A Computational Model for Synchronous Motion Imitation by Humans: The Mirror Controller Applied on Stepping Motions
Mehdi Benallègue, Pierre-Brice Wieber, Abderrahmane Kheddar, Bernard Espiau

To cite this version:

HAL Id: inria-00567736
https://inria.hal.science/inria-00567736
Submitted on 21 Feb 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
A Computational Model for Synchronous Motion Imitation by Humans: The Mirror Controller Applied on Stepping Motions

Mehdi Benallegue, Pierre-Brice Wieber, Abderrahmane Kheddar and Bernard Espiau

Abstract—We propose a new computational model describing the motion imitation by humans, and which parallels the direct matching hypothesis in Neuroscience. To illustrate our computational model, we present a scheme where a robot imitates another robot achieving stepping motions on horizontal plane. The idea is to map an observed action onto the robot’s motor representation of the same action: the motion controller. We propose a simple state observed based implementation of this approach. Simulation results, where a HRP-2 robot tries to imitate another HRP-2 robot, show that imitation can be realized with virtually no delay, in perfect synchrony.

I. INTRODUCTION

Biological mechanisms of motion imitation in humans and animals are subject to increasingly active studies in Neuroscience, giving birth to the direct matching hypothesis, which “holds that we understand actions when we map (...) the observed action onto our motor representation of the same action” [13]. A potential neurophysiologic grounding of this hypothesis has been evidenced in groups of neurons of the premotor cortex called mirror neurons, that activate selectively both during the execution of given actions and during the observation of the same actions executed by others [13]. Recent studies suggest then that “the goal of an action might be more important for mirror activations than the way in which the action is performed” [3]1...

Many questions have been raised by this hypothesis. For example, how much “a tight match between the observed movements and the observer’s motor programs to be activated” [3] is required? Robots have been already used as alternate avatars for human being, and the effect of a potential mismatch between a human and a robot motor control has been investigated from the point-of-view of the human mirror neuron system [3]. However these questions can not be assessed easily in their original biological setting and require a computational context to be investigated.

Our approach aims at (i) understanding motion imitation of human by humans, and (ii) devise a computational model that can serve humanoid (robotics or virtual avatars) imitation of human motions. To do so, we first define a parallel imitation scheme: a humanoid robot (imitator) observing a motion performed by another robot (demonstrator) using a set of sensors, and imitating this motion on-line.

First of all, following the assumptions of the direct matching theory, the imitator needs to map the demonstrator’s motion onto it’s own conception of the way to achieve the same task. This conception relies trivially on the imitator’s motion controller. Our computational model makes the imitator project the demonstrator’s motions on its own motion controller, i.e. the imitator assumes that the demonstrator is controlled by the exact duplicate of the imitator’s controller. In order to underline this biological equivalence, we propose to name our imitation scheme, the mirror controller.

In order to achieve the synchronous imitation goal, the imitator has to “interpret”, on-line, the demonstrator’s motion in the control-space of its own motion controller. This interpretation reduces to extract the input parameters which give the closest motion to the demonstrator’s action. The good news is that in some cases, this parameter identification problem can be tackled on-line with a very small delay, as will be discussed in the Section III.

Of course, imitation raises many other non-trivial questions such as “what, how, when and who to imitate” [1], but we will focus here on this very specific problem: the imitator must execute a sequence of steps as shown by the demonstrator. The next section presents the walking motion scheme that the imitator would imitate; in section III the imitation method is described. Simulation results are presented in IV and in-depth discussion and conclusion are in section V and VI.

II. PREDICTIVE CONTROL SCHEME FOR WALKING

Let’s consider the case where the motion to imitate is a walking motion on a horizontal plane without obstacles. Instead of a simple cyclic walk, steps can be chosen randomly with any step-length: e.g. a simple chaotic dance without any predefined style. One of the most efficient walking motion controllers for humanoid robots is based on predictive control and is able to cope with the simple chaotic dance proposed as a scenario of our study. It relies on a Linear Quadratic Regulation (LQR) of the motion of the Center of Mass (CoM) \( \xi_C \) of the robot [6].

