

# Design and Validation of a Virtual Player for Studying Interpersonal Coordination in the Mirror Game

Chao Zhai<sup>1</sup>, Francesco Alderisio<sup>1</sup>, Piotr Słowiński<sup>2</sup>, Krasimira Tsaneva-Atanasova<sup>2</sup>, and Mario di Bernardo<sup>1,3</sup>

<sup>1</sup>Department of Engineering Mathematics, University of Bristol, Merchant Venturers' Building, BS8 1UB, United Kingdom

<sup>2</sup>College of Engineering, Mathematics and Physical Sciences, University of Exeter, United Kingdom

<sup>3</sup>Department of Electrical Engineering and Information Technology, University of Naples Federico II, 80125 Naples, Italy

September 22, 2015

## Abstract

The mirror game has been recently proposed as a simple, yet powerful paradigm for studying interpersonal interactions. It has been suggested that a virtual partner able to play the game with human subjects can be an effective tool to affect the underlying neural processes needed to establish the necessary connections between the players, and also to provide new clinical interventions for the rehabilitation of patients suffering from social disorders. Inspired by the motor processes of the central nervous system (CNS) and the musculoskeletal system in the human body, in this paper we develop a novel interactive cognitive architecture based on nonlinear control theory to drive a virtual player (VP) to play the mirror game with a human player (HP) in different configurations. Specifically, we consider two cases: the former where the VP acts as leader and the latter where it acts as follower. The crucial problem is to design a feedback control architecture capable of imitating and following or leading a human player (HP) in a joint action task. Movement of the end-effector of the VP is modeled by means of a feedback controlled Haken-Kelso-Bunz (HKB) oscillator, which is coupled with the observed motion of the HP measured in real time. To this aim, two types of control algorithms (adaptive control and optimal control) are used and implemented on the HKB model so that the VP can generate human-like motion while satisfying certain kinematic constraints. A proof of convergence of the control algorithms is presented in the paper together with an extensive numerical and experimental validation of their effectiveness. A comparison with other existing designs is also discussed, showing the flexibility and the advantages of our control-based approach.

# 1 Introduction

The emergence of coordinated behavior between humans is a common phenomenon in many areas of human endeavor. Examples include improvisation theater, group dance, music playing, team sports and parade marching [1]. At the core of the interaction between the players lies a fundamental feedback mechanism where each player adapts his/her motion in response to the observed movement of the other.

To study this intriguing phenomenon, the mirror game has been recently proposed as a simple, yet effective paradigm. In its simplest formulation, the mirror game features two people imitating each other's movements at high temporal and spatial resolution [2]. The game can be played in different experimental conditions: the former where one of the players leads and the other has to follow the leader movement (Leader-Follower condition); the latter where the two players create joint synchronized movement (Joint Improvisation condition).

The theory of similarity in social psychology suggests that people prefer to team up with others possessing similar morphological and behavioral features, and that interpersonal coordination is enhanced if their movement shares similar kinematic features [3, 4]. Further evidence suggests that motor processes caused by interpersonal coordination are strictly related to mental connectedness. To be specific, motor coordination between two people contributes to social attachment [5].

As suggested in [5], coordination games can therefore be used to help people suffering from social disorders to improve their social skills. Also they can be effectively exploited in social robotics to enhance attachment, coordination and rehabilitation during human-robot interactions [6]. For this reason, it has been proposed that creating a VP or avatar able to coordinate its motion with that of a HP can be extremely useful to study the onset of coordination and how it is affected by similarity/dissimilarity between the players' motion characteristics [7]. A VP can also be used for diagnostics and rehabilitation of patients suffering from social disorders as recently proposed in [8].

The aim of this paper is the design of a novel interactive cognitive architecture (ICA) based on nonlinear control theory able to drive a VP to play the mirror game with a human either as a leader or as a follower. Specifically, the goal is that of designing a cognitive architecture able to drive the motion of the VP interacting with a HP in real-time while exhibiting certain desired kinematic features. When playing as a follower, the ICA needs to guarantee that, while exhibiting the desired movement properties, the VP tracks as closely as possible the motion of the human leader. When playing as the leader, the ICA needs instead to generate new interesting motion. In both cases, it is crucial for the VP to engage with the HP by producing human-like response in terms of kinematics (maximum acceleration, velocity profile etc) and delay times. In this paper we take the view that the design of such an architecture is fundamentally a nonlinear control design problem where given some reference input the architecture has to drive the VP onto a desired motion which is a function of the movement of the human player being sensed during the game. In particular, the ICA can be integrated into the humanoid robot

to achieve the desired dual-arm coordination [9].

We explore two different approaches, one based on adaptive control, the other on optimal control. Our control architecture mimics the two fundamental actions which have been suggested to be at the core of the emergence of motor coordination between two or more effectors in biological systems: feedback and feedforward [10–12]. Specifically, the motor system is able to correct the deviation from the desired movement with the aid of feedback control, whilst feedforward control allows it to reconcile the interdependency of the involved effectors and preplan the response to the sensory incoming information [11–13].

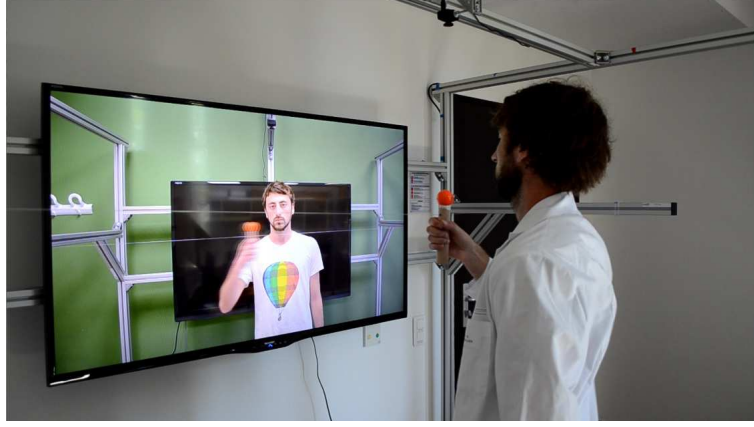
It is shown experimentally that the proposed control architectures are able to effectively drive the VP to play the mirror game while generating motion with desired kinematic properties. In particular, we use the concept of Individual Motor Signature (IMS) recently proposed in [14, 15] to characterize the motion of an individual player and evaluate how similar/dissimilar the motion of two different individuals is. Following our approach we are able to show that the VP driven by the cognitive architecture presented in the rest of this paper can play the mirror game either as a leader or a follower while exhibiting a desired IMS.

Relevant previous work in the literature includes the generation of human-like movement [16], the development of a mathematical model to explain the coordination dynamics observed experimentally in the mirror game [2], and the Human Dynamic Clamp paradigm proposed in [7, 17, 18] where the use of a virtual partner driven by appropriate mathematical models is proposed to study human motor coordination. These previous approaches will be used to investigate and compare the performance of the novel strategy presented in this paper. We wish to emphasize that the control algorithms developed and validated in what follows can be also effectively used for trajectory planning to enhance human-robot coordination in joint interactive tasks.

The rest of the paper is organized as follows. The mirror game set-up, problem statement and motor signature are discussed in Section 2 before presenting the schematic of the proposed cognitive architecture in Section 3. The feedback control strategies at the core of the ICA are developed and analyzed in Section 4 and 5. The experimental validation of the control algorithms is presented in Section 6 where experimental results are discussed showing the effectiveness of the proposed strategies. A comparison with other existing approaches is also carried out. Finally, conclusions and suggestions for future work are drawn in Section 7.

## 2 The Mirror Game Problem

Investigation of interpersonal coordination requires appropriate experimental paradigms. A typical paradigm recently proposed in the literature is the mirror game, which involves two people imitating each other’s movements at high temporal and spatial resolution [2]. It can be played in two different conditions: *Leader-Follower condition*, where the follower attempts at tracking the leader motion as accurately



**Figure 1** Experimental set-up of the mirror game between a VP and a HP at the University of Montpellier, France (see [8] for further details).

as possible, and *Joint Improvisation condition*, where the players jointly coordinate and synchronize their movements without any of the two being designated as leader or follower.

Our set-up is inspired by the one in [2]. Specifically, a small orange ball is mounted onto a string, which the HP can move back and forth along the string itself. In the meanwhile, the VP on the opposite screen moves its own ball on a parallel string with the same length (see Fig. 1). In this implementation of the mirror game, two players (a HP and a VP) are required to move their respective ball back and forth and synchronize their movement. Here, we assume that the game is played in a Leader-Follower condition, where the HP is the leader and the VP (robot or computer avatar) is the follower trying to track the leader movement. However, the VP can opt to act as the leader as well.

The position of the ball moved by the HP is detected by a camera. A feedback control strategy then needs to be designed in order to generate the trajectory of the ball moved by the VP so as to track the movement of the ball controlled by the HP. Such a trajectory can then be provided to the on-board controllers of the VP (robot or computer avatar) as the desired trajectory for its end effector.

