Author:
Lombardi, Maria

Title:
Modelling, control and design of autonomous artificial avatars in human motor coordination task

Modelling, control and design of autonomous artificial avatars in human motor coordination tasks

By

MARIA LOMBARDI

Department of Engineering Mathematics
UNIVERSITY OF BRISTOL

A dissertation submitted to the University of Bristol in accordance with the requirements of the degree of DOCTOR OF PHILOSOPHY in the Faculty of Engineering.

APRIL 9TH, 2020

Word count: 33,160 words

Abstract

Human coordination is a phenomenon that takes place in numerous daily activities, such as simple oral communication, walking in a crowd, clapping within an audience and when performing more complex coordinated activities such as playing in team sports or in musical ensembles. Unveiling the mechanisms that lead people to coordinate, adjust their movements properly, reach and maintain a stable coordinated behaviour represents a key challenge, both from a psychological and from a control point of view. Addressing this challenge is crucial, for example, to control artificial cyber-agents able to interact with people to perform common joint tasks. This thesis is concerned with the problem of designing an autonomous artificial agent able to move in a natural way in coordination with one or more humans. This is particularly relevant in the context of healthcare applications. Indeed, the use has been proposed of artificial agents coordinating their movements with those of patients suffering from social or motor disorders. Specifically, it has been shown that an artificial agent moving its end-effector with certain human kinematic properties could provide innovative and efficient rehabilitation strategies. In this thesis, human behaviour is studied through a simple yet effective coordination paradigm, where participants are asked to synchronise their hand motion. Keeping the same motor task, artificial agents with different control strategies are designed to interact with human participants so as to produce coordinated motion in different configurations. Different control approaches including those based on reinforcement learning are explored and validated via numerical simulations and experiments confirming the effectiveness of the proposed control architectures. The results of some additional work on the implementation of an exergame for motor rehabilitation of patient after stroke is also reported together with the analysis of leadership emergence in walking groups.

AUTHOR'S DECLARATION

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's Regulations and Code of Practice for Research Degree Programmes and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in the dissertation are those of the author.

SIGNED: ..... DATE: .....

TABLE OF CONTENTS

List of Tables

Table **Page**

List of Figures

Figure **Page**

List of Abbreviations

Abbreviation	Full name
AI	Artificial Intelligence
ADS	Autism Spectrum Disorders
ANN	Artificial Neural Network
AWING	Avatar folloWING
CDF	Cumulative Density Function
CG	Complete Graph
CP	Cyber Player
CV	Circular Variance
DOF	Degree Of Freedom
DPI	Data Processing Inequality
DQN	Deep Q-Network
EMD	Earth Mover's Distance
FFT	Fast Fourier Transform
GMM	Gaussian Mixture Model
GMR	Gaussian Mixture Regression
GWS	Global Wavelet Spectrum
HKB	Haken Kelso Bunz
HMI	Human Machine Interaction
HP	Human Player
HRI	Human Robot Interaction
IMS	Individual Motor Signature
IPD	Inter-Personal Distance
JI	Joint Improvisation

Abbreviation	Full name
LF	Leader Follower
MC	Markov Chain
MDP	Markov Decision Process
MDS	Multi-Dimensional Scaling
PbD	Programming by Demonstration
PDF	Probability Density Function
PG	Path Graph
PAWING	PAtient followING
RG	Ring Graph
RL	Reinforcement Learning
RMSE	Root Mean Square Error
ROM	Range Of Motion
RP	Relative phase
RPE	Relative Position Error
SC	Solo Condition
SEM	Standard Error of the Mean
SG	Star Graph
STFT	Short Time Fourier Transform
TDDC	Time-dependent Delay Directional Correlation
TDSC	Time-dependent Delay Speed Correlation
TL	Time Lag
TP	Target Player
VF	Virtual Follower
VL	Virtual Leader
VP	Virtual Player
VT	Virtual Trainer
VQ	Vector Quantisation

CHAPTER 1

INTRODUCTION

Investigating and analysing the dynamics of groups of systems interacting with each other is a relatively new area of study and is the subject of much ongoing research [133, 183]. Such interacting systems can be studied through the formalism of complex networks, where each system can be modelled as a single node with its own dynamics while the interaction among systems can be easily represented as edges in a graph. The importance of understanding how the behaviour of a single node changes when interconnected with other agents can be found in the numerous contexts that can be easily modelled according to the paradigm of complex networks. For instance, distribution and transportation infrastructures (roadways, electric power grids, airlines), social networks (Internet, telephone networks), biological systems (gene regulator networks, protein interaction networks) and many others, see e.g. [17].

Several interesting phenomena can emerge from the collective behaviour of an interacting group of agents that cannot be explained in terms of the dynamics of the single agent [4, 33]. Among these phenomena, the emergence of synchronisation and movement coordination within a group of people engaged in joint tasks represents a challenging example to study. Human coordination is simply ubiquitous and is involved in many daily activities. Indeed, people need to interact with each other and coordinate their motion to perform common joint tasks, like simple oral communication, walking together, handling objects and to perform more specific coordinated activities such as for example dance, sport or theatre [59, 75, 131, 169]. It has been shown that the motion of each individual is characterised by some kinematic motor features, or biomarkers, that make his/her motion unique and distinguishable from any other [179].

Unveiling such human kinematic features and the mechanisms that lead people to coordinate, adjust their movements to those of the others, reach and maintain a stable coordinated behaviour represents a key challenge, both from a psychological and from a control point of view.

Addressing this challenge is essential, for example, to control and program artificial agents (e.g., robots or avatars) able to interact with humans to perform some joint motor task while exhibiting desired "human-like" kinematic features. Investigating how to design an autonomous artificial agent able to move in a natural and safe way in coordination with one or more humans is a fundamental and yet open problem in social robotics to improve the quality of human-robot interaction. Major open questions in the design of artificial agents (or avatars) in joint motor tasks with humans are:

i) how to endow an artificial entity with the kinematic features characterising human behaviour;

ii) how to develop control architectures enabling artificial agents to adjust their motion and interact with a human partner in joint motor tasks;

iii) whether it is possible to build a mathematical model able to capture, describe and unfold the adaptive mechanism of a human while engaged in a common task;

iv) how to extend the study of human motor coordination from dyadic to group scenario (e.g., how a human changes his/her behaviour when engaged in a group, who coordinates his/her motion with whom, etc.)

Addressing all these points is a cumbersome challenge due to the complexity of human behaviour, the unknown mechanisms underlying human coordination and the lack of formal models in the literature that can be implemented in an artificial agent. Again, human coordination has rarely been explored in the existing literature in the case of human ensembles (more results can be found for the dyadic interaction where only two individuals are involved).

The work presented in the rest of this thesis is an effort spent trying to tackle the aforementioned challenging issues. In so doing, human behaviour is studied through a simple yet effective coordination paradigm, known as the *mirror game* [136], where participants are asked to synchronise an oscillatory hand motion while being visually coupled. Keeping the same motor task, artificial agents with different control strategies are designed to interact with human participants so as to produce coordinated motion in a number of different configurations [178, 179, 206, 207]. Crucially, we want to design a virtual agent able to successfully interact with humans while exhibiting kinematic motor features resembling human ones.

An interesting application is also explored in the healthcare context. Indeed, it has been suggested that the availability of artificial agents able to interact with one or more individuals can allow further development of both diagnostic and therapeutic strategies for patients with social and motor disorders [21]. In this direction, an effort has been spent to develop new home therapies based on serious games, which are less expensive, more efficient and able to guarantee better working conditions for clinicians [70, 97, 205]. Examples of such *exergames* (exercise + game) have been developed for postural, neglect and stroke rehabilitation [13, 30, 148]. Whereas the

current exergames consist in a series of exercises that the patient has to perform, the innovative contribution of this thesis is to consider the use of an adaptive virtual agent that assists the patient coordinating its motion with her/him during the therapy and can change the features of its behaviour in successive sessions of the game for rehabilitation purposes.

1.1 Thesis outline

We start this thesis by providing an extensive overview of the literature on emerging phenomena in human joint interaction in Chapter 2. Firstly, we discuss synchronisation and human motor coordination reviewing some of the works found in the current literature both in dyadic interpersonal coordination and in human ensembles. Secondly, we present the phenomenon of leadership that can emerge from a group of humans interacting with each other, underlying how most results in the literature deal only with the case of animal groups. Finally, we introduce the need for an autonomous virtual agent able to interact with humans in coordination motor tasks.

In Chapter 3 we give a broad overview of existing results on the design of artificial agents, tasks and solution strategies addressed in the literature. In particular we focus on joint motor coordination tasks between humans and artificial agents reporting some sample works existing in the literature, analysing advantages and disadvantages and proposing our approach to overcome them.

We then begin our investigation on the design of a cognitive architecture able to drive a virtual agent (or avatar) in motor coordination tasks with humans in Chapter 4. Specifically, we start from the control architecture proposed in [12] and based on a controlled nonlinear oscillator. Even if such control solution has been proved to solve the coordination problem in the mirror game task, we point out that its main drawback is the lack of "human kinematic features" in the avatar's motion. To overcome such issue, Markov chains are introduced as mathematical models able to capture and artificially reproduce the kinematic features of human motions. Such a model is then embedded in the described control scheme as a stochastic reference generator. The approach is validated experimentally by comparing the motion signals generated by the Markov models with those of humans and by making the virtual agent, having the Markov chain in the control loop, perform dyadic interactions with human partners.

In Chapter 5 we discuss advantages and disadvantages of the proposed control architecture. We find that the control architecture in Chapter 4 is able to solve the tracking problem in dyads but requires many trial-and-error attempts to tune the controlled system dynamics in order to make the virtual avatar play like a human. Furthermore, the architecture presented in Chapter 4 cannot be easily extended to a group scenario. To overcome this problem, we design a different control architecture in Chapter 5, based on learning techniques and in particular on reinforcement learning in order to make the avatar learn how to play directly from humans.

While in Chapter 5 we test the ability of the AI based architecture in dyadic scenarios, in

Chapter 6 we present its extension and validation in group scenarios. Also in the latter chapter we report both in-silico and experimental validation explaining carefully the experimental setup and the experimental platform used to carry out the experiments.

In Chapter 7 we present an *exergame* designed for the rehabilitation of people suffering from motor impairments, e.g., after a stroke. The innovation of such an exergame relies on the presence of an adaptive virtual avatar that assists the patient performing and coordinating its motion with him/her in order to complete the exercise together.

Chapter 8 reports further work done for a side project regarding the study of the leadership emerging in human ensembles. Specifically, we use a small group of walking people, as a scenario of interest, to detect leader-follower patterns spontaneously established in the group and unveil whether a leader emerges in the group when no roles are assigned.

Finally, a summary of our results and discussion of their implications, followed by directions for future work, are presented in Chapter 9.

1.2 Publications

Part of the work presented in this thesis has been published in the following papers:

- **Lombardi M.**, Warren W.H., & di Bernardo M. (2020). *Leadership emergence in walking groups*. Submitted to Scientific Reports (Chapter 8);
- **Lombardi M.**, Liuzza D., & di Bernardo M. (2019). *Using learning to control artificial avatars in human motor coordination tasks*. Submitted to IEEE Transactions on Robotics (Chapter 5);
- **Lombardi M.**, Liuzza D., & di Bernardo M. (2019). *Deep learning control of artificial avatars in group coordination tasks*. Proceedings of the 2019 IEEE International Conference on Systems, Man and Cybernetics (SMC), pp. 724 - 729, Bari, Italy (Chapter 6);
- **Lombardi M.**, Liuzza D., & di Bernardo M. (2018). *Generation and classification of individual behaviours for virtual players control in motor coordination task*. Proceedings of the 2018 European Control Conference (ECC), pp. 2374-2379, Limassol, Cyprus, doi: 10.23919/ECC.2018.8550321 (Chapter 4);
- Alderisio F., **Lombardi M.**, & di Bernardo M. (2018). *Emergence of leadership in complex networks and human groups*. Proceedings of the 2018 IEEE International Symposium on Circuits and Systems (ISCAS), Florence, Italy, doi: 10.1109/ISCAS.2018.8351745;
- Alderisio F., **Lombardi M.**, Fiore G., & di Bernardo M. (2017). *A novel computer-based set-up to study movement coordination in human ensembles*. Frontiers in Psychology 8(967), doi: 10.3389/fpsyg.2017.00967 (Chapter 4 and Chapter 5).

CHAPTER 2

BACKGROUND ON HUMAN MOTOR COORDINATION

In this chapter we give an overview of the current state of the art on human motor coordination in joint tasks. Specifically, in Section 2.1 we review the emergence of motor coordination as a nonlinear phenomenon that arises when two or more people have to coordinate their motion in order to perform a shared task. In this context, we introduce the mirror game which is largely used as paradigmatic task to study interpersonal coordination. In Section 2.2 we discuss the emergence of leadership in human groups, placing emphasis on the fact that the mechanisms underlying such a phenomenon are still unclear and under investigation. Then in Section 2.3, we introduce the importance of designing an autonomous agent able to interact with humans in joint motor tasks. In particular, we review different real applications in which reaching a good level of synchronisation between humans and artificial agents has a great relevance. Finally, in Section 2.4 we summarise the content of the chapter.

2.1 Synchronisation in joint dynamic tasks

People interacting with each other tend to synchronise and coordinate their movements. Indeed, people often need to coordinate during their daily activities in order to perform a joint task, that spans from a simple walk together, or handling objects to performing more complex choreography in dance or theatre [59, 75, 131, 169] (Figure 2.1). Over the past few decades, understanding the complex mechanisms underlying the emergence of interpersonal coordination has been the goal of many researchers from different fields, spanning from human psychology and movement science to mathematics and engineering [120, 158, 159, 170, 194].

The emergence of interpersonal coordination is a complex phenomenon characterised by highly non-linear dynamics and complicated cognitive and motor processes [158, 159, 174]. The emergence of synchronisation is possible thanks to a feedback mechanism (visual or auditory)

(a) Soldiers during a march

(b) Synchronised swimming

(c) Music ensemble

(d) Rowers in team race

FIGURE 2.1. People coordinating their movements in different group activities: **(a)** marching soldiers (*Image credit*: `https://www.theguardian.com`), **(b)** synchronised swimming (*Image credit*: `http://www.baystategames.org`), **(c)** playing in an orchestra (*Image credit*: `https://www.justdial.com`), **(d)** rowing in team race (*Image credit*: `https://www.pressherald.com`).

between the participants, according to which they adjust their movements based on those of the others in order to complete some common task.

Synchronisation may be intentional or unintentional depending on whether the individuals involved in the interaction explicitly want to coordinate their movements or not. Unintentional synchronisation was explored in [159], where experiments were conducted in which pairs of participants were instructed to complete an interpersonal puzzle task while swinging handheld pendulums. Furthermore, the effects of visual and verbal interaction on interpersonal movement coordination were analysed. Higher degree of coordination was observed when the pairs were visually coupled, whereas verbal interaction alone was not found to be sufficient for unintentional coordination.

Despite the synchronisation, studies reveal that each individual moves differently from each other, for example faster/slower, harder/smoother, etc. [171]. In the context of the European

project AlterEgo [3], it has been proved that these differences are described by an *Individual Motor Signature* (IMS), an observable able to capture the subtle differences (or similarities) in the way humans move [179]. Specifically, the probability density function (PDF) of velocity profiles exhibited by an individual, while performing the task of the mirror game, has been proposed as the IMS of that person. Such a task has been often presented as a paradigmatic scenario to study the emergence of social motor coordination between two individuals imitating each other's movements (more details are in Section 2.1.1). Moreover, this signature does not change over time and is, in principle, unique for every individual, thus making it possible to distinguish one from another.

Sharing similar kinematic features has a main effect on movement coordination between people. In fact, in social psychology it was suggested that people prefer to team up with others possessing similar morphological and behavioural features, spontaneously producing higher level of coordination despite not being explicitly asked to [105]. Furthermore, it was discussed that acting in synchrony with others can produce positive emotions increasing social attachments among group members and leading the entire group to remain cohesive [203]. These observations have led to the development of a founded *theory of similarity* which predicts that the level of synchronisation in joint actions is enhanced if the participants are similar in terms of morphology and movement dynamics and are willing to match their behaviours [21, 209].

While the emergence of synchronisation between two individuals performing a joint task has been extensively studied and a lot of examples are present in the literature [59, 100, 169, 171], social interaction and motion coordination in a group of several people still need to be accurately investigated. Such an investigation can be cumbersome and complex due to the lack of an appropriate mathematical framework and to the variety of scenarios that can be considered, the countless types of activities people might be involved with (limb or finger movements, walking, head movements, or more in general music or sport activities), the many different ways in which participants can interact and communicate with each other and the different ways they can be physically located with respect to each other while performing the specified task (e.g., while being in the same physical location or not).

Among the limited examples in the existing literature that tackle the problem of movement coordination in a group of several agents, we find studies on a group of people sitting on rocking chairs [156], on a cross-cultural choir workshop [88], on rowers in a race [45, 204], on walking pedestrian groups [127], on group behaviour in dance [63, 140]. In [39] the emergence of unintentional synchronisation in a group setting was investigated and it was studied if such a synchronisation also extended to heart and respiratory rhythms. Specifically, the group was asked to move their arms in a predetermined way without receiving any explicit instructions about synchronisation. It was observed that spontaneous motor synchronisation in that group scenario involved also the cardiovascular and respiratory rhythms, which were enhanced when an external auditory rhythm (metronome and music) was provided to all group members. Others

results on multi-player synchronisation investigated how the pattern of visual interactions among people in the group affects their coordination levels. In [10] it was shown computationally and experimentally that synchronisation depends on the structure of visual interconnections (who looked at whom) between group members that can be enhanced if the group shares similar individual motion characteristics.

In addition to the results mentioned above, many others are available in the literature on human motor coordination. In this thesis, we exploit some of these results to guide an artificial agent in the interaction with human partners, both in dyads and in a group scenario, focusing our attention on which kinematic features to provide it with in order to have it move with features typical of the human behaviour. To investigate human coordination and deploy a proper artificial agent, we take the mirror game as paradigmatic joint task, largely used in the literature and described in detail below.

2.1.1 Mirror game as paradigmatic task to study human coordination

Firstly proposed in [136], the mirror game has been largely used over the last few years to study the complex phenomenon of interpersonal human coordination. Usually played by actors, musician and dancers, in the mirror game several players have to imitate each other enjoying the game and creating synchronised body motions. In its simplest formulation the mirror game involves only two people mirroring each other's hand movements by moving their own handle back and forth horizontally along parallel strings (one-dimensional motion) as depicted in Figure 2.2. It can be played in three different experimental conditions:

- *Leader-Follower* (LF): in which the player, designated as leader, has to drive the game creating the motion, while the other player, designated as follower, tries to follow the leader's trajectory.

- *Joint Improvisation* (JI): in this condition there is no designation of leader and follower, but both the players can move as they want trying to synchronise their hands in an unique movement.

- *Solo Condition* (SC): in this condition the player is asked to generate a trajectory in isolation, so as the kinematic features of his/her own dynamics could spontaneously emerge.

Despite its simplicity, the mirror game was found to be a powerful tool in many research studies. For example, in [32] the mirror game was used to study the synchronisation ability of people suffering from autism spectrum disorders (ASD). Specifically, they aimed at assessing whether individuals with autism can attain a highly synchronised motion when playing with an expert improviser. Moreover, [178] used the mirror game to characterise the different motion patterns of patients suffering from schizophrenia building also a diagnostic classifier based on such motion patterns. Finally, [137] employed the mirror game to investigate whether periods of

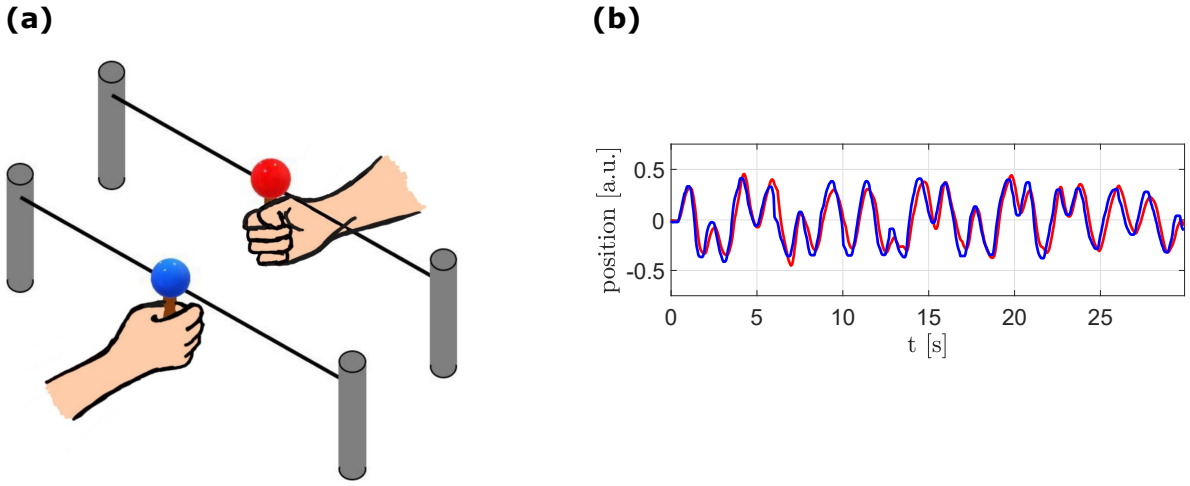


FIGURE 2.2. Illustration of the task of the mirror game. **(a)** Two players' hands are illustrated while moving a vertical handle along a constrained string. **(b)** Example of two recorded position time series of a trial of the mirror game in one dimension.

high synchronisation in joint improvisation between two expert improvisers had a distinctive influence in physiological activities (e.g., increase of players' heart rate).

2.2 Leadership emergence in groups

Another interesting phenomenon of great importance can emerge from the collective behaviour of a group of agents interacting with each other, that is leadership [4, 33]. Leadership plays a crucial role in determining the success or failure of several activities both in animals (e.g., protecting the animal group against predatory attacks, locating food resources) and human groups (e.g., steering opinion dynamics, playing music or sport).

Many studies exist in the literature addressing coordination and leadership emergence in groups of animals [42, 130, 155, 210], where the problem is tackled by studying how agents, informed on a migration route or food position, influence the collective behaviour. Limited results are instead available on leadership emergence in human ensembles [47, 61, 162]. Examples include consensus decision making [28, 60, 61], the role of a conductor in an orchestra [47, 77], group behaviour in dance [111, 140], among many others.

Understanding leadership emergence in groups of humans can be relevant in many applications. For example, during the evacuation from a building, the presence of a leader with global knowledge of the structure can safely guide people towards the exit avoiding fatal consequences [143]. In sports, athletes regard leaders who provide moral support and positive feedback as a crucial component towards the effective functioning of the team [65]. Past studies have observed that, in some situations, the leader is spontaneously and unconsciously elected by the

(a) Flock of birds

(b) Marching band

FIGURE 2.3. Examples of leadership emergence both **(a)** in animal (*Image credit*: `https://www.civilsocietyhowto.org`) and **(b)** in human groups (*Image credit*: `http://lifevesting.com`).

other members of the group according to his/her attitude or qualities, e.g., charisma, confidence, friendliness, effective speaking [182]. Also, it has been observed that such a spontaneous leader is characterised by specific behavioural or physiological traits such as a common tendency to act first [101]. A key element for the emergence of leadership is the presence of coordination. Indeed, leader-follower patterns take place in coordinated group movements where individuals characterised by specific behavioural traits, behave as leaders by increasing their tendency to act first.

In this thesis preliminary results are provided on leadership emergence in human ensembles. Our study is part of a broad project in collaboration with Prof. William Warren (Brown University in Providence, Rhode Island) aiming at investigating the emergence of leadership in groups of walking crowds. In this thesis, the goal is to assess whether leadership emerges and whether it facilitates group coordination and cohesiveness. Our aim is to unveil whether, when no designated roles are assigned, a leader spontaneously emerges in the group, and how such emergence depends on the structure of the interconnections among the group members.

2.3 The need for an autonomous virtual agent

Controlling the emergence of human coordination and leadership can be exploited as a strategy to help people to improve their social skills and promote social attachment and group cohesion [203]. It has been suggested that a virtual partner able to play the mirror game with human subjects can be an effective tool to both study interpersonal coordination between two or more individuals, and also to provide new clinical interventions for the rehabilitation of patients suffering from social disorders [21, 178]. People suffering from social disabilities that accompany schizophrenia, autism, or social phobia have trouble interacting with others even in the most common situations.

(a) Social rehabilitation

(b) Motor rehabilitation

FIGURE 2.4. Examples of robot-assisted rehabilitation both for **(a)** social disease (*Image credit*: `https://tecreview.tex.mx`) and **(b)** motor impairment (*Image credit*: `https://wwww.designboom.com`).

Studies have shown that one of the causes of this impairment is that their individual motor signatures are different from that of other people [195].

Taking advantage from the observation in [179] that individuals with similar signatures reach higher level of synchronisation, one rehabilitation strategy of this kind of patient consists in making them interact with another participant (virtual) whose signature starts from being similar to that of the patients themselves and then, gradually, ends up being similar to that of healthy humans. In this way, patients are unconsciously guided towards the direction of the desired movement features [21]. The virtual player can be properly parameterised as required so as to explore the influence of the interaction on the patient, which is not possible in ordinary human-human interaction [59].

The design of the virtual agent may enhance the development of home therapies [53, 93], replacing traditional rehabilitation strategies that can be too expensive as they consist in daily sessions of exercise carried out from the patient with the real time supervision of the therapist. So-called *exergames* aim at embedding therapeutic repetitive tasks in enjoyable and interactive games. Many exergames have been developed for postural, neglect and schizophrenia rehabilitation [13, 30, 148], consisting in a series of static exercises (pop the balloon, catch the fruit and so on) but none of them is a really “interactive” game. Hence, an important application of the research described in this thesis is to develop a *smart* exergame with an adaptive virtual agent that drives and interact in real-time with the patient during each rehabilitation session. The virtual agent will have the role of supervising, advising and correcting the patient while exercising, mimicking what a real therapist would do. Wrong movements make the exercise useless or even dangerous. Analysing the continuous stream of motion data, the avatar will be able to diagnose the level of disease, adapt the therapy to the patient, evaluate the progresses done, identify dangerous movements in real time and provide immediate feedback to the patient

to correct his/her posture. Designing and implementing such an exergame is a cumbersome and ambitious project. In this thesis, and in particular in Chapter 7, we present the design of an avatar able to lead the patient to complete post-stroke motor tasks and adapt its behaviour to the progresses/regresses of the patient during the therapy.

2.4 Summary

In this chapter we gave an overview of the current literature on human coordination both in dyads and in group scenarios. In particular, we discussed the human coordination as main phenomenon (spontaneous and not) that occurs in many daily activities. We illustrated the phenomenon of human coordination through several examples studied in literature (rocking chairs, walking pedestrians, swinging pendulums). We emphasised the concept of the Individual Motor Signature that characterises uniquely the human motion and can be measured easily during session of the mirror game.

We introduced also the leadership, another interesting phenomenon emerging from human interaction. Unveiling the mechanisms under leadership emergence is still an open problem, due to its complexity and many factors that influence the role of the leader (task, individual dynamics, interaction patterns, coupling mechanisms).

Finally, we discussed the importance of having a virtual partner able to interact with humans. Exploiting the synchronisation phenomenon, such virtual agent can be used as tool to propose new rehabilitation strategies introducing an innovation in the way of seeing the social robotics. Further details on human-robot interaction, what currently exists in literature and how these artificial agents can interact safely with humans are presented in the next chapter.

CHAPTER 3

ARTIFICIAL AGENTS FOR JOINT MOTOR COORDINATION TASKS: ANALYSIS, DESIGN AND IMPLEMENTATION

In Chapter 2 we gave an overview on human coordination explaining how it emerges in dyadic and group interaction, reporting the unanswered questions in the understanding of such phenomenon and underlying the importance of having artificial agents able to interact with humans. This chapter aims at analysing the design of artificial agents giving a broader review on Human-Machine Interaction (HMI). In the rest of the thesis, we will refer to the artificial agent also as "avatar", "robot" or "virtual agent". Specifically in Section 3.1 we introduce what is intended for "Human-Robot Interaction" (HRI) and explain why human-robot interaction is an emerging field of great importance both for the industrial and the academic community. In addition we try to summarise the different perspectives, the key themes and the open challenges regarding HRI reporting a brief but exhaustive taxonomy of the field. Emerging application domains of HRI are then listed in Section 3.2 to underline the importance artificial agents can have in human activities. After such overview, in Section 3.3 we focus on "socially interactive robots" in joint motor coordination tasks with humans discussing critically few key examples existing in literature. A summary is then given in Section 3.4.

3.1 Human-robot interaction as emerging field

Human-robot interaction is the field of study that deals with the problem of understanding and designing the interaction between one or more humans and one or more artificial agents. HRI is a cumbersome problem due to its interdisciplinary nature requiring contributions from engineering, psychology, cognitive science, computer science, linguistics, and other disciplines [58, 81, 86, 114].

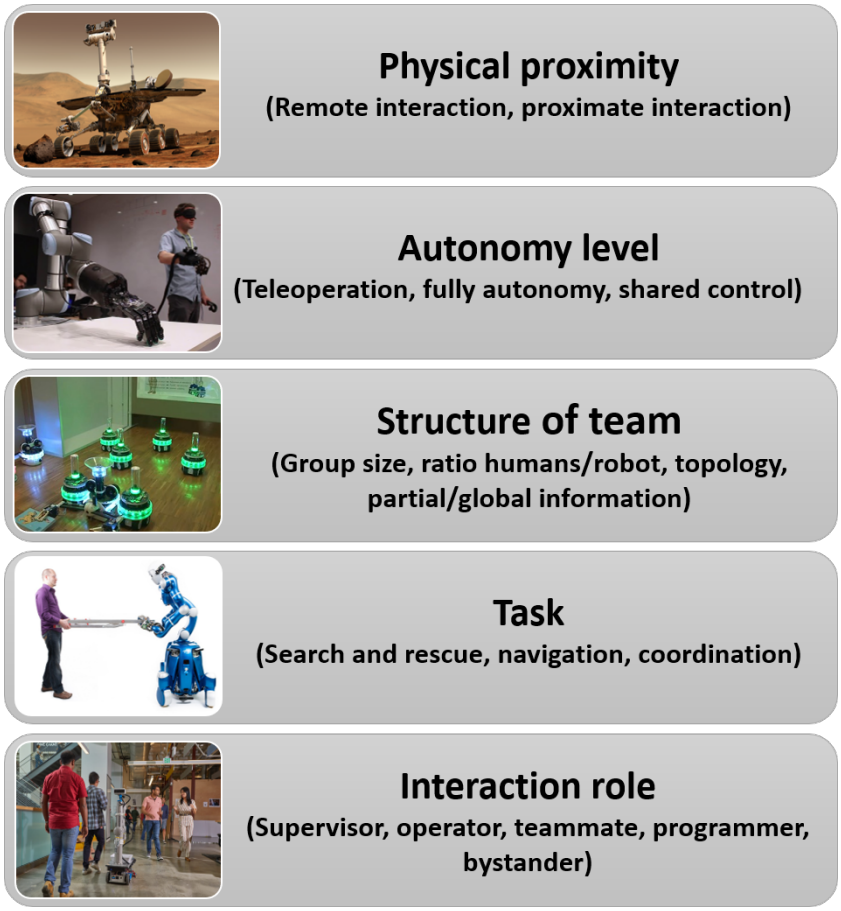


FIGURE 3.1. High-level taxonomy showing the 5 key features characterised human-robot interaction. They are: physical proximity, autonomy level, structure of team, task, interaction role. For each level a symbolic picture is reported.

3.1.1 High-level taxonomy

In order to have a clear taxonomy of the problem, we can identify some key features to take into consideration that shape the form of interaction between humans and robots (see Figure 3.1):

1. **Physical proximity.** Interaction is generally separated in two categories: remote and proximate interaction [79]. In *remote interaction* the human and the robot are separated spatially or even temporally. Representative examples are teleoperation where the human may control the robot by remote while being in a different place [176], and the autonomous Mars Rovers that are separated from Earth both in space and in time [89]. In *proximate interaction* the human and the robot are required to interact while collocating spatially close to each other. This type of interaction includes service and social robots that may be in the same room of humans or even have physical interaction with them.

2. **Autonomy level.** There are numerous formal definitions of autonomy in the literature

[22, 82, 122, 168]. Qualitatively, the autonomy level measures the percentage of time that the artificial agent is carrying out its task on its own without the intervention of a human operator. In [175], Sheridan proposed a scale for measuring automation (or the autonomy of a system) with ten levels that are accompanied by criteria against which a given system can be evaluated. According to such scale, teleoperated robots that are fully controlled by an operator, usually from a distance, have a *level of autonomy (LOA)* equal to unity. At the other end of the scale there are fully autonomous robots, e.g., robots that perform unassisted tours in (possibly) unknown environment and delivery robots, with a *level of autonomy* equal to ten. In between there is the so called *shared control*, where humans and robots have different responsibilities in a shared task.

3. **Structure of the team**. The HRI problem is clearly not restricted only to the case of a single human with a single robot (dyadic interaction), but the interaction can involve, possibly mixed, teams of multiple humans and multiple robots. For example, in search and rescue operation, robots are typically managed by two or more people, each with a different role in the team [129, 173]. The problem becomes harder in the case of multi-agent systems since additional factors arise that need to be handled. Such factors include the group size, the ratio of humans to robots, the communication topology, whether each agent has global or partial information about the team, and whether the robot subgroup is homogeneous or heterogeneous (different artificial agents in the same team may differ in implementation and role). Besides the structural properties of the group, the agents can interact with each other in order to fulfil a common goal (cooperative system) [90, 151] or have non-aligned goals acting only to maximise their own gains (competitive system) [119, 132, 151].

4. **Task.** The introduction of assistive robots allows humans to perform tasks that they could not carry out before, simplifying the task for the human. Indeed, the nature of the task to be accomplished determines the type of interaction. Examples of different types of tasks are *urban search and rescue* where the rescue robots assist humans during the exploration of unknown environments, or even substitute them if the situation turns out to be dangerous [114], e.g., searching for victims in urban disaster environment, entering unstable structures; and *coordination tasks* where humans and robots need to work together on a shared task, such as removing objects from a table [135], sawing a wood log [145], passing objects from one to another [62], assembling more complex structures [128], just to cite a few examples.

5. **Interaction roles.** As said above, the nature of the task shapes the type of the interaction and in particular the role assumed by the human when interacting with a robot. In [172], five different roles are identified, that are supervisor, operator, teammate, programmer and bystander. A *supervisor* role is that of monitoring the behaviour of the artificial agent, and controlling it if needed, with respect to a goal. For example, a supervisor of an unmanned

vehicle may tell the robot where it should move, then the robot plans and carries out its assigned task without any further human interventions. An *operator* needs to interact more with the robot, teleoperating or changing its behaviour. A *teammate* (or peer) works with a robot with equal roles in order to accomplish a common task. An example of this would be performing a joint motor task (e.g., assembling a workpiece, transferring of objects from one point to another). A *programmer* needs to physically change the robot's hardware or software. Finally, a *bystander* does not interact directly with any robot but s/he needs to be aware of it because they are in the same physical space (e.g., a person who walks into a room with a robot vacuum cleaner needs to be able to avoid the robot safely).

3.2 HRI applications

Several application areas can be identified combining the key features of human-robot interaction listed in the previous section. Many of these areas have been studied for decades, raising the interest of many fields of research. They are discussed in details in what follows:

- **Assistive and rehabilitation robotics.** The main goal of this application domain is to help people whose ability to perform daily activities is limited by disease or impairment. Examples are stroke survivors with neurological injuries [18, 81, 121], elderly populations who have age-related challenges (e.g., muscle weakness) [149, 150, 163, 198], people suffering from social disorders such as schizophrenia, autism and a variety of degenerative diseases affecting cognition and motor control [15, 27, 38, 87]. Assistive robotics places the robot or the avatar in close proximity to humans, requiring to interact in a peer-like role and with an appropriate level of trust. Many ethical considerations and trust issues arise when delegating assisting artificial agents to interact with someone with motor or mental challenges, as is discussed in [187]. In recent years, artificial agents have been proposed as a mean for innovative and personalised rehabilitation plans. Indeed, successful outcomes were reported in children with cognitive impairments (ASD) who may lack social skills [38]. Still in this direction, the work in [81] aimed at developing a socially assistive robots to autonomously supervise stroke patients in clinics during their therapeutic walking. Such social agents were able to accompany and monitor the patient during the rehabilitation, provide feedback and encourage the patient to repeat the exercises regularly. A major challenge in assistive and rehabilitation robotics is to develop reliable, safe and robust social agents, aware that each human is different and flexible enough to deal with changing human needs.

- **Search and rescue.** The development of unmanned search and rescue robots (USAR) was motivated by their potential use in urban and natural disasters (e.g., fire rescue, hurricanes, earthquakes) in which USARs take part in several tasks assisting human rescuers [129]. Examples of such tasks are: i) *search* is the exploration of a building, cave or

wilderness aimed at finding someone, for example victims of a disaster; ii) *reconnaissance and structural inspection* helps human rescuers to understand the environment and to develop a general knowledge of the area of interest; iii) *serving for a team member* is a task in which the robot works side-by-side with rescuers to assist and speed up the operations. Recently effort from many researchers has been spent in making rescue robots intelligent enough so as to avoid the need for constant supervision, allowing in this way the establishment of a peer-to-peer interaction in human-robot teams.

- **Industry.** Industrial robots were built to be flexible machines, equipped with various sensors in order to perform a wide range of industrial manufacturing tasks [16, 86, 104]. In principle, such robots were kept away from humans in a separate workspace and deployed to replace humans in performing various repetitive/hazardous and tedious manufacturing tasks with high accuracy [96]. Recently, due to the technological progress, Industrial Robotics is rapidly evolving so that robots not only share the same workspace with humans, but are also considered useful side-by-side collaborators in shared work [58, 86]. An industrial collaborative robot is also called as *industrial cobot*, a term first introduced in [144] and used to underline the direct interaction between a robot and a human. Nowadays, cobots have been largely adopted in several industrial lines such as food-processing industry [189], aerospace exploration [26], health industry, automotive, construction industry [73] and assembly systems [23, 52]. Despite many results exist in this field, several challenges and open issues still need to be addressed. For example in a highly advanced industrial system, cobots are required to detect human presence and communicate with them through body language or speech processing and be fault tolerant so as to carry out the task despite any accidental failure. In addition, other open challenges regard security and safety as well as scalability and the cobot learning ability.

- **Space exploration.** Robots have long been part of space exploration. Indeed, they are sent by humans to places where the environment is too far or hostile (e.g., Moon, Mars, Venus), or used in construction tasks to build the international space station [95, 181]. The reason why robots are very suitable for space environments is that they can be specifically designed to fit the task and the environmental constraints, such as the temperature of the target environment, radiation, lack of the atmosphere and the Earth magnetic field, presence of caves and craters, etc. Successful missions in space robotics include “Soviet Lunokhods” which were used to explore the surface of the Moon [34] and “Spirit”, the rover that landed successfully on Mars as part of the NASA’s Mars Exploration mission [110]. Despite the numerous successful missions, several key challenges still exist in the field [68, 69]. One of them is enabling the robot to perform its tasks as autonomously as possible requesting human assistance only when necessary, due to the big difference in space and time between the ground control and the robot. The way of communication is also an open problem. For the success of the mission, ground control needs to communicate with in-situ

robots efficiently and clearly using for example voice-based commands and gestures. At the same time, the robot is required to be able to ask the right questions and interact with the human as a peer.

- **Other application domains** of HRI are numerous besides those cited above. These domains include entertainment, military and police, autonomous driving, educational robotics, home use and so on [35, 123, 154].

3.3 Focus on joint motor coordination task between humans and artificial agents

So far we have presented HRI in its general terms, showing the main features that characterise human-robot interaction and how they need to be combined to design an artificial agent suitable for a specific application domain. In this section we will focus on a particular class of artificial agents, that is the class of "assistive robots", or more generally *socially interactive robots* [31, 48, 49, 69]. As previously discussed, such class includes agents able to interact with humans in peer-to-peer roles exhibiting "human social" features. Socially interactive robots behaving as partners are important for domains in which they are required to fulfil specific tasks in cooperation with humans. In this thesis, joint motor coordination tasks are considered and analysed with the aim of developing an autonomous artificial agent able to reach successful motor coordination while exhibiting human-like kinematic biomarkers.

When humans work together, motor coordination usually emerges spontaneously without considerable effort or any explicit communication intent during task coordination. Nevertheless, expectations and motion planning are significantly important for a successful coordination, and they often seem to be transparent to the coordinating individuals. From a control point of view, the emergence of coordination while performing a joint motor task is a phenomenon characterised by non-linear dynamics in which an individual has to predict what the other is going to do and adjust his/her movements in order to complement those of the other so as to achieve precise and accurate spatial and temporal correspondence. If we consider that each partner can take initiatives, the adaptation performed during a collaborative task is surely a bilateral process [64, 196]. The cognitive process underlying the ability of humans to coordinate their movements with others is still not well understood and, therefore significant research effort is aimed at building agents "biologically inspired" from humans, trying to emulate human coordination skills by imitation or demonstration.

Widely known examples of motor coordination tasks include collaborative object lifting and transportation [6, 64, 164, 184], object handling [62, 94], object placing [108, 126, 192], object swinging [54, 55, 141], and sawing tasks [145, 146]. In what follows, we further detail some key results taken from the current literature with the aim of both presenting the problem of

coordination in several types of tasks and discussing the different proposed solutions to such problem.

3.3.1 Pick and place task: agent based on human dynamics

Always in the context of joint dyadic coordination, in [108] an algorithm is proposed to drive an avatar in real time joint-action "pick-and-place" tasks with a human agent (PAPAc). To perform such a task, two participants (human/virtual) stood at a table wearing virtual reality headsets while their head and hand movements were tracked by position sensors and shown in a virtual environment (Figure 3.3.2). The final goal of the task is to take an object appearing one side of the table to the target location on the other side. Both participants could pick up the object and decide whether to take it to the target location on their own or to pass it to their partner that then completes the task. The authors' approach consisted of four main steps: *i)* performing experimental sessions of the pick and place task with human dyads in order to investigate the dynamics of joint action behaviour; *ii)* analysing the main aspects emerging from human behaviour and deriving an ad-hoc mathematical model capturing it (top-down approach) [106, 107, 109]; *iii)* encapsulating such a mathematical model in the avatar [107, 108]; *iv)* validating the avatar performing dyadic sessions of the task with a human partner [106, 108].

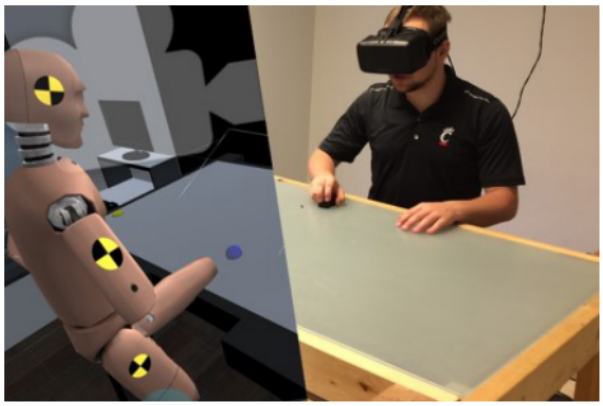
The mathematical model (a set of ordinary differential equations) was built putting together three main features of human behaviour: *i)* the trajectory dynamics of the participant's hand when moving from one point to another; *ii)* the decision dynamics that characterises the participant's choice to complete the task alone or pass the object to the partner; and *iii)* the location where the participant would choose to release the object when passing it to their partner. The authors found that this approach yields a human-based agent that successfully interacts with a human partner completing all the proposed trials. Furthermore human-virtual trials reported an average percentage of pick-and-place decisions (pass/not pass the object) and an average completion time that were not statistically different from those reported in human-human trials (see [108] for more details).

3.3.2 Wood sawing task: hybrid control approach

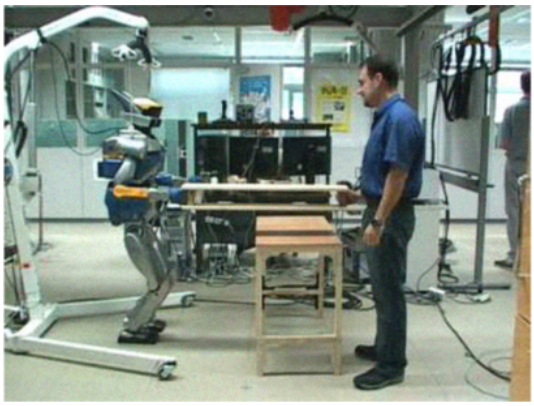
Another interesting example of coordination between a human and a robot was addressed in [145] and then in [146], where a dyadic wood sawing task was investigated (Figure 3.3.2). As the authors explained in [146], such kind of task can be split into two different phases. In the first phase, the human has to pull the saw along the motion direction while the robot has to be compliant and follow the human. In the second phase when the blade reaches the edge, the robot starts to pull the saw and the human partner stays compliant. During the task, both human and robot should apply a vertical downward force in order to push the saw teeth into the material.

In order to accomplish this task, the robot used in the paper (a KUKA Lightweight Robot [25]) was driven by an hybrid impedance/force controller. Specifically, the robot received EMG

(a) Pick and place task



(b) Object lifting task



(c) Wood sawing task

FIGURE 3.2. Dyadic coordination between a human and a robot/avatar during different types of tasks: **(a)** pick and place task (*Image credit*: [108]), **(b)** object lifting task (*Image credit*: [64]) and **(c)** wood sawing task (*Image credit*: `https://hri.iit.it`)

signals from the human's shoulder (representing his/her muscle contractions), estimated online the human stiffness and regulated its own stiffness accordingly. Such stiffness was used in the impedance controller to reach a desired position along the motion direction. Furthermore, a PID controller was used as a force controller to ensure that the contact between the blade and the wood was maintained. More complex cognitive aspects were left to the human and did not involve the robot, e.g., position of incision along the beam, orientation of the cut and execution frequency. The approach was validated experimentally showing that the robot was able to compensate the human's movement and cut the wood log in different scenarios, for example when the human stopped his motion during the task or changed his/her oscillation frequency.

