OPEN ACCESS

University of BRISTOL

Peer reviewed version

Link to published version (if available):
10.1109/ECC.2016.7810289

Link to publication record in Explore Bristol Research
PDF-document

University of Bristol - Explore Bristol Research
General rights

Comparing Different Control Approaches to Implement a Human-like Virtual Player in the Mirror Game

Francesco Alderisio[1], Dario Antonacci[2], Chao Zhai[1], and Mario di Bernardo[1,2]

Abstract— In this work we attempt to create a human-like virtual player able to play the mirror game with a human player in real time. A control architecture is developed in order to drive such virtual player to lead or track a human player during the game while maintaining a certain degree of similarity to its individual signature. The virtual player is able to exhibit diverse kinematic characteristics by integrating the relevant intrinsic dynamics. Two alternative control strategies are presented here to implement its cognitive architecture, namely PD control and a receding horizon optimal control strategy. To validate and compare the performance of these control strategies, we establish a benchmark based on experimental data collected from two human players so as to evaluate the human-like performance of the virtual player when playing together with a human. Experimental validation is provided showing the advantages and disadvantages of using different control strategies and models to drive the virtual player.

I. INTRODUCTION

Synergetic movements of two or more people mirroring each other frequently occur in many activities such as handling objects, manipulating a common workpiece, dancing and choir singing ([1], [2], [3] and references therein). In order to understand the mechanisms that underly the coordination between two humans, interpersonal coordination and synchronization between their motion have been extensively studied over the past few decades [4], [5], [6]. In particular, it has been suggested that the kinematic features characterizing the individual motion of players when they are not interacting with any other partner (solo trials) can be used to define the so called *kinematic* or *motor signature*, which is unique for each individual and does not vary much over time [7]. Motor signatures are well represented by the probability density function (PDF) of velocity profiles [8].

Recently, in the context of the EU Project *AlterEgo* [9], coordination tasks and synchronization games have been used as effective rehabilitation methods in order to help people suffering from social disabilities (that accompany schizophrenia, autism, or social phobia) get over their social anxiety and improve their social skills. In particular the *mirror game*, first introduced in [10] to study social

interaction between two human participants, has been employed as a simple yet powerful paradigm to implement such rehabilitation therapy. It can be played in two different experimental conditions: *Leader-follower* (LF), in which one player is designated as leader and the other one tries to follow his/her movements, and *Joint Improvisation* (JI), in which both players have to imitate each other and create interesting movements without any designation of leader and follower.

The use of a virtual player (VP) substituting one of the players in the game has been highlighted as an effective way forward to study the onset of coordination and to define a novel rehabilitation therapy [11], [12], [13], [14], [15]. An interesting control problem is then the design of an appropriate strategy to drive the VP to play the game with the human player (HP) while allowing it to exhibit some desired human-like kinematic characteristics described by their motor signature.

Some models to drive the motion of a VP interacting with a HP were investigated in [16], but in a different context and without any characterization of the motion of the former in terms of its kinematic signature. The aim of this paper is to present an informed comparison of different control strategies and models to drive the VP such that its motion exhibits human-like features. Firstly, we focus on understanding whether nonlinear dynamics are needed to replicate the behavior of a HP, or if a simpler linear model can be used instead. Then, we contrast the performance of the optimal control strategy presented in [13], [14] against that of a simpler PD control strategy. In so doing, we develop a robust methodology to evaluate the human-like performance of the VP in terms of its ability to reproduce a given kinematic signature and to guarantee an acceptable coordination level with the human player. To this aim, we compare all the possible combinations of inner dynamics models and control approaches. Finally, we propose some improvements to the existing control strategies that can be used to enhance their effectiveness and perform experiments to validate the comparison.