To solve this control problem so that the VP motion presents similar features to the motion of a human player, we need to choose an appropriate model of the VP motion that can then be controlled using a nonlinear feedback strategy. To this purpose, here we use the Haken-Kelso-Bunz oscillator, which was first proposed in [19] as a model able to capture the observations made in experiments on human bimanual coordination. The model consists of two nonlinearly coupled nonlinear oscillators described by

$$\ddot{z} + (\alpha\dot{z}^2 + \beta z^2 - \gamma)\dot{z} + \omega^2 z = [a + b(z - w)^2](\dot{z} - \dot{w}) \quad (1)$$

where  $z, \dot{z}$  represent the position and velocity of finger 1,  $w, \dot{w}$  the position and velocity of finger 2 (modeled by a replica of the equation above obtained by swapping

$w, \dot{w}$  with  $z, \dot{z}$ );  $a$  and  $b$  are the coupling parameters and  $\alpha, \beta, \gamma$  and  $\omega$  characterize the response of each uncoupled finger when subject to some reference signal. However, it is worth pointing out that, other than describing intrapersonal motor coordination, the HKB model has been also used to describe interpersonal motor coordination involving two different people [20, 21]. In particular, the HKB oscillator has been suggested in the literature as a paradigmatic example of human motor coordination [7, 19]. Solving the mirror game can then be formulated as the following control problem. Given a nonlinear HKB oscillator of the form

$$\begin{cases} \dot{x} = y \\ \dot{y} = -(\alpha y^2 + \beta x^2 - \gamma)y - \omega^2 x + u \end{cases} \quad (2)$$

where  $x$  and  $\dot{x}$  refer to the position and velocity of the end effector of the VP, respectively, and  $u$  is an external control input, the problem is to design a feedback controller  $u$  such that  $x(t)$  achieves bounded asymptotic tracking of the position of the HP, while expressing some desired kinematic features.

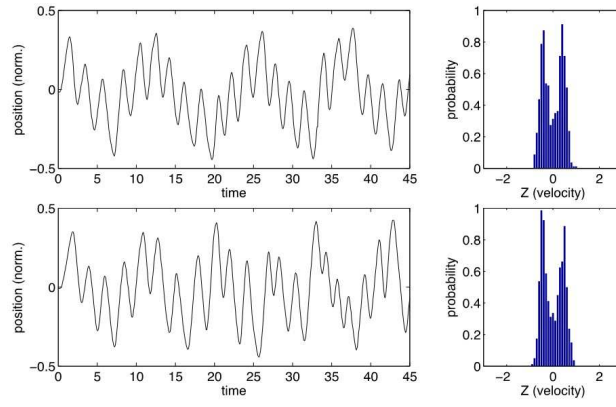
As metrics to characterize the kinematic properties of the motion of an individual playing the game we use the concept of individual motor signature (IMS), recently introduced in [14, 15]. It has been shown that the IMS is time invariant and is unique for each player. It is defined in terms of the velocity profile (or distribution) of the player's motion in solo trials. To quantify how similar or dissimilar the signatures of two different players are, we use the earth mover's distance (EMD) between any two probability distribution functions (PDF) of their velocity time series [15, 22]. The EMD between two PDFs  $p_1$  and  $p_2$  can be computed as follows

$$EMD(p_1, p_2) = \int_Z |CDF_{p_1}(z) - CDF_{p_2}(z)| dz$$

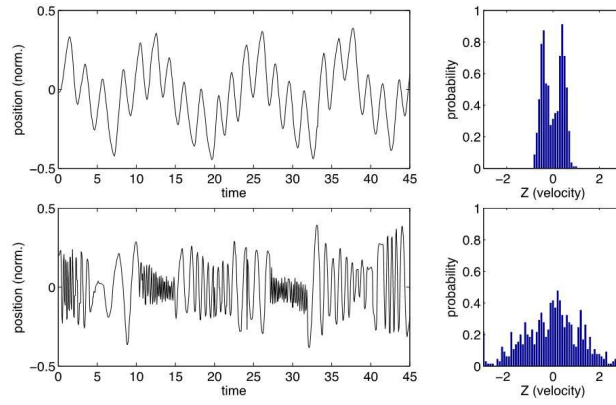
where  $Z$  denotes the integration domain, and  $CDF_{p_i}(z)$  denotes the cumulative distribution function of the distribution  $p_i, i \in \{1, 2\}$ . Fig.2(a) shows the position time series of the same HP and the corresponding PDF of velocity in two solo trials. It is visible that the two PDFs of velocity time series resemble each other in terms of their shape, and the EMD between them is 0.024. In contrast, the two PDFs of velocity time series in Fig.2(b) differ remarkably from each other, and the value of EMD is 0.604, which confirms the qualification of the PDF of velocity time series in solo trials as individual motor signature.

### 3 Design of the Cognitive Architecture

We design the cognitive architecture of the VP so as to replicate the main processes involved in making a human being play the mirror game (see Fig. 3). The visual system detects the ball's position on the string and generates visual signals, which are then transmitted to the central nervous system (CNS including brain and spinal cord). Several parts of the CNS (such as ventral horn, cerebellum and motor cortex) use an internal model to predict the kinematic characteristics of the other player's

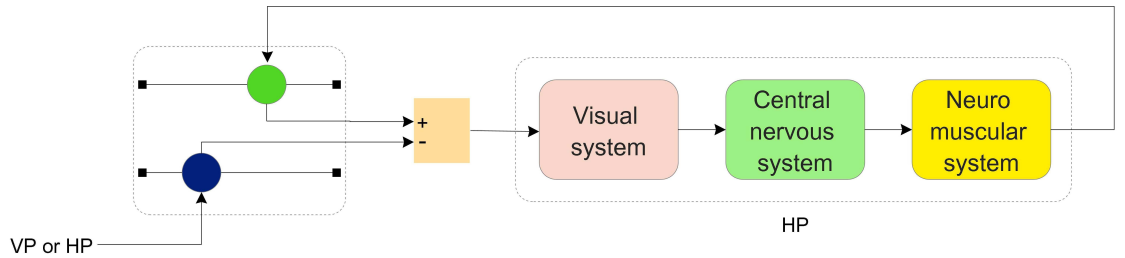


(a) Same player



(b) Two different players

**Figure 2** Position time series and the corresponding PDF of velocity time series for the same player (a) and two different players (b).



**Figure 3** Motor coordination between two players in mirror game.

motion and generate the neural impulses that control the extension and contraction of muscles. Finally, the neuromuscular system activates and coordinates the muscles involved in generating the hand movements.

This architecture is mapped onto the real-time control schematics shown in Fig. 4 whose blocks are briefly described below.

- A *camera* is used to detect the position of the HP, say  $r_p$ ;
- A *filtering and velocity estimation block* is used to filter the position data acquired by the camera via a low pass filter and to estimate the velocity of the HP (reference) via the simple formula

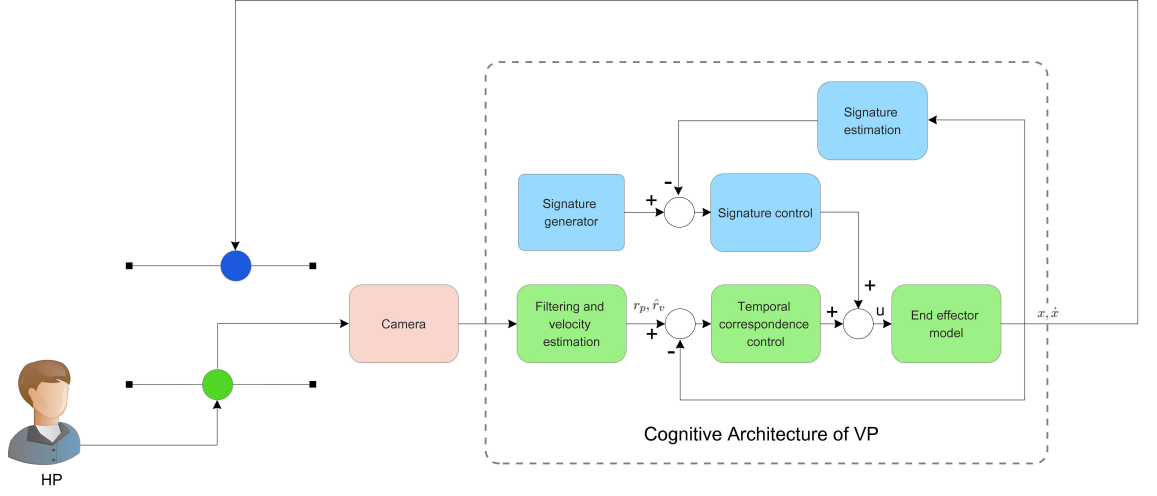
$$\hat{r}_v(t) = \frac{r_p(t_k) - r_p(t_{k-1})}{T} \quad t \in [t_k, t_{k+1}] \quad (3)$$

where  $k \in \mathbb{N}^*$ , and  $T = t_k - t_{k-1}$  denotes the sampling period of the camera. The estimated velocity is then used to predict the HP position over the next interval by using the expression:

$$\hat{r}_p(t) = r_p(t_k) + \hat{r}_v(t)(t - t_k), \quad t \in [t_k, t_{k+1}] \quad (4)$$

As an alternative, we could adopt a nonlinear observer to provide a better prediction of the reference velocity; for example, the nonlinear extended observer in [23]. Here we find that such a complication is unnecessary to solve the problem of interest and therefore choose to use the simple yet effective estimation strategy discussed above.

- At the core of the architecture lie the two blocks *Temporal Correspondence Control* and *Signature Control*. The former is designed to regulate the end effector model so that its motion tracks that of the HP with varying degrees of dynamic similarity. Specifically, it aims at minimizing the position error between the time series of the HP and that of the VP. The latter block uses the prerecorded velocity time series of a reference HP with the desired IMS (velocity profile) in order to generate the avatar trajectory with desired kinematic features. In particular, the aim of the signature controller is that of reducing the distance (computed in terms of EMD) between the velocity distribution of the VP and that of some reference HP it aims at replicating the motion characteristics of.
- The prerecorded velocity trajectory of a reference HP playing solo representing the desired IMS is stored in the *Signature generator block* while the signature of the avatar motion is estimated by the *Signature estimation block*.
- The *end effector model* is used to generate the avatar motion via an appropriate feedback control scheme. As mentioned before, we use the HKB oscillator to describe the dynamics of the end effector.



**Figure 4** Block diagram of the cognitive architecture of VP. Green blocks allow for the control of temporal correspondence between the VP and the HP, and blue blocks take into account the desired individual motor signature.

- Finally, the output of the cognitive architecture (position and velocity  $x$  and  $\dot{x}$ ) is used as the reference motion for the VP.