3.3.3 Object lifting task: programming by demonstration

A completely different task was studied in [64], where the Programming-by-Demonstration (PbD) approach was used to solve the problem of human-robot coordination in object lifting tasks. According to PbD, an artificial system (precisely the HRP-2 humanoid robot [98]) learned from several instances of the lifting task performed by a "teacher", and became capable of reproducing

the human skill needed to perform such a task. Specifically, during the teaching phase, a human demonstrator (the "teacher") teleoperated the robot in order to lift the object with a human partner along the vertical direction while keeping it horizontal. During the demonstration, the robotic system used probabilistic methods to learn the human behaviour, in particular Gaussian Mixture Models (GMMs) to extract the main features from the task and Gaussian Mixture Regression (GMR) to reproduce it [24].

In [64] it is also claimed that during a dyadic collaborative interaction, human partners regulate their behaviour between two extreme theoretical conditions: pure leading and pure following. In order to study this human skill, two different scenarios were implemented, one in which the robot behaved as follower and the other where it behaved as leader. Two models were built during the teaching phase and then combined in the validation phase in order to make the robot able to adapt its behaviour in both conditions depending on the human partner.

The programmed robot was validated experimentally with different human partners confirming that it was able to recognise a leader or a follower behaviour of the partner and adapt its own behaviour in a complementary way. Nevertheless, the authors also showed that such robot failed when the task was performed with a different velocity and force profile; this was due to the fact that the robot learned from only two human partners and, thus, the dataset available was not general enough.

3.3.4 Limitations of the previous approaches

All the previous approaches to solve interaction tasks have some key limitations when employed to implement the dynamics of an artificial agent interacting with a human partner. In the hybrid control approach, even though the robot works in real-time, it is considered a mere follower that has to adapt its motion to that of the human without taking any responsibilities in the task completion. On the other hand, the approach implemented in the lifting task (PbD), in which the robot has a decision making mechanism (balancing between leading and following) aims at emulating human behaviour but is unable to adapt its behaviour to a changing environment (e.g., different partner, performing the task to a different velocity). Finally, PAPAc is able to complete the pick and place task with different partners showing human cognitive mechanisms. Nevertheless, the top-down approach implemented in PAPAc has the main drawback of capturing the "average decision making" of humans while performing the pick-and-place task. Indeed, a mathematical model was built ad-hoc to capture the main features emerging from the experiments, completely ignoring the differences characterising each individual (e.g. different individuals could take different decisions in the same situation).

With the aim of overcoming all these drawbacks, in the rest of this thesis we propose our own implementation of a control-based cognitive architecture able to make the artificial agent interact with another agent. In particular, we choose the mirror game as a paradigmatic coordination task (see details in Section 2.1.1) for its versatility. Indeed, the mirror game allows us *i)* to study

motor coordination between two agents while leading, following or joint improvising (that is a combination of leading and following); *ii)* to extract and analyse individual kinematic features that uniquely characterise different individuals; and *iii)* to easily extend the interaction from dyads to groups.

We formulate the problem as that of designing a multi-objective feedback control strategy able to drive an artificial avatar in performing joint oscillatory motor tasks in dyads as well as in human ensembles. Our cognitive architecture has to guarantee a bounded tracking error between the players and ensure that desired human-like kinematic features are assigned to the virtual player. To the best of our knowledge, very few results exist in literature on human group interaction and even less on multi-agent systems where robots/avatars and humans have to collaborate to perform a joint motor task.

3.4 Summary

In this chapter we gave a brief overview of human-robot interaction as an emerging field both in both industrial and in academic context. We first discussed the key features that define human-robot interaction (e.g., physical proximity, autonomy level and so on) expounding a high level taxonomy. Furthermore, we listed the main application domains that require an artificial agent interacting with one or more humans. Examples included assistive robotics, search and rescue, industry, and many others. Then, we focused on human-robot interaction in motor coordination tasks explaining the fascinating phenomenon of synchronisation emerging from motor interaction and the challenges behind the development of an avatar able to coordinate its motion with that of one or more humans. Three key examples of different motor tasks were discussed together with the different strategies used to solve them. Finally we made a comparison between all these strategies pointing out advantages and disadvantages.

In the next chapter, we begin our study of coordination between humans and artificial agents starting with the dyadic case, and providing different control solutions to drive an avatar to interact with another partner in different conditions.

CHAPTER 4

DESIGN OF AUTONOMOUS VIRTUAL AGENTS VIA NONLINEAR CONTROL STRATEGY

The main objective of this chapter, whose content has been published in [116, 118], is to design and implement an autonomous artificial agent via nonlinear control strategies, here referred also with virtual player or avatar, able to synchronise its movements with those of a human in a dyadic task. In so doing, the mirror game (introduced in Section 2.1.1) was selected as significant case-of-study. Specifically, in Section 4.1 we describe the cognitive architecture, proposed in [207], able to drive the artificial agent during a dyadic session of the mirror game both in Leader-Follower (LF) and in Joint-Improvisation (JI) interaction. After having discussed the main drawbacks of the previous architecture, in the Section 4.2 we completely re-design the "architecture block" giving the human-like reference by means of learning strategies. Such a learning-based reference generator block is put back into the nonlinear controlled architecture in order to overcome the discussed drawbacks of the previous architecture. In particular, we show step-by-step the design of an improved virtual agent able to interact with humans autonomously and in a more natural way. Then, we move to Section 4.3 to validate experimentally our new Virtual Player (VP) in dyadic session of the mirror game with a Human Player (HP). A summary of the results is provided in Section 4.4.

4.1 Virtual player's architecture

In this section we describe briefly the cognitive architecture designed for the virtual player and proposed in [207, 208]. Such an architecture is grounded in the nonlinear control theory, characterising the participants' motion through a nonlinear mathematical model and considering their interaction as a bi-directional feedback. It has been shown that the presence of feedback is

fundamental for the emergence of the human interpersonal coordination [139].

It is required that such a virtual player (or avatar) plays the mirror game with a human player both in the leader-follower and joint-improvisation configurations while exhibiting desired kinematic features similar to those of humans.

The cognitive architecture is shown Figure 4.1 and is mainly composed of two parts:

- *an inner dynamics*, representing how the virtual player moves when not interacting with any other participant. This is modelled using a nonlinear Haken-Kelso-Bunz oscillator (HKB);
- *a control strategy*, that takes care of the coupling with another agent while making it behave in a human-like fashion. Such a control algorithm allows to reach two objectives: 1) temporal correspondence, that is the minimisation of the position error between HP and VP, and 2) similarity to a desired motor signature in order to guarantee that the kinematic behaviour is similar to that of a human. To solve this problem, we consider an optimal control approach.

4.1.1 Haken-Kelso-Bunz model

The Haken-Kelso-Bunz model, proposed in [83], was the first mathematical formalism capable to describe the observations made in human rhythmic coordination experiments relying on the concept of synergy and theory of nonlinear oscillators. Firstly used to explain the intra-personal coordination, such as the coordination between the oscillation of the two index fingers, it became a general model largely adopted in the literature to capture both dyadic interaction between two different subjects [14, 19, 72, 158, 171] and group interaction in human ensembles [8].

The model is described by the following equation:

$$\ddot{x} + \left(\alpha x^2 + \beta \dot{x}^2 - \gamma\right)\dot{x} + \omega^2 x = u, \tag{4.1}$$

where x, $\dot{x}$ and $\ddot{x}$ represent the position, velocity and acceleration of the VP correspondingly, u is the control input modelling the interaction with the other agent, ω is the natural oscillation frequency of the generated motion when u is set to zero, α, β and γ are parameters characterising the nonlinear damping term.

4.1.2 Nonlinear optimal control

The control input u is chosen following an optimal control solution aiming at minimising a cost function that depends both on the position error between HP and VP and the velocity error between that of the virtual agent and the reference individual motor signature. Specifically such a cost function, clarified in (4.3), consists in three terms: 1) a term modelling the position mismatch between the virtual agent and the human agent; 2) a term modelling the error between the

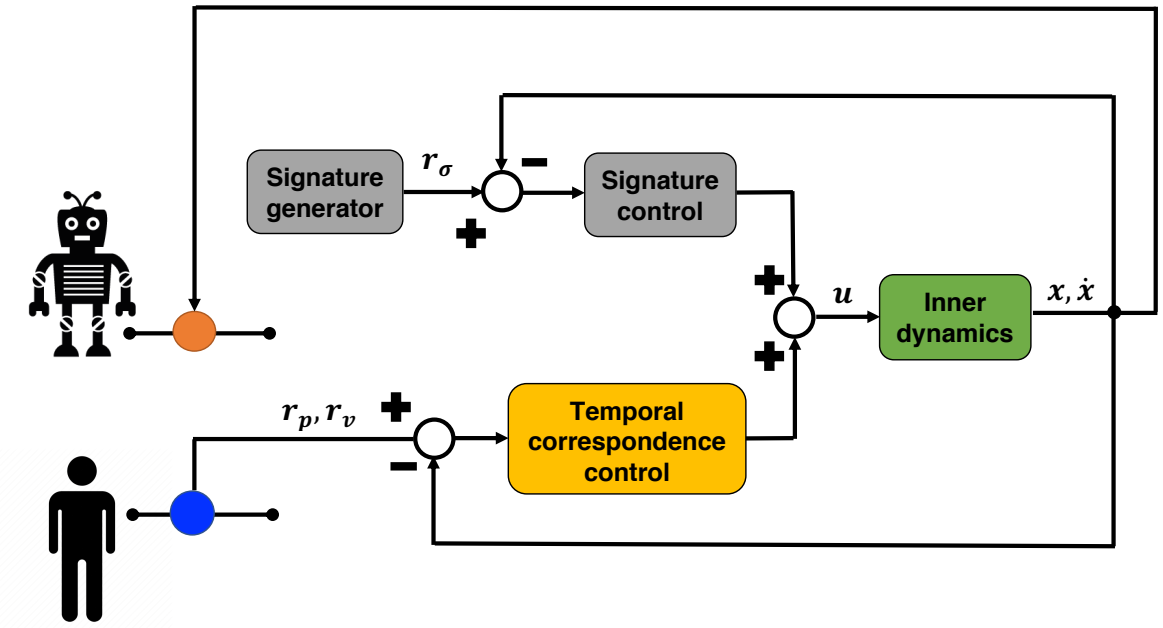


FIGURE 4.1. Interactive cognitive architecture designed for the virtual player. The green block models the inner dynamics of the virtual player, the yellow block minimises the position and velocity error between the virtual and human player, whereas the grey blocks confer a desired IMS on the virtual player. Blue and orange circles represent the end effector's position of the human and the virtual agent; $x, \dot{x}$ are position and velocity of the virtual agent, whereas r_p, r_v are measured position and estimated velocity of the human partner; r_σ is the velocity signal reference characterising the desired human motion; u is the control input.

velocity of the VP and the desired reference signature to be exhibited and 3) a term minimising the control effort. The resulting optimising problem constrained to the dynamics in (4.1) is:

$$\min_{u \in R} J \tag{4.2}$$

with:

$$J(t_k) = \frac{1}{2}\theta_p \left(x(t_{k+1}) - r_p(t_{k+1})\right)^2 + \frac{1}{2}\int_{t_k}^{t_{k+1}} \left(1-\theta_p\right)\left(\dot{x}(\tau) - r_\sigma(\tau)\right)^2 + \eta u(\tau)^2 \, d\tau, \tag{4.3}$$

evaluated in each finite time interval $[t_k, t_{k+1}]$. Specifically, $x, \dot{x}$ are position and velocity of the VP, r_p is the measured position time series of the HP, r_σ is the reference velocity corresponding to the desired motor signature of the VP, η is a positive weight used to tune the control energy, and finally θ_p is a constant parameter in $[0,1]$ that makes the VP more responsive to the position mismatch with the human. Tuning properly the weight θ_p, the virtual agent can act more as a leader ($\theta_p \to 0$) or as a follower of the human player ($\theta_p \to 1$).

Signature generator The signature generator is a block of great importance in the cognitive architecture depicted in Figure 4.1. Indeed, it is responsible to generate realistic synthetic velocity profiles taken as reference IMS and is the core of the "human-like" property characterising the virtual player's motion. In the cognitive architecture proposed in [207] and described above, a pre-recorded human velocity trajectory is used as reference input to the control law.

4.1.3 Limitation of the previous architecture

The use of a pre-recorded motor signal as human-like reference input to the control law represents the main drawback of the presented architecture. Indeed, these references have been collected off-line, recording the kinematics of several human players performing the mirror game in the solo condition. The use of a pre-recorded signature makes the VP's motion deterministic and repetitive between consecutive game sessions (especially in leader mode) due to the fact that the controller always has the same reference input.

A different point of view is introduced by artificial intelligence techniques. Through observational learning it is possible to have an avatar that learns how to move simply *observing* a human playing more than once. The aim of the observational learning is to derive, through examples of the real process, some common and characterising features of how a human plays the mirror game in order to generalise them and build an internal description model used by the virtual agent. Having such an internal model, the virtual agent will be able to generate as many reference inputs as required having the same features as those used during the learning phase.

Here, we introduce a stochastic modelling approach based on artificial intelligence techniques, specifically on Markov Chains (MC), that is able to observe the player's motion and capture the key features of his/her IMS in order to generate new IMS having the same properties. Then, this model will be embedded in the control scheme as stochastic signature generator. In this way the virtual player will be able to generate synthetic reference signals autonomously to use during a game session having a desired IMS. To better understand the Section 4.1.3, more theoretical details about the Markov chains are given in what follows.

Markov chains. A Markov model is a stochastic model used to describe randomly changing system [134, 152]. For any given system, a Markov model consists of a list of finite possible states in which the system can be, a set of transition paths between two states with the corresponding probability, and a set of possible observations as output for each state. If the states are observable, they correspond to the model outputs and the Markov model is specified by the so-called Markov chain.

Denoting by N the number of all possible states and by $s_k = i$ with $i \in [1,...,N]$ that the system is in the state i at time instant k, a Markov chain is fully characterised by:

- an initial state s_0;

- a transition matrix $A := [a_{ij}]$ where $a_{ij} := P(s_{k+1} = j | s_k = i)$ is the probability of being in the state j at time instant $k+1$ given the state i at time instant k;

- the property of "limited horizon assumption" specifying that the probability of being in a state at time k depends only on the state at time $k-1$:

$$P(s_k | s_{k-1}, s_{k-2}, ..., s_0) = P(s_k | s_{k-1}) \tag{4.4}$$

4.2 Generation of synthetic individual motor signature

As discussed above, to successfully interact with and learn from humans in cooperative modes, virtual agents need a mechanism for recognition, characterising and emulating human motion. Thanks to Markov chains, we can generate new IMS in real-time building an autonomous virtual agent that has previously learnt to move as a human player having an own signature.

4.2.1 Modelling

The procedure for MC-based modelling consists of the following three steps:

1. *Data collection and pre-processing*: it is essential that the raw data recorded from a human player performing sessions of the mirror game in solo condition are represented in a homogeneous and invariant form before starting to build the MC model. In so doing, we consider the Short Time Fourier transform (STFT) and the Vector Quantization (VQ) techniques [7];

2. *Markov model training*: the pre-processed data are encoded in the MC in order to find the right parameters for the transition matrix A;

3. *Data movement generation*: having the MC model defined in the previous step, new synthetic movement data are generated sharing the same kinematic properties as those of the input data.

Step 1 - Data collection and pre-processing. A Hamming window of a certain width is first used to partition the sampled input signal into a finite set of frames ("signal windowing"). To prevent loss of information and in order to have the minimum distortion of the signal, two consecutive Hamming windows are overlapped by $\frac{3}{4}$ of the window's width. In this way, the sum of the sequence of the windows is a resulting flat-top window [85]. Then, the STFT is performed for every data segment so that a vector of Fourier transform's coefficients, or "feature vector", can be associated to it. The STFT can be viewed as a "local spectrum" of the sampled position signal $x(t_n)$ in an analysis window long N samples. Formally, the short time Fourier transform on the

m-th window is defined as:

$$X_m(\Pi) = \sum_{n=1}^{N} x(t_n)\Pi(t_n - mN)e^{-\jmath t_n \Pi}, \tag{4.5}$$

where $\Pi(t_n)$ is the window function (in our case, the Hamming window).

Since the MC is a discrete model, the set of feature vectors obtained from the STFT is then processed through a vector quantizer in order to get a finite set of representatives, called "symbols". A vector quantizer is completely decided by a code-book, which consists of N fixed prototype vectors generated by an appropriate algorithm. In this work, the Lloyd's algorithm is used to produce the VQ codebook [166]. Also known as Voronoi iteration, it is an algorithm for partitioning the Euclidean space into well-shaped and uniformly sized convex cells (called Voronoi cells). Concretely, it repeatedly finds the centroid of each partition and then re-partitions the space according to which of these centroids the inputs are closest to. Since the operation of quantization inevitably causes distortion between the original data and the quantized ones, the VQ has to map each feature vector to one of the prototype vectors of the codebook in order to minimise their squared distortion error. The indices of the N prototype vectors (integers from 0 to $N-1$) are used as symbols in the output of the discrete Markov chain. The complete process is depicted in Figure 4.2.

Step 2 - Markov model training. The features, encoded in symbols, are used to build the Markov chain representing the human IMS to which those features belong. Building a Markov chain means deciding the elements of the transition matrix A, whose elements are the probability to move from one state to another. The states of the MC model correspond to the codebook symbol set, i.e., the model has as many states as symbols. Despite its simplicity in the training, having as many states as symbols was found to be sufficient to correctly capture the variability detected in the human's motion. The Baum-Welch algorithm [152] is used on an extensive experimental data set to estimate the model parameters, which is essentially based on a frequential approach aimed at finding the maximum likelihood estimate of the coefficients given a set of observed feature vectors.

Step 3 - Data movement generation. Thanks to its stochastic nature, the MC itself is able to generate a random sequence of symbols according to the probabilities included in the transition matrix. Having this new sequence of symbols, it is possible to reconvert it into a function of time through reverse pre-processing. To be more specific, the generated sequence of symbols is dequantized with the same codebook used for the forward pre-processing in order to be mapped into a sequence of prototype vectors of Fourier transform coefficients. Then, the inverse STFT is applied to each vector of coefficients using the overlap-add (OLA) method [44] that works on the discrete convolution of the signal. The main advantage of this method is the possibility to reconstruct a smooth position time series over time, removing the discontinuities obtained from

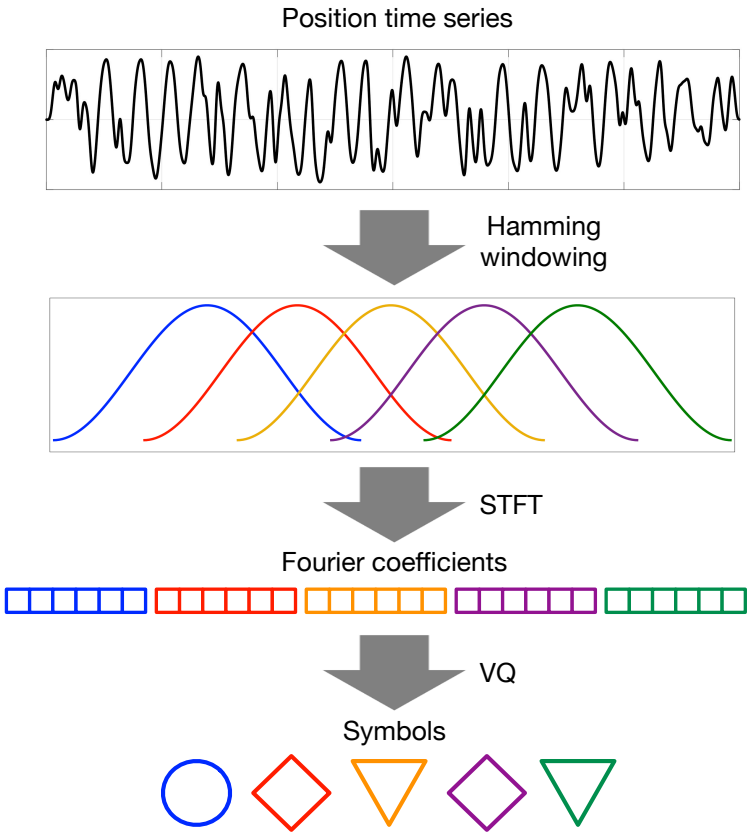


FIGURE 4.2. The recorded position time series (in black) is windowed in several data segments through the Hamming window. The STFT is performed for each segment in order to extract a feature vector represented as a cells array and then quantized in a symbol. For a better understanding, each symbol is represented conceptually as a geometric shape. Different colours are used to highlight the pre-processing steps for different frames of the signal.

simply concatenating two random symbols. The generated position data are then filtered with a zero-phase 2^{nd} order lowpass Butterworth filter having the cut-off frequency 10 Hz.

4.2.2 Similarity metrics

Before moving to Section 4.2.3, the different analysis tools considered to validate correctly the MC-based signature generator are explained:

1. *Probability Density Function* (PDF) of the velocity of the VP that defines the exhibited IMS.

2. *Skewness* and *kurtosis*: used to assess if the velocity trajectories generated by the MC have the features of a human movement. Skewness s and kurtosis k are respectively the 3^{rd} and 4^{th} moment of a curve. Let us suppose of having the curve $f(t)$ with mean value μ and

variance σ defined on $T = [t_1, t_2]$, the 3^{rd} and 4^{th} moments are evaluated respectively as:

$$s = \frac{1}{\sigma^{\frac{3}{2}}} \int_{t_1}^{t_2} \left(t - \mu\right)^3 f(t)\, dt \tag{4.6}$$

$$k = \frac{1}{\sigma^2} \int_{t_1}^{t_2} \left(t - \mu\right)^4 f(t)\, dt \tag{4.7}$$

Skewness indicates asymmetry in acceleration and deceleration, while kurtosis provides information about uniformity of the maximal velocity. Low kurtosis means that an object is quickly accelerating and decelerating and keeps constant velocity in between, whereas high kurtosis means that the object is accelerating slowly, and after reaching maximum velocity it almost immediately starts to slow down [179].

3. *Earth Mover's Distance* (EMD): also known as Wasserstein metric, it is used to quantify the similarity between the PDF of the velocity of the signal generated by the MC and that of the corresponding human player on which the MC has been trained. In case of univariate probability distribution, the EMD is given by the area of the difference between their cumulative distribution functions (CDF) [40, 179]:

$$EMD(PDF_1(z), PDF_2(z)) = \int_Z |CDF_1(z) - CDF_2(z)|\, dz, \tag{4.8}$$

where both CDFs have support in Z. The EMD is normalized to its maximum value; because the quantity $|CDF_1(z) - CDF_2(z)| \leq 1$, the maximal EMD is given by the length of the support Z. Moreover, the two PDFs are more similar the more the EMD tends to zero, meaning that the two velocity profiles are exactly equal and so a perfect synchronization is achieved. Viceversa, if the two velocity profiles have no overlap the EMD is equal to Z.

The EMD has been proved to be more robust than the other existing techniques to compare and measure the difference between two distributions. In particular, the EMD has been proved to be more efficient and robust than histogram matching techniques because it avoids quantisation and binning problems typical of histograms. Furthermore, it can be computed efficiently and applied to variable-length representations of distributions [165]. Examples of histogram matching techniques are the Kolmogorov-Smirnov distance, quadratic-form distance, histogram intersection, to cite a few.

4. *Similarity space*: the multidimensional scaling (MDS) [29, 178] is used to study the relations between the velocity profiles of the players. MDS allows us to model the players' motion as points in an abstract geometric space, called "similarity space" reducing the dimensionality of the data but preserving as much information as possible.

To build the similarity space, first of all we compute the EMD between all the analysed velocity PDFs in order to write the distance matrix $D := \left[d_{ij} = EMD\left(PDF_i; PDF_j\right)\right]$. MDS is used to reduce the cardinality of each row of the matrix D to just two coordinates (x

and y), corresponding to the two most significant eigenvalues of the distance matrix D. Using MDS, we guarantee that the Euclidean distances between elements in the similarity space are a good approximation of the EMD between velocity profiles; the closer the points are in the similarity space, the more similar their velocity profiles are. Further analysis conducted in [178] revelead that the x-coordinate of the similarity space is correlated to the mean absolute velocity, whereas the y-coordinate to the mean value of the kurtosis of the velocity distribution.

4.2.3 Validation

In this section, we present the validation of the proposed methodology to generate synthetic motor signatures. Specifically, we carried out experiments with the following setup:

- *Participants*: a total of 6 people participated (1 female and 5 males). All the participants were right handed and none of them has physical or mental disabilities. All of them took part in the experiments voluntarily, signing an informed consent in accordance with the Declaration of Helsinki. Any information obtained in this study remained confidential and participants identity is kept anonymous.

- *Experimental task*: each participant was asked to carry out 30 different trial of 30 seconds each playing in solo condition mode. The given instruction was to oscillate the index of the preferred hand over a position sensor in a spontaneous way from left to right in order that his/her individual motor signature could emerge.

- *Experimental platform*: all trials were performed through Chronos, a house-made software tool developed to study movement coordination [12]. In particular, the one-dimensional motion of the player is recorded through a leap motion controller (position sensor) [193] and sent to a laptop computer. The human trajectory is then visualised on the computer screen as a blue solid circle (Figure 4.3).

Each position time series, also referred simply as “trial”, was recorded at a sampling rate of 10 Hz, interpolated then to 100 Hz with a shape preserving spline interpolation [80] in order to guarantee some constrains presented in the given data, such as the monotonicity and the convexity. Such a signal is windowed with a Hamming window of 60 samples with an overlap between two consecutive windows of 45 samples. The number of levels in the codebook and so the states of the MC was chosen heuristically. Precisely, we tested different numbers of levels spanning from 32 to 256 in order to find a quantizer that minimises the distortion error with the minimum number of levels. We found that this trade-off was met by a quantizer with 256 levels. In Figure 4.4, the IMS of a human player is compared to different synthetic IMSs generated by Markov models built with codebooks with different cardinalities.

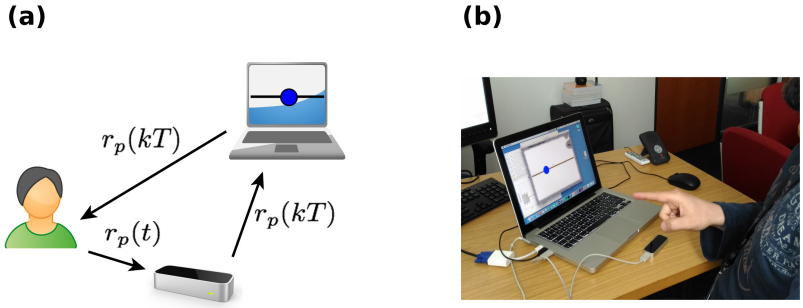


FIGURE 4.3. **(a)** Experimental equipment needed to carry out sessions of the mirror game in solo condition. The position of the human's finger $r_p(t)$ is captured by the leap motion and sent to the laptop computer. The measured position $r_p(kT)$, with the sampling period T, is displayed on the screen as a blue circle. **(b)** Picture of a human playing the mirror game.

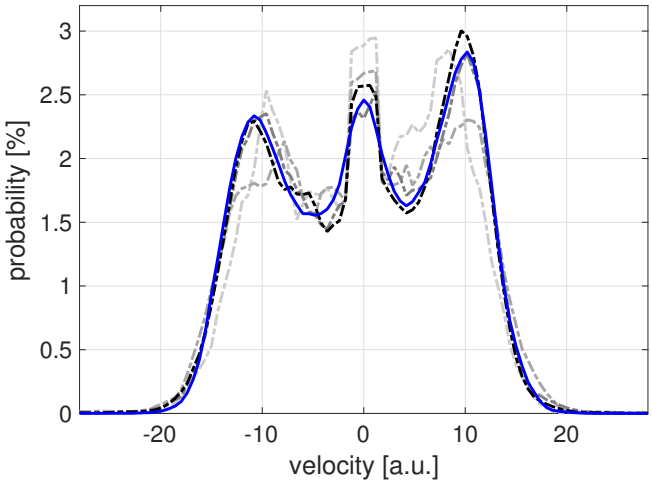


FIGURE 4.4. A continuous blue line represents the IMS of the human player over 30 trials, whereas dashed lines spanning from light grey to black are the synthetic IMS generated by the MC with a codebook respectively with 32, 64, 128 and 256 levels.

At the end of the training, 6 different MCs have been derived, one characterising each player. In what follows we compare 30 new motion signals generated by the MC with the motion signals belonging to the HP, with which the MC was trained. For a better graphic visualisation, our results are shown for just 3 players out of the 6 involved.

Figure 4.5 shows the velocity PDF of the signatures recorded by three different human players with the corresponding virtual velocity PDF obtained by generating synthetic signatures with the built Markov models. It is possible to appreciate that the Markov chain is able to capture the

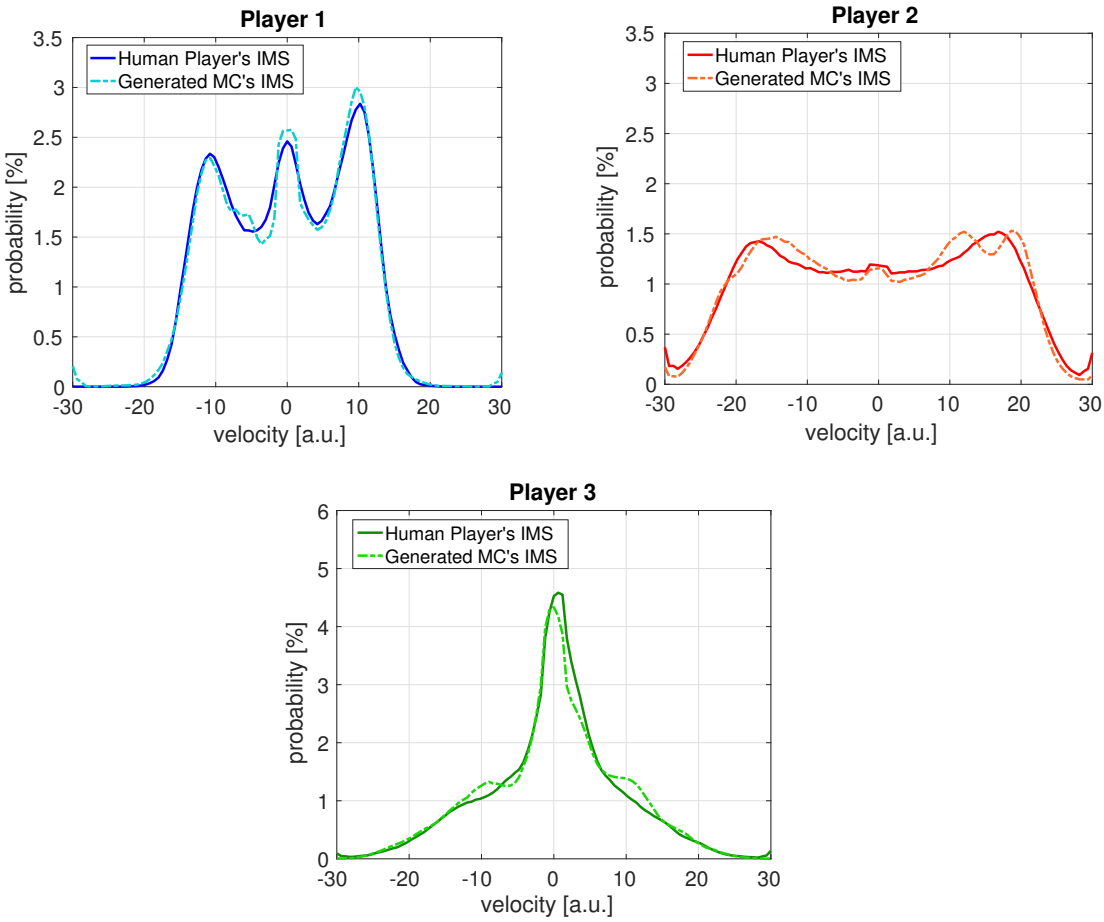


FIGURE 4.5. Velocity distribution of the signatures belonging to three different human players (solid line in blue, red and green) and those generated artificially by the corresponding Markov models (dashed line in light blue, orange and light green).

differences in the motor kinematic belonging to different players.

Before evaluating the skewness and the kurtosis the velocity time series is divided in segments. Each segment is the part of the velocity time series between two consecutive turning points, representing that the player is moving in only one direction (from left to right or from right to left). In this way we can distinguish between positive and negative velocity that correspond to the two different directions. To perform a meaningful comparison, the velocity segments have been time re-scaled to a common support (in our case $[0,1]$) and normalised in order that the area under the curve is unitary. Velocity segments with less than 20 samples have been ignored because of the probability of noise. For each velocity segment so pre-processed, we evaluate the skewness and kurtosis using (4.6) and (4.7).

In Figure 4.6 for each player the velocity segments are divided in positive and negative segments (representing the two directions of the movement), and each of them is plotted as a dot in the skewness-kurtosis plane. We observe that the generated movements have similar

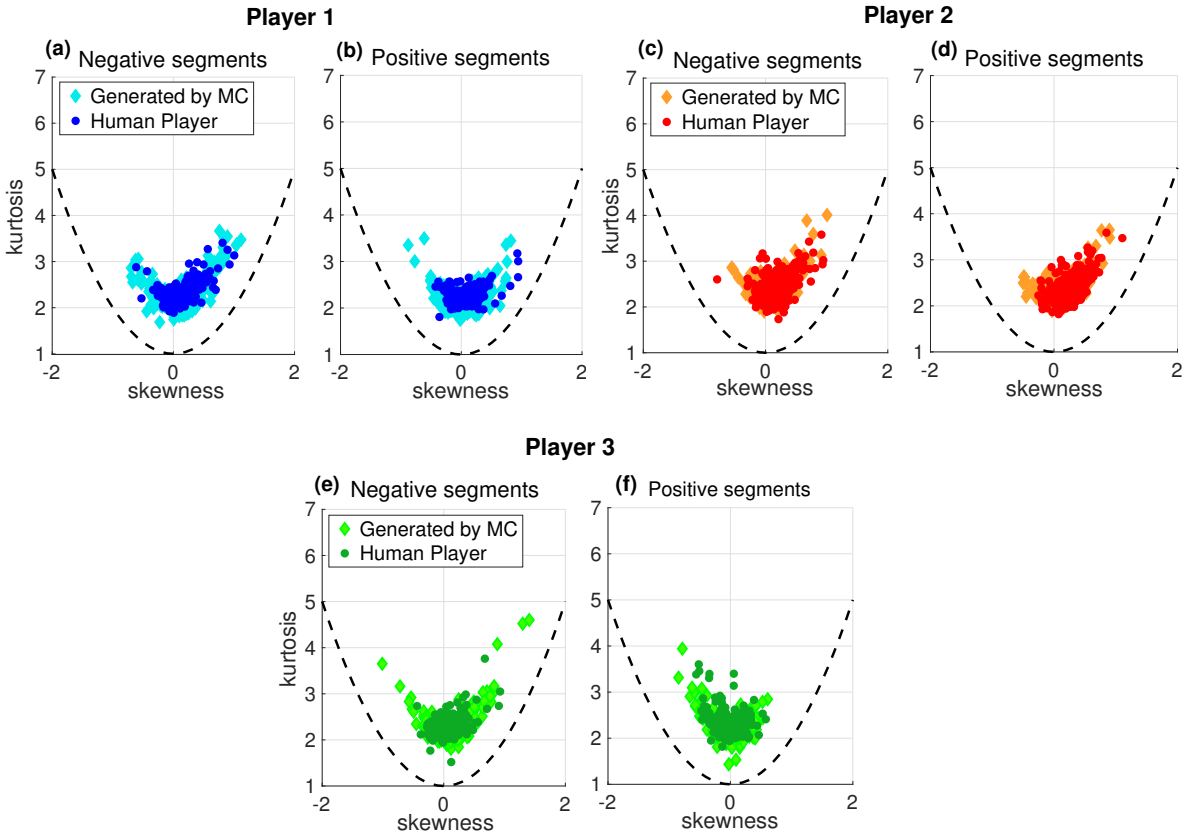


FIGURE 4.6. Skewness-kurtosis plane, having the skewness on the x-axis and the kurtosis on the y-axis. Two different plots for each player are used to divide the positive **(a),(c),(e)** from the negative velocity segments **(b), (d), (f)**. The segments belonging to the three human players are respectively represented as blue, red and green dots, whereas those generated by the corresponding MCs are represented as light blue, orange and light green diamonds. A dashed black line shows the theoretical relationship stating that the values of skewness s and kurtosis k must be above the function $k \leq s^2 + 1$ [177].

characteristics to those of real human players since they are located in the same area.

As third comparison metric, the similarity space is depicted in Figure 4.7 for all 6 players that took part in the experiment. Each velocity distribution extracted by a single trial is represented as a point in such abstract space. All the points belonging to the same agent (human or MC) are encircled by an ellipse corresponding to 0.7 mass of bivariate normal distribution fitted to the points in the space. Each ellipse can be considered as a characteristic region identifying the motor kinematic features of a single individual. It is possible to appreciate how the characteristic regions of the signature generated by the MC are included in or overlapped by the regions of the respective human players, showing that the MC approach is able to generate synthetic signatures that well reproduce the human kinematic features of the real ones.

The metrics used in this Section, even though they are effective and sufficient for our analysis,

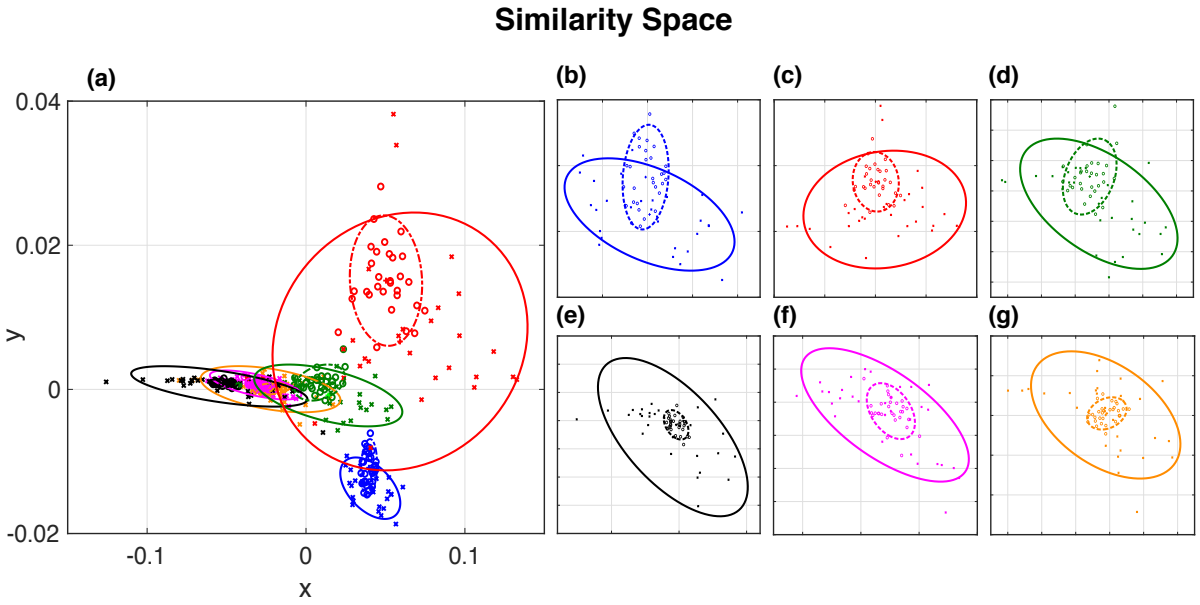


FIGURE 4.7. Each trial is represented as a point in the plane, an ellipse with the same colour encircles all the points belonging to the same agent (human player or Markov model). Empty circles identify the trials performed by the human players, while crosses identify those generated by the Markov models. Different colours are used to identify different players with the respectively models. **(a)** All the regions belonging to the 6 human players (solid line) are represented together with those defined by the Markov chains (dashed line). **(b)-(g)** Six different plots are depicted individually for each couple human player - Markov model.

belong to a qualitative validation. Quantitatively, ANOVA can be performed to assess whether the points in the similarity space belonging to the MC and the corresponding human player came from the same distribution (each point corresponds to the velocity profile of a session of the mirror game in solo condition). An absence of statistical difference implies that the velocity PDF (or IMSs) have been generated by the same player. Performing such a statistical analysis is a mathematical challenge due to the limited studies existing in the literature addressing ANOVA for high dimensional means [37].

4.3 MC-based virtual player in the control loop

In the previous section, we have developed and validated a method based on Markov chains to capture and reproduce the kinematic features of the movement of a human player playing the mirror game in solo condition. In what follows, we close the control loop embedding the Markov chain model in the wider architecture of the virtual player replacing the pre-recorded signature (described in Section 4.1). In this way, the artificial agent can play dyadic sessions of the mirror

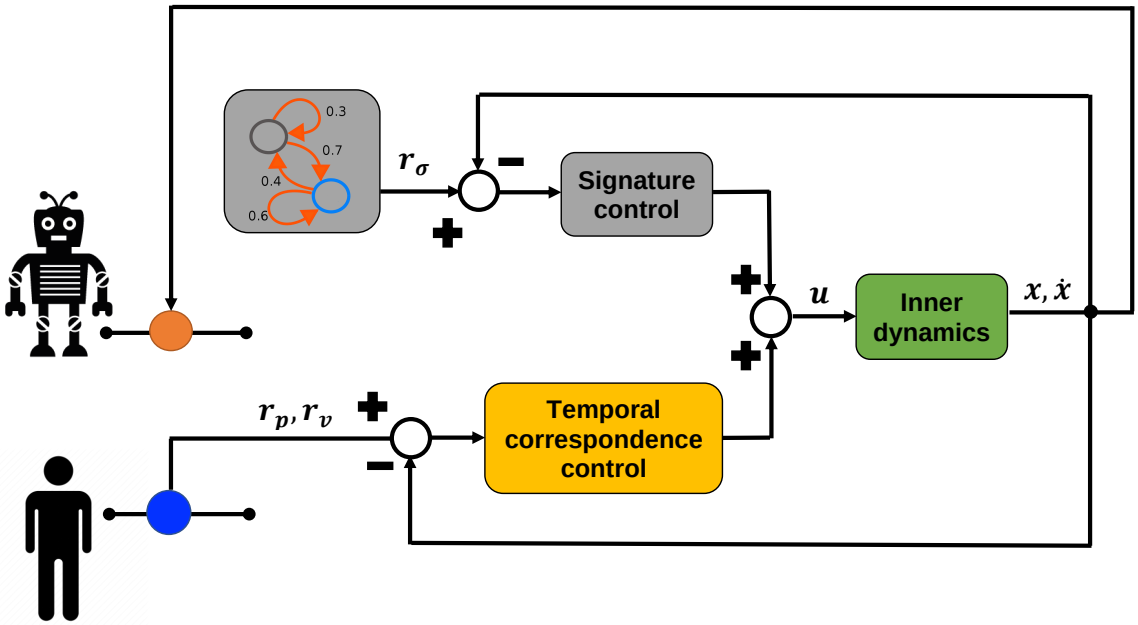


FIGURE 4.8. Interactive architecture designed for the VP, where $x, \dot{x}$ are position and velocity of the virtual agent, r_p, r_v are measured position and estimated velocity of the human player, u is the control action moving the VP's end effector and r_σ is the stochastic motor signature generated by the Markov model and given in input to the control law as reference.

game autonomously with any human player both in leader and in follower configuration. By means of the Markov chain, we use a stochastic model to generate reference signals in real time and that are always different due to its stochastic nature. This overcomes the limitations of the previous architecture that used pre-recorded movement data to generate the reference signal in velocity. The cognitive architecture so improved is in Figure 4.8.

4.3.1 Tracking metrics

To evaluate the performance of the virtual player embedding a Markov chain as signature generator, when it is engaged in a dyadic leader-follower session with a human player, we evaluated different metrics:

1. *Root mean square error* (RMSE) between the position time series of the VP (x) and that of the HP (r_p) defined as:

$$RMSE = \sqrt{\frac{1}{N}\sum_{k=1}^{N}\left(r_{pk} - x_k\right)^2}, \tag{4.9}$$

where N is the number of samples of the player time series and k is the single time instant. Lower values of RMSE imply a good position tracking between the two players along the trial.

2. *Relative position error* (RPE) is used to verify whether the follower effectively tracks the leader at the time t [179]. Supposing that the virtual player is the follower and the human player the leader, it is defined as:

$$RPE = \begin{cases} \big(r_p(t) - x(t)\big)\, sgn\big(r_v(t)\big) & \text{if } sgn\big(r_v(t)\big) = sgn\big(\dot{x}(t)\big) \neq 0 \\ \big|r_p(t) - x(t)\big|, & \text{otherwise} \end{cases} \tag{4.10}$$

where r_p and x (r_v and $\dot{x}$) are positions (and velocities) of the human player and the virtual player respectively. Positive values of the RPE indicate that the latter is behind the former.

3. *Relative phase* (RP), defined as the difference between the phase signals of the two players:

$$\Delta\Phi = \Phi_{HP} - \Phi_{VP}, \tag{4.11}$$

where $\Phi_{HP/VP}$ is the phase of the HP and the VP respectively and estimated by means of the Hilbert transform of the position time series as explained in [102]. According to the definition of phase leadership - that the leader is the player with the greatest phase [11] - we use the sign of the relative phase to distinguish the leader from the follower. Indeed, a positive $\Delta\Phi$ means that the HP is moving ahead of the VP and so the HP is the leader. Viceversa, a negative value of $\Delta\Phi$ means that the VP is leading the HP.

4. *Circular variance* (CV) is used to quantify the level of coordination between the two players. The CV is formally computed as follows [103]:

$$CV = \left\| \frac{1}{N} \sum_{k=1}^{N} e^{J\Delta\Phi_k} \right\| \qquad \in [0,1], \tag{4.12}$$

where N is the number of the samples, Φ_k is the relative phase at time instant k and $\|\cdot\|$ denotes the 2-norm. Higher the circular variance is, higher the coordination is.

5. *Time lag* (TL) describes the amount of time shift that achieves the maximum cross-covariance between two position time series. It can be interpreted as the average reaction time of the player in the mirror game [138].

4.3.2 Validation

In Figure 4.9 and 4.10, the virtual player is validated both in follower and in leader configuration respectively while playing a dyadic session with a human player. Here we test the VP with embedded the Markov model trained on the player 2 (on the 6 models available), but any other Markov model can be used.