II. PROBLEM STATEMENT

The mirror game between a VP and a HP can be formulated as a multi-objective control problem [13], [14]. Given the dynamical system which describes the movement of the end-effector of the VP

$$\dot{X} = f(X, u, v) \tag{1}$$

with output equation

$$y = h(X) \tag{2}$$

[1]Francesco Alderisio, Chao Zhai and Mario di Bernardo are with Department of Engineering Mathematics, University of Bristol, Merchant Venturers' Building, Woodland Road, Clifton, BS8 1UB Bristol, United Kingdom. `f.alderisio@bristol.ac.uk`, `zhaichao@amss.ac.cn`, `m.dibernardo@bristol.ac.uk`

[2]Dario Antonacci is with Department of Electrical Engineering and Information Technology, University of Naples Federico II, 80125 Naples, Italy. Mario di Bernardo is also with Department of Electrical Engineering and Information Technology, University of Naples Federico II, 80125 Naples, Italy. `da.antonacci@studenti.unina.it`, `mario.dibernardo@unina.it`

where $X \in \mathbb{R}^2$ represents the state vector of the VP, $y \in \mathbb{R}$ its output, $u \in \mathbb{R}$ its control input and v is a vector containing the parameters describing the dynamical features of the VP, our aim is to develop a feedback control strategy

$$u = u(X, r_p, v, \theta) \tag{3}$$

so that the position error between the VP end-effector position and that of the HP, say r_p, is bounded

$$|y - r_p| \leq \varepsilon, \qquad t \in [0, t_{end}] \tag{4}$$

while at the same time a *similarity index* $\rho > 0$ (defined later in Section IV) satisfies

$$\rho(X) < \mu \tag{5}$$

Here, $\rho(X)$ denotes how similar the kinematic properties of the VP are to some reference value, θ is a vector containing tunable control parameters in the feedback controller, $\varepsilon > 0$ denotes the upper bound of the position error and $\mu > 0$ an upper bound for ρ during each trial of duration t_{end}.

Depending on the particular type of game to be performed, the VP needs to either track the HP while guaranteeing a predetermined kinematic signature (follower model), or lead him/her by generating a trajectory based on the signature of a given human player, while at the same time taking into account the behavior of the human follower to keep him/her engaged in the game (leader model). Hence, the main challenge is to develop a control architecture that enables a VP to behave like and to interact with a HP (following or leading him/her, according to the current type of game session) while guaranteeing a desired kinematic signature.

III. STRUCTURE OF THE VIRTUAL PLAYER

The cognitive architecture of our VP is mainly composed of two parts: a model of its intrinsic dynamics and an appropriate control algorithm monitoring the motion of the HP and driving that of the VP accordingly.

A. Intrinsic Dynamics

The first objective of this work is to find which of the following models, that describe a VP performing the mirror game, can better reproduce the behavior of a HP playing in a leader-follower condition:

- double integrator (DI)

$$\ddot{x} = u \tag{6}$$

- damped harmonic oscillator (DHO)

$$\ddot{x} + a\dot{x} + bx = u \tag{7}$$

- HKB oscillator [4]

$$\ddot{x} + (\alpha x^2 + \beta \dot{x}^2 - \gamma)\dot{x} + \omega^2 x = u \tag{8}$$

with x, $\dot{x}$ and $\ddot{x}$ representing position, velocity and acceleration of the VP, respectively, and u being the control signal modelling the interaction with the other player (see Section III-B).

B. Control Strategy

The second goal is understanding which control strategy best models the interaction between VP and HP, that is which one can replicate more accurately the interaction between two human participants performing a mirror game session. The control algorithm allows for two objectives: maximizing the temporal correspondence with the other player (e.g. minimizing the position error between HP and VP) and the kinematic similarity of the motion generated by the VP (e.g. minimizing the similarity index with a desired kinematic signature). We contrast two alternative strategies.

1) Optimal Control: The first control solution we consider is an optimal control strategy that aims at minimizing a certain cost function given by the sum of three terms: position error between the two players, error between the virtual player velocity and its reference kinematic signature, and control energy. Such cost function reads:

$$J(t_k) = \frac{1}{2}\theta_p(x(t_{k+1}) - \hat{r}_p(t_{k+1}))^2$$
$$+ \frac{1}{2}\int_{t_k}^{t_{k+1}} (1 - \theta_p)(\dot{x}(\tau) - \dot{r}_\sigma(\tau))^2 + \eta u(\tau)^2 d\tau \tag{9}$$