In what follows we focus on the design of the feedback control strategies that drive the cognitive architecture. We derive and compare two different types of controllers. First we develop an adaptive algorithm able to control the temporal correspondence between the VP and the HP during the game (green blocks in Fig.4). Then, we consider an optimal controller to solve simultaneously the multi objective control problem of tracking the trajectory of the HP while preserving the features of the desired IMS of interest (both green and blue blocks in Fig.4). For both strategies a proof of convergence is given before presenting numerical and experimental investigation of their performance.

## 4 Adaptive Control of Temporal Correspondence

To solve the control problem of temporal correspondence, we propose an adaptive controller based on the end effector model shown in (2). Specifically, we choose the nonlinear controller given by

$$u = \underbrace{[a(t) + b(t)(x - r_p)^2]}_{\text{Coordination}} (\dot{x} - \hat{r}_v) - \underbrace{C_p e^{-\delta(\dot{x} - \hat{r}_v)^2}}_{\text{Temporal Correspondence}} (x - r_p) \quad (5)$$

where  $r_p$  is the position of the HP,  $\hat{r}_v$  is the estimated velocity,  $C_p$  and  $\delta$  are constant parameters while the coupling parameters  $a$  and  $b$  are updated according to the



**Table 1** Adaptive Control Algorithm

---

1:	set $k = 1$ and running time $T_s$
2:	<b>while</b> ( $time < T_s$ )
3:	detect the position $r_p(t_k)$ of HP
4:	estimate the velocity $\hat{r}_v(t_k)$ of HP with (3)
5:	generate the control signal $u$ with (5)
6:	update coupling parameters $a$ and $b$ with (6) and (7)
7:	obtain the position $x$ and velocity $\dot{x}$ of VP by solving (2)
8:	$k = k + 1$
9:	<b>end while</b>

---

adaptive laws:

$$\dot{a} = -e^{-2a} [(x - r_p)(y - \hat{r}_v) + \eta_a(x - r_p)^2] - \eta_a \quad (6)$$

and

$$\dot{b} = \frac{y - \hat{r}_v}{e^{2b}} [\omega^2 x + (\alpha y^2 + \beta x^2 - \gamma)y - \eta_a(y - \hat{r}_v) - u] - \eta_a \quad (7)$$

where  $\eta_a$  is a positive constant. Note that the control law (5) consists of two complementary terms. The first has the same structure as the one of the coupling proposed in [19] to model the interaction between two HPs, albeit with the introduction of adaptive parameters to account for variability between different HPs. The second term, depending on the fixed parameters  $C_p$  and  $\delta$ , deals with the position error when the velocity mismatch approaches zero and hence the first term decays to zero. When  $|\dot{x} - \hat{r}_v|$  is relatively large, the coupling term of the HKB equation instead dominates and motor coordination between the two players becomes more pronounced during the mirror game.

Theoretical analysis of the adaptive control algorithm in Table 1 is given in what follows below.

#### 4.1 Convergence analysis

**Theorem 4.1.** *The adaptive feedback controller (5) ensures that the solution of the controlled HKB model (2) satisfies*

$$|x(t) - r_p(t)| \leq e^{\eta_a T} \sqrt{\frac{2\varepsilon}{e^{2\eta_a T} - 2}} + \frac{2}{e^{\eta_a T}} \sqrt{E(0)}, \quad t \in [0, +\infty)$$

if  $\eta_a$  is chosen so that

$$\eta_a > \frac{\ln 2}{2T}$$

where

$$E(0) = \frac{1}{2} \left[ (x(0) - r_p(0))^2 + (y(0) - \hat{r}_v(0))^2 + e^{2a(0)} + e^{2b(0)} \right]$$

and

$$\varepsilon = \sup_{k \in \mathbb{N}^*} (T^2 + 1) (\hat{r}_v(kT) - \hat{r}_v(kT - T))^2.$$

Proof: Choose the energy-like function

$$E \triangleq \frac{1}{2} \left[ (x - r_p)^2 + (y - \hat{r}_v)^2 + e^{2a} + e^{2b} \right] \quad (8)$$

Note that  $\hat{r}_v$  is fixed in each sampling interval  $[kT, (k+1)T), k \in \mathbb{N}^*$ . Then the time derivative of  $E$  along the trajectories of (2) with  $u$  defined in (5) is given by

$$\begin{aligned} \dot{E} &= (x - r_p)(\dot{x} - \dot{\hat{r}}_v) + (y - \hat{r}_v)\dot{y} + e^{2a}\dot{a} + e^{2b}\dot{b} \\ &= (x - r_p)(y - \hat{r}_v) - (y - \hat{r}_v) [(\alpha y^2 + \beta x^2 - \gamma)y + \omega^2 x - u] \\ &\quad + (y - \hat{r}_v) [\omega^2 x + (\alpha y^2 + \beta x^2 - \gamma)y - \eta_a(y - \hat{r}_v) - u] - \eta_a e^{2b} \\ &\quad - [(x - r_p)(y - \hat{r}_v) + \eta_a(x - r_p)^2] - \eta_a e^{2a} \\ &= -\eta_a(x - r_p)^2 - \eta_a(y - \hat{r}_v)^2 - \eta_a e^{2a} - \eta_a e^{2b} \\ &= -2\eta_a E, \quad t \in [kT, (k+1)T) \end{aligned}$$

Solving the above differential equation yields

$$E(t) = e^{-2\eta_a(t-kT)} E(kT), \quad t \in [kT, (k+1)T) \quad (9)$$

Moreover, at the sampling point  $kT$  we have

$$\begin{aligned} &E(kT) - E^-(kT) \\ &= \frac{1}{2} [(x - r_p(kT))^2 - (x - r_p(kT - T) - \hat{r}_v(kT - T)T)^2 \\ &\quad + (y - \hat{r}_v(kT))^2 - (y - \hat{r}_v(kT - T))^2] \\ &\leq (r_p(kT) - r_p(kT - T) - \hat{r}_v(kT - T)T)^2 \\ &\quad + (\hat{r}_v(kT) - \hat{r}_v(kT - T))^2 + E^-(kT) \\ &= (1 + T^2)(\hat{r}_v(kT) - \hat{r}_v(kT - T))^2 + E^-(kT) \end{aligned}$$

which is equivalent to

$$E(kT) \leq \varepsilon + 2E^-(kT) \quad (10)$$

where

$$\varepsilon = \sup_{k \in \mathbb{N}^*} (T^2 + 1) (\hat{r}_v(kT) - \hat{r}_v(kT - T))^2$$

and

$$E^-(kT) = \lim_{t \searrow kT} E(t)$$

Evaluating (9) and (10) at  $t = (k+1)T$  and nesting the inequalities backwards till  $t = 0$ , we get

$$\begin{aligned}
E(kT) &\leq \varepsilon \left[ 1 + \frac{2}{e^{2\eta_a T}} + \left( \frac{2}{e^{2\eta_a T}} \right)^2 + \dots + \left( \frac{2}{e^{2\eta_a T}} \right)^{k-1} \right] \\
&\quad + \left( \frac{2}{e^{2\eta_a T}} \right)^k E(0) \\
&= \varepsilon \frac{1 - \left( \frac{2}{e^{2\eta_a T}} \right)^k}{1 - \frac{2}{e^{2\eta_a T}}} + \left( \frac{2}{e^{2\eta_a T}} \right)^k E(0) \\
&\leq \frac{\varepsilon}{1 - \frac{2}{e^{2\eta_a T}}} + \frac{2}{e^{2\eta_a T}} E(0), \quad \forall k \in \mathbb{N}^*
\end{aligned} \tag{11}$$

when the inequality  $\eta_a > \frac{\ln 2}{2T}$  holds. Moreover, combining (9) with (11), we get

$$E(t) \leq \frac{\varepsilon}{1 - \frac{2}{e^{2\eta_a T}}} + \frac{2}{e^{2\eta_a T}} E(0), \quad t \in [0, +\infty)$$

which clearly implies

$$|x(t) - r_p(t)| \leq e^{\eta_a T} \sqrt{\frac{2\varepsilon}{e^{2\eta_a T} - 2}} + \frac{2}{e^{\eta_a T}} \sqrt{E(0)}, \quad t \in [0, +\infty)$$

■

**Remark 4.1.** It is easy to demonstrate that the coupling parameters  $a$  and  $b$  are upper bounded with the proposed adaptive laws.

**Remark 4.2.** Since  $r_p(t) \in [0, l], \forall t \geq 0$  and  $|\hat{r}_v(t)| \leq \frac{l}{T}$ , the following inequality holds

$$\varepsilon = \sup_{k \in \mathbb{N}^*} (T^2 + 1) (\hat{r}_v(kT) - \hat{r}_v(kT - T))^2 \leq \frac{4l^2(1 + T^2)}{T^2}$$

where  $l$  refers to the length of the string. Generally, the upper bound for the position error is relatively conservative. When the velocity of the HP is small,  $\varepsilon$  is small as well, and the estimation for the position error is accurate enough. In addition, taking the limit of (11) as  $kT \rightarrow \infty$  and combining it with (8), the position error between the two players satisfies the following inequality:

$$\limsup_{kT \rightarrow +\infty} |x(t) - r_p(t)| \leq e^{\eta_a T} \sqrt{\frac{2\varepsilon}{e^{2\eta_a T} - 2}}$$

Similarly, we can estimate the velocity error as

$$|\dot{x}(t) - \hat{r}_v(t)| \leq e^{\eta_a T} \sqrt{\frac{2\varepsilon}{e^{2\eta_a T} - 2}} + \frac{2}{e^{\eta_a T}} \sqrt{E(0)}, \quad t \in [0, +\infty)$$

While being effective in achieving bounded tracking of the HP position, the control approach derived so far is unable to explicitly guarantee that the generated motion follows a desired IMS. Therefore we consider a different scheme based on optimal control.

