typically movement artefacts) and above 500 Hz (typically white noise). To suppress the power-line interference specifically, notch filters are used, able to selectively remove the components at a single frequency (50 Hz in Europe).

Depending on the application, other processing steps may be applied. The **sEMG** signal carries rich information about the neuromuscular system and finds application across a wide range of contexts. In clinical neurophysiology, it is used to detect conditions such as neuropathies, myopathies, or motor neuron diseases. In biomechanics and sport science, the amplitude of the **sEMG** envelope is indicative of muscle activation timing and force level, as well as a means to study the onset of muscle fatigue. In rehabilitation engineering, **sEMG** can be used to detect motor intent, driving an external active assistive device such as a prosthesis, orthosis, or

exoskeleton. This last application is known as myoelectric control, and it is the focus of this thesis; it is therefore described in more detail below, but only after explaining why it is needed (1).

1.4 Spinal Cord Injury

Spinal Cord Injury (SCI) is defined as damage to the spinal cord resulting in temporary or permanent changes in its normal function (6). The spinal cord, as mentioned in the previous section, is the primary pathway for motor commands travelling from the brain to the muscles; when this pathway is interrupted or damaged, the consequences for voluntary movement control can be severe, depending on the level and position of the lesion. The spinal cord is organised into segments, and each segment innervates a specific region of the body. The segments are grouped into four regions: cervical (C1–C8), thoracic (T1–T12), lumbar (L1–L5), and sacral (S1–S5). Fig. 1.4 shows these different segments. The cervical segments innervate the neck, shoulders, arms, and hands; more particularly, C5 and C6 control the biceps and wrist extensors, while C7 and C8 control the triceps, wrist flexors, and finger muscles (6). The thoracic segments innervate the trunk and chest muscles, while the lumbar and sacral segments innervate the lower limbs. As a general rule, if the injury occurs at a certain level of the spine, signal transmission is compromised from that point downward; for instance, an injury at C4 will affect both arms and legs (tetraplegia), while an injury at T1 or below will spare the upper limbs, impairing only the lower ones (paraplegia). Injuries at the cervical level are unfortunately the most common, as well as the most severe. For patients with tetraplegia, impaired hand function is arguably one of the most disabling motor deficits, causing a substantial loss of independence in the common **Activities of Daily Livings (ADLs)** (7). The loss of the ability to grasp, release, and manipulate objects affects virtually every aspect of life, and is consistently ranked by patients themselves as one of the highest priorities for functional rehabilitation.

However, even when a person with **SCI** is unable to produce visible movement below the level of the injury, the lesion is rarely a complete anatomical transection of the spinal cord; instead, most injuries result from contusions that leave some spared neural connections (7). While this residual activity is insufficient to generate enough force to produce movement, it can still be detected using **EMG**. This spared neural activity therefore represents the bridge that makes myoelectric control a possible solution for the recovery of motor function and, eventually, gradual rehabilitation.

1.5 Myoelectric Control

Myoelectric control refers to the detection of the residual voluntary activity of a muscle through **sEMG**, used as input to drive an external assistive device and

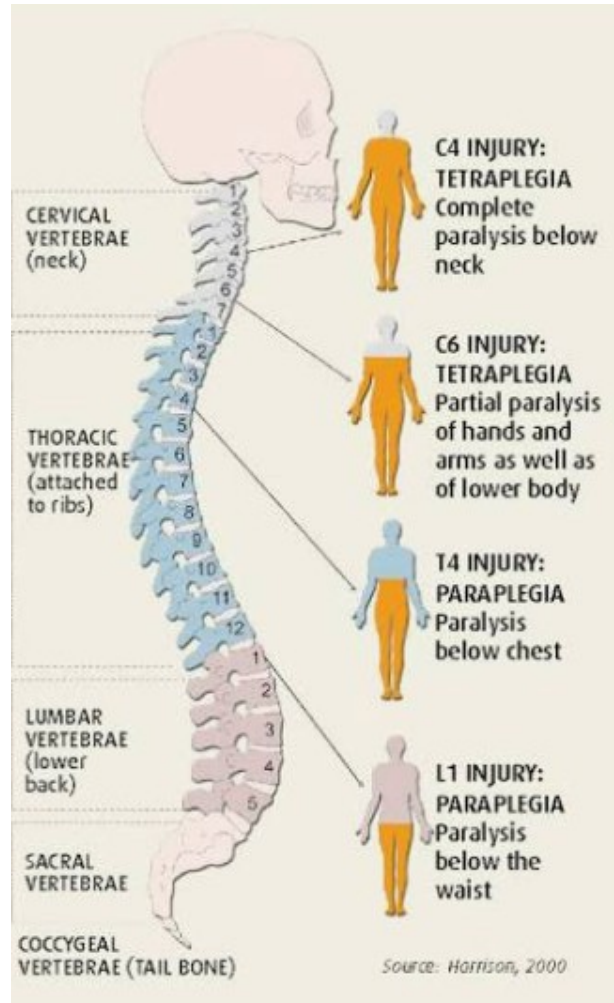


Figure 1.4. Segmental organisation of the spinal cord and the body regions innervated by each level (cervical, thoracic, lumbar, sacral). (8).

thereby bypass the motor impairment. This mechanism can be regarded as a form of [Brain-Computer Interface \(BCI\)](#), the muscle acting as an amplifier, the mirror of the brain's intention to move. The earliest and still most widely deployed commercial approach is direct control, also known as threshold-based or amplitude-based control (9). Here the envelope of the [EMG](#) signal, obtained by rectifying and low-pass filtering the raw signal, is compared against a threshold to determine whether the muscle is contracting or not. In the simplest configuration, two bipolar electrodes placed over two antagonist muscles (e.g. biceps and triceps) produce two independent control channels, each mapped to a movement of the prosthesis (e.g. flexion and extension). This approach is the most intuitive; however, the number of controllable [Degree of Freedoms \(DoFs\)](#) must equal the number of electrodes

(and thus the number of muscles) used: with two electrodes and two muscles, at most two independent commands can be generated, which is sufficient for the hand opening/closing problem but little more. An alternative strategy for obtaining more DoFs consists in using more than one amplitude threshold: with two thresholds, for example, the EMG amplitude range is divided into three segments, each associated with a movement of the prosthesis. This method is, however, less intuitive, as it requires the user to contract the same muscle at different force levels to perform different movements. Another type of myoelectric control uses pattern recognition. Rather than relying on the raw amplitude from a small number of channels, pattern recognition extracts features from multi-channel EMG recordings and trains a classifier to identify the intended movement from the spatial and temporal pattern of muscle activation across all channels (10). This approach can in principle distinguish a large number of different movements, allowing control over more DoFs of the actuator. However, pattern-recognition-based models have their own challenges: a classifier trained in one session may degrade in another due to electrode shift, fatigue, or changes in the user’s posture (11), and the mapping between EMG patterns and intended movements is not always stable across conditions. One of the most clinically relevant challenges for daily use, and the one directly addressed in this work, is the limb-position effect: EMG patterns recorded during a movement can change when the arm is held in a posture different from the one used during calibration, because of electrode shift and crosstalk from other muscles recruited in the new posture. A classifier trained to recognise an activation with the arm at the side might fail to recognise the same movement when it is performed with the hand raised overhead. This challenge, and the strategies proposed to address it, are reviewed in detail in the following section.

1.6 State of the Art - Literature Review

The history of myoelectric control for prostheses and orthoses spans over half a century, with the first commercial products appearing in the 1960s. The most established commercial approach, exemplified by devices from Ottobock, relies on the direct amplitude-based control discussed in the previous section (13). In recent years, mainly thanks to the continued and increasingly rapid development of artificial intelligence, pattern-recognition-based approaches have become more frequent. The improved quality of electrodes, and their ever-smaller dimensions, have made high-density sEMG (HD-sEMG) practical, acquiring more information and giving deep learning algorithms the opportunity to capture complex relationships within the data. The common practice for pattern-based control consists in dividing the real-time EMG into windows (each containing a fixed number of samples), computing one or more features for every window and every channel, and feeding the resulting multivariate vector to a classifier, which outputs a gesture

label. The extracted features are commonly in the time domain, typically [Root Mean Square \(RMS\)](#), [Mean Absolute Value \(MAV\)](#), [Waveform Length \(WL\)](#), or zero crossings. The whole pipeline described corresponds to supervised learning: it requires a labelled dataset of [EMG](#) recordings for each gesture to be discriminated, recorded during a calibration session. The calibration dataset is first fed to the classifier, which learns to associate feature patterns with gesture classes through a process called training. Then, at inference time, the trained model predicts the intended gesture from each incoming window in real time. Some classifiers require only the most recent window, and therefore have the window length as their inference latency; others require features from consecutive windows to obtain a wider temporal context of how the signal is evolving. The most historically prevalent classifier for this task is [Linear Discriminant Analysis \(LDA\)](#) which, assuming each class follows a Gaussian distribution, finds a linear boundary in the feature space that best separates the gesture classes, assigning each feature vector to the class whose distribution it most likely belongs to. [LDA](#) is computationally efficient, requires relatively little training data, and has demonstrated robust performance in controlled settings (14). Another well-established method is the [Support Vector Machine \(SVM\)](#), which finds the optimal hyperplane that maximises the margin, i.e. the gap between the boundary and the closest data points of each class (the support vectors), and is able to model non-linear class boundaries; for this reason it tends to outperform [LDA](#) when recording conditions vary, being more robust to violations of the underlying distributional assumptions. More recently, non-linear machine learning approaches such as deep neural networks and decision trees have been applied to myoelectric control, capturing more complex class boundaries than linear classifiers. Their ability to model complex, non-linear patterns in the training data has recently made machine learning the preferred choice of classifier. However, the downside is that these kinds of algorithms are, metaphorically, black boxes: given their complexity and depth, it is not always easy to understand the mathematical relationship between input and output, or to verify that the classifier has correctly learnt the intended patterns; when errors occur, it is correspondingly harder to identify their cause. Comparative studies suggest that the performance advantage of more complex models depends heavily on the amount of training data and the evaluation conditions; in matched, single-posture settings, even [LDA](#) achieves near-ceiling accuracy (14). Despite the variety of classifiers proposed, no single algorithm consistently dominates across all conditions; rather, performance depends on many factors simultaneously: the feature set, the number of gesture classes, the number of training repetitions, and, crucially, the conditions under which the classifier is evaluated. When conditions are matched (same posture, same session), many classifiers achieve similarly high accuracies. The real difference appears under mismatched conditions, such as posture changes, where robustness rather than peak accuracy becomes the critical metric (10). The ensemble of challenges currently inhibiting the real-world application of myoelectric

control is referred to as confounding factors (11). Confounding factors are activities, actions, or circumstances that are common in real life but difficult to reproduce in a laboratory context; for this reason, practitioners do not always account for them during offline performance analyses, partly owing to the increased data collection time and complexity involved. This explains why many algorithms achieve high offline performance while true online robustness has yet to be achieved (16).

This sensitivity to arm posture, known as the limb-position effect, has been documented since the early 2010s (15) and remains one of the primary barriers to the clinical deployment of pattern-recognition-based myoelectric control. In the review by Campbell et al., when natural variations in arm position were not accounted for during training, classification accuracy showed an average decrease of 12.8% (12). Proposed solutions to the limb-position effect broadly fall into three categories. The first is multi-posture training: including calibration data from multiple arm positions to expose the classifier to posture-induced variability (15). This improves robustness but substantially increases the calibration burden, a significant barrier for clinical populations. The second is sensor fusion: augmenting the EMG signal with Inertial Measurement Unit (IMU) data to provide explicit posture information, allowing the classifier to condition its predictions on the current arm orientation (15). This approach adds hardware complexity and requires the user to wear additional sensors. The third is self-calibrating or adaptive classifiers that update their parameters online during use, reducing drift over time without requiring explicit multi-posture calibration (10). Each of these strategies introduces trade-offs in calibration burden, hardware requirements, or algorithmic complexity. More recently, Eddy et al. demonstrated that Dynamic Time Warping (DTW) for discrete myoelectric control may be a robust choice for an algorithm that selectively detects the target movement while ignoring confounding factors that might otherwise falsely trigger the actuator (20); this solution is studied in detail in the next subsection.

1.6.1 Dynamic Time Warping

Dynamic Time Warping (DTW) is a sequence alignment algorithm originally developed for spoken word recognition (17); it computes the optimal alignment between two time series even when they differ in speed or timing, producing a distance measure that is small when the two sequences have a similar shape and large when they do not. Unlike the classifiers described above, which treat each window independently and assume the signal is stationary within it, DTW explicitly looks at the temporal evolution of the signal; this makes it particularly suited for detecting discrete, dynamic gestures with a characteristic temporal profile inside the continuous stream of activity and human movements. In the context of myoelectric control, DTW was studied in depth by Eddy et al. In a first work (11), they evaluated how robust different algorithms are to three real-world confounding factors (limb

position, cross-day use, and gesture elicitation speed), comparing DTW against a majority-vote LDA classifier (MVLDA) and two LSTM neural networks across a set of dynamic gestures, recorded with an 8-channel, 200 Hz forearm armband. When tested on unseen arm positions without any multi-posture training, DTW’s accuracy degraded only mildly (from 95.6% at the calibration posture to 86.7–92.7% across four unseen postures), while MVLDA collapsed far more severely (from 81.3% to 61.1–70.0%). The authors concluded that it is exactly DTW’s sensitivity to the temporal dynamics of a gesture, rather than to its static amplitude, that gives it an implicit robustness to posture changes, a property that frame-by-frame classifiers cannot exploit. In a second work (20), the same group applied DTW to the detection of a “wake gesture”: a finger snap used to toggle a myoelectric system between an active and a sleep mode, with the goal of avoiding false activations during everyday activities. During calibration, twenty one-second templates of the finger snap are recorded. At inference, a one-second window of live EMG is compared via DTW, every 50 ms, against each of the twenty templates, using the DTW alignment cost as the distance measure; the twenty distances are then averaged into a single value and compared against a threshold. If the average distance is below the threshold, the window is recognised as a finger snap. A majority vote over five consecutive predictions (250 ms) is used to confirm the detection and reduce false activations. The threshold itself is not arbitrary but is derived from the spread of the pairwise distances among the twenty templates (their mean plus a tunable number of standard deviations) and then fine-tuned on a one-minute recording of everyday activity to push the false positive rate toward zero. With this configuration, the system achieved a median false positive rate of 0% during activities such as walking, typing, writing, driving, and using a phone, at the cost of a false negative rate of 18.9% and 24.9% in the two evaluation tasks, meaning the user occasionally had to repeat the snap to toggle the system. This work takes the DTW template-matching idea as a starting point to achieve robustness and adapts it to the aim of this project; the specific design choices and a more detailed explanation of what DTW actually does will be addressed in Chapter 2.

1.6.2 MyoGestic Framework and CatBoost Classifier

MyoGestic is the open-source EMG interfacing framework developed at the n-squared lab of Friedrich-Alexander-Universität Erlangen-Nürnberg (18), and it is the platform on which the present work was built. It was designed to enable the rapid adaptation of machine learning models to the residual motor capabilities of individuals with neural lesions, and has been validated in real time on participants with SCI, stroke, and limb amputation, decoding multiple spared motor dimensions within a single clinical session. The framework provides a graphical user interface that orchestrates the whole pipeline: connecting to the EMG device, plotting the EMG signal in real time, recording the calibration data, training the

model, and running the real-time prediction. It is the same software used for all the experiments described in this thesis, into which the algorithm proposed here was integrated. MyoGestic was designed to be paired with a variety of [EMG](#) devices, so as to be a universal interface for [EMG](#) acquisition regardless of the hardware used. In the state of the art described in this section, the equipment used was the same as in this work: a probe paired with a 32-channel HD-EMG bracelet, further described in Section 2.1. The method developed by Sîmpetru et al. (18) consists in using a pattern-recognition-based classifier called CatBoost (19). CatBoost is an open-source supervised machine learning algorithm belonging to the family of gradient-boosting methods: it builds an ensemble of decision trees sequentially, with each new tree correcting the errors of the previous ones. It has been shown to be robust and accurate especially when trained on relatively little data (a crucial property in a clinical setting, where long calibration sessions are impractical for patients), and to perform particularly well on tabular data such as the [RMS](#) feature vectors used here. In the MyoGestic protocol, calibration consists of recording a short segment of [EMG](#) for each movement ([DoF](#)) the system has to decode, with a guiding virtual hand shown on screen; these recordings, approximately 30 seconds per movement, are then featurised to obtain the training data, using windows of 180 ms with a 9 ms shift. The user interface allows different time-domain features to be selected, though [RMS](#) remains the default and most commonly used choice in the framework (18). The labelled 32-dimensional feature vectors form the training set for the CatBoost classifier, which learns to map each vector to the corresponding hand state. At inference time, the trained classifier predicts the hand state from each incoming feature vector at the same 9 ms rate (a streaming rate of 111 Hz). When this thesis work began, the CatBoost classifier was the standard decoding algorithm of the MyoGestic framework, and it therefore represents the state of the art of the laboratory and the natural baseline against which the algorithm proposed in this thesis is compared. Importantly, however, the CatBoost classifier is a continuous controller (11), designed to map every window to an output, producing a constant stream of predictions; in addition, it was designed to decode as many degrees of freedom as possible in a way that minimises calibration burden, rather than to robustly detect discrete gestures and reject confounding factors, as [DTW](#) does. The Spatial DTW algorithm proposed in this thesis, by contrast, is a discrete state-machine controller built specifically around robustness to non-target activity and arm posture. A direct comparison between the two is therefore not entirely fair, as they pursue different goals; nonetheless, CatBoost is retained as the reference baseline, being the established decoding algorithm of the laboratory when this work began.

1.7 Aim of the project

Finally, the work presented in this thesis starts from the clinical need introduced above: a posture-robust myoelectric controller for a hand orthosis, intended for patients with [SCI](#) who have lost the ability to voluntarily open and close their hand. To address this problem in a simpler way the focus was only on single-degree-of-freedom (1-DoF); in this setting, the controller does not need to decode many degrees of freedom — it needs to do one thing reliably: detect the intention to open or close the hand, and never trigger when that intention is absent, regardless of arm position or of the other movements the user performs throughout the day. Whereas most myoelectric control research, including the CatBoost classifier of MyoGestic, aims to maximise the number of controllable [DoFs](#), the key insight motivating this work is that for an orthosis used in daily life an algorithm that is highly accurate across all arm positions and virtually never produces a false activation — even with a single degree of freedom — may be far more valuable than one that offers many degrees of freedom but cannot be trusted to stay silent when it should. Rather than building yet another classifier, this work therefore explores the possibility of a state-machine controller, in which the system occupies one of two states (hand open or hand closed) and changes state only when the corresponding target movement is detected. To drive the state machine, [DTW](#) was chosen, since detecting only the target movement while ignoring everything else requires a method sensitive to the temporal shape of the gesture rather than to the static amplitude of the signal; the algorithm is then extended with a runtime spatial correction mechanism specifically aimed at the limb-position effect.

Chapter 2

Materials and Methods

The software framework described in this chapter was implemented in Python (3.12) inside the open-source MyoGestic platform (18), which provides the graphical user interface through which the experimenter controls acquisition, calibration, training and real-time prediction, as well as the underlying device-interface and recording infrastructure. The Spatial DTW algorithm, the calibration and online protocols, and the analysis pipeline were developed for this thesis and integrated into the platform.

2.1 EMG Data Acquisition

SEMG was recorded using the HD20MM1602B 32-channel monopolar bracelet (2×16 array, OT Bioelettronica S.r.l., Torino, Italy), made of a flexible printed circuit board with dry gold-coated electrodes, paired with the Muovi probe (sampling frequency 2000 Hz; 16-bit resolution; OT Bioelettronica S.r.l., Torino, Italy). Since the bracelet operates in a monopolar configuration, a reference electrode was additionally required, generally placed in a bony part of the body where no significant electrical activity happens. The EMG signal captured by the electrodes is sent through a Wi-Fi connection from the probe to the laboratory laptop, where it can be visualised, processed in real time, and saved to disk using a custom script written in Python. No GPU is required in any step of the pipeline. More details regarding the experimental setup are provided in Section 2.3.2.

2.1.1 Real-Time Acquisition

The EMG signal acquired by the electrodes is sampled by the Muovi amplifier at 2000 Hz (2000 samples per second). However, the probe is not built to send every singular sample to the laptop in real-time; instead it buffers 18 samples per channel before transmitting each chunk wirelessly to the laboratory laptop over a Wi-Fi link via a Transmission Control Protocol (TCP) connection, resulting in an



Figure 2.1. Sync-station containing four Muovi probes (in gray).



Figure 2.2. EMG bracelet with 32 electrodes (divided in two rows of 16).

effective transmission rate of 111 Hz. Each chunk is first filtered and then appended to the continuous EMG signal saved to disk.

2.1.2 Filtering

Every 18-sample chunk arriving from the Muovi amplifier is passed through three real-time processing stages: a single-differential spatial filter, and two temporal filters, band-pass and notch.

Single Differential Spatial filter. The 32 monopolar channels of the bracelet are organised into a 16×2 grid (two rows of 16 electrodes running circumferentially around the forearm). A single-differential derivation is applied along the longitudinal axis of the bracelet by subtracting the proximal row from the distal row:

$$\text{diff}[i] = \text{ch}[i + 16] - \text{ch}[i], \quad i = 0, \dots, 15.$$

This step suppresses common-mode interference and rejects volume-conducted activity originating from deeper or more distant muscles, at the cost of halving the channel count from 32 to 16. From this point onward, all downstream stages consider the channel count to be 16 rather than 32.

Band-pass filter. A 4th-order Butterworth band-pass filter with cut-off frequencies of 20 Hz and 500 Hz is applied to each channel, to remove motion artefact and baseline drift at low frequencies and out-of-band noise at higher frequencies. The 20 Hz lower cut-off is conservative with respect to the physiological EMG band (typically 10–500 Hz) and was chosen to suppress residual motion artefact during arm repositioning, this being one of the main concerns of the framework.