A. Controller taking relative positions of steps as input parameters

The original LQR in [6] doesn’t ensure that the resulting walking motion is stable in all cases, so we consider here a safer version, similar to what has already been proposed in [14], which ensures explicitly that the Center of Pressure
(CoP) $z_k$ is always right inside the support surface. But with this controller, the rhythm of the steps is imposed, so the imitator will be able to achieve and imitate only dance motions with a prescribed rhythm. The control parameters $p_k$ can be taken to be the relative positions of the steps. We define the predictive horizon as the current and next relative step positions.

$$p_k = \begin{pmatrix} \delta_{i_k} \\ \delta_{i_{k+1}} \end{pmatrix}$$  \hspace{1cm} (1)

where $j_k$ is the step number at instant $t_k$.

The dynamic state of the robot includes the 2D position, velocity and acceleration of the CoM

$$c_k = \begin{pmatrix} \xi_k \\ \dot{\xi}_k \\ \ddot{\xi}_k \end{pmatrix}$$  \hspace{1cm} (2)

and we are interested in motions with a piece-wise constant jerk, so that we can consider the linear dynamics with the sampling time $T$:

$$c_{k+1} = Ac_k + b\ddot{\xi}_k$$  \hspace{1cm} (3)

with

$$A = \begin{bmatrix} 1 & T & T^2/2 \\ 0 & 1 & T \\ 0 & 0 & 1 \end{bmatrix}$$  \hspace{1cm} and \hspace{1cm} $$b = \begin{bmatrix} T^3/6 \\ T^2/2 \\ T \end{bmatrix}. \hspace{1cm} (4)$$

Reducing the multi-body dynamics of the robot into a single point-mass moving at a constant height $h_{CoM}$ [14], the position of the CoP can be obtained with a simple linear function:

$$z_k = \ddot{\xi}_k - \frac{h_{CoM}}{g} \dddot{\xi}_k$$  \hspace{1cm} (5)

where $g$ is the gravity.

To ensure dynamic stability, the CoP lies within the support area $S$. The robustness of the walking motion is highest when the CoP’s reference $z^*_k$ is in the middle of the supporting foot during single support, and move smoothly to the next supporting foot during double support [6]. With this choice, this reference trajectory depends in a linear way on the relative position of the steps $p_k$. $S$ can be assumed to have a polygonal shape so that the CoP position constraint can be written as linear inequalities.

With this setting, our LQR scheme minimizes the jerk of the CoM and the deviation of the CoP with respect to a reference $z^*_k$ over a horizon of $N$ future samples while maintaining the CoP within the support surface:

$$\min_{\dddot{\xi}_i, i = k:k+k+N} \sum_{i=k}^{k+N} \left( \frac{\alpha}{2} \| \dddot{\xi}_i \|^2 + \frac{\beta}{2} \| z_{i+1} - z^*_{i+1} \|^2 \right)$$  \hspace{1cm} (6)

under the constraint $z_{k+1} \in S$  \hspace{1cm} (7)

A recursive application of the dynamics (3) allows computing the position $z_k$ of the CoP over the whole horizon from the jerks $\dddot{\xi}_k$ over this horizon and the state $x_k$ at the beginning of this horizon. The solution $\dddot{\xi}_k$ to the LQR (6)-(7) can be expressed then as a function of $x_k$ and $p_k$. Considering the linear structure of this constrained LQR, this solution appears to be a piece-wise linear function. In the most frequent case the CoP $z_k$ stays away from the border of the support surface thanks to the centered reference $z^*$. This allows to keep for the longest periods the linear property of the solution [14]:

$$\dddot{\xi}_k = d^T(P \dddot{c}_k - V_k p_k - W_k q_k)$$  \hspace{1cm} (8)

Combined with the dynamics (3), it relates piece-wise linearly the evolution of the state $x_k$ to the control parameters $p_k$. The dynamic state $x_k$ of the robot during walking control should also include at least the instant position $\gamma_k$ of its feet, which are the limbs involved into ground contact, together with their velocity and acceleration. We propose to steer the motion of the feet by a smooth polynomial interpolation between its current state and the desired state at the end of the step: a position given by the parameter $p_k$ (position of the steps), with a vanishing velocity and acceleration.