## 5 Optimal Temporal Correspondence and Signature Control

The second approach we propose is based on optimal control, which is an effective framework to allow for movement coordination and reconcile target tracking and individual motor signature [24]. We assume that in the mirror game the motion of the VP can be decomposed into a series of goal-directed movements (see Fig. 5) influenced by both the position of the HP and the desired individual motor signature. Thus, we formulate the problem of driving the end effector motion as described by (2) on a finite time interval  $[t_k, t_{k+1}]$  as the dynamic optimization problem

$$\min_{u \in R} J \quad (12)$$

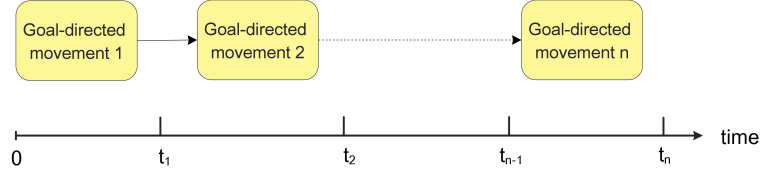
where

$$J = \underbrace{\frac{\theta_p}{2} (x(t_{k+1}) - \hat{r}_p(t_{k+1}))^2}_{\text{Temporal Correspondence}} + \underbrace{\frac{1}{2} \int_{t_k}^{t_{k+1}} \theta_\sigma (\dot{x}(\tau) - r_\sigma(\tau))^2}_{\text{Similarity}} + \eta_m \int_{t_k}^{t_{k+1}} u(\tau)^2 d\tau \quad (13)$$

with the constraint  $\theta_p + \theta_\sigma = 1$  and  $\theta_p, \theta_\sigma, \eta_m > 0$  being tunable control parameters. Here,  $\hat{r}_p(t_{k+1})$  denotes the estimated position of the HP at time  $t_{k+1}$  (see (4) for further details), while  $r_\sigma$  refers to a prerecorded velocity time series representing the desired motor signature. For the sake of simplicity, the optimization interval  $[t_k, t_{k+1}]$  in the cost function (13) corresponds to the sampling interval. The above cost function mainly consists of three terms. The first term aims at minimizing the mismatch between the position time series of the HP and that of the VP. The second term takes care of making the velocity profile of the VP as close as possible to the reference one (motor signature). The last term guarantees boundedness of the control effort. In particular, the idea behind this cost function is that the human-like movement of the VP emerges from the integration of three different goals related to temporal correspondence, motor signature and control energy expenditure, respectively. Notice that the VP acts as a leader when  $\theta_p$  is close to 0, since the term related to the position error  $x - \hat{r}_p$  in the cost function is negligible and the only aim of the VP is to exhibit the motion characterized by the desired motor signature. Instead, the VP behaves as a follower if  $\theta_p$  is close to 1 as in this case the controller aims solely at minimizing the mismatch between the HP and VP terminal positions.

### 5.1 Convergence Analysis

To prove stability of the optimal control algorithm in Table 2, we focus on proving the boundedness of the position error between the reference input and the output of



**Figure 5** Movement of the VP end effector in the mirror game

**Table 2** Optimal Control Algorithm

---

1: set $k = 1$ and running time $T_s$
2: <b>while</b> ( $time < T_s$ )
3:     detect the position $r_p(t_k)$ of HP
4:     estimate the position $\hat{r}_p(t_{k+1})$ of HP with (4)
5:     generate the control signal $u$ by solving (12)
6:     obtain the position $x$ and velocity $\dot{x}$ of VP by solving (2)
7: $k = k + 1$
8: <b>end while</b>

---

the cognitive architecture. In particular, optimality of the cost function is guaranteed in each optimization interval if the damped harmonic oscillator is adopted as end effector model instead of the HKB oscillator. Since both the reference position  $r_p$  and the desired velocity  $r_\sigma$  are bounded, we assume  $r_p \in [\underline{r}, \bar{r}]$  and  $r_\sigma \in [\underline{v}, \bar{v}]$ .

**Theorem 5.1.** *The optimal control algorithm applied to the HKB oscillator (2) with cost function (13) ensures bounded position error between the HP and the VP.*

Proof: First of all, we need to demonstrate that there exists a limit cycle in the HKB oscillator

$$\begin{cases} \dot{x} = y \\ \dot{y} = -(\alpha x^2 + \beta y^2 - \gamma)y - \omega^2 x \end{cases} \quad (14)$$

Choose the energy-like function as follows

$$V(x, y) = \frac{\omega^2 x^2 + y^2}{2}$$

The time derivative of  $V(x, y)$  along the trajectory of the HKB oscillator (14) is given by

$$\begin{aligned} \dot{V}(x, y) &= \omega^2 x \dot{x} + y \dot{y} \\ &= \omega^2 x y - (\alpha x^2 + \beta y^2 - \gamma)y^2 - \omega^2 x y \\ &= -(\alpha x^2 + \beta y^2 - \gamma)y^2 \end{aligned}$$

Define

$$r_{max} := \max \left( \sqrt{\frac{\gamma}{\alpha}}, \sqrt{\frac{\gamma}{\beta}} \right), r_{min} := \min \left( \sqrt{\frac{\gamma}{\alpha}}, \sqrt{\frac{\gamma}{\beta}} \right)$$

and construct a region  $\mathfrak{R}$  as follows (see Figure 6)

$$\mathfrak{R} := \{(x, y) \in \mathbb{R}^2 : c_1 \leq V(x, y) \leq c_2\}$$

where the positive constants  $c_1$  and  $c_2$  satisfy

$$r_{min} = \max \left( \sqrt{\frac{2c_1}{\omega^2}}, \sqrt{2c_1} \right), r_{max} = \min \left( \sqrt{\frac{2c_2}{\omega^2}}, \sqrt{2c_2} \right)$$

Clearly,  $\mathfrak{R}$  contains no stationary points of the system. Indeed, the only stationary point of the system is  $(x, y) = (0, 0)$ , but this stationary point is located outside of the region  $\mathfrak{R}$ . Moreover,  $\dot{V}(x, y) \geq 0$  when  $V(x, y) = c_1$  and  $\dot{V}(x, y) \leq 0$  when  $V(x, y) = c_2$ . According to the Poincare-Bendixson theorem, we can conclude that the HKB oscillator (14) has a limit cycle in  $\mathfrak{R}$ .

Let  $J^*$  denote the value of the cost function (13) with the optimal control algorithm in each time interval, and let  $J_0$  represent the value of the corresponding cost function when  $u = 0$ . Since  $u$  aims at minimizing the value of the cost function for all  $k \in \mathbb{N}^*$ , we can write

$$J^* \leq J_0 = \frac{\theta_p}{2} (x(t_{k+1}) - \hat{r}_p(t_{k+1}))^2 + \frac{\theta_\sigma}{2} \int_{t_k}^{t_{k+1}} (\dot{x}(\tau) - r_\sigma(\tau))^2 d\tau$$

Recall that  $r_p$  is bounded, which indicates that  $\hat{r}_p(t_{k+1})$  is bounded according to equation (3) and (4). Moreover,  $r_\sigma(\tau)$  is bounded as well, and note that  $x(t)$  and  $\dot{x}(t)$  are also bounded since the trajectory of the HKB oscillator converges to the limit cycle in  $\mathfrak{R}$ . Thus, we can claim that  $J_0$  is bounded for  $k \in \mathbb{N}^*$ , which implies boundedness of  $J^*$  and as a consequence of the position error between the VP and the HP.  $\blacksquare$

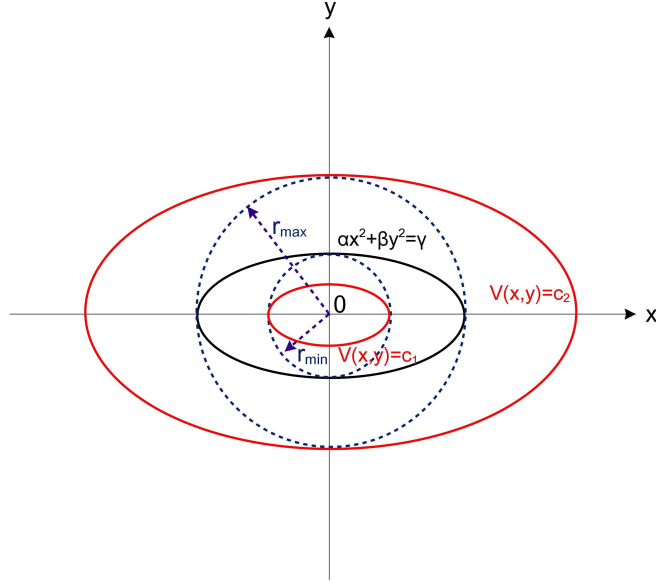
**Remark 5.1.** *It is demonstrated that the bound on the tracking error  $|x(t_{k+1}) - \hat{r}_p(t_{k+1})|$  converges to 0 as  $\theta_p \rightarrow 1$  and  $\eta_m \rightarrow 0$ . Similarly, the bound of the velocity error  $|\dot{x}(t_{k+1}) - r_\sigma(t_{k+1})|$  goes to 0 if  $\theta_\sigma \rightarrow 1$ ,  $\eta_m \rightarrow 0$  and  $r_\sigma(t_k) = y(t_k)$  (see Supplementary Material for the detailed analysis).*

The analytical solution for the optimization problem (12) is available if a linear damped harmonic oscillator is employed as the end effector model, and Pontryagin's minimum principle provides necessary and sufficient conditions to solve the minimization problem [25].

**Corollary 5.1.** *Given the linear system*

$$\ddot{x} + a\dot{x} + bx = u$$

*the optimal control approach guarantees convergence to the optimum solution over each subinterval.*



**Figure 6** Illustration on the construction of region  $\mathfrak{R}$ . The black ellipse is described by the equation  $\alpha x^2 + \beta y^2 = \gamma$ , and the region  $\mathfrak{R}$  refers to the ring-shaped area bounded by two red ellipses corresponding to  $V(x,y) = c_1$  and  $V(x,y) = c_2$ , respectively.