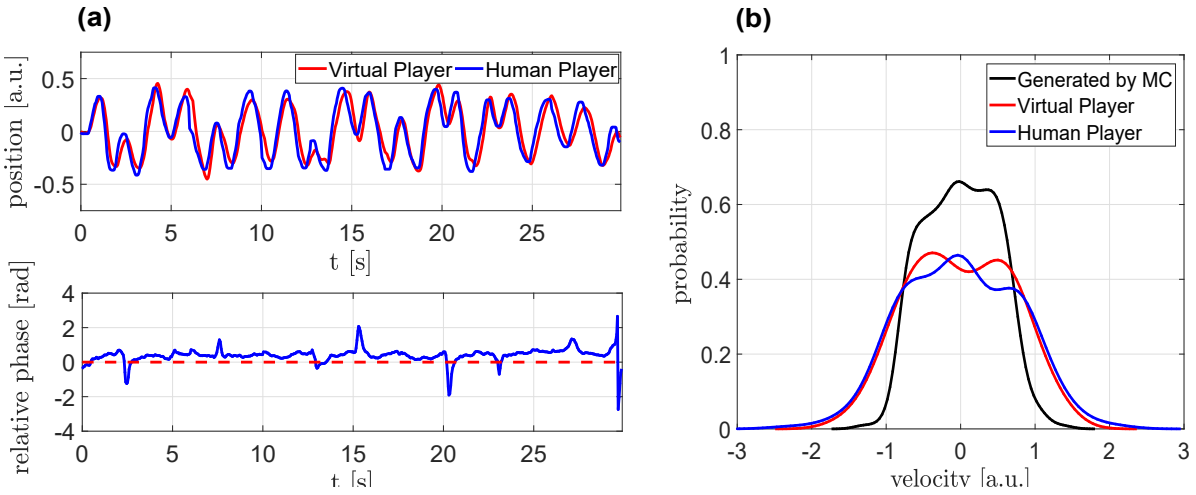


FIGURE 4.9. Dyadic session of the mirror game having the virtual player as follower (in red) and the human player as leader (in blue). In black the individual motor signature generated by the Markov chain model. **(a)** Position time series, relative phase and **(b)** the velocity distribution are shown for a single trial of 30 seconds long.

In follower configuration (Figure 4.9), the control architecture has its parameters heuristically set as follows: $\alpha = 1, \beta = 1, \gamma = -1, \omega = 0.1$ for the dynamics; $\eta = 10^{-4}, \theta_p = 0.9$ and a sampling period of $dt = 0.03$s for the optimal control law. When the VP acts as a follower, the relative phase is prevalently positive meaning that the human player is a phase leader ($\Delta\Phi = 0.394 \pm 0.408$), the circular variance is 0.933 showing a high level of synchronisation between the players, and the RMS error between the position trajectories is only 0.112. Furthermore since the virtual player acts as follower, it is able to adapt its behaviour to the HP, making its PDF more similar to that of HP partially ignoring that one generated by the Markov chain, EMD(HP, VP) = 0.02 and EMD(Ref, VP) = 0.006.

In leader configuration, we have $\alpha = 1, \beta = 2, \gamma = -1, \omega = 0.8$ for the dynamics; $\eta = 10^{-4}, \theta_p = 0.1$ and a sampling period of $dt = 0.03$s for the control law. In this case the CV reaches 0.868, the RMS error of the position is 0.122, and the relative phase is negative in average ($\Delta\Phi = -0.664 \pm 0.574$) meaning that this time the phase leader is the VP. Furthermore, the virtual player takes into account more its kinematic characteristic than that of the human partner, driving in this way the human player to behave in a similar way, EMD(Ref, VP) = 0.03 and EMD(HP,VP) = 0.02.

4.4 Summary

In this chapter we described the interactive cognitive architecture, proposed in [207], able to drive the virtual player in dyadic sessions of the mirror game with a human partner both in

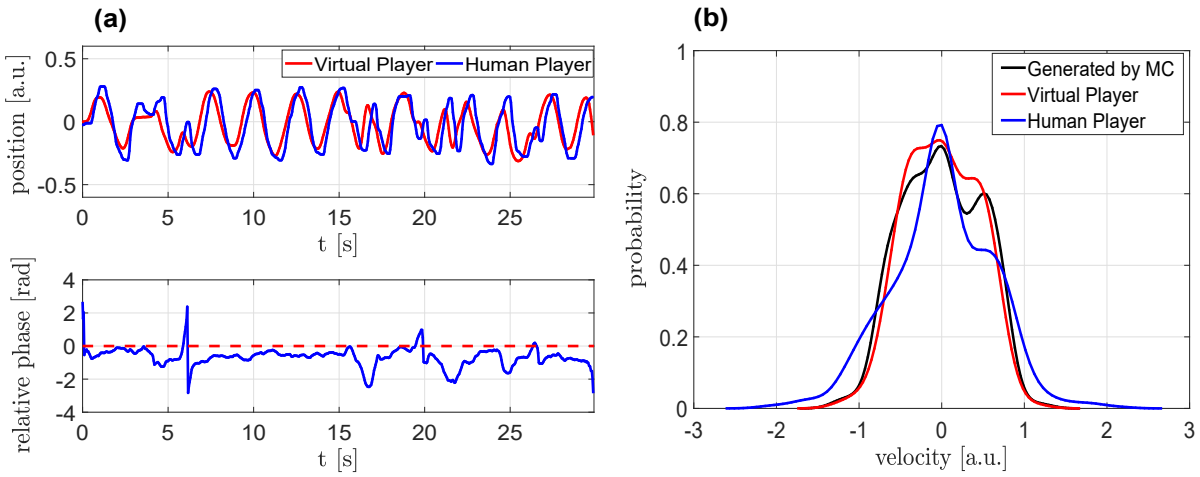


FIGURE 4.10. Dyadic session of the mirror game having the virtual player as leader (in red) and the human player as follower (in blue). In black the individual motor signature generated by the Markov chain model. **(a)** Position time series, relative phase and **(b)** the velocity distribution are shown for a single trial of 30 seconds long.

leader and in follower configuration. We explained the mathematics behind this architecture presenting the Haken-Kelso-Bunz oscillator used as inner dynamics and the nonlinear optimal control law. We analysed deeper such an architecture and noticed that failed on the generation of a good individual motor signature. Indeed, few instances of velocity profiles recorded offline cannot be used to model the complexity of the human natural motor behaviour, and so neither to model the human identity of the virtual agent.

To overcome this lack, we proposed the use of a stochastic Markov model as signature generator trained on a human player performing solo sessions of the mirror game. The Markov model approach was found to be able to build an internal mathematical model capturing the motor kinematic features that uniquely identify an individual. As last step, such a signature generator was embedded in the cognitive architecture designed for the virtual player replacing in this way the pre-recorded signature used previously. We validated the MC-based virtual agent making it play dyadic sessions with humans both in leader and in follower configuration. By means of several metrics, we showed that our improved architecture was able to provide desired kinematic characteristic and at the same time reach high level of coordination with the partner in both conditions.

The reference generator model developed in this chapter can also be used as a signal classifier. Specifically, using it in a "reversed" way allows to properly classify a motor signature as belonging to a specific player model stored in a models database. More details on this point can be found in Appendix A of this thesis.

CHAPTER 5

MACHINE LEARNING APPROACH TO DESIGN THE VIRTUAL AGENT

This Chapter is devoted to the design of a new architecture based on machine learning techniques capable of driving an artificial agent in a joint coordination task with humans. Specifically, in Section 5.1 we explain why another strategy is needed for the avatar replacing in this way the cognitive architecture proposed in Chapter 4 (mainly based on nonlinear control) and how machine learning can provide such a solution. To understand better the proposed approach, a brief background of the learning techniques used in this work, and in particular of reinforcement learning, is given in Section 5.2. We then move to Section 5.3, that is the core of this chapter, to present our own design of a new "smart" player capable to learn how to move in a human-like way observing humans performing the motor coordination task of interest. The proposed control approach is validated both in-silico and experimentally. Section 5.4 is entirely dedicated to the latter validation, in which we describe in details the setup used to carry out the experiments and present the results obtained when the AI-based agent is engaged in dyadic interactions with several humans. The main results are then summarised in Section 5.5, in which we wish to highlight the advantages and the innovation of our AI-based virtual agent. The content of this Chapter has been published in [118].

5.1 The need of a new smart agent

In Chapter 4 we proposed a stochastic cognitive architecture for the virtual player based on the use of Markov chain as signature generator. Such a signature generator was built on real human players performing solo sessions of the mirror game in order to capture their individual motor signatures. We showed that the designed virtual player was able to coordinate its motion with that of the human player it was trained upon both in leader and in follower configuration keeping

a desired kinematic characteristic given by the Markov chain.

Embedding the artificial IMS generator as a stochastic component in the architecture of the virtual player solved the problem of having pre-recorded human trajectories matching the desired properties of the reference IMS. Nevertheless, the virtual player's behaviour still depends on the deterministic nature of the controller and the model often leading to unnatural tracking behaviour. In [206, 207], it has been proved that in order to have such a virtual player behaving like a human when engaged in dyadic trials of the mirror game, a careful tuning of the parameters characterising the inner dynamics and the control law is required and many trial-and-error attempts have to be performed to find a proper trade-off between them. Furthermore, it is worthy to underline that such a "tuning" approach is required every time the virtual player has to "replicate" the kinematic features of the human player while interacting with different partners.

To overcome the mentioned issue and avoid the burden of the tuning, here we propose a completely different approach based on machine learning techniques. By means of reinforcement learning we will be able to design a new artificial player, called Cyber-Player (CP), completely data-driven that does not need neither any explicit mathematical model nor any off-line tuning parameters. Indeed, we will see that the cyber player can learn directly observing players while performing the motor coordination task with others and mimic their motor features.

To better understand the design of the CP, some key concepts about the reinforcement learning approach are given in the next section.

5.2 Introduction to reinforcement learning

Reinforcement Learning (RL) belongs to the wider area of machine learning, and is largely used in robotics, gaming and navigation for its flexibility and adaptability to many problems. The idea behind it takes inspiration from the natural world. As a child learns how to behave in several situations by experience (trying to do something and seeing what happens), in the same way the artificial agent adopts a trial and error approach to deal with an unknown world and learn the best action that it can do in any situations.

The core of the reinforcement learning algorithm is made up of:

- the *agent*: it is any artificial entity (for example a robot, avatar or a generic software algorithm) able to perceive information, elaborate them, make decisions and carry out actions accordingly;
- the *environment*: anything that is external to the agent belongs to the environment;
- the *reward*: it is a quantity that somehow "measures" how good taking an action has been in a specific situation.

The agent continuously interacts with the environment, making actions that cause changes to the environment. The response from the environment is then perceived by the agent, elaborated

in terms of reward (or punishment) according to if that action was good or not, and finally translated in knowledge on the explored environment (also called experience) in terms of what is better to do in a specific situation. The experience obtained influences the process of decision making, the more experience the agent has the better it behaves. Goal of the reinforcement learning algorithm is to find an optimal strategy (decision making) that maximises the sum of all received rewards.

5.2.1 Finite Markov decision process

Reinforcement learning [166, 186] can be formalised mathematically using the finite Markov Decision Process (MDP) [202]. A MDP is a discrete time stochastic control process, suitable for modelling decision making in situations where the outcomes are unknown a priori. It is characterised by a quadruple $\langle \mathbb{X}, \mathbb{U}, p, r \rangle$ where:

- $\mathbb{X}$ is a finite set of states, in which the environment (or process or system) can be (also termed as state-space);
- $\mathbb{U}$ is a finite set of actions that the agent can take to influence the surrounding environment (also termed as action-space);
- the function $p : \mathbb{X} \times \mathbb{X} \times \mathbb{U} \to [0,1]$ is called *state-transition probability* and describes the probability to reach the state j from the state i choosing the action u;
- the function $r : \mathbb{X} \times \mathbb{U} \to \mathbb{R}$ computes the expected reward for any state-action pair.

To use the MDP as mathematical tool, we assume that the environment can be described through a Markov process and so satisfy the Markov property of limited horizon assumption (already defined in Section 4.1.3).

Denoting by r_k the immediate expected reward received at time instant k, the agent's goal is to maximise the sum of the rewards r_k for all time steps k, so as to maximise the so called *expected cumulative reward*, denoted with R. In general the learning process might take a very long time, making the sum of all rewards infinitely high. To deal with this problem, the concept of discounting is introduced. This approach means that the future rewards that are more than one time step away from the current one are not fully considered but are scaled by a constant (discount factor). The farther the reward is in the future, the higher is the applied discount factor. The expected cumulative reward at time step k can be evaluated as follows:

$$R_k = r_{k+1} + \gamma r_{k+2} + \gamma^2 r_{k+3} + \cdots = \sum_{m=0}^{\infty} \gamma^m r_{k+m+1}, \tag{5.1}$$

where the discount factor $\gamma \in [0,1]$. If $\gamma < 1$, the series in (5.1) has a finite value as long as the reward sequence R_k is bounded, whereas if $\gamma \to 0$ the agent is considered "myopic" because it takes into account only the immediate reward, discarding the future.

The learning agent uses the reward signal to find an optimal strategy (or policy) that maximises it. Formally, a policy $\pi : \mathbb{X} \rightarrow \mathbb{U}$ is a mapping from any state to the probability of selecting each possible action. That policy changes as long as the agent learns by its experience until is reaches the optimal policy π^*.

Having a determined policy, it is possible to define the *action-state function* $q_\pi(\mathbf{x}, u)$ as the expected return evaluated taking the action $u \in \mathbb{U}$ in the system's state $\mathbf{x} \in \mathbb{X}$ and then following the policy π. Specifically for each time instant k we have:

$$q_\pi(\mathbf{x}, u) = \mathbb{E}_\pi[R_k | \mathbf{x}_k = i, u_k] = \mathbb{E}_\pi \left[\sum_{m=0}^{\infty} \gamma^m r_{k+m+1} | \mathbf{x}_k = i, u_k \right], \tag{5.2}$$

where $\mathbf{x}_k = i$ specifies that the system is in the state i at time k, γ is the discount factor and $\mathbb{E}$ represents the mathematical expected value.

Given (5.2), the goal of reinforcement learning can be rephrased by saying that the optimal policy π^* has to maximise the state-action function in order to get the optimal value $q^*(\mathbf{x}, u)$ for each pair $(\mathbf{x}, u)$. The optimal state-action function can be reformulated in a recursive way obtaining the optimal Bellman equation [112, 186]:

$$q^*(\mathbf{x}, u) = \mathbb{E}\left[r_{k+1} + \gamma \max_u q^*(\mathbf{x}_{k+1}, u_{k+1}) | \mathbf{x}_k = i, u_k \right]. \tag{5.3}$$

We can conclude that mathematically solving a reinforcement learning problem means to find a policy that maximises the Bellman equation for each starting state and for each action. This results in a system of $N_\mathbf{x} \times N_u$ recursive equations where $N_\mathbf{x}$ is the total number of possible states and N_u the number of the actions. Explicitly solving this system of equations could be a cumbersome problem, or even intractable especially if the model of the environment is unknown. Several numerical and iterative approaches have been proposed to solve a Markov decision process, such as dynamic programming, Monte-Carlo methods and temporal difference learning. In our work we used the Q-learning, one of the most famous algorithm belonging to the temporal difference learning algorithms, described below, because of its double advantage to work online and to not require any knowledge of the Markov process.

5.2.2 Q-learning

The Q-learning strategy, proposed by Watkins in 1989, is defined as an off-policy temporal difference control algorithm that iteratively tries to find an approximation q of the optimal action-value function q^* [186, 199]. It is called "off-policy" because the update of q does not depend directly on the control policy chosen, simplifying in this way the analysis and the convergence proof of the algorithm. However, all that is required for the correct convergence is that all state-action pairs continue to be updated according to the following rule:

$$q_{k+1}(\mathbf{x}_k, u_k) = q_k(\mathbf{x}_k, u_k) + \alpha \left[r_{k+1} + \gamma \max_{u_{k+1} \in U} q_k(\mathbf{x}_{k+1}, u_{k+1}) - q_k(\mathbf{x}_k, u_k) \right] \tag{5.4}$$

where $(\mathbf{x}_k, u_k)$ is the state-action pair at time instant k, $q_{k+1}(\mathbf{x}_k, u_k)$ is the function approximation at time $k+1$, r_{k+1} is the immediate reward expected at time $k+1$, α is the learning rate and γ is the discount factor.

Since the convergence of the Q-learning algorithm is independent of the control policy, the ϵ-greedy policy can be chosen [186]. Specifically, at each interaction the ϵ-greedy policy takes the best known action with $(1-\epsilon)$ probability (exploitation), whereas with ϵ probability it takes a random action (exploration). The value of ϵ follows a monotonic decreasing function (e.g., $1/t$), since as time increases the exploration phase is replaced by the exploitation phase.

Summarising, we can identify a sequence of steps in which the agent interacts with the environment until that the function q converges to the optimal q^*:

- the agent looks at the environment (variable states of interest) and takes an action u in the set of all possible actions $\mathbb{U}$ according to its control policy π. This action causes the transition to a new state within the environment;
- the agent observes the new state $\mathbf{x}$ of the environment from the set of all possible states $\mathbb{X}$ in which the environment can be;
- the agent receives a reward (or penalty) r for the action taken in the previous step;
- according to the received reward, the agent updates the estimation of the q function following (5.4);
- after that, the previous steps are repeated until the end of the learning phase.

Next we describe in details how each of these steps was performed to design the cyber player able to interact with a human partner.

5.3 Control synthesis of the Cyber player

In Section 5.1 we highlighted the importance of having a cyber player that learns directly by observing human players engaged in coordination tasks. The aim of this section is to develop an artificial agent (cyber player), following the reinforcement learning approach, able to learn the way in which a specific human player (or target player) would move while playing a dyadic session of the mirror game both as follower and as leader. Note that the cyber player will be not trained to be a perfect follower (or leader) focusing on the minimisation of the tracking error, but will mimic the target human player with all his/her motor features (delay, reaction time, level of synchronisation, and so on).

5.3.1 Training

The training phase is extremely important in the reinforcement learning approach. As the goal of the RL algorithm is to make the CP able to play the game while mimicking the kinematic features

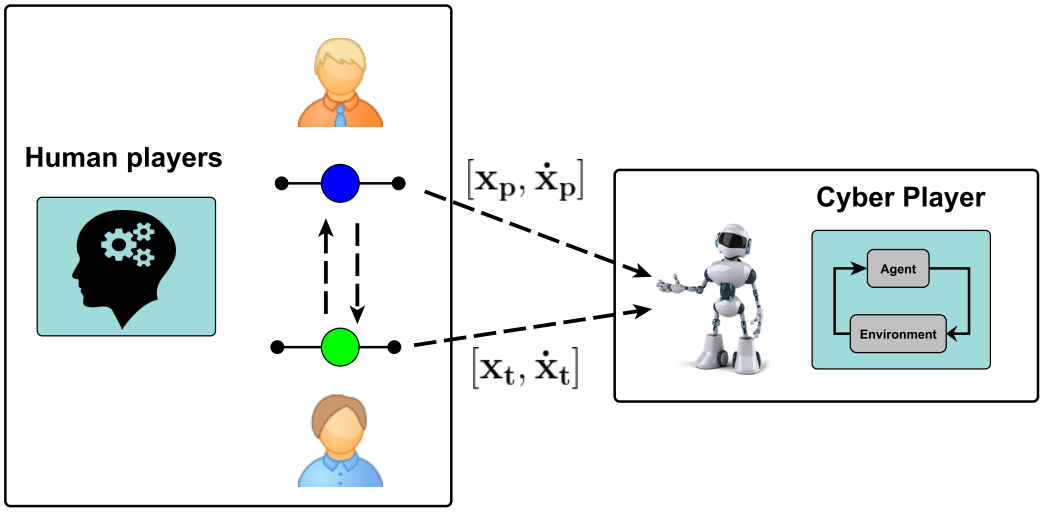


FIGURE 5.1. The Cyber player is trained observing directly two human players playing sessions of the mirror game in a follower-leader configuration. $[(x_t, \dot{x}_t)]$ is made of position and velocity of the target human player (green sphere), while $[(x_p, \dot{x}_p)]$ is made of position and velocity of the partner human player (blue sphere).

(IMS) of a target human player, in the training phase position and velocity time series of such a human player should be ideally collected during several live sessions of the mirror game (Figure 5.1). Such collected data then should be used to feed the cyber player and so, simultaneously to update the estimation of the function Q using (5.4). As learning typically requires a large dataset, real data from human players that allow the cyber player to converge towards a viable control solution might be difficult to collect.

To overcome this problem, here we propose a practical way of training the CP by using synthetic data generated by two "virtual trainers" (VT) playing the mirror game against each other. Each virtual trainer is nothing more than the nonlinear virtual player driven by the optimal architecture embedding as artificial signature generator block the IMS Generator synthesised as described in Chapter 4. In this dyadic set-up, the behaviour of two human players is mimicked by two virtual trainers parameterized and endowed with Markov chains built on human data in order to exhibit the desired human-like behaviour (Figure 5.2).

Note that to synthesise the IMS Generator only a small dataset obtained during sessions of the game where the HP plays in solo condition is needed. By embedding the IMS Generator in a model-based VT we can then generate much larger synthetic datasets from VTs playing dyadic session of the game that can be used to train the CP.

Spelling out the learning process described in Section 5.2.2 to our specific case, we have to define:

- the **state space** $\mathbf{x} := [x, \dot{x}, x_p, \dot{x}_p]$ where $[x, \dot{x}]$ are position and velocity of the CP while

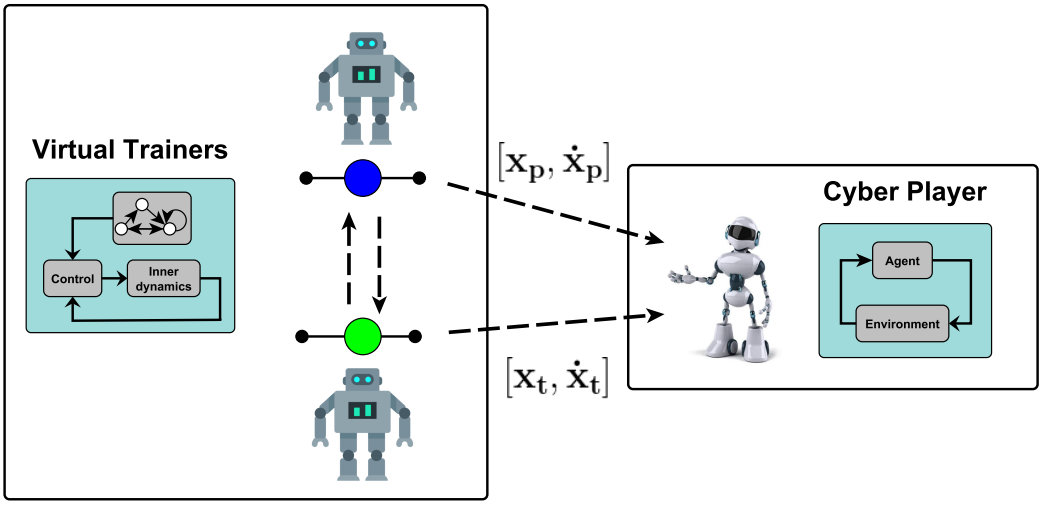


FIGURE 5.2. Synthetic data are generated running dyadic sessions of the mirror game between two virtual trainer, one in leader and the other in follower configuration. Each virtual trainer is driven by the control architecture with embedded the Markov chain which captures the individual kinematic feature of the HP to be replaced. While the two VTs are playing together, the values of position (x_t, x_p) and velocity $(\dot{x}_t, \dot{x}_p)$ are given in input to the reinforcement learning algorithm in order to train the cyber player. The vector $[(x_t, \dot{x}_t)]$ is made of position and velocity of the target VT (green sphere), while $[(x_p, \dot{x}_p)]$ is made of position and velocity of the partner VT (blue sphere).

$[x_p, \dot{x}_p]$ are position and velocity acquired from the partner player. Since we work with a finite Markov decision process, the state space has to be properly discretised. In particular the position values are in the range $[-0.5, 0.5]$ sampled with a step size of 0.2, whereas the velocity values span in $[-1, 1]$ with a step size of 0.4;

- the **action space** is the set of all possible control input u, and so acceleration values, that can be imparted to the end effector of the CP. Such acceleration spans from negative to positive values in order to move from left to right and viceversa. To estimate the action space and the range of acceleration values to be discretised, experiments were performed making two human players performing the task. Analysing the data, we found that the maximum value of acceleration reached by humans is $|u_{max}| = 4$. That range was then discretised using 9 different values of control input, heuristically chosen as trade-off between the learning time and the quality of discretisation;
- the **policy** π is an ϵ-greedy policy, already explained in Section 5.2.2;
- the **reward function** is defined as

$$r := -(x - x_t)^2 - 0.1(\dot{x} - \dot{x}_t)^2 - \eta u^2, \tag{5.5}$$

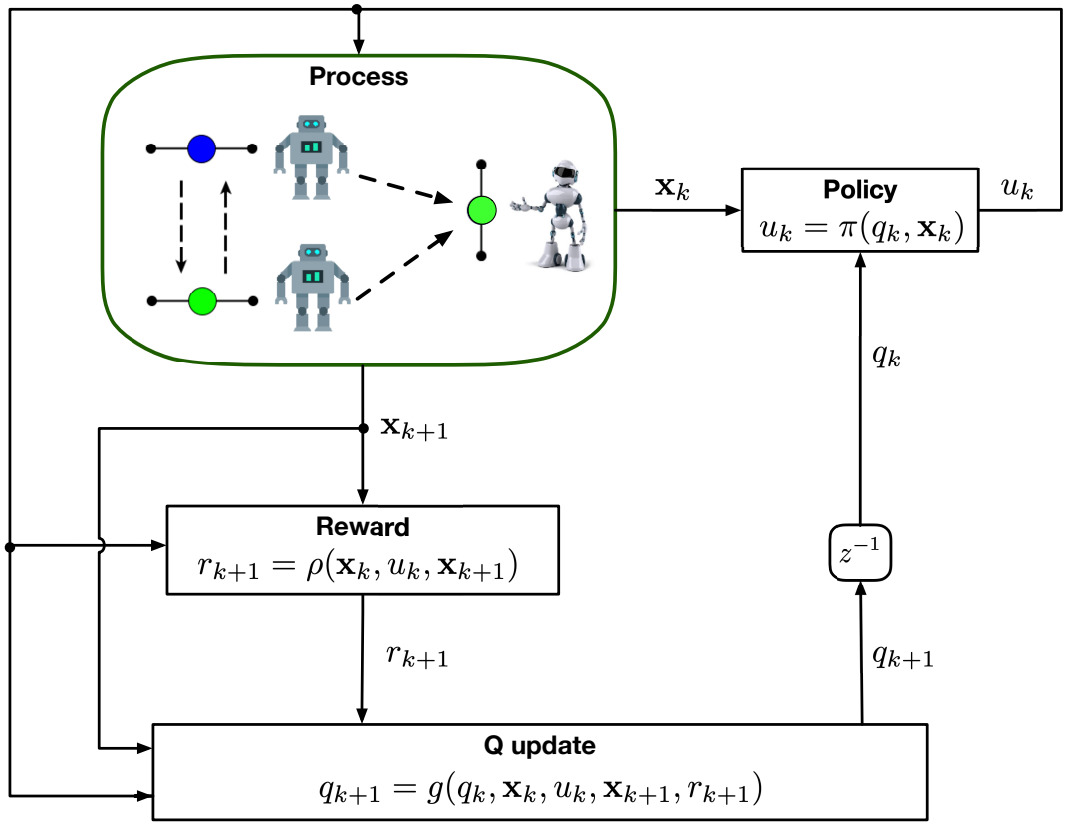


FIGURE 5.3. Learning architecture to train a cyber player to play dyadic sessions of the mirror game. The game is simulated making two MC-based virtual trainer playing the mirror game, one in leader and the other in follower configuration. The Q-learning controller chooses a control input u according to the Q tables and the process state. The CP and the process evolve in a new state and the CP receives an instant reward. The latter is used then to update iteratively the Q-table updating consequently also the control law (policy). In the depicted scenario, the CP is learning to emulate the VT with the green handle.

where $[x_t, \dot{x}_t]$ are position and velocity of the target player to emulate (while playing with the partner player), η is a constant weight for the minimisation of the control energy u^2. Such a reward function aims at minimising the error in position and in velocity between the CP and the target player, such that the CP mimics its behaviour;

- the learning rate and the discount factor are set both to 0.8.

To implement the Q-learning algorithm, as a first step we define a matrix $Q \coloneqq [q_{ij}]$, also referred to as the Q-table, where the states are listed on the rows and the actions on the columns. Each element q_{ij} of the matrix represents the "value" of the action j taken while being in the state i, such a value is also termed the q-value of the state-action pair. At the beginning of the learning process the matrix Q is initialised with random values and then, modified at each

iteration until convergence. Figure 5.3 depicts the reinforcement learning algorithm. Specifically at each iteration k, the CP observes position and velocity of both players and takes an action u_k according to its policy rule π, that is an ϵ-greedy policy. After the action, the CP and the game evolve to a new state $\mathbf{x}_{k+1}$ and the CP receives the reward $r_{k+1} = \rho(\mathbf{x}_k, u_k, \mathbf{x}_{k+1})$ regarding the action taken at the previous time k. According to the reward received, the CP updates the value of the matrix Q corresponding to the pair $(\mathbf{x}_k, u_k)$ following the rule (5.4).

5.3.2 In-silico validation

We trained the cyber player by generating synthetic data of dyadic sessions of the mirror game between two VTs in leader-follower configuration. Each virtual trainer was controlled using the MC embedded cognitive architecture described in Chapter 4. Being VTs characterised by a nonlinear dynamical system and an optimal control law, the dynamics and control parameters were heuristically selected in order to replicate the kinematic features of the corresponding human player when engaged in a leader-follower session. Furthermore, the Markov chains to be embedded in the VTs were built on the data collected from the human players as explained in Chapter 4.

The in-silico validation was carried out in two different phases:

1. in this initial phase the cyber player was trained to emulate the target VT while always playing with the same partner player;

2. the cyber player was trained using a set of partner VTs, each one characterised by an own Markov chains, and validated with a further VT that did not belong to the training set. In this way, the CP was trained to be general enough in order to emulate the target VT independently from the specific player it is engaged with.

Training with one partner VT. As test-bed a VT endowed with the Markov chain model of human player 1 was set as VT leader (VT_1), while the Markov chain model of player 5 is used for the VT follower (VT_5). The CP was trained to emulate the VT_5 acting as follower. Training was performed in real time while two VTs were playing against each other and took approximately 7 hours ($\approx$ 2000 trials 30 seconds long). Once the training was completed, the performance was evaluated comparing 20 trials 60 seconds long of the mirror game between the virtual leader (VL) and the cyber player with 20 trials of the same length between the same virtual leader and the virtual follower (VF) used during the training.

In what follows, the performance of the cyber player is evaluated in terms of relative phase, circular variance and RMS position error (metrics defined in Section 4.3.1). In Figure 5.4 the position time series and the relative phase signal are reported for a trial played between the virtual trainer leader and the virtual trainer follower [panel (a)] and the same virtual trainer leader with the cyber player [panel (b)]. In both cases the average of the signal phase is negative,

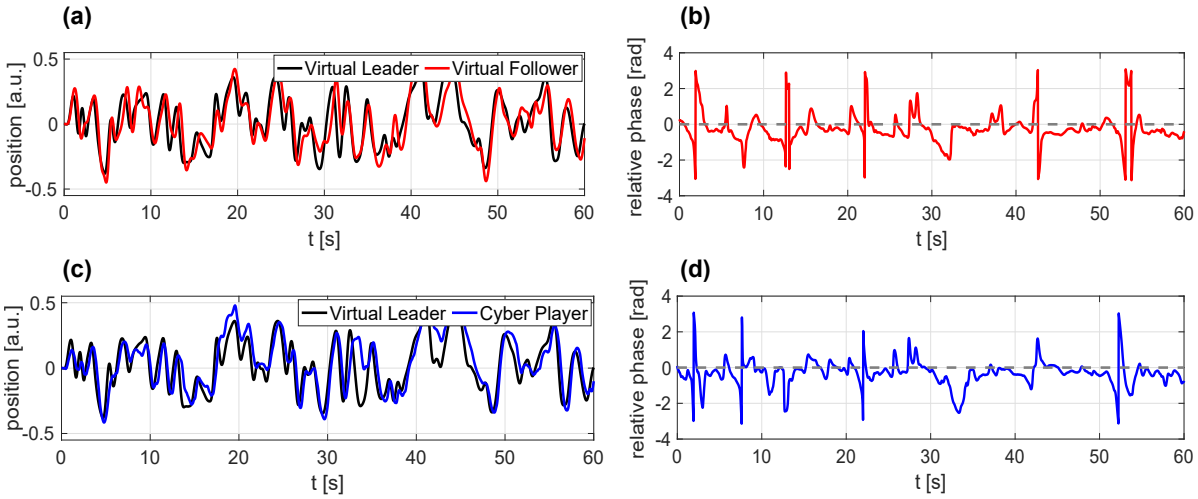


FIGURE 5.4. Position time series are reported for one-trial example between **(a)** the virtual trainer leader (in black) and the virtual trainer follower (in red) and between **(c)** the virtual trainer leader (in black) and the cyber player (in blue). The corresponding relative phase signal is drawn in **(b)** and **(d)**. A dashed grey line indicates when the relative phase between the two players is null (perfect synchronisation).

meaning that both the players (virtual and cyber) are playing effectively in follower mode. Specifically the relative phase is $\Delta\Phi = -0.29 \pm 0.1$ for the VT follower, whereas it is $\Delta\Phi = -0.42 \pm 0.1$ for the cyber player. Other metrics are the circular variance to quantify the level of coordination during the game with $CV(VL,VF) = 0.83 \pm 0.04$ and $CV(VL,CP) = 0.88 \pm 0.04$, and the root mean square of the position error with $RMSE(VL,VF) = 0.104 \pm 0.01$ and $RMSE(VL,CP) = 0.108 \pm 0.01$.

In Figure 5.5, the velocity PDF and the similarity space are depicted for the same set of 20 trials. In particular the distance between each pair of velocity PDFs is quantitatively described by $EMD(VL,VF) = 0.0024$ and $EMD(VL,CP) = 0.0033$, meaning that during the game the CP (in blue) adapted its own distribution to that of the leader (in black) as the VF (in red) (for details about the EMD and the similarity space read Section 4.2.2). In the panel (b), the characteristic regions of the VF (in red) and of CP (in blue) while playing with the VL (in black) are depicted in the similarity space. Mathematically we computed the overlap A_{ij} between the ellipse i and the ellipse j as the ratio of the area of the intersection and the total area of the two ellipses. In this way, a complete overlap between the two ellipses corresponds to $A_{ij} = 1$, whilst $A_{ij} = 0$ corresponds to a complete separation. The overlapping equal or greater than 0.6 for all pairs of players ($A_{VF,VL} = 0.74$, $A_{CP,VL} = 0.60$, $A_{CP,VF} = 0.62$) confirms the ability of the cyber player to track the leader exhibiting the same kinematic feature of the virtual follower.

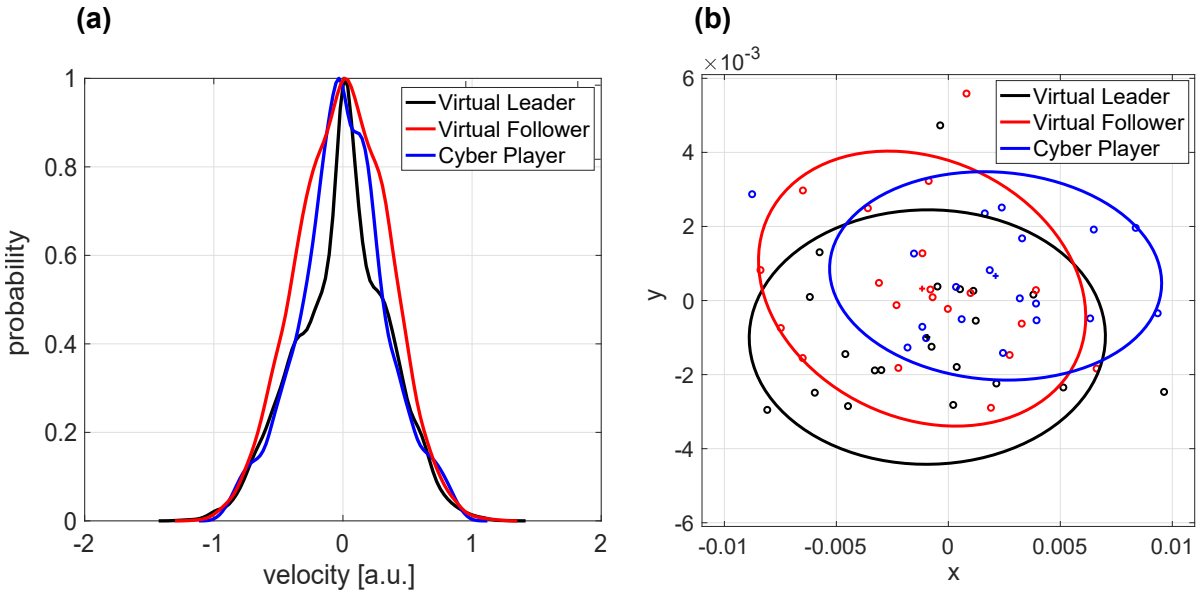


FIGURE 5.5. **(a)** Velocity probability density function and **(b)** similarity space built taking 20 sessions of the mirror game in leader-follower mode. Each point in the similarity space represents the velocity PDF obtained from a single trial. An ellipse encircles all the trial belonging to the same player. In black the virtual trainer leader (VL), in red the virtual trainer follower (VF) and in blue the cyber player (CP).

Training extended to several VT partners. So far we have shown that the CP is able to play the game with the virtual trainer leader which was also used for its training. In the following phase, we validate the CP using a set of virtual trainer leaders characterised by different Markov chains. Specifically we performed the training with different VT leaders based on the MCs built on the human players 1,2,3 and 4 (VT_1, VT_2, VT_3, VT_4). A further VT_6 leader (representing the human player 6) that does not belong to the training set was used to test the "generalisation ability" of the AI-based CP. In this case, the training took approximately 5 days (playing ≈ 24000 trials 30 seconds long of the mirror game).

The results are summarised in Figure 5.6, where the position trajectories of the VF (VT_5) and the CP are shown while performing a session of the mirror game with a VL used during the training (VT_2) [panel (a)] and a VL which was not used in the training set (VT_6) [panel (b)]. In both cases, the CP is able to track the trajectory of the leader showing a general enough skill to track any leader. To assess that the CP is able to play in both conditions of the mirror game (leader and follower), we trained the CP to emulate the virtual trainer 5 while acting as leader as well (players 1,2,3,4 and 6 were set coherently as followers). Position time series of a dyadic session between the CP and VL against the VF included in the training set are depicted in Fig. 5.6 [panel (c)] while those against the VF not included in the set are in [panel(d)].

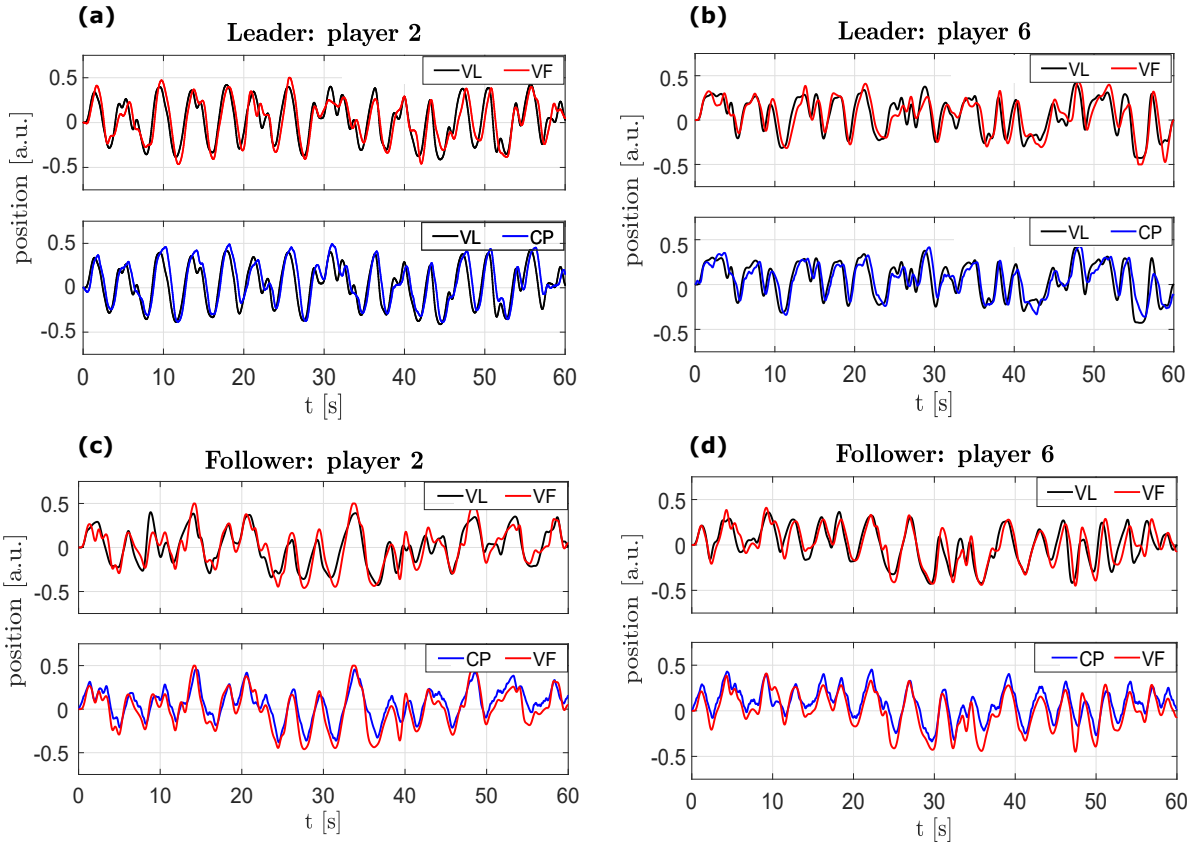


FIGURE 5.6. Position time series of a dyadic session with a virtual leader/follower included in the set of players used during the training, player 2 in the panel **(a)** and a virtual leader/follower not included in the set, player 6 in the panel **(b)**. The CP (in blue) plays both with the VF (in red) and with the VL (in black).

In Table 5.1 a quantitative analysis is provided, where EMD, CV and RMS of the position error between the CP and VLs (VFs) modelled on human players 1,2,3,4 and 6 are given and respectively compared to those obtained when the VF (VL) modelled on player 5 is used. The lack of any statistically significant difference between the CP and the VT it was trained upon confirms that the training phase has been successful.

5.4 Experimental validation

5.4.1 Experimental setup

Participants. A total of 5 people participated in the experiments: 4 females and 1 male. In particular 1 out of 5 participants was left handed. Participants were PhD students and Postdoctoral Researcher Associates from University of Bristol and University of Naples “Federico

	Following	**EMD**	**CV**	**RMS**
VT_1	VT_5	0.002	0.76 ± 0.06	0.11 ± 0.01
	CP	0.008	0.80 ± 0.05	0.11 ± 0.01
VT_2	VT_5	0.003	0.86 ± 0.06	0.11 ± 0.01
	CP	0.004	0.86 ± 0.06	0.12 ± 0.01
VT_3	VT_5	0.003	0.85 ± 0.04	0.12 ± 0.01
	CP	0.007	0.88 ± 0.04	0.12 ± 0.01
VT_4	VT_5	0.004	0.74 ± 0.07	0.10 ± 0.01
	CP	0.004	0.79 ± 0.07	0.10 ± 0.01
VT_6	VT_5	0.003	0.83 ± 0.05	0.11 ± 0.01
	CP	0.002	0.86 ± 0.04	0.11 ± 0.01

	Leading	**EMD**	**CV**	**RMS**
VT_1	VT_5	0.003	0.83 ± 0.04	0.11 ± 0.01
	CP	0.007	0.88 ± 0.03	0.12 ± 0.01
VT_2	VT_5	0.003	0.82 ± 0.03	0.12 ± 0.01
	CP	0.007	0.90 ± 0.03	0.11 ± 0.01
VT_3	VT_5	0.003	0.81 ± 0.05	0.12 ± 0.01
	CP	0.007	0.90 ± 0.03	0.12 ± 0.01
VT_4	VT_5	0.003	0.86 ± 0.05	0.10 ± 0.01
	CP	0.007	0.89 ± 0.03	0.11 ± 0.01
VT_6	VT_5	0.003	0.84 ± 0.04	0.10 ± 0.01
	CP	0.007	0.90 ± 0.03	0.11 ± 0.01

TABLE 5.1. Earth mover's distance (EMD), circular variance (CV) and root mean square of the position error (RMS) are reported for the CP/VT_5 while playing with each of the other players. Two different scenarios are considered: CP as follower and as leader.

II". The players were assigned numbers from 1 to 5 (HP_1, HP_2, HP_3, HP_4, HP_5).

This study was carried out according to the principle expressed in the Declaration of the Helsinki. All subjects gave written informed consent for both experiments participation and publication of identifying images.

Experimental task. 4 participant (HP_1, HP_2, HP_3, HP_4) were asked to perform 8 HP-HP trials of the mirror game each lasting 60 seconds in leader-follower condition with the 5-th participant (HP_5) acting as a follower.

HP_5 was then substituted with the CP (driven by the Q-learning algorithm). Once again, 8

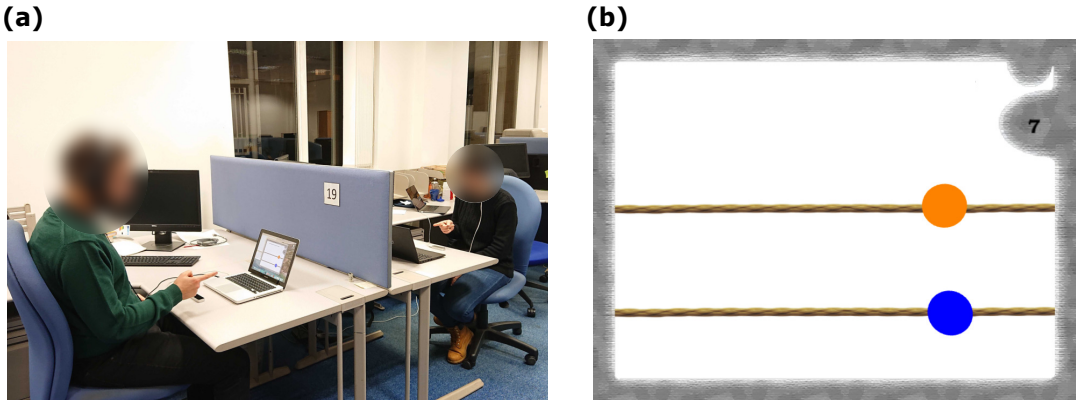


FIGURE 5.7. **(a)** Human performs a dyadic trial moving his/her index finger over a Leap motion controller. **(b)** User interface seen by the human player. Blue circle represents his/her motion, while orange circle that of the partner (human or virtual).