Notch filter. A Butterworth notch is applied at the European power-line fundamental of 50 Hz and its next seven harmonics (100, 150, 200, 250, 300, 350,

and 400 Hz, i.e. all harmonics within the 0–500 Hz pass-band), with the aim of removing the noise caused by coupling between the subject and the power line at frequencies not covered by the band-pass filter. Since 10–500 Hz is the bandwidth of the physiological EMG signal, this harmonic notch filtering effectively removes part of the signal’s own information; however, in most everyday circumstances subjects are near electronic devices connected to the power line (e.g. lights, computers, TVs), and proximity alone can noticeably alter the EMG signal if left unfiltered. The choice of aggressive notch filtering is therefore a deliberate trade-off, favouring robustness across environments over the small loss of physiological information it entails.

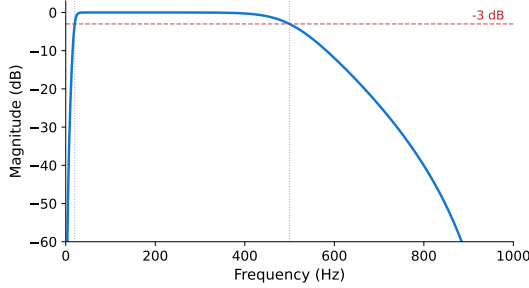


Figure 2.3. Frequency response of the 4th-order Butterworth band-pass filter (20–500 Hz).

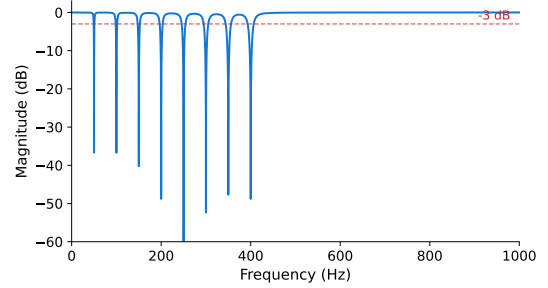


Figure 2.4. Frequency response of the notch filter bank (50 Hz and its harmonics up to 400 Hz).

Both the band-pass and the notch (showed in Fig. 2.3 and 2.4) are implemented as stateful [Infinite Impulse Response \(IIR\)](#) filters, which are causal filters, meaning the filter’s output at any instant depends only on the present and past samples, never on the future – a requirement for real-time filtering. [IIR](#) filters satisfy this requirement by computing each output sample as a recursive combination of past inputs and past outputs; the resulting memory is updated sample by sample and persists between consecutive 18-sample chunks, eliminating the per-buffer edge artefacts that a stateless filter operating on each buffer alone would otherwise introduce. The price of this recursive structure is a required warm-up time: when the filter is first enabled, its internal memory contains no real samples, so the first outputs are biased toward zero until enough true samples have entered the recurrence. This warm-up period, however, concerns only the first ~ 100 ms of EMG immediately after filtering is enabled, making it a negligible artefact given that filtering is activated before any acquisition is saved to disk: the streaming data are simply discarded during the warm-up window.

2.2 Spatial Dynamic Time Warping Algorithm

The myoelectric control algorithm is structured as a binary state machine, with the two states being hand OPEN and hand CLOSED. Transitions between the two states are triggered exclusively by the detection, in the EMG signal, of a target gesture, based on template matching through a spatially-corrected adaptation of DTW (Spatial DTW) presented in this work. DTW aligns two signals according to their temporal shape, accounting for small differences in speed and timing that are intrinsically part of human muscle activation in repeated gestures. It has been shown by Eddy et al. to be robust to the limb-position effect (11) and well suited to detecting a discrete, dynamic gesture while rejecting the non-target muscular activity produced during ADLs (20). In (20), DTW was used to detect a finger snap as a wake-up gesture, toggling a binary ON/OFF control of a prosthesis; in this work, this approach is extended to a 1-DoF, posture-robust open/close controller, by maintaining two separate template sets and searching for the target gesture within the set opposite to the current state. Two classes of gestures are therefore defined, corresponding to the two template sets: the act of closing the hand from an open-hand position, which triggers OPEN→CLOSED, and the act of opening the hand from a closed-hand position, which triggers CLOSED→OPEN. These templates are pre-recorded offline before the online operation of the algorithm, and are cut to be one second long, capturing the onset of the movement (further details on the templates are given in Section 2.2.2). During real-time prediction, as in (20), the 16-channel stream is held in a one-second circular buffer (2000 samples per channel), matching the template length. Every 50 ms, the Spatial DTW distance is computed between the live EMG buffer and every template of the class opposite to the current state; if the distance obtained is smaller than a class-specific threshold for five consecutive updates (250 ms), a change of state is triggered.

2.2.1 Feature Extraction

Within each one-second buffer, a sliding window of 96 ms with a 32 ms increment is applied, and the RMS value is computed independently for each of the 16 channels in each window, yielding a feature matrix of 29 frames $\times$ 16 channels per buffer:

$$\text{RMS}_{c,k} = \sqrt{\frac{1}{W} \sum_{n=0}^{W-1} x_c[(k-1) \cdot S + n]^2}$$

where x_c is the signal of channel c , $W = 192$ is the window length in samples, $S = 64$ is the shift, and $k = 1, \dots, 29$ is the window index. The same procedure is applied offline to every template. A shorter window and increment were chosen with respect to (20) to take advantage of the higher sampling frequency of the probe used in this work. As a consequence, the effective buffer and template length

used in computation is 992 ms rather than a full second: with 29 frames of 96 ms spaced 32 ms apart, the last window ends at sample 1984 of the 2000-sample buffer, leaving the final 8 ms of each buffer unused by the feature extraction. The choice of **RMS** as the feature is motivated by its computational simplicity and its monotonic relationship with motor-unit recruitment in the underlying muscles (1). The feature registry also supports other time- and frequency-domain features (waveform length, mean absolute value, zero crossings, slope-sign changes, mean and median frequency, spectral entropy, etc.), selectable through the user interface. **RMS** was selected after a pilot comparison, reported in the Template Study tool, as one of the most stable features.

2.2.2 Templates Extraction

The templates are recorded during a separate calibration session performed before the online prediction phase, and constitute the foundation of the whole system's performance. They represent the target movement: either the act of opening the hand from a closed state, or closing the hand from an open state. During these acquisitions, EMG is recorded continuously throughout the session; the subject is shown a visual cue indicating when to perform each movement, and the timestamp of the cue is saved alongside the recording. The templates are then extracted offline: an algorithm searches, within a fixed neighbourhood around each cue timestamp, for the sample marking the start of the one-second template window, using a max-energy criterion:

$$k^* = \arg \max_k \sum_{n=k}^{k+L-1} \text{RMS}^2[n] \quad (2.1)$$

where $L = 2000$ samples (one second) is the template window length, and the RMS envelope is computed over a 50 ms sliding window across the active region of the recording (the closing region for the CLOSED template, the opening region for the OPEN template). Since the template captures the physiological EMG translation of the hand transitioning from open to closed (or vice versa), the preparation for the movement and its earliest onset are a critical part of the signal. For this reason, the index found by the max-energy criterion is always anticipated by 200 samples (100 ms) to ensure this pre-activation is included in the segment.

Each template is reviewed in the *Template Study* dialog of MyoGestic, which computes silhouette scores and intra- and inter-class cosine similarities between spatial profiles, flagging outliers (silhouette < 0 or a "gap-crosser" criterion) for optional manual discard before model creation.

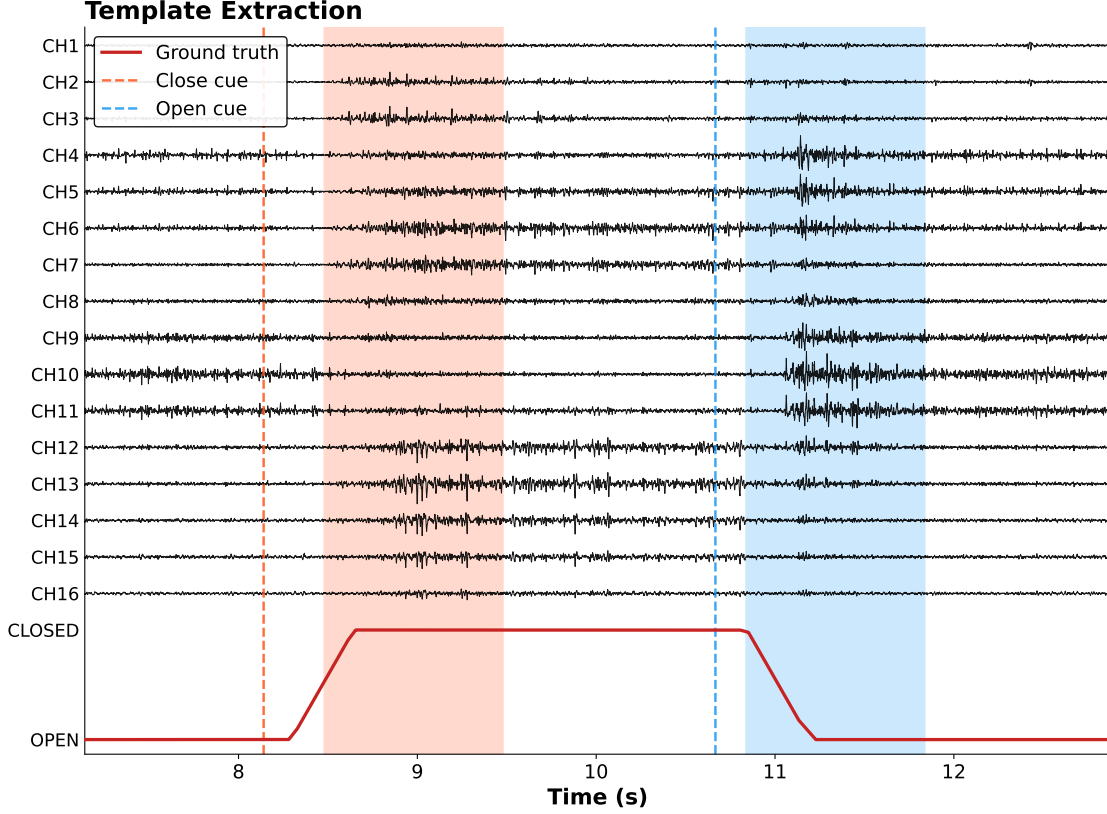


Figure 2.5. Template extraction. The two one-second windows (red: closed template, blue: open template) are positioned automatically by the max-energy algorithm. The -200 ms shift makes sure the windows are placed before the beginning of the activation, capturing also the preparation for the movement. An important layer in the differentiation of the two templates is the intrinsic difference between this two activations, with different channels having higher and lower amplitudes, representing the different muscles of the fingers flexors and extensors.

2.2.3 Cosine DTW

DTW is a distance measure originally developed in the speech-recognition community to compare time series that share the same underlying content but differ in their timing (17). Given two feature matrices, one extracted from the live buffer and one from a template, **DTW** searches for the optimal point-to-point alignment between the two sequences and returns the cumulative distance along this optimal path. The search is performed by building a cost matrix D of size $(N+1) \times (M+1)$, where N is the number of feature windows of the live signal and M of the template (here $N = M = 29$, see Section 2.2.1). Row i indexes window i of the live signal $\mathbf{x}$ and column j indexes window j of the template $\mathbf{y}$. The matrix is initialised with $D[0,0] = 0$ and $D[i,0] = D[0,j] = +\infty$ for $i, j \geq 1$, and is then built iteratively via

the recurrence

$$D[i, j] = d(\mathbf{x}_i, \mathbf{y}_j) + \min \left(D[i-1, j], D[i, j-1], D[i-1, j-1] \right), \quad i, j \geq 1,$$

where $d(\cdot, \cdot)$ is the multivariate cosine distance, discussed below. The recurrence allows window i of the live signal to be aligned with window $j-1$, j , or $j+1$ of the template (and vice versa), so that the optimal alignment can stretch or compress one sequence relative to the other instead of being forced into a strict frame-by-frame correspondence. The final DTW distance between the live signal and the template is then the value of the cumulative alignment cost at the matrix endpoint, $D[N, M]$. This timing flexibility is exactly the property that makes DTW well suited to EMG comparison. The same volitional movement, repeated by the same subject, is never executed with identical timing: the rate at which a contraction ramps up, the duration of the peak, and the relaxation phase all vary slightly from trial to trial, even when the intended movement and the resulting kinematics are the same. A naive whole-sequence Euclidean distance, i.e. the sum of frame-by-frame distances along the matrix diagonal, with no warping permitted, would penalise these benign timing differences as if they were genuine class differences. DTW, by contrast, absorbs the intrinsic trial-to-trial timing variability of human muscle activation. The per-cell base distance $d(\mathbf{x}_i, \mathbf{y}_j)$ that populates each cell of the cost matrix compares a single feature window of the live signal with a single feature window of the template, both represented as 16-dimensional vectors (one value per differential channel). The multivariate cosine distance is defined as

$$d_{\cos}(\mathbf{x}_i, \mathbf{y}_j) = 1 - \frac{\mathbf{x}_i \cdot \mathbf{y}_j}{\|\mathbf{x}_i\| \cdot \|\mathbf{y}_j\|}, \quad \mathbf{x}_i, \mathbf{y}_j \in \mathbb{R}^{16}.$$

This was preferred over a per-cell Euclidean distance because, by measuring the angle between the two channel vectors rather than their magnitude, it does not require any normalisation of the data and makes the distance computation insensitive to the overall contraction strength at any given instant, a property that matters when comparing signals across sessions or subjects whose absolute EMG amplitude may differ for reasons unrelated to the subject's intent, such as skin-electrode impedance, environmental conditions, or fatigue. Because the overall distance is not dependent on EMG amplitude, no normalisation is applied to the signal, in line with the design choice of keeping the data as close to the raw physiological signal as possible. The final alignment cost is a dimensionless value derived from the sum of the cosine distances along the alignment path; it represents the cost of aligning the two time series. Since the cosine distance takes values between 0 and 1, the maximum possible alignment cost is theoretically $N = 29$. One might expect that if the opening and closing movements are performed at a comparable speed, the alignment cost between them should be low; this is not the case, because the

two movements produce different spatial patterns, depending as they do on different groups of muscles that are spatially separated and located close to different electrodes of the bracelet (see Fig. 2.5). Opening the hand primarily involves the extensor muscles of the forearm, while closing involves the flexors. As a result, for each window the cosine distance is higher when the spatial pattern of the two 16-dimensional vectors differs, yielding a higher overall alignment cost than when the same movement is compared to itself. The DTW distance defined above is therefore intrinsically a joint measure of temporal and spatial similarity. The algorithm is implemented in JIT-compiled Numba code, which compiles the relevant functions to machine code at runtime, substantially speeding up the computation.

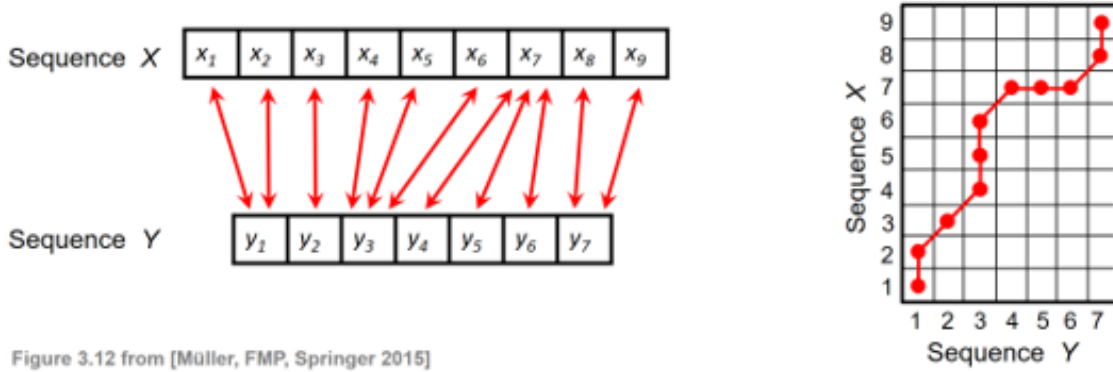


Figure 2.6. Illustration of DTW alignment between two sequences X and Y . *Left*: the two sequences with the optimal point-to-point alignment found by DTW; in the context of this work, each red arrow corresponds to the per-window cosine distance $d_{\cos}(\mathbf{x}_i, \mathbf{y}_j)$ contributing to the cumulative alignment cost. *Right*: the corresponding warping path through the cost matrix, showing how the alignment stretches or compresses one sequence relative to the other. Reproduced from (21).

2.2.4 Spatial Correction

The most common false positives are mainly caused by co-contractions, electrode shift, and non-target movements that recruit overlapping muscle groups (like wrist flexion/extension), as well as arm-movement artefacts. With the goal of increasing the robustness of the decision-making, another filter stage was introduced, the Spatial Correction, where the algorithm performs a direct comparison between the spatial activation pattern of the live signal and the spatial pattern of each template, penalising the DTW distance whenever the two do not agree. For each template, a 16-channel spatial profile $\mathbf{p}_i \in \mathbb{R}^{16}$ is computed once at template-creation time as the L2-normalised per-channel RMS over the second half of the template’s EMG

segment:

$$\mathbf{p}_i = \frac{\mathbf{r}_i}{\|\mathbf{r}_i\|_2}, \quad \mathbf{r}_i[c] = \sqrt{\frac{1}{T} \sum_{n=1}^T x_{c,i}^2[n]},$$

capturing the template's "spatial signature", which channels are active and in what relative proportion, independent of the absolute amplitude. The RMS is computed only in the latter half of the second because it is the part where the hold phase is expected, and the muscles relative to the movement are more active, while the first half was designed to capture the onset and preparation for the movement. At inference time, the same profile $\mathbf{p}_{\text{live}}$ is computed over the second half of the live 1-second buffer, and the spatial similarity between the live signal and template i is the cosine similarity between the two profiles:

$$\sigma_i = \max(0, \min(1, \mathbf{p}_{\text{live}} \cdot \mathbf{p}_i)) \in [0, 1].$$

The raw DTW distance D_i of that template is then penalised based on the value of σ_i , via a saturating scaling function:

$$w(\sigma) = \begin{cases} 1 & \sigma \geq \theta \\ b((1-\sigma/\theta)^{1/k}) & \sigma < \theta \end{cases} \quad D_i^* = \frac{D_i}{w(\sigma_i)}.$$

The function, shown in Fig. 2.7, is parameterised by a saturation threshold θ , a

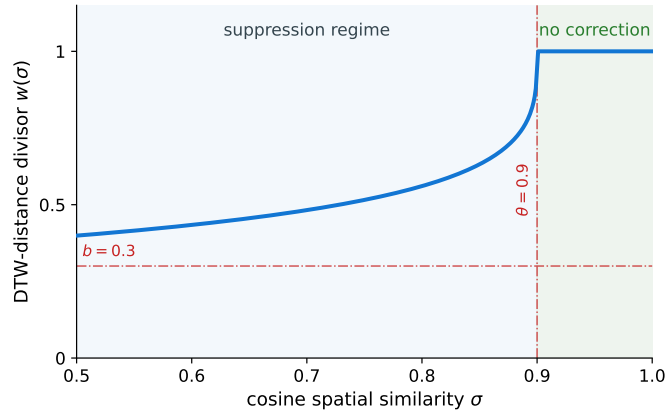


Figure 2.7. The saturating scaling function $w(\sigma)$ used to penalise the DTW distance based on the cosine spatial similarity σ . Above the saturation threshold $\theta = 0.9$ (no correction regime), no penalty is applied. Below it (suppression regime), the divisor decreases smoothly toward the baseline $b = 0.3$ as $\sigma \rightarrow 0$, inflating the DTW distance by up to a factor of $1/b \approx 3.3$ for completely orthogonal spatial profiles.

baseline b , and a sharpness exponent k . Above the saturation threshold ($\sigma_i \geq \theta$)

the live spatial pattern is judged to match the template’s signature closely enough, and no penalty is applied ($w = 1$, $D_i^* = D_i$), so it depends only on the DTW distance. Below the threshold the penalty grows as the spatial mismatch increases: as $\sigma_i \rightarrow 0$ the scaling factor approaches the baseline b , inflating the DTW distance by up to a factor of $1/b$. The exponent k shapes how steeply the penalty ramps between saturation and the floor, and its specific value used in this thesis is given in the caption of Fig. 2.7.

2.2.5 The state-machine framework

The output of the algorithm is governed by a finite state machine, with the two states being OPEN and CLOSED: at any moment the algorithm is in one of the two, corresponding to the most recently committed hand configuration. The detection of a transition between the two states is the only event the pipeline is allowed to declare; everything described in Sections 2.2.3–2.2.4 ultimately serves this binary decision. For example, if the current state is OPEN, the algorithm searches for the transition in the real-time EMG data: it computes the Spatial DTW distance between the last second of signal and every template of the opposite class (in this case, the CLOSED templates), every 50 ms. The predicted state does not change until the distance between the real-time signal and the templates becomes small enough for the algorithm to consider it a real transition, at which point the prediction changes from OPEN to CLOSED. How small the distance should be is governed by two factors: the distance aggregation method and a threshold. From the current state, a transition to the opposite state is triggered whenever the class scalar associated with the opposite state falls below the corresponding per-class threshold (the procedure for setting these thresholds is described in Section 2.2.5). Once a transition is committed, a refractory period prevents an immediate back-transition into the previous state. In the experiments reported in this thesis the refractory period was set to 1 s, a value chosen to match the natural minimum interval between two voluntary open–close events while still suppressing accidental back-and-forth oscillations on edge cases. The duration is configurable for clinical use, where a longer dwell time between transitions may be desirable.