Proof: According to the fundamental theorem of the calculus of variations, we need to examine the second variation of the given cost function in order to establish the optimum. From the conclusions in [25], the second variation of the cost function (13) is given by

$$\begin{aligned} \delta^2 J &= \theta_p [\delta x(t_{k+1})]^2 \\ &+ \int_{t_k}^{t_{k+1}} \begin{pmatrix} \delta X & \delta u \end{pmatrix} \begin{pmatrix} H_{XX} & H_{Xu} \\ H_{uX}^T & H_{uu} \end{pmatrix} \begin{pmatrix} \delta X \\ \delta u \end{pmatrix} dt \end{aligned}$$

where  $H$  is the Hamiltonian

$$H(X, u, \lambda) = \frac{1}{2} \theta_s (\dot{x} - r_\sigma)^2 + \frac{1}{2} \eta u^2 + \lambda^T \begin{pmatrix} y \\ -ay - bx + u \end{pmatrix}$$

with  $X = [x, \dot{x}]^T = [x, y]^T$  and  $\lambda = [\lambda_1, \lambda_2]^T$ . Rewrite the linear system in matrix form as follows

$$\dot{X} = AX + Bu$$

where

$$A = \begin{pmatrix} 0 & 1 \\ -b & -a \end{pmatrix}, \quad B = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Let  $X = X^* + \delta X$  and  $u = u^* + \delta u$ , where  $X^*$  and  $u^*$  denote optimal state and optimal control, respectively. Since  $\dot{X}^* = AX^* + Bu^*$ , we get

$$\delta \dot{X} = A \delta X + B \delta u \tag{15}$$

where  $\delta X = [\delta x, \delta \dot{x}]^T$ . Thus, it follows from  $H_{Xu} = H_{uX} = [0 \ 0]^T$ ,  $H_{uu} = \eta > 0$  and

$$H_{XX} = \begin{pmatrix} 0 & 0 \\ 0 & \theta_s \end{pmatrix} \geq 0$$

that

$$\begin{aligned} \delta^2 J &= \theta_p [\delta x(t_{k+1})]^2 + \int_{t_k}^{t_{k+1}} \delta X(t)^T H_{XX} \delta X(t) + \eta (\delta u(t))^2 dt \\ &= \theta_p [\delta x(t_{k+1})]^2 + \int_{t_k}^{t_{k+1}} \theta_s (\delta \dot{x}(t))^2 + \eta (\delta u(t))^2 dt \\ &\geq 0 \end{aligned}$$

Moreover,  $\delta^2 J = 0$  is equivalent to  $\delta x(t_{k+1}) = 0$ ,  $\delta \dot{x}(t) = 0$  and  $\delta u(t) = 0$  for all  $t \in [t_k, t_{k+1}]$ , which yields  $\delta x(t) = \delta x(t_k) = 0$  from (15). This corresponds to the optimal solution  $X^*$  and the optimal control  $u^*$ . Therefore, we conclude that the optimal control ensures the minimum value of the cost function (13) for the linear system in each time interval. ■

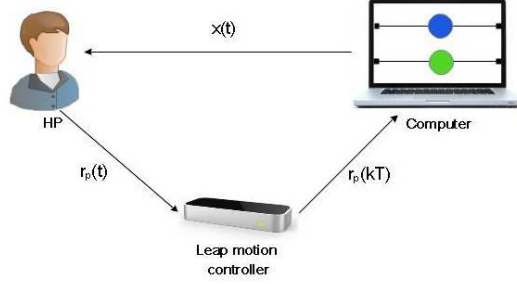
## 6 Validation

In this section we experimentally validate our control algorithms on a simple, yet effective, set-up implemented at the University of Bristol, UK. Experimental data of human-human interaction is used to evaluate the matching performance of the VP and a comparison with existing VP models is provided as well.

### 6.1 Experimental Set-up

The employed set-up was developed for measuring motions of players in the one-dimensional mirror game. A human participant is required to join the game while interacting with a VP (implemented on a laptop computer). In order to detect the position of his/her hand, a leap motion controller [27] is employed (see Fig. 7). The leap motion controller and the laptop computer are both placed on a table whose height is around 70cm. The HP is required to wave his/her hand horizontally over the leap motion controller at a vertical distance of approximately 50cm. Indeed, at this distance the horizontal resolution of the device is maximum and it is able to measure the hand position within a range of 60cm. The position of the hand of the HP within this interval is mapped into the interval  $[-0.5, 0.5]$  and visualized on the computer screen as a green solid circle, while the position of the VP is visualized as a blue solid circle. The adaptive control algorithm described in Section 4 is implemented with Euler method in Matlab (version R2012b), whilst the solver “bvp4c” is adopted to handle the optimization problem of the optimal control algorithm presented in Section 5. Players can be either standing or seated. After the game is initialized, there is a 2s wait before data recording begins and the game starts. This initial delay is used to allow the HP to place his/her hand over the leap





**Figure 7** Experimental set-up of the mirror game between a HP and a VP. The position of the human fingertip  $r_p(t)$  is detected by a leap motion controller, and the sampled position  $r_p(kT)$  is sent to the computer, while the position  $x(t)$  of the VP is generated by implementing the control algorithm. Two balls are shown on the computer screen, which describe the end effector positions of the HP (green ball) and the VP (blue ball), respectively.

motion controller. Human players are not instructed before playing the game, but they are just told to act as a leader and let the VP follow them during a 60s round. The case where the VP acts as a leader is also tested experimentally.

## 6.2 Measures

The temporal correspondence between the VP and the HP is evaluated according to the following indexes: the root mean square (RMS), the relative position error (RPE) [15], the circular variance (CV) [28] and the time lag (TL) [29].

1. RMS: The root mean square of the position error between two players describes the tracking accuracy of the follower in the mirror game.

$$RMS = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_{1,i} - x_{2,i})^2}$$

where  $n$  is the number of sampling steps in the simulation, and  $x_{1,i}$  and  $x_{2,i}$  denote the positions of the leader and the follower at the  $i$ -th sampling step, respectively.

2. RPE: The relative position error is a measure of how well the follower was tracking the leader in the mirror game [15]. Positive values of the RPE indicate that the follower is indeed behind the leader.

$$RPE = \begin{cases} (x_1(t) - x_2(t)) \operatorname{sgn}(\dot{x}_1(t)), & \text{if } \operatorname{sgn}(\dot{x}_1(t)) = \operatorname{sgn}(\dot{x}_2(t)) \neq 0; \\ |x_1(t) - x_2(t)|, & \text{otherwise.} \end{cases}$$

where  $x_1(t)$  and  $x_2(t)$  ( $\dot{x}_1(t)$  and  $\dot{x}_2(t)$ ) are the positions (velocities) of the leader and follower at time  $t$ , respectively.

3. CV: The circular variance is used to quantify the coordination level between two players

$$CV = \left\| \frac{1}{n} \sum_{k=1}^n e^{i\Delta\Phi_k} \right\| \in [0, 1]$$

where  $\Delta\Phi_k$  represents the relative phase between two players at the  $k$ -th sampling step,  $n$  refers to the total number of time steps and  $\|\cdot\|$  denotes the 2-norm.

4. TL: The time lag describes the shifted time that achieves the maximum cross-covariance of the two time series. It is sensitive to the changes of motion direction and hence can be interpreted as the average reaction time of the player in the mirror game [29].

### 6.3 Results

#### 6.3.1 VP driven by the ICA based on adaptive control

The parameters for the HKB equation and the adaptive feedback controller (AFC) in (5) are set heuristically as follows:  $\alpha = 10$ ,  $\beta = 20$ ,  $\gamma = -1$ ,  $\omega = 0.1$ ,  $a(0) = -5$ ,  $b(0) = -5$ ,  $C_p = 40$  and  $\delta = 0.25$ . In our implementation the sampling time is  $T = 0.1$ s and therefore  $\eta_a = 30 > \frac{\ln 2}{2T} \simeq 0.35$ . In particular, the values of all the previous parameters have been chosen so that the response of the HKB oscillator to several sinusoidal signals with different frequencies is qualitatively the same as the one of a HP trying to track the same references. Moreover, it is worth pointing out that the initial values of  $a$  and  $b$  influence the performance of the avatar only at start-up.