HP-CP trials lasting 60 seconds each were performed in the leader-follower condition for each pair. During the experimental session, the human partners were not informed about the change of the player 5. The given instruction to the players HP_1, HP_2, HP_3, HP_4 was to be the leader but still taking care to be followed by the partner player, while the instruction given to HP_5 was simply to follow the partner player.

Experimental platform. Experiments were performed through CHRONOS, home-made java platform recently developed to study movement coordination which was presented in [12, 115]. CHRONOS is a software platform running over a dedicated WLAN network. It allows to perform dyadic trials of the mirror game by making players see their own movement and that of the other player as moving objects visualised on the screen of their laptop, removing any social interaction through visual or auditory coupling between the participants (Figure 5.7).

Furthermore and of main interest for this work, CHRONOS allows to join one or more virtual agents in the game sessions. Since the joint task is performed through a computer setup, the participants are not aware whether the traces of the objects on their screen are driven by another player or by the virtual agent.

During the experiments, participants were asked to coordinate their motion waving the index finger of their own preferred hand from left to right (and viceversa) along a direction required to be as straight as possible. Leap motion position sensors [193] connected to the laptops were used to store the trajectories provided by the players during all the trials, respectively. Furthermore, each participant was required to wear headphones through which white noise was transmitted in order to eliminate possible auditory couplings with the partner.

To avoid delays in the communication between the laptops, a sampling frequency of 10 Hz was found to be low enough to guarantee good communication performance while acquiring a

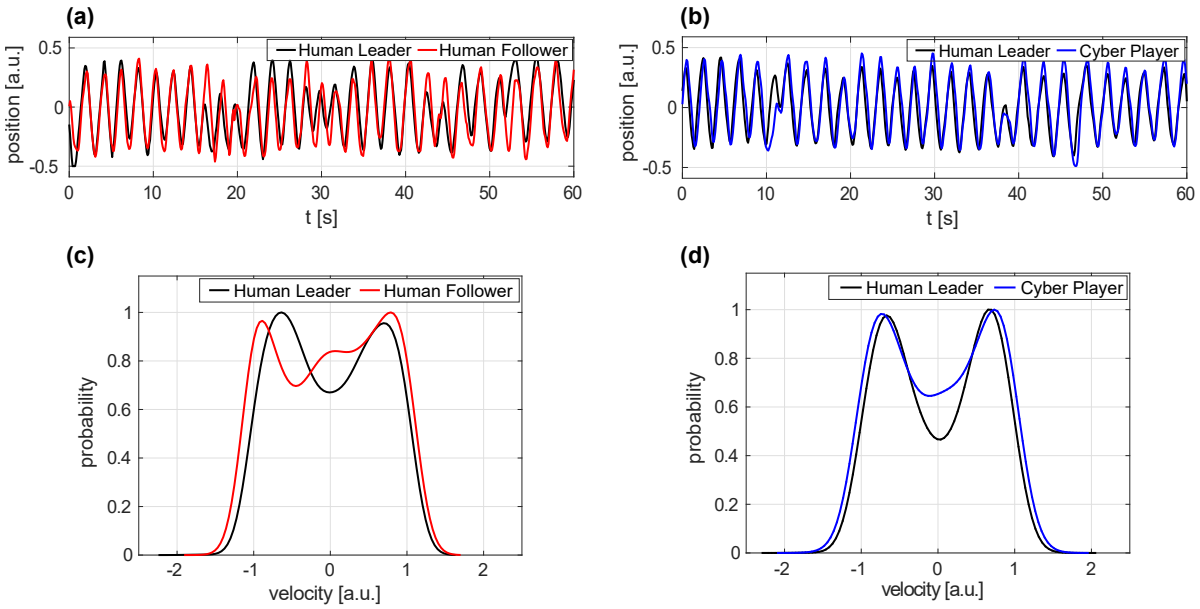


FIGURE 5.8. **(a)** Position time series and **(c)** velocity PDF of a leader-follower trial between the human leader (in black) and the human follower (in red). **(b)** Position time series and **(d)** velocity PDF of a leader-follower trial between the same human leader (in black) and the cyber player in follower mode (in blue). Each trial is 60 seconds long.

sufficiently rich time series to be analysed. An upsampling to 100 Hz was performed a posteriori before the analysis.

5.4.2 Validation

The metrics employed to analyse and quantify the participants' performance are those described in Section 4.2.2 and in Section 4.3.1.

In Figure 6.9 a representative position time series and IMS of a mirror game session where a human leader and a human follower play against each other is compared with the time series and IMS acquired during a session where the CP is made to play the game as a follower against the same human leader.

Quantitatively, for each pair of players we evaluated the distance between the corresponding PDFs by means of EMD. Other indices considered in the comparison are related to the level of synchronisation measured in terms of circular variance (CV) and RMS. Results are reported in Table 5.2, where the mean and the standard deviation are reported over the total number of trials for each metric and for each pair of players. Furthermore a paired T-test was performed at 95% confidence level with $N-1$ degrees of freedom. A p-value>0.05 was computed for each pair of players showing that there is no significant difference between the behaviour of the cyber

Human Leader	Metric	Human Follower	Cyber Player	Paired T-test
HP_1	EMD	0.01	0.01	
	CV	0.9 ± 0.05	0.9 ± 0.06	$t(7) = -0.081$, $p = 0.938$
	RMS	0.18 ± 0.03	0.14 ± 0.06	$t(7) = 1.562$, $p = 0.162$
HP_2	EMD	0.007	0.006	
	CV	0.84 ± 0.07	0.89 ± 0.05	$t(7) = -2.105$, $p = 0.073$
	RMS	0.15 ± 0.04	0.15 ± 0.05	$t(7) = -0.347$, $p = 0.739$
HP_3	EMD	0.006	0.005	
	CV	0.92 ± 0.01	0.93 ± 0.06	$t(7) = -0.496$, $p = 0.635$
	RMS	0.17 ± 0.03	0.15 ± 0.05	$t(7) = 1.25$, $p = 0.251$
HP_4	EMD	0.009	0.004	
	CV	0.86 ± 0.06	0.87 ± 0.03	$t(7) = 0.933$, $p = 0.382$
	RMS	0.16 ± 0.03	0.12 ± 0.02	$t(7) = 2.340$, $p = 0.052$

TABLE 5.2. Earth mover's distance (EMD), circular variance (CV) and root mean square of the position error (RMS) are reported for each pair of players. The human players are numbered from 1 to 5, player 5 is the human follower while players $1-4$ are human leaders, CP is the cyber player. For each pairs paired T-test with 7 degrees of freedom is reported both for CV and RMS metric.

player when playing the game and that of the human player whose Markov Chain model was used by the VT during the training stage. This implies that the CP has learnt successfully to adapt its IMS and its behaviour to that of the HP as desired.

The Paired-T test was performed to assess the absence of any statistical difference between the performance of the CP and the human follower. Another interesting analysis is performing Repeated Measures ANOVA where the experiments are repeated several times in different consecutive days. This improves the precision of the experimental validation by reducing the size of the error variance in the statistical test.

5.5 Summary

In this chapter we addressed the problem of synthesising an autonomous artificial agent able to coordinate its movements and perform a joint motor task in dyadic interactions. Specifically, we showed how to design and train such an agent through the reinforcement learning approach in order to emulate the way of behaving of a target player used for the training. The introduction of a new artificial agent, called cyber player, was motivated by avoiding the explicit mathematical model that could be very complicated and requiring a fine tuning of its parameters. In case of dyadic interaction, we found that the Q-learning algorithm was effective at solving the problem of achieving motor coordination between the CP and the partner player.

As all the methods belonging to the machine learning, the main drawback is that they require a dataset large enough to find all the correspondences and build an internal model. When this is not available we proposed that model-based artificial agents (or Virtual Trainers) endowing the IMS generator derived in Chapter 4 can be used during the training to generate as much synthetic data as needed for the Q-learning algorithm to converge. We wish to highlight that the use of such virtual trainers has a dual advantage:

1. the synthesis of the IMS Generator requires only a small dataset obtained making humans play sessions of the mirror game in solo condition (in our study we collected only 30 trials of 30 seconds each). By embedding such an IMS Generator in the model-based VT, we can exploit it to generate much larger synthetic datasets from VTs playing dyadic session of the game that can be used to train the CP;

2. by tuning few virtual trainers in order to make them emulate the behaviour of the corresponding human players (in our study we set 5 VTs), we can train the cyber player to be general enough to play the mirror game also with other players that were not used for its training without requiring further lengthy offline parameters tuning.

In the next chapter, we will address the problem of extending the cyber player to group scenarios. Specifically, we will use a multi-player version of the mirror game as coordination task to design a cyber player able to join such a group. The main drawback for this problem is that the state space to be explored by the learning algorithm may become quickly too large to be dealt with by tabular methods such as the Q-learning algorithm. A possible solution is the use of different learning strategies such as deep reinforcement learning, which can be more apt to tackle the challenging case of multi-player task.

CHAPTER 6

MACHINE LEARNING APPROACH TO DESIGN THE VIRTUAL AGENT: FROM DYADS TO GROUPS

So far the problem of designing an artificial agent able to coordinate itself with a human in dyadic sessions of the mirror game has been tackled. In this chapter, we move to another open question, that is of investigating how an artificial agent can be programmed to integrate its movement with that of others in a group scenario. Moving from dyads to groups is a much more complex and ambitious open problem to address, due to the lack of studies existing in the literature on group coordination (as discussed in Chapter 2).

With the aid of graph theory, the aim of this chapter is to present a control architecture based on learning techniques to drive artificial agent to perform a joint task in a multi-agent scenario. We tackle this problem by considering a multi-player version of the mirror game as scenario of interest, firstly proposed in [10], where a group of participants is asked to oscillate a finger sideways synchronising with the other group members. Extending the Q-learning approach to multi-agent systems is cumbersome as the approach becomes computationally expensive with the growth of the system state due to the addition of other players. To overcome this limitation, in what follows we use the deep reinforcement learning [113, 124, 186], combining the reinforcement learning strategy with the powerful generalisation capabilities of neural networks.

Before discussing the design of the cyber player in a group, some fundamental concepts about graph theory and deep reinforcement learning are given in Section 6.1 and Section 6.2. Then, in Section 6.3 we propose our control strategy capable to drive the cyber player in group joint oscillatory task, followed by in-silico validation. To further validate the cyber player we performed our own experiments involving humans in group interactions. Specifically, in Section 6.4 the results of the experimental validation are given showing the effectiveness of the cyber player trained upon a target group member. Finally, a summary is given in Section 6.5. The content of

this chapter has been partially published in [117].

6.1 Elements of graph theory

A graph [201] is a tuple $G = (V, E)$ defined by a set of *nodes* $V = \{1, ..., N\}$ and a set of *edges* E. A graph is said *undirected* if the edge $(i, j) \in E \iff (j, i) \in E$. In an undirected graph, two nodes i and j are said to be *neighbours* or *adjacent* if a *link* or *edge* $(i, j) \in E$ exists.
We call *adjacent matrix* the matrix $A = \{a_{ij}\} \in R^{NxN}$, where

$$a_{ij} = \begin{cases} > 0, & \text{if } (i, j) \text{ are neighbours} \\ = 0, & \text{otherwise} \end{cases} \tag{6.1}$$

A *weight* w can be associated to each edge of the graph and it is an indicator of the strength of interaction between the two nodes linked by the considered edge. In particular, a graph is said *unweighted* if all edges of the matrix have the weight equals to 1. A graph is *connected* if there exists a path between any two of its nodes.

The configuration of all edges interconnecting the nodes in the graph determines the graph topology. In this chapter we consider the following well known topologies:

- Complete graph: each node is connected to all the others;
- Ring graph: each node is connected to only two others, called neighbours. In case of networks with more than three nodes, any pair of two neighbouring nodes cannot share their neighbours;
- Path graph: similar to the ring graph, but two nodes (qualified as external) are connected only with one neighbour (not the same);
- Star graph: one node (qualified as central) is connected to all the others (qualified as peripheral), which in turn are supposed to be connected only to the central one.

6.2 Deep reinforcement learning

As explained in Chapter 5, solving a problem with the classical Q-learning approach [199] means to iteratively explore all possible combinations between the set $\mathbb{X}$ of possible states and the set $\mathbb{U}$ of possible actions in order to evaluate them in terms of action-value function (or Q-function) in a tabular form. As this is too computationally expensive in our group scenario, we use the deep learning control approach shown in Figure 6.1. Specifically:

- the **state space** is defined as $\mathbf{x} := \left[x, \dot{x}, \bar{x}, \dot{\bar{x}}\right]$ where $[x, \dot{x}]$ are position and velocity of the cyber player (CP), while $\left[\bar{x}, \dot{\bar{x}}\right]$ are mean position and mean velocity of the players connected to the target player (TP);

- the **action space** is made of 9 different values of acceleration in the range $[-4,4]$, which was empirically chosen looking at typical human accelerations in the same conditions;
- the **reward function** is selected as $\rho := -(x-x_t)^2 - 0.1(\dot{x}-\dot{x}_t)^2 - \eta u^2$ where $[x_t, \dot{x}_t]$ are position and velocity of the target player, u is the control action to be performed, while the constant parameter η tunes the control effort;
- the **policy** π according to which the CP chooses the action to take in a specific state is an ϵ-greedy policy as in [186].

As a deep learning technique, in particular we exploit the Deep Q-network (DQN) strategy. The motivation behind such a choice lies in the fact that the DQN is somehow the "closest" extension of the Q-learning approach we adopted in Chapter 5 to large state-action spaces such as the one studied here. More specifically, an artificial neural network (ANN) is used to approximate the optimal action-value function Q^* defined as:

$$Q^*(\mathbf{x},u) = \max_{\pi} \mathbb{E}\left[\sum_{m=0}^{\infty} \gamma^m r_{k+m+1} | \mathbf{x}_k = \mathbf{x}, u\right], \tag{6.2}$$

which maximises the expected value of the sum of the rewards r discounted by a positive factor $\gamma < 1$, obtained taking the action u in the state $\mathbf{x}$ following the policy π at any time instant k.

Training an ANN in order to approximate a desired function (in our case the Q-function) means finding the vector of network weights θ of the connections between the neurons, iteratively evaluated by back-propagation algorithms in order to minimise a loss function. The loss function is used to measure the error between the actual and the predicted output of the neural network (e.g., mean squared error) (see [124] for further details).

Contrarily to what is done in traditional supervised learning with ANN where the predicted output is well defined, in the deep Q-network approach the loss function is iteratively changed because the predicted output itself depends on the network parameters θ_k at every instant k. Namely, the loss function is chosen as:

$$L_k(\theta_k) = \mathbb{E}\left[\left(r_k + \gamma \max_{u_{k+1}} Q(\mathbf{x}_{k+1}, u_{k+1}, \theta_{k-1}) - Q(\mathbf{x}_k, u_k, \theta_k)\right)^2\right], \tag{6.3}$$

which represents the mean squared error between the current estimated Q-function and the approximate optimal action-value function.

The neural network we considered to approximate the action-value function Q in (6.2) is designed as a feedforward neural network with (Figure 6.1):

- an *input layer* with 4 different nodes, one for each state variable $\left[x, \dot{x}, \bar{x}, \dot{\bar{x}}\right]$;
- *two hidden layers*, empirically found, made of 64 and 32 nodes each implementing a sigmoidal activation function;

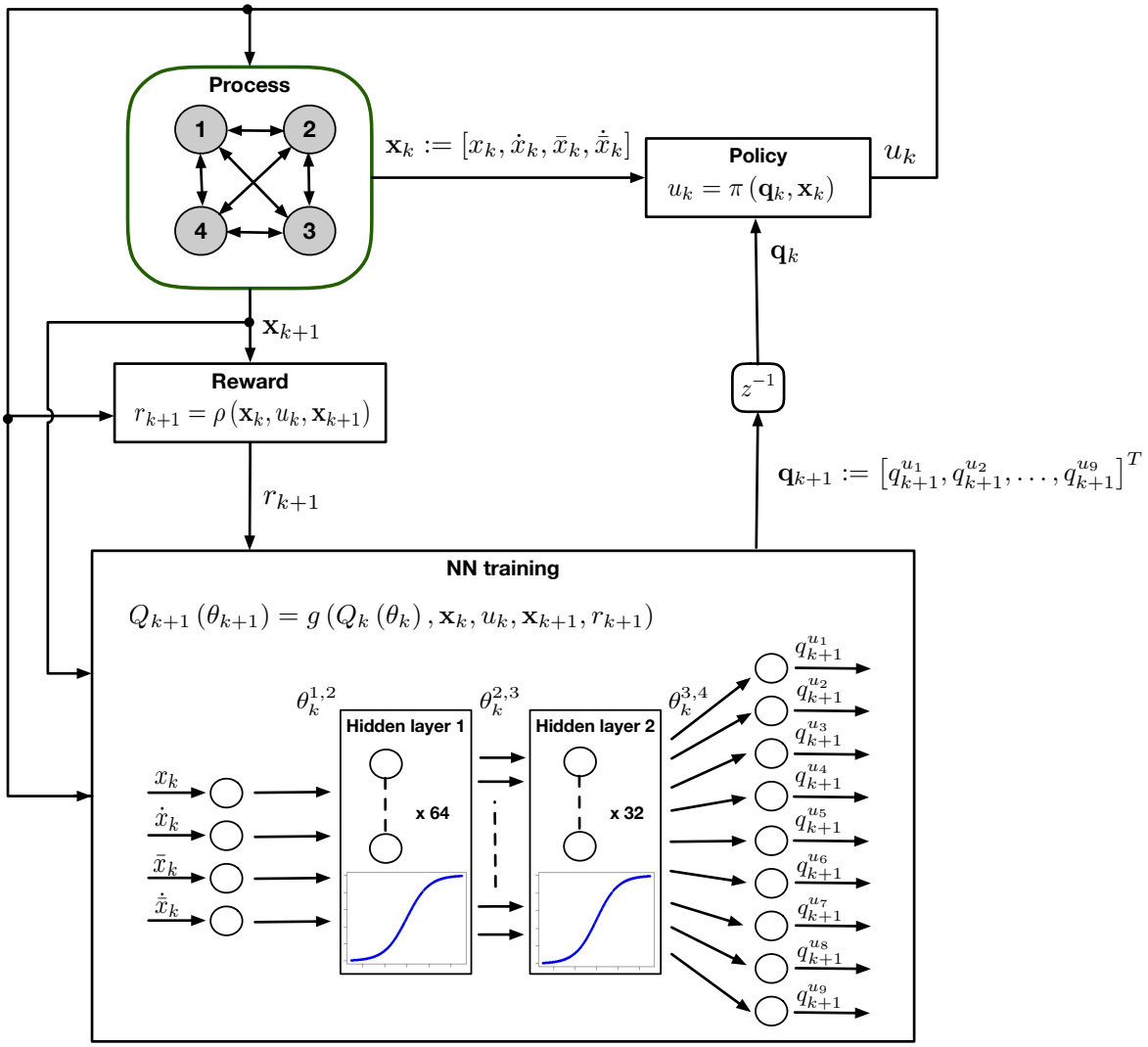


FIGURE 6.1. Block scheme of the deep Q network algorithm. At each iteration, the RL controller chooses the control u according to the current neural network and the system's state. The process evolves in a new state generating the reward r. The reward, previous and current state are then used as a new sample to train the neural network.

- an *output layer* with 9 different nodes, one for each action available in the set $\mathbb{U}$. In the DQN, the network output returns the estimated action-value q^u for each possible action $u \in \mathbb{U}$ in a single shot reducing in this way the time needed for the training. Then, the action corresponding to the maximum q-value (neuron's output) is chosen as the next control input.

The specific structure of the NN adopted (number of levels and number of nodes) is the result

of several tuning trials where other possible configurations have been explored.

Reinforcement learning is known to be unstable or even to diverge when a nonlinear function approximator, such as an ANN, is used to estimate the Q-function [124, 186]. Such an instability is caused by: *i)* the presence of correlation in the sequence of observed states and *ii)* the presence of correlation between the current estimated Q and the target network, resulting in the loss of the Markov property. To solve this problem, the correlation in the observation sequence is removed by introducing the *experience replay mechanism*, where the observed states used to train the ANN are not taken sequentially but are sampled randomly from a circular buffer [124]. Also, the correlation between the current estimate of the function Q and the target optimal one Q^* is reduced updating the latter at a slower rate instead of at each iteration.

To implement the DQN as describe above, a feedforward neural network needs to be initialized with random values and the experience replay mechanism is implemented instantiating an empty circular buffer in order to store the system's state at each iteration. Then at each iteration:

1. the CP observes the process' state $\mathbf{x}_k$ at time instant k and performs an action u_k according to the policy π, that is an ϵ-greedy policy;

2. the process evolves into a new state $\mathbf{x}_{k+1}$ and the CP receives the reward r_{k+1} that measures how good taking the action u in the state $\mathbf{x}_k$ has been;

3. the new sample $\langle \mathbf{x}_k, u_k, \mathbf{x}_{k+1}, r_{k+1} \rangle$ is added to the circular buffer and a random batch taken from it is used to train the NN. The training is done through the gradient descend back-propagation algorithm with momentum [186] so as to tune the network's weights θ in order to minimize the loss function (6.3). We denote the network's weights between the layer n and $n+1$ at instant k as $\theta_k^{n,n+1}$.

The steps above are repeated until convergence is achieved according to the "termination criterion":

$$\left\| RMS_{TP,\bar{x}} - RMS_{CP,\bar{x}} \right\| \leq \epsilon, \tag{6.4}$$

where $RMS_{i,\bar{x}}$ is the root mean square error between the position of the player $i \in [TP, CP]$ and the mean position of the group, while ϵ is a non-negative parameter (details about the RMS are given in Section 4.3.1).

6.3 Control synthesis of the Cyber Player

Ideally, data used to train the CP should be extracted from real human players playing group sessions of the mirror game. As done also in the case of dyadic interaction, in the absence of a sufficiently large experimental dataset the data needed to feed the CP during the training can be generated synthetically making several virtual trainers, modelling human players, perform

sessions of the mirror game against each other. Being a multi-player scenario, we adapted the cognitive architecture described in Section 4.3 not only to drive the virtual trainer in leader-follower sessions of the mirror game but also in the joint improvisation case. This kind of interaction differs from a leader-follower configuration as no prior roles are assigned to the players but they move together to create a coordinated motion. The virtual trainer in its more general form is explained in the next section.

6.3.1 Virtual trainer in joint improvisation

As done for the leader-follower interaction, we model the motion of the virtual trainer by means of a nonlinear HKB oscillator of the form (details about the HKB in Section 4.1.1):

$$\ddot{x} + \left(\alpha x^2 + \beta \dot{x}^2 - \gamma\right)\dot{x} + \omega^2 x = u, \tag{6.5}$$

where $x, \dot{x}$ and $\ddot{x}$ are position, velocity and acceleration of the VT end effector respectively, α, β, γ are positive empirically tuned damping parameters while ω is the oscillation frequency.

The optimal controller described in Section 4.1.2 and designed for the mirror game in leader-follower condition, is extended to model the task in joint-improvisation, resulting in a more general cost function of the form [206]:

$$\begin{aligned} \min_u J(t_k) = \frac{\theta_p}{2}\left(x(t_{k+1}) - r_p(t_{k+1})\right)^2 + \frac{\theta_\sigma}{2}\int_{t_k}^{t_{k+1}} \left(\dot{x}(\tau) - r_\sigma(\tau)\right)^2 d\tau + \\ + \frac{\theta_v}{2}\int_{t_k}^{t_{k+1}} \left(\dot{x}(\tau) - r_v(\tau)\right)^2 d\tau + \frac{\eta}{2}\int_{t_k}^{t_{k+1}} u(\tau)^2 \, d\tau, \end{aligned} \tag{6.6}$$

where r_p, r_v is the position and the velocity time series of the partner player, r_σ is the reference signal modelling the desired human motor signature and generated by the Markov chain, η tunes the control effort, t_k and t_{k+1} represent the current and the next optimisation time instant. $\theta_p, \theta_s, \theta_v$ are positive control parameters satisfying the constraint $\theta_p + \theta_s + \theta_v = 1$. By tuning appropriately these parameters, it is possible to change the VT configuration making it act as a leader, follower or joint improviser [206, 207]. In the most trivial cases:

- $\theta_p = 1, \theta_\sigma = 0, \theta_v = 0$: the VT behaves as a perfect follower aiming at minimising only the position error with the other agent;
- $\theta_p = 0, \theta_\sigma = 1, \theta_v = 0$: the VT behaves as a blind leader aiming at minimising the velocity error with the reference signal completely ignoring the other agent;
- $\theta_p = 0, \theta_\sigma = 0, \theta_v = 1$: the VT copies the velocity of the other agent without considering the position mismatch.

In the case of a multi-player scenario, r_p and r_v are taken as the mean value of the position and the velocity of the target player's neighbours, that is:

$$r_p := \frac{1}{M}\sum_{j=1}^{M} x_j; \quad r_v := \frac{1}{M}\sum_{j=1}^{M} \dot{x}_j, \tag{6.7}$$

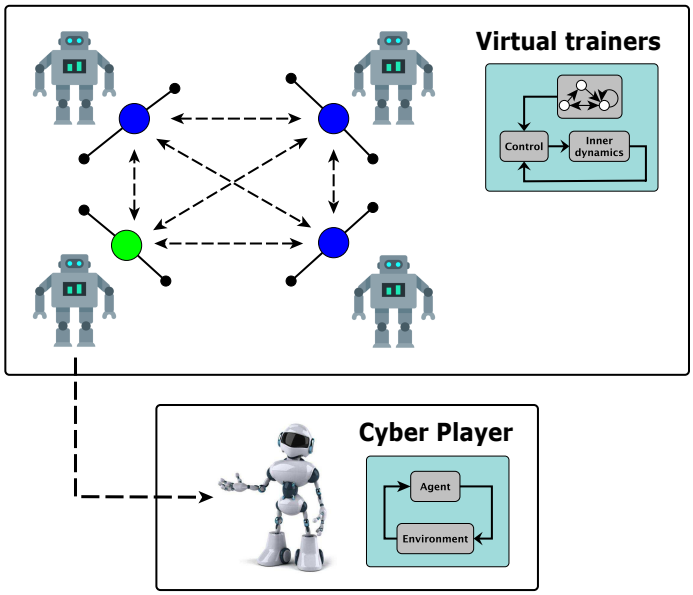


FIGURE 6.2. Architecture used to train the CP to mimic the target player (ball in green). The group is simulated by 4 VTs playing the mirror game in joint improvisation and driven by the cognitive architecture based on optimal control and Markov chains. The CP, driven by a deep reinforcement learning approach, receives in real time the data from the group and learns how to interact in order to replace the target player.

where M is the number of neighbours and x_j and $\dot{x}_j$ are the position and the velocity of the jth neighbour.

6.3.2 Training

To train the CP to coordinate its movements in the group like the virtual trainer target does, a group of 4 different virtual trainers interconnected in a complete graph were used (Figure 6.2). Without loss of generality, VT_4 was taken as the target player that the deep learning driven CP has to mimic. The experience replay in the CP algorithm was implemented with a buffer of $200,000$ elements, batches of 32 sampled states were used to train the feedforward neural network at each iteration. A target network updated every 150 time steps was considered in the Q-function, with a discount factor $\gamma = 0.95$ and a learning rate $\alpha = 0.1$.

The training stage was carried out on a Desktop computer having an Intel Core i7-6700 CPU, 16 GB of RAM and 64-bit Windows operating system. It took less than $1,000$ trials of 500 observations each to converge (around 3 hours) according to the criterion (6.4). In Figure 6.3 the training curve is reported showing for each trial the RMS error between the VT_4 and the group (in blue), and between the CP and the group (in red); in black the cumulative reward function over trials.

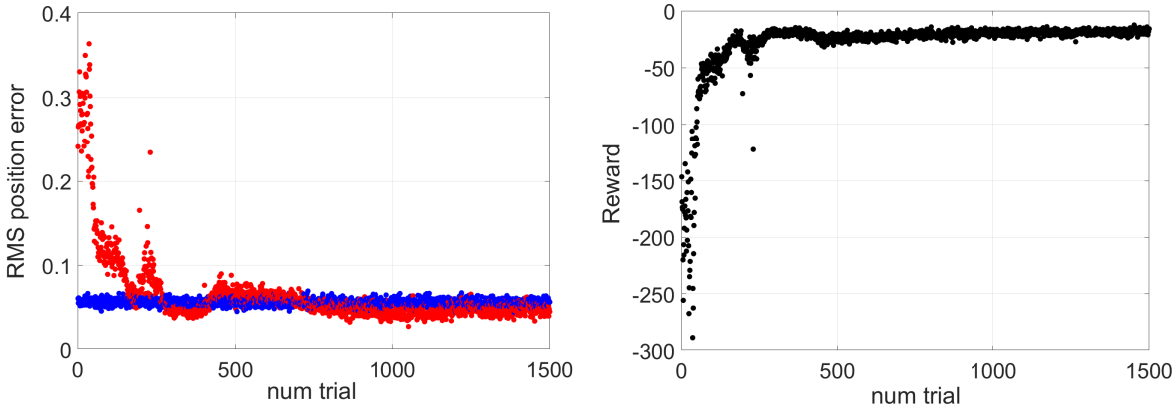


FIGURE 6.3. Training curve and cumulative reward showing the convergence of the algorithm. The root mean square error in position is reported for each trial both for the target VT (in blue) and the CP (in red). The cumulative reward is also reported in black for each trial.

6.3.3 Coordination metrics

In order to quantify and analyse the coordination in a network of more than two agents, we use some of the metrics proposed in the context of human group coordination [156] in addition to the metrics already provided in Section 4.3.1 used to evaluate the quality of the tracking.

Let $x_k(t) \in R \; \forall t \in [0,T]$ be the continuous time series representing the motion of each agent, with $k \in [1,N]$, where N is the number of agents and T is the duration of the trial.

Let $x_k(t_i) \in R$, with $k \in [1,N]$ and $i \in [1,N_T]$, be the respective discrete time series of the k-th agent, obtained after sampling $x_k(t)$, where N_T is the number of time steps and $\Delta T := \frac{T}{N_T}$ is the sampling period.

Let $\theta_k(t) \in [-\pi,\pi]$ be the phase of the k-th agent, which can be estimate by making use of the Hilbert transform of the signal $x_k(t)$ [102]. In (6.8) and (6.9) we define the *cluster phase* or *Kuramoto order parameter* both in its complex form $\Psi'(t) \in \mathbb{C}$ and in its real form $\Psi(t) \in [-\pi,\pi]$ which can be regarded as the average phase of the group at time t:

$$\Psi'(t) := \frac{1}{N}\sum_{k=1}^{N} e^{j\theta_k(t)}, \tag{6.8}$$

$$\Psi(t) := \tan^{-1}\frac{\Im(\Psi'(t))}{\Re(\Psi'(t))}. \tag{6.9}$$

Let $\phi_k(t) := \theta_k(t) - \Psi(t) \in [-\pi,\pi]$ be the relative phase between the k-th participant and the group phase at time t. We can define the relative phase between the k-th participant and the

group averaged over the time interval $[0,T]$ both in its complex form $\bar{\phi}'_k \in \mathbb{C}$ and its real form $\bar{\phi}_k \in [-\pi,\pi]$ as follows:

$$\bar{\phi}'_k := \frac{1}{T}\int_0^T e^{j\phi_k(t)}\,dt \simeq \frac{1}{N_T}\sum_{i=1}^{N_T} e^{j\phi_k(t_i)}, \tag{6.10}$$

$$\bar{\phi}_k := \tan^{-1}\frac{\Im(\bar{\phi}'_k)}{\Re(\bar{\phi}'_k)}. \tag{6.11}$$

In order to quantify the degree of synchronisation for the k-th agent within the group, we define the parameter ρ_k in (6.12) that gives information on how much the k-th agent is synchronised with the average trend of the group: the closer ρ_k is to 1, the better is the synchronisation of the k-th agent.

$$\rho_k := \left|\bar{\phi}'_k\right| \quad \in [0,1]. \tag{6.12}$$

The coordination level of the entire group at time t can be quantified as follows:

$$\rho_g(t) := \frac{1}{N}\left|\sum_{k=1}^{N} e^{j[\phi_k(t)-\bar{\phi}_k]}\right| \quad \in [0,1], \tag{6.13}$$

where N is the number of individuals, $\phi_k(t)$ is the relative phase between the k-th participant and the group phase at time t and $\bar{\phi}_k \in [-\pi,\pi]$ is the same relative phase averaged over the duration of the experiment T. The closer $\rho_g(t)$ is to 1, the better the synchronisation level of the group is at time t. Such synchronisation function can be averaged over the whole interval $[0,T]$ in order to have an estimate of the mean coordination level of the group:

$$\rho_g := \frac{1}{T}\int_0^T \rho_g(t)\,dt \simeq \frac{1}{N_T}\sum_{i=1}^{N_T}\rho_g(t_i) \quad \in [0,1]. \tag{6.14}$$

6.3.4 In-silico validation

To show that the cyber player is effectively able to emulate the target virtual trainer when engaged in a group, we validated its performance when interacting over several topologies, different from the one used during the training. Specifically, we trained the CP over a complete graph and validated also with a ring, path and star graph as shown in Figure 6.4.

In Figure 6.5 the validation is reported both for the topology used during the training (complete graph) and for one of those not used (path graph). The performance of the CP was evaluated by comparing its behaviour with that of the target VT it was trained upon. The CP (in red) successfully tracks the mean position of the group (dashed line in black) being able to mimic

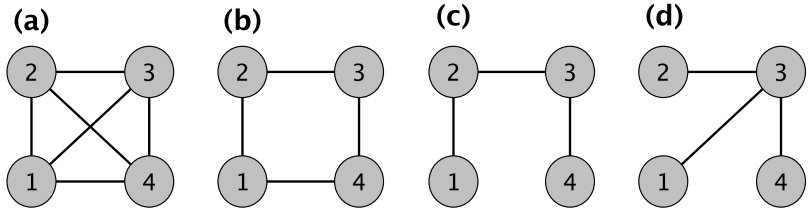


FIGURE 6.4. Network topologies implemented to validate the CP when playing in group scenario. Specifically the group consists in 4 different players represented as nodes in the network, whereas the visual coupling between two players is represented by an edges between the corresponding two nodes. During the training, the CP has to mimic the player labelled with the number 4. (**a**) Complete graph. (**b**) Ring graph. (**c**) Path graph. (**d**) Star graph.

the target player it was trained to imitate (in blue) when it is connected with the others through a complete topology [panel (a)]. The RPE (defined in Section 4.3.1) was also evaluated between the target VT and the mean position of the neighbours and compared with the relative position error between the CP and the same mean position [panel (c)]. Both the errors are very small and with comparable mean values, less than 0.1. Similar considerations can be done for behaviour of the CP when the group topology is a path graph, not used during the training, as shown in panels (b) and (d).

The key features of the motion of the CP and the VT were captured by the relative phase, RMS position error and time lag (metrics defined in Section 4.3.1). Such metrics were evaluated performing 20 trials of 60 seconds and reporting both the mean value and the standard deviation for all the tested network topologies (Table 6.1). It is possible to notice that all indeces show a remarkable degree of similarity between the motion of the CP and that of the target VT.

In Figure 6.6 the group level synchronisation is reported for each topology. Despite the different topologies, the presence of the CP does not alter the group dynamics when the CP is substituted to the VT it was trained upon. We notice that the level of coordination varies with the topology, confirming what found in [12]. Specifically, in [12] as confirmed by Figure 6.6, the complete and the star graph were found to be associated with the highest level of synchronisation.

6.4 Experimental validation

6.4.1 Experimental setup

Participants. A total of 4 people participated in the experiments (4 females, all right-handed). Participants were PhD students and Postdoctoral Researcher Associates from University of Bristol. The players were assigned numbers from 1 to 4 (HP_1, HP_2, HP_3, HP_4).

This study was carried out according to the principle expressed in the Declaration of the

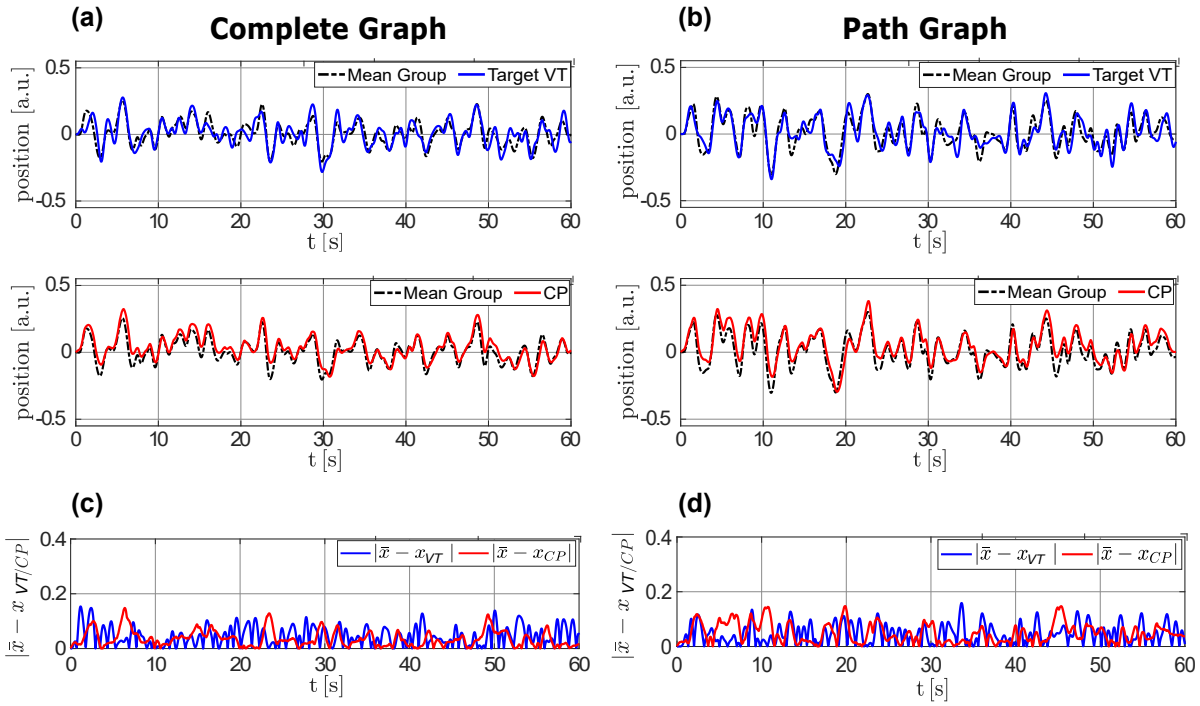


FIGURE 6.5. **(a)-(b)** The position trajectories of the target VT (in blue) and of the CP (in red) are reported together with the mean position of the neighbours (in dashed back). **(c)-(d)** Relative position error evaluated between the mean position and the position of the target VT (in blue) and the CP respectively. Two different topologies are depicted: **(a)-(c)** the complete graph and **(b)-(d)** the path graph.

Helsinki. All subjects gave written informed consent for both experiments participation and publication of identifying images.

Experimental task. All participant were asked to perform 8 group trials of the mirror game each lasting 60 seconds in each implemented topology. HP_4 was then substituted with the CP (driven by the DQN learning algorithm). Once again, 8 group trials lasting 60 seconds each were performed for each topology.

The given instruction to the HPs was to try to coordinate their own motion with that of the others. The players did not have any knowledge about the interaction patterns, to whom they are connected, whether one of the traces is driven by the cyber player. The visual pairings implemented for the group are: complete graph, ring graph, path graph and star graph (details are in Section 6.1). All of them are described by undirected graph, participant i sees the motion of the participant j if and only if the participant j sees the motion of the participant i too.

Experimental platform. All the experiments were performed through CHRONOS. Already used in dyadic experiments in Chapter 5, CHRONOS also allows groups to perform coordination

	Metric	CP	Target VT
Complete graph	Relative phase	-0.168 ± 0.051	-0.05 ± 0.095
	RMS position error	0.044 ± 0.005	0.056 ± 0.004
	Time lag	-0.114 ± 0.016	-0.015 ± 0.036
Path graph	Relative phase	-0.23 ± 0.079	-0.159 ± 0.098
	RMS position error	0.056 ± 0.005	0.06 ± 0.004
	Time lag	-0.104 ± 0.015	-0.068 ± 0.024
Ring graph	Relative phase	-0.181 ± 0.061	-0.033 ± 0.132
	RMS position error	0.046 ± 0.003	0.054 ± 0.004
	Time lag	-0.101 ± 0.015	-0.02 ± 0.027
Star graph	Relative phase	-0.204 ± 0.051	-0.211 ± 0.107
	RMS position error	0.055 ± 0.003	0.062 ± 0.004
	Time lag	-0.104 ± 0.015	-0.08 ± 0.029

TABLE 6.1. Metrics are reported for 20 trials of the multiplayer mirror game in complete graph, path graph, ring graph and star graph. The relative phase, the RMS position error and the time lag are evaluated between the mean of the neighbours with the CP (first column) and with the target VT (second column).

tasks. CHRONOS extends the mirror game to a multiplayer scenario, allowing each participant to run the game on his/her laptop where s/he can move the position of an object visualised on computer screen and see traces of the objects moved by the others connected to their own laptops over Internet. Furthermore, the software allows to implement different interaction patterns among the participants showing on each screen only the traces of a designed subset of players in the group (decided a priori by an administrator).

Being the game implemented on different laptops connected over internet, CHRONOS removes any form of social interaction among the participants that inevitably affects the level of coordination. Indeed, body movements, friendship relationships, shared feelings and level of hierarchy have a significant effect on how each individual moves in the ensemble choosing his/her preferred partner(s) to interact the most with [56, 57, 99, 142]. Removing social interaction, the participants are connected solely through visual couplings making it possible to assess the impact of the structure of interconnections on the level of coordination. Since the joint task is performed through a computer setup, the participants are not aware whether the traces of the objects on their screen are driven by another human player or by the artificial agent. In so doing, we can easily analyse the ability of the avatar in merging the group without being recognised by the other members.

As in the case of dyadic experiments, leap motion position sensors [193] connected to the laptops were used to store the trajectories provided by the players during all the trials, respectively. Furthermore, each participant was required to wear headphones through which white noise

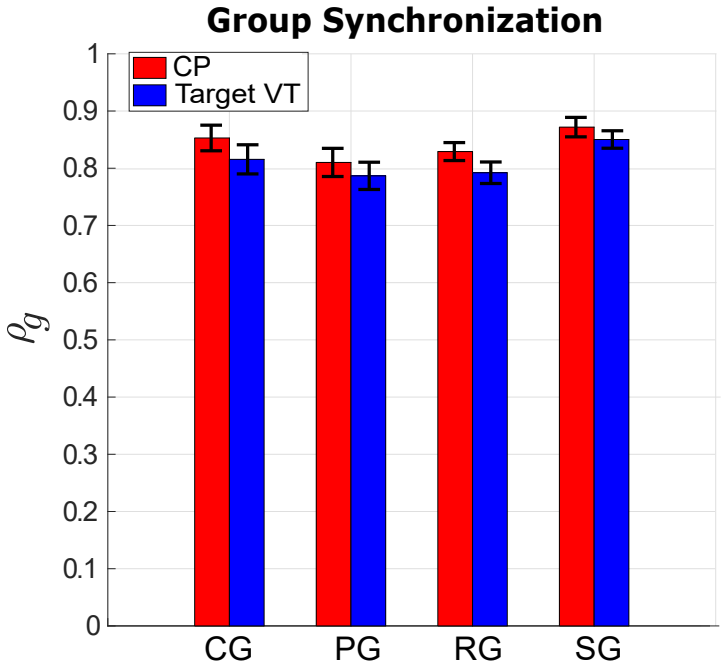


FIGURE 6.6. Histogram reporting the group synchronisation level reached by the group both with the VP target (blue bars) and with the CP (red bars). Different topologies were implemented during the validation: complete graph (CG), path graph (PG), ring graph (RG) and star graph (SG) with player 3 as center node.

was transmitted in order to eliminate possible auditory couplings with the others. A schematic drawing representing CHRONOS architecture is shown in Figure 6.7.

6.4.2 Validation

The metrics employed to analyse and quantify the participants' performance as well as the level of coordination are those described in Section 6.3.3.

Group synchronisation levels ρ_g were evaluated for each implemented topology both for the group of only humans and for the group with the cyber player and averaged across the trials. Their means and standard deviations over the total number of participants in both groups are shown for each topological structure in Figure 6.8.

Paired T-test was performed to understand if replacing the player 4 with the cyber player affects somehow the synchronisation level reached by the group. Specifically, the statistical analysis at 95% confidence level with 7 degrees of freedom reported: complete graph $t(7) = -1.425, p = 0.197$; path graph $t(7) = -0.466, p = 0.655$; ring graph $t(7) = -0.522, p = 0.618$; star graph $t(7) = -1.491, p = 0.180$. The results revealed that the presence of the cyber player in the group has no statistically significant effects on the synchronisation level in none of the interconnected networks.

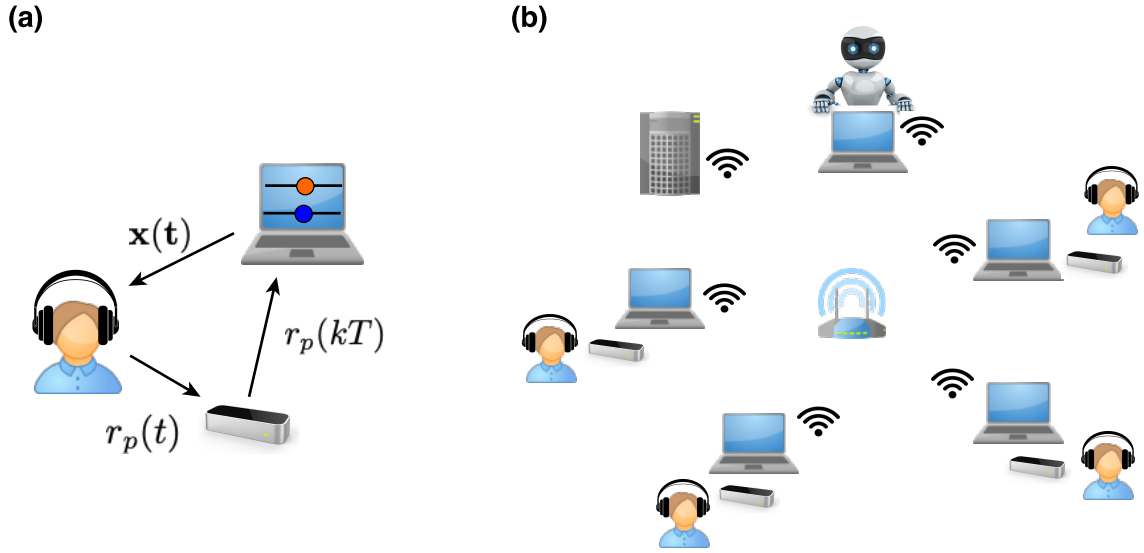


FIGURE 6.7. **(a)** The position of the human's finger $r_p(t)$ is recorded by a leap motion controller. Sampled $r_p(kT)$ is then sent to the laptop and shown as a blue circle on the screen. The vector of the other participants' position **x(t)** to whom the human is connected appears as a set of orange circles. **(b)** Simplified CHRONOS architecture consisting of N human players wearing headphones and M artificial agents. All players are connected over internet through a WiFi router. The server implements the structure of interconnections between the players, handling the exchange of the players' trajectories with each other.