Distance Aggregation

At every inference cycle, the algorithm produces one spatially-corrected DTW distance $D_{\text{class},i}^*$ for each template i of the opposite class. These are then reduced to a single scalar by averaging the three smallest:

$$D_{\text{class}} = \frac{1}{3} \sum_{i \in S_3(\text{class})} D_{\text{class},i}^*,$$

where $\mathcal{S}_3(\text{class})$ is the set of three indices whose $D_{\text{class},i}^*$ values are the lowest within the class. Averaging only over the three most similar templates enables the use of a wider and more diverse repertoire of templates for the same gesture, for example spanning many arm postures: since only the best-matching templates contribute to the score, adding templates of the same movement from further postures enriches the posture coverage without penalising the decision. While DTW alone has already been shown to be robust to the limb-position effect (11), the spatial correction layer introduces a filter that could in principle penalise templates recorded in other postures, due to their different muscle activation patterns. Provided at least three templates are recorded per position, however, the aggregate can still be satisfied entirely by the single position that best matches the live signal, while the remaining templates, being spatially inconsistent with the current posture, are inflated by the spatial correction and left out of the average. This decoupling of the score from the overall composition of the set is what allows the template library to grow across postures while a single per-class threshold continues to separate target from non-target activity. Averaging across all templates would instead dilute the score whenever only a few of them spatially match the current posture. Averaging the three smallest values avoids this dilution and, in doing so, also removes any strict numerical limit on how many templates can be included in the bank: the only practical constraint is the computational time required to evaluate a single DTW distance, since every additional template adds one more distance computation per inference cycle.

Threshold Calibration

The threshold used to trigger a change of state is the last layer influencing the decision-making, and its value can dramatically affect the performance of the algorithm. Since the two classes have two different sets of templates, the two thresholds assume different values. As in (20), the algorithm computes the Spatial DTW distance between every unique combination of templates within the same class, and each threshold is initialised as

$$T_{\text{class}} = \mu_{\text{class}} + s_{\text{class}} \cdot \sigma_{\text{class}},$$

with μ_{class} the mean, σ_{class} the standard deviation, and s_{class} the tunable number of standard deviations away from the mean. This value allows fine-tuning of the threshold to reduce false positives. Having two separate thresholds also gives the possibility of adding hysteresis to the state transitions: for example, a smaller T_{open} adds robustness when the hand is closed, increasing the selectivity for the gesture that opens the hand. Reducing a threshold is, however, a trade-off, as it also increases the number of missed events and the response latency. In the experimental phase, the thresholds were both calibrated offline against recordings of non-target activity, tuned to suppress false activations while preserving true detections, a

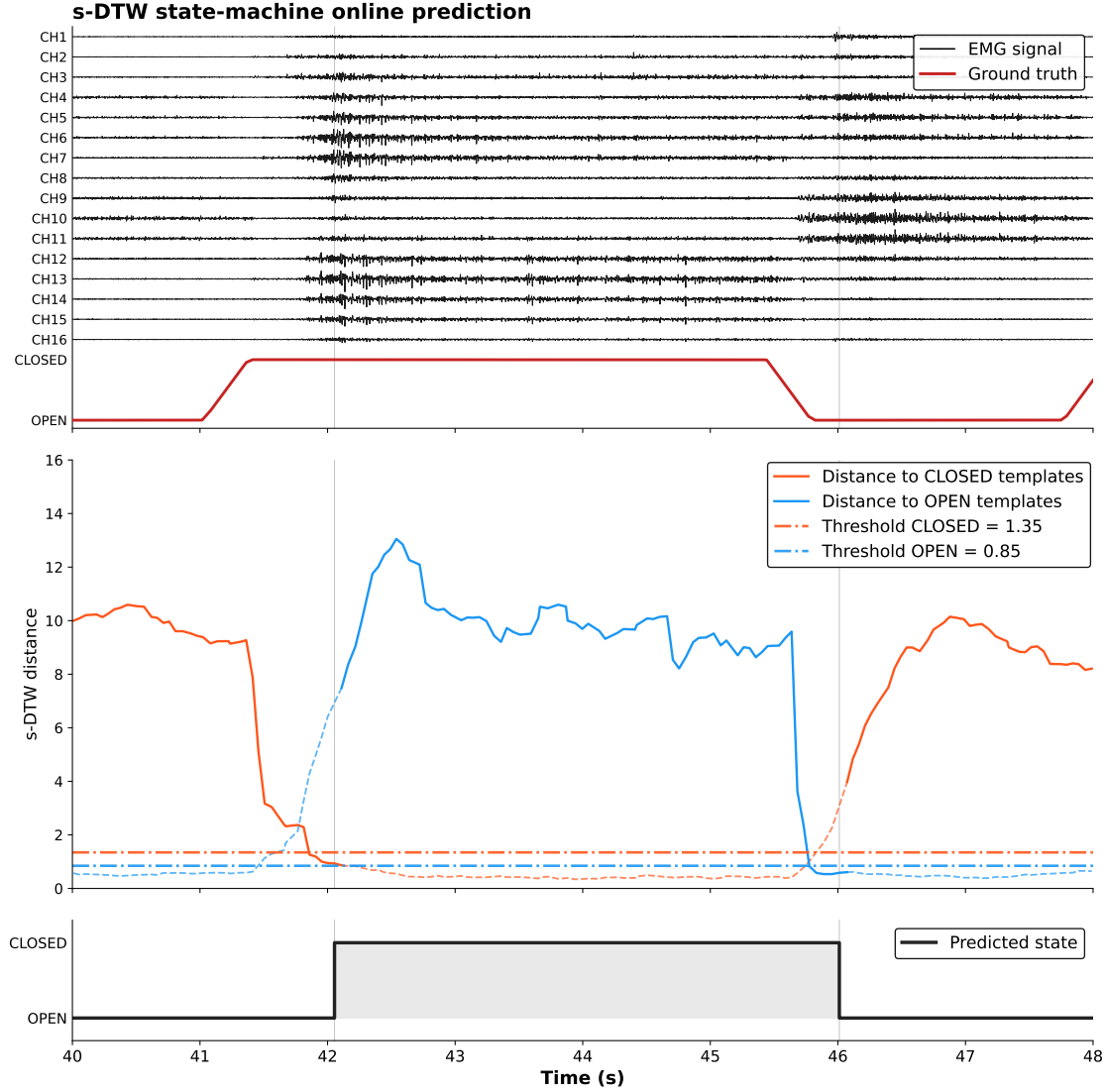


Figure 2.8. The state-machine logic driving the final prediction. Whenever the distance between EMG and templates of the opposite class goes below the per-subject threshold, for five consecutive times, the change of state is triggered. In the real-time trials the algorithm was computing only the Spatial DTW distances to the opposite state; this figure also shows the distances with the same-state templates, drawn with dashed lines.

procedure whose data collection is described in 2.3.5 and whose calibration formula is given in 2.3.7. Across the cohort, the default value $s_{\text{class}} = 3$ was retained for the majority of subjects.

2.3 Experimental Phase

In order to test the real-time performance of the prediction of the algorithm, a single session consisting of three calibration steps (approx. 20 minutes) and four testing phases (approx. 45 minutes) was designed. The study involved a total of 16 healthy subjects and 4 patients. Tests on 2 of the 4 patients were conducted with an earlier stage of the code and are therefore not reported in the official experimental phase, since their performance could not be quantified. As the experiment involved moving the arm across different postures in space, the patients followed a simplified version of the protocol, involving only movements of their paralysed hand.

2.3.1 Population

The algorithm testing involved sixteen healthy subjects (S1–S16; age between 33 and 24 years old, 9 women, 7 men) recruited on a voluntary basis from the university population, and two spinal cord injury patients (P1, P2), recruited through Dr. med. Matthias Ponfick at the Kliniken Rummelsberg, Germany. Both patients had sustained a spinal cord lesion resulting in impaired hand movement. P1 was able to close the hand but could not fully open it, extending all the fingers, while P2 was able to perform both movements, though with visible effort and a reduced range of motion, especially in the interphalangeal joints. All participants provided written informed consent, and the study was conducted in accordance with the institutional ethics committee.

2.3.2 Laboratory Setup

The subject was seated on a rigid cube stool, without backrest or armrests, in front of a desk with the laptop and a supplementary monitor connected via HDMI. The arm carrying the bracelet was kept unsupported throughout the session, with the elbow free, so that the subject could move it through ten experimental positions (described in Section 2.3.3) without any physical constraint imposed by the laboratory environment.

The supplementary screen displayed two interfaces in parallel:

- the [Virtual Hand Interface \(VHI\)](#), which animates two three-dimensional hands: the control hand and the predicted hand. The first, transparent, mirrors the gesture the subject is asked to perform; under the hypothesis that the subject is always following it, this hand represents the [Ground Truth \(GT\)](#). The latter, pink-coloured, shows the predicted state when the algorithm is running, giving real-time feedback to the subject. The [VHI](#) code lives in Unity and was inherited from the MyoGestic framework.



Figure 2.9. Laboratory setup: the subject was seated on the cube stool, with the laptop (on the right) showing the UI interface to control the experiment and plot the real-time EMG signal, and the supplementary monitor (left) which showed the Position Interface on the right side and the Virtual Hand Interface on the left side. On the desk, different sizes of the EMG-bracelets, and the reference electrode.

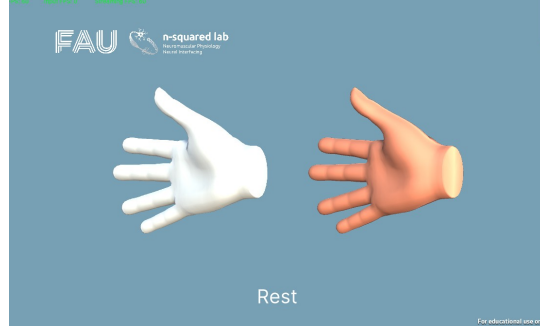


Figure 2.10. Virtual Hand Interface (VHI): the transparent control hand and the pink predicted hand, shown on the supplementary monitor.

- the [Position Interface \(PI\)](#), which displays an image representing the arm posture the subject was expected to assume before performing the gesture of opening or closing the hand. During some tasks involving wrist movements, this interface was also used to show a zoom of the forearm, to display the wrist movement to perform. More details regarding the different postures are given in the next sections. This interface was built from scratch, after designing the images of the positions.

The bracelet was placed on the dominant forearm, approximately 3–5 cm distal to the elbow crease, with a reference electrode on the back of the hand. The electrode array spanned the full circumference of the forearm with uniform spacing. No prior preparation was applied to the skin before placing the bracelet; this choice prioritises clinical and daily-life practicality. The laptop screen displayed the framework’s [User Interface \(UI\)](#), likewise inherited and modified from MyoGestic. Once the probe was connected to the laptop via Wi-Fi and data streaming was started, the [UI](#) also plotted the live EMG signal, so that it was always possible to inspect it in order to verify correct functioning. Through the [UI](#), all phases and settings of the experiment were controlled. Two cameras were used to record the experiment: a personal Fujifilm X-S10 (Fujifilm Corporation, Tokyo, Japan) and the laboratory’s Sony α 7 (Sony Corporation, Tokyo, Japan). The latter, used to record both the calibration and testing phases, was placed vertically behind the subject, capturing

the upper body, the arm carrying the bracelet, the hand, and the two screens all in one frame. The former captured the complementary lateral angle and was turned on only during the testing phases. The videos were synchronised offline by aligning the common audio spikes. The full video of the experiment serves as proof that the subjects were following the control hand, making the [GT](#) reliable. The videos were also inspected to extract the final results of the experiment.

2.3.3 Positions Overview

As one of the aims of the project is a robust algorithm across different arm postures, a set of 10 positions was used throughout the experimental phase. The positions were selected to span a representative sample of the upper-limb workspace during [ADL](#), so that the calibration and testing protocols would expose the algorithm to the kind of postural variability a user would encounter in clinical and home use. Each position was assigned a letter as reference (A–J). The full list, the letter, the physiological description, and the associated [ADL](#) task are reported in Table 2.1. The images of the positions displayed to the subjects through the [PI](#) are shown in Fig. 2.11. While the wrist orientation was kept neutral by default for most of the positions, three intentionally introduce a non-neutral wrist orientation: position B uses a slight pronation typical of grasping a vertical handle, position C uses full supination characteristic of receiving an object or holding a plate, and position E uses pronation matching a keyboard or mouse posture. Despite the possibility of being redundant, more templates were given to the frontal portion of space, as these would be the most frequent postures for a patient (e.g. drinking, eating). Moreover, this deliberate variation in the wrist axis introduces an additional source of postural variability beyond shoulder and elbow configuration, and stresses the algorithm on the forearm-rotation dimension that pure shoulder/elbow positions would not exercise.

2.3.4 Templates Acquisition

The first step of the calibration phase was the acquisition of the templates on which the whole algorithm bases its prediction. The template acquisition protocol asks the subject to perform the target movement, opening and closing of the hand, across the ten arm positions described in Table 2.1. A total of 80 templates was recorded per subject: 40 open and 40 closed, with 4 templates per class for each position (8 templates per position in total, 4 open and 4 closed). For each of the ten positions, the subject was guided through two visits, each of which contained two open-close cycles of the hand, yielding the 4 templates per class per position. The order in which the positions were visited was randomised within each visit round, so that subjects did not all experience the same position sequence and any order effect averaged out across the cohort, and to average out the eventual

Label	Short name	ADL task	Wrist
A	Side Rest	Arm relaxed at side (carrying a bag)	Neutral
B	Table Reach	Reaching a bottle on a table, elbow 90°	Slight pron.
C	Palm-Up Receive	Receiving an object, holding a plate	Supine
D	Forward Reach	Opening a door handle, elbow extended	Neutral
E	Desk Side	Using a mouse or keyboard, elbow 90°	Prone
F	Lateral Pass	Passing an object sideways, elbow extended	Neutral
G	Hand to Mouth	Drinking from a glass	Neutral
H	Hand to Head	Combing hair, brushing teeth	Neutral
I	Overhead Reach	Reaching a high shelf, arm raised	Neutral
J	Cross-Body	Fastening a seatbelt, reaching opposite side	Neutral

Table 2.1. The ten arm positions used in the experimental phase, with the corresponding short name, ADL task each posture represents, and wrist orientation. Positions B, C, and E deliberately introduce a non-neutral wrist orientation to vary the forearm-rotation axis beyond what shoulder and elbow configuration alone provide.

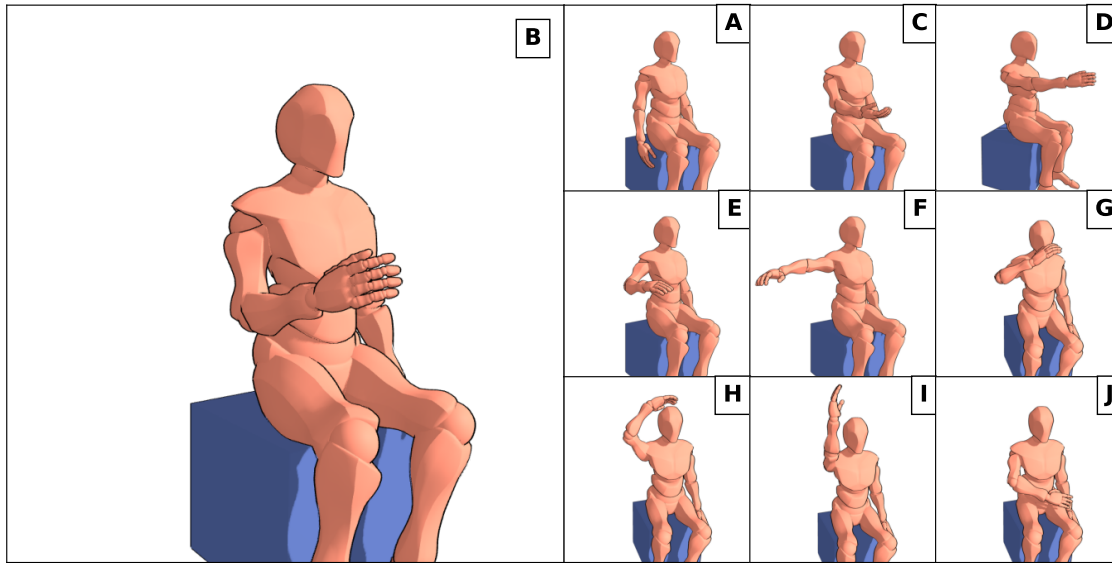


Figure 2.11. The 10 ADL positions chosen for the experiment, which were shown to the subjects through the Position Interface.

effects of fatigue on the last positions visited. The two interfaces (Virtual Hand and Position) never changed simultaneously. First, the [PI](#) showed the target arm posture the subject had to achieve; once in position, the operator started the two cycles, and the control hand started moving. When the two cycles ended, the position changed automatically in the interface, and the subject was required to

move to that position to record two more cycles. Each open-close cycle animated by the control hand consisted of the same two movements that are the object of this framework: the transition from open to closed and the transition from closed to open. The animation followed a four-phase sequence, hold open, closing, hold closed, opening, with configurable durations stored in the recording metadata; for the present study the four durations were set to 2.00 s, 0.55 s, 2.00 s, and 0.55 s, respectively. A visit to a position therefore lasted approximately 12 s. The corresponding ground-truth ramp signal (a continuous value between 0 and 1 indicating the expected hand state at every instant) was streamed in parallel and saved alongside the EMG, so that the offline template-extraction algorithm could later locate the onset of each open and close transition, using the procedure described in Section 2.2.2. The templates were stored in the calibration pickle with a label encoding both the position letter and the visit suffix (`_12` for the first visit, `_34` for the second), so that the position of origin of each template could later be recovered during online and offline analysis.

The choice of four templates per class per position follows directly from the distance aggregation method introduced in Section 2.2.5. The average of the three closest templates works under the implicit assumption that, when the subject performs a movement in position P , the three lowest DTW distances will likely be those of the templates acquired in the same position P . For this reason, at least three templates per position are needed, and four per position were finally chosen as a safety margin, in case any one of them is later judged to be an outlier and discarded, as described in Section 2.2.2.

2.3.5 Wrist Distractor Recording

The second component of the calibration phase was a single short (less than one minute) recording in which the subject was asked to perform a sequence of non-target movements, while keeping the overall state of the hand constant. The recording was performed at position B (table reach) and consisted of two blocks executed in sequence: in the first block the subject held the hand open, in the second block the subject held the hand closed, the transition between the two blocks being an in-place open→close animated by the Control Hand. Within each block the subject performed the six wrist distractors, pronation, supination, wrist flexion, wrist extension, radial deviation and ulnar deviation, once each, separated by a neutral rest. The full event sequence and the per-event durations are reported in Table 2.2. During the recording, the **VHI** and the **PI** played distinct roles: the Control Hand always showed the required hand state, open during the first block, closed during the second block, with the in-place open→close transition animated at the block boundary; the **PI** displayed the photograph of the wrist gesture the subject was expected to perform at each instant, switching in sync with each audio cue.

The purpose of this auxiliary recording was not to provide further templates to

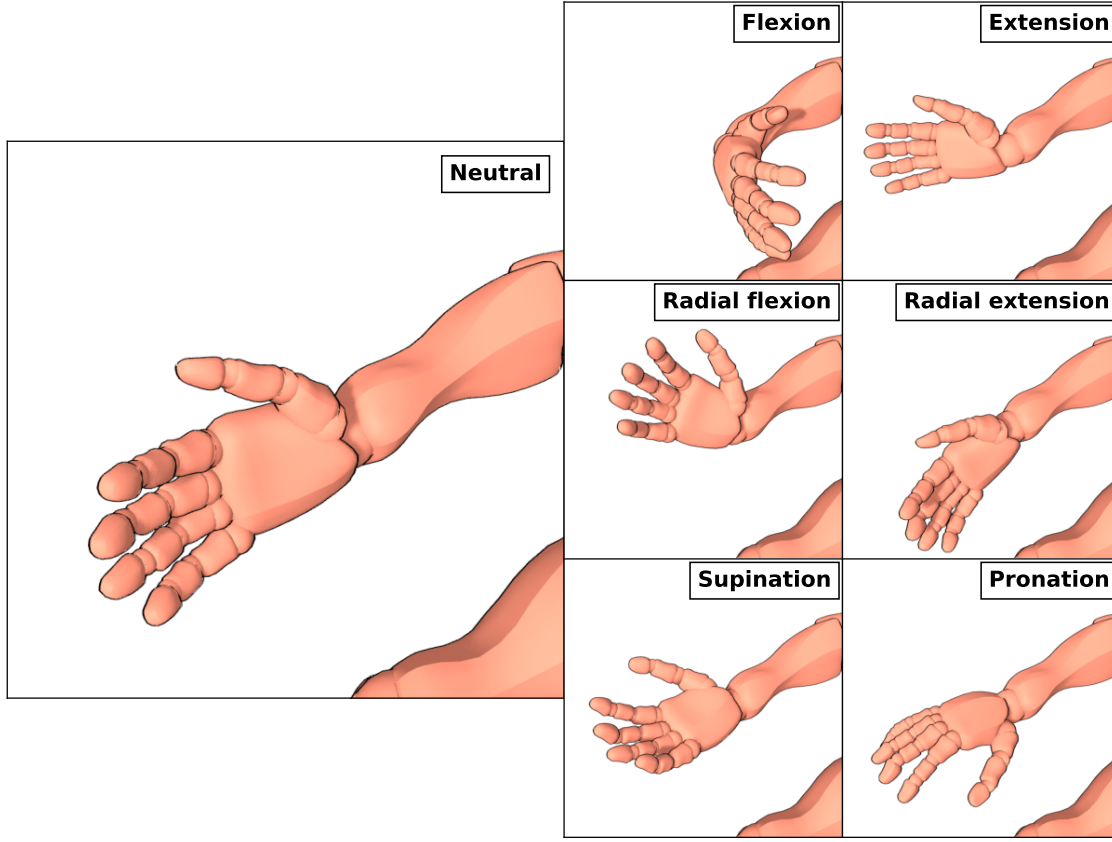


Figure 2.12. The wrist position displayed in the Position Interface for wrist false positive calibration and the other wrist distractor experiment.

the algorithm, but to give a per-subject empirical reference of how close the DTW distance came to crossing the OPEN or CLOSED threshold, these movements being a hypothetical source of false positives. Wrist movements are the highest false-positive risk for a 1-DoF hand controller because the wrist flexor compartment of the forearm is anatomically adjacent to the finger flexor compartment, and the bracelet does not separate them: a voluntary wrist flexion produces a channel-activation pattern that partially overlaps with the templates of a hand-closing transition, and could, in principle, drive the DTW distance below the OPEN-state threshold even when no hand closure was intended. The same concern applies, with smaller amplitude, to wrist extension against the OPEN templates and to the rotational and deviation gestures against both classes. The wrist-distractor recording is therefore an additional input to the offline threshold-calibration step during the final model creation, described in Section 2.2.5; no template is extracted from it, and the recording itself is not used during the online testing phase.