The reactive-predictive controller (RPC) proposed in [2] is also implemented to compare its performance against that of our adaptive feedback controller when considering the same input trajectory from the human leader. Following the scheme presented in [2] to implement the RPC, the dynamics of the VP is described by the following system:

$$\ddot{x} = \sum_{i=1}^5 A_i \omega_i \cos(\omega_i t) + f$$

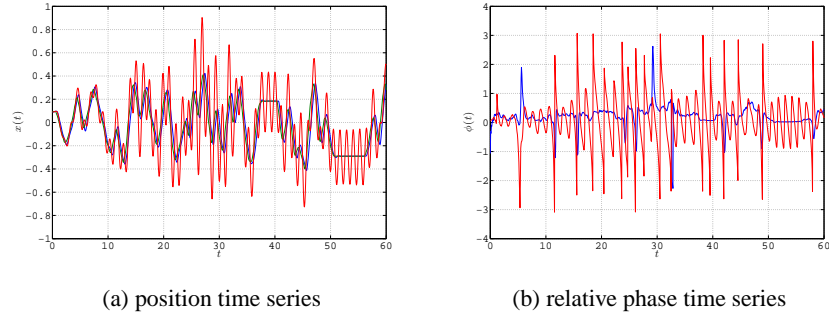
where  $x \in \mathbb{R}$  represents the position of the avatar and

$$\dot{f} = k(\hat{r}_v - \dot{x}), \quad k > 0$$

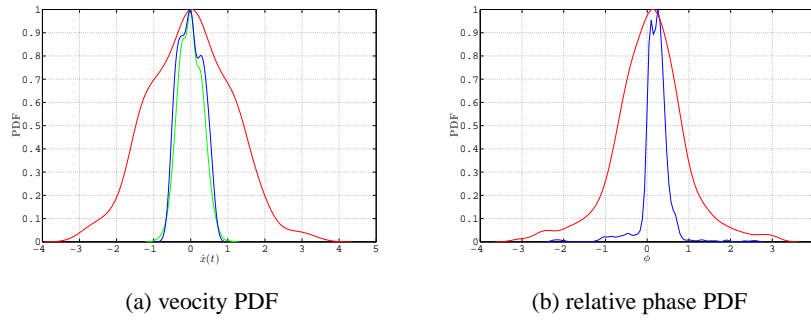
with the parameters  $A_i$  being estimated adaptively as

$$\dot{A}_i = \lambda \left[ \hat{r}_v - \sum_{i=1}^5 A_i \sin(\omega_i t) \right] \sin(\omega_i t), \quad \lambda > 0$$

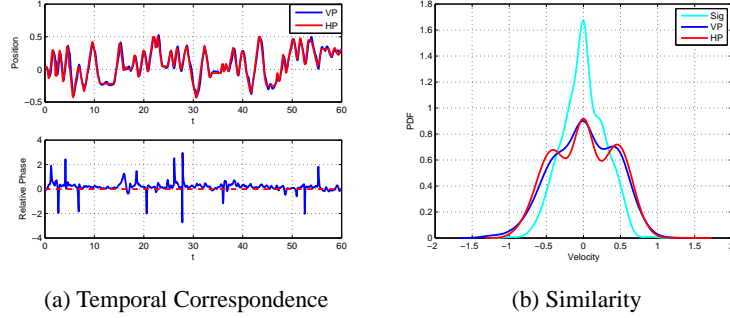
As suggested in [2], in this case the parameters are chosen as follows:  $\omega_1 = 0.025$ ,  $\omega_2 = 0.05$ ,  $\omega_3 = 0.075$ ,  $\omega_4 = 0.1$ ,  $\omega_5 = 0.125$ ,  $\lambda = 0.01$ ,  $k = 30$  and  $A_i(0) = 0, \forall i = 1, \dots, 5$ .



**Figure 8** Time series of the position (a) and of the relative phase (b) between the human leader and the VP; blue (AFC), red (RPC), green (human leader)



**Figure 9** Distributions of the velocity (a) and the relative phase (b) between the human leader and the VP; blue (AFC), red (RPC), green (human leader)

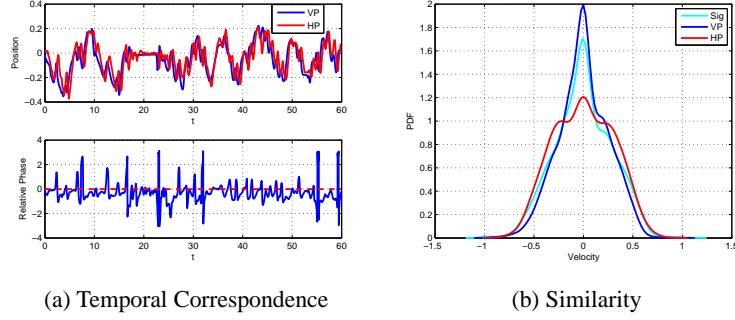


**Figure 10** Time evolution of positions and relative phase (a) and PDF of velocities (b) while the VP is driven by the optimal control and acts as follower in the mirror game

To compare the performance of the two algorithms, we plot the time series of both the position and the relative phase (see Fig. 8) together with the distributions of velocity and relative phase of the human and the VP (see Fig. 9). In particular, the relative phase between the two players is defined as  $\Delta\phi = \phi_{HP} - \phi_{VP}$ , where  $\phi_{HP}$  and  $\phi_{VP}$  are the phases of the HP and the VP, respectively. In addition, the phase is estimated according to the method proposed in [30]. Note that positive values of  $\Delta\phi$  correspond to the avatar following the HP during the game.

We can observe that, when using the reactive-predictive controller, the position of the VP presents oscillations away from the human participant position not only when he/she is moving, but also when he/she is still. Such an oscillatory feature does not appear when using the adaptive feedback controller. In general, both the position error  $e = x - r_p$  and the velocity error  $\dot{e} = \dot{x} - \dot{r}_v$  turn out to be higher when using the RPC strategy. When using the adaptive feedback controller, the position error remains smaller never exceeding 0.2, while it can become as high as 0.8 when using RPC. Similarly, the velocity error never exceeds 0.62 for the AFC, while it goes up to a maximum of 3 for the RPC.

Moreover, when using AFC, the relative phase time series is much closer to 0 than that obtained when using RPC, meaning that with our proposed algorithm it is possible for the VP to better synchronize with the human leader. Such results are confirmed by the relative phase distributions obtained when using both the algorithms, as shown in Fig. 9(b). Finally, the difference in the velocity distributions of the HP and the VP is much more evident when RPC is used, confirming that our strategy better captures the features of the HP and is therefore able to replicate more accurately the kinematic properties observed in human motor coordination in the context of the mirror game.

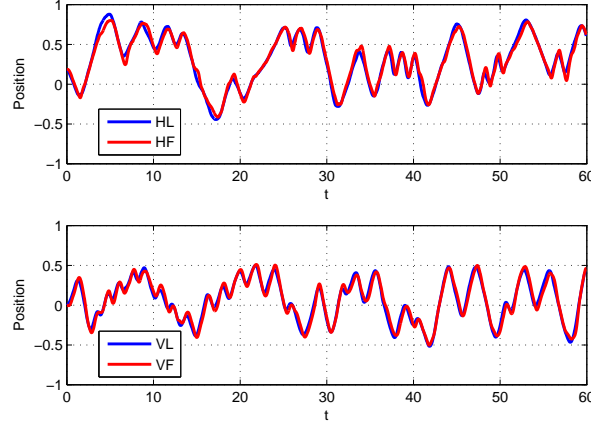


**Figure 11** Time evolution of positions and relative phase (a) and PDF of velocities (b) while the VP is driven by the optimal control and acts as leader in the mirror game

### 6.3.2 VP driven by the ICA based on optimal control

The parameters of the VP are set heuristically as follows:  $\alpha = 1$ ,  $\beta = 1$ ,  $\gamma = 1$ ,  $\omega = 1$ ,  $\eta_m = 10^{-4}$  and the sampling period  $T = t_{k+1} - t_k = 0.03s$ . In order for the VP to play the mirror game as a follower, we set the control parameters  $\theta_p = 0.9$  and  $\theta_\sigma = 0.1$ , which makes the VP play in a follower configuration (minimizing the position mismatch more than the signature EMDs). As we can see from the top panel in Fig.10(a), the VP performs quite well as a follower during the game; indeed, the root mean square (RMS) of the position error is equal to 0.057. In order to distinguish the leader from the follower in the game, we also calculate the relative phase between the HP and the VP. From the bottom panel in Fig. 10(a) we can observe that the majority of relative phase is positive, meaning that the VP is following the HP in the game for most of the time. The circular variance (CV) is also calculated to take into account the coordination level between two players. The CV between the HP and the VP is 0.95, which indicates a high coordination level. As for the distribution of the velocity, we can see in Fig. 10(b) that the VP signature (blue line) is closer to that of the HP (red line) than the desired motor signature (cyan line). This is due to the choice of the control parameters in the cost function (13) that render the strategy able to minimize more the position error between the players. The measured EMDs at the end of the trial are given as follows:  $EMD(Sig, HP) = EMD(Sig, VP) = 0.017$  and  $EMD(VP, HP) = 0.005$ .

The VP can be enabled to play the game as a leader by changing the control parameters setting  $\theta_p = 0.1$  and  $\theta_\sigma = 0.9$ . Experimental results are shown in Fig. 11. The RMS of the tracking error is 0.08, and the CV between the two players is 0.81. As depicted in the bottom panel of Fig. 11(a), the majority of the relative phase time series is negative, meaning that now the VP is leading the HP during the game for most of the time. In contrast to the previous case, the velocity distributions shown in Fig. 11(b) confirm that the VP is now matching well the desired



**Figure 12** Position time series of the HP pair (upper panel) and the corresponding VP pair (lower panel) in the mirror game.

signature (velocity profile). In this case the trade off is slightly larger but still the relevant EMDs are given as follows:  $EMD(Sig, HP) = EMD(Sig, VP) = 0.004$  and  $EMD(VP, HP) = 0.008$ .

### 6.3.3 Interaction between two VPs

As mentioned before, our final goal is to create a customized VP able to “replicate” the kinematic features of a given HP in the mirror game. To test how well the VP can replicate the features of a given HP, we carried out the following experiment. First of all, two HPs are required to play the mirror game in a Leader-Follower condition. Then the signatures of the human leader (HL) and the human follower (HF) are fed into a virtual leader (VL) and a virtual follower (VF), respectively. Finally, we make the VL ( $\theta_p = 0.43$ ) and the VF ( $\theta_p = 0.92$ ) play the mirror game together. In Fig. 12, the upper panel shows the time evolution of the position trajectories of the HL and the HF, while the lower panel presents the position trajectories of the corresponding VL and VF. It appears that the VL and the VF succeed in matching the kinematic characteristics of the HL and the HF in terms of the RMS value of their position error [ $RMS(HL, HF) = 0.0466$  and  $RMS(VL, VF) = 0.0497$ ] and the time lag between the two players [ $TL(HL, HF) = 0.09$  and  $TL(VL, VF) = 0.09$ ]. In addition, Table 3 gives the values of EMDs and describes the matching results quantitatively.  $Sig_L$  and  $Sig_F$  represent the signatures of HL and HF when playing solo, respectively. It shows that the proposed approach allows to replicate effectively the game dynamics between two humans playing the mirror game via two coupled VPs.