This statistical analysis aimed at to assess the absence of statistical differences between the behaviour of the CP and the real one of the human, it was trained to upon, for each different network topology. Nevertheless, other statistical analysis can be used to estimate the variation in the data. For example, ANOVA can be performed to assess whether the network topology influences the level of synchronisation. Furthermore, ANOVA allows to assess whether the effect due to the interaction between the network topology and the presence of the cyber player has some statistical relevance.

In addition to the group synchronisation metric that averages over the trials and over the group members, the levels of dyadic synchronisation ρ_d between the HP_4 and the others were also evaluated and respectively compared to the values of dyadic synchronisation obtained between the CP and the other group members (Figure 6.9). We reported the values of dyadic synchronisation for each implemented topology showing high levels of synchronisation with small standard deviation in complete, ring and star graph for both human and cyber player. A larger variance was observed when the human player performed the task while interconnected in the path graph and reflected respectively in the cyber player while playing in the same topology.

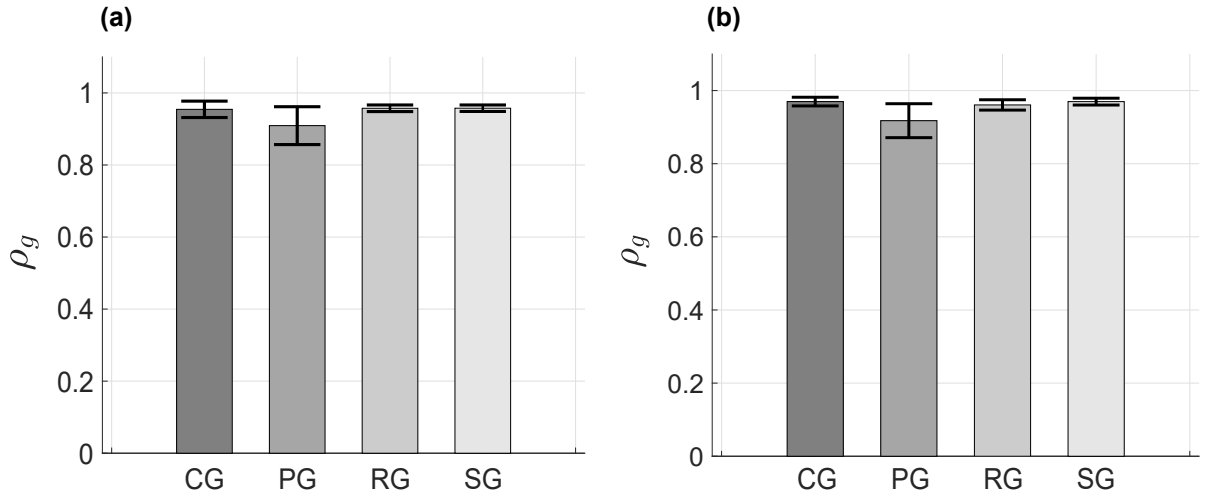


FIGURE 6.8. Group synchronisation ρ_g evaluated for different topologies (CG: complete graph, PG: path graph, RG: ring graph, SG: star graph) both in a network of 4 humans **(a)** and in a network of 3 humans and 1 virtual agent **(b)**. Different scale of grey refers to the different network structures. Each bar is representative of the mean value averaged over the 8 trials for each topology, while the black error bar shows its averaged standard deviation.

Furthermore, player 4 registered higher level of synchronisation with player 3 as well as the cyber player. These results suggest that the cyber player driven by the deep reinforcement learning approach is able to well coordinate its motion with those of the others in different group scenarios reproducing the same motor behaviour of the human that it is emulating.

6.5 Summary

In this chapter we addressed the problem of synthesising an autonomous artificial agent able to coordinate its movements and perform the mirror game task in a group setting. To achieve our goal, we introduced a deep reinforcement learning algorithm in which a feedforward neural network was used to approximate the nonlinear action-value function. To train the cyber player we used the same trick as in the dyadic case, explained in the previous chapter, to generate enough synthetic data to allow the DQN algorithm to converge towards the optimal solution. Specifically we made several virtual trainers (tuned and embedded with the IMS Generator to emulate human players) play the mirror game over the network.

After the in-silico validation, we further validated the CP performing our own experiments making several humans play sessions of the mirror game with and without the cyber player. The experiments were carried out through CHRONOS, developed by me to study both human-human and human-avatar coordination. We briefly described CHRONOS architecture underlying its

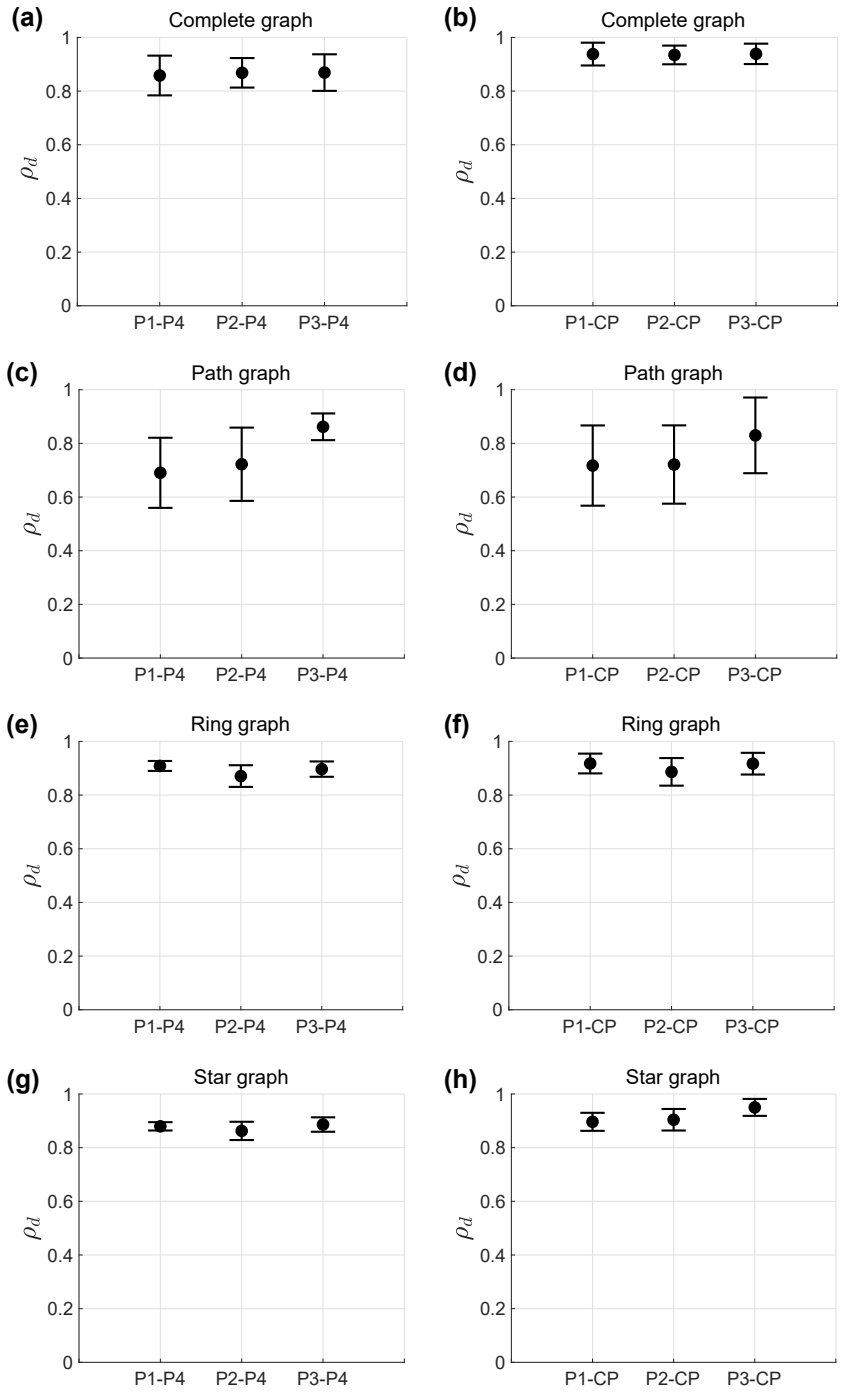


FIGURE 6.9. Dyadic synchronisation ρ_d evaluated between the human player 4 and each of the other players in **(a)** complete, **(c)** path, **(e)** ring and **(g)** star graph. Besides, dyadic synchronisation ρ_d evaluated between the CP, replacing the human, and each of the other players in **(b)** complete, **(d)** path, **(f)** ring and **(h)** star graph. Dots and error bars represent mean and standard deviation of ρ_d over 8 trials.

indispensable value in implementing the important requirements during the experiments: *i)* removing the social interaction between the participants and *ii)* introducing easily the avatar in the group without that the other participants have knowledge of that.

We evaluated the performance of the CP in terms of group synchronisation and dyadic synchronisation reached with the partners and comparing these values with those characterising the human player it was emulating. We found that the cyber player was able to play sessions of the mirror game exhibiting desired human features in group interaction becoming in this way a good surrogate of the human it was trained upon.

CHAPTER 7

THE VIRTUAL PLAYER IN A REAL APPLICATION: FIRST STEPS TO BUILD AN EXERGAME FOR MOTOR DISEASES

In this chapter we deploy the virtual player in a real application. The work presented in what follows was carried out under the project INDAGO in collaboration with Prof. Pasquale Arpaia (University of Naples "Federico II") and Giovanni D'Addio (Fondazione Salvatore Maugeri, Care and Research Institute) aimed at developing an innovative platform for the monitoring of, diagnosis of and rehabilitation from motor neuron diseases. The main novelty and contribution with respect to previous literature is the design and the implementation of an exergame for post-stroke therapy having an interactive avatar able to propose autonomously therapeutic motor exercises personalised to the patient's needs. Key elements on which the exergame is based are *i)* establishing motor coordination between the patient and the avatar and *ii)* designing a mathematical model of the voluntary motion of the arm capturing the main features that distinguish the motion of a healthy person from that of a patient.

Specifically, in the rest of this chapter an exhaustive motivation for INDAGO is given in Section 7.1 explaining the state of the art of current home therapies, discussing their limitations and proposing the novelty approach of our rehabilitation platform. In Section 7.2 we give more details on the Hogan model (used to described arm movements) and describe the motor task selected as a therapeutic exercise. Then, we describe the proposed rehabilitation platform analysing the architecture and the steps of the therapeutic process that such platform aims at implementing. Then, Section 7.4 is completely dedicated to the design of the adaptive avatar, presenting the overall cognitive architecture able to drive it in dyadic coordination tasks with the patient. A preliminary validation of the proposed exergame is presented in Section 7.5. The software implementation of the exergame in virtual reality is explained in Section 7.6 showing a typical use case between the patient and the platform. Finally, a summary of the results is

provided in Section 7.7.

7.1 Motivation

Stroke is one of the leading causes for motor disability worldwide (stroke events have increased by 30% in the last 5 years [191]). The most common deficit after a stroke is the hemiparesis of the contralateral upper limb, of which muscle weakness, changes in muscle tone, joint laxity, and impaired motor control are physical manifestations [43]. These impairments induce disabilities in common daily activities such as reaching, picking up objects, and holding objects, thus contributing to poor quality of life. To date, rehabilitation training is the most effective way to reduce motor impairments in stroke patients [84].

In the current healthcare model, post-stroke rehabilitation consists of daily and systematic exercises carried out by the patient under the direct supervision of clinicians. In particular, two different phases can be identified in the rehabilitation process: *i) acute phase*, in which the patient resides at the hospital so that initial rehabilitation is part of the hospital treatment, and *ii) long-term outpatient rehabilitation*, in which the patient has to regularly carry out daily rehabilitation exercises (typically in specialised centers) aiming at restoring strength, coordination, and mobility as closely to normal as possible [147]. Due to high costs, low resources and the increment of the number of stroke patients, the support to outpatient rehabilitation is progressively shrinking. Furthermore, daily commuting to rehabilitation centers can be infeasible for those people that live in peripheral areas. For these reasons, in the last few years much effort from many research areas has been spent on the development of home therapies.

Examples of recent home therapies are the so called *exergames* (exercise + game). Exergames let patients exercise while playing games that hide therapeutic repetitive tasks needed for rehabilitation under the world of fantasy. Many exergames have been developed for postural, neglect and schizophrenia rehabilitation [13, 30, 148], so as to demonstrate that exergame-based therapy is really effective [51, 167]. Typically, existing exergames consist of a series of "static" exercises (e.g., pop the balloon, catch the fruit, collect the hay and so on) that the patient has to carry out, often without a direct supervision of the specialist [13, 71]. Feedback mechanisms are particularly important in this kind of rehabilitation, where wrong movements or postures make the exercise not only useless but even dangerous. Very limited examples of exergames exist in which a feedback system is present in the form of a virtual assistant [147]. The main drawback of existing virtual assistants is that they are "external" agents, not directly involved in the exercise, and having as a main task only that of inviting the patient to the therapy, summarising the results and encouraging him/her to carry on.

The main goal of this work is to propose a preliminary version of an innovative and "smart" exergame for motor rehabilitation with an adaptive avatar assisting the patient along all the therapeutic process. Specifically, such virtual agent is required to be able to drive the patient

during the exercise (and so performing the exercise with him/her) adapting its motion to that of the patient. Also, the agent must be able to analyse the continuous stream of motion data coming from a proper sensor system, and *i)* diagnose the level of the disease, *ii)* evaluate the progresses/regresses made by the patient and *iii)* personalise the therapy to the specific patient.

Since the avatar is required to diagnose the level of motor impairment of the upper limb of the patient, the Hogan model is taken here as a reference mathematical model describing theoretically "healthy" human movements. Firstly proposed by Hogan in 1984 [91], it is largely used by clinicians themselves to describe voluntary arm movements. It has been observed that the motion of patients after stroke has characteristic patterns that differ from those of a healthy individual and described by the Hogan model. Specifically, a healthy individual has been observed to have a single-peaked velocity profile while moving his/her upper limb, whereas a patient is characterised by a multi-peaked velocity profile according to the level of impairment. The main aim of our avatar will be to propose motor exercises aimed at reducing the differences in the velocity profiles between the motion of the patient under therapy and the one of a healthy individual captured by the Hogan model.

7.2 Design of the rehabilitation task

7.2.1 Reaching motor task

We selected reaching tasks as representative motor exercises, largely used by clinicians in the traditional motor rehabilitation from stroke [41, 190]. In detail, the patient is asked to sit so as to have his/her forearm comfortably placed on a table. Keeping the elbow pivoted at a point, the exercise consists in moving the forearm from a point to another along a circular arc of a certain length. The elbow, forearm and hand of the patient are considered as a unique rigid body (Figure 7.1).

Reaching tasks have been used also as experimental setups in many studies to understand the kinematic and the dynamic aspects of voluntary human arm movements. Some of them identified some common kinematic features and patterns that characterise voluntary human movements [5, 74, 125, 180]. In particular, it has been observed that humans, when moving their arm from one point to an other, tend to generate straight and smooth trajectories. Smooth trajectories imply no discontinuities in the acceleration and so in a characteristic single-peaked, bell-shaped velocity profile.

Hogan model. In [91], Neville Hogan proposed the first mathematical model (now known as the "Hogan model") to describe the principles underlying voluntary arm movements. In particular, it is assumed that maximising the smoothness can be modelled by minimising the mean square jerk of the motion, defined mathematically as the third derivative of the position with respect to time. Such a problem can be formalised in terms of optimal control problem having the following

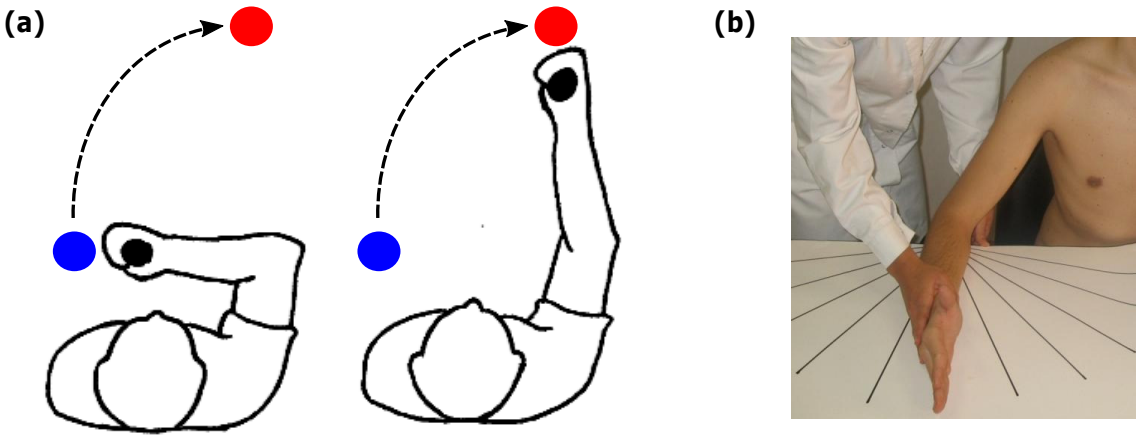


FIGURE 7.1. **(a)** Illustration representing a human performing the reaching task at the beginning and at the end of the task. Blue circle represents the start angle position, whereas the red circle represents the final angle position. A dashed arrow indicates the trajectory that the human has to follow. **(b)** Real scenario in which human is asked to move the forearm in an angular motion assisted by the clinician.

cost function to minimise:

$$C = \frac{1}{2}\int_0^T \dddot{\theta}\, dt, \tag{7.1}$$

where $\dddot{\theta}$ is the jerk of the trajectory and T is the duration of the movement. A sufficient set of boundary conditions consists of:

$$\begin{cases} \theta(0) = \theta_0, \\ \theta(T) = \theta_f, \\ \dot{\theta}(0) = 0, \\ \dot{\theta}(T) = 0, \end{cases} \tag{7.2}$$

where θ_0 and θ_f are the start and final angular position respectively.

Hogan proved that the trajectory $\theta(t)$ solving the optimal control in (7.1), and having as set of boundary conditions those in (7.2), is a fifth order polynomial in time, completely independent of the physical system generating the motion. Such a fifth order polynomial has the form:

$$\theta(t) = \theta_0 + (\theta_f - \theta_0)\left[10\left(\frac{t}{T}\right)^3 - 15\left(\frac{t}{T}\right)^4 + 6\left(\frac{t}{T}\right)^5\right] \qquad 0 \le t \le T, \tag{7.3}$$

where, once again, θ_0 and θ_f are the start and final angular position respectively, and T is the duration of the movement.

The Hogan model is largely used in the current rehabilitation medicine as a reference model describing the human arm movements. Also in this thesis it is used as reference during the therapy sessions to describe the motion of a healthy individual.

7.2.2 Rehabilitation process

In this section, we illustrate in detail the phases of the proposed therapeutic process which is assisted by the avatar. In the *preliminary phase*, the patient is asked to perform the reaching task several times without being assisted (but in presence of the clinician) trying to reach his/her maximum angular extension, so as to evaluate his/her Range of Motion (ROM). The Range of Motion is defined as the maximum range through which a joint of the body (in our case it is the elbow) can be moved measured in degrees of a circle.

Contextually, the patient is also asked to perform specific exercises aiming at evaluating some physical parameters characterising the patient's forearm, e.g., inertia, friction, stiffness (such parameters will be used to model the dynamics of the forearm in the avatar). Such parameters can be evaluated either experimentally or by using the personal details of the patient (age, weight, height). For example, an experimental method to evaluate the inertia of the patient's forearm is that of applying a short mechanical impulse to the elbow of the patient and measuring its acceleration. The inertia can be evaluated dividing the impulse's width by the acceleration.

After this preliminary phase, each *therapeutic session* consists of 2 main stages:

Stage 1. The avatar has to follow the patient (AWING phase: Avatar folloWING): in this phase the patient is asked to perform the reaching motor task several times without being assisted. The angle of the proposed tasks span within the ROM and each of them is repeated three times by default. The number of required tasks having different angles is decided by the clinician. During this phase, the avatar behaves as a perfect follower, "observing" the patient's motion and extracting the main "features" of his/her motion so as to parameterise an internal mathematical model of the patient and evaluate the optimal trajectory given by the Hogan model. Having the motion of the patient and the optimal one, the avatar can identify the differences between the two motor profiles through some performance metrics and decide the level of disability of the patient.

Stage 2. The patient is asked to follow the avatar (PAWING phase: Patient folloWING): according to the performance metrics evaluated in the previous stage, the avatar decides which exercises to propose to the patient. In this phase the avatar has to be adaptive and lead the patient in performing the exercise together ensuring to be followed by him/her. Each exercise is repeated three times by default. Also in this case, the total number of required tasks having different angles is decided by the clinician.

The session described above is repeated iteratively until the rehabilitation goal is reached. The goal of each session is to reduce the distance between the velocity profile of the patient and the velocity profile of the Hogan model. As the therapy proceeds, the several velocity peaks

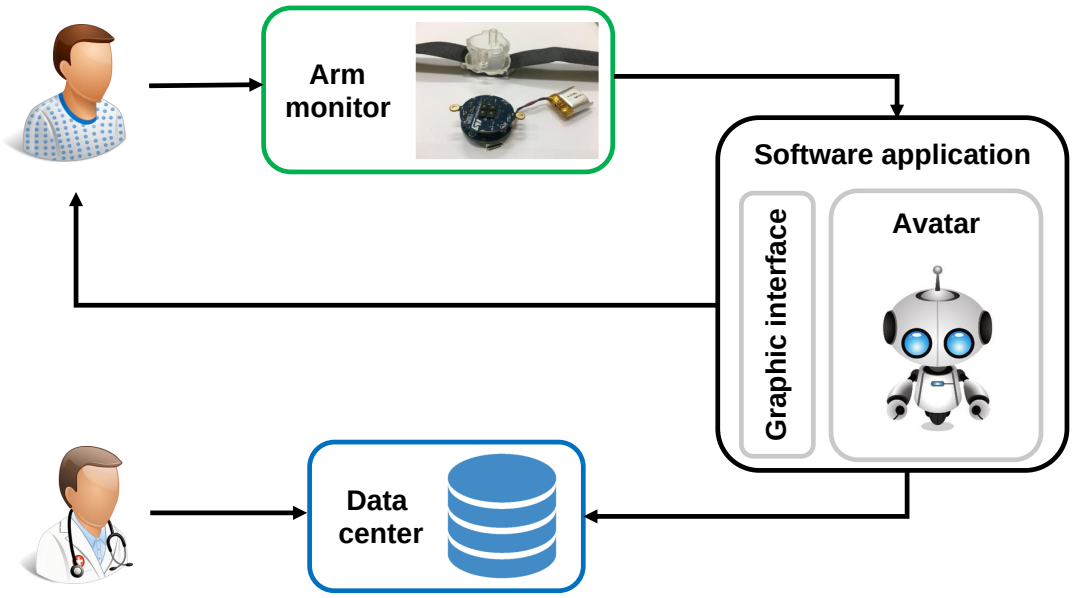


FIGURE 7.2. Architecture of the rehabilitation platform. The patient is constantly monitored by an arm monitor that acquires angular position and angular velocity of his/her forearm during the exercise. Measured position and velocity are sent to the avatar that adapts its behaviour to that of the patient. Data are also stored into a database in order to be checked by clinicians.

characterising the profile of the patient at the beginning of the rehabilitation have to become closer and closer to each other until merging in a single-peaked curve.

7.3 Architecture of the rehabilitation platform

The rehabilitation platform designed to support the therapeutic process described in Section 7.2 mainly comprises three modules (see Figure 7.2):

1. *an arm monitor*: composed of a set of sensors to measure the angular position and velocity of the patient's forearm during the movement. In particular the iNEMO inertial module of the ST-Microelectronics was used as a wireless inertial measurement units (IMU), which integrates different types of sensors (accelerometer, gyroscope and magnetometer). Furthermore, it is easy to wear (like a wristwatch) and integrates a battery lasting 17 days (Figure 7.3). The arm monitor was implemented in collaboration with Anna Lauria, PhD student at University of Naples "Federico II";

FIGURE 7.3. Arm monitor used in the rehabilitation platform used to acquire the angular position and velocity of the human. Lightweight and battery powered, the arm monitor can be worn like a wristwatch.

2. *a software application*: mainly divided into two parts, one implementing the adaptive avatar and all logic behind the proposed rehabilitation strategy, and an other implementing the graphic interface directly interacting with the patient (the latter was implemented both for virtual reality systems using Oculus Rift S and for Windows operative systems);

3. *a storage facility*: storing the data acquired during the exercises (e.g. position and velocity time series of the patient and avatar), personal information of the patient and the avatar settings.

The patient is required to wear the arm monitor placing it on his/her forearm before starting the therapy session. After the calibration of the sensors, the raw signal acquired by the gyroscope is filtered through an anti-aliasing lowpass filter with cutoff frequency at 8 Hz in order to get the angular velocity. The angular position is obtained from the velocity by means of the Kalman filter. The measured position and velocity are then sent to the software application in real time via Bluetooth.

The avatar, implemented in the software application, adapts its behaviour in response to the data coming from the sensors in order to coordinate its motion with that of the patient and drive him/her towards the final point of the trajectory. The patient can visualise both movements (their own and that of the avatar) through a proper graphic interface (see Figure 7.4). At the end of the task, the software application evaluates performance metrics in order to asses the current degree of the disabilities of the patient, compares the patient's motion with the "healthy" trajectory given by the Hogan model and store all information in the data center (Figure 7.2).

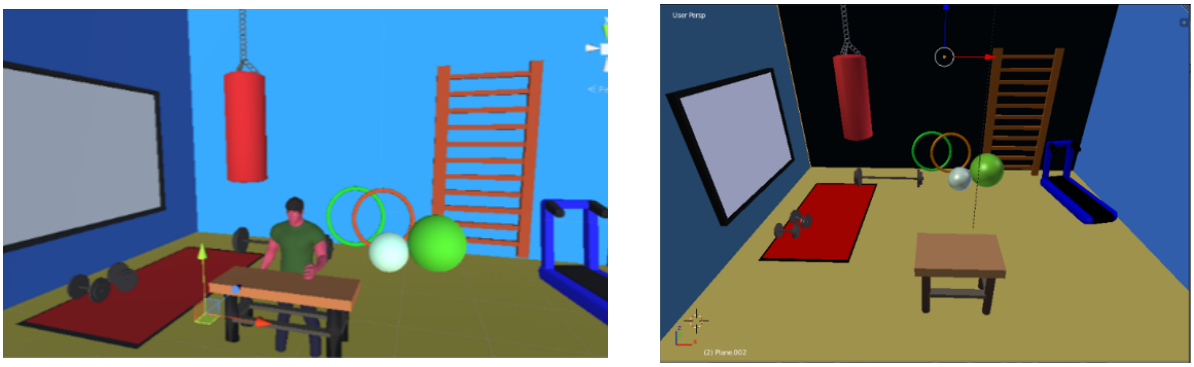


FIGURE 7.4. Graphic interface seen by the patient during the exercise. The environment is implemented in VR (with Unity) and simulates a typical gym at a therapy centre. The avatar is morphologically similar to a human in order to enhance the social sense of community and so the coordination with the patient.

7.4 Cognitive architecture of the avatar

In this section we design the cognitive architecture of the avatar so as to implement the rehabilitation strategy explained previously and help the patient during his/her recovery. In particular, the avatar is required to dynamically adapt its motor behaviour in order to accomplish a dual goal: *i)* lead the patient's motion towards a healthy kinematic profile and *ii)* adapt its motion to that of the patient (e.g., slowing down or following the motion of the patient) in order to complete the motor task together.

The designed architecture is mapped on the real-time control schematics shown in Figure 7.5 whose blocks are briefly described below:

1. *inner dynamics block*: represents how the avatar moves in the absence of any coupling with the patient. An angular harmonic oscillator is employed to described such dynamics;

2. *a reference signal generator*: generated by the mathematical model proposed by Hogan in [91] to describe voluntary arm movements;

3. *a control strategy*: generates the movement of the avatar in response to that of the patient and is mainly composed of two blocks, "temporal correspondence control" and "reference control". The former is designed to track the motion of the patient minimising the position error between the motion of the avatar and that of the patient, the latter uses the reference signal in order to generate the avatar trajectory with a desired healthy kinematic features. In particular the aim of the reference control is that of reducing the distance between the velocity profile of the avatar and that obtained from the Hogan model.

We will next describe each of these blocks in greater detail.

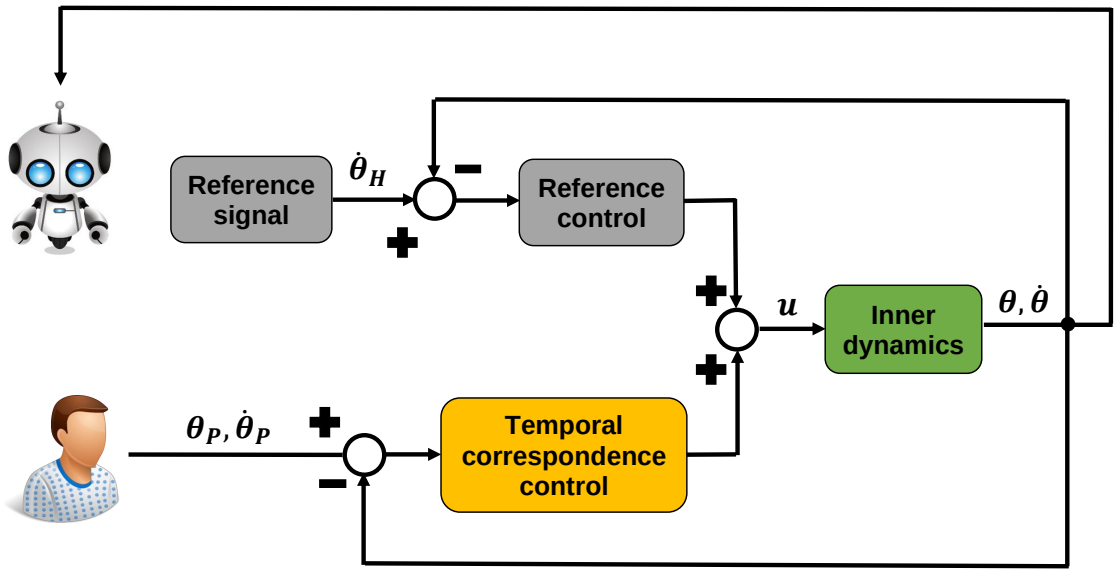


FIGURE 7.5. Cognitive architecture designed for the assisting avatar. The green block models the inner dynamics of the avatar, the yellow block minimises the position error between the avatar and human, whereas the grey blocks confer a desired kinematic features on the virtual agent. $\theta, \dot{\theta}$ are angular position and angular velocity of the avatar, whereas $\theta_P, \dot{\theta}_P$ are measured angular position and estimated angular velocity of the human partner; $\dot{\theta}_H$ is the angular velocity signal reference generated by the Hogan model; u is the control input.

7.4.1 Inner dynamics

The motion of the avatar is modelled through a linear harmonic oscillator describing the angular position of the avatar end-effector. The model is of the form:

$$I\ddot{\theta} + B\dot{\theta} + K\theta = u \qquad 0 \leq \theta \leq \theta_{ROM}, \tag{7.4}$$

where $\theta, \dot{\theta}, \ddot{\theta}$ are forearm angular position, velocity and acceleration respectively. The constant I is the forearm inertia while B and K represent friction and stiffness (evaluated ad-hoc for each patient by the clinicians in the preliminary phase of the rehabilitation). The signal u is the time-varying control input given to the system that models the physical inputs coming from the alpha neurons being responsible for the muscles contraction.

The damped harmonic oscillator was chosen as inner dynamics for its simplicity, which makes such a model particularly appealing in describing the arm motion, as reported in the literature [66, 91, 108, 157]. Any more sophisticated model of human arm dynamics [20, 83, 153] can be chosen without influencing the conclusions drawn in this work.

7.4.2 Reference signal generator

In addition to the measured position and velocity, a reference signal is given in input to the avatar as optimal velocity profile towards which guiding the patient. Such a signal, provided by the Hogan model described in Section 7.2.1, represents the theoretical velocity profile performed by a healthy human while moving the hand from a point to an other.

Taking the time derivative of the $\theta(t)$ defined in (7.3) we obtain the velocity reference signal required by the control strategy, that is:

$$\dot{\theta}_H(t) = \frac{1}{T^5}\left[30\,t^2\left(\theta_f - \theta_0\right)(T-t)^2\right] \qquad 0 \le t \le T, \tag{7.5}$$

where θ_0 and θ_f are the start and final position respectively, and T is the duration of the movement.

7.4.3 Control strategy

The time-varying input to the dynamics in (7.4) are provided by a multi-objective optimal control law evaluated on a sequence of finite time intervals $[t_k, t_{k+1}]$. In each of these windows, the proposed optimal control solution aims at minimising a cost function that depends both on the position mismatch between the avatar and the patient, and on the distance between the velocity profile of the avatar and the desired one. Formally, we formulate the control of the avatar, constrained by the dynamics in (7.4), as the following optimisation problem:

$$\min_{u \in \mathbb{R}} \Phi, \tag{7.6}$$

where

$$\Phi = \frac{1}{2}\alpha_p\left(\theta(t_{k+1}) - \theta_P(t_{k+1})\right)^2 + \frac{1}{2}\int_{t_k}^{t_{k+1}} \alpha_s\left(\dot{\theta}(\theta(\tau)) - \dot{\theta}_H(\theta_H(\tau))\right)^2 + \eta u(\tau)^2\, d\tau, \tag{7.7}$$

with the constraint $\alpha_p, \alpha_s \ge 0$ and $\alpha_p + \alpha_s = 1$. Specifically, θ_P is the angular position measured from the human, $\theta, \dot{\theta}$ are the angular position and angular velocity of the avatar's end-effector, $\theta_H, \dot{\theta}_H$ are the angular position and angular velocity taken as reference from the Hogan model. η is a tunable positive weight to minimise the control energy whereas $[t_k, t_{k+1}]$ is the optimisation interval.

Choice of the parameters α_p and α_s

The constant parameters $\alpha_p, \alpha_s \in [0,1]$ are used to determine how much the avatar takes into account the motion of the patient rather than the reference velocity signal. If $\alpha_p \simeq 0$ the avatar acts as a "blind" leader, ignoring the motion of the patient and only aiming at minimising the velocity error with the reference signal. Viceversa, if $\alpha_p \simeq 1$ the avatar behaves as a perfect follower matching its position with that of the patient. Any value of α_p in between will make

the motion of the avatar a compromise between tracking the human motion and the reference velocity of the Hogan model, allowing to implement different types of leader/follower behaviour.

The proper tuning of these parameters is the basis for a successful rehabilitation strategy. As long as the therapy proceeds, α_p and α_s are required to be tuned in order to adapt the motion of the avatar to the progresses (or regresses) of the patient.

Let us define the index of smoothness J as a quantity measuring the smoothness of the motion of the patient. Formally it is defined as in [92]:

$$J = \frac{T^3}{\left(\max\dot{\theta}\right)^2} \int_0^T \dddot{\theta}^2 \, dt, \tag{7.8}$$

where T is the time duration of the movement.The index J is a measure of the shape of the movement without any dependency on duration and amplitude. Having such an index of smoothness, we can easily define another useful index quantifying the progresses made by the patient, that is the "index of ability" I_A. Such an index is defined as:

$$I_A(n) = \frac{J_H(n)}{J_P(n)} \qquad \in [0,1], \tag{7.9}$$

where J_P is the index of smoothness of the patient in session n evaluated by using (7.8) and J_H is the ideal index of smoothness given by the Hogan model for session n. Under the hypothesis that J_H is the optimal value of smoothness for moving along a certain curve, $I_A = 1$ is the perfect case in which the patient's smoothness is exactly equal to that of the reference motion ($J_P = J_H$), otherwise the greater the distance is of the patient's trajectory from the reference the more the index tends to zero ($I_A \to 0$).

The parameters α_p and α_s are then defined in terms of how much the index I_A varies between two consecutive sessions. Specifically:

$$\begin{cases} \alpha_s(n) = \alpha_s(n-1) + \Delta I_A = \alpha_s(n-1) + [I_A(n-1) - I_A(n-2)] & \in [0,1] \\ \alpha_p(n) = 1 - \alpha_s(n) \end{cases} \tag{7.10}$$

in which the parameter α_s in the session n is increased (or decreased) of a certain quantity proportional to the derivative of the index I_A. The parameter α_p is then set as complementary of α_s.

7.5 Validation

In this section we test the ability of the proposed avatar to adapt its behaviour in different scenarios. Specifically, the avatar is required to lead or follow the human partner adjusting their parameters according to his/her degree of disability. We wish to emphasise here that this is currently a work in progress. Therefore, the test reported in this thesis are preliminary and more accurate validations are currently under investigation.

7.5.1 Synthetic data

To this end, we constructed a series of synthetic motion profiles mimicking a typical point-to-point motion, as done in [92]. The function we used is the sum of three terms: *i)* a linear ramp, *ii)* a rescaled sinusoid with a period equal to the movement duration, and *iii)* a rescaled sinusoid with a period equal to an integer m divisor of the movement duration. Mathematically, we choose:

$$\theta_P(t) = A\left(\frac{t}{T} - \sin\left(\frac{2\pi t}{T}\right)\frac{1-b}{2\pi} - \sin\left(\frac{2m\pi t}{T}\right)\frac{b}{2m\pi}\right), \tag{7.11}$$

where A is the movement amplitude, T is the movement duration, and b is a constant defining the deviation from a smooth cycloidal function. The constant n defines the shape of the movement. Specifically, with $n = 1$ the movement has a single-peaked velocity profile, otherwise with $m > 1$ the movement results in a multi-peaked velocity profile with m peaks. Conceptually, the presence of these peaks in the velocity aims at simulating a series of hesitations in the movement typically exhibited by individuals with impairments.

In Figure 7.6 the movement of a patient under therapy was simulated mathematically modulating function (7.11). Specifically, in the Awing phase the function (7.11) was parameterised with $m = 3$ and $b = 0.8$ in order to simulate two patient's hesitations (trajectory with three peaks). The avatar's parameters in (7.4) were set as $K = 0.64, B = 0.4, I = 0.014$ as done in [91], while the controller was tuned to be a pure follower ($\alpha_p = 1$, $\alpha_s = 0$). In this phase as explained in Section 7.2, the avatar does not influence the patient's movement but only follows it in order to evaluate his/her index of smoothness J and compare it with the index of smoothness J_H computed by the Hogan model along the same trajectory. Quantitatively, with this setting we have $J_P = 11$ and $J_H = 6.58$ resulting in $I_A = 0.6$. In the next phase of Pawing the progresses of the patient were simulated in two exercises. In the first exercise the patient was simulated to be slightly improved setting $m = 2$ and $b = 0.5$ (one hesitation, trajectory with two peaks). According to (7.10) the avatar was set as $\alpha_p = 0.8$ and $\alpha_s = 0.2$ showing a behaviour that is a trade-off between the reference coming from the Hogan model and the patient's motion. Quantitatively, we have $J_P = 8.7$ and so $I_A = 0.75$. Once again, according to (7.10) in the next exercise $\alpha_p = 0.65$ and $\alpha_s = 0.35$, while the patient was simulated with $m = 1$ and $b = 0.1$ (no hesitation, one single peak). In this last exercise, $J_P = 7.35$ and $I_A = 0.9$ simulating the case where a patient has reached a healthy profile.

7.5.2 Preliminary experimental trial

After the mathematical test showed above, a human was asked to experience the rehabilitation platform while simulating the motion of a potential patient suffering of motor impairments (Figure 7.7). The signal acquired by the platform was filtered by means of a zero-phase low pass Butterworth filter with cutoff frequency of 2 Hz. Specifically, in the phase of Awing the human was asked to produce motion with hesitations showing an index of smoothness $J_P = 13.14$ and so

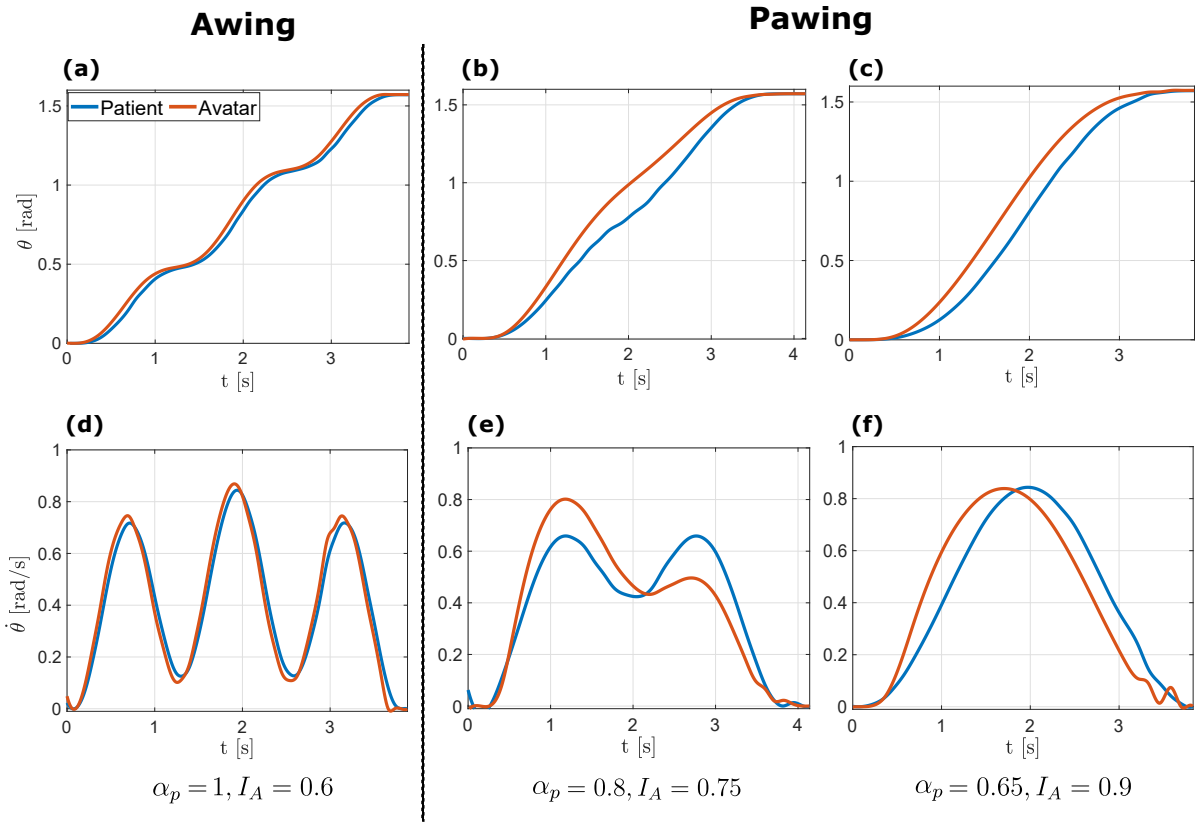


FIGURE 7.6. Mathematical simulation of the rehabilitation process showing the patient (in blue) assisted by the avatar (in red) during the phases of Awing and Pawing. **(a)-(d)** Position and velocity time series of the patient (in red) and the avatar (in blue) are reported for an Awing trial having the maximum angle set at 90 degrees. The patient was simulated to have two hesitations during the task ($m = 3$). **(b)-(e)** Position and velocity time series are reported for a Pawing trial that goes from 0 to 90 degree. The patient was simulated to have only one hesitation ($m = 2$, velocity profile with two peaks). **(c)-(f)** Position and velocity time series are reported for a Pawing trial simulating the patient with no hesitation ($m = 1$, velocity profile with one single peak). Furthermore, for each trial the weight of the avatar's control law α_p and the index of ability I_A are reported showing that while one decreases the other one increases simulating the progresses of the patient.

an index of ability $I_A = 0.5$. In this phase the avatar behaves as a pure follower with $\alpha_p = 1$. Next, an "improved" trajectory was simulated by the same human whereas the avatar was endowed also with a leader component ($\alpha_p = 0.8$, $\alpha_s = 0.2$). Quantitatively, we registered $J_P = 10.65$ and $I_A = 0.62$. Finally, in the last exercise the human performed the exercise naturally resulting in a healthy single-peaked velocity profile with $J_P = 7.78$ ($I_A = 0.85$). The control parameter α_p was set to 0.68 according to (7.10). The avatar was able to adapt its behaviour in different scenarios showing its effectiveness in assisting the patient along the rehabilitation.