Table 2.2. Wrist calibration event sequence at position B (table reach). Each block consists of the six wrist distractor gestures interleaved with neutral rest periods. Between the two blocks an in-place pen→close transition is animated by the Control Hand. The order of the gestures is the same across subjects.

Block 1 — hand kept OPEN (starting state: OPEN)		Block 2 — hand kept CLOSED (after in-place open→close transition)	
#	Event (duration)	#	Event (duration)
1	wrist neutral (2.0 s)	1	wrist neutral (2.0 s)
2	pronation (1.5 s)	2	pronation (1.5 s)
3	wrist neutral (2.0 s)	3	wrist neutral (2.0 s)
4	supination (1.5 s)	4	supination (1.5 s)
5	wrist neutral (2.0 s)	5	wrist neutral (2.0 s)
6	wrist flexion (1.5 s)	6	wrist flexion (1.5 s)
7	wrist neutral (2.0 s)	7	wrist neutral (2.0 s)
8	wrist extension (1.5 s)	8	wrist extension (1.5 s)
9	wrist neutral (2.0 s)	9	wrist neutral (2.0 s)
10	radial deviation (1.5 s)	10	radial deviation (1.5 s)
11	wrist neutral (2.0 s)	11	wrist neutral (2.0 s)
12	ulnar deviation (1.5 s)	12	ulnar deviation (1.5 s)
13	wrist neutral (2.0 s)	13	wrist neutral (2.0 s)

2.3.6 Data for CatBoost

The third component of the calibration phase was a pair of 40-second recordings during which the subject held a constant hand state, open in the first recording, closed in the second, and sequentially traversed all ten arm positions, holding each one for four seconds. For these recordings the subject followed only the [PI](#), which displayed each position for four seconds before changing to the next one. The order was from A to J, for every subject. Each recording is therefore a single continuous segment of EMG with a stable hand state, distributed evenly across the ten positions. The two recordings together provided the labelled dataset used to train a CatBoost classifier, the state of the art already implemented inside the MyoGestic framework and explained in Section [1.6.2](#). Although this data was not used during any online prediction, it will be used for offline studies, as a 1-DoF classifier applied offline to the same data as the online tests performed, in order to have a direct comparison between this method and the Spatial DTW algorithm. The two 40-second recordings provide the same total duration of raw EMG ($40 \text{ s} \times 2 = 80 \text{ s}$) that the Spatial DTW templates collectively represent ($40 \text{ templates per class} \times 1 \text{ s per template} \times 2 \text{ classes} = 80 \text{ s}$), so the two algorithms are placed on a comparable footing in terms of the raw amount of EMG data used for training. This does not, however, imply a comparable calibration burden for the subject:

in a clinical-deployment scenario, the CatBoost calibration is considerably faster and less demanding, approximately 40 seconds per class, against the roughly fifteen minutes required to complete the full Spatial DTW template-acquisition protocol across all ten positions. The working hypothesis of this comparison, evaluated in 3.7.1, is nevertheless that the Spatial DTW algorithm will achieve better robustness against false positives and a higher true-positive rate than the CatBoost baseline, despite this heavier calibration burden, which is precisely the trade-off the position-aware design is meant to win.

2.3.7 Model Creation

At the end of the three calibration recordings (template acquisition, wrist-distractor recording, CatBoost continuous-state recording) the bracelet was left in place and the raw recordings were stepped through to create the per-subject Spatial DTW model used during the testing phase. The pipeline consists of four sequential steps that progressively assemble this model; none of them requires additional EMG acquisition.

Template extraction. The ten template-acquisition recordings (Section 2.3.4) were processed as described in Section 2.2.2, yielding the 80 retained templates per subject after outlier review in the Template Study dialog.

Spatial profiles and intra-class statistics. For each retained template, the spatial profile of Section 2.2.4 was computed and stored alongside the template, together with the intra-class mean μ_C and standard deviation σ_C of pairwise DTW distances for each class, providing the baseline statistics for the next step.

Threshold calibration. The per-class thresholds were initialised at the default preset of Section 2.2.5, then refined against the wrist-distractor recording (Section 2.3.5): the model, in this provisional state, was used to replay the recording offline, producing the per-class DTW distance traces $D_{\text{open}}(t)$ and $D_{\text{closed}}(t)$. The recording was segmented into positive events — the in-place OPEN→CLOSED transition performed between the two blocks, during which the algorithm should fire, and negative events (the six wrist-distractor gestures), during which the algorithm should not fire. For each class $C \in \text{OPEN, CLOSED}$ the minimum DTW distance reached during any negative event was identified,

$$D_{\min}^-(C) = \min_{t \in \text{neg}} D_C(t),$$

and the per-class threshold was then set to the midpoint between the intra-class preset $\mu_C + 3\sigma_C$ and $D_{\min}^-(C)$, subject to a hard floor at the intra-class preset itself:

$$T_C = \max \left(\mu_C + 3\sigma_C, \frac{(\mu_C + 3\sigma_C) + D_{\min}^-(C)}{2} \right).$$

This rule places the threshold equidistantly between the template-based notion of "normal" intra-class variability and the empirically observed worst-case distractor,

while preventing it from drifting below the intra-class preset even when the wrist recording happens to be benign.

Saving and locking. The calibrated thresholds were written back to the model file, and the resulting model, templates, their spatial profiles, the intra-class statistics (μ_C , σ_C), the spatial correction parameters, and the calibrated thresholds (T_{OPEN} , T_{CLOSED}), was serialised to a single `.pkl` model file and loaded into the online prediction pipeline. From this point until the end of the session no template re-extraction, no spatial-profile re-fit, no threshold re-tuning, and no per-subject parameter change was performed. The testing-phase results therefore reflect the generalisation of a single, fixed per-subject calibration to all subsequent test conditions, not a continuously adapted model.

2.3.8 Testing Phase

The testing phase began immediately after the model-creation step described in Section 2.3.7, with the bracelet still in place and the Spatial DTW model loaded into the online prediction pipeline. Once the model was loaded, the predicted hand of the **VHI** was made visible to the subject for the first time, and a brief familiarisation period followed: the prediction pipeline ran live but no data were saved, so that the subject could observe the predicted hand mirror their voluntary open–close contractions and become comfortable with the visual feedback loop before the start of the actual experiment. This familiarisation period was also used to set up the second lateral camera, operated only during the testing phase. The testing phase consisted of four blocks performed in immediate succession, separated only by short rests for fatigue management: the main experiment (Section 2.3.9), the wrist-distractor test (Section 2.3.10), the finger-distractor test (Section 2.3.11), and the mug-holding test. The total duration of the testing phase varied between 30–45 minutes per subject.

2.3.9 Main Experiment

The main experiment is the core test of the position-robustness claim: it evaluates whether the algorithm correctly detects open-close transitions across the ten ADL positions, while being robust during movements of the arm. The structure of the recording recalls the template-acquisition phase described in Section 2.3.4: each position was visited two times, in a randomised order, and in each visit the subject was required to perform transitions from open to closed or vice versa. Again, the **PI** was showing the position, and only once the subject was on site were the transitions started manually, displayed with the Control Hand of the **VHI**; once finished, the position was automatically updated. Differently, this testing phase was conceived as a single, continuous recording, logging the full prediction also during the repositioning of the arm from one position to the next. At each visit to

a position, the subject was asked to perform three transitions; after the third, the subject was required to move the arm to the next ADL posture while maintaining the hand in the last state (e.g. if the last transition of the visit was open→closed, they had to keep the hand closed). The odd number of transitions ensured that the repositionings were alternated once with the hand open and once with the hand closed. The randomisation order of the positions was a permutation that ensured the subject arrived once in the open state, to leave with the hand closed, and the other time in the closed state, to leave with the hand open. For each single transition, the opening/closing time of the Control Hand remained 0.55 s, while the hold-open and hold-closed phases were randomised each time between 1.75 and 4.00 s, so as not to let the subject adapt to a fixed rhythm of opening and closing and anticipate the movement. The main experiment therefore has a double aim: first, to test whether all the cued transitions are detected correctly by the real-time prediction of the model; second, to check whether, during changes of position, the predicted state remained unchanged, meaning no false positives during movements of the arm. Each recording contained a total of 60 transition cues (3 transitions per visit $\times$ 2 visits per position $\times$ 10 positions) and 19 repositioning windows (the beginning of the recording was not counted, and the recording always started with the Control Hand open, so a total of 10 repositionings with the hand closed and 9 with the hand open). Each subject performed this test twice, with a short rest period to reduce eventual fatigue effects, for a total of 120 transitions to be detected, and 38 repositionings (20 closed, 18 open) per subject. The subject could visualise the real-time prediction from the predicted hand of the [VHI](#), right next to the Control Hand, for the whole experimental phase. To quantify the performance of the prediction, each transition was considered correctly detected only if the prediction tracked the ground truth (the Control Hand) at the cue, and if, in the time interval between that cue and the next one, no other transitions were detected. This means that, for example, if the Control Hand closed and the predicted hand also closed (due to the correct movement of the subject) but then re-opened right afterwards, before the reopening of the Control Hand, the detection was considered not correct, but an *instability* (counted as a false positive, FP). The performance metric is therefore per-event based. Besides the instability, the other form of error is the *missed event*, corresponding to a missed prediction of transition within a two-cue interval; this error is counted as a false negative (FN).

The main metric extracted will be referred to as the [Correct Detection Rate \(CDR\)](#), the ratio of all transitions correctly detected (true positives, TP) to the total number of cued transitions retained for analysis; formally, the Jaccard index (or intersection-over-union):

$$\text{Correct Detection Rate} = J = \frac{|TP|}{|TP \cup FN \cup FP_{\text{cue}}|} = \frac{TP}{TP + FN + FP_{\text{cue}}}, \quad (2.2)$$

The Correct Detection Rate is reported alongside the standard *F1-score*, which

combines the per-class *recall* (the fraction of cued transitions that produced at least one correct firing) and *precision* (the fraction of fired transitions that were correct):

$$\text{Recall} = \frac{TP}{TP + FN}, \quad \text{Precision} = \frac{TP}{TP + FP}, \quad (2.3)$$

$$F_1 = 2 \cdot \frac{\text{Recall} \cdot \text{Precision}}{\text{Recall} + \text{Precision}}. \quad (2.4)$$

The behaviour of the algorithm during the inter-position repositioning windows is reported through a separate metric, the *Arm Movements False Positives Rejection Rate* (AMFPRR), which quantifies the fraction of repositioning windows during which the algorithm correctly maintained the current hand state without firing any spurious transition:

$$\text{AMFPRR} = \frac{TN_{\text{repos}}}{TN_{\text{repos}} + FP_{\text{repos}}} = 1 - \frac{FP_{\text{repos}}}{N_{\text{repos}}}, \quad (2.5)$$

where N_{repos} is the total number of repositioning windows in the recording, FP_{repos} is the number of those windows that contained at least one spurious state transition, and TN_{repos} is the number of clean repositioning windows.

2.3.10 Wrist Movements False Positives Tests

The second part of the testing phase aimed to evaluate the algorithm’s ability to ignore voluntary wrist movements while the subject maintains a constant hand state. The wrist flexors of the forearm are anatomically adjacent to the finger flexors, and the bracelet does not separate them spatially: a voluntary wrist flexion produces a channel-activation pattern that partially overlaps with the templates of the hand-closing transition and can eventually cause false positives. This test therefore probes one of the most clinically relevant false-positive risks for the algorithm: spurious state transitions caused by upper-limb movements that the user might perform without intending to actuate the orthosis. The structure of this experiment partially recalls the wrist-distractor calibration recording described in Section 2.3.5, with some key differences. The test was performed at two ADL positions, position B (table reach) and position G (hand to mouth), chosen as the two most representative postures for daily activities involving the hand. At each position, the subject performed two wrist-distractor blocks of opposite hand states, separated by an in-place open↔close transition animated by the Control Hand. The dual-feedback mechanic, Control Hand cueing the hand state, PI displaying the wrist-gesture photograph in sync with each audio cue, was the same as in the calibration recording (Section 2.3.5). The block order was inverted between the two positions (OPEN→CLOSED at B, CLOSED→OPEN at G), so that the inter-position repositioning from B to G was performed with the hand held constantly

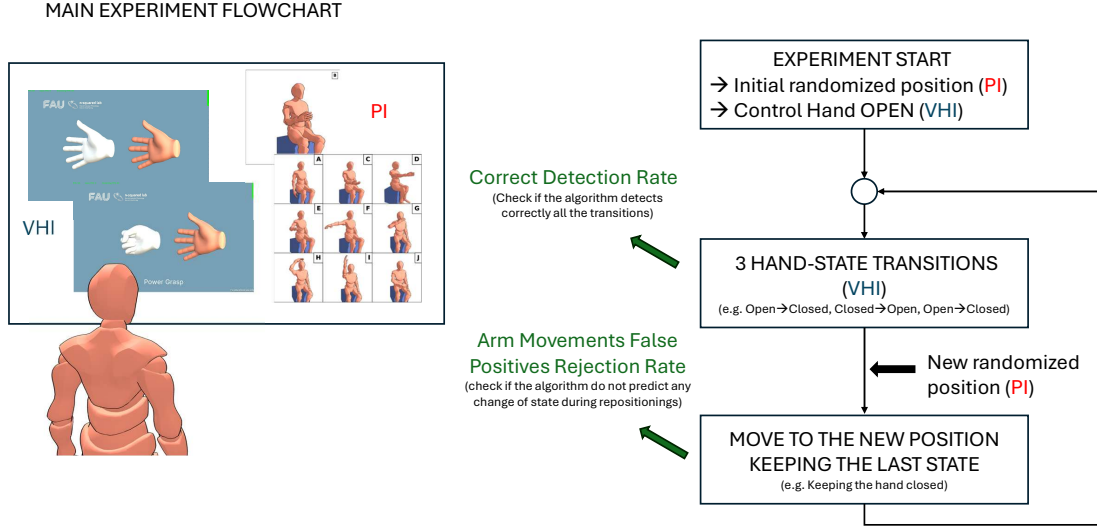


Figure 2.13. Flowchart and setup of the main experiment. Left: subject setup, seated in front of the monitor displaying the Position Interface and the Virtual Hand Interface. Right: flowchart of the experiment, with the two rates being measured to extract the results.

closed, requiring no additional voluntary state change. The full per-round structure of the test is summarised in Table 2.3.

Each block consisted of a sequence of the same six wrist gesture types as the wrist-distractor calibration recording (Section 2.3.5, Table 2.2): pronation, supination, wrist flexion, wrist extension, radial deviation, and ulnar deviation. In this testing phase, however, wrist flexion and wrist extension were each performed twice, since, based on the initial assumption, they carry the highest false-positive risk against the closing and opening templates respectively; the other four gestures appeared once each. The resulting per-block sequence contained 8 gestures, interspersed with short rests as the subject returned to the neutral wrist position. The order of the gestures was randomised at every block, under the constraint that the two flexion instances and the two extension instances did not appear adjacently. Each gesture was held for 2.0 s and each neutral rest lasted 2.5 s. The full test

Table 2.3. Pipeline of one round of the wrist-distractor test. Each round consists of four wrist-distractor blocks performed at two ADL positions, separated by in-place open $\leftrightarrow$ close transitions (animated by the Control Hand) and one inter-position repositioning (performed freely by the subject). The block order is inverted between the two positions so that the subject’s hand state is constant during the B $\rightarrow$ G repositioning.

Step	Position	Hand state	Description
1	B	OPEN	Block B _{OPEN} — 8 wrist gestures
↓	(in place)	OPEN $\rightarrow$ CLOSED	Change of state
2	B	CLOSED	Block B _{CLOSED} — 8 wrist gestures
↓	B $\rightarrow$ G	CLOSED	Change of position (hand held closed)
3	G	CLOSED	Block G _{CLOSED} — 8 wrist gestures
↓	(in place)	CLOSED $\rightarrow$ OPEN	Change of state
4	G	OPEN	Block G _{OPEN} — 8 wrist gestures

yielded a total of 32 gestures evenly distributed across the open and closed hand states (8 gestures $\times$ 2 blocks $\times$ 2 positions). The full test was repeated for two rounds for subjects S5–S16, yielding a total of 64 wrist-gesture events, which should not have produced any change of state in the real-time prediction. The prediction pipeline ran uninterrupted throughout the recording, exactly as in the main experiment. The algorithm was expected to maintain the current hand state, in agreement with the Control Hand of the VHI, across the entire gesture sequence: any state transition fired during a wrist-gesture window was treated as a false positive. As before, the predicted hand of the VHI was visible to the subject for the duration of the test; the subject was instructed to mirror only the cued hand state shown by the Control Hand at the block boundaries, and not to adapt their muscle activations based on the predicted hand. For this phase, performance is quantified by the *Wrist Distractor Rejection Rate* (WDRR), defined as the fraction of wrist-gesture events during which the algorithm correctly maintained the hand state without firing any spurious transition:

$$\text{WDRR} = \frac{TN_{\text{wrist}}}{TN_{\text{wrist}} + FP_{\text{wrist}}} = 1 - \frac{FP_{\text{wrist}}}{N_{\text{wrist}}}, \quad (2.6)$$

2.3.11 Fingers Movements False Positives Test

The finger-distractor test was included to probe the most demanding false-positive scenario for the algorithm. Wrist movements, tested in Section 2.3.10, share part of their innervation with the hand flexors, but can still be spatially separated to some extent, since the wrist flexor compartment occupies a different region of the forearm than the finger flexors. Individual finger movements, on the other hand, engage exactly the muscles whose activation patterns the templates were trained to detect:

the act of closing the hand is itself a synergistic combination of single-finger flexions, and an isolated movement of one finger produces a channel-activation pattern that is partially indistinguishable from a full hand closure. Finger distractors are therefore the hardest case for the Spatial DTW, and the FP rates reported here are expected to be more pronounced than those of the wrist test. The test was performed at the same two ADL positions used in Section 2.3.10, position B (table reach) and position G (hand to mouth), and followed the same overall block structure: two blocks of opposite hand states at each position, separated by an in-place open↔close transition animated by the Control Hand, with the block order inverted between the two positions exactly as in Table 2.3. In this case, the PI was used only to display the current position, while the source of ground truth returned to being the Control Hand. For the first time since the start of the experiment, the Control Hand showed something different from the two OPEN and CLOSED states of the 1-DoF project: single-finger movements. The subject was required to follow the exact movements shown in the VHI; the pool of gestures was therefore split into two based on the overall, expected state.

- **OPEN block:** five single-finger flexions (thumb, index, middle, ring, pinky) plus the pinch gesture (thumb–index flexion) performed twice, for a total of seven finger-gesture events per block. The pinch is doubled because it is by far the most common voluntary finger movement performed from an open hand in daily life. Moreover, particular attention was given to the pinch gesture, as it was considered the most likely gesture to be added next to the framework if attempting to add one more DoF to the hand orthosis. It is therefore important that the algorithm can fully discern the pinch gesture from a closing gesture starting from the open state.
- **CLOSED block:** five single-finger extensions (thumb, index, middle, ring, pinky), for five finger-gesture events per block. The closed-state counterpart of the pinch, a thumb-index extension from a closed hand, has no daily-life equivalent and was deliberately excluded to keep the test focused on physiologically plausible motions. Even some gestures, like opening only the ring or pinky finger while keeping the rest of the hand closed, are not considered ADL; however, it was considered interesting to check the robustness against those distractors.

The order of the gestures within each block was randomised at every block; no adjacency constraint was applied, since no gesture is repeated within a single block. Each finger-gesture event consisted of five sub-phases: a 0.3 s reaction pause, a 0.9 s animation showing the gesture being performed, a 1.5 s hold during which the subject mirrored the gesture, a 0.9 s animation returning the Control Hand to neutral, and a 0.6 s rest before the next event, for a total of 4.2 s per gesture event. The recording ran again uninterrupted, and S1, S2, S4 performed this stage

only once, while for S3, S5–S16 it was repeated for two rounds, yielding a total of 48 finger-distractor events per subject ((7 OPEN-block + 5 CLOSED-block) $\times$ 2 positions $\times$ 2 rounds). Performance is quantified by the *Finger Distractor Rejection Rate* (FDRR), defined analogously to the WDRR of Section 2.3.10, but separated for each finger:

$$\text{FDRR} = \frac{TN_{\text{finger}}}{TN_{\text{finger}} + FP_{\text{finger}}} = 1 - \frac{FP_{\text{finger}}}{N_{\text{finger}}}, \quad (2.7)$$

where N_{finger} is the total number of finger-gesture events in the recording, FP_{finger} is the number of those events that contained at least one spurious state transition, and TN_{finger} is the number of clean gesture events.

2.3.12 The Mug-Holding Experiment

The fourth and final block of the testing phase was the mug-holding experiment, designed to complement the open-loop cued protocols with a closed-loop, goal-directed real-life scenario. In the three preceding tests the Predicted Hand was visible on the screen but had no influence on the subject’s behaviour, who was instructed to mirror only the Control Hand. In the mug-holding experiment, by contrast, the Control Hand was inactive, no audio cues were played, no reference movement was animated, and the only visual feedback on screen was the Predicted Hand of the [VHI](#), which reflected in real time the algorithm’s interpretation of the subject’s intent. The task was structured as a continuous sequence of steps performed without interruption: starting from a relaxed posture with the hand open, the subject reached for a real mug placed on the table in front (approximately at position B), closed the hand around the mug handle, lifted the mug, and brought it close to the mouth (approximately at position G), faking the act of drinking; before returning the mug to the table, the subject was free to perform any movement that came to mind, some examples being cheering, fake-pouring liquid into another imaginary glass, waving, or dancing. Once the subject returned the mug to the table, they finally opened the hand to release the handle. No time constraint was imposed on any of the five steps, and the subject was free to perform the sequence at a natural pace. No formal quantitative metric is computed for this test. Success was evaluated qualitatively by visual inspection of the prediction recording and of the video footage, by checking whether the subject was able to complete the full five-step sequence without spurious hand-state changes that would have compromised the task. The predicted hand should close only once the hand was fully wrapped around the handle, and should remain closed for the whole time in which the mug is lifted; finally, it should open only once the mug is back on the table and the human hand is opening, releasing the handle. The mug-holding test serves as a usability sanity check rather than a primary quantitative result, and is included to demonstrate that the algorithm’s behaviour remains intuitive

and usable in a continuous-control, real-life scenario outside the structured cued protocols. The mug used in the experiment is a traditional Bavarian mug, made of ceramic, weighing 1109(1) g, Fig. 2.14 shows a photo of it.