**Table 3** Matching performance of VPs in terms of EMD.

$EMD(Sig_L, HL)$	0.010	$EMD(Sig_F, HF)$	0.007
$EMD(Sig_L, VL)$	0.006	$EMD(Sig_F, VF)$	0.006
$EMD(HL, HF)$	0.0034	$EMD(VL, HL)$	0.0052
$EMD(VL, VF)$	0.0031	$EMD(VF, HF)$	0.0053

**Table 4** Indexes of temporal correspondence

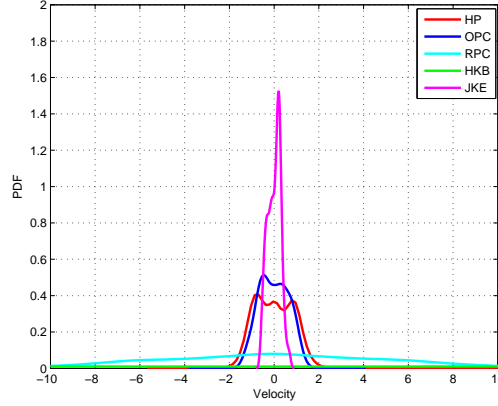
	HP	OPC	HKB	RPC	JKE
RPE	0.0914	0.0816	2.1767	0.3838	0.1467
CV	0.3011	0.1002	0.9400	0.7602	0.4859
TL	0.2035	0.1274	1.5192	0.0428	-0.9674

### 6.3.4 Comparison with existing models

In order to compare the VP models, we need to establish a benchmark, which describes the general kinematic characteristics of human participants in the mirror game. To this aim, 5 human participants were asked to track a prerecorded reference signal, and indexes of temporal correspondence were recorded to represent a benchmark of typical human dynamics. The existing VP models were then enabled to track the same reference. Corresponding indexes were computed for the VPs and compared with the benchmark to evaluate the proposed control algorithms. We adopted the following models to drive the VP: our strategy based on optimal control (OPC), Haken-Kelso-Bunz model (HKB) [17], reactive-predictive control (RPC) [2] and Jirsa-Kelso excitator (JKE) [7]. The parameter setting of existing VP models is the same as that in the relevant references. In Table 4, we show the benchmark of temporal correspondence and performance indexes of the corresponding VP models. On the whole, our algorithm performs best in terms of matching the benchmark among all the VP models. In addition to temporal correspondence, we also computed how similar the VP signature is to that of the HP during the benchmark experiment (see Fig.13). We find again that the optimal control strategy developed in this paper is the best in terms of replicating the human-like movement with an estimated  $EMD(HP, OPC) = 0.0184$ .

## 7 Conclusions

We presented the novel design of an interactive cognitive architecture able to drive a virtual player to play the mirror game against a human player. Two strategies were developed. The first, based on adaptive control, was shown to be effective



**Figure 13** PDF of velocity time series for different VP models.

to achieve temporal correspondence between the motion of the virtual player and that of the human individual. Convergence of the algorithm was proved. It was noticed that the adaptive control strategy does not allow the VP to exhibit some desired kinematic features (individual motor signature) of a given human player. To overcome this limitation, a different strategy based on the iterative solution of an appropriate optimal control problem was proposed. After proving boundedness and convergence of this additional approach, its effectiveness was tested experimentally. It was shown that the proposed strategy is able to drive the VP so as to play the game both as leader or follower while matching well the individual motor signature of a given individual. Finally, a comparison with other existing models was carried out confirming the effectiveness of the proposed approach. We wish to emphasize that our approach opens the possibility of making VPs, each modeling a different individual, play against each other and produce *in silico* experiments. This can reduce the cost and time of carrying out joint action experiments and can be effectively used to test different human-machine interaction scenarios via the mirror game.

## Acknowledgements

This work was funded by the European Project AlterEgo FP7 ICT 2.9 - Cognitive Sciences and Robotics, Grant Number 600610. The authors wish to thank Prof. Benoit Bardy, Dr. Ludovic Marin and Dr. Robin Salesse at EUROMOV, University of Montpellier, France for all the insightful discussions and for collecting some of the experimental data that is used to validate the approach presented in this paper.



## Supplementary Material

The solution of the optimal control algorithm in each time interval  $[t_k, t_{k+1}]$  can be transformed into a boundary value problem by applying Pontryagin's minimum principle [25]. We start by constructing the Hamiltonian as follows

$$H(X, u, \lambda) = \frac{1}{2}\theta_\sigma(\dot{x} - r_\sigma)^2 + \frac{1}{2}\eta_m u^2 + \lambda^T \begin{pmatrix} -(\alpha x^2 + \beta y^2 - \gamma)y - \omega^2 x + u \\ y \end{pmatrix}$$

where  $X = [x, \dot{x}]^T = [x, y]^T$  and  $\lambda = [\lambda_1, \lambda_2]^T$ . Using the minimum principle gives optimal open loop control

$$u^* = \operatorname{argmin}_{u \in R} H(X^*, u, \lambda) = -\eta_m^{-1} \lambda^T \begin{pmatrix} 0 \\ 1 \end{pmatrix} = -\eta_m^{-1} \lambda_2$$

and optimal state equation

$$\dot{X}^* = \nabla_\lambda H = \begin{pmatrix} -(\alpha x^{*2} + \beta y^{*2} - \gamma)y^* - \omega^2 x^* - \eta_m^{-1} \lambda_2 \\ y^* \end{pmatrix}$$

with initial condition  $X(t_k) = [x(t_k), \dot{x}(t_k)]^T$  and optimal costate equation

$$\dot{\lambda} = -\nabla_X H = \begin{pmatrix} \lambda_2(2\alpha x^* y^* + \omega^2) \\ \lambda_2(\alpha x^{*2} + 3\beta y^{*2} - \gamma) - \lambda_1 - \theta_\sigma(y^* - r_\sigma) \end{pmatrix}$$

with the terminal condition

$$\lambda(t_{k+1}) = \begin{pmatrix} \theta_p(x^*(t_{k+1}) - \hat{r}_p(t_{k+1})) \\ 0 \end{pmatrix}$$

Let  $\tilde{x}$  denote the approximation of the optimal solution  $x^*$ , then it is feasible to estimate the position error between the VP and the HP based on the collocation method as.

$$|x^* - \hat{r}_p| = |x^* - \tilde{x} + \tilde{x} - \hat{r}_p| \leq |x^* - \tilde{x}| + |\tilde{x} - \hat{r}_p|$$

Notice that  $|x^* - \tilde{x}|$  is negligible due to the high approximation accuracy of numerical methods [26]. In particular, considering that normally the optimal solution  $x^*$  is not available, the approximate solution  $\tilde{x}$  exactly corresponds to the position of the VP in the simulation. Thus, we mainly focus on the estimation of  $|\tilde{x} - \hat{r}_p|$ . For simplicity, we define  $\tilde{x}(t) = a_0 + a_1(t - t_k) + a_2(t - t_k)^2$ ,  $\lambda_1(t) = b_0 + b_1(t - t_k) + b_2(t - t_k)^2$  and  $\lambda_2(t) = c_0 + c_1(t - t_k) + c_2(t - t_k)^2$ , where  $a_i, b_i$  and  $c_i, i \in \{0, 1, 2\}$  are unknown constants and  $t \in [t_k, t_{k+1}]$ . Substituting  $\tilde{x}(t)$ ,  $\lambda_1(t)$  and  $\lambda_2(t)$  into the above optimal state equation and costate equation at the boundary points yields the linear matrix equation

$$A_k X_k = B_k \tag{16}$$

where

$$A_k = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \theta_p & \theta_p T & \theta_p T^2 & -1 & -T & -T^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & T & T^2 \\ 0 & 0 & 2 & 0 & 0 & 0 & \eta_m^{-1} & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & -(2\alpha x(t_k)y(t_k) + \omega^2) & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & -(\alpha x(t_k)^2 + 3\beta y(t_k)^2 - \gamma) & 1 & 0 \\ \theta_p & T\theta_p + \theta_\sigma & T(T\theta_p + 2\theta_\sigma) & 0 & 0 & 0 & 0 & 1 & 2T \\ 0 & 0 & 0 & 0 & 1 & 2T & 0 & 0 & 0 \end{pmatrix}$$

and

$$B_k = \begin{pmatrix} x(t_k) \\ y(t_k) \\ \theta_p \hat{r}_p \\ 0 \\ -(\alpha x(t_k)^2 + \beta y(t_k)^2 - \gamma)y(t_k) - \omega^2 x(t_k) \\ 0 \\ -\theta_\sigma(y(t_k) - r_\sigma(t_k)) \\ \theta_p \hat{r}_p + \theta_\sigma r_\sigma(t_{k+1}) \\ 0 \end{pmatrix}, \quad X_k = \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ b_0 \\ b_1 \\ b_2 \\ c_0 \\ c_1 \\ c_2 \end{pmatrix}$$

Solving equation (16) determines the vector of unknown constants

$$X_k = A_k^{-1} B_k$$

Thus, we obtain the approximate solution

$$\tilde{x}(t) = x(t_k) + y(t_k)(t - t_k) + \frac{\mathcal{N}}{\mathcal{D}}(t - t_k)^2, \quad t \in [t_k, t_{k+1}]$$

where

$$\begin{aligned} \mathcal{N} = & 2T \left[ \left( \frac{r_\sigma(t_k) + r_\sigma(t_{k+1})}{2} - y(t_k) \right) \theta_\sigma + (\hat{r}_p - x(t_k) - Ty(t_k)) \theta_p \right] \\ & - \eta_m \left( \frac{T^2 \omega^2}{2} + \alpha T^2 x(t_k)y(t_k) + \alpha T x(t_k)^2 + 3\beta T y(t_k)^2 - \gamma T + 2 \right) \\ & \cdot [(\alpha x(t_k)^2 + \beta y(t_k)^2 - \gamma)y(t_k) + \omega^2 x(t_k)] \end{aligned}$$

and

$$\mathcal{D} = 2T^2(\theta_p T + \theta_\sigma) + 2\eta_m \left( \frac{T^2 \omega^2}{2} + \alpha T^2 x(t_k)y(t_k) + \alpha T x(t_k)^2 + 3\beta T y(t_k)^2 - \gamma T + 2 \right)$$