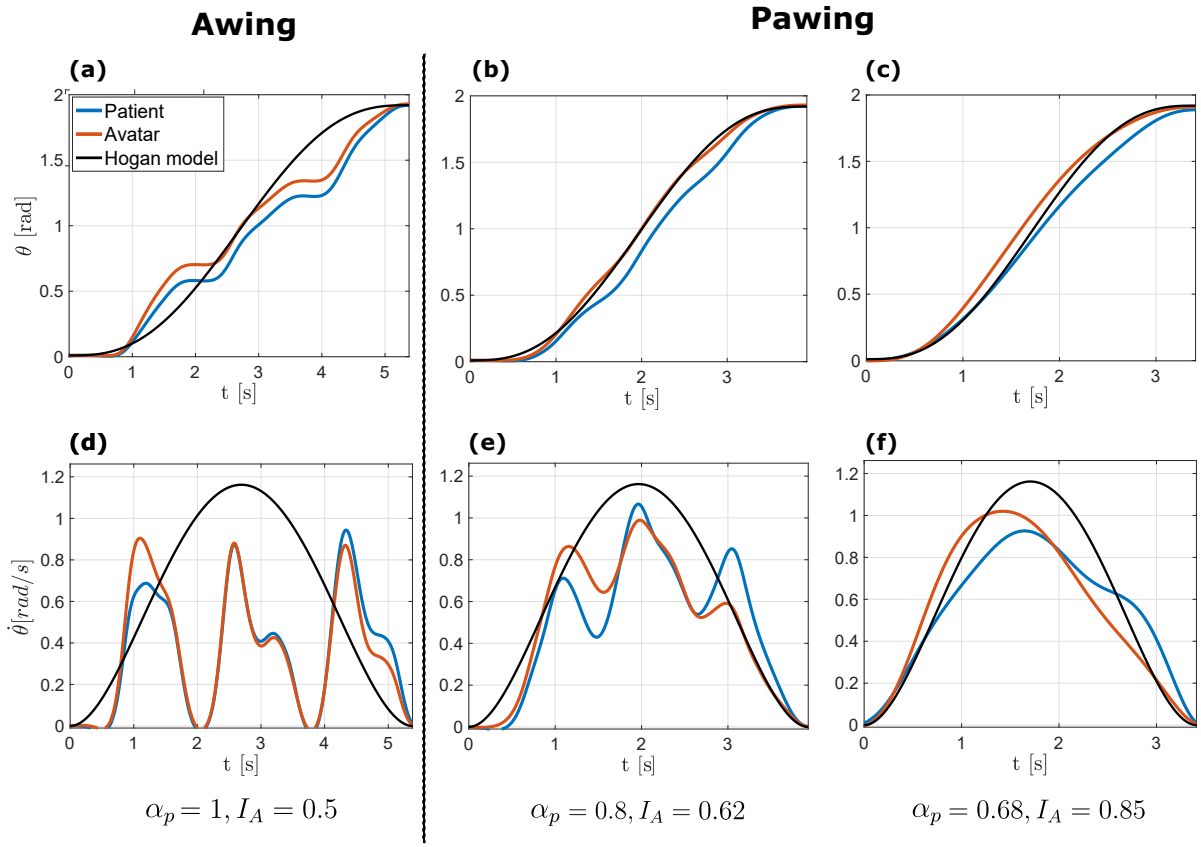


FIGURE 7.7. Experimental simulation of the rehabilitation process performed by a human emulating the motion of a potential patient (in blue) assisted by the avatar (in red). **(a)-(d)** Position and velocity time series of the human (in red), the avatar (in blue) and the reference of the Hogan model (in black) are reported for an Awing trial having the maximum angle set at 110 degrees. The human was asked to stop several times during the trials resulting in a three-peaks velocity profile. **(b)-(e)** Position and velocity time series are reported for a Pawing trial that goes from 0 to 110 degree. The human was asked to move smoother and follow the avatar. **(c)-(f)** Position and velocity time series are reported for a Pawing trial in which the human was asked to perform the task normally following the avatar. Furthermore, the weight of the avatar's control law α_p and the index of ability I_A are reported showing that while one decreases the other one increases as the progresses made by the human.

7.6 Software design of the exergame

The exergame presented in this work was implemented in virtual reality using Unity, a cross-platform game engine used by the world largest game companies [188]. The choice of using virtual reality allows the patient to directly interact with the avatar and the surrounding environment. Moreover, the avatar itself can be visualised with human appearances enhancing in this way

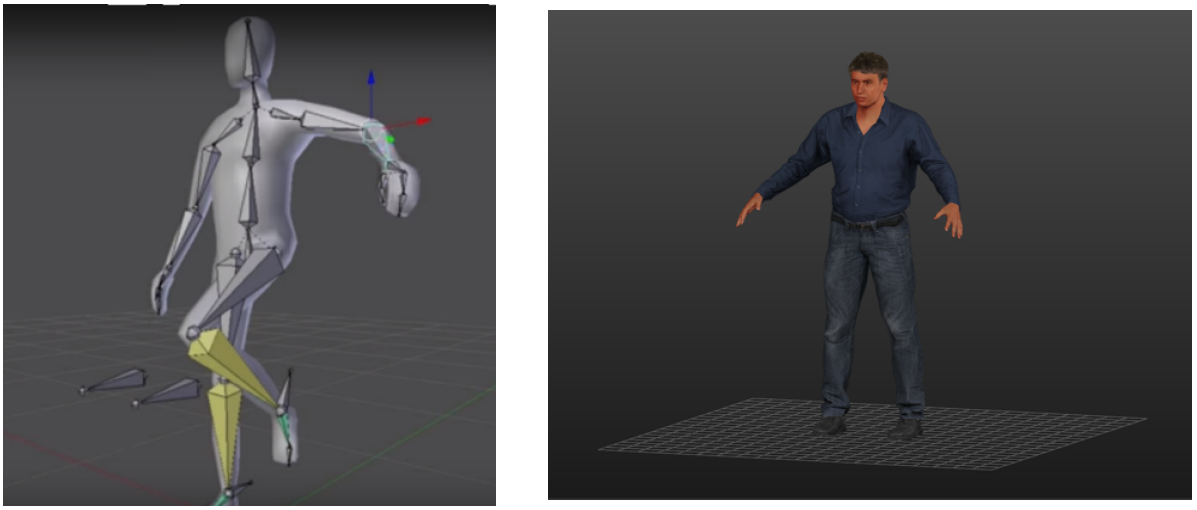


FIGURE 7.8. Prototyping and human rendering of the virtual avatar in Blender and MakeHuman environments.

the social attachment with the patient. Through special devices (e.g., headset) used to simulate our senses, the patient can become active part in this virtual world becoming able to perform any rehabilitation exercise. The main features of the virtual reality implementation are for the environment to be *responsive* and *immersive*. It should be *responsive* because it needs to change in real time when the person explores his/her surroundings or interacts with objects in the virtual environment, and *immersive* in the sense that the user feels his/her presence in a non-physical world (e.g. synchronisation between the sounds and the vision).

We will next describe the main steps carried out to build the software application. The application was designed in collaboration with Mr. Ernesto Erra from the University of Naples "Federico II" and the interface is currently in Italian (English translation is provided in the captions describing the user interface).

7.6.1 Prototyping and rendering

The graphic design of the virtual avatar mainly consists of two steps. In the first step, the avatar's skeleton was build using Blender [1], open source software for modelling and graphic rendering. Specifically, the shape of the avatar was drawn by using the technique of the "mirroring". Such a technique consists in drawing the polygonal mesh of only half character and then duplicating (or mirroring) it around an axis of symmetry (e.g. vertical axis). Each part of the avatar's skeleton can be animated independently in order to make the character perform different types of task. In the second step, the software MakeHuman [2] was used for the human rendering of the avatar. In so doing the physical human like features of the avatar need to be specified, such as age, gender, height, hair colour, dressing, etc. (see Figure 7.8).

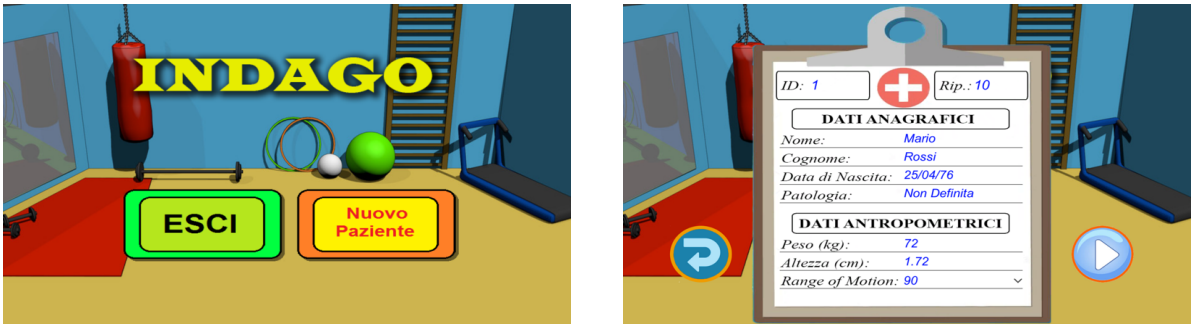


FIGURE 7.9. New patient storage. The patient enters data to appropriately store his/her personal details in the database. *Italian-English panels translation*: esci (exit), nuovo paziente (new patient), nome (firstname), cognome (lastname), data di nascita (date of birth), patologia (illness), peso (weight), altezza (height).

FIGURE 7.10. Sensor calibration. Before to start the exercise, the patient is asked to adjust his/her forearm and wait for the successful calibration of the sensors. Instructions appear directly on the application panel: "please, place your forearm as shown in the figure ensuring to have the elbow fixed on the table". *Italian-English panels translation*: esci (exit).

7.6.2 Use case of the software application

In this section we describe a possible "use case" consisting of a sequence of user's interactions with the software application and reported in what follows.

1) Add a new patient. At the first start of the software application, the patient is asked to write some of his/her personal details in an apposite form, for example firstname, lastname, weight, height, etc. (see Figure 7.9).

2) Setup and sensor calibration. After having clicked on the "Play" button, a panel shows to the patient how to place his/her forearm before to start the exercise (Figure 7.10). Four LEDs

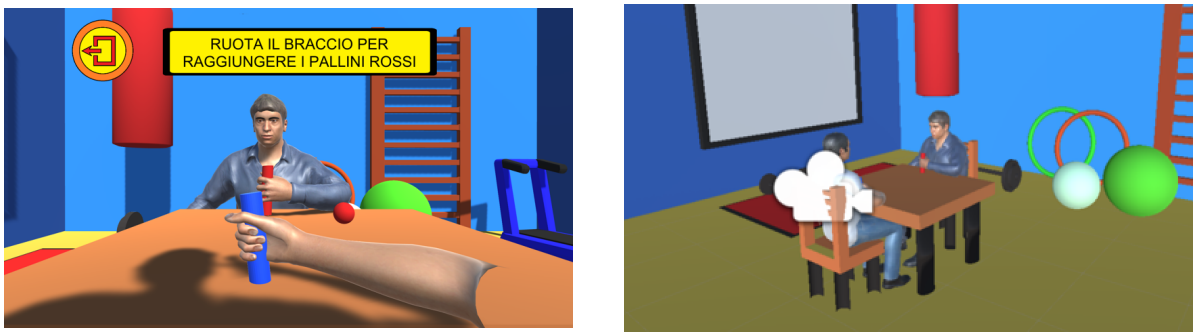


FIGURE 7.11. Gaming screen showing the patient's forearm while holding a blue cylindrical handle. The avatar is seated in front of the patient while holding a red cylindrical handle. Patient and avatar are shown seated at the table. Instructions appear on the screen explaining the task: "please, rotate your forearm to reach the red sphere".

indicate the status of the sensor calibration. Flashing LEDs mean that the calibration is ongoing. Once LEDs are turned off, the patient is ready to start the rehabilitation session.

3) Rehabilitation session. In this step, the patient can see himself/herself seated at the table with the avatar sitting in front of him/her (Figure 7.11). Each rehabilitation session consists of one phase of AWING and p phases of PAWING. In each phase (Awing or Pawing) q different angles are proposed to the patient randomly chosen between zero and his/her range of motion; each of them is repeated three times. The required number p of Pawing phases and the number q of different angles need to be properly set by the clinician. The aim of the reaching task is to move the forearm from the initial position to a target position marked with a red sphere. Instructions on how to perform the task are written on the game screen both in the Awing and in Pawing phase.

4) Feedback system. During the Pawing phase, a feedback system is implemented in real time to encourage the patient to continue with the exercise trying to do his/her best (see Figure 7.12). At the same time, the feedback system is designed to be as least intrusive as possible so that it does not distract the patient from the game. Specifically, a colour-coded smile is shown on the top corner of the game screen according to the level of motor synchronisation measured in real time between the patient and the avatar. A low level of synchronisation is coded with a red smile, a medium level with a yellow smile and a high level with a green smile.

At the end of each Pawing phase, a score is given in feedback to the patient as the average of the scores obtained in each exercise. Such a score is evaluated as percentage of the index of the ability of the patient evaluated according to (7.9). Higher the score is, better the patient has performed the exercise.

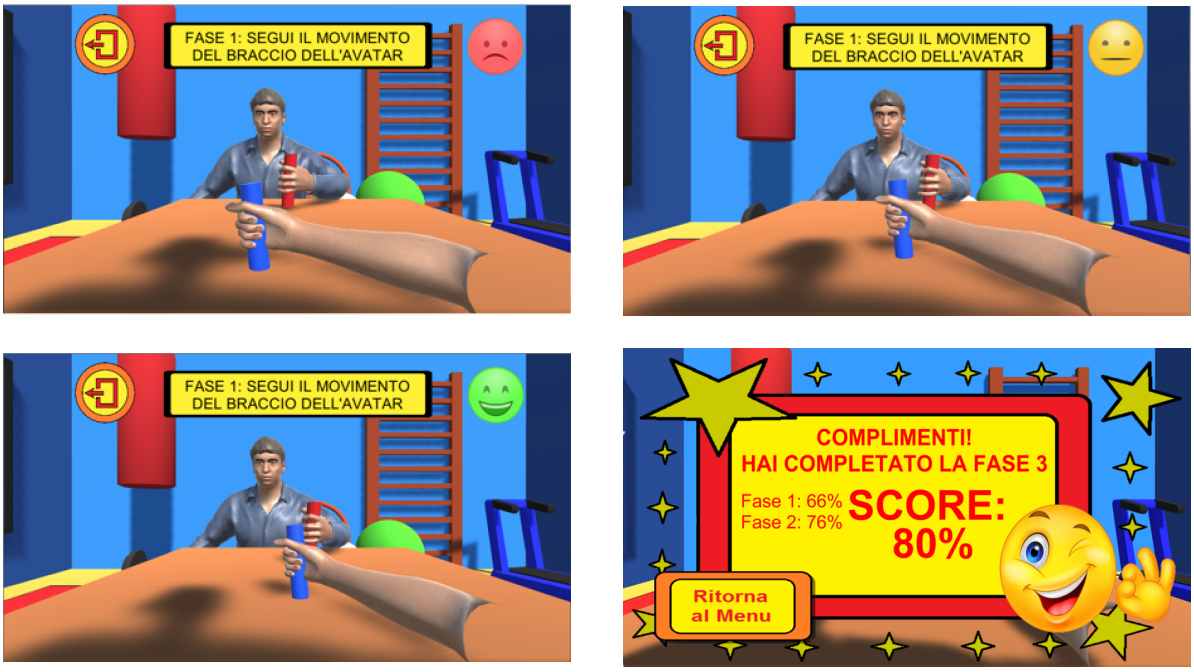


FIGURE 7.12. Feedback system implemented to encourage the patient to perform the exercise with the avatar. During the exercise, a coloured smile appears in the top right corner of the game screen encoding the level of synchronisation between the patient and the avatar: low (red), medium (yellow), high (green). Instructions appear on the screen explaining the task: "please, rotate your forearm following that of the avatar". At the end of the Pawing phase, a percentile score is shown as mean of the score obtained in each exercise. *Italian-English panels translation*: ritorna al menu (home).

7.7 Summary

In this chapter we presented a first version of an exergame developed under the project INDAGO as a rehabilitation platform for post-stroke survivors. Specifically, we highlighted the need of having available computer-aided platforms for the *long-term outpatient rehabilitation phase* helping the patients to carry out rehabilitation exercises comfortably at home assisted by the avatar and remotely monitored by experts.

The main contribution of our rehabilitation platform is the presence of an adaptive avatar able to guide the patient during the whole rehabilitation plan and propose adequate exercises according to his/her level of disability. Reaching motor task was used as motor exercise to propose to the patient during the rehabilitation. The avatar was required to coordinate its movements with that of the patient and lead him/her towards a "healthy" movement profile defined by the Hogan model, used to describe voluntary arm movements.

We presented a preliminary validation of the exergame only focused to test the functionality of the avatar in both phases (AWING and PAWING). We showed that the avatar is able to coordinate its movements during the reaching tasks being a leader that ensures to be followed by the patient. Furthermore, the avatar is also able to adapt its weight parameters automatically according to the index of smoothness J and the index of ability I_A evaluated at the end of each session of rehabilitation. Clinical validation of the exergame involving both healthy people and patients is currently under investigation and part of future work.

CHAPTER 8

FURTHER WORK: LEADERSHIP EMERGENCE IN SMALL WALKING CROWDS

Leadership is a complex yet fascinating phenomenon that plays an essential role in several activities, as for example economics, politics, music and sport. Leadership has been mostly investigated in the case of animal groups, where the collective movement is steered by informed agents in order to locate food sources or avoid predatory attacks. For example in this context, results assess that as group size became larger, only a small proportion of informed individuals is needed to guide the group with a given accuracy. Furthermore, information is transferred from experienced to naive individuals as function of information differences between each single individual with its local neighbours. On the other hand, only a few works exist addressing the problem of leadership emergence in human ensembles, e.g., pedestrian walking, group dance. An open research question in this field seeks to understand the mechanisms emerging in a human group when no leader role is designated a priori.

In this chapter we report the work done in collaboration with Prof. William Warren (Brown University in Providence, Rhode Island) with the aim of investigating the emergence of leadership in human ensembles. In particular, we study the emergence of leadership in a group of people walking together, focusing on the detection of leader-follower relationships that spontaneously are established within the group in order to identify the leading group member.

To do so, we carried out experiments and defined metrics suitable to detect the leader-follower patterns in the scenario of interest, both described in Section 8.1. Then, in Section 8.2 we discuss our preliminary experimental results and shed light on whether the topological structure of the visual pairings among the group members has an effect on leadership, and whether individual locomotion behaviour affects its emergence. Finally, a summary of the main findings and related implications are drawn in Section 8.3.

Additional details on the experimental protocol as well as further experimental results can be found in Appendix B.

8.1 Experimental setup and methods

8.1.1 Task description

Two groups of 4 participants were tested in separated sessions as part of a larger study. The research protocol was approved by Brown University's Institutional Review Board, in accordance with the Declaration of Helsinki. Informed consent was obtained from all participants, who were compensated for their time. All participants had normal or corrected vision, and none reported a motor impairment. To refer to a single pedestrian we use the notation "P$x.y$" where x is the group and y is a sequential number identifying the specific member, for example P1.3 indicates the pedestrian 3 in group 1.

Before each trial, participants were positioned at the vertices of a square configuration marked on the floor ("front left" (FL), "front right" (FR), "back left" (BL), "back right" (BR)), at one of three initial inter-personal distances (IPD of 1, 2, or 4m) (see Figure 8.1). They were instructed to change direction or speed two times as they walked together across the room (about 20m), and to "try to stay together with your neighbours".

Nine different conditions are possible:

- no changes (CC);
- changes in heading direction combining "turn left" and "turn right" in a sequence of two instructions (left-left (LL), right-right (RR), left-right (LR), right-left (RL));
- changes in speed combining "speed up" and "slow down" in a sequence of two instructions (slow-slow (SS), fast-fast (FF), slow-fast (SF), fast-slow (FS)).

The experimental design was thus 2 Conditions (heading direction, speed) × 4 Sequences × 3 IPD (1, 2, 4 m), with one trial in each of the 24 conditions plus three control trials. Data were collapsed across sequence and IPD for analysis, yielding 12 trials per analysed condition.

At the beginning of each trial, a sequence instruction was given, then the group started walking upon a verbal "ready, begin" command. Before the next trial, participants shuffled their positions in the starting configuration. Each group received one practice trial, followed by a block of 12 heading change trials, and a block of 12 speed change plus 3 control trials, for a total of 27 test trials. Trial order was randomised within blocks, and block order was counterbalanced across the two groups. A test session lasted about 40 min. Further details about each single trial are reported in Appendix B.

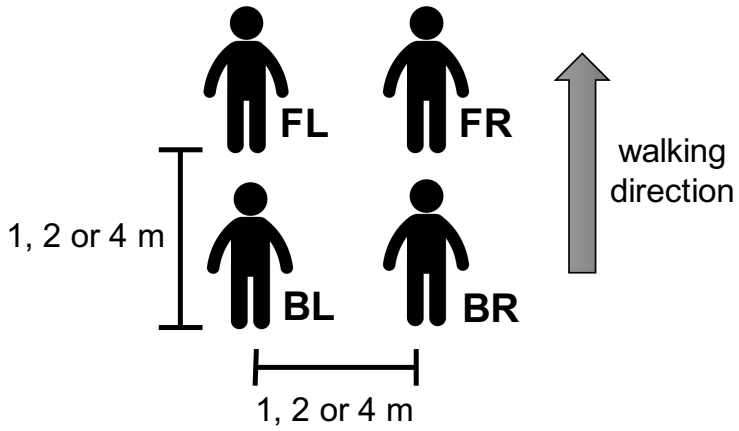


FIGURE 8.1. Group arrangement during the experiments. Participants are placed at the vertices of a square (FL, FR, BL, BR) with an initial inter-personal distance (IPD) between them of 1, 2 or 4 metres.

8.1.2 Data acquisition and analysis

The experiments were carried out at Brown University in a large hall with a 14m × 20m tracking area marked on the floor. Head position was recorded at 60 Hz with a 16-camera infrared-reflective motion capture system (Qualisys Oqus). Each participant wore a bicycle helmet having 5 markers placed on 30 – 40cm stalks in a unique configuration, so that each helmet could be identified.

The time series of head positions of all four pedestrians were recorded in three dimensions, but only two dimensions (xy-plane) were analysed further. Time series were filtered using a forward and backward 4th-order low-pass Butterworth filter to reduce tracker error and gait oscillations, before differentiating to obtain time series of speed and heading. A 1.0 Hz cutoff was used to reduce anterior-posterior oscillations on each step before computing speed, and a 0.6 Hz cutoff for reducing medial-lateral oscillations on each stride before computing heading. Files were truncated to eliminate endpoint error.

Leadership was then evaluated by assessing the difference in heading direction at each time between each pair of pedestrians by adapting to the case of interest the time-dependent delayed direction correlation (C^d), proposed in [76] as a way to characterise the leadership emergence in groups of trawling bats. Summing the percentage of time in which a specific pedestrian is predicted as leader, we were able to reconstruct a network representing the influence that a specific pedestrian has on any other. Similarly, leadership in speed was evaluated by adapting the previous algorithm using time-dependent delayed speed correlation C^s instead.

Other metrics have been proposed in the literature to determine the leadership in a group of interacting agents. Other approaches used, for example, information-theory based tools as causation entropy to infer casual relationship between the group members with the aim of

extracting the underlying interaction network [185]. Such information-theory based tools can be used in a complementary way to the cross-correlation metric in order to extract different information from the data.

8.1.3 Leadership metrics

Time-dependent delay directional correlation (TDDC). Using time-dependent delay directional correlation (TDDC), we characterised leadership in the walking group in terms of the heading direction of each individual, spotting whether some of the participants adjusted their own direction to that of the others and exhibited a follower role. Accordingly, a leader is defined as the participant whose heading direction influences the most that of the others in the sense that the other participants match their headings with that of the leader after a time delay corresponding to the natural human visual-locomotor delay.

Denoting by $v_k \in R^2$ the velocity vector of pedestrian k in the plane, we compute:

$$h_{ij}(t,\tau) = \frac{v_i(t)}{||v_i(t)||} \frac{v_j(t+\tau)}{||v_j(t+\tau)||} \quad \in [-1,1], \tag{8.1}$$

which represents the scalar product between the heading direction of pedestrian i at time t and that of pedestrian j at delayed time $t+\tau$.

Formally, TDDC is defined as follows:

$$C^d_{ij}(t,\tau) = \frac{1}{2\omega+1} \sum_{k=-\omega}^{\omega} h_{ij}(t+k\Delta t,\tau), \tag{8.2}$$

where Δt is the sampling time and $\omega > 0$ is a suitably selected constant. In order to compensate for high-frequency noise and measurement errors, $h_{ij}(t,\tau)$ is averaged over a symmetric time-window $2\omega\Delta t$. As the measured human visual-locomotor delay to the leader's change in heading is approximately equal to 1000 ms and the sampling time for the collected data is equal to $\Delta t = \frac{1}{60Hz}$, we have that $2\omega\Delta t > 1s$ and so ω was set to 40.

To identify the leader, $C^d_{ij}(t,\tau)$ was evaluated for each pair of pedestrians (i,j) at every time instant t along the trial and for different values of delay τ. If the value of $\tau_{ij}(t)$ maximising the heading direction correlation is positive then we say that pedestrian i leads pedestrian j at time instant t, or otherwise if that value is negative.

A left-right trial carried out by group 1 with an initial inter-personal distance of 4m is taken as a representative example to illustrate the analysis based on time-dependent directional correlation metric. Figure 8.2(d) shows the heat maps characterising values of $C^d_{i,j}(t,\tau)$ for the whole duration of the trial with different time delay τ between each pair $(i,\ j)$ of pedestrians. Different colours are used to indicate different values of TDDC, with blue representing $C^d_{ij}(t,\tau) \leq 0.5$ and red representing $C^d_{ij}(t,\tau) = 1$. Dashed white lines represent level curves where $C^d_{ij}(t,\tau) = 0.95$, whereas the dashed black line represents $\tau(t) = 0$. The solid green line on top of the heat map represents the value of $\tau(t)$ maximising the heading direction correlation between two pedestrians.

Time-dependent delay speed correlation (TDSC). To assess leadership roles in speed-change trials, we define:

$$s_{ij}(t,\tau) = \big|\, ||v_i(t)|| - ||v_j(t+\tau)|| \,\big| \tag{8.3}$$

which represents the difference between the speed of pedestrian i at time t and that of pedestrian j at delayed time $t+\tau$. The time-dependent speed correlation (TDSC) is defined as:

$$C^s_{ij}(t,\tau) = \frac{1}{2\omega+1} \sum_{k=-\omega}^{\omega} s_{ij}(t+k\Delta t,\tau), \tag{8.4}$$

where the parameter ω is the length of the symmetric window used to average $s_{ij}(t,\tau)$ in order to compensate for noise and measurements errors as done before. As the measured human visual-locomotor delay to the leader's change in speed is approximately equal to 500 ms, we have set ω to 20. In this case pedestrian i is said to lead pedestrian j at time t if the value of $\tau_{ij}(t)$ minimising the speed correlation is positive, or otherwise if it is negative.

A fast-slow trial carried out by group 1 with an initial inter-personal distance of 2m is taken as a representative example to illustrate the analysis of the data based on time-dependent speed correlation metric. Figure 8.3(d) shows the heat maps characterising values of $C^s_{i,j}(t,\tau)$ for the whole duration of the trial with different time delay τ between each pair $(i,\ j)$ of pedestrians. Different colours are used to indicate different values of TDSC, with blue representing $C^s_{ij}(t,\tau) = 0$ and red representing $C^s_{ij}(t,\tau) \geq 0.5$. Dashed white lines represent level curves where $C^s_{ij}(t,\tau) = 0.05$, whereas the dashed black line represents $\tau(t) = 0$. The solid green line on top of the heat map represents the value of $\tau(t)$ minimising the speed correlation between two pedestrians.

Individual Percentile leadership index. Knowledge of $\tau_{ij}(t)$, maximising/minimising the heading direction/speed correlation, for each trial and pair of pedestrians allows to quantitatively estimate how much a participant acts as a leader in the group. Specifically we define the *individual percentile leadership index* for each participant as the overall average percentage of time for which the computed $\tau_{ij}(t)$ is positive (Figure 8.2(e) and Figure 8.3(e)).

Network reconstruction. Knowledge of $\tau_{ij}(t)$ allows also to infer leader-follower interactions between each pair of group members. Each pedestrian is represented as a node in the network and the influence that each pedestrian has on the others is represented as an edge (an edge directed from node i to node j means that pedestrian j is influenced by pedestrian i). The strength of this influence is captured by the edge weight $w_{i,j} \in [0,1]$ representing the normalised percentage of time in which pedestrian i leads pedestrian j. The network were reconstructed repeatedly on different time windows in order to evaluate also the evolution of the interactions over time. A total of 5 consecutive time windows were considered (Figure 8.2(f) and Figure 8.3(f)).

The topologies inferred were then refined by means of DPI techniques (data processing inequality) and by a further thresholding to remove false positive links between nodes, as

suggested in [9]. Specifically: i) link whose weight $w_{i,j}$ is null are discarded from the reconstructed network, and ii) non-zero weights $w_{i,j}$ are checked among all triplets of connected nodes so that, on the basis of their intensity, one of the triplet can be possibly regarded as a false connection and set to zero (so that the corresponding link is removed). For example, let us consider three nodes i, j and k. If the weight $w_{i,k}$ between the nodes i and k is lower than $w_{i,j}$ between the nodes i and j, and is also lower than $w_{j,k}$ between the nodes j and k, then the weight $w_{i,k}$ is set to zero and the link between the two nodes i and k is removed.

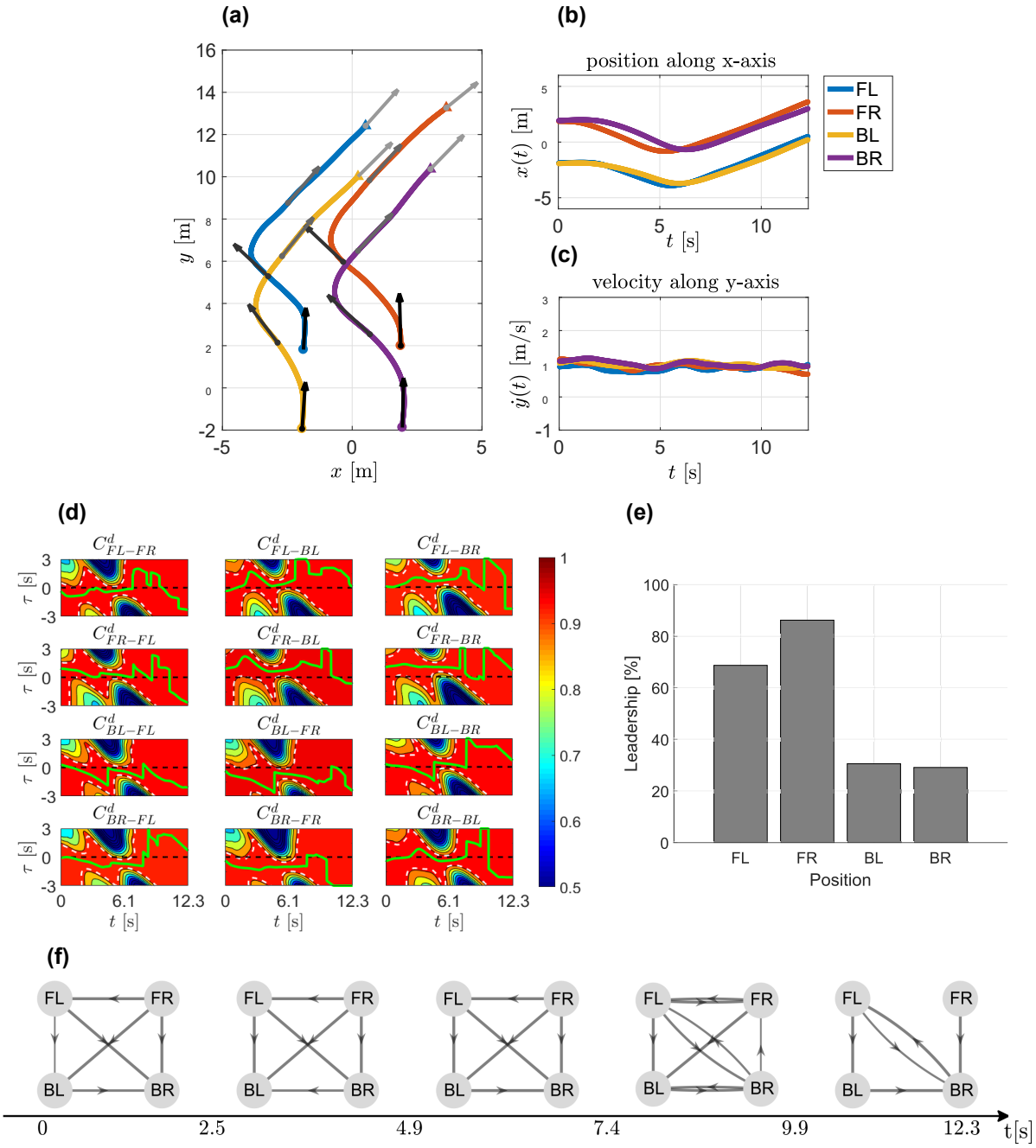


FIGURE 8.2. **Left-right representative trial**. **(a)** Pedestrians' trajectories in the xy-plane with different colours. Velocity vectors are represented by an arrow each 3 seconds, the colour goes from black to grey as time increases.**(b-c)** Position along x-axis and velocity along y-axis. Pedestrians are identified by his/her position (FL, FR, BL, BR). **(d)** TDDC evaluated for each pair of pedestrians. The x-axis represents time instant t, whereas the y-axis represents different values of the time-delay τ. **(e)** Percentile leadership. Each grey bar refers to a different member in the group. **(f)** Network reconstruction on 5 different time windows.

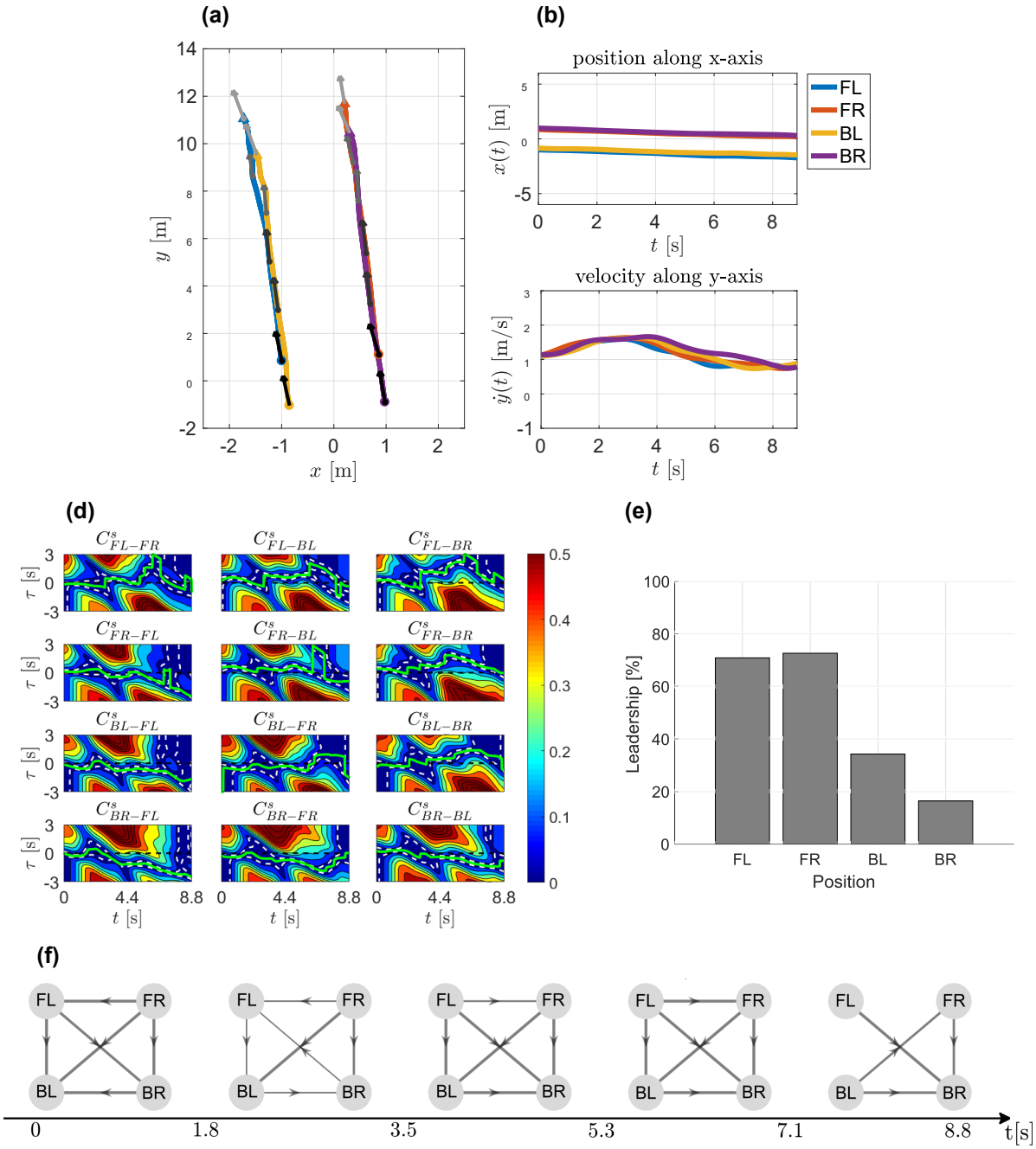


FIGURE 8.3. **Fast-slow representative trial (a)** Pedestrians' trajectories in the xy-plane with different colours. Velocity vectors are represented by an arrow each 3 seconds, the colour goes from black to grey as time increases.**(b-c)** Position along x-axis and velocity along y-axis. Pedestrians are identified by his/her position (FL, FR, BL, BR). **(d)** TDSC evaluated for each pair of pedestrians. The x-axis represents time instant t, whereas the y-axis represents different values of the time-delay τ. **(e)** Percentile leadership. Each grey bar refers to a different member in the group. **(f)** Network reconstruction on 5 different time windows.

8.2 Results

To investigate leadership emergence in walking groups, we consider the simplest scenario in which four participants were asked to walk together across the room. Our aim is to unveil whether, when no designated roles are assigned, a leader spontaneously emerges in the group. Specifically, we adopt a time-dependent cross correlation analysis to unfold the leader-follower relationships established within the group by instructed all participants to change their heading direction or speed twice during every experiment. We then evaluating the *individual percentile leadership* index (formally defined in Section 8.1.3) to identify the group member whose change of locomotion (heading direction or speed) is readily followed by other group members for the most of the time. Such a participant is then labelled as the group leader.

In what follows, the data were analysed to assess the effect on leadership emergence of (i) pedestrian position in the group and (ii) individual locomotion behaviour.

8.2.1 Effect of pedestrian position in the group

Because of positional layout used during the experiments, we expected one of the pedestrians in the front row to act as the leader due to the fact that they cannot be visually coupled to the participants in the back row.

Figure 8.4 shows the influence on the *individual percentile leadership* index of the positions of the participants, averaged over all trials, in the two conditions (heading direction and speed). Results effectively confirm that participants in the front left and front right positions have the highest value of the leadership index.

This is also confirmed by the formal statistical analysis through ANOVA. In particular a 2(Condition) × 4(Position) × 3(IPD) Mixed ANOVA was performed considering the factors Condition and IPD as "between-subjects factors", and the factor Position as "within-subjects factor". The additional factor Group was considered as random effect. The analysis revealed a significant main effect of Position on the measured leadership ($F(3,117) = 19.865, p < 0.05, \eta^2 = 0.337$).

The Condition main effect and the Condition × Position interaction were not in themselves significant, having respectively $F(1,39) = 1.334, p = 0.255, \eta^2 = 0.033$ and $F(3,117) = 0.196, p = 0.899, \eta^2 = 0.005$. Furthermore, neither the IPD main effect ($F(2,39) = 0.133, p = 0.876, \eta^2 = 0.007$) nor the IPD × Position interaction ($F(6,117) = 0.906, p = 0.493, \eta^2 = 0.044$) were significant. Finally, the Condition × IPD × Position interaction reported $F(6,117) = 1.301, p = 0.262, \eta^2 = 0.063$. This shows that the factor Position alone dominates percentile leadership.

Since the factor Position was statistically significant, we performed Bonferroni post-hoc test for a pairwise comparison in order to understand in details which positions contribute to the leadership emergence. We found that the differences between the front positions and the back positions are statistically significant ($p < 0.05$) while the differences within the two positions in

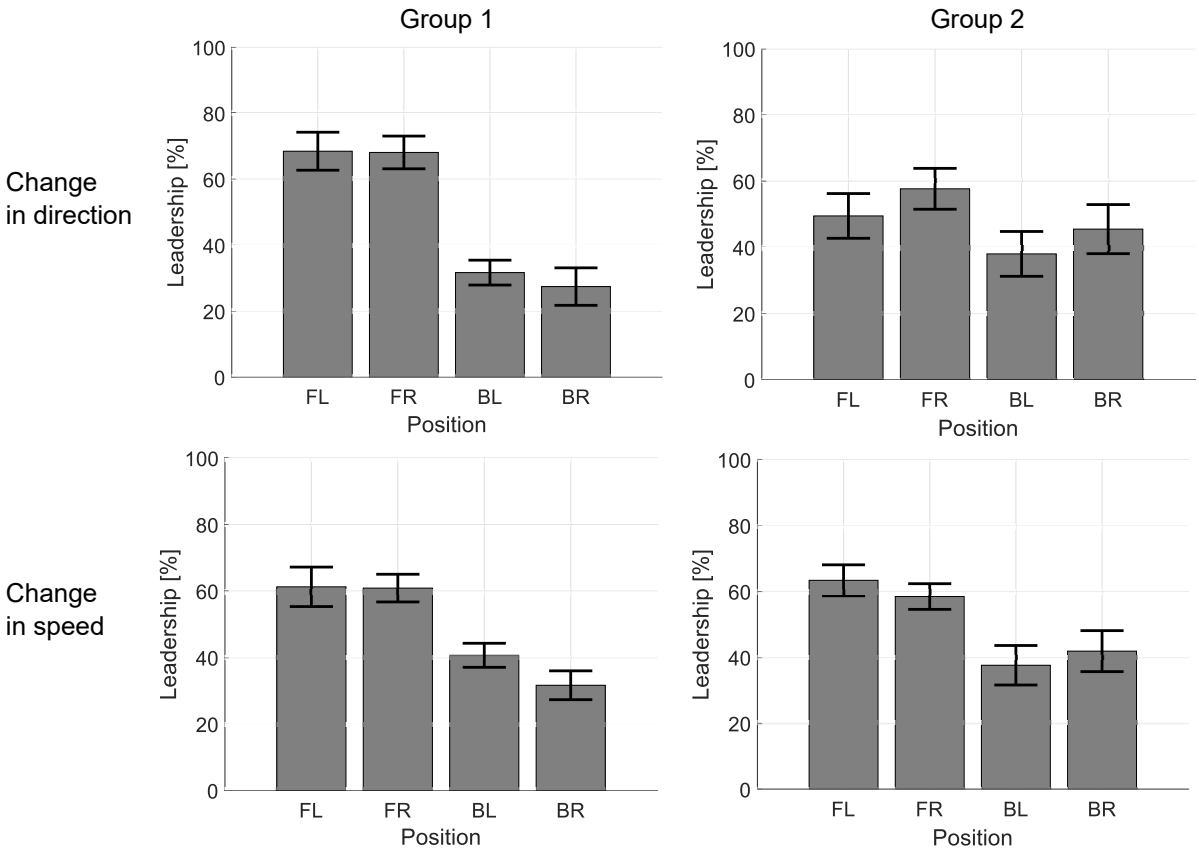


FIGURE 8.4. Four bar charts are reported, 2 groups × 2 conditions (speed and heading direction). Each height's bar represents the mean value of the leadership grouped by position over the number of trials, whereas grey error bars are the standard error of the mean. FL = front left, FR = front right, BL = back left, BR = back right.

the front and within the two in the back have a $p = 1$ and so not statistically significant.

Our findings confirm the observations made in [46, 160, 162] that leadership depends on local visual couplings between group members that try to match their speed and heading direction accordingly.

8.2.2 Effect of individual locomotion behaviour

The same data are replotted by participant in Figure 8.5, which shows the mean value and standard error of the mean (SEM) of the individual leadership index for each pedestrian, where we noticed some participants to have consistently higher indices than the others. As we were interested in assessing the pedestrians individually, we performed a statistical ANOVA analysis on both groups separately. The values of the *individual percentile leadership* index were evaluated and averaged over different trials grouping them by specific participant. A one-way ANOVA was

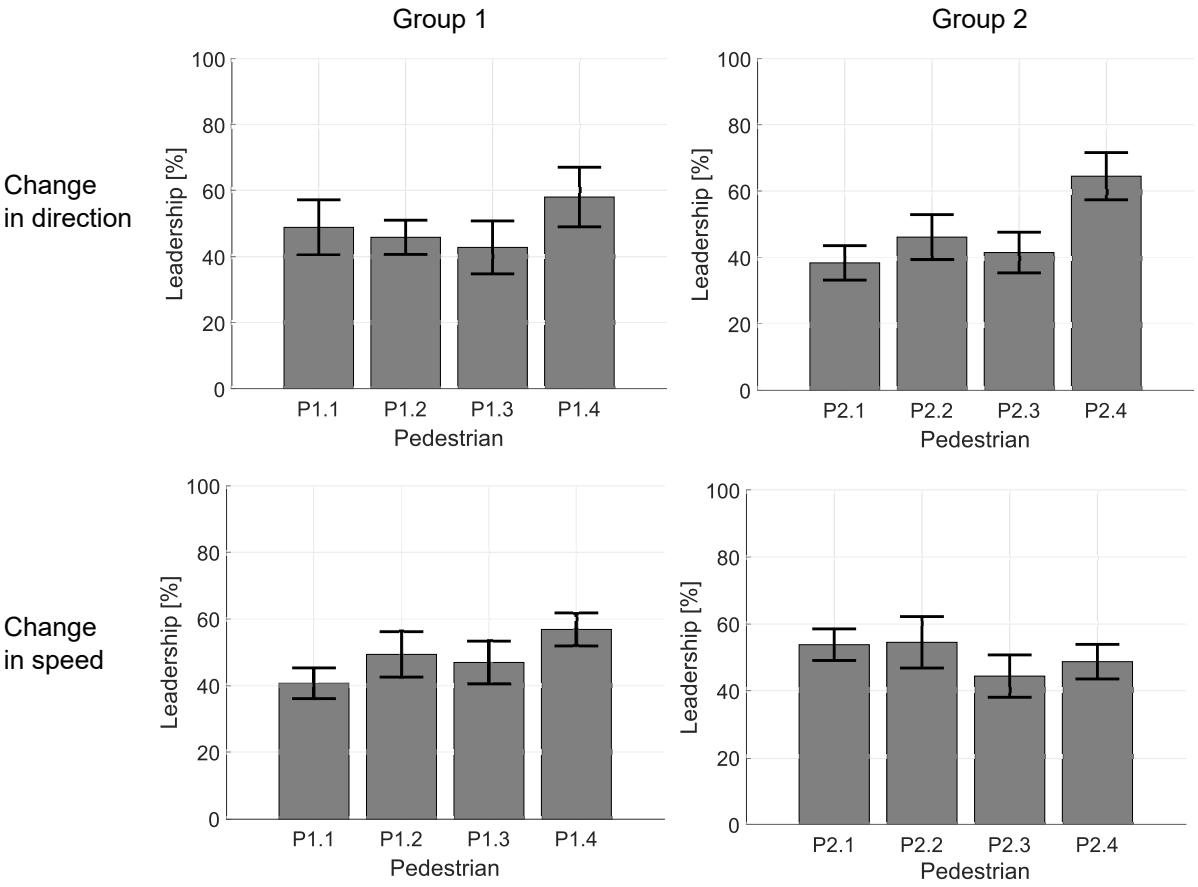


FIGURE 8.5. Four bar charts are reported, 2 groups × 2 conditions (speed and heading direction). Each height's bar represents the mean value of the leadership grouped by a specific pedestrian over the number of trials independently from his/her position, whereas grey error bars are the standard error of the mean.

performed on each group revealing that the factor Pedestrian has a main effect only in group 2 for direction-change trials ($F(3,24) = 4.663, p < 0.05, \eta^2 = 0.368$).