Here, $\hat{r}_p$ is the estimated human player position, $\dot{r}_\sigma$ is the VP's desired motor signature obtained by differentiating the position time series r_σ recorded during a human player solo trial, η is a positive tunable weight used to increase/decrease the control energy, t_k and t_{k+1} represent the current and the next optimization time instant (with $T_P = t_{k+1} - t_k$ representing the optimization time interval), and $\theta_p \in [0, 1]$ is a positive weight that can be used to make the VP take care more of temporal correspondence or similarity index, respectively. Indeed by setting $\theta_p = 1$, the second term in the cost function is equal to zero and the control signal only minimizes position error and control energy, thus making the system behave like a perfect follower. Vice versa, by setting θ_p to 0, we obtain a perfect leader. Motor signature and estimated human player's position and velocity are respectively computed as $\dot{r}_\sigma(t_k) = \frac{r_\sigma(t_k) - r_\sigma(t_{k-1})}{T}$, $\hat{r}_p(t_k) = r_p(t_{k-1}) + \hat{r}_v(t_{k-1})T$ and $\hat{r}_v(t_{k-1}) = \frac{r_p(t_{k-1}) - r_p(t_{k-2})}{T}$, where r_p is the position of the human player, $\hat{r}_v$ is his/her estimated velocity and $T = t_k - t_{k-1}$ is the sampling time.

2) PD Control: The second control strategy we take into account is a much simpler PD control, which aims at minimizing both the position error between the two players and the error between the virtual player's velocity and its reference motor signature, respectively:

$$u = K_p(r_p - x) + K_\sigma(\dot{r}_\sigma - \dot{x}) \tag{10}$$

Here, K_p and K_σ are two positive tunable gains representing the importance given to each of the two error terms.

In order to make a fair comparison between the results obtained with the two different control strategies, we first set the parameters of the cost function of the optimal controller, then we use the conclusions of Proposition 1 below to tune the PD control parameters accordingly.

Proposition 1. *Supposing that the sampling time T and the optimization interval T_P are small enough, the optimal control and the PD control solutions act almost in the same way on the virtual player if, for each of the models used (DI, DHO or HKB), the gains of the PD controller are set as follows:*

$$K_p = \begin{cases} \eta^{-1}T\theta_p, & VP: DI \\ \frac{\eta^{-1}T\theta_p}{1+aT+bT^2}, & VP: DHO \\ \frac{\eta^{-1}T\theta_p}{1+(\alpha x^2+3\beta y^2-\gamma)T+(2\alpha xy+\omega^2)T^2}, & VP: HKB \end{cases} \quad (11)$$

$$K_\sigma = \begin{cases} \eta^{-1}T(1-\theta_p), & VP: DI \\ \frac{\eta^{-1}T(1-\theta_p)}{1+aT+bT^2}, & VP: DHO \\ \frac{\eta^{-1}T(1-\theta_p)}{1+(\alpha x^2+3\beta y^2-\gamma)T+(2\alpha xy+\omega^2)T^2}, & VP: HKB \end{cases} \quad (12)$$

Proof. Let us first assume to employ the double integrator $\ddot{x} = u$ as end-effector model of the VP and let us consider the cost function (9). We formulate the Hamiltonian as

$$H(X,u,\lambda) = \frac{1}{2}\theta_\sigma(\dot{x}-\dot{r}_\sigma)^2 + \frac{1}{2}\eta u^2 + \lambda^T \begin{pmatrix} \dot{x} \\ u \end{pmatrix} \quad (13)$$

where $X = [x,\dot{x}]^T$, $\lambda = [\lambda_1, \lambda_2]^T$ and $\theta_\sigma = 1-\theta_p$. Applying the Pontryagin's minimum principle, we get the optimal control

$$u = -\eta^{-1}\lambda^T \begin{pmatrix} 0 \\ 1 \end{pmatrix} = -\eta^{-1}\lambda_2 \quad (14)$$

and the optimal costate equation

$$\dot{\lambda} = -\nabla_X H = \begin{pmatrix} 0 \\ -\lambda_1 - \theta_\sigma(\dot{x}-\dot{r}_\sigma) \end{pmatrix} \quad (15)$$

with the terminal condition

$$\lambda(t_{k+1}) = \begin{pmatrix} \theta_p(x(t_{k+1})-\hat{r}_p(t_{k+1})) \\ 0 \end{pmatrix} \quad (16)$$