B. Controller taking CoM velocity as an input parameter

Let’s consider a second controller that no longer supports an input of relative positions of steps, but directly the desired walking speed expressed as an on-line reference velocity $\dddot{\xi}^*_k$. A method that generates walking motions by tracking this reference trajectory is devised in [5]. Such an approach implies to have a control on steps position, a fortiori when a stepping rhythm is imposed. Hence, the controller determines on-line the steps position $p_{k+1}$, these positions can be expressed using the CoP reference trajectory $z^*_k$ centered on the support foot. So the new LQR scheme minimizes the deviation of the velocity from the reference, centering the position of the feet in the CoP and slightly smoothes the trajectory by minimizing the jerk.

$$\min_{\dddot{\xi}_i, i = k:k+k+N} \sum_{i=k}^{k+N} \left( \frac{\alpha}{2} \| \dddot{\xi}_i \|^2 + \frac{\beta}{2} \| \dddot{\xi}_{i+1} - \dddot{\xi}^*_{i+1} \|^2 + \frac{\gamma}{2} \| z_{i+1} - z^*_{i+1} \|^2 \right)$$  \hspace{1cm} (9)

under the constraint $z_{k+1} \in S$  \hspace{1cm} (10)

This optimization computes automatically the CoM jerk trajectory and the steps positions. This problem can be expressed as a canonical Quadratic Program. (Details may be found in [5].)

III. A WALKING MOTION IMITATION BASED ON STATE OBSERVER

We describe the imitation implementation using the mirror controller. Let’s define the system’s dynamics with $x$ the full dynamic state of a demonstrator including the center of mass ($\xi_k$)’s dynamics and the feet position $\gamma_k$, velocity and acceleration

$$x_k = (\xi_k \quad \dot{\xi}_k \quad \ddot{\xi}_k \quad \gamma_k \quad \dot{\gamma}_k \quad \ddot{\gamma}_k)^T$$  \hspace{1cm} (11)

and $p$ control parameters in its sensorimotor control system that we’re interested in, defined in equation (1). Denoting $y$ the measures obtained from some sensors on its body, we have the following relationships between the corresponding sampled signals:

$$x_k = f_k(x_{k-1}, p_k),$$  \hspace{1cm} (12)

$$y_k = h_k(x_k) + v_k,$$  \hspace{1cm} (13)
where \( h_k \) is the sensors’ behavior, \( v_k \) a Gaussian noise, and \( f_k \) the closed loop dynamics of the demonstrator, meaning its dynamics including its sensorimotor control system described in the previous section. While imitating motions online, we have no idea on the value and evolution of the putative control parameters \( p \). For some predictable tasks, we may have some prior knowledge on these parameters and how they evolve, but in the case of general stepping motions, without prior knowledge, we probably can’t do much better than suppose that they follow some random walk:

\[
p_k = p_{k-1} + w_{k-1} \tag{14}
\]

with some Gaussian noise \( w_k \). The mirror controller approach means that the imitator’s conjecture \( \hat{p}_k \) on the demonstrator’s control parameters \( p_k \) will be based on the imitator’s own closed loop dynamics, including its own sensorimotor control system. Considering only local corrections to these conjectures, we are led to a classical state observer design:

\[
\hat{x}_k = f_k(\hat{x}_{k-1}, \hat{p}_k) + L_k(y_{k-1} - \hat{y}_{k-1}), \tag{15}
\]

\[
\hat{y}_k = h_k(\hat{x}_k), \tag{16}
\]

\[
\hat{p}_k = \hat{p}_{k-1} + K_k(y_{k-1} - \hat{y}_{k-1}) \tag{17}
\]

where \( \hat{x}_k \) is the imitator’s conjecture on the state of the demonstrator, steered by the closed loop dynamics \( f_k \) of the imitator, and \( L_k \) and \( K_k \) are observer gains. We show in the next section that in some cases, convergence of \( \hat{x}_k \) to \( x_k \) is difficult to obtain using this observer dynamics. However, are we interested in this convergence? What should be the goal of our imitation scheme? The goal we propose to the imitator is to match the control parameters \( p_k \), without focusing more precisely on details of the motion \( x_k \). But the evolution (17) of the conjecture \( \hat{p}_k \) is driven by the difference \( y_k - \hat{y}_k \) which itself depends on the difference between the states \( x_k \) and \( \hat{x}_k \), so the convergence of the state \( \hat{x}_k \) and the convergence of the control parameter \( \hat{p}_k \) appear to be tightly coupled. However, reintroducing the dynamics (12) and (15) in the sensors’ behaviors (13) and (16), we have

\[
y_k = h_k(f_k(x_{k-1}, p_k) + v_k \tag{18}
\]

\[
\hat{y}_k = h_k(f_k(\hat{x}_{k-1}, \hat{p}_k)) \tag{19}
\]

where we can see that an option could be to select a sensor system that is more sensitive to differences between the underlying control parameters \( p_k \) and \( \hat{p}_k \) than it is to the differences between the states \( x_k \) and \( \hat{x}_k \). That way, the sensors’ behaviors (13) and (16) could be replaced by

\[
y_k = g_k(p_k) + v_k \tag{20}
\]

\[
\hat{y}_k = g_k(\hat{p}_k) \tag{21}
\]

and the convergence of the control parameter \( \hat{p}_k \) is decoupled from the convergence of the state \( \hat{x}_k \). This is definitely a strong hypothesis, but we will see in the following section that it’s not unrealistic.