Figure 2.14. The traditional Bavarian ceramic mug used in the mug-holding experiment.

2.3.13 Tests on Patients

Due to the impairment of their hand and the rest of the body, the SCI patients performed a simplified version of the experiment, with a subject-specific number of trials performed until fatigue was reached. Only the test from the main phase (described in Section 2.3.9) was administered, maintaining the three-transitions-per-visit structure but performed entirely in the single, comfortable position found at the beginning of the experimental phase. This test serves as a proof of concept, to establish whether the Spatial DTW algorithm can also work on EMG data from a clinical population, while placing greater emphasis on the results from the healthy population for the purposes of future studies.

Chapter 3

Results

The results presented in this chapter were first extracted from the prediction recordings with an automatic algorithm, which checked the alignment between the binary ground-truth signal of the Control Hand and the binary prediction signal of the algorithm, expecting the latter to be a shifted version of the former due to the human reaction time summed with the algorithm latency. After this first screening, each automatically-generated verdict was then manually verified against the synchronised video recording, and corrected when necessary, according to the rules described in the previous sections. Two classes of events are discarded from the confusion matrix that drives the metrics. First, if video inspection revealed that the subject did not correctly follow the virtual ground-truth hand (for example, mistakenly opening the hand during a change of position with the hand held closed), the event was attributed to human error and excluded from the analysis. Second, if the predicted hand was already in the target state before the start of a transition cue, as a consequence of a missed event or instability at the previous cue, the current cue was discarded, since no meaningful detection could occur; the earlier cue responsible for the premature state change was still scored as an error in its own right.

3.1 Main Experiment Results

Across the 16 healthy subjects of the main cohort (32 recordings; 1905 retained cued transitions and 608 repositioning windows), Spatial DTW reached a Correct Detection Rate of 98.4%, a Recall of 99.6%, and a Precision of 98.0%, for a cohort F1-score of 98.8%. Fig. 3.1 shows the averaged Correct Detection Rate per subject, highlighting that for 8 subjects out of 16 all the transitions shown at screen during the experiment were correctly detected; in Fig. 3.2 the same results are averaged per position, showing that all the cues from position A and I were detected correctly, while the most difficult resulted to be position H and J. Fig. 3.3 and Fig. 3.4

report the error breakdown, again first per subject and per positions, distinguishing between errors caused by missed events and by instabilities of the prediction. This results highlight that the most frequent cause of error were instabilities rather than missed events; however, almost half of all the instabilities observed are attributed to S2 alone, whose unstable prediction were observed always in the transition from closed to open, while keeping the open state at the end. Finally, Fig. 3.5 shows the nature of the different errors, pooled together, confirming the most frequent was instability during the transition from closed to open: the opening was correctly detected most of the time, but the prediction was not robust enough and the state oscillated back to closed even though the subject was keeping the hand open. Regarding the second aim of this main experiment, the robustness test against ADL arm movements, out of the 608 repositioning windows across the trials, the algorithm mistakenly predicted a change of state only 16 times, for a cohort Arm Movements False Positives Rejection Rate of 97.4%. Fig. 3.6 shows a per-subject summary of those errors. Once again, most of the errors came from S2 while in the open state. The pie diagram in Fig. 3.7 shows that, in general, the open state was the unstable one; only S13 had problems with false positives while keeping the hand closed, as will also be shown in the rest of the experiments.

3.1.1 Latency

To compute the algorithm latency, only the correctly detected events were considered. The latency was first computed, as a first approximation, by subtraction between the temporal index of the cued transition of the ground truth (the Control Hand of the VHI) and the temporal index of the associated predicted transition. This computation gave a median of 1.12 s (IQR [0.92, 1.37] s) and a mean of 1.22 ± 0.48 s across all subjects. This latency will be referred to as the *Behavioural Latency*, as it contains not only the true algorithm latency but also the subject's reaction time. Since the effective algorithm latency is the time difference between the algorithm's prediction of the movement and the moment at which the subject actually starts moving the hand, the same max-energy onset-detection formula used to extract the templates (Equation 2.1 of Section 2.2.2) was used to find the point at which the subject started moving their hand, based on their EMG activity. A per-channel RMS envelope was computed over a 50 ms sliding window in the interval following the cue, and the start of the 1-second window with the highest cumulative RMS² across the active channels was taken as the activation onset. Unlike the template-extraction procedure, no 200 ms back-shift was applied. The algorithm latency thus computed had a cohort median of 0.53 s (IQR [0.32, 0.80] s), with a mean of 0.61 ± 0.41 s. Fig. 3.8 shows the latency box-plots for each subject, while Fig. 3.9 pools the latencies across subjects and reports the box-plots per position. Across positions, the algorithm latency was relatively uniform, with per-position medians ranging from 0.46 s (position F, lateral pass) to 0.62 s (position H, hand to

Per-subject Correct Detection Rate

S1	118/119	99.2
S2	102/115	88.7
S3	117/120	97.5
S4	120/120	100.0
S5	120/120	100.0
S6	118/119	99.2
S7	120/120	100.0
S8	113/116	97.4
S9	114/119	95.8
S10	120/120	100.0
S11	118/119	99.2
S12	120/120	100.0
S13	114/118	96.6
S14	120/120	100.0
S15	120/120	100.0
S16	120/120	100.0
Total	1874/1905	98.4

Figure 3.1. Per-subject summary of the results of the main experiment. The middle column reports the number of transitions correctly detected out of all the events retained for the study, for each subject, pooling together all the positions.

head); no arm posture was associated with a markedly slower algorithm response. Per-subject medians ranged from 0.34 s (S12, S14, S15) to 1.30 s (S13), with the cohort distribution skewed by S13's slower reaction profile.

Per-position Correct Detection Rate

A	191/191	100.0
B	187/190	98.4
C	189/191	99.0
D	188/191	98.4
E	186/190	97.9
F	190/192	99.0
G	189/192	98.4
H	183/189	96.8
I	190/190	100.0
J	181/189	95.8
Total	1874/1905	98.4

Figure 3.2. Per-position summary of the results of the main experiment. The middle column reports the number of transitions correctly detected out of all the events retained for the study, for each position, pooling together events from different subjects.

3.2 Robustness to Wrist Movements False Positives

Across the 16 healthy subjects, the wrist-distractor test yielded a cohort Wrist Distractor Rejection Rate of 98.0% for the OPEN state and 99.1% for the CLOSED state, out of 448 wrist-gesture events each. Fig. 3.10 and Fig. 3.11 show the per-gesture breakdown of the rejection rate for the OPEN and CLOSED states respectively, reporting the fraction of correctly rejected events for each of the six wrist distractor types. Fig. 3.12 and Fig. 3.13 report the same rejection rate broken down per subject, for the OPEN and CLOSED states respectively. Finally, Fig. 3.14 and Fig. 3.15 show the distribution of the false positives across the six gesture types, pooling all subjects together, for the OPEN and CLOSED states respectively. The CLOSED state showed a higher rejection rate than the OPEN state, with the hand mistakenly opening from wrist movements only four times, and only for S13. For the OPEN state, no specific wrist movement showed a particular tendency to cause

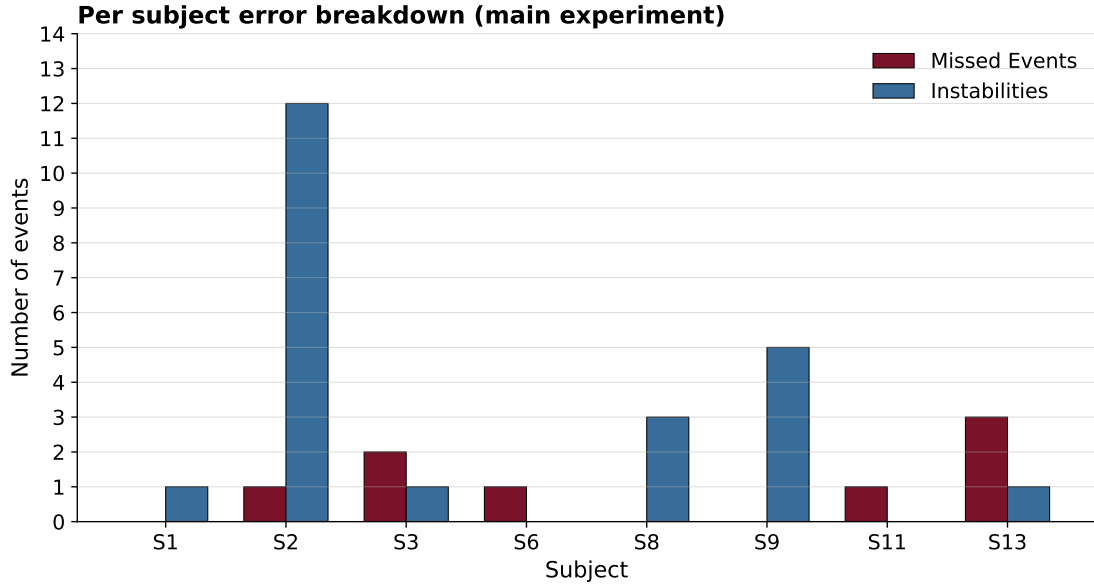


Figure 3.3. Per-subject error breakdown. Each subject has two bars, one counting the errors due to missed events (red), one counting the instabilities (blue).

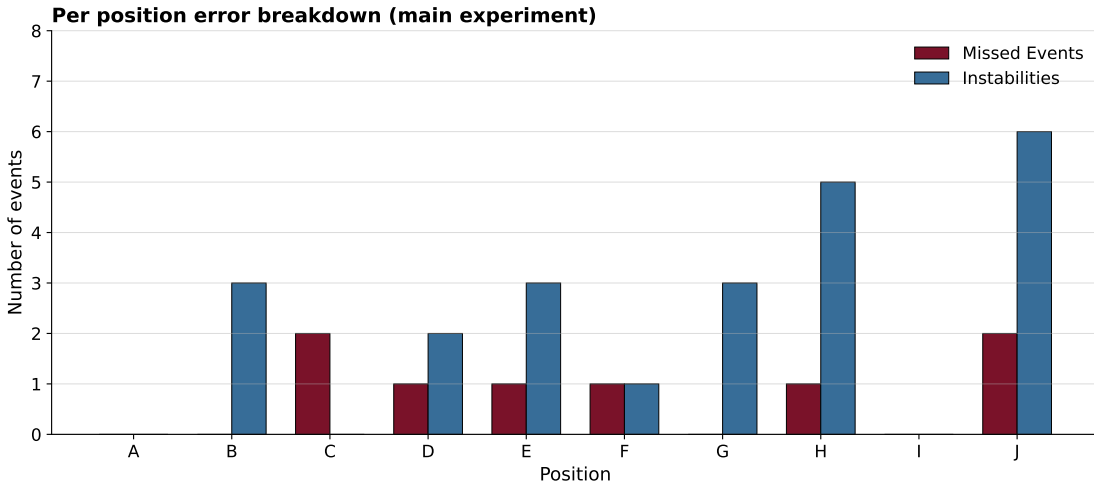


Figure 3.4. Per-position error breakdown. Each position has two bars, one counting the errors due to missed events (red), one counting the instabilities (blue).

false positives, however, as reported in 3.12, the open state rejection rate is inflated by S2 unstable open state, being responsible for six out of the nine total false positives observed in the open state.

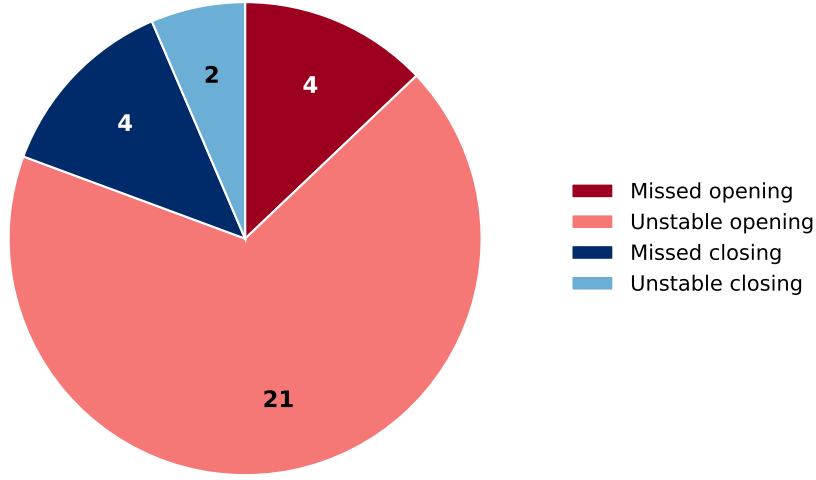


Figure 3.5. Frequency error breakdown. Out of all the errors encountered in the main experiment (across all subjects), this pie diagram breaks down which were the most frequent errors.

3.3 Robustness to Finger Movements False Positives

The finger-distractor test yielded a cohort Finger Distractor Rejection Rate of 96.1% for the OPEN state and 78.4% for the CLOSED state. Again, Figs. 3.16 and 3.17 break these rates down by individual finger, for the OPEN and CLOSED states respectively, while Figs. 3.18 and 3.19 report the same rejection rates per subject. The OPEN state proved overall more robust to false positives caused by finger movements: when the subject’s hand was open, performing a pinch, or flexing the thumb, the index, or the middle finger, never caused the predicted hand to close throughout the whole experiment. For finger movements performed with the hand closed, only for S3 the hand never mistakenly opened, while all other subjects had some errors. Figs. 3.20 and 3.21 show how the false positives are distributed across the five fingers, pooling all subjects together, for both classes. The results clearly show that ring and pinky movements were, for both classes, the main source of false positives; extending the ring finger while keeping the hand closed alone generated an opening of the predicted hand 36 times out of 54, for a ring-movement rejection rate of only 33.3% when the hand was closed.

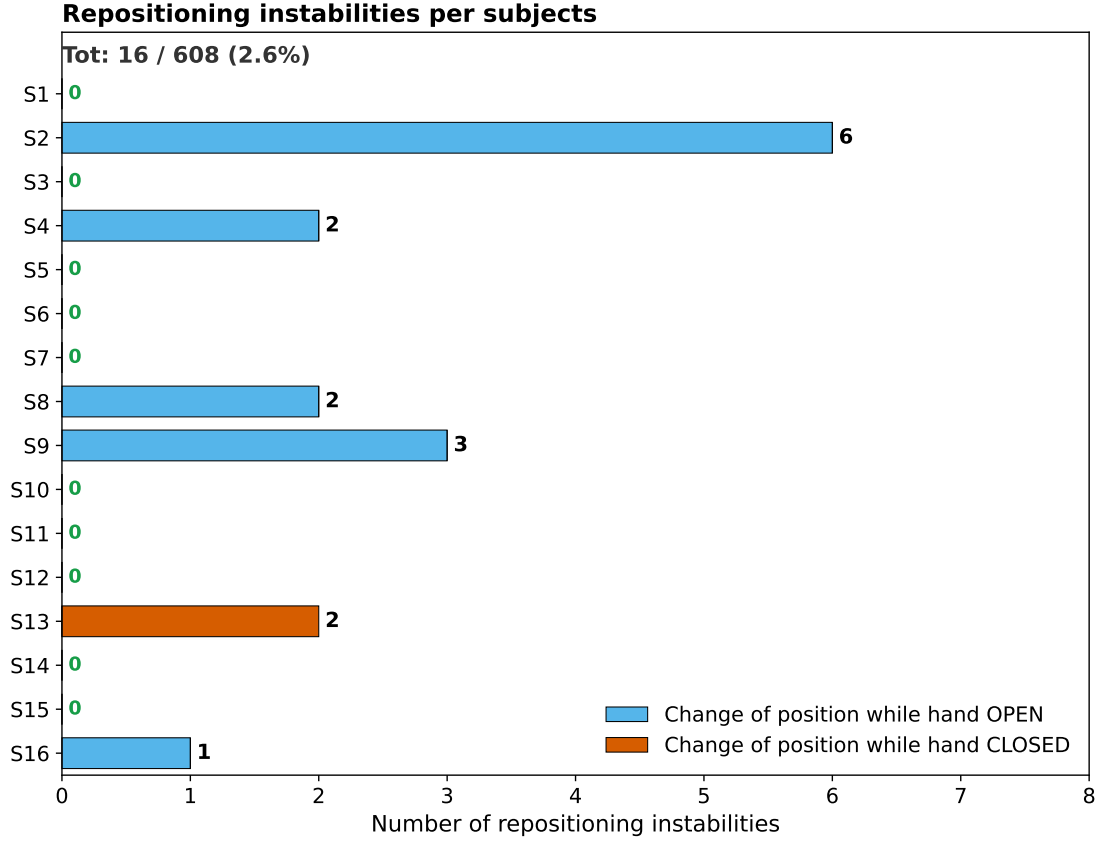


Figure 3.6. Repositioning errors breakdown: how the repositioning errors were distributed across the subjects.

3.4 Mug Holding Results

The mug-holding experiment was evaluated qualitatively, as described in Section 2.3.12. 12 of the 16 subjects completed the full five-step sequence without any spurious hand-state change on their first attempt; subjects 4, 15, and 16 required a second trial before passing, changing their grip on the mug. Fig. 3.22 reports, per subject, which of the additional free-choice activities were attempted and whether the predicted hand remained stable throughout. Only S13 was not able to complete the ADL gesture of drinking and putting down the mug without the algorithm mistakenly predicting an opening, even after a second attempt. Both S5 and S6, when performing the experiment again after a week (more details in the next section), passed again this test.

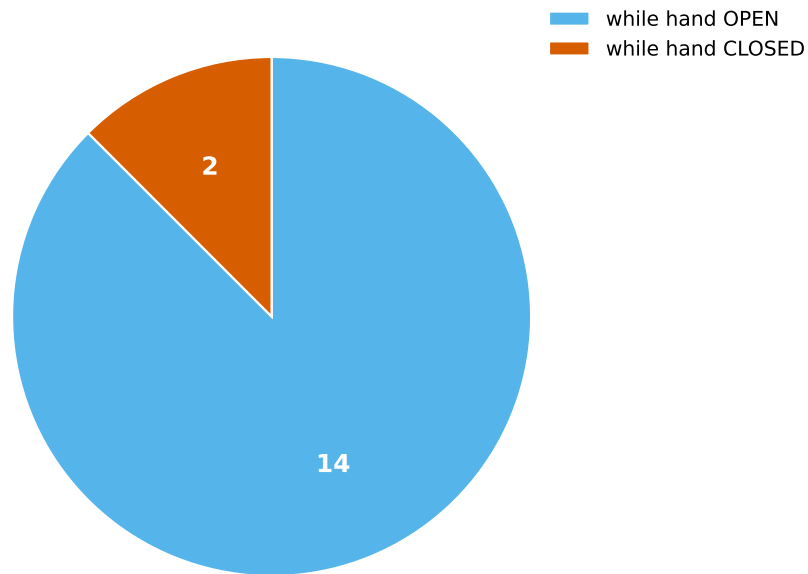
Repositioning false positives

Figure 3.7. Repositioning errors pie diagram. Out of all the repositioning errors, this diagram shows which occurred during the OPEN and CLOSED states. The total number of events is 608.

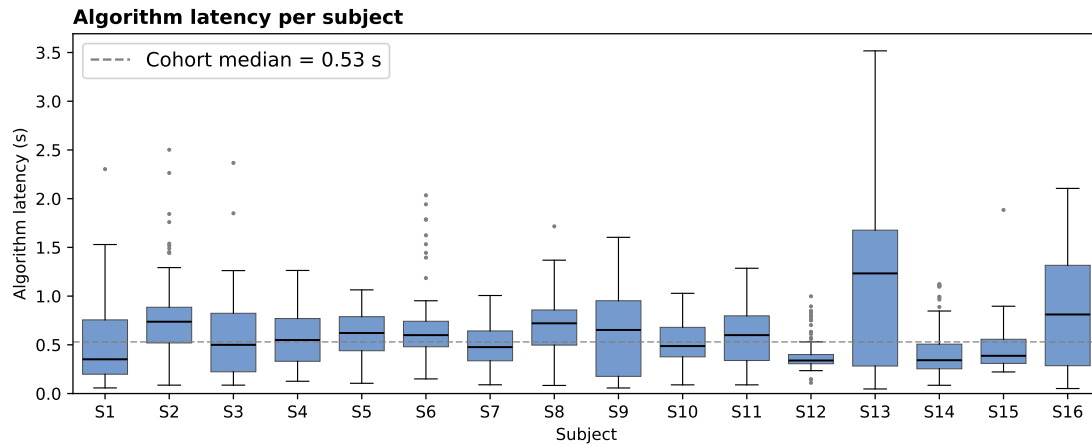


Figure 3.8. Algorithm latency across subjects, computed only for correctly detected transitions.