Since $\dot{\lambda}_1 = 0$ and $\lambda_1(t_{k+1}) = \theta_p(x(t_{k+1})-\hat{r}_p(t_{k+1}))$, we get $\lambda_1(t_k) = \theta_p(x(t_{k+1})-\hat{r}_p(t_{k+1}))$. Thus,

$$\dot{\lambda}_2 = -\theta_p(x(t_{k+1})-\hat{r}_p(t_{k+1})) - \theta_\sigma(\dot{x}-\dot{r}_\sigma) \quad (17)$$

Moreover, assuming that the sampling time T is small and equal to the optimization interval T_P, we obtain

$$\frac{\lambda_2(t_{k+1})-\lambda_2(t_k)}{T} \approx \dot{\lambda}_2$$
$$= -\theta_p(x(t_{k+1})-\hat{r}_p(t_{k+1})) - \theta_\sigma(\dot{x}-\dot{r}_\sigma) \quad (18)$$

Knowing that $\lambda_2(t_{k+1}) = 0$, we get

$$\lambda_2(t_k) = T\theta_p(x(t_{k+1})-\hat{r}_p(t_{k+1})) + T\theta_\sigma(\dot{x}-\dot{r}_\sigma) \quad (19)$$

Hence, the optimal control can be written as:

$$u = \eta^{-1}T\theta_p(\hat{r}_p(t_{k+1})-x(t_{k+1})) + \eta^{-1}T\theta_\sigma(\dot{r}_\sigma-\dot{x}) \quad (20)$$

Now, assuming that the sampling time is sufficiently small, we have $\hat{r}_p(t_{k+1}) \approx r_p(t_k)$ and $x(t_{k+1}) \approx x(t_k)$, which leads to the PD controller

$$u = K_p(r_p-x) + K_\sigma(r_\sigma-\dot{x}) \quad (21)$$

where $K_p = \eta^{-1}T\theta_p$ and $K_\sigma = \eta^{-1}T\theta_\sigma$.

If the linear model $\ddot{x}+a\dot{x}+bx = u$ is used as end-effector model of the VP instead of the double integrator, reasoning in an analogous way we can derive

$$u = \frac{\eta^{-1}T}{1+aT+bT^2}[\theta_p(r_p-x)+\theta_\sigma(\dot{r}_\sigma-\dot{x})] \quad (22)$$

when the sampling time T is quite small and equal to optimization interval. Thus, the gains of the PD controller should be set as follows

$$K_p = \frac{\eta^{-1}T\theta_p}{1+aT+bT^2}, \quad K_\sigma = \frac{\eta^{-1}T\theta_\sigma}{1+aT+bT^2} \quad (23)$$

We omit the proof for the HKB oscillator for the sake of brevity. Furthermore, for the sake of simplicity, since the control parameters for the HKB oscillator would be time varying quantities depending on x and y, we set them equal to the constant ones used for the harmonic oscillator. $\square$

As a first step we tested both the control approaches while keeping all the parameters constant, then to further improve the performance of the VP in terms of its human-like behavior, we developed two strategies to adapt the optimal control cost function during the game. In particular:

1) we updated the value of θ_p at each time step depending on the current errors in terms of temporal correspondence with the HP and kinematic similarity with the desired signature;
2) we varied the optimization interval T_P depending on the trend of the VP motor signature.

IV. BENCHMARK DEFINITION

In order to measure how human-like the VP is, we first find a quantitative way of representing the behavior of a human player playing the mirror game in a leader-follower condition. In so doing, we make two human players perform three trials of the mirror game, each one lasting $60s$, from which we can evaluate some metrics to represent their level of interaction. We then compute their mean values in order to represent the average behavior shown in the game, that we shall refer to as the *human benchmark*. At this point we replace either of the two human players with a virtual agent (both in a leader and in a follower condition) and repeat the three $60s$ trials, so that we are able to evaluate again average values representing the interaction between VP and HP. We will say that the VP plays in a human-like way if the metrics obtained from the VP-HP interaction are similar to those obtained after the HP-HP interaction.

We chose four macroscopic observables as metrics, two measuring temporal correspondence and two kinematic similarity:

- Circular variance (CV), a measure of synchronization between the players based on the difference between their relative phases:

$$CV = \left| \frac{1}{n}\sum_{i=1}^{n} e^{i\Delta\Phi_i} \right| \quad (24)$$

where n is the number of sampling points in the simulation, $\Delta\Phi_i$ represents the relative phase between the two players at the i-th sampling step, and $|\cdot|$ denotes the absolute value. Note that $CV \in [0,1]$, where $CV = 1$ corresponds to perfect phase synchronization between the players and $CV = 0$ otherwise.