Then we can compute

$$\begin{aligned}
|\tilde{x}(t_{k+1}) - \hat{r}_p(t_{k+1})| &= \lim_{t \rightarrow t_{k+1}} |\tilde{x}(t) - \hat{r}_p(t_{k+1})| \\
&= |x(t_k) + Ty(t_k) + \frac{\mathcal{N}}{\mathcal{D}}T^2 - \hat{r}_p(t_{k+1})| \\
&\leq T^2(1 - \theta_p) \frac{|2(x(t_k) - \hat{r}_p(t_{k+1})) + T(r_\sigma(t_k) + r_\sigma(t_{k+1}))|}{|\mathcal{D}|} + \eta_m \frac{|\mathcal{L} \cdot \mathcal{M}|}{|\mathcal{D}|}
\end{aligned} \tag{17}$$

where

$$\mathcal{L} = \frac{T^2\omega^2}{2} + \alpha T^2 x(t_k)y(t_k) + \alpha T x(t_k)^2 + 3\beta Ty(t_k)^2 - \gamma T + 2$$

and

$$\mathcal{M} = 2(x(t_k) + Ty(t_k) - \hat{r}_p(t_{k+1})) - T^2(y(t_k)(\alpha x(t_k)^2 + \beta y(t_k)^2 - \gamma) + \omega^2 x(t_k))$$

Since  $\hat{r}_p$ ,  $r_\sigma$ ,  $\mathcal{D}$ ,  $\mathcal{L}$  and  $\mathcal{M}$  are all bounded, it follows from inequality (17) that the bound on the tracking error  $|\tilde{x}(t_{k+1}) - \hat{r}_p(t_{k+1})|$  converges to 0 as  $\theta_p \rightarrow 1$  and  $\eta_m \rightarrow 0$ . Similarly, we can estimate the velocity error between the VP and the reference signal encoding the desired signature as follows

$$\begin{aligned}
|\dot{\tilde{x}}(t_{k+1}) - r_\sigma(t_{k+1})| &= \lim_{t \rightarrow t_{k+1}} |\dot{\tilde{x}}(t) - r_\sigma(t)| \\
&= |y(t_k) + \frac{2\mathcal{N}}{\mathcal{D}}T - r_\sigma(t_{k+1})| \\
&\leq (1 - \theta_\sigma) \frac{2T^2|T(y(t_k) - r_\sigma(t_{k+1})) + 2(\hat{r}_p(t_{k+1}) - x(t_k) - Ty(t_k))|}{|\mathcal{D}|} \\
&\quad + \theta_\sigma \frac{2T^2|r_\sigma(t_k) - y(t_k)|}{|\mathcal{D}|} + 2\eta_m \frac{|\mathcal{L} \cdot \mathcal{P}|}{|\mathcal{D}|}
\end{aligned} \tag{18}$$

where

$$\mathcal{P} = y(t_k) - r_\sigma(t_{k+1}) - T[(\alpha x(t_k)^2 + \beta y(t_k)^2 - \gamma)y(t_k) + \omega^2 x(t_k)]$$

According to inequality (18), the bound of the velocity error goes to 0 if  $\theta_\sigma \rightarrow 1$ ,  $\eta_m \rightarrow 0$  and  $r_\sigma(t_k) = y(t_k)$ .

## References

- [1] Johnstone, K. (1987) *Impro: Improvisation and the Theatre* (Theatre Arts Book, NY).
- [2] Noy, L., Dekel, E., Alon, U. (2011). The mirror game as a paradigm for studying the dynamics of two people improvising motion together. *Proceedings of the National Academy of Sciences*, 108(52), 20947-20952.

- [3] Folkes, V. S. (1982). Forming relationships and the matching hypothesis. *Personality and Social Psychology Bulletin*, 8(4), 631-636.
- [4] Lakens, D., Stel, M. (2011). If they move in sync, they must feel in sync: Movement synchrony leads to attributions of rapport and entitativity. *Social Cognition*, 29(1), 1-14.
- [5] Wiltermuth, S. S., Heath, C. (2009). Synchrony and cooperation. *Psychological Science*, 20(1), 1-5.
- [6] Hussain, S., Xie, S. Q., Jamwal, P. K. (2013). Adaptive impedance control of a robotic orthosis for gait rehabilitation. *IEEE Transactions on Cybernetics*, 43(3), 1025-1034.
- [7] Dumas, G., de Guzman, G. C., Tognoli, E., Kelso, J. S. (2014). The human dynamic clamp as a paradigm for social interaction. *Proceedings of the National Academy of Sciences*, 111(35), E3726-E3734.
- [8] <http://www.euromov.eu/alterego/>
- [9] Liu, Z., Chen, C., Zhang, Y., Chen, C. L. P. (2015). Adaptive neural control for dual-arm coordination of humanoid robot with unknown nonlinearities in output mechanism. *IEEE Transactions on Cybernetics*, 45(3), 521-532.
- [10] Jordan, M. I., Wolpert, D. M. (1999). Computational motor control.
- [11] Drop, F. M., Pool, D. M., Damveld, H. J., van Paassen, M. M., Mulder, M. (2013). Identification of the feedforward component in manual control with predictable target signals. *IEEE Transactions on Cybernetics*, 43(6), 1936-1949.
- [12] Laurence, V., Pool, D. M., Damveld, H. J., Van Paassen, M. R. M., Mulder, M. (2015). Effects of controlled element dynamics on human feedforward behavior in ramp-tracking tasks. *IEEE Transactions on Cybernetics*, 45(2), 253-265.
- [13] Desmurget, M., Grafton, S. (2000). Forward modeling allows feedback control for fast reaching movements. *Trends in cognitive sciences*, 4(11), 423-431.
- [14] Słowiński, P., Rooke, E., di Bernardo, M., Tanaseva-Atanasova, K. (2014). Kinematic characteristics of motion in the mirror game. *Proceedings of IEEE International Conference on Systems, Man and Cybernetics*, San Diego, California, USA, 748-753, October 5-8, 2014.
- [15] Słowiński, P., Zhai, C., Alderisio, F., Salesse, R. N., Gueugnon, M., Marin, L., Benoit, B. G., di Bernardo, M., Tsaneva-Atanasova, K. (2015). Dynamic similarity promotes interpersonal coordination in joint-action. Available arXiv preprint: [arxiv.org/abs/1507.00368](https://arxiv.org/abs/1507.00368)/euromov.eu/alterego.

- [16] Zhang, Z., Beck, A., Magnenat-Thalmann, N. (2014). Human-like behavior generation based on head-arms model for robot tracking external targets and body parts. *IEEE Transactions on Cybernetics*, 45(8), 1390-1400.
- [17] Kelso, J. S., de Guzman, G. C., Reveley, C., Tognoli, E. (2009). Virtual partner interaction (VPI): exploring novel behaviors via coordination dynamics. *PLoS ONE*, 4(6), e5749.
- [18] Hölldampf, J., Peer, A., Buss, M. (2009). Virtual partner for a haptic interaction task. *Human Centered Robot Systems* (pp. 183-191). Springer Berlin Heidelberg.
- [19] Haken, H., Kelso, J. S., Bunz, H. (1985). A theoretical model of phase transitions in human hand movements. *Biological cybernetics*, 51(5), 347-356.
- [20] Fuchs, A., Jirsa, V. K., Haken, H., Kelso, J. S. (1996). Extending the HKB model of coordinated movement to oscillators with different eigenfrequencies. *Biological Cybernetics*, 74(1), 21-30.
- [21] Schmidt, R. C., Richardson, M. J. (2008). Dynamics of interpersonal coordination. In *Coordination: Neural, behavioral and social dynamics* (pp. 281-308). Springer Berlin Heidelberg.
- [22] Levina, E., Bickel, P. (2001). The earth mover's distance is the Mallows distance: Some insights from statistics. *Proceedings of the 8th IEEE International Conference on Computer Vision* (Vol. 2, pp. 251-256), 2001.
- [23] Wang, W., Gao, Z. (2003). A comparison study of advanced state observer design techniques. *Proceedings of the American Control Conference* (Vol. 6, pp. 4754-4759), 2003.
- [24] Diedrichsen, J., Shadmehr, R., Ivry, R. B. (2010). The coordination of movement: optimal feedback control and beyond. *Trends in cognitive sciences*, 14(1), 31-39.
- [25] Naidu, D. S. (2002). *Optimal control systems* (Vol. 2). CRC press.
- [26] De Boor, C., Swartz, B. (1973). Collocation at Gaussian points. *SIAM Journal on Numerical Analysis*, 10(4), 582-606.
- [27] <https://www.leapmotion.com>.
- [28] Kreuz, T., Mormann, F., Andrzejak, R. G., Kraskov, A., Lehnertz, K., Grassberger, P. (2007). Measuring synchronization in coupled model systems: A comparison of different approaches. *Physica D: Nonlinear Phenomena*, 225(1), 29-42.
- [29] Orfanidis, S.J. (1996) *Optimum Signal Processing: An Introduction*. 2nd Edition. New York: McGraw-Hill.

- [30] Kramann, B., Cimponeriu, L., Rosenblum, M., Pikovsky, A., Mrowka, R. (2008). Phase dynamics of coupled oscillators reconstructed from data. *Physical Review E*, 77(6), 066205.