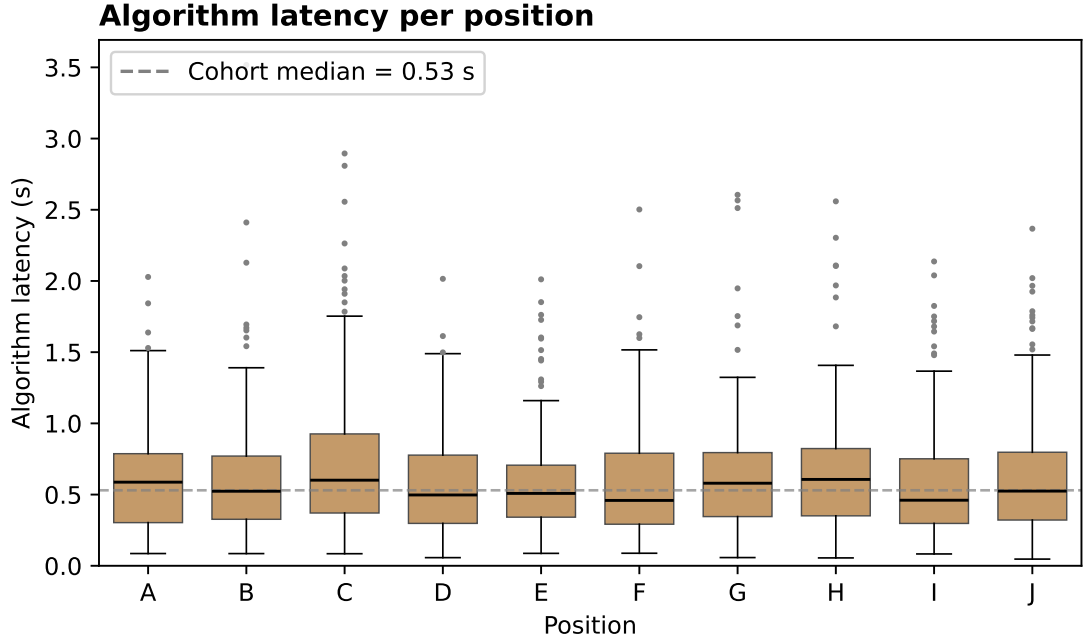


Figure 3.9. Algorithm latency across positions. Values from all subjects are pooled per position to check whether detection was slower for certain postures.

Wrist movements rejection rate - hand OPEN

	Flexion	Extension	Supination	Pronation	Radial-flex	Radial-ext	Total (%)
FPs / Total	2/112	1/112	1/56	2/56	2/56	1/56	9/448
Rejection	98.2%	99.1%	98.2%	96.4%	96.4%	98.2%	98.0%

Figure 3.10. Wrist distractors - per-gesture robustness for the OPEN state.

3.5 Temporal Stability

As a pilot study, S5 and S6 repeated the same three experimental phases approximately a week after the first session (S5 after six days, S6 after a week), using the same calibration data (same model, same templates) as the first session. The bracelet was worn on the same arm, in approximately the same position as in the first episode: during the first session, photographs of the setup were taken, and on the second attempt the bracelet was placed using those photographs as reference. After a period used to adjust the position if needed and to check that the prediction was correctly responding to the subject's movements, the same original

Wrist movements rejection rate - hand CLOSED

	Flexion	Extension	Supination	Pronation	Radial-flex	Radial-ext	Total (%)
FPs / Total	0/112	3/112	0/56	0/56	1/56	0/56	4/448
Rejection	100.0%	97.3%	100.0%	100.0%	98.2%	100.0%	99.1%

Figure 3.11. Wrist distractors - per-gesture robustness for the CLOSED state.

Per subject wrist rejection rate - hand OPEN

S1	0/16	100.0%
S2	6/16	62.5%
S3	0/16	100.0%
S4	0/16	100.0%
S5	0/32	100.0%
S6	0/32	100.0%
S7	0/32	100.0%
S8	0/32	100.0%
S9	1/32	96.9%
S10	0/32	100.0%
S11	0/32	100.0%
S12	1/32	96.9%
S13	0/32	100.0%
S14	0/32	100.0%
S15	0/32	100.0%
S16	1/32	96.9%
Total	9/448	98.0%

Figure 3.12. Wrist distractors - per-subjects robustness for the OPEN state.

Per subject wrist rejection rate - hand CLOSED

S1	0/16	100.0%
S2	0/16	100.0%
S3	0/16	100.0%
S4	0/16	100.0%
S5	0/32	100.0%
S6	0/32	100.0%
S7	0/32	100.0%
S8	0/32	100.0%
S9	0/32	100.0%
S10	0/32	100.0%
S11	0/32	100.0%
S12	0/32	100.0%
S13	4/32	87.5%
S14	0/32	100.0%
S15	0/32	100.0%
S16	0/32	100.0%
Total	4/448	99.1%

Figure 3.13. Wrist distractors - per-subjects robustness for the CLOSED state.

thresholds and spatial corrections were used, and the subject proceeded to perform all the experimental phases again. For the main experiment, both subjects had no repositioning errors in the original recording, and the same held true for this second trial. Pooling S5 and S6 together, the F1-score was 99.8% in the original session and 99.4% in the second trial, a change driven entirely by S6, since S5 completed the experimental part with 100% accuracy in both sessions, while S6 had a -1.7% overall loss of performance (4 errors, all missed events), and a change in latency reported in Fig. 3.28. Fig. 3.23 shows a breakdown of this main experiment. The results of the robustness-against-false-positives experiments are shown in Figs. 3.24 and 3.25 for the wrist distractors, and in Figs. 3.26 and 3.27 for the finger distractors.

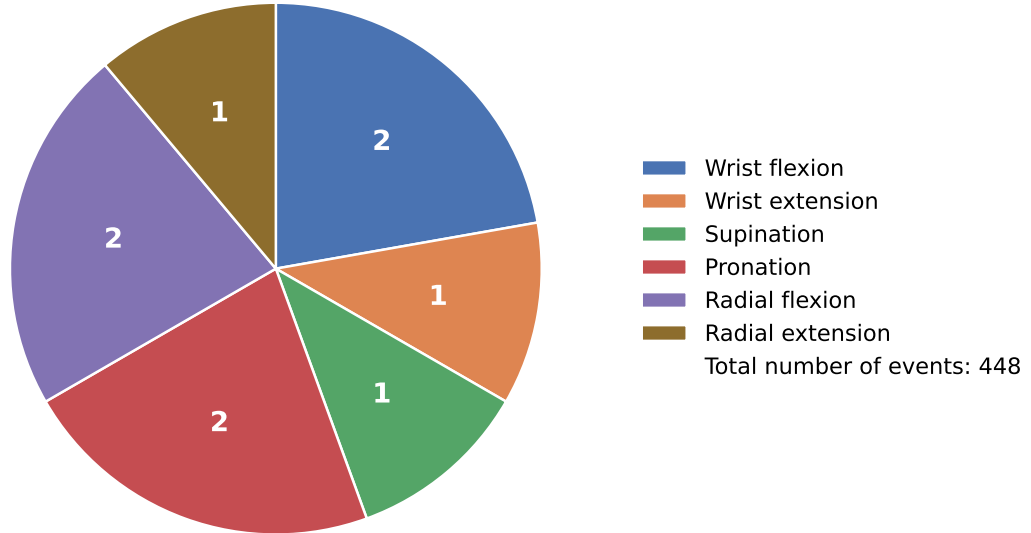
False positives breakdown — hand OPEN (wrist test)

Figure 3.14. Wrist distractors - False positives breakdown for the OPEN state.

Fig. 3.28 gives an insight into the change in latency between the two trials, showing a slightly greater latency on the week-after trial

3.6 Patients Tests

Two SCI patients, P1 and P2, performed the simplified protocol described in Section 2.3.13, without changing arm posture and until they signaled to be fatigued. For P1, 28 of 31 cued transitions were correctly detected (90.3% Correct Detection Rate), with the 3 errors consisting of 2 missed events and 1 instability, and a median algorithm latency of 0.58 s. For P2 22 of 25 cued transitions (88.0% Correct Detection Rate) were correctly detected, with all 3 errors being missed events, and a median algorithm latency of 0.83 s. The complete breakdown is shown in Fig 3.29.

3.7 Offline Results

The results presented in the previous sections were all obtained online, in real time, during the experimental sessions. The remainder of this chapter reports offline analyses, meaning running offline predictions on the same EMG data recorded from the

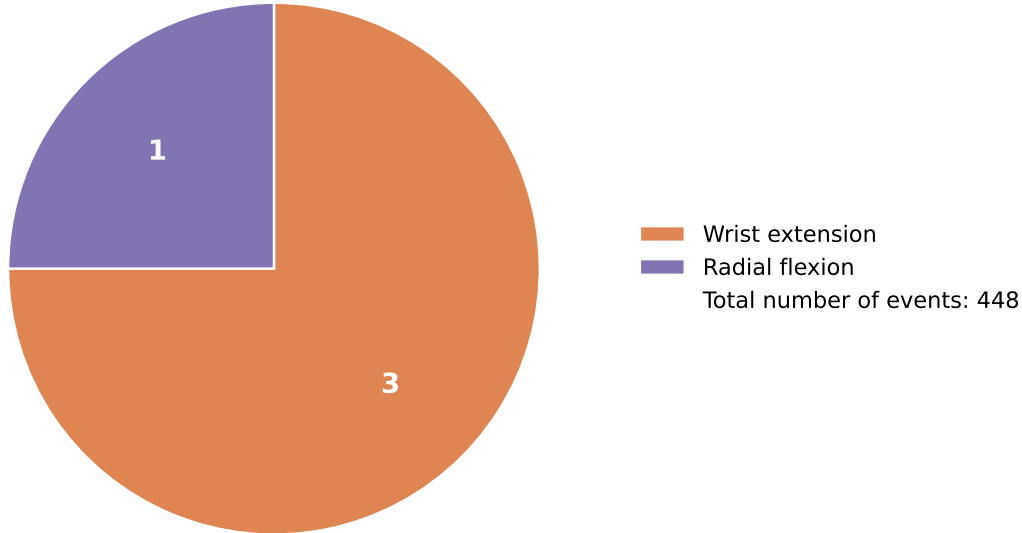
False positives breakdown — hand CLOSED (wrist test)

Figure 3.15. Wrist distractors - False positives breakdown for the CLOSED state.

Finger movements rejection rate - hand OPEN

	Pinch	Thumb	Index	Middle	Ring	Pinky	Total (%)
FPs / Total	0/116	0/57	0/58	0/58	5/59	11/58	16/406
Rejection	100.0%	100.0%	100.0%	100.0%	91.5%	81.0%	96.1%

Figure 3.16. Finger distractors - per-gesture robustness for the OPEN state.

experimental sessions while using either a different classifier (Section 3.7.1), or changing some conditions in the tested classifier, particularly removing different components of the pipeline to assess their real influence in the overall framework (Section 3.7.2). All offline analyses use the same 16-subject cohort and the same event pool as the online results reported above; unlike the online results, however, they were not subjected to the manual video-verification and event-discarding procedure described at the beginning of this chapter, and the total number of events retained may therefore differ slightly from the corresponding online figures.

Finger movements rejection rate - hand CLOSED

	Thumb	Index	Middle	Ring	Pinky	Total (%)
FPs / Total	2/56	1/54	13/53	36/54	7/56	59/273
Rejection	96.4%	98.1%	75.5%	33.3%	87.5%	78.4%

Figure 3.17. Finger distractors - per-gesture robustness for the CLOSED state.

Per subject finger rejection rate - hand OPEN

S1	3/14	78.6%
S2	0/14	100.0%
S3	0/28	100.0%
S4	0/14	100.0%
S5	1/28	96.4%
S6	0/28	100.0%
S7	0/28	100.0%
S8	1/28	96.4%
S9	4/28	85.7%
S10	0/28	100.0%
S11	0/28	100.0%
S12	0/28	100.0%
S13	1/28	96.4%
S14	0/28	100.0%
S15	4/28	85.7%
S16	2/28	92.9%
Total	16/406	96.1%

Figure 3.18. Finger distractors - per-subjects robustness for the OPEN state.

Per subject finger rejection rate - hand CLOSED

S1	1/10	90.0%
S2	2/8	75.0%
S3	0/20	100.0%
S4	2/9	77.8%
S5	4/20	80.0%
S6	1/20	95.0%
S7	5/19	73.7%
S8	4/11	63.6%
S9	2/20	90.0%
S10	2/19	89.5%
S11	5/19	73.7%
S12	5/20	75.0%
S13	8/19	57.9%
S14	4/20	80.0%
S15	9/19	52.6%
S16	5/20	75.0%
Total	59/273	78.4%

Figure 3.19. Finger distractors - per-subjects robustness for the CLOSED state.

3.7.1 CatBoost Comparison

The CatBoost classifier described in Sections 1.6.2 and 2.3.6 was evaluated offline on the same recordings collected during the testing phase, for a direct comparison against the Spatial DTW algorithm. On the main experiment, CatBoost reached a cohort Correct Detection Rate of 87.4% (1611/1844 cued transitions), against the 98.4% achieved by Spatial DTW, a difference ranging from -6.9 to -15.6 percentage points depending on the arm position (Fig. 3.30). On the wrist-distractor test, CatBoost achieved a rejection rate of 53.1% for the OPEN state (210/448 events) and 65.8% for the CLOSED state (153/448 events), against 98.0% and 99.1% for

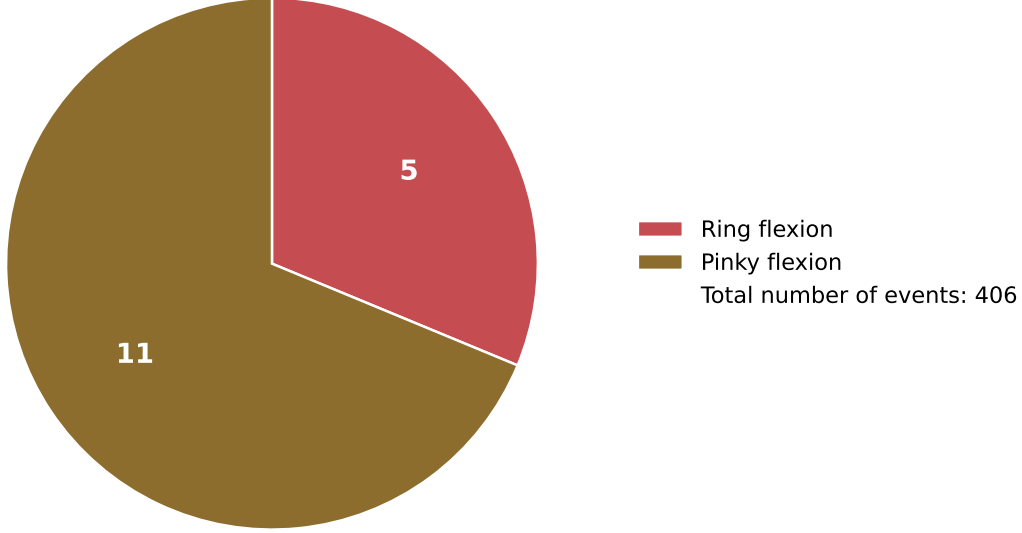
False positives breakdown — hand OPEN (finger test)

Figure 3.20. Finger distractors - False positives breakdown for the OPEN state.

Spatial DTW respectively, corresponding to a difference of -44.9 and -33.3 percentage points (Figs. 3.31 and 3.32). Wrist flexion was by far the most problematic single gesture: with the hand open, it produced a false closing in 98 of 112 events, an 87.5% false-positive rate corresponding to a rejection rate of only 12.5%. On the finger-distractor test, CatBoost achieved a rejection rate of 70.5% for the OPEN state (122/413 events) and 30.6% for the CLOSED state (179/258 events), against 96.1% and 78.4% for Spatial DTW respectively, corresponding to a difference of -25.6 and -47.8 percentage points (Figs. 3.33 and 3.34).

3.7.2 Ablation Study

To assess the individual contribution of each component of the Spatial DTW pipeline, three ablation studies were performed offline on the same testing-phase recordings: removing the spatial correction layer, reducing the number of templates per position, and reducing the number of calibrated arm positions. In each case, the modified pipeline was run on the same offline event pool as its corresponding full-algorithm baseline, so that the reported difference isolates the effect of the single component removed or reduced. Because this offline baseline was not subjected to the manual video-verification procedure described at the start of this chapter,

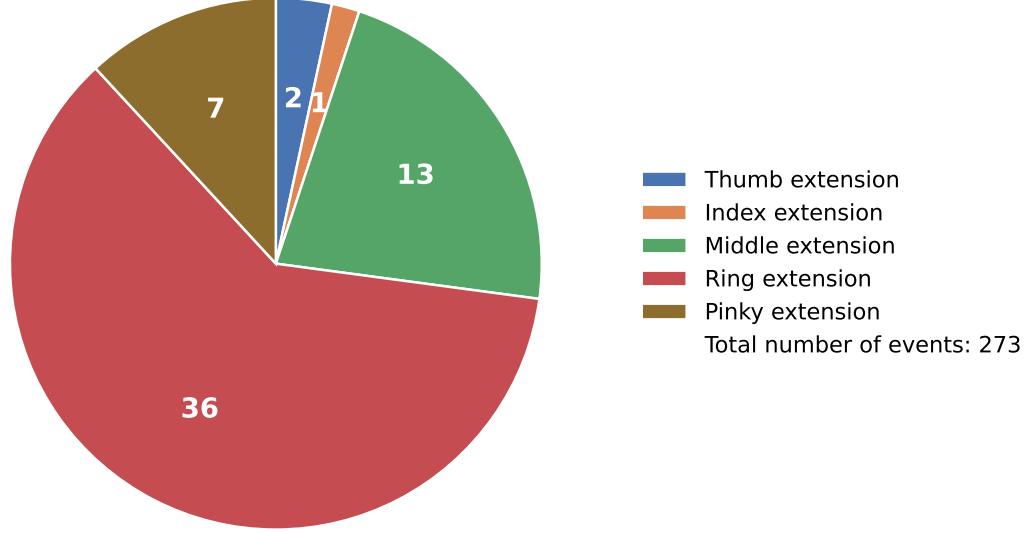
False positives breakdown — hand CLOSED (finger test)

Figure 3.21. Finger distractors - False positives breakdown for the CLOSED state.

all Δ values in this section are therefore computed against the offline baseline and cannot be obtained by subtracting from the online results.

Influence of the Spatial Correction

Removing the spatial correction layer, and relying on the cosine DTW distance alone, reduced the cohort Correct Detection Rate on the main experiment from 98.4% to 91.9% (1728/1880 cued transitions), a drop of 6.5 percentage points, with the largest per-position losses at F (lateral pass, -10.1pp) and I (overhead reach, -5.8pp), and the smallest at G (hand to mouth, -2.1pp) (Fig. 3.35). On the wrist-distractor test, the rejection rate dropped to 89.3% for the OPEN state (-8.2pp) and 93.5% for the CLOSED state (-4.3pp) (Figs. 3.36 and 3.37). On the finger-distractor test, the rejection rate dropped to 85.7% for the OPEN state (-8.5pp) and 70.1% for the CLOSED state (-4.1pp), with the ring finger remaining the single worst-performing gesture in both conditions (29.4% rejection rate with the hand closed) (Figs. 3.38 and 3.39).

Mug experiment results

Subject	Grasp	Lift	Drink	Pour	Cheer	Wave	Release	Test result
S1	✓	✓	✓		✓		✓	Passed
S2	✓	✓	✓	✓	✓	✓	✓	Passed
S3	✓	✓	✓	✓	✓	✓	✓	Passed
S4	✓	✓	✓		✓		✓	Passed*
S5.1	✓	✓	✓	✓	✓		✓	Passed
S6.1	✓	✓	✓	✓	✓	✓	✓	Passed
S7	✓	✓	✓	✓	✓	✓	✓	Passed
S8	✓	✓	✓	✓	✓	✓	✓	Passed
S9	✓	✓	✓	✓	✓	✓	✓	Passed
S10	✓	✓	✓	✓	✓	✓	✓	Passed
S11	✓	✓	✓	✓	✓	✓	✓	Passed
S12	✓	✓	✓	✓	✓	✓	✓	Passed
S13	✓	✓	✓	✓	X	X		Failed
S14	✓	✓	✓	✓	✓	✓	✓	Passed
S15	✓	✓	✓	✓	✓	✓	✓	Passed*
S16	✓	✓	✓	✓	✓	✓	✓	Passed*
S5.2	✓	✓	✓	✓	✓	✓	✓	Passed
S6.2	✓	✓	✓	✓	✓	✓	✓	Passed

Figure 3.22. Mug experiment per subject breakdown. The table shows which activities did every subject attempt during the mug experiment. The tick means the subject did the activity with the prediction never mistakenly opening during the whole recording. The X means the algorithm mistakenly predicted an opening while the subject was doing that specific activity. If in a cell no tick is present it means the subject did not do that activity. On the last column, Passed* means the subject did not pass the experiment at first, but did pass it on a second trial.

Templates Reduction

Two offline analyses assessed whether the four templates per position, per class, used in the online experiments are all necessary. The first analysis checked, for every event in the main experiment, whether the cued arm position's own templates were among the three closest matches used by the distance-aggregation rule (Section 2.2.5). Across correctly detected events, the cued position was absent from the top-3 closest templates in 46.7% of cases; across error events (missed and unstable), it was absent in 44.8% of cases. Fig. 3.40 shows this breakdown per cued position. The second analysis re-decoded all recordings while progressively reducing the number of templates retained per position and class, from the full

Temporal stability - main experiment per session breakdown											
S5.1	12/12	12/12	12/12	12/12	12/12	12/12	12/12	12/12	12/12	12/12	100.0
S5.2	11/11	12/12	12/12	12/12	12/12	12/12	12/12	12/12	12/12	12/12	100.0
Δ %	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
S6.1	12/12	12/12	12/12	12/12	12/12	12/12	12/12	12/12	12/12	10/11	99.2
S6.2	12/12	11/12	12/12	12/12	12/12	12/12	10/11	12/12	12/12	10/11	97.5
Δ %	0.0	-8.3	0.0	0.0	0.0	0.0	-9.1	0.0	0.0	0.0	-1.7
	A	B	C	D	E	F	G	H	I	J	Total (%)

Figure 3.23. Comparison of the performance in the second trial of the main experiment for S5 and S6, divided per position. The last row of each table reports the difference in false-positive rejection rate with respect to the first test. No repositioning errors occurred, matching the first experiment. The first event of the second trial of S5 happened when the subject wasn't ready therefore the event was discarded.