- Relative position error (RPE), that measures the difference in position in relation to the direction of the players' movement:

$$RPE_{x_1,x_2}(t) = \begin{cases} (x_1(t) - x_2(t))sgn(v_1(t)), & \\ \qquad sgn(v_1(t)) = sgn(v_1(t)) \neq 0 \\ |x_1(t) - x_2(t)|, & otherwise \end{cases}$$
(25)

  where x_i and v_i represent position and velocity of the i-th player [8].

- Earth mover's distance (EMD), a measure of the difference between the PDFs of two velocities:

$$EMD(p_1, p_2) = \int_Z |CDF_{p1}(z) - CDF_{p2}(z)|dz \quad (26)$$

  where p_i is the PDF of the velocity of the i-th player and CDF_{p_i} is the cumulative distribution function (CDF) of p_i [17]. In what follows we will consider two different EMDs: *1)* the EMD between leader and follower as a measure of the difference between the velocities of the two players when interacting together, and *2)* the EMD between the velocity profile exhibited by each player during the game and the respective kinematic signature exhibited when playing solo.

In order to define the human benchmark, we recorded the trajectories during three trials of HP-HP interaction, where one player (*player1*) acted as leader, and the other (*player2*) as follower. In order to get their motor signatures, both the players were separately asked to first perform three solo trials, each one lasting $60s$, trying to create interesting motions. All the parameters described above were computed, and their mean values were used to obtain two benchmarks, one describing the typical behavior of a HP playing as a leader, the other that of the HP playing as a follower:

- CV_b: mean value of the circular variances between the phases of leader and follower;
- RPE_b: mean value of the relative position errors between leader and follower;
- EMD_{LF}: mean value of the earth mover's distance between the PDFs of the velocities of leader and follower during the HP-HP interaction;
- EMD_{PS}: mean value of the earth mover's distance between the PDFs of the velocity of leader (or follower) during the HP-HP interaction and his/her own kinematic signature.

V. EXPERIMENTAL VALIDATION

A. Parameter Setting

To compare the models highlighted in the previous section, we set their parameters so that their dynamics is comparable:

- The dynamics of the DI are compared with those of the DHO and HKB where all parameters are set to 1 (we will term these models as *DHO*1 and *HKB*1, respectively);
- When comparing the DHO with an HKB we consider identified parameter sets that make the two systems exhibit similar responses in time and frequency to a signal with a rich harmonic spectrum. The parameters we have obtained via trial and error techniques are $a = 1.5$ and $b = 3$ for the DHO and $\alpha = 3$, $\beta = 7$, $\gamma = 0.2$ and $\omega = 1.75$ for the HKB oscillator. We will refer to these models as *DHO*2 and *HKB*2, respectively.

Using these inner dynamics parameters and setting the control parameters of the optimal controller cost function as $\eta = 10^{-4}$, $T = 0.05$, $T_P = 0.1$ for both the follower and the leader model, $\theta_p = 0.85$ for the follower and $\theta_p = 0.15$ for the leader, it is possible to obtain the values shown in Table I for the PD control signal's weights K_p and K_σ by making use of the relations in (11) and (12).

B. Experimental Set-up of the Mirror Game

The human leader's trajectories have been recorded by using a Leap Motion Controller®, a motion sensor used to capture the movements of their hands, while for the human follower a touchpad has been employed. The two participants have been asked to sit in front of a laptop monitor and move their index finger horizontally in a range of $40cm$ (monitor width); the leader was asked to move approximately $30cm$ above the Leap Motion Controller®. Their positions were displayed on the screen during all the game session (see [11] for more details on the experimental set-up).