The convergence of \( \hat{x}_k \) to \( x_k \) using the observer dynamics (15) appears now to be not required, so we have no real need for the gain \( L_k \), which can therefore vanish:

\[
L_k = 0. \tag{22}
\]

The interest of doing so is that the dynamics (15) boils down then to the original closed loop dynamics of the imitator, so when the control parameters \( \hat{p}_k \) are fed to the imitator controller, its dynamics will be exactly the dynamics (15), and the state \( \hat{x}_k \) will be exactly the state of the imitator.

![Diagram](image)

Fig. 1. A simple stochastic process representing the evolution of the control parameters \( p \) (the block \( S \) is a simple time shift), and the Extended Kalman Filter which is used to identify the value of this parameter with the help of sensor signals \( y \).

We end up having a simple stochastic process (14) and (20) for which we will consider a classical Extended Kalman Filter (EKF) implementation for the observer (17) and (21) (Figure 1).

IV. SIMULATION RESULTS

A. Robots having the same controller

We simulated a demonstrator and imitator with two HRP-2 robots [7], both controlled by the motion controller presented in Section II-A, which we denote controller A. The demonstrator realizes random sequences of steps at a known constant rhythm, one step every 0.8 s. The sensors we simulate consist of accelerometers. It is not easy to observe the demonstrator’s \( x_k \) using accelerometers because of the unavoidable drift generated by a numeric integration; but this convergence is not our primary concern. We simulate three 400Hz noisy bi-axial accelerometers, one on the waist and one on each foot; in order to be the most sensitive possible to the state (the CoM and the feet) variations which are directly linked to the underlying input parameter \( p_k \) described in 1. The imitator realizes the sequence of steps identified by this observer without delay.

The imitator reproduced the stepping of the demonstrator with good precision, but most of all, with virtually no delay. In a first sequence, only forward and backward steps with random lengths were considered, and one sees in Figures 2 and 3 the evolution in time of the control parameters \( p_k \) of the demonstrator in thin blue line and of the control parameters \( \hat{p}_k \) of the imitator in thick red line. The sensors have a white noise of 3cm/s² standard deviation. The length of the
current step is shown in Figure 2, and it appears to be tracked very closely by the imitator (1mm of mean error). But the most striking feature is that the length of the current step appears each time to be precisely identified ever since the very beginning of the step: the thick red line curve matches the thin blue line curve without delay. It’s interesting then to have a look at Figure 3 where the length of the next step to be executed is shown: we can see that the imitator can even make a guess on the next step before it begins to be executed.

More complex sequences are undertaken then, with steps taken in random directions, and we can see in figure 5 that the imitation is of good quality (23mm of mean error), although less than when only steps forward were considered.

B. Robots having different controller

In a second scenario, the same scheme is reproduced, except that we equip the demonstrator by the controller B, presented in Section II-B. The signals given by the sensors are now very different from those given by a robot equipped with controller A, and having the same input parameters. Fig. 4 shows a comparison between these output signals without any noise (a stronger mismatch is obtained for the acceleration of the waist). When the controllers are different, the relationship described in equation (12) are hypothetical in the sense that the imitator has no actual knowledge about the dynamics of the demonstrator:

\[
x_k = \bar{f}_k(x_{k-1}, p_k), \tag{23}
\]

\[
y_k = h_k(x_k) + v_k, \tag{24}
\]

By using the state observer (15) - (17), the closed loop dynamics \( \hat{x}_k \) of the imitator and the closed loop dynamics \( f_k \) of the demonstrator do not match perfectly, even in cases where the imitator’s controller is designed to reproduce the demonstrator’s motions as accurately as possible; so a convergence of \( \hat{x}_k \) to \( x_k \) is not plausible.
We reproduce the same imitation scheme on that new motion controller. Fig. 6 shows the observation result of the next step length, in the ideal case of non noisy sensors: the next step prediction is very bad, leading to a wrong imitation. This bad behavior is not surprising, because the sensor network is too sensitive to the state $x_k$, while its values are generated by mismatching closed loop dynamics. We propose then to remove the accelerometer on the waist. The interest of doing so is to reduce the sensitivity of sensors system to $x_k$ while keeping the effect of the most important input parameter: the current step position (sensors on the ankles). The drawback we meet is to loose any possibility to anticipate the next step position. In order to keep the imitation on-line, we assume the next step to be equal to the current one. Fig. 7 shows that the imitator successfully find the current step position of the demonstrator (mean error 2.6mm) without any delay even with a higher noise level (10cm/s$^2$ of standard deviation).