Temporal stability - wrist movements rejection - hand OPEN							
S5.1	0/8	0/8	0/4	0/4	0/4	0/4	100.0
S5.2	0/8	0/8	0/4	0/4	0/4	0/4	100.0
Δ %	0.0	0.0	0.0	0.0	0.0	0.0	0.0
S6.1	0/8	0/8	0/4	0/4	0/4	0/4	100.0
S6.2	0/8	0/8	0/4	0/4	0/4	0/4	100.0
Δ %	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	Flex	Ext	Sup	Pron	Rad-flex	Rad-ext	Total (%)

Figure 3.24. Comparison of the performance in the second trial of the wrist distractors experiment for S5 and S6, with the hand OPEN. The last row of each table reports the difference in false-positive rejection rate with respect to Fig. 3.10.

four down to one. The first templates to be removed from the set for each position were the ones progressively closer to being outliers, following the evaluation criteria described in Section 2.2.2. The cohort Correct Detection Rate was 98.4% with all four templates, 98.3% with three ($\Delta - 0.1\text{pp}$), 97.0% with two ($\Delta - 1.4\text{pp}$), and 93.6% with one ($\Delta - 4.8\text{pp}$) (Fig. 3.41).

Temporal stability - wrist movements rejection - hand CLOSED

S5.1	0/8	0/8	0/4	0/4	0/4	0/4	100.0
S5.2	0/8	0/8	0/4	0/4	0/4	0/4	100.0
Δ %	0.0	0.0	0.0	0.0	0.0	0.0	0.0
S6.1	0/8	0/8	0/4	0/4	0/4	0/4	100.0
S6.2	0/8	0/8	0/4	0/4	0/4	0/4	100.0
Δ %	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	Flex	Ext	Sup	Pron	Rad-flex	Rad-ext	Total (%)

Figure 3.25. Comparison of the performance in the second trial of the wrist distractors experiment for S5 and S6, with the hand CLOSED. The last row of each table reports the difference in false-positive rejection rate with respect to Fig. 3.11.

Temporal stability - finger movements rejection - hand OPEN

S5.1	0/8	0/4	0/4	0/4	1/4	0/4	96.4
S5.2	0/8	0/4	0/4	0/4	0/4	1/4	96.4
Δ %	0.0	0.0	0.0	0.0	+25.0	-25.0	0.0
S6.1	0/8	0/3	0/4	0/4	0/5	0/4	100.0
S6.2	0/8	0/4	0/4	0/4	0/4	0/4	100.0
Δ %	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	Pinch	Thumb	Index	Middle	Ring	Pinky	Total (%)

Figure 3.26. Comparison of the performance in the second trial of the finger distractors experiment for S5 and S6, with the hand OPEN. The last row of each table reports the difference in false-positive rejection rate with respect to Fig. 3.16.

Positions Reduction

A separate offline analysis identified five arm positions, A, B, F, G, and I, as the most influential (and frequent) in the state-machine decisions. The main experiment was then re-decoded offline using only the templates calibrated at these five positions, instead of all ten. This reduced the cohort Correct Detection Rate from

Temporal stability - finger movements rejection - hand CLOSED						
S5.1	0/4	0/4	0/4	4/4	0/4	80.0
S5.2	0/4	0/4	0/4	4/4	1/4	75.0
Δ %	0.0	0.0	0.0	0.0	-25.0	-5.0
S6.1	0/4	0/4	1/4	0/4	0/4	95.0
S6.2	0/4	0/4	3/4	1/4	0/4	80.0
Δ %	0.0	0.0	-50.0	-25.0	0.0	-15.0
	Thumb	Index	Middle	Ring	Pinky	Total (%)

Figure 3.27. Comparison of the performance in the second trial of the finger distractors experiment for S5 and S6, with the hand CLOSED. The last row of each table reports the difference in false-positive rejection rate with respect to Fig. 3.17.

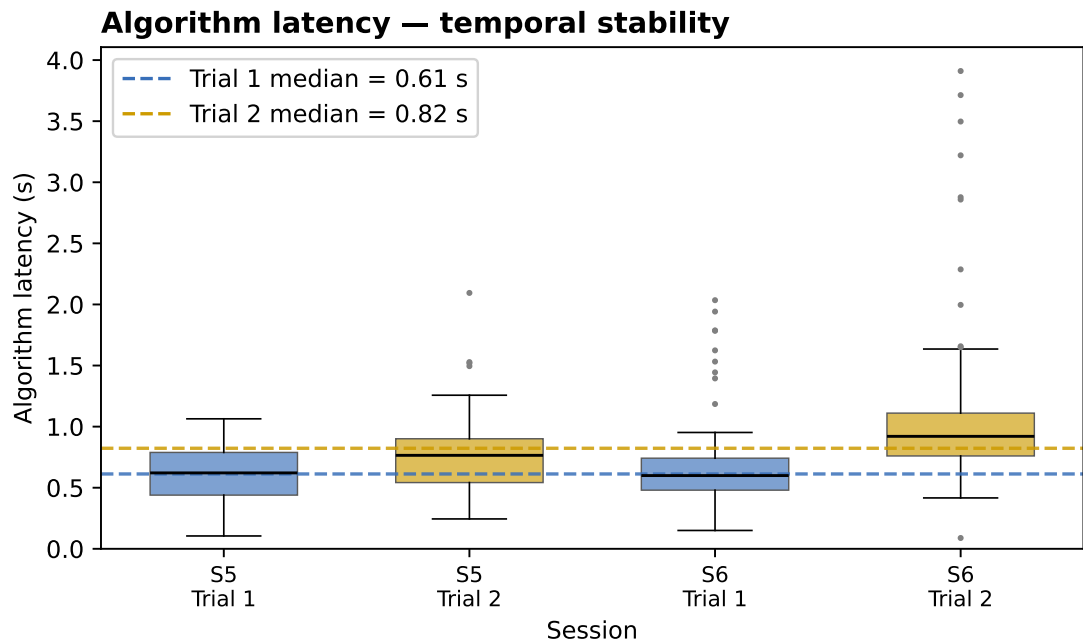


Figure 3.28. Latency comparison. Changes in the algorithm latency distribution between the second trials of the main experiment for S5 and S6.

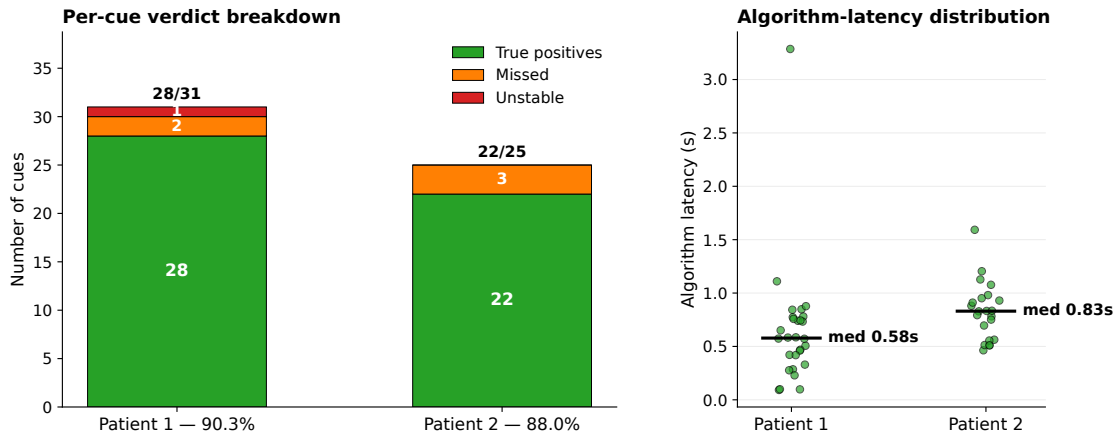


Figure 3.29. Patients results. The protocol for patients was a simplified version, without changes of position. Each patient kept following the ground truth until declared to be fatigued. Left: breakdown of the prediction of all the cues the patient followed. Right: algorithm latency distribution in every event

CatBoost comparison - Main experiment Correct Detection Rate per arm position

	A	B	C	D	E	F	G	H	I	J	Total (%)
TP / Total	163/185	167/186	162/187	172/188	149/181	159/187	163/183	155/184	168/183	153/180	1611/1844
CDR (%)	88.1	89.8	86.6	91.5	82.3	85.0	89.1	84.2	91.8	85.0	87.4
Δ % (vs sDTW)	-11.9	-8.6	-12.3	-6.9	-15.6	-13.9	-9.4	-12.6	-8.2	-10.8	-11.0

Figure 3.30. CatBoost comparison, main experiment Correct Detection Rate per arm position, against the Spatial DTW baseline.

98.4% to 96.0% (1796/1870 cued transitions), a decrease of 2.5 percentage points, with the largest per-position loss at H (hand to head, -4.5pp) and the smallest at B (table reach, -0.5pp) (Fig. 3.42).

CatBoost comparison - Wrist movements rejection rate - hand OPEN

	Flex	Ext	Sup	Pron	Rad-flex	Rad-ext	Total (%)
FP / Total	98/112	49/112	8/55	9/56	13/57	33/56	210/448
Rejection (%)	12.5	56.2	85.5	83.9	77.2	41.1	53.1
Δ % (vs sDTW)	-85.7	-42.9	-12.8	-12.5	-19.2	-57.1	-44.9

Figure 3.31. CatBoost comparison, wrist movements rejection rate per gesture, hand OPEN, against the Spatial DTW baseline.

CatBoost comparison - Wrist movements rejection rate - hand CLOSED

	Flex	Ext	Sup	Pron	Rad-flex	Rad-ext	Total (%)
FP / Total	21/113	57/111	23/56	9/56	24/56	19/56	153/448
Rejection (%)	81.4	48.6	58.9	83.9	57.1	66.1	65.8
Δ % (vs sDTW)	-18.6	-48.7	-41.1	-16.1	-41.1	-33.9	-33.3

Figure 3.32. CatBoost comparison, wrist movements rejection rate per gesture, hand CLOSED, against the Spatial DTW baseline.

CatBoost comparison - Finger movements rejection rate - hand OPEN

	Pinch	Thumb	Index	Middle	Ring	Pinky	Total (%)
FP / Total	16/118	7/58	19/59	21/59	29/60	30/59	122/413
Rejection (%)	86.4	87.9	67.8	64.4	51.7	49.2	70.5
Δ % (vs sDTW)	-13.6	-12.1	-32.2	-35.6	-39.9	-31.9	-25.6

Figure 3.33. CatBoost comparison, finger movements rejection rate per gesture, hand OPEN, against the Spatial DTW baseline.

CatBoost comparison - Finger movements rejection rate - hand CLOSED

	Thumb	Index	Middle	Ring	Pinky	Total (%)
FP / Total	22/52	28/52	36/48	44/52	49/54	179/258
Rejection (%)	57.7	46.2	25.0	15.4	9.3	30.6
Δ % (vs sDTW)	-38.7	-52.0	-50.5	-17.9	-78.2	-47.8

Figure 3.34. CatBoost comparison, finger movements rejection rate per gesture, hand CLOSED, against the Spatial DTW baseline.

Spatial-correction ablation - Main experiment Correct Detection Rate per arm position

	A	B	C	D	E	F	G	H	I	J	Total (%)
TP / Total	175/190	178/190	176/188	177/190	167/185	167/188	185/192	161/183	179/190	163/184	1728/1880
CDR (%)	92.1	93.7	93.6	93.2	90.3	88.8	96.4	88.0	94.2	88.6	91.9
Δ % (vs sDTW)	-7.9	-4.7	-5.3	-5.3	-7.6	-10.1	-2.1	-8.8	-5.8	-7.2	-6.5

Figure 3.35. Spatial-correction ablation, main experiment Correct Detection Rate per arm position, against the full Spatial DTW algorithm.

Spatial-correction ablation - Wrist movements rejection rate - hand OPEN

	Flexion	Extension	Supination	Pronation	Radial-flex	Radial-ext	Total (%)
FP / Total	16/140	11/140	10/69	8/70	10/71	5/70	60/560
Rejection (%)	88.6	92.1	85.5	88.6	85.9	92.9	89.3
Δ % (vs sDTW)	-10.0	-6.4	-11.6	-7.1	-9.9	-4.3	-8.2

Figure 3.36. Spatial-correction ablation, wrist movements rejection rate per gesture, hand OPEN, against the full Spatial DTW algorithm.

Spatial-correction ablation - Wrist movements rejection rate - hand CLOSED

	Flexion	Extension	Supination	Pronation	Radial-flex	Radial-ext	Total (%)
FP / Total	4/139	14/139	4/70	4/70	7/70	3/69	36/557
Rejection (%)	97.1	89.9	94.3	94.3	90.0	95.7	93.5
Δ % (vs sDTW)	-2.9	-5.8	-1.4	-4.3	-7.1	-4.3	-4.3

Figure 3.37. Spatial-correction ablation, wrist movements rejection rate per gesture, hand CLOSED, against the full Spatial DTW algorithm.

Spatial-correction ablation - Finger movements rejection rate - hand OPEN

	Pinch	Thumb	Index	Middle	Ring	Pinky	Total (%)
FP / Total	12/142	10/70	6/71	11/71	14/72	18/71	71/497
Rejection (%)	91.5	85.7	91.5	84.5	80.6	74.6	85.7
Δ % (vs sDTW)	-6.3	-10.0	-5.6	-14.1	-9.7	-7.0	-8.5

Figure 3.38. Spatial-correction ablation, finger movements rejection rate per gesture, hand OPEN, against the full Spatial DTW algorithm.

Spatial-correction ablation - Finger movements rejection rate - hand CLOSED

	Thumb	Index	Middle	Ring	Pinky	Total (%)
FP / Total	9/68	9/68	22/65	48/68	13/69	101/338
Rejection (%)	86.8	86.8	66.2	29.4	81.2	70.1
Δ % (vs sDTW)	-4.4	-2.9	-6.2	-1.5	-5.8	-4.1

Figure 3.39. Spatial-correction ablation, finger movements rejection rate per gesture, hand CLOSED, against the full Spatial DTW algorithm.

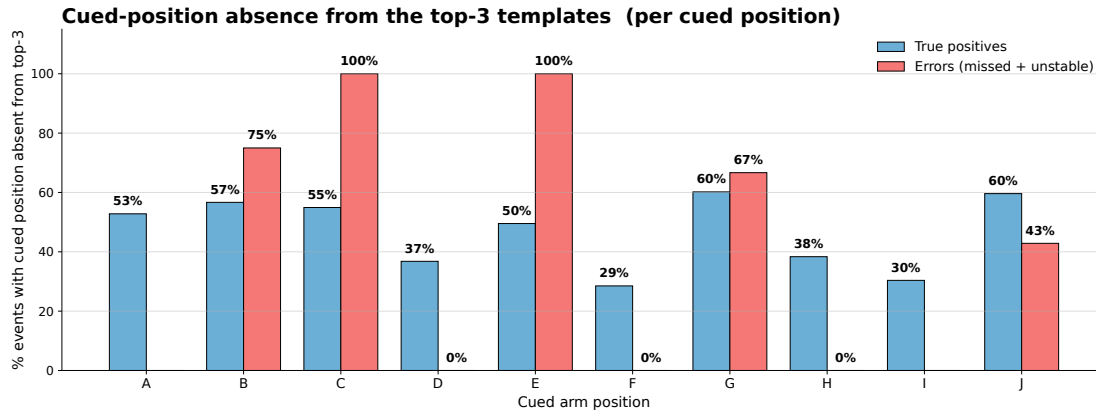


Figure 3.40. Fraction of events, per cued arm position, for which the cued position's own templates were absent from the three closest templates used in the distance aggregation, separated by correctly detected events (true positives) and errors (missed and unstable). Note that several positions have very few error events (e.g. position D: 1/1, position I: 0 errors), so the percentages for these positions should be interpreted with caution.

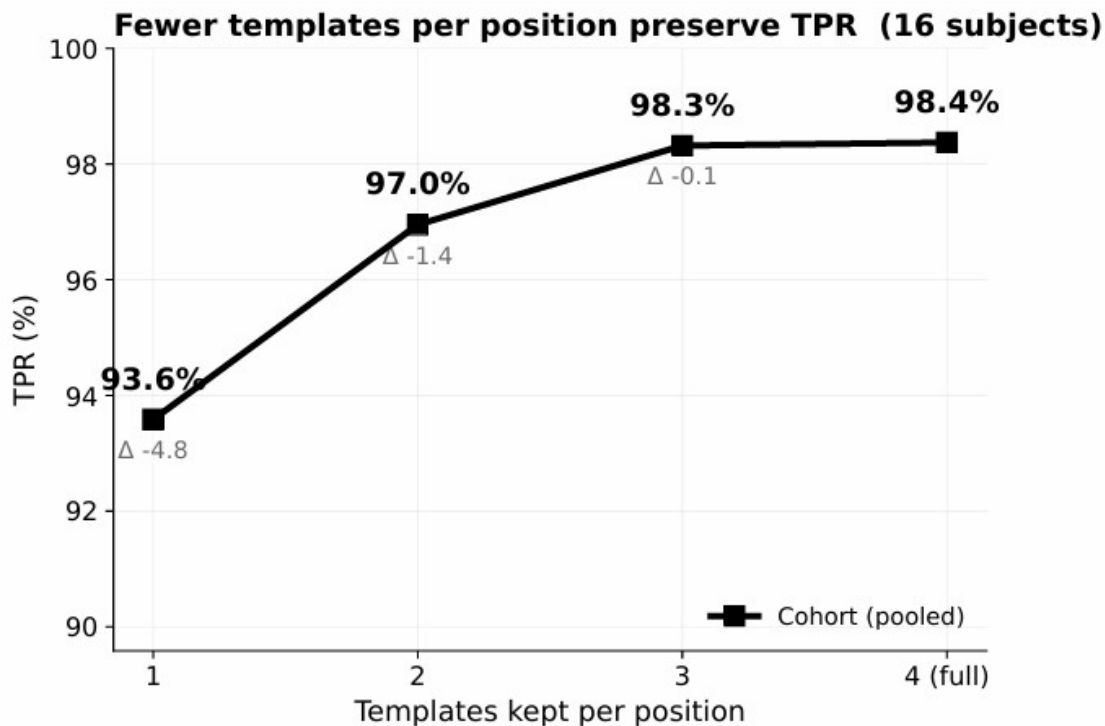


Figure 3.41. Cohort Correct Detection Rate as a function of the number of templates retained per position and class, from the full four down to one.

ABFGI protocol - Phase 2a Correct Detection Rate per arm position

	A	B	C	D	E	F	G	H	I	J	Total (%)
TP / Total	183/187	185/189	178/185	179/190	179/185	185/190	186/191	168/183	182/188	171/182	1796/1870
CDR (%)	97.9	97.9	96.2	94.2	96.8	97.4	97.4	91.8	96.8	94.0	96.0
Δ % (vs sDTW)	-2.1	-0.5	-3.3	-4.2	-1.7	-2.1	-1.1	-4.5	-3.2	-2.3	-2.5

Figure 3.42. ABFGI protocol, main experiment Correct Detection Rate per arm position, using templates from only five arm positions (A, B, F, G, I), against the full ten-position Spatial DTW algorithm.

Chapter 4

Discussion

This thesis presented the novel Spatial DTW, a posture-robust, sEMG-only algorithm for 1-DoF myoelectric control of a hand orthosis, and validated it across 16 healthy subjects, ten ADL arm positions, and two SCI patients. With the main experiment, the algorithm achieved a cohort Correct Detection Rate of 98.4% and an Arm Movements False Positives Rejection Rate of 97.4%, a strong number on its own, given the total number of events (1905 and 608 respectively). Robustness to non-target wrist and finger movements was likewise high, and the algorithm substantially outperformed a CatBoost-based continuous classifier, the existing state of the art within the MyoGestic framework, particularly on distractor rejection. Reduced-calibration variants retained most of this performance, and a pilot one-week retest showed only minor degradation. Two SCI patients achieved 90.3% and 88.0% Correct Detection Rates under a simplified protocol. The remainder of this chapter interprets these results, discusses the algorithm’s failure modes, and situates the findings within the broader myoelectric control literature.

4.1 Robustness Across Activities of Daily-Living

The main experiment’s 98.4% cohort CDR reflects strong posture generalisation, and the per-position error breakdown (Fig. 3.4) shows this is not driven by a small number of catastrophic positions: errors are distributed across most positions, with pos. J (cross-body) and H (hand to head) showing the highest instability counts. J’s relatively poorer performance is plausibly related to it being kinematically the most distinct of the ten positions, a full cross-body reach differs substantially from the other nine, which are largely front-facing or lateral movements, and may therefore be less well covered by templates recorded at neighbouring postures. On the other hand, position H combines shoulder abduction and elbow flexion, which together could be a source of different EMG patterns or increased electrode shift. Two of the four positions with an extended elbow, A (rest) and I (overhead reach), were in

fact the only fully error-free positions in the main experiment; although the other two extended-elbow positions, D (forward reach) and F (lateral pass), still showed a small number of errors, the elbow-flexion may be a contributing factor alongside the compound shoulder-elbow demands specific to position H. The whole experimental design tried to highlight position B (table reach) and G (hand to mouth), the only two positions in which the wrist and finger distractor were tested, and no remarkable instabilities occurred in this two postures; an important result, as this two positions combined represent the act of grasping something on a table and bring it to the mouth, i.e. the acts of drinking and eating, two of the most important ADL activities. Throughout the experiments, spurious spikes in the EMG signal were noted during the sessions for S8, S9, S13, and S16, likely due to loss of contact from one or more electrodes. Since the single-differential derivation is performed in real time, if one of the two longitudinally adjacent electrodes suddenly lost skin contact, its signal shifted to zero, and the difference between the two channels became the monopolar signal of the electrode still in contact. This produced a spike in the EMG, as the monopolar signal has higher amplitude, containing more information as well as common-mode noise. This could plausibly cause latency or missed detections, and is likely a contributing factor to S13's higher latency. S13 was also the only subject to ever produce instabilities during arm movements and wrist movements in the CLOSED state, and the source of this outlier result could be related to its EMG signal quality. The same S13 was in fact the only subject to not pass the mug-holding experiment.