C. Comparison Method

We simulated all the possible combinations of inner dynamics and control signals described in Section III, numbering ten different possibilities given by five models (*DI*, *DHO1*, *DHO2*, *HKB1*, *HKB2*) and two control signals (optimal control and PD control). For each combination we ran three simulations, one for each motor signature the virtual player was fed with. We used two different methods to carry out the experiments:

- in order to evaluate the behavior of the VP as follower, we made it follow each of the three position time series recorded during the HP-HP interaction from *player1* (human leader), while feeding it with each of the three kinematic signatures evaluated from *player2* (human follower) when playing solo;

TABLE I
PD CONTROL PARAMETERS

Condition	K	DI	DHO1	DHO2	HKB1	HKB2
FOLLOWER	K_p	425	403.8	392.6	403.8	392.6
	K_σ	75	71.26	69.28	71.26	69.28
LEADER	K_p	75	71.26	69.28	71.26	69.28
	K_σ	425	403.8	392.6	403.8	392.6

- in order to assess the performance of the VP as leader, we fed it with each of the three kinematic signatures evaluated from *player1* when playing solo, and asked *player2* to follow the VP in three game sessions.

Then we evaluated CV, RPE and the two EMDs on these trajectories and, after each triplet of simulations, we computed their mean values and compared them with the human benchmark in order to see which combination of model and controller makes the VP behave more similarly to a HP playing in an analogous condition.

D. Results and Analysis

In order to evaluate the results, we computed four percentage quantities:

$$\Delta CV = 10^2 \left|1 - \frac{CV_V}{CV_b}\right|, \ \Delta RPE = 10^2 \left|1 - \frac{RPE_V}{RPE_b}\right|$$

$$\Delta EMD_{VH} = 10^2 \left|1 - \frac{EMD_{VH}}{EMD_{LF}}\right|, \ \Delta EMD_{VS} = 10^2 \left|1 - \frac{EMD_{VS}}{EMD_{PS}}\right|$$

with CV_b, RPE_b, EMD_{LF} and EMD_{PS} representing the benchmark quantities described above, and with CV_V, RPE_V, EMD_{VH} and EMD_{VS} representing the corresponding metrics obtained when replacing one of the two HPs with a VP, respectively. These percentage quantities define how the VP playing the game compares with the human benchmark in each of the four metrics we chose to characterize the motion. We use the relative diagrams shown in Fig.1 to summarize the results.
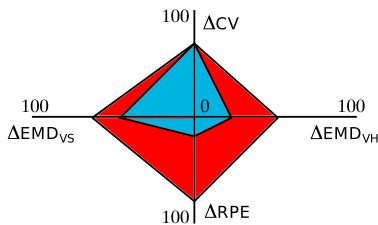

Fig. 1. Two different virtual players' relative diagrams represented in blue and red, overlapped. On each axis we plot one of the percentage quantities so that the area defined by the diagram is representative of how similar or dissimilar the VP's motion is to the human benchmark. Smaller areas, denoted with ΔA, will correspond to better matching with the human benchmark (blue diagram in this case).

The results of 60 simulations (10 possible combinations of inner dynamics and control for the follower model and 10 for the leader model, 3 trials for each combination) with constant values for the parameters of the two control signals are presented in Table II.

Looking at ΔA as an overall index of similarity to the human-like behavior, these results point out that the virtual player behaving the most similarly to the human follower is the one obtained by combining *HKB2* with the PD control, while if the aim is to reproduce the behavior of a human leader, *HKB2* combined with the optimal control seems to be the best choice. Other combinations are also possible if the aim is to specifically minimize some of the four metrics more than the others.

E. Control Design Improvements

In order to improve the control performance, we present two possible alternatives to the design of the optimal control law.

1) *Updating θ_p*: Since increasing the value of θ_p during the game results in a better tracking of the other player's position (smaller RPE_V) and a worse tracking of the motor signature (larger EMD_{VS}) and vice versa, we propose to switch θ_p in real time during the game to larger/smaller values according to the current values of RPE_V and EMD_{VS} during the HP-VP interaction. We used three different values for θ_p depending on the four possible situations presented in Fig. 2.
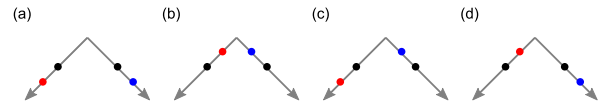

Fig. 2. Four different scenarios for RPE_V and EMD_{VS} during a game session. In (a) and (b) both RPE_V (red dot) and EMD_{VS} (blue dot) are larger/smaller than the human benchmark (black dot), respectively, hence increasing or decreasing θ_p will move one of the two metrics away from it. In (c) a higher value of θ_p brings both the metrics closer to the desired value. In (d) a lower value of θ_p is needed to get both the metrics closer to the benchmark.