In these trials, pedestrian P2.4 became the leader independently of his/her position in the group as shown in Figure 8.6 (details about leader-follower patterns inferred from the data are reported in Appendix B. This was confirmed by conducting a Bonferroni post-hoc test to perform a pairwise comparison. The result shows that the behaviour of pedestrian P2.4 has a significant influence on the behaviour of the whole group ($p < 0.05$). The same behaviour is not present in the speed-change condition.

This discrepancy is probably due to the fact that in the case of direction-change trials pedestrian P2.4 moved into the peripheral field of view of all the others while in the speed-change trials the same pedestrian, when located in the back positions, remains behind the front pedestrians.

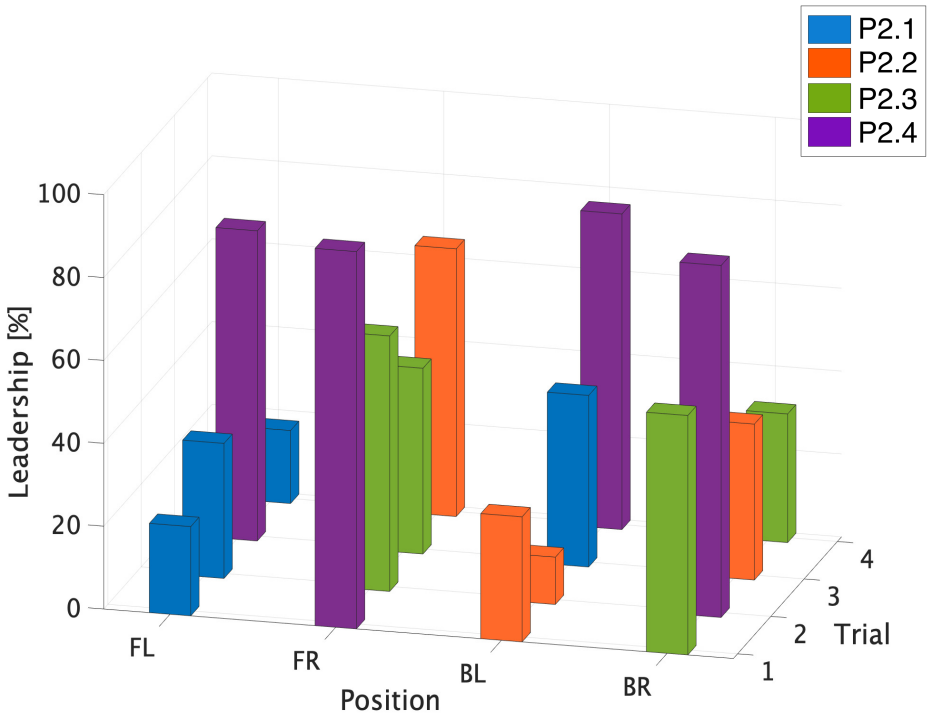


FIGURE 8.6. Percentile leadership of four sample trials are reported having the pedestrian P2.4 (in purple) in each of position. The x-axis represents the trial, the y-axis the position (Front Left, Front Right, Back Left, Back Right), whereas the z-axis reports the value of leadership in percentage. Each pedestrian is identified by different colour.

Our data shows that in some situations individual locomotion behaviour of a participant can be so influential as to dominate over the group independently of the visual interaction pattern existing with the others. By deciding to always turn first, such a participant exploited the mechanism of collective motion - in which participants match the heading of neighbours in their field of view [161] - to seize the leadership (*self-elected leader*).

8.3 Summary

In this chapter we tackled the problem of leadership emergence in human ensembles trying to unveil whether, when no designed roles are assigned, a leader spontaneously emerges in the group. In so doing, we considered the simplest scenario in which four participants were asked to walk side by side.

Our preliminary analysis showed that the combination of the positions of the pedestrians in the group and their individual behaviour contribute in a statistically significant manner to define the leadership role. In particular, our findings showed, as expected, that participants walking in

the front row, which were not visually coupled to those walking behind, were usually emergent as leaders. Counterintuitively, though, we also found that in a significant number of trials the same participant assumed a leadership role initiating a group change of direction irrespective of his/her position in the group.

These findings demonstrate that leadership emergence in walking groups is a complex phenomenon that does not only depend on geometric conditions, such as the leader location in the group, but also on other features related to the individual locomotor behaviour of the participants; for example, the initiation of sudden movements or changes of direction of some individual that when perceived by the neighbours propagate to the rest of the group. This observation might be relevant to understand how coordination in walking groups emerges in those situations, such as evacuations or some sport scenarios, where verbal communication among group members is hidden.

The methodology we adopted in this work is general enough to be adapted to other scenarios and larger walking groups. We wish to emphasise that it can also be used to obtain information on the specific leader-follower relationships between group members and their changes over time. This again can be an extremely useful information to possess in those cases where it is desirable to influence the group dynamics, as for example when guiding the group towards a safety exit in emergency situations.

CHAPTER 9

CONCLUSIONS

In this thesis, the design and control of an artificial agent able to perform coordination task in human-robot or human-avatar interactions has been discussed. It has been shown that the availability of such an artificial agent can be extremely important in many common applications (e.g., assistive robots for the elderly population, motor rehabilitation, educational and entertainment purposes). The work presented addressed the following set of key research issues:

1. Understand how the individual kinematic features characterising the human behaviour in motor joint tasks can be captured by a mathematical framework.
2. Investigate how to transfer such kinematic features in the cognitive architecture of an artificial agent in order to be able to reproduce them autonomously while interacting with humans.
3. Develop control architectures enabling the virtual player to learn how to interact with the human partner in dyadic scenario directly from data (in-silico and experimentally).
4. Extend the control architecture of the virtual player from dyadic to group interaction in the multi-player version of the mirror game.
5. Apply the architecture to develop an exergame for motor rehabilitation of stroke patients.

9.1 Summary of the main results

After introducing the thesis motivation and the key challenges in Chapter 1, a detailed literature review was given in Chapter 2 and Chapter 3. In particular, in Chapter 2 the main results

were reviewed on human synchronisation and motor coordination both in dyadic interpersonal coordination and in human ensemble underlying the need of having an autonomous artificial agent able to coordinate its movements with human partners. In Chapter 3 we presented the problem of human-robot interaction giving a broad overview of the solutions and the tasks addressed in the literature and focusing on assistive robots in joint motor tasks with humans.

We began our investigation on the design of the virtual agent in Chapter 4. Specifically we presented a cognitive architecture based on a controlled nonlinear oscillator capable to drive the artificial agent in the dyadic task of the mirror game both in the leader-follower and in the joint improvisation condition. The main contribution of this chapter is the introduction of Markov chains as a mathematical framework to capture and model the individual kinematic features of human behaviour from experimental data. Such a mathematical framework was exploited as a signature generator and embedded in the nonlinear architecture of the virtual player in order to produce artificial motion with a desired IMS. Results obtained by making the MC-based virtual player play game sessions with humans were discussed. One of the main advantages of such a strategy lies in the resulting reference generated motion, which both reproduces the human-like IMS feature and results in an not pre-programmed/recorded human reference. Indeed, the generated motion also reproduces the randomness and the non-perfect repetitiveness similar to that of a human.

Reinforcement learning techniques, and in particular the Q-learning algorithm, was introduced in Chapter 5 as a solution to the main drawbacks of the nonlinear control architecture discussed in Chapter 4, e.g., the lack of an explicit mathematical model capturing the mechanisms underlying the emergence of human coordination, and the need for carefully tuning the nonlinear model parameters in order to make the virtual player move human-like. Such an approach allowed us to design a new artificial agent (called cyber player) capable to learn how to move in a human-like fashion directly observing data coming from humans while performing the motor coordination task of interest. In this chapter, we solved the problem of the learning approach requiring a large amount of training data, by synthesising one or more "virtual trainers" able to generate as much synthetic data as needed during the training phase. Such VTs were driven by the model-based control architecture from previous literature aptly modified by embedding in its core the IMS generator described and discussed in Chapter 4. The use of the virtual trainer had a dual advantage: *i)* only a small dataset of solo sessions performed by humans (e.g., 30 trials) was required to embed the signature generator in the model-based artificial avatar and generate much larger synthetic datasets; *ii)* by tuning few virtual trainers in order to emulate the behaviour of a set of human players, the cyber player was trained to be general enough to play the mirror game with any human partner both in the leader and in the follower condition.

Motor coordination in a group was tackled in Chapter 6. Therein, we investigated the problem of integrating the AI-driven artificial agent in a group performing the multi-player version of the mirror game. We found that the Deep Q-network algorithm was suitable to address such a

problem with a high accuracy. The CP was validated both in-silico and by performing experiments involving groups of 4 human players connected with each other through different topological patterns. The experiments were carried out through CHRONOS [12, 115], a software developed by us to study both human-human and human-avatar coordination.

We concluded our investigation on assistive avatars implementing the virtual player in a real application. Specifically, in Chapter 7 we presented a simple exergame designed for the rehabilitation of people suffering from post-stroke impairments. The novelty of our approach relied on the presence of an adaptive virtual avatar that supervises the patient along the rehabilitation therapy proposing personalised motor coordination exercises. The preliminary validation showed the potential impact of the exergame on the design of innovative rehabilitation strategies that allows the patient to carry out assisted rehabilitation exercises comfortably at home. The activity is a work in progress and an extended validation needs to be carried out performing massive campaigns both with healthy people and patients in order to assess its effectiveness.

Finally, Chapter 8 was dedicated to the study of leadership emergence in small walking groups. The leadership is a phenomenon strictly correlated with the presence of coordination established in a human group performing a common activity.

Our analysis showed that using time-lag method can unfold leadership emergence in human group. We found that the combination of the positions of the pedestrians in the group and their individual locomotor behaviour contributed in a decisive manner to define the leadership role. In particular, participants walking in the front row were usually emergent leaders due to the fact that the visual coupling only connects the back row with the front row. In this context, we highlighted also the importance of the locomotor behaviour characterising each pedestrian leading often to a self-elected leader. This attitude to anticipate and lead the movements of others was observed to be independent of the position in the group.

9.2 Possible applications and future work

The availability of an autonomous artificial agent able to perform coordination task with humans is of crucial importance in the context of human-machine interaction. Such virtual agents provide a novel and promising tool to be used in the human dynamic clamp setting proposed in [59] as a method to study social interaction and movement coordination among humans. In particular, virtual agents allow us to change their exhibited kinematic features at will becoming in this way an essential tool to explore the important principle in social psychology suggesting that behavioural similarity between people enhances their coordination and social rapport [67, 179].

Other than allowing to investigate human movement coordination and its relationship with socio-psychological factors, the cyber player designed in this thesis can be effectively used to implement the rehabilitation strategies for patients affected by motor and social disorders that

were suggested in [21, 178]. Therein, it is suggested that for rehabilitation purposes the patient should be made to play the mirror game with an artificial avatar that starts by playing sessions of the game while exhibiting the same IMS of the patient and is then gradually made to change its IMS to habituate the patient to coordinate its motion with players moving in a different way from her/him. Our approach allows to easily realise such an avatar. Indeed, the patient IMS could be captured by our model-based Markov Chain approach. The resulting IMS generator can then be used to synthesise virtual trainers able to train the cyber player to play the game as the patient during the initial sessions. Further VTs with dissimilar IMS from those of the patient can then be used to re-train the CP for later sessions of the rehabilitation exercise implementing the vision first proposed in [21, 178, 179] where the IMS of the avatar is changed from that of the patient during the first mirror game sessions to become that of a healthy individual towards the end of the rehabilitation [21, 59]. In this way, the patient is habituated over successive trials of the game with the CP to coordinate his/her motion better and better before moving to trials involving other humans. The potential application of the CP in a clinical setting is beyond the scope of this thesis, which is focused on the development of the IMS generator, the virtual trainers and the RL based CP.

Extending the control of the cyber player from dyadic to a group setting is useful not only to implement group rehabilitation activities (proved to be apt to enhance the effectiveness of the rehabilitation if performed together with individual programs [197]) but also to unveil the mechanisms underlying group coordination. Understanding how different topological patterns can change the quality of coordination reached by the group and how the dynamics of each member of the group influences the others will allow us to introduce one or more avatars in the network in order to entrain a subset of the members in the ensemble to promote certain collective dynamics. Formally, in control theory this strategy is the so called *pinned control strategy* [36, 50], according to which a network is forced to achieve a desired synchronisation state by directly influencing only a fraction of nodes properly selected from the network.

Other scenarios where the design of artificial agents leads to the improvement of humans' quality of life include: industrial applications, where robots and human operators are required to work in cooperation in close proximity in order to maximise performance and safety; human-robot teams in search and rescue operations, e.g., after urban or natural disasters. Also, the ability of the artificial agent to be an adaptive leader can be exploited in educational robotics to help children in developing abilities that promote autonomy, problem solving skills and creativity.

Extensions and improvements of the work presented in this thesis include the possibility of extending our study to a more complex task (so far we have considered a simple and mono-dimensional oscillatory task). For example, the task can be extended to a multi DOF problem where the position of the player end-effector is a point $[x, y, z]$ in 3D Euclidean space.

In addition, social coupling between two individuals can emerge in collaborative tasks under different modes of perceptual information (e.g., auditive, tactile). While the mechanism of visual

coupling established between humans has been investigated in many existing works, the effects of auditory interaction have been rarely explored. Starting from that, a further improvement will be that of building an autonomous artificial agent that has also an auditory feedback besides the visual one. Auditory and visual information might function in a complementary fashion to support each other in order to make the agent able to perceive the surrounding environment through vision and sound.

For group scenarios, different paths can be followed to improve the architecture of the cyber player presented in this thesis. For example, in Chapter 6 the DQN implemented in the cyber player takes as input the average of the positions of its neighbours. It is clear that this is a simplification (although reasonable) and that much more complex mechanisms can be hidden in human cognition when engaged in different scenarios. A solution could be using learning techniques to make the cyber player extract the main motor features directly from the raw data taking the whole state of its neighbours. The main drawback of these problems is that the state space may became quickly too large and be dependent on the number of the neighbours (not scalable and not general). Future studies may be so performed in this direction.

Finally, CHRONOS can be extended into an on-line environment in order to enable a higher number of experiments and connect in a faster way a larger group of people in game sessions over the Internet. This will allow to have a distributed coordination facility without requiring any additional space or experimental burden.

The clinical trial of the exergame proposed in Chapter 7 is left to future work as it requires access to patients. In addition, several improvements and extensions can be implemented in the rehabilitation platform. For example, the Range of Motion, currently evaluated by the clinicians in the preliminary phase of the therapy, could be included in the exergame proposing to the patient different target angles to reach and evaluating the maximum angular excursion supported by him/her. Furthermore, techniques of system identification can be also implemented to evaluate the physical and mechanical parameters characterising the motion of the human forearm (e.g., stiffness, inertia and friction). Such parameters can be directly embedded in the dynamics of the avatar in order to have a model tuned on the patient.

APPENDIX A

MARKOV CHAIN METHODOLOGY TO BUILD A CLASSIFIER

In this appendix, whose content has been published in [116], we exploit the Markov Chain methodology (described in Chapter 4) to classify different Individual Motor Signatures according to some properties. This is particularly relevant as IMS anomalies are exploited to diagnose patients affected by schizophrenia providing new biomarkers for this mental disorder. In particular, a classification methodology was presented in [178] which was based on the characterisation of neuromotor markers extracted from the recordings of participants spontaneous hand motion while playing the mirror game in solo condition. Three different features were extracted from the collected data and, a posteriori, a majority rule was applied by the classifier to take the diagnostic decision. Even though this proposed scheme is innovative for the use of biomarkers in comparison with the previous literature [78, 200], it still suffers from several limitations, such as the possibility of providing reliable online classification. In this appendix we propose a classification methodology based on Markov chains able to overcome such drawback, and reach also a better accuracy than that presented in [178].

Specifically, in Section A.1 we describe the design of the classifier to distinguish between different individuals by analysing their IMS and diagnose if an individual suffers from mental disorders such as schizophrenia. Then, in Section A.2 with the aim of making a good comparison we validate our approach on the same human motion data in solo condition adopted in [178].

A.1 Classifier design

Given a dataset, it is possible to divide it into different classes based on some measures of inherent similarity or distance. The classification procedure of a new observation is the problem of identifying to which of these classes it belongs, based on "how similar" this new observation is

with respect to the members of different classes. A *classifier* is the entity that implements such a classification process and assigns the data to a class.

In this appendix, we exploit the Markov Chain methodology (described in Chapter 4) to build an IMS classifier by tuning and assigning a dedicated MC model to each class we are interested in. Specifically, we evaluate the problem of recognising healthy from schizophrenic players by means of their signature profile. A total of two Markov chain models are needed, one representing the class *control* made of all signatures coming from healthy people, and the other representing the class *patient* made of all signatures coming from schizophrenic people.

Figure A.1 depicts the block diagram of the classifier for control/patient recognition. The classifier takes as input one or more position time series of a certain player. For each of them, it consults its base of knowledge and takes a decision. More specifically the classifier evaluates the probability of having the sequence in input as output of each MC model and chooses the class as the one corresponding to the model that maximises the probability. Formally, we compute the probability of having a specific sequence of state $\vec{s}$ belonging to a Markov model λ_i as

$$P(\vec{s},\lambda_i) = \prod_{t=1}^{T} P(s_t|s_{t-1};A_i) = \prod_{t=1}^{T} a^{(i)}_{s_{t-1}s_t}, \tag{A.1}$$

where $\vec{s}$ is the sequence of the states, T is the length of the sequence, s_t is the current state at time instant t and $A_i := [a^{(i)}_{ij}]$ is the transition matrix for the model λ_i. After having evaluated the maximum of all the probabilities as $max_{\lambda_i}(P(\vec{s}|\lambda_i))$, where N is the number of the models λ_i in the classifier's base of knowledge, the time series is assigned to the most probable class. If more than one time-series belonging to the same class are presented in input, a second level of classification takes the final decision according to a majority rule.

A great advantage of the MC-based classifier is that it is able to work completely online while the HP is performing the mirror game in solo condition. Indeed, the classifier acquires the samples of the position signal in real time, buffers a minimal window of samples needed to evaluate a meaningful FFT, and then starts evaluating the classification probability of the data. Note that although during the game the length of the position time-series being acquired increases, the computational complexity of the classification exercise remains constant ($O(N)$ where N is the constant number of buffered samples).

A.2 Classifier validation

Participants and experimental task. To validate our classifier, we used the same dataset adopted in [178] and collected with the collaboration of the patients from the University Department of Adult Psychiatry (CHRU Montpellier, France). A total of 59 participants took part to the experiments (29 controls and 30 patients). Each participant was asked to perform 4 trials lasting 60 seconds of the mirror game in solo condition, in order that his/her IMS could emerge.

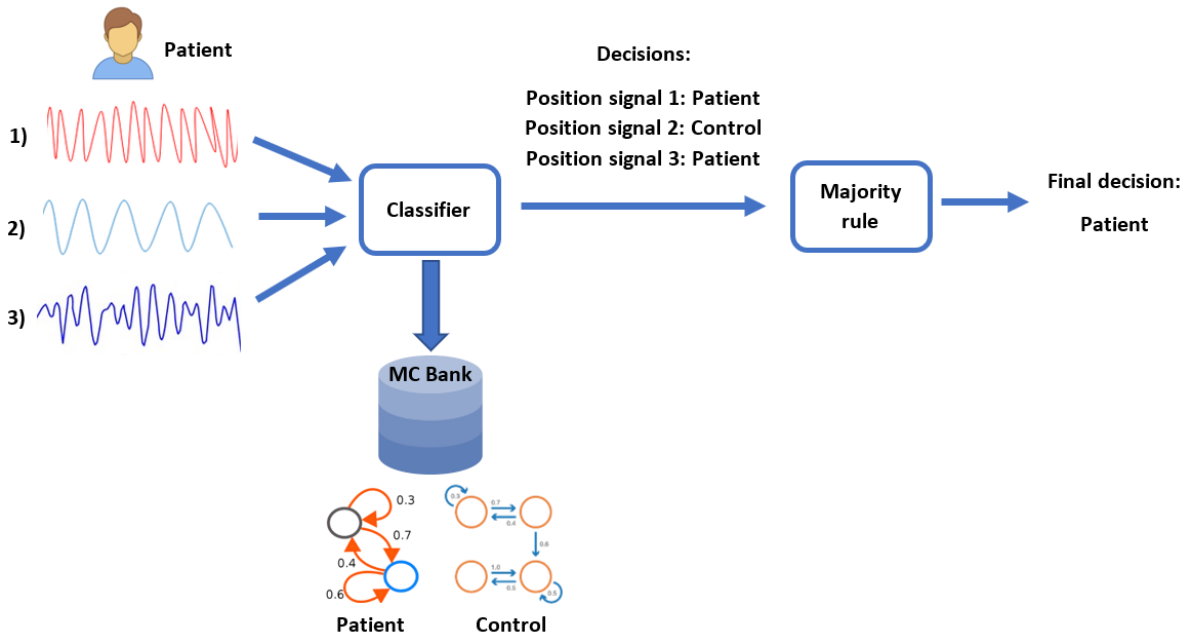


FIGURE A.1. Classifier with two levels of decisions. In the first level the classifier assigns the choice patient/control to each position time series according to the Markov chain model that has in its knowledge base. In the second level, the final decision is taken following a majority rule. The position time series are acquired by sessions of the mirror game in solo condition performing by the same player.

Metrics. We defined the metrics used in this work in terms of 4 sets: *i) true negative* (TN) are controls classified as controls; *ii) false negative* (FN) are the patients classified as controls; *iii) true position* (TP) are patients classified as patients; *iv) false positive* (FP) are controls classified as patients. Specifically, we define:

$$Accuracy := \frac{(TP+TN)}{N} \qquad Sensitivity := \frac{TP}{(TP+FN)}$$

(A.2)

$$Specificity := \frac{TN}{(TN+FP)} \qquad Precision := \frac{TP}{(TP+FP)}$$

Classification. To build the Markov models we followed the same steps described in Chapter 4. Specifically, each trial (position time series) of the dataset was interpolated to 100 Hz during the preprocessing and then windowed with a Hamming window of 200 samples with an overlapping of 150 samples. The signal was quantized with a codebook of 256 levels/symbols and so the trained MC is made up of 256 states (one per symbol). The two modelled MCs represent one the class “patient” and the other the class “control”.

Table A.1 shows the comparison between our results and the ones in [178]. The main difference

between the two methodologies lies in the number of features adopted. Indeed, our method exploits only one feature, based on the STFT, whereas [178] uses a total of three features, combines them and chooses according the majority. The three features are: the global wavelet spectrum (GWS), the ΔP_0 distribution of the lengths of movement segments (lengths between two turning points) and the distribution of $(c_1, c_2, c_3, c_4, c_5)$, polynomial coefficients of the stochastic model used to describe hand motion. Furthermore, the method in [178] requires to acquire the whole signal before processing it, while our method works intrinsically online.

Results of classification based on SFTF									
Feature	**Ctrls/Pts**	**TN**	**FP**	**TP**	**FN**	**Accuracy**	**Sensitivity**	**Specificity**	**Precision**
STFT-MC	29/30	29	0	30	0	1	1	1	1
Existing results of classification based on biomarkers									
(c_4, c_3)	29/30	23	6	21	9	0.7458	0.7000	0.7931	0.7778
(c_4, c_2)	29/30	25	4	21	9	0.7797	0.7000	0.8621	0.8400
(c_3, c_2)	29/30	21	8	24	6	0.7627	0.8000	0.7251	0.7500
ΔP_0	29/30	21	8	24	6	0.7627	0.8000	0.7251	0.7500
GWS	29/30	23	6	21	9	0.7458	0.7000	0.7931	0.7777
Majority	29/30	28	1	27	3	0.9322	0.9000	0.9655	0.9642

TABLE A.1. Comparison between our STFT-MC methodology and the existing results in [178] regarding the classification of people suffering from social disorders or not. Data are extracted from solo condition experiments of 29 controls (Ctrls) and 30 patients (Pts). For each feature the number of true negative (TN), false positive (FP), true positive (TP), false negative (FN) is reported and using them the performance in terms of Accuracy, Sensitivity, Specificity and Precision are evaluated.

APPENDIX B

FURTHER DETAILS ON THE LEADERSHIP EMERGENCE IN WALKING GROUPS

In this appendix we provide further details on Chapter 8. Specifically, in Section B.1 some tables supplementing the information about the design of experimental trials are reported. Then, in Section B.2 we confirm our findings providing further evidence about the behavioural leadership.

B.1 Further details on the design of experiments

A total of eight participants took part voluntarily to the experiment. In particular they were divided in two groups of 4 people each, named group 1 and group 2. We denoted by P$x.y$ pedestrian y in group x. Using this notation group 1 consisted of pedestrians P1.1, P1.2, P1.3 and P1.4, while group 2 of pedestrians P2.1, P2.2, P2.3 and P2.4.

All pedestrians of each group were asked to perform both speed-change and direction-change trials. At the beginning of each trial we varied the pedestrians' positions, the initial inter-personal distance between them and the sequence of the given instructions. In particular, a trial consisted of a sequence of any combination of two commands, turn right/turn left if the trial changes in heading direction or speed up/slow down if the trial changes in speed. Details about each trial are reported in Table B.1 and B.2 respectively for the group 1 and 2. A total of 3 trials out of 54 were discarded.

Trial	**Condition**	**IPD** [m]	**Group**	**Pedestrian** [FL, FR, BL, BR]
1	direction (left, left)	4	1	[P1.1, P1.4, P1.2, P1.3]
2	direction (left, right)	2	1	[P1.2, P1.3, P1.1, P1.4]
3	direction (left, right)	4	1	[P1.4, P1.1, P1.3, P1.2]
4	direction (right, left)	2	1	[P1.1, P1.2, P1.4, P1.3]
5	direction (left, left)	2	1	[P1.3, P1.4, P1.2, P1.1]
6	direction (right, right)	1	1	[P1.4, P1.2, P1.3, P1.1]
7	direction (right, right)	2	1	[P1.1, P1.2, P1.3, P1.4]
8	direction (left, left)	1	1	[P1.1, P1.4, P1.3, P1.2]
9	direction (right, left)	4	1	[P1.2, P1.3, P1.4, P1.1]
10	direction (right, right)	4	1	[P1.1, P1.4, P1.3, P1.2]
11	direction (right, left)	1	1	[P1.4, P1.1, P1.3, P1.2]
12	control	2	1	[P1.2, P1.3, P1.4, P1.1]
13	control	4	1	[P1.3, P1.1, P1.4, P1.2]
14	speed (fast, slow)	2	1	[P1.4, P1.2, P1.1, P1.3]
15	speed (slow, slow)	2	1	[P1.3, P1.2, P1.1, P1.4]
16	control	1	1	[P1.1, P1.4, P1.3, P1.2]
17	speed (slow, fast)	1	1	[P1.1, P1.3, P1.2, P1.4]
18	speed (fast, slow)	4	1	[P1.2, P1.4, P1.3, P1.1]
19	speed (fast, fast)	4	1	[P1.3, P1.2, P1.4, P1.1]
20	speed (slow, slow)	1	1	[P1.1, P1.2, P1.3, P1.4]
21	speed (fast, fast)	2	1	[P1.4, P1.3, P1.2, P1.1]
22	speed (slow, fast)	1	1	[P1.2, P1.4, P1.1, P1.3]
23	speed (fast, fast)	1	1	[P1.1, P1.3, P1.4, P1.2]
24	speed (slow, slow)	4	1	[P1.4, P1.1, P1.3, P1.2]
25	speed (fast, slow)	1	1	[P1.2, P1.1, P1.3, P1.4]
26	speed (slow, fast)	2	1	[P1.4, P1.3, P1.2, P1.1]

TABLE B.1. Details about the trials carried out by the group 1. Trial condition, initial inter-personal distance (IPD) between participants, and participants' positions are specified.

Trial	**Condition**	**IPD** [m]	**Group**	**Pedestrian** [FL, FR, BL, BR]
1	control	2	2	[P2.4, P2.2, P2.1, P2.3]
2	speed (fast, slow)	4	2	[P2.1, P2.2, P2.4, P2.3]
3	speed (slow, slow)	1	2	[P2.2, P2.1, P2.4, P2.3]
4	speed (slow, fast)	2	2	[P2.1, P2.3, P2.2, P2.4]
5	speed (slow, fast)	4	2	[P2.2, P2.4, P2.1, P2.3]
6	speed (fast, slow)	1	2	[P2.1, P2.3, P2.4, P2.2]
7	control	1	2	[P2.4, P2.1, P2.2, P2.3]
8	speed (slow, fast)	1	2	[P2.2, P2.4, P2.1, P2.3]
9	speed (fast, fast)	2	2	[P2.3, P2.4, P2.2, P2.1]
10	speed (fast, fast)	1	2	[P2.3, P2.1, P2.4, P2.2]
11	speed (slow, slow)	4	2	[P2.4, P2.2, P2.3, P2.1]
12	speed (fast, fast)	4	2	[P2.1, P2.4, P2.2, P2.3]
13	speed (slow, slow)	2	2	[P2.2, P2.1, P2.4, P2.3]
14	control	4	2	[P2.4, P2.2, P2.1, P2.3]
15	speed (fast, slow)	2	2	[P2.2, P2.4, P2.1, P2.3]
16	direction (left, right)	4	2	[P2.1, P2.4, P2.2, P2.3]
17	direction (left, right)	2	2	[P2.1, P2.3, P2.2, P2.4]
18	direction (right, right)	4	2	[P2.1, P2.4, P2.3, P2.2]
19	direction (left, left)	4	2	[P2.2, P2.1, P2.4, P2.3]
20	direction (left, left)	1	2	[P2.4, P2.3, P2.1, P2.2]
21	direction (left, right)	1	2	[P2.3, P2.2, P2.1, P2.4]
22	direction (right, left)	4	2	[P2.3, P2.2, P2.1, P2.4]
23	direction (right, left)	2	2	[P2.2, P2.3, P2.4, P2.1]
24	direction (left, left)	2	2	[P2.1, P2.4, P2.2, P2.3]
25	direction (right, right)	1	2	[P2.1, P2.2, P2.4, P2.3]

TABLE B.2. Details about the trials carried out by the group 2. Trial condition, initial inter-personal distance (IPD) between participants, and participants' positions are specified.

B.2 Additional evidence of behavioural leadership

In order to confirm our findings in the main text, in this section we show further evidence that the pedestrian P2.4 effectively assumes the leadership independently of his/her position in the group. In Figure B.1, B.2, B.3, B.4 four sample trials are visualised, where the pedestrian is respectively on the front right, back right, front left and back left position. It is easy to appreciate that the participant P2.4 moves first in all conditions acting in this way as leader even if he/she is located in the back positions of the group. Furthermore, in the network reconstruction the same pedestrian has only outgoing edges meaning that the participant under investigation influences the others but is seldom influenced.

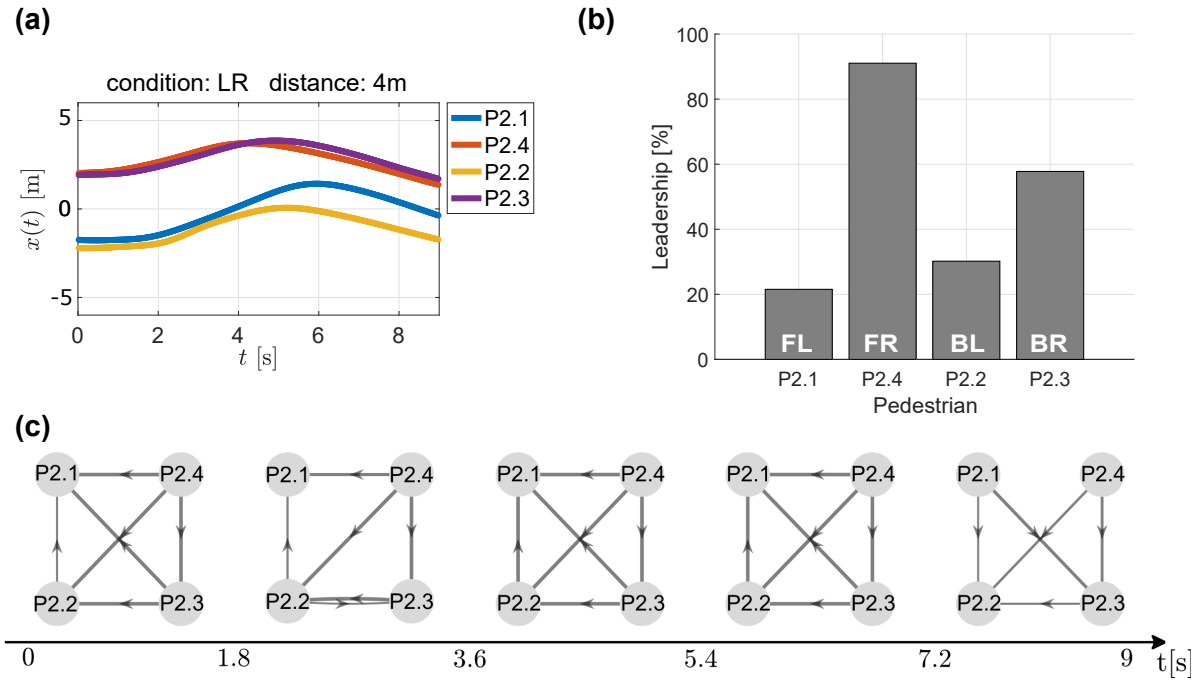


FIGURE B.1. Pedestrian P2.4 in the front right position. Direction-change trial implementing the sequence "LR" (turn left, turn right) and having an initial distance between participants of 4 metres. **(a)** *Velocity time series*. Pedestrians' velocity time series in the xy-plane where the x-axis represents the time in seconds and the y-axis the horizontal component of the velocity vector. Different pedestrians are characterised by different colours and identified by the label P$x.y$, where x is the group and y is the specific member of the group. **(b)** *Percentile leadership.* Each grey bar represents the percentage of leadership taken by the participant in the group during the trial. The labels FL, FR, BL, BR on each bar refer to the position occupied by the corresponding pedestrians and they are "front left, front right, back left, back right". **(c)** *Network reconstruction.* Each node represents a participant. The interaction between them are graphically represented through an arrow with different width, the thicker the arrow the stronger the interaction. Five different networks are reconstructed over different time-windows.

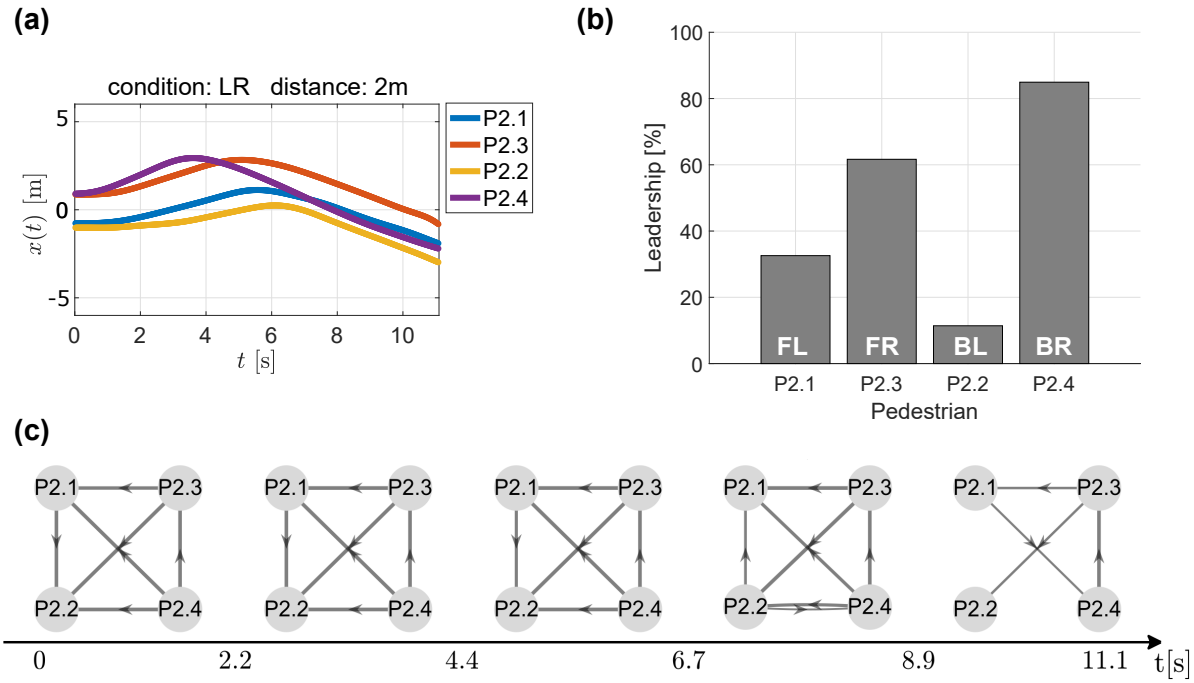


FIGURE B.2. Pedestrian P2.4 in the back right position. Direction-change trial implementing the sequence "LR" (turn left, turn right) and having an initial distance between participants of 2 metres. **(a)** *Velocity time series*. Pedestrians' velocity time series in the xy-plane where the x-axis represents the time in seconds and the y-axis the horizontal component of the velocity vector. Different pedestrians are characterised by different colours and identified by the label P$x.y$, where x is the group and y is the specific member of the group. **(b)** *Percentile leadership.* Each grey bar represents the percentage of leadership taken by the participant in the group during the trial. The labels FL, FR, BL, BR on each bar refer to the position occupied by the corresponding pedestrians and they are "front left, front right, back left, back right". **(c)** *Network reconstruction.* Each node represents a participant. The interaction between them are graphically represented through an arrow with different width, the thicker the arrow the stronger the interaction. Five different networks are reconstructed over different time-windows.

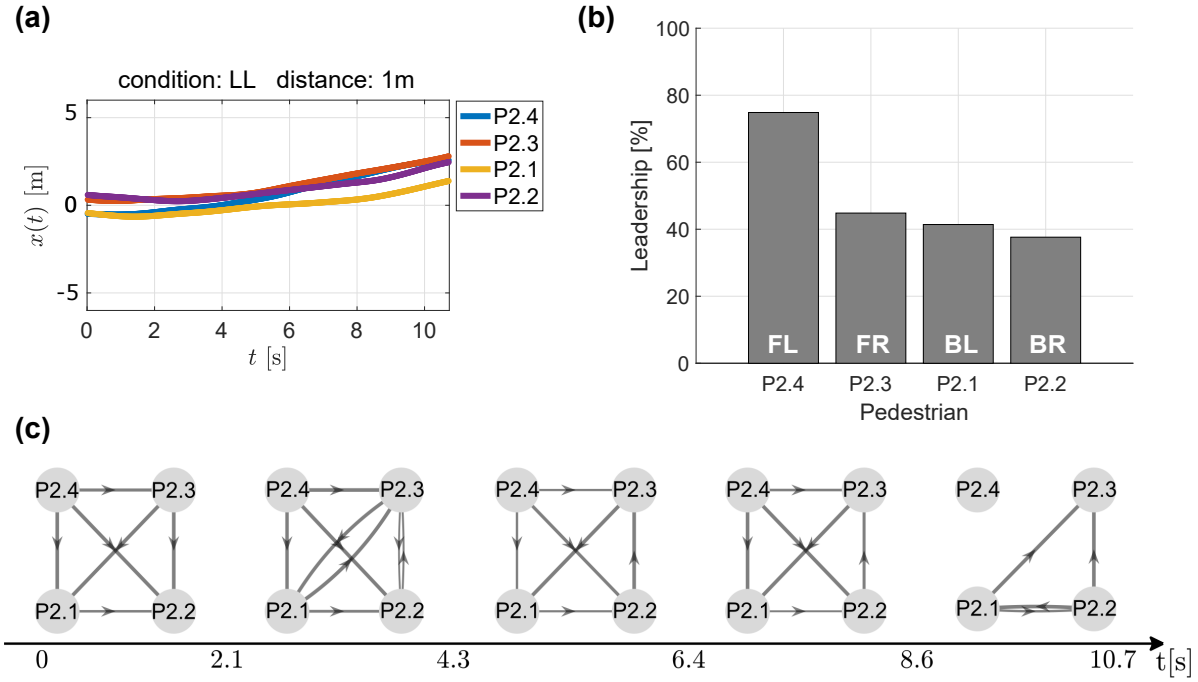


FIGURE B.3. Pedestrian P2.4 in the front left position. Direction-change trial implementing the sequence "LL" (turn left, turn left) and having an initial distance between participants of 1 metre. **(a)** *Velocity time series*. Pedestrians' velocity time series in the xy-plane where the x-axis represents the time in seconds and the y-axis the horizontal component of the velocity vector. Different pedestrians are characterised by different colours and identified by the label P$x.y$, where x is the group and y is the specific member of the group. **(b)** *Percentile leadership.* Each grey bar represents the percentage of leadership taken by the participant in the group during the trial. The labels FL, FR, BL, BR on each bar refer to the position occupied by the corresponding pedestrians and they are "front left, front right, back left, back right". **(c)** *Network reconstruction.* Each node represents a participant. The interaction between them are graphically represented through an arrow with different width, the thicker the arrow the stronger the interaction. Five different networks are reconstructed over different time-windows.

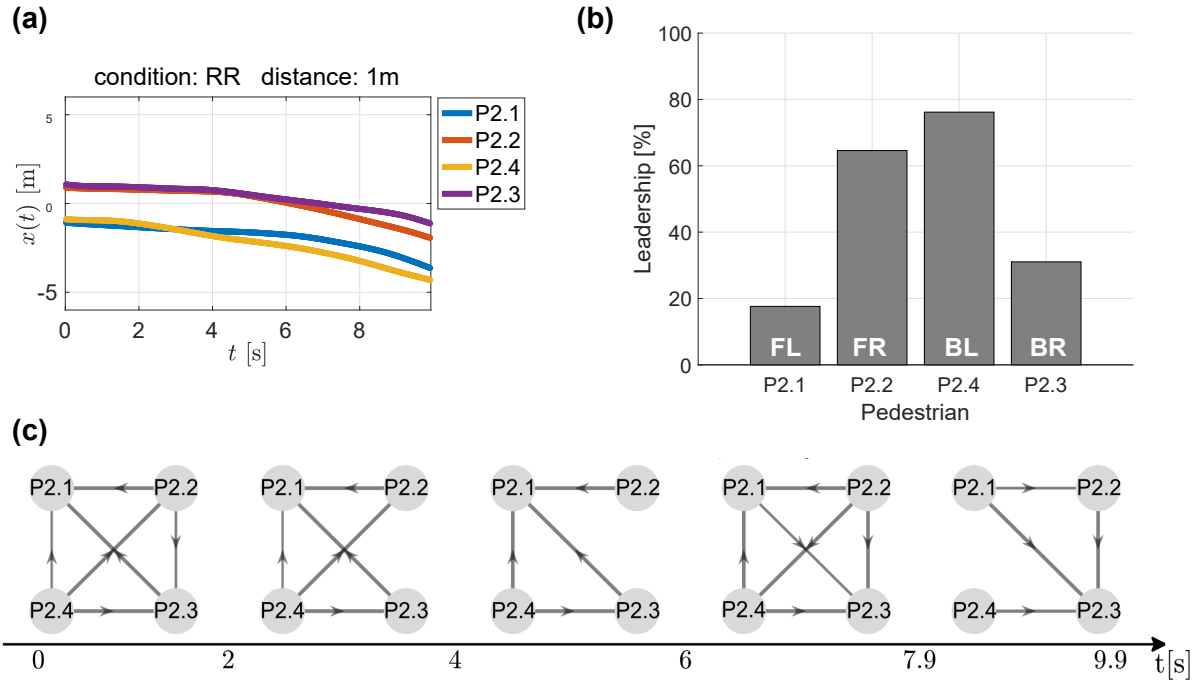


FIGURE B.4. Pedestrian P2.4 in the back left position. Direction-change trial implementing the sequence "RR" (turn right, turn right) and having an initial distance between participants of 1 metre. **(a)** *Velocity time series*. Pedestrians' velocity time series in the xy-plane where the x-axis represents the time in seconds and the y-axis the horizontal component of the velocity vector. Different pedestrians are characterised by different colours and identified by the label Px.y, where x is the group and y is the specific member of the group. **(b)** *Percentile leadership.* Each grey bar represents the percentage of leadership taken by the participant in the group during the trial. The labels FL, FR, BL, BR on each bar refer to the position occupied by the corresponding pedestrians and they are "front left, front right, back left, back right". **(c)** *Network reconstruction.* Each node represents a participant. The interaction between them are graphically represented through an arrow with different width, the thicker the arrow the stronger the interaction. Five different networks are reconstructed over different time-windows.

BIBLIOGRAPHY

[1] *Website of Blender*.
https://www.blender.org.

[2] *Website of MakeHuman: open source tool for making 3D characters*.
http://www.makehumancommunity.org.

[3] *Website of the European project AlterEgo*.
http://www.euromov.eu/alterego/homepage.

[4] C. T. ABDALLAH AND H. G. TANNER, *Complex networked control systems: Introduction to the special section*, IEEE Control Systems Magazine, 27 (2007), pp. 30–32.

[5] W. ABEND, E. BIZZI, AND P. MORASSO, *Human arm trajectory formation.*, Brain: a journal of neurology, 105 (1982), pp. 331–348.

[6] D. J. AGRAVANTE, A. CHERUBINI, A. BUSSY, P. GERGONDET, AND A. KHEDDAR, *Collaborative human-humanoid carrying using vision and haptic sensing*, in 2014 IEEE international conference on robotics and automation (ICRA), 2014, pp. 607–612.

[7] N. AHMED AND K. R. RAO, *Orthogonal transforms for digital signal processing*, Springer Science & Business Media, 2012.

[8] F. ALDERISIO, B. G. BARDY, AND M. DI BERNARDO, *Entrainment and synchronization in networks of rayleigh–van der pol oscillators with diffusive and haken–kelso–bunz couplings*, Biological cybernetics, 110 (2016), pp. 151–169.