During repositionings, out of 16 errors, 14 occurred when the hand was in the OPEN state (Fig. 3.7); this, paired with the results of the wrist-distractor test (Section 2.3.10), confirms the OPEN state was the less robust of the two. It is important to note, however, that 38.2% of all errors pooled together from the main experiment (18 out of 47) came from instabilities during the OPEN state of S2, who was also the subject achieving the lowest rejection rate of wrist movements while keeping the hand open (62.5%, accounting for 6 of the total 8 errors from all subjects in this phase). Reviewing the videos of the experiment, it can be observed that the subject performed the task of opening the hand and keeping it open with what appeared to be a near-continuous, isometric finger extension, or a co-contraction, almost over-extending the fingers. This behaviour could plausibly explain the OPEN state's lower robustness, if the co-contraction produces an EMG pattern more similar to the activation seen during hand closing. No quantitative EMG-pattern analysis was performed on S2's recordings, however, leaving this an open question.

The wrist- and finger-distractor results point to two qualitatively different failure mechanisms. Wrist movements were rejected well overall, consistent with the premise that the wrist flexor and extensor compartments, while adjacent to the finger flexors, are not spatially and temporally identical to them, even if their position in the forearm overlaps with the finger flexors and extensors. Pilot experiments

carried out during the early stages of this work, using the CatBoost classifier already available within MyoGestic (Section 1.6.2), had flagged wrist movements as a major source of false positives; Spatial DTW, by contrast, was found to reject these movements almost entirely.

Finger movements were markedly harder, particularly with the hand closed (Fig. 3.17), where the ring finger was the single worst-performing gesture (33.3% rejection rate). This is plausibly explained by the limited functional independence of the ring finger relative to the other digits: movements of the thumb, index, and middle finger are typically more individuated than movements of the little or ring fingers (22), consistent with the ring finger’s extension being mediated largely through the same shared extensor mechanism responsible for whole-hand opening, rather than through an independent muscle. Isolating ring extension without partially recruiting this shared mechanism is therefore mechanically difficult, which plausibly explains why the resulting activation pattern is so often mistaken for a genuine hand opening. It is also worth noting that isolated ring-finger extension while the rest of the hand remains closed is not itself a common activity of daily living; this gesture was included in the test battery specifically to probe a worst-case scenario, and its low rejection rate should be weighed accordingly.

The mug-holding experiment (Sections 2.3.12 and 3.4) showed that the algorithm performs correctly even while the hand holds a heavy object, such as the ceramic mug used, without the predicted hand opening. This is a behavioural consequence of the design choice to detect discrete events rather than a continuous prediction: even though holding a mug is intrinsically a muscle activation the algorithm was not explicitly trained for, as long as the activation produced by holding the mug differs from the one produced by opening the hand, which this experiment demonstrates, the predicted hand does not open. However, since the templates were recorded without the subjects holding any object, i.e. "grasping the air", the algorithm trained like that is expected to detect a grasp of a real object only if the object is thin enough so that the movement recreated is similar to the closed templates. It has not been quantified, but in pilot experiments, grabbing a glass, a wide object which does not require the hand to fully close, resulted in most of the times in a missed closing. The EMG-activation generated by simply closing the hand and grasping an actual, wide object is therefore different.

4.2 Comparison with States of the Art

Spatial DTW outperformed the CatBoost baseline by a wide margin on every metric. On the wrist-distractor test, CatBoost’s rejection rate for wrist flexion in the OPEN state was only 12.5% (Fig. 3.31), meaning a voluntary wrist flexion produced a false hand closure in roughly nine out of every ten instances, against a 98.2% rejection rate for the same gesture under Spatial DTW (Fig. 3.10). This gap

is large enough that it should not be read simply as a better classifier, but as a consequence of a fundamental design difference: CatBoost judges each 9 ms window independently, with just a post-prediction filter for some sustained temporal or spatial consistency, while Spatial DTW requires agreement across a full second of signal, five consecutive votes, and spatial consistency with the calibrated templates before committing to a state change. This outperformance was, to some extent, expected: pilot experiments carried out during the early stages of this work, using the CatBoost classifier already available within MyoGestic (Section 1.6.2), had flagged wrist movements as a major source of false positives; Spatial DTW, by contrast, was built to reject these non-target movements and did almost entirely. CatBoost was designed within MyoGestic to maximise the number of decodable degrees of freedom under a continuous-control paradigm, not to reject non-target activity, so the comparison primarily illustrates the cost of that design choice rather than establishing Spatial DTW’s superiority over discrete, temporally-aware methods in general. It is also worth reiterating that the comparison is not symmetric in calibration burden: even though the actual amount of data used by the two algorithms is the same, CatBoost requires only a session of approximately 80 seconds, against roughly 15 minutes of session for the full Spatial DTW protocol, a trade-off the position-aware design of this thesis is explicitly intended to justify, and one that an automatic template-management strategy (discussed below) could further mitigate. A more demanding comparison would benchmark Spatial DTW against other posture-robust or discrete-gesture architectures explicitly designed for false-positive rejection, such as the wake-gesture DTW system of (20) or the LSTM-based architectures evaluated by (11); this was beyond the scope of the present work and is left for future comparison.

4.3 Algorithm and Design Choices

A direct mechanistic explanation for the algorithm’s general posture robustness comes from the template-proximity analysis (Fig. 3.40): in 46.7% of correctly detected events overall, the cued arm position’s own templates were not among the three closest matches used by the distance-aggregation rule (Section 2.2.5), and a comparable proportion (44.8%) of error events had the cued position absent from top-3. This indicates that the algorithm does not depend on strict same-position template matching to succeed; templates recorded at nearby or kinematically similar postures appear to substitute adequately for one another. This interpretation is consistent with the modest cost of the five-position ABFGI protocol (Fig. 3.42, a 2.5-percentage-point drop in CDR): if templates share information across postures, a sparser but well-distributed set of calibration positions can cover much of the same operating space as the full ten. The templates-reduction analysis showed that the number of templates per position can be reduced from four to three at

essentially no cost (a 0.1-percentage-point drop), while reducing further to two or one incurs increasing losses (1.4 and 4.8 percentage points respectively). Together, these analyses might suggest the current calibration protocol is more conservative than strictly necessary for the level of performance achieved. Regarding the influence of the spatial correction, removing this layer reduced the main-experiment CDR by 6.5 percentage points overall, and produced comparable losses in distractor rejection. Its contribution was not uniform: the benefit was consistently larger for the OPEN-state distractor tests than for the CLOSED-state finger-distractor test, and the ring finger remained the dominant source of error even with the correction active. This asymmetry is consistent with the correction’s underlying mechanism: it penalises templates whose spatial activation pattern disagrees with the live signal, and is therefore most effective against confounds that are spatially distinct from the target movement, as is largely the case for wrist rotations and deviations, but offers comparatively little protection against confounds that share substantial muscle synergy with the target gesture itself. While one concern regarding this layer was the possibility that it would overtook the Dynamic Time Warping mechanism underlaying, this study instead demonstrates the spatial correction’s overall contribution was nonetheless consistently positive, never degrading performance relative to plain cosine DTW, confirming it as an additive layer that complements rather than competes with the base DTW distance.

4.4 Clinical Deployment

The two SCI patients achieved Correct Detection Rates of 90.3% and 88.0% (Fig. 3.29), meaningfully lower than the healthy cohort’s 98.4%. Part of this gap is expected and not directly comparable to the healthy-subject results: the patient protocol used a single fixed posture rather than the full ten-position battery, and both sessions were shortened by fatigue. The remaining gap likely reflects genuinely less consistent voluntary EMG activity in patients with impaired hand function. It is worth noting, however, that of the total 6 errors, 5 were missed events, with the patients attempting the movement but no transition being detected; in terms of robustness, missed events represent a more benign failure mode than instabilities or false positives, since they do not risk an unintended actuation of the orthosis. These results should be read as a proof of concept that the Spatial DTW algorithm works on EMG from SCI patients, and as a pilot step toward an eventual clinical deployment of this algorithm.

The one-week pilot retest (Fig. 3.28) showed only a small overall change in performance, driven almost entirely by S6, whose second-trial latency distribution is visibly wider and shifted higher than the first, while S5’s distribution changed only marginally between sessions. The sample size (two subjects) is too small to draw general conclusions about long-term calibration stability but, again these

results are promising and suggest that a single calibration session can remain valid over short timescales without requiring a substantial re-tuning.

Two of the algorithm’s tunable parameters offer a direct and concrete route toward improving robustness for an eventual clinical use without any major change to the underlying method. First, since the per-class thresholds are independently calibrated (Section 2.2.5), a clinical deployment could deliberately favour missed events over false positives by lowering the thresholds further, especially for the closed state: once the subject grasps an object, raising the specificity over the sensitivity reduces the the possibility of dropping the object due to a false opening, at the cost of requiring the user to occasionally repeat the gesture of opening when it actually wants to. Second, although reducing the number of templates was shown not to produce drastic performance losses, this should not be read as a reason to minimise the template bank in a real deployment: since the distance-aggregation rule selects only the three smallest per-class distances, additional templates beyond this minimum do not dilute the decision, provided the two classes remain well separated. A larger and more diverse template bank, spanning more postures, grip variations, or movement speeds, could in principle only improve robustness at negligible cost to accuracy. Since it’s being proven less templates can still achieve >90% accuracies, an hypothetical final product could be initialized with a first calibration consisting in a limited number of templates (to reduce the calibration burden); some transitions will likely still be detected and, since the templates and the real-life gestures are the same, a system which recognizes the correctly detected activations could automatically add them to the set of templates, continuously adding more templates, improving the performance and keeping it high over time. The same automatic system could at the same time be responsible of deleting the most outdated or less used templates to keep the set balanced.

The algorithm latency was expected to be around one second, since the algorithm intrinsically relies on similarity between one-second-long windows. The average latency being substantially smaller than a second demonstrates the value of extracting templates that capture not only the movement itself but also its onset, the preparation phase anticipated by 200 ms during template extraction (Section 2.2.2). A plausible explanation for this sub-second detection lies in the warping flexibility of DTW itself: while a template is composed of roughly 20% onset and 80% sustained activation, the live buffer at the moment of detection is likely composed of the opposite proportion, mostly residual rest or onset signal, with only a small, recently accumulated fraction of genuine activation at its trailing edge. Because DTW does not require a strict sample-by-sample correspondence, the alignment path can compress the template’s activation segment against this small but growing fraction of real signal in the live buffer, allowing the cumulative distance to fall below threshold before a full second of movement has actually occurred in real time.

A further consideration for real-world use concerns the nature of the calibration

templates themselves. As noted in the discussion of the mug-holding experiment, the templates used in this experiment were recorded during unresisted, “grasping air” movements, whereas an orthosis actuating against a real object requires the user to sustain what is closer to an isometric contraction against genuine mechanical resistance. The EMG activity underlying a sustained isometric contraction can differ from that of a dynamic, unresisted movement, in amplitude, in the balance between agonist and antagonist activation, and potentially in spatial recruitment pattern, a factor any clinical calibration protocol would need to account for.

4.4.1 Limitations - Future Work

The patient cohort was small (two participants) and tested under a simplified, single-posture protocol. The temporal-stability analysis involved only two subjects over a single one-week interval. Several subject-specific observations in this chapter, particularly regarding S2 and S13, are based on qualitative impressions from the experimental sessions rather than quantified signal-level analysis, and would benefit from further investigation of the raw EMG data. The mug-holding experiment was evaluated qualitatively and is intended as a usability check rather than a quantitative validation. The current calibration templates, recorded without resistance, may not fully anticipate the isometric contraction conditions of real orthosis actuation, as discussed above.

The latency of the algorithm is higher than the delay generally considered optimal in the literature for continuous myoelectric control, approximately 100–125 ms, with controllability degrading noticeably beyond this range (23). However, half a second can be considered good enough for the 1-DoF discrete control of an orthosis that helps patients open and close their hand, a task that rewards a correct, stable decision over an instantaneous one. The higher latency is accepted as a trade-off to maximise the robustness of the decision, since the state machine requires sustained agreement across a full second of signal and five consecutive votes before committing to a change of state (Section 2.2). However, given how the whole framework was designed around this trade-off, extending the algorithm to multi-DoF continuous control would be considerably more difficult: a continuous controller cannot tolerate the same sub-second commitment delay without becoming visibly unresponsive, so achieving both robustness and the shorter latency required for continuous control would likely require a substantially different decision architecture.

As previously mentioned, the algorithm’s architecture itself is a pipeline of several sequential stages, temporal alignment, spatial correction, aggregation, and thresholding, each involving its own subjective design choices, such as the specific saturating function used in the spatial correction layer (Section 2.2.4). The result is therefore a sum of different choices, in terms of values of different parameters of each layer, manually (and where possible semi-automatically) tuned with the goal of achieving the best performance possible. The offline results demonstrates each

layer is not dominantly influencing the performance, instead contributes to improve the accuracy and robustness of the framework. On the other hand, each layer is dependent from choices (e.g. value of the thresholds, template extraction methods, length, spatial correction function parameters) which in this exact experiment were fixed to a value but are yet to be optimized: each combination of different values could affect the resulting performance, and there might be combination of tunings which can obtain even better results. Considering that an important requirement for a limb prosthesis or orthosis is its simplicity in usage and setting, a more elegant alternative would optimize this value so that the parameter tuning becomes objective rather than subjective. For example, defining a single, jointly-optimised functional capable of aligning all channels simultaneously while integrating their spatial relationships directly into the alignment cost, rather than treating temporal and spatial similarity as separate, sequentially-combined stages, would already reduce the amount of parameters derived from the cosine-DTW and spatial correction layers alone.

Several other directions can be the natural follow-up to this work. Embedding the algorithm directly on the orthosis hardware, rather than running it on a laboratory laptop, is a necessary step toward untethered daily use. As discussed above, an automatic template manager, periodically discarding the most outdated or outlier templates and replacing them with recent correct detections, could reduce the need for full manual recalibration over time. Similarly, an automatic thresholding mechanism, adjusting the per-class thresholds in response to changing conditions such as skin-electrode impedance, could improve robustness without requiring manual retuning. A future calibration protocol may also need to include representative isometric holds against real or simulated resistance, rather than relying solely on unresisted movement templates, to fully anticipate the EMG conditions of real orthosis use. A larger and more diverse patient cohort, tested across multiple postures, would be needed to properly assess the postural-robustness claim in the target clinical population. Finally, extending the framework beyond a single degree of freedom, for instance to include a pinch gesture, would move the system closer to the multi-gesture control needed for full clinical utility while preserving the robustness from the discrete gesture detection approach.

Chapter 5

Conclusion

This thesis addressed two of the most persistent barriers to the clinical deployment of myoelectric hand orthoses, the limb-position effect and robustness to false activations, by proposing Spatial DTW, a posture-robust, state-machine-based 1-DoF myoelectric control algorithm built on a spatially-corrected adaptation of Dynamic Time Warping. The algorithm was implemented within the MyoGestic framework and validated across 16 healthy subjects, ten activities-of-daily-living arm positions, and two spinal cord injury patients.

Taken together, these results support Spatial DTW as a viable, lightweight alternative to more data-hungry or sensor-heavy approaches to myoelectric control: posture-robust, resistant to most non-target activity, requiring only surface EMG, and workable on clinical signals from spinal cord injury patients. Relative to the DTW-based approaches it builds upon, Spatial DTW extends single-posture, single-gesture designs with a multi-position, multi-template calibration bank and a spatial correction layer explicitly targeting confounds that share musculature with the target gesture, and achieves comparable or superior robustness on a substantially more demanding, bidirectional control task. Its main limitations, restriction to a single degree of freedom, dependence on fixed calibration templates and thresholds, and calibration on unresisted movements that may not fully anticipate the isometric contraction required to actuate a loaded orthosis, define a clear set of next steps: embedding the algorithm on the orthosis hardware for untethered use, an automatic template-management mechanism to ease long-term recalibration, and an adaptive thresholding scheme responsive to changing acquisition conditions. Addressing these would move Spatial DTW from a validated proof of concept toward a deployable component of a clinically usable neuroprosthetic hand orthosis.

Acknowledgements

I would like to express my gratitude to all the member of n-squared lab at Friedrich-Alexander-Universität Erlangen-Nürnberg for welcoming me and giving me the opportunity to carry out this thesis project within their group, which resulted in a truly wonderful experience, both professionally and personally.



Supplementary Material

Video recordings of the experimental sessions described in this thesis are available upon request.

Bibliography

- [1] R. Merletti and D. Farina, Eds., *Surface Electromyography: Physiology, Engineering, and Applications*. Hoboken, NJ, USA: John Wiley & Sons / IEEE Press, 2016. DOI: 10.1002/9781119082934.
- [2] A. L. Hodgkin and A. F. Huxley, “A quantitative description of membrane current and its application to conduction and excitation in nerve,” *The Journal of Physiology*, vol. 117, no. 4, pp. 500–544, 1952. DOI: 10.1113/jphysiol.1952.004764.
- [3] A. Compston and A. Coles, “Multiple sclerosis,” *The Lancet*, vol. 372, no. 9648, pp. 1502–1517, 2008. DOI: 10.1016/S0140-6736(08)61620-7.
- [4] C. E. Shannon, “Communication in the presence of noise,” *Proceedings of the IRE*, vol. 37, no. 1, pp. 10–21, 1949. DOI: 10.1109/JRPROC.1949.232969.
- [5] R. Merletti and G. Cerone, “Tutorial. Surface EMG detection, conditioning and pre-processing: Best practices,” *Journal of Electromyography and Kinesiology*, vol. 54, 102440, 2020. DOI: 10.1016/j.jelekin.2020.102440.
- [6] S. C. Kirshblum, S. P. Burns, F. Biering-Sørensen, W. Donovan, D. E. Graves, A. Jha, M. Johansen, L. Jones, A. Krassioukov, M. J. Mulcahey, M. Schmidt-Read, and W. Waring, “International standards for neurological classification of spinal cord injury (revised 2011),” *The Journal of Spinal Cord Medicine*, vol. 34, no. 6, pp. 535–546, 2011. DOI: 10.1179/204577211X13207446293695.
- [7] Daniela Souza Oliveira, Matthias Ponfick, Dominik I Braun, Marius Osswald, Marek Sierotowicz, Satyaki Chatterjee, Douglas Weber, Bjoern Eskofier, Claudio Castellini, Dario Farina, Thomas Mehari Kinfe, Alessandro Del Vecchio, *A direct spinal cord-computer interface enables the control of the paralysed hand in spinal cord injury.*, Brain, Volume 147, Issue 10, October 2024, Pages 3583–3595, <https://doi.org/10.1093/brain/awae088>
- [8] Z. Bayram-Weston, M. Andrade-Sienz, and J. Knight, “Nervous system 4: the peripheral nervous system – spinal nerves,” *Nursing Times*, vol. 118, no. 6, pp. 41–44, 2022.

- [9] P. Parker, K. Englehart, and B. Hudgins, “Myoelectric signal processing for control of powered limb prostheses,” *Journal of Electromyography and Kinesiology*, vol. 16, no. 6, pp. 541–548, 2006. DOI: 10.1016/j.jelekin.2006.08.006.
- [10] E. Scheme and K. Englehart, “Electromyogram pattern recognition for control of powered upper-limb prostheses: State of the art and challenges for clinical use,” *Journal of Rehabilitation Research and Development*, vol. 48, no. 6, pp. 643–659, 2011. DOI: 10.1682/JRRD.2010.09.0177.
- [11] E. Eddy, E. Campbell, S. Bateman, and E. Scheme, “Understanding the influence of confounding factors in myoelectric control for discrete gesture recognition,” *Journal of Neural Engineering*, vol. 21, 036015, 2024. DOI: 10.1088/1741-2552/ad4915.
- [12] E. Campbell, A. Phinyomark, and E. Scheme, “Current trends and confounding factors in myoelectric control: Limb position and contraction intensity,” *Sensors*, vol. 20, no. 6, 1613, 2020. DOI: 10.3390/s20061613.
- [13] P. Geethanjali, “Myoelectric control of prosthetic hands: State-of-the-art review,” *Medical Devices: Evidence and Research*, vol. 9, pp. 247–255, 2016. DOI: 10.2147/MDER.S91102.
- [14] K. Englehart and B. Hudgins, “A robust, real-time control scheme for multi-function myoelectric control,” *IEEE Transactions on Biomedical Engineering*, vol. 50, no. 7, pp. 848–854, 2003. DOI: 10.1109/TBME.2003.813539.
- [15] A. Fougner, E. Scheme, A. D. C. Chan, K. Englehart, and Ø. Stavdahl, “Resolving the limb position effect in myoelectric pattern recognition,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 19, no. 6, pp. 644–651, 2011. DOI: 10.1109/TNSRE.2011.2163529.
- [16] E. Eddy, E. J. Scheme, and S. Bateman, “A framework and call to action for the future development of EMG-based input in HCI,” in *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, Hamburg, Germany, 2023, pp. 1–23. DOI: 10.1145/3544548.3580962.
- [17] H. Sakoe and S. Chiba, “Dynamic programming algorithm optimization for spoken word recognition,” *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 26, no. 1, pp. 43–49, 1978. DOI: 10.1109/TASSP.1978.1163055.
- [18] Raul C. Sîmpetru et al. (2025) *MyoGestic: EMG interfacing framework for decoding multiple spared motor dimensions in individuals with neural lesions*. Sci.Adv.11,eads9150.DOI:10.1126/sciadv.ads9150
- [19] *CatBoost: unbiased boosting with categorical features*
<https://doi.org/10.48550/arXiv.1706.09516> <https://catboost.ai/docs/en>

- [20] Ethan Eddy et al 2025 J. Neural Eng. 22 016006. *EMG-based wake gestures eliminate false activations during out-of-set activities of daily living: an online myoelectric control study*. <https://doi.org/10.1088/1741-2552/ada4df>
- [21] M. Müller, *Fundamentals of Music Processing: Audio, Analysis, Algorithms, Applications*. Cham, Switzerland: Springer, 2015. DOI: 10.1007/978-3-319-21945-5.
- [22] C. Häger-Ross and M. H. Schieber, “Quantifying the independence of human finger movements: Comparisons of digits, hands, and movement frequencies,” *The Journal of Neuroscience*, vol. 20, no. 22, pp. 8542–8550, 2000. DOI: 10.1523/JNEUROSCI.20-22-08542.2000.
- [23] T. R. Farrell and R. F. Weir, “The optimal controller delay for myoelectric prostheses,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 15, no. 1, pp. 111–118, 2007. DOI: 10.1109/TNSRE.2007.891391.