Namely, we propose to set:

$$\theta_p(t) = \begin{cases} \theta_{max}, \ RPE_V(t) > RPE_b, \ EMD_{VS}(t) < EMD_{PS} \\ \theta_{min}, \ RPE_V(t) < RPE_b, \ EMD_{VS}(t) > EMD_{PS} \\ \frac{\theta_{max}+\theta_{min}}{2}, \ \text{otherwise} \end{cases}$$

(27)

where θ_{max} is equal to 1 for a virtual follower and to 0.3 for a virtual leader, while θ_{min} is equal to 0.7 for a virtual follower and to 0 for a virtual leader.

2) *Varying T_P*: Depending on their own motor signature, human players can find it easier/harder to anticipate their partners' movements in a mirror game session based on their current position and velocity. Thus we

TABLE II

VIRTUAL PLAYER RESULTS

Control	Metric	DI	DHO1	DHO2	HKB1	HKB2
OPTIMAL CONTROL *Follower*	ΔCV	10%	9%	9%	0.2%	0.2%
	ΔRPE	17%	18%	19%	4%	4%
	ΔEMD_{VH}	46%	37%	30%	36%	36%
	ΔEMD_{VS}	40%	42%	43%	46%	46%
	ΔA	1172	1089	1030	182	182
PD CONTROL *Follower*	ΔCV	3%	4%	4%	4%	3%
	ΔRPE	3%	3%	3%	3%	3%
	ΔEMD_{VH}	2%	4%	6%	7%	0.6%
	ΔEMD_{VS}	47%	47%	48%	48%	47%
	ΔA	155	172	187	201	147
OPTIMAL CONTROL *Leader*	ΔCV	10%	13%	14%	22%	35%
	ΔRPE	70%	66%	61%	106%	90%
	ΔEMD_{VH}	642%	632%	575%	15%	0.4%
	ΔEMD_{VS}	71%	69%	82%	83%	81%
	ΔA	28182	27675	24519	6305	5114
PD CONTROL *Leader*	ΔCV	72%	61%	41%	62%	60%
	ΔRPE	102%	87%	100%	88%	81%
	ΔEMD_{VH}	15%	32%	19%	18%	17%
	ΔEMD_{VS}	79%	65%	69%	78%	75%
	ΔA	8192	7224	6165	7192	6556

TABLE III

ADAPTIVE CONTROL PARAMETERS

Condition	Metric	Updating θ_p	Constant θ_p
FOLLOWER Adaptive θ_p	ΔCV	12%	0.2%
	ΔRPE	13%	4%
	ΔEMD$_{VH}$	5%	36%
	ΔEMD$_{VS}$	4%	46%
	ΔA	112	182

Condition	Metric	Varying T_P	Constant T_P
LEADER Adaptive T_P	ΔCV	6%	35%
	ΔRPE	103%	90%
	ΔEMD$_{VH}$	23%	0.4%
	ΔEMD$_{VS}$	59%	81%
	ΔA	4486	5114

propose to adapt the optimization interval T_P by setting it to the temporal distance between two consecutive zero-crossings of the signature (Fig. 3) and, in order to prevent T_P from becoming too large, we set an upper bound $T_{P_{max}} = 0.2s$, which has been shown to be equal to the reaction time of a human player [18].
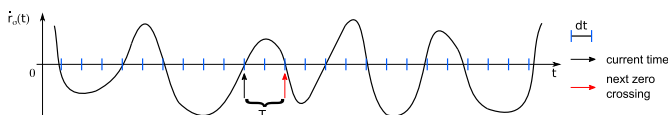

Fig. 3. Time varying optimization interval. T_p is updated by setting it to the value of the temporal distance between two consecutive zero crossings of the motor signature.

The results obtained by adapting the control parameters of the cost function during the game session are presented in Table III. We tested these strategies only on *HKB2* since this system has been proven to be the best choice when using the optimal control. If we evaluate the human-like performance of the VP by the measure of the area ΔA, such results show an improvement of 38% in the virtual follower behavior when using the updating strategy for θ_p, and an improvement of 12% in the virtual leader when using the time-varying strategy for T_P.