[9] F. ALDERISIO, G. FIORE, AND M. DI BERNARDO, *Reconstructing the structure of directed and weighted networks of nonlinear oscillators*, Physical Review E, 95 (2017), p. e042302.

[10] F. ALDERISIO, G. FIORE, R. N. SALESSE, B. G. BARDY, AND M. DI BERNARDO, *Interaction patterns and individual dynamics shape the way we move in synchrony*, Scientific Reports, 7 (2017), p. e6846.

[11] F. ALDERISIO, M. LOMBARDI, AND M. DI BERNARDO, *Emergence of leadership in complex networks and human groups*, in 2018 IEEE International Symposium on Circuits and Systems (ISCAS), 2018, pp. 1–5.

[12] F. ALDERISIO, M. LOMBARDI, G. FIORE, AND M. DI BERNARDO, *A novel computer-based set-up to study movement coordination in human ensembles*, Frontiers in Psychology, 8 (2017), p. e967.

[13] I. AMADO, L. BRÉNUGAT-HERNÉ, E. ORRIOLS, C. DESOMBRE, M. DOS SANTOS, Z. PROST, M.-O. KREBS, AND P. PIOLINO, *A serious game to improve cognitive functions in schizophrenia: a pilot study*, Frontiers in psychiatry, 7 (2016), p. e64.

[14] P. G. AMAZEEN, R. SCHMIDT, AND M. T. TURVEY, *Frequency detuning of the phase entrainment dynamics of visually coupled rhythmic movements*, Biological Cybernetics, 72 (1995), pp. 511–518.

[15] C. ANDERSON-HANLEY, K. TURECK, AND R. L. SCHNEIDERMAN, *Autism and exergaming: effects on repetitive behaviors and cognition*, Psychology research and behavior management, 4 (2011), p. 129.

[16] A. ANGERER, A. HOFFMANN, A. SCHIERL, M. VISTEIN, AND W. REIF, *Robotics api: Object-oriented software development for industrial robots*, (2013).

[17] A. ARENAS, A. DÍAZ-GUILERA, J. KURTHS, Y. MORENO, AND C. ZHOU, *Synchronization in complex networks*, Physics reports, 469 (2008), pp. 93–153.

[18] S. ATES, C. J. HAARMAN, AND A. H. STIENEN, *Script passive orthosis: design of interactive hand and wrist exoskeleton for rehabilitation at home after stroke*, Autonomous Robots, 41 (2017), pp. 711–723.

[19] D. AVITABILE, P. SŁOWIŃSKI, B. BARDY, AND K. TSANEVA-ATANASOVA, *Beyond in-phase and anti-phase coordination in a model of joint action*, Biological cybernetics, 110 (2016), pp. 201–216.

[20] A. BABIARZ, A. CZORNIK, J. KLAMKA, M. NIEZABITOWSKI, AND R. ZAWISKI, *The mathematical model of the human arm as a switched linear system*, in 2014 19th International Conference on Methods and Models in Automation and Robotics (MMAR), 2014, pp. 508–513.

[21] B. G. BARDY, R. N. SALESSE, M. GUEUGNON, Z. ZHONG, J. LAGARDE, AND L. MARIN, *Movement similarities and differences during social interaction: The scientific foundation of the alterego european project*, in 2014 IEEE International Conference on Systems, Man, and Cybernetics (SMC), 2014, pp. 772–777.

[22] G. BEAVERS AND H. HEXMOOR, *Types and limits of agent autonomy*, in International Workshop on Computational Autonomy, Springer, 2003, pp. 95–102.

[23] R. BERNHARDT, D. SURDILOVIC, V. KATSCHINSKI, G. SCHRECK, AND K. SCHRÖER, *Next generation of flexible assembly systems*, in International Conference on Information Technology for Balanced Automation Systems, Springer, 2008, pp. 279–288.

[24] A. BILLARD, S. CALINON, R. DILLMANN, AND S. SCHAAL, *Robot Programming by Demonstration*, Springer Berlin Heidelberg, 2008.

[25] R. BISCHOFF, J. KURTH, G. SCHREIBER, R. KOEPPE, A. ALBU-SCHÄFFER, A. BEYER, O. EIBERGER, S. HADDADIN, A. STEMMER, G. GRUNWALD, ET AL., *The kuka-dlr lightweight robot arm-a new reference platform for robotics research and manufacturing*, in ISR 2010 (41st international symposium on robotics) and ROBOTIK 2010 (6th German conference on robotics), 2010, pp. 1–8.

[26] W. BLUETHMANN, R. AMBROSE, M. DIFTLER, S. ASKEW, E. HUBER, M. GOZA, F. REHNMARK, C. LOVCHIK, AND D. MAGRUDER, *Robonaut: A robot designed to work with humans in space*, Autonomous robots, 14 (2003), pp. 179–197.

[27] L. BOCCANFUSO, S. SCARBOROUGH, R. K. ABRAMSON, A. V. HALL, H. H. WRIGHT, AND J. M. O'KANE, *A low-cost socially assistive robot and robot-assisted intervention for children with autism spectrum disorder: field trials and lessons learned*, Autonomous Robots, 41 (2017), pp. 637–655.

[28] M. BOOS, J. PRITZ, S. LANGE, AND M. BELZ, *Leadership in moving human groups*, PLoS computational biology, 10 (2014), p. e1003541.

[29] I. BORG, P. J. GROENEN, AND P. MAIR, *Applied multidimensional scaling*, Springer Science & Business Media, 2012.

[30] N. A. BORGHESE, M. PIROVANO, R. MAINETTI, AND P. LANZI, *Iger: An intelligent game engine for rehabilitation*, in Converging Clinical and Engineering Research on Neurorehabilitation, Springer, 2013, pp. 947–950.

[31] C. BREAZEAL, *Toward sociable robots*, Robotics and autonomous systems, 42 (2003), pp. 167–175.

[32] R.-S. BREZIS, L. NOY, T. ALONY, R. GOTLIEB, R. COHEN, Y. GOLLAND, AND N. LEVIT-BINNUN, *Patterns of joint improvisation in adults with autism spectrum disorder*, Frontiers in psychology, 8 (2017), p. 1790.

[33] F. BULLO, *Lectures on Network Systems*, Kindle Direct Publishing, 1.3 ed., 2019.
With contributions by J. Cortes, F. Dorfler, and S. Martinez.

[34] W. CARRIER, III, *Soviet rover systems*, in Space Programs and Technologies Conference, 1992, p. e1487.

[35] D. CATLIN AND J. WOOLLARD, *Educational robots and computational thinking*, in Proceedings of 4th International workshop teaching robotics, teaching with robotics & 5th International conference robotics in education, 2014, pp. 144–151.

[36] G. CHEN, *Pinning control and synchronization on complex dynamical networks*, International Journal of Control, Automation and Systems, 12 (2014), pp. 221–230.

[37] S. X. CHEN, J. LI, P.-S. ZHONG, ET AL., *Two-sample and anova tests for high dimensional means*, The Annals of Statistics, 47 (2019), pp. 1443–1474.

[38] P. CHEVALIER, J.-C. MARTIN, B. ISABLEU, C. BAZILE, AND A. TAPUS, *Impact of sensory preferences of individuals with autism on the recognition of emotions expressed by two robots, an avatar, and a human*, Autonomous Robots, 41 (2017), pp. 613–635.

[39] E. CODRONS, N. F. BERNARDI, M. VANDONI, AND L. BERNARDI, *Spontaneous group synchronization of movements and respiratory rhythms*, PLoS One, 9 (2014), p. e107538.

[40] S. COHEN AND L. GUIBAS, *The earth mover's distance: Lower bounds and invariance under translation*, tech. rep., Stanford University, CA: Department of Computer Science, 1997.

[41] K. C. COLLINS, N. C. KENNEDY, A. CLARK, AND V. M. POMEROY, *Kinematic components of the reach-to-target movement after stroke for focused rehabilitation interventions: Systematic review and meta-analysis*, Frontiers in neurology, 9 (2018), p. e472.

[42] I. D. COUZIN, J. KRAUSE, N. R. FRANKS, AND S. A. LEVIN, *Effective leadership and decision-making in animal groups on the move*, Nature, 433 (2005), pp. 513–516.

[43] S. C. CRAMER, G. NELLES, R. R. BENSON, J. D. KAPLAN, R. A. PARKER, K. K. KWONG, D. N. KENNEDY, S. P. FINKLESTEIN, AND B. R. ROSEN, *A functional mri study of subjects recovered from hemiparetic stroke*, Stroke, 28 (1997), pp. 2518–2527.

[44] R. CROCHIERE, *A weighted overlap-add method of short-time fourier analysis/synthesis*, IEEE Transactions on Acoustics, Speech, and Signal Processing, 28 (1980), pp. 99–102.

[45] L. CUIJPERS, P. PASSOS, A. MURGIA, A. HOOGERHEIDE, K. LEMMINK, AND H. DE POEL, *Rocking the boat: does perfect rowing crew synchronization reduce detrimental boat movements?*, Scandinavian journal of medicine & science in sports, 27 (2017), pp. 1697–1704.

[46] G. C. DACHNER AND W. H. WARREN, *Behavioral dynamics of heading alignment in pedestrian following*, Transportation Research Procedia, 2 (2014), pp. 69–76.

[47] A. D'AUSILIO, L. BADINO, Y. LI, S. TOKAY, L. CRAIGHERO, R. CANTO, Y. ALOIMONOS, AND L. FADIGA, *Leadership in orchestra emerges from the causal relationships of movement kinematics*, PLoS ONE, 7 (2012), p. e35757.

[48] K. DAUTENHAHN, *Design spaces and niche spaces of believable social robots*, in Proceedings. 11th IEEE International Workshop on Robot and Human Interactive Communication, 2002, pp. 192–197.

[49] ——, *Roles and functions of robots in human society: implications from research in autism therapy*, Robotica, 21 (2003), pp. 443–452.

[50] P. DELELLIS, M. DI BERNARDO, AND M. PORFIRI, *Pinning control of complex networks via edge snapping*, Chaos: An Interdisciplinary Journal of Nonlinear Science, 21 (2011), p. 033119.

[51] J. E. DEUTSCH, D. ROBBINS, J. MORRISON, AND P. G. BOWLBY, *Wii-based compared to standard of care balance and mobility rehabilitation for two individuals post-stroke*, in 2009 Virtual Rehabilitation International Conference, 2009, pp. 117–120.

[52] A. M. DJURIC, R. URBANIC, AND J. RICKLI, *A framework for collaborative robot (cobot) integration in advanced manufacturing systems*, SAE International Journal of Materials and Manufacturing, 9 (2016), pp. 457–464.

[53] B. H. DOBKIN, *A rehabilitation-internet-of-things in the home to augment motor skills and exercise training*, Neurorehabilitation and neural repair, 31 (2017), pp. 217–227.

[54] P. DONNER AND M. BUSS, *Cooperative swinging of complex pendulum-like objects: Experimental evaluation*, IEEE Transactions on Robotics, 32 (2016), pp. 744–753.

[55] P. DONNER, F. CHRISTANGE, J. LU, AND M. BUSS, *Cooperative Dynamic Manipulation of Unknown Flexible Objects: Joint Energy Injection Based on Simple Pendulum Fundamental Dynamics*, International Journal of Social Robotics, 9 (2017), pp. 575–599.

[56] R. DUARTE, D. ARAÚJO, L. FREIRE, H. FOLGADO, O. FERNANDES, AND K. DAVIDS, *Intra- and inter-group coordination patterns reveal collective behaviors of football players near the scoring zone*, Human Movement Science, 31 (2012), pp. 1639–1651.

[57] R. DUARTE, B. TRAVASSOS, D. ARAÚJO, AND M. RICHARDSON, *The influence of manipulating the defensive playing method on collective synchrony of football teams*, Performance analysis of sport IX, (2013), pp. 91–96.

[58] B. R. DUFFY, *Anthropomorphism and the social robot*, Robotics and autonomous systems, 42 (2003), pp. 177–190.

[59] G. DUMAS, G. C. DE GUZMAN, E. TOGNOLI, AND J. S. KELSO, *The human dynamic clamp as a paradigm for social interaction*, Proceedings of the National Academy of Sciences, 111 (2014), pp. E3726–E3734.

[60] J. R. DYER, C. C. IOANNOU, L. J. MORRELL, D. P. CROFT, I. D. COUZIN, D. A. WATERS, AND J. KRAUSE, *Consensus decision making in human crowds*, Animal Behaviour, 75 (2008), pp. 461–470.

[61] J. R. DYER, A. JOHANSSON, D. HELBING, I. D. COUZIN, AND J. KRAUSE, *Leadership, consensus decision making and collective behaviour in humans*, Philosophical Transactions of the Royal Society B: Biological Sciences, 364 (2009), pp. 781–789.

[62] A. EDSINGER AND C. C. KEMP, *Human-robot interaction for cooperative manipulation: Handing objects to one another*, in RO-MAN 2007-The 16th IEEE International Symposium on Robot and Human Interactive Communication, 2007, pp. 1167–1172.

[63] M. ELLAMIL, J. BERSON, J. WONG, L. BUCKLEY, AND D. S. MARGULIES, *One in the dance: musical correlates of group synchrony in a real-world club environment*, PloS One, 11 (2016), p. e0164783.

[64] P. EVRARD, E. GRIBOVSKAYA, S. CALINON, A. BILLARD, AND A. KHEDDAR, *Teaching physical collaborative tasks: Object-lifting case study with a humanoid*, (2009), pp. 399–404.

[65] M. A. EYS, T. M. LOUGHEAD, AND J. HARDY, *Athlete leadership dispersion and satisfaction in interactive sport teams*, Psychology of Sport and Exercise, 8 (2007), pp. 281–296.

[66] T. FLASH AND N. HOGAN, *The coordination of arm movements: an experimentally confirmed mathematical model*, Journal of neuroscience, 5 (1985), pp. 1688–1703.

[67] V. S. FOLKES, *Forming relationships and the matching hypothesis*, Personality and Social Psychology Bulletin, 8 (1982), pp. 631–636.

[68] T. FONG, I. NOURBAKHSH, C. KUNZ, L. FLUCKIGER, J. SCHREINER, R. AMBROSE, R. BURRIDGE, R. SIMMONS, L. HIATT, A. SCHULTZ, ET AL., *The peer-to-peer human-robot interaction project*, in Space 2005, 2005, p. e6750.

[69] T. FONG, C. THORPE, AND C. BAUR, *Collaboration, dialogue, human-robot interaction*, in Robotics Research, Springer, 2003, pp. 255–266.

[70] C. T. FREEMAN, E. ROGERS, A.-M. HUGHES, J. H. BURRIDGE, AND K. L. MEADMORE, *Iterative learning control in health care: Electrical stimulation and robotic-assisted upper-limb stroke rehabilitation*, IEEE Control Systems Magazine, 32 (2012), pp. 18–43.

[71] M. J. FU, J. S. KNUTSON, AND J. CHAE, *Stroke rehabilitation using virtual environments*, Physical Medicine and Rehabilitation Clinics, 26 (2015), pp. 747–757.

[72] A. FUCHS, V. K. JIRSA, H. HAKEN, AND J. S. KELSO, *Extending the hkb model of coordinated movement to oscillators with different eigenfrequencies*, Biological cybernetics, 74 (1996), pp. 21–30.

[73] E. GAMBAO, M. HERNANDO, AND D. SURDILOVIC, *A new generation of collaborative robots for material handling*, in ISARC. Proceedings of the International Symposium on Automation and Robotics in Construction, vol. 29, 2012, p. 1.

[74] A. P. GEORGOPOULOS, J. F. KALASKA, AND J. T. MASSEY, *Spatial trajectories and reaction times of aimed movements: effects of practice, uncertainty, and change in target location*, Journal of Neurophysiology, 46 (1981), pp. 725–743.

[75] M. GILBERT, *Walking together: A paradigmatic social phenomenon*, MidWest studies in philosophy, 15 (1990), pp. 1–14.

[76] L. GIUGGIOLI, T. J. MCKETTERICK, AND M. HOLDERIED, *Delayed response and biosonar perception explain movement coordination in trawling bats*, PLoS computational biology, 11 (2015), p. e1004089.

[77] D. GLOWINSKI, P. COLETTA, G. VOLPE, A. CAMURRI, C. CHIORRI, AND A. SCHENONE, *Multi-scale entropy analysis of dominance in social creative activities*, in Proceedings of the 18th ACM international conference on Multimedia, ACM, 2010, pp. 1035–1038.

[78] D. C. GOFF, K. ROMERO, J. PAUL, M. M. PEREZ-RODRIGUEZ, D. CRANDALL, AND S. G. POTKIN, *Biomarkers for drug development in early psychosis: current issues and promising directions*, European Neuropsychopharmacology, 26 (2016), pp. 923–937.

[79] M. A. GOODRICH, A. C. SCHULTZ, ET AL., *Human–robot interaction: a survey*, Foundations and Trends® in Human–Computer Interaction, 1 (2008), pp. 203–275.

[80] J. A. GREGORY, *Shape preserving spline interpolation*, tech. rep., Brunel University, Uxbridge, UK, 1985.

[81] H.-M. GROSS, A. SCHEIDIG, K. DEBES, E. EINHORN, M. EISENBACH, S. MUELLER, T. SCHMIEDEL, T. Q. TRINH, C. WEINRICH, T. WENGEFELD, ET AL., *Roreas: robot coach for walking and orientation training in clinical post-stroke rehabilitation—prototype implementation and evaluation in field trials*, Autonomous Robots, 41 (2017), pp. 679–698.

[82] J. GUNDERSON AND L. GUNDERSON, *Intelligence ≠ autonomy ≠ capability*, Performance Metrics for Intelligent Systems, PERMIS, (2004).

[83] H. HAKEN, J. S. KELSO, AND H. BUNZ, *A theoretical model of phase transitions in human hand movements*, Biological cybernetics, 51 (1985), pp. 347–356.

[84] S. M. HATEM, G. SAUSSEZ, M. DELLA FAILLE, V. PRIST, X. ZHANG, D. DISPA, AND Y. BLEYENHEUFT, *Rehabilitation of motor function after stroke: a multiple systematic review focused on techniques to stimulate upper extremity recovery*, Frontiers in human neuroscience, 10 (2016), p. e442.

[85] G. HEINZEL, A. RÜDIGER, AND R. SCHILLING, *Spectrum and spectral density estimation by the discrete fourier transform (dft), including a comprehensive list of window functions and some new at-top windows*, tech. rep., Max Plank Institute, 2002.

[86] A. HENTOUT, M. AOUACHE, A. MAOUDJ, AND I. AKLI, *Human–robot interaction in industrial collaborative robotics: a literature review of the decade 2008–2017*, Advanced Robotics, 33 (2019), pp. 764–799.

[87] C. L. HILTON, K. CUMPATA, C. KLOHR, S. GAETKE, A. ARTNER, H. JOHNSON, AND S. DOBBS, *Effects of exergaming on executive function and motor skills in children with autism spectrum disorder: A pilot study*, American Journal of Occupational Therapy, 68 (2014), pp. 57–65.

[88] T. HIMBERG AND M. THOMPSON, *Group synchronization of coordinated movements in a cross-cultural choir workshop*, in ESCOM 2009: 7th Triennial Conference of European Society for the Cognitive Sciences of Music, 2009.

[89] G. HIRZINGER, B. BRUNNER, J. DIETRICH, AND J. HEINDL, *Rotex-the first remotely controlled robot in space*, in Proceedings of the 1994 IEEE international conference on robotics and automation, 1994, pp. 2604–2611.

[90] G. HOFFMAN AND C. BREAZEAL, *Collaboration in human-robot teams*, in AIAA 1st Intelligent Systems Technical Conference, 2004, p. 6434.

[91] N. HOGAN, *An organizing principle for a class of voluntary movements*, Journal of Neuroscience, 4 (1984), pp. 2745–2754.

[92] N. HOGAN AND D. STERNAD, *Sensitivity of smoothness measures to movement duration, amplitude, and arrests*, Journal of motor behavior, 41 (2009), pp. 529–534.

[93] M. HUBER, B. RABIN, C. DOCAN, G. C. BURDEA, M. ABDELBAKY, AND M. R. GOLOMB, *Feasibility of modified remotely monitored in-home gaming technology for improving hand function in adolescents with cerebral palsy*, IEEE Transactions on information technology in biomedicine, 14 (2010), pp. 526–534.

[94] M. HUBER, M. RICKERT, A. KNOLL, T. BRANDT, AND S. GLASAUER, *Human-robot interaction in handing-over tasks*, Proceedings of the 17th IEEE International Symposium on Robot and Human Interactive Communication, RO-MAN, (2008), pp. 107–112.

[95] T. HUNTSBERGER, G. RODRIGUEZ, AND P. S. SCHENKER, *Robotics challenges for robotic and human mars exploration*, in Robotics 2000, 2000, pp. 340–346.

[96] A. ISMA AND B. BRAHIM, *Time-dependant trajectory generation for tele-operated mobile manipulator*, in 2015 3rd International Conference on Control, Engineering & Information Technology (CEIT), 2015, pp. 1–5.

[97] S. JEZERNIK, R. G. WASSINK, AND T. KELLER, *Sliding mode closed-loop control of fes controlling the shank movement*, IEEE transactions on biomedical engineering, 51 (2004), pp. 263–272.

[98] K. KANEKO, F. KANEHIRO, S. KAJITA, H. HIRUKAWA, T. KAWASAKI, M. HIRATA, K. AKACHI, AND T. ISOZUMI, *Humanoid robot hrp-2*, in IEEE International Conference on Robotics and Automation, 2004. Proceedings. ICRA '04. 2004, vol. 2, 2004, pp. 1083–1090.

[99] S. KAUFFELD AND R. A. MEYERS, *Complaint and solution-oriented circles: Interaction patterns in work group discussions*, European Journal of Work and Organizational Psychology, 18 (2009), pp. 267–294.

[100] J. S. KELSO, *Dynamic patterns: The self-organization of brain and behavior*, MIT press, 1995.

[101] A. J. KING, D. D. JOHNSON, AND M. VAN VUGT, *The origins and evolution of leadership*, Current Biology, 19 (2009), pp. 911–916.

[102] B. KRALEMANN, L. CIMPONERIU, M. ROSENBLUM, A. PIKOVSKY, AND R. MROWKA, *Phase dynamics of coupled oscillators reconstructed from data*, Physical Review E, 77 (2008), p. e066205.

[103] T. KREUZ, F. MORMANN, R. G. ANDRZEJAK, A. KRASKOV, K. LEHNERTZ, AND P. GRASSBERGER, *Measuring synchronization in coupled model systems: A comparison of different approaches*, Physica D: Nonlinear Phenomena, 225 (2007), pp. 29–42.

[104] J. KRÜGER, T. K. LIEN, AND A. VERL, *Cooperation of human and machines in assembly lines*, CIRP annals, 58 (2009), pp. 628–646.

[105] D. LAKENS AND M. STEL, *If they move in sync, they must feel in sync: Movement synchrony leads to attributions of rapport and entitativity*, Social Cognition, 29 (2011), pp. 1–14.

[106] M. LAMB, R. W. KALLEN, S. J. HARRISON, M. DI BERNARDO, A. MINAI, AND M. J. RICHARDSON, *To pass or not to pass: Modeling the movement and affordance dynamics of a pick and place task*, Frontiers in psychology, 8 (2017), p. e1061.

[107] M. LAMB, T. LORENZ, S. J. HARRISON, R. W. KALLEN, A. MINAI, AND M. J. RICHARDSON, *Behavioral dynamics and action selection in a joint action pick-and-place task*, in CogSci 2017, vol. 1, Cognitive Science Society, 2017, pp. 2506–2511.

[108] ——, *PAPAc: A Pick and Place Agent Based on Human Behavioral Dynamics*, Proceedings of the 5th International Conference on Human Agent Interaction - HAI '17, (2017), pp. 131–141.

[109] M. LAMB, R. MAYR, T. LORENZ, A. A. MINAI, AND M. J. RICHARDSON, *The paths we pick together: a behavioral dynamics algorithm for an hri pick-and-place task*, in Companion of the 2018 ACM/IEEE International Conference on Human-Robot Interaction, 2018, pp. 165–166.

[110] P. C. LEGER, A. TREBI-OLLENNU, J. R. WRIGHT, S. A. MAXWELL, R. G. BONITZ, J. J. BIESIADECKI, F. R. HARTMAN, B. K. COOPER, E. T. BAUMGARTNER, AND M. W. MAIMONE, *Mars exploration rover surface operations: Driving spirit at gusev crater*, in 2005 IEEE International Conference on Systems, Man and Cybernetics, vol. 2, 2005, pp. 1815–1822.

[111] N. E. LEONARD, G. F. YOUNG, K. HOCHGRAF, D. T. SWAIN, A. TRIPPE, W. CHEN, K. FITCH, AND S. MARSHALL, *In the dance studio: An art and engineering exploration of human flocking*, in Controls and Art, Springer, 2014, pp. 27–49.

[112] F. L. LEWIS, D. VRABIE, AND K. G. VAMVOUDAKIS, *Reinforcement learning and feedback control: Using natural decision methods to design optimal adaptive controllers*, IEEE Control Systems Magazine, 32 (2012), pp. 76–105.

[113] Y. LI, *Deep reinforcement learning: An overview*, arXiv preprint arXiv:1701.07274, (2017).

[114] Y. LIU AND G. NEJAT, *Robotic urban search and rescue: A survey from the control perspective*, Journal of Intelligent & Robotic Systems, 72 (2013), pp. 147–165.

[115] M. LOMBARDI, *Chronos: a tool to study synchronization and coordination in human ensembles*.
`https://dibernardogroup.github.io/Chronos/index.html`.

[116] M. LOMBARDI, D. LIUZZA, AND M. DI BERNARDO, *Generation and classification of individual behaviours for virtual players control in motor coordination tasks*, in 2018 European Control Conference (ECC), 2018, pp. 2374 – 2379.

[117] ——, *Deep learning control of artificial avatars in group coordination tasks*, in 2019 IEEE International Conference on Systems, Man and Cybernetics (SMC), 2019, pp. 724 – 729.

[118] ——, *Using learning to control artificial avatars in human motor coordination tasks*, Accepted to IEEE Transaction on Robotics, (2019).

[119] S. LUKE ET AL., *Genetic programming produced competitive soccer softbot teams for robocup97*, Genetic Programming, 1998 (1998), pp. 214–222.

[120] K. L. MARSH, M. J. RICHARDSON, AND R. C. SCHMIDT, *Social connection through joint action and interpersonal coordination*, Topics in Cognitive Science, 1 (2009), pp. 320–339.

[121] M. J. MATARIĆ, J. ERIKSSON, D. J. FEIL-SEIFER, AND C. J. WINSTEIN, *Socially assistive robotics for post-stroke rehabilitation*, Journal of NeuroEngineering and Rehabilitation, 4 (2007).

[122] A. MEYSTEL, *Multiresolutional autonomy*, in Proceedings of the 1998 IEEE International Symposium on Intelligent Control (ISIC) held jointly with IEEE International Symposium on Computational Intelligence in Robotics and Automation (CIRA) Intell, 1998, pp. 516–519.

[123] X. A. MIRO, R. KUHN, AND L. BRAYDA, *Interactive personalized robot for home use*, Mar. 25 2008.
US Patent 7,349,758.

[124] V. MNIH, K. KAVUKCUOGLU, D. SILVER, A. A. RUSU, J. VENESS, M. G. BELLEMARE, A. GRAVES, M. RIEDMILLER, A. K. FIDJELAND, G. OSTROVSKI, ET AL., *Human-level control through deep reinforcement learning*, Nature, 518 (2015), p. 529.

[125] P. MORASSO, *Spatial control of arm movements*, Experimental brain research, 42 (1981), pp. 223–227.

[126] A. MÖRTL, T. LORENZ, AND S. HIRCHE, *Rhythm patterns interaction - Synchronization behavior for human-robot joint action*, PLoS ONE, 9 (2014).

[127] M. MOUSSAÏD, N. PEROZO, S. GARNIER, D. HELBING, AND G. THERAULAZ, *The walking behaviour of pedestrian social groups and its impact on crowd dynamics*, PloS one, 5 (2010), p. e10047.

[128] T. MUNZER, M. TOUSSAINT, AND M. LOPES, *Preference learning on the execution of collaborative human-robot tasks*, in 2017 IEEE International Conference on Robotics and Automation (ICRA), 2017, pp. 879–885.

[129] R. R. MURPHY, *Human-robot interaction in rescue robotics*, IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 34 (2004), pp. 138–153.

[130] M. NAGY, Z. ÁKOS, D. BIRO, AND T. VICSEK, *Hierarchical group dynamics in pigeon flocks*, Nature, 464 (2010), pp. 890–893.

[131] Z. NÉDA, E. RAVASZ, Y. BRECHET, T. VICSEK, AND A.-L. BARABÁSI, *Self-organizing processes: The sound of many hands clapping*, Nature, 403 (2000), p. 849.

[132] A. L. NELSON, E. GRANT, G. BARLOW, AND M. WHITE, *Evolution of complex autonomous robot behaviors using competitive fitness*, in IEMC'03 Proceedings. Managing Technologically Driven Organizations: The Human Side of Innovation and Change (IEEE Cat. No. 03CH37502), 2003, pp. 145–150.

[133] M. E. NEWMAN, *The structure and function of complex networks*, SIAM review, 45 (2003), pp. 167–256.

[134] P. NICOLAS, *Understanding Markov chains, examples and applications*, Springer, 2013.

[135] S. NIKOLAIDIS, S. NATH, A. D. PROCACCIA, AND S. SRINIVASA, *Game-theoretic modeling of human adaptation in human-robot collaboration*, in 2017 12th ACM/IEEE International Conference on Human-Robot Interaction (HRI, 2017, pp. 323–331.

[136] L. NOY, E. DEKEL, AND U. ALON, *The mirror game as a paradigm for studying the dynamics of two people improvising motion together*, Proceedings of the National Academy of Sciences, 108 (2011), pp. 20947–20952.

[137] L. NOY, N. LEVIT-BINUN, AND Y. GOLLAND, *Being in the zone: physiological markers of togetherness in joint improvisation*, Frontiers in human neuroscience, 9 (2015), p. 187.

[138] S. J. ORFANIDIS, *Optimum signal processing: an introduction*, Macmillan publishing company, 1988.

[139] P. E. W. OUWEHAND AND C. L. E. PEPER, *Does interpersonal movement synchronization differ from synchronization with a moving object?*, Neuroscience letters, 606 (2015), pp. 177–181.

[140] K. ÖZCIMDER, B. DEY, R. J. LAZIER, D. TRUEMAN, AND N. E. LEONARD, *Investigating group behavior in dance: an evolutionary dynamics approach*, in 2016 American Control Conference (ACC), 2016, pp. 6465–6470.

[141] I. PALUNKO, P. DONNER, M. BUSS, AND S. HIRCHE, *Cooperative suspended object manipulation using reinforcement learning and energy-based control*, in 2014 IEEE/RSJ International Conference on Intelligent Robots and Systems, 2014, pp. 885–891.

[142] P. PASSOS, K. DAVIDS, D. ARAÚJO, N. PAZ, J. MINGUÉNS, AND J. MENDES, *Networks as a novel tool for studying team ball sports as complex social systems*, Journal of Science and Medicine in Sport, 14 (2011), pp. 170–176.

[143] N. PELECHANO AND N. BADLER, *Modeling Crowd and Trained Leader Behavior during Building Evacuation*, IEEE Computer Graphics and Applications, 6 (2006), pp. 80–86.

[144] M. PESHKIN AND J. E. COLGATE, *Cobots*, Industrial Robot: An International Journal, 26 (1999), pp. 335–341.

[145] L. PETERNEL, T. PETRIČ, E. OZTOP, AND J. BABIČ, *Teaching robots to cooperate with humans in dynamic manipulation tasks based on multi-modal human-in-the-loop approach*, Autonomous robots, 36 (2014), pp. 123–136.

[146] L. PETERNEL, N. TSAGARAKIS, AND A. AJOUDANI, *A human-robot co-manipulation approach based on human sensorimotor information*, IEEE Transactions on Neural Systems and Rehabilitation Engineering, 25 (2017), pp. 811–822.

[147] M. PIROVANO, R. MAINETTI, G. BAUD-BOVY, P. L. LANZI, AND N. A. BORGHESE, *Intelligent game engine for rehabilitation (IGER)*, IEEE Transactions on Computational Intelligence and AI in Games, 8 (2014), pp. 43–55.

[148] M. PIROVANO, E. SURER, R. MAINETTI, P. L. LANZI, AND N. A. BORGHESE, *Exergaming and rehabilitation: A methodology for the design of effective and safe therapeutic exergames*, Entertainment Computing, 14 (2016), pp. 55–65.

[149] M. E. POLLACK, *Intelligent technology for an aging population: The use of ai to assist elders with cognitive impairment*, AI magazine, 26 (2005), p. e9.

[150] M. E. POLLACK, L. BROWN, D. COLBRY, C. OROSZ, B. PEINTNER, S. RAMAKRISHNAN, S. ENGBERG, J. T. MATTHEWS, J. DUNBAR-JACOB, C. E. MCCARTHY, ET AL., *Pearl: A mobile robotic assistant for the elderly*, in AAAI workshop on automation as eldercare, 2002, pp. 85–91.

[151] Y. QUINONEZ, J. DE LOPE, AND D. MARAVALL, *Cooperative and competitive behaviors in a multi-robot system for surveillance tasks*, in International Conference on Computer Aided Systems Theory, Springer, 2009, pp. 437–444.

[152] L. R. RABINER, *A tutorial on hidden markov models and selected applications in speech recognition*, Proceedings of the IEEE, 77 (1989), pp. 257–286.

[153] I. RAÑÓ AND I. IOSSIFIDIS, *Modelling human arm motion through the attractor dynamics approach*, in 2013 IEEE International Conference on Robotics and Biomimetics (ROBIO), 2013, pp. 2088–2093.

[154] A. RASOULI AND J. K. TSOTSOS, *Autonomous vehicles that interact with pedestrians: A survey of theory and practice*, IEEE transactions on intelligent transportation systems, (2019).

[155] S. G. REEBS, *Can a minority of informed leaders determine the foraging movements of a fish shoal?*, Animal Behaviour, 59 (2000), pp. 403–409.

[156] M. RICHARDSON, R. L. GARCIA, T. D. FRANK, M. GREGOR, AND K. L. MARSH, *Measuring group synchrony: a cluster-phase method for analyzing multivariate movement time-series*, Frontiers in physiology, 3 (2012), p. e405.

[157] M. J. RICHARDSON, R. W. KALLEN, P. NALEPKA, S. J. HARRISON, M. LAMB, A. CHEMERO, E. SALTZMAN, AND R. C. SCHMIDT, *Modeling embedded interpersonal and multiagent coordination.*, in COMPLEXIS, 2016, pp. 155–164.

[158] M. J. RICHARDSON, K. L. MARSH, R. W. ISENHOWER, J. R. GOODMAN, AND R. C. SCHMIDT, *Rocking together: Dynamics of intentional and unintentional interpersonal coordination*, Human movement science, 26 (2007), pp. 867–891.

[159] M. J. RICHARDSON, K. L. MARSH, AND R. SCHMIDT, *Effects of visual and verbal interaction on unintentional interpersonal coordination.*, Journal of Experimental Psychology: Human Perception and Performance, 31 (2005), p. 62.

[160] K. RIO AND W. H. WARREN, *The visual coupling between neighbors in real and virtual crowds*, Transportation Research Procedia, 2 (2014), pp. 132–140.

[161] K. W. RIO, G. C. DACHNER, AND W. H. WARREN, *Local interactions underlying collective motion in human crowds*, Proceedings of the Royal Society B: Biological Sciences, 285 (2018).

[162] K. W. RIO, C. K. RHEA, AND W. H. WARREN, *Follow the leader: Visual control of speed in pedestrian following*, Journal of Vision, 14 (2014).

[163] N. ROY, G. BALTUS, D. FOX, F. GEMPERLE, J. GOETZ, T. HIRSCH, D. MARGARITIS, M. MONTEMERLO, J. PINEAU, J. SCHULTE, ET AL., *Towards personal service robots for the elderly*, in Workshop on Interactive Robots and Entertainment, vol. 25, 2000, p. 184.

[164] L. ROZO, D. BRUNO, S. CALINON, AND D. G. CALDWELL, *Learning optimal controllers in human-robot cooperative transportation tasks with position and force constraints*, in 2015 IEEE/RSJ international conference on intelligent robots and systems (IROS), 2015, pp. 1024–1030.

[165] Y. RUBNER, C. TOMASI, AND L. J. GUIBAS, *The earth mover's distance as a metric for image retrieval*, International journal of computer vision, 40 (2000), pp. 99–121.

[166] S. J. RUSSELL AND P. NORVIG, *Artificial intelligence: a modern approach*, Malaysia; Pearson Education Limited,, 2016.

[167] G. SAPOSNIK, M. MAMDANI, M. BAYLEY, K. THORPE, J. HALL, L. COHEN, AND R. TEASELL, *Clinical trial protocols effectiveness of virtual reality exercises in stroke rehabilitation (evrest): rationale, design, and protocol of a pilot randomized clinical trial assessing the wii gaming system*, International Journal Of Stroke, 5 (2010), pp. 47–51.

[168] G. N. SARIDIS, *Hierarchically Intelligent Machines*, World Scientific, 2001.

[169] R. SCHMIDT, L. NIE, A. FRANCO, AND M. J. RICHARDSON, *Bodily synchronization underlying joke telling*, Frontiers in human neuroscience, 8 (2014), p. 633.

[170] R. SCHMIDT, M. J. RICHARDSON, C. ARSENAULT, AND B. GALANTUCCI, *Visual tracking and entrainment to an environmental rhythm.*, Journal of Experimental Psychology: Human Perception and Performance, 33 (2007), p. 860.

[171] R. C. SCHMIDT AND M. J. RICHARDSON, *Dynamics of interpersonal coordination*, in Coordination: Neural, behavioral and social dynamics, Springer, 2008, pp. 281–308.

[172] J. SCHOLTZ, *Theory and evaluation of human robot interactions*, in 36th Annual Hawaii International Conference on System Sciences, 2003. Proceedings of the, 2003, pp. 10–pp.

[173] J. SCHOLTZ, J. YOUNG, J. L. DRURY, AND H. A. YANCO, *Evaluation of human-robot interaction awareness in search and rescue*, in IEEE International Conference on Robotics and Automation, 2004. Proceedings. ICRA'04. 2004, vol. 3, 2004, pp. 2327–2332.

[174] J. SCOTT KELSO, K. G. HOLT, P. RUBIN, AND P. N. KUGLER, *Patterns of human interlimb coordination emerge from the properties of non-linear, limit cycle oscillatory processes: Theory and data*, Journal of motor behavior, 13 (1981), pp. 226–261.

[175] T. B. SHERIDAN, *Humans and automation: System design and research issues.*, Human Factors and Ergonomics Society, 2002.

[176] T. B. SHERIDAN AND W. L. VERPLANK, *Human and computer control of undersea teleoperators*, tech. rep., Massachusetts Inst of Tech Cambridge Man-Machine Systems Lab, 1978.

[177] J. SHOHAT, *Inequalities for moments of frequency functions and for various statistical constants*, Biometrika, (1929), pp. 361–375.

[178] P. SŁOWIŃSKI, F. ALDERISIO, C. ZHAI, Y. SHEN, P. TINO, C. BORTOLON, D. CAPDEVIELLE, L. COHEN, M. KHORAMSHAHI, A. BILLARD, R. SALESSE, M. GUEUGNON,

L. MARIN, B. G. BARDY, M. DI BERNARDO, S. RAFFARD, AND K. TSANEVA-ATANASOVA, *Unravelling socio-motor biomarkers in schizophrenia*, npj Schizophrenia, 3 (2017), p. 8.

[179] P. SŁOWIŃSKI, C. ZHAI, F. ALDERISIO, R. SALESSE, M. GUEUGNON, L. MARIN, B. G. BARDY, M. DI BERNARDO, AND K. TSANEVA-ATANASOVA, *Dynamic similarity promotes interpersonal coordination in joint action*, Journal of The Royal Society Interface, 13 (2016), p. e20151093.

[180] J. SOECHTING AND F. LACQUANITI, *Invariant characteristics of a pointing movement in man*, Journal of Neuroscience, 1 (1981), pp. 710–720.

[181] M. STIEBER, C. TRUDEL, AND D. HUNTER, *Robotic systems for the international space station*, in Proceedings of International Conference on Robotics and Automation, vol. 4, 1997, pp. 3068–3073.

[182] R. M. STOGDILL, *Personal factors associated with leadership: A survey of the literature*, The Journal of psychology, 25 (1948), pp. 35–71.

[183] S. H. STROGATZ, *Exploring complex networks*, nature, 410 (2001), pp. 268–276.

[184] J. STÜCKLER AND S. BEHNKE, *Following human guidance to cooperatively carry a large object*, IEEE-RAS International Conference on Humanoid Robots, (2011), pp. 218–223.

[185] J. SUN, D. TAYLOR, AND E. M. BOLLT, *Causal network inference by optimal causation entropy*, SIAM Journal on Applied Dynamical Systems, 14 (2015), pp. 73–106.

[186] R. S. SUTTON AND A. G. BARTO, *Reinforcement learning: An introduction*, MIT press, 2018.

[187] W. TAGGART, S. TURKLE, AND C. D. KIDD, *An interactive robot in a nursing home: Preliminary remarks*, Towards Social Mechanisms of Android Science, 2005 (2005), pp. 56–61.

[188] U. TECHNOLOGIES, *Website of Unity*.
`https://unity.com`.

[189] R. TEMPLER, H. NICHOLLS, AND T. NICOLLE, *Robotics for meat processing–from research to commercialisation*, Industrial Robot: An International Journal, 26 (1999), pp. 290–296.

[190] G. T. THIELMAN, C. M. DEAN, AND A. GENTILE, *Rehabilitation of reaching after stroke: task-related training versus progressive resistive exercise*, Archives of physical medicine and rehabilitation, 85 (2004), pp. 1613–1618.

[191] T. TRUELSEN, B. PIECHOWSKI-JÓŹWIAK, R. BONITA, C. MATHERS, J. BOGOUSSLAVSKY, AND G. BOYSEN, *Stroke incidence and prevalence in europe: a review of available data*, European journal of neurology, 13 (2006), pp. 581–598.

[192] T. TSUMUGIWA, R. YOKOGAWA, AND K. HARA, *Variable impedance control with virtual stiffness for human-robot cooperative peg-in-hole task*, in IEEE/RSJ international conference on intelligent robots and systems, vol. 2, 2002, pp. 1075–1081.

[193] ULTRALEAP, *Website of Leap motion controller*. `https://www.ultraleap.com/product/leap-motion-controller/`.

[194] M. VARLET, L. MARIN, J. LAGARDE, AND B. G. BARDY, *Social postural coordination.*, Journal of Experimental Psychology: Human Perception and Performance, 37 (2011), p. 473.

[195] M. VARLET, L. MARIN, S. RAFFARD, R. C. SCHMIDT, D. CAPDEVIELLE, J.-P. BOULENGER, J. DEL-MONTE, AND B. G. BARDY, *Impairments of social motor coordination in schizophrenia*, PloS One, 7 (2012), p. e29772.

[196] C. VESPER, S. BUTTERFILL, G. KNOBLICH, AND N. SEBANZ, *A minimal architecture for joint action*, Neural Networks, 23 (2010), pp. 998–1003.

[197] A. VESTRI, F. PERUCH, S. MARCHI, M. FRARE, P. GUERRA, S. PIZZIGHELLO, S. MENEGHETTI, A. NUTBROWN, AND A. MARTINUZZI, *Individual and group treatment for patients with acquired brain injury in comprehensive rehabilitation*, Brain injury, 28 (2014), pp. 1102–1108.

[198] K. WADA, T. SHIBATA, T. SAITO, AND K. TANIE, *Robot assisted activity for elderly people and nurses at a day service center*, in Proceedings 2002 IEEE International Conference on Robotics and Automation, vol. 2, 2002, pp. 1416–1421.

[199] C. J. WATKINS AND P. DAYAN, *Q-learning*, Machine learning, 8 (1992), pp. 279–292.

[200] C. S. WEICKERT, T. W. WEICKERT, A. PILLAI, AND P. F. BUCKLEY, *Biomarkers in schizophrenia: a brief conceptual consideration*, Disease markers, 35 (2013), pp. 3–9.

[201] D. B. WEST ET AL., *Introduction to graph theory*, vol. 2, Prentice hall Upper Saddle River, NJ, 1996.

[202] C. WHITE, *Markov decision processes*, Springer, 2001.

[203] S. S. WILTERMUTH AND C. HEATH, *Synchrony and cooperation*, Psychological science, 20 (2009), pp. 1–5.

[204] A. M. WING AND C. WOODBURN, *The coordination and consistency of rowers in a racing eight*, Journal of sports sciences, 13 (1995), pp. 187–197.

[205] C. J. WINSTEIN, D. K. ROSE, S. M. TAN, R. LEWTHWAITE, H. C. CHUI, AND S. P. AZEN, *A randomized controlled comparison of upper-extremity rehabilitation strategies in acute stroke: a pilot study of immediate and long-term outcomes*, Archives of physical medicine and rehabilitation, 85 (2004), pp. 620–628.

[206] C. ZHAI, F. ALDERISIO, P. SŁOWIŃSKI, K. TSANEVA-ATANASOVA, AND M. DI BERNARDO, *Design of a virtual player for joint improvisation with humans in the mirror game*, PloS one, 11 (2016), p. e0154361.

[207] ——, *Design and validation of a virtual player for studying interpersonal coordination in the mirror game*, IEEE transactions on cybernetics, 48 (2017), pp. 1018–1029.

[208] C. ZHAI, F. ALDERISIO, K. TSANEVA-ATANASOVA, AND M. DI BERNARDO, *A novel cognitive architecture for a human-like virtual player in the mirror game*, in 2014 IEEE International Conference on Systems, Man, and Cybernetics (SMC), 2014, pp. 754–759.

[209] Z. ZHAO, R. N. SALESSE, M. GUEUGNON, R. C. SCHMIDT, L. MARIN, AND B. G. BARDY, *Moving attractive virtual agent improves interpersonal coordination stability*, Human Movement Science, 41 (2015), pp. 240–254.

[210] A. ZIENKIEWICZ, D. A. BARTON, M. PORFIRI, AND M. DI BERNARDO, *Leadership emergence in a data-driven model of zebrafish shoals with speed modulation*, The European Physical Journal: Special Topics, 224 (2015), pp. 3343–3360.