V. DISCUSSION

Our mirror controller is different from traditional imitation techniques; namely those based on motion pattern recognition/classification or any method that uses joint or operational space tracking, sometimes combined with retargeting [4], dynamic filtering [15] or learning techniques. Because these methods are based on reproducing human-motion tracking at the trajectory level control, they can not deal properly with the differences between the mechanical system of the human being and the humanoid robot’s one. Thus, these methods have to solve several problems such as singularities, self-collisions, joint dynamic limits and of course stability in the case of whole-body imitation. For example, in all the previous works targeting the same objective of imitating walking or dancing motions [2], [8], [9], [11], [12], the contact phases between the feet and the ground need to be extracted first from the sensor signals. It is only then that the steps’ lengths and timings can be deduced, so the length of a step can be identified only once it has been completed.

Two of the previous works which achieved experiments on real robots and tried to reach the limits of their approaches acknowledge that the motion parameters they extract from the original motion may not always be safely applicable directly on the robot. But most of all, they acknowledge that an automatic adaptation of these parameters is not always possible, potentially requiring a series of trials and errors before obtaining a safe imitation of the motion [11], [12]; completely automatic and timely imitation appears to be out of reach of these approaches.

In our approach, if parameters of the controller can be early identified and tracked, they are provided to the controller of the robot which then will achieve the task taking its own constraints into account. It can also synchronize in real-time on the human’s task and even predict and anticipate the action to be done. By doing so, imitation, human’s action anticipation and subsequently human-robot interaction, synchronization and more can be made (i) easily, and more importantly, (ii) in a robust way, since we use directly the embedded controllers of the robot, which is considered responsible for allowing versatility and expressiveness while ensuring safety of the resulting motions. This opens paths to synchronous imitation, learning and interaction.

Automatic and timely imitation has already been reached, but by simplifying the whole sequence of steps to binary decisions, go forward or backward, left or right [9]. On the opposite, the two approaches discussed earlier were targeting expressive imitation, retaining as many details as possible from the original motion. An opposition seems to lie between the expressivity and the safety of the imitation process.

In our mirror controller approach, these two questions are supposed to be tackled together at the level of the controller,
which is considered responsible for allowing versatility and expressivity while ensuring safety of the resulting motions. With the controller presented in Section II, the results of Section IV lie somewhere in the middle between the results in [9] and the results in [11], [12]; complex step sequences are correctly imitated, what would not be possible in [9], but only the step positions are imitated, and not more subtle joint motions as in [11], [12]. The controller of the robot naturally shapes its capacities of imitation, whatever the approach considered for the imitation. But our approach puts this controller even more in the core of the problem.

Within the scope of building a computational model for the direct matching hypothesis, the Mimesis framework introduced in [10] shares fundamental aspects with our mirror controller. A Hidden Markov Model (HMM) is used to generate the motions of a humanoid robot, and this same HMM is used symmetrically to identify which motions a human is doing, by an application of dynamic programming (Viterbi algorithm). This allows imitating a motion even from partial observations: partial in time since the motion doesn’t need to be observed entirely to be identified and reproduced, and partial in space since it’s not necessary either to observe the motion of the whole body. The question of the potential delay between the observation and the imitation was not discussed, but this approach is definitely able to identify and imitate a motion before it’s finished since only partial observations are necessary. How much information can be grasped in advance will depend however on subtle details in the design of the HMM (which variables, parameters and transitions). The most significant difference between this method and ours is probably the use of a HMM where we insist on using directly the motion controller of the robot. The problem is that the motion generated by the HMM may not be feasible, especially considering the tight stability constraints that rule walking and dancing motions. Indeed, the imitation of walking motions has been approached with this method [10], but “in the air”, not on the ground, so without having to take into account any dynamic feasibility constraint. One of the big advantages however of using a HMM is that there are straightforward ways to