VI. CONCLUSIONS

We presented combinations of different models and control strategies that can be used to drive a virtual agent able to lead or follow a human playing the mirror game while exhibiting human-like motion features defined through a benchmark dataset. We proposed the use of a set of new metrics to assess the performance of the two control strategies considered in this paper and suggested some improvements that can lead to better performance. We found that the best combination of model and control law depends on the game conditions and the metrics that one wants to minimize, although the HKB model used in conjunction with an optimal control law seems to be the most versatile and effective strategy to drive the virtual player. The analysis and the implementation of an appropriate control strategy capable of driving the VP to perform joint improvisation with a HP are currently being explored in [15], while networks of HKB oscillators describing coordination in a group of several people are the subject of investigation in [19].

REFERENCES

[1] P. Valdesolo, J. Ouyang, and D. DeSteno, "The rhythm of joint action: Synchrony promotes cooperative ability," *Journal of Experimental Social Psychology*, vol. 46, no. 4, pp. 693–695, 2010.
[2] A. Mörtl, T. Lorenz, B. N. Vlaskamp, A. Gusrialdi, A. Schubö, and S. Hirche, "Modeling inter-human movement coordination: synchronization governs joint task dynamics," *Biological cybernetics*, vol. 106, no. 4-5, pp. 241–259, 2012.
[3] T. Lorenz, A. Mörtl, and S. Hirche, "Movement synchronization fails during non-adaptive human-robot interaction," in *Proceedings of the 8th ACM/IEEE International Conference on Human-Robot Interaction*. IEEE Press, 2013, pp. 189–190.
[4] H. Haken, J. S. Kelso, and H. Bunz, "A theoretical model of phase transitions in human hand movements," *Biological cybernetics*, vol. 51, no. 5, pp. 347–356, 1985.
[5] R. Schmidt and M. T. Turvey, "Phase-entrainment dynamics of visually coupled rhythmic movements," *Biological cybernetics*, vol. 70, no. 4, pp. 369–376, 1994.
[6] M. Varlet, L. Marin, J. Lagarde, and B. G. Bardy, "Social postural coordination," *Journal of Experimental Psychology: Human Perception and Performance*, vol. 37, no. 2, pp. 473–483, 2011.
[7] P. Słowiński, E. Rooke, M. di Bernardo, and K. Tanaseva-Atanasova, "Kinematic characteristics of motion in the mirror game," in *Proceedings of the 2014 IEEE International Conference on Systems, Man and Cybernetics (SMC)*. IEEE, 2014, pp. 748–753.
[8] 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*, vol. 13, no. 116. doi: 10.1098/rsif.2015.1093, 2016.
[9] Euromov.eu, "What is AlterEgo project ?", 2013. [Online]. Available: http://www.euromov.eu/alterego/project. [Accessed: 30- Jul- 2015].
[10] 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*, vol. 108, no. 52, pp. 20 947–20 952, 2011.
[11] 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 *Proceedings of the 2014 IEEE International Conference on Systems, Man and Cybernetics (SMC)*. IEEE, 2014, pp. 754–759.
[12] ——, "Adaptive tracking control of a virtual player in the mirror game," in *Proceedings of the 2014 IEEE 53rd Annual Conference on Decision and Control (CDC)*. IEEE, 2014, pp. 7005–7010.
[13] ——, "A model predictive approach to control the motion of a virtual player in the mirror game," in *Proceedings of the 2015 IEEE 54th Annual Conference on Decision and Control (CDC)*. IEEE, 2015, pp. 3175–3180.
[14] C. Zhai, F. Alderisio, P. Słowiński, K. Tsaneva-Atanasova, and M. di Bernardo, "Design and validation of a virtual player for studying interpersonal coordination in the mirror game," *arXiv preprint arXiv:1509.05881*, 2015.
[15] ——, "Modeling joint improvisation between human and virtual players in the mirror game," *arXiv preprint arXiv:1512.05619*, 2015.
[16] 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*, vol. 111, no. 35, pp. E3726–E3734, 2014.
[17] E. Levina and P. Bickel, "The earth mover's distance is the mallows distance: some insights from statistics," in *Proceedings of the 8th IEEE International Conference on Computer Vision (ICCV)*, vol. 2. IEEE, 2001, pp. 251–256.
[18] R. J. Kosinski, "A literature review on reaction time," *Clemson University*, vol. 10, 2008.
[19] 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*, 2016. DOI: 10.1007/s00422-016-0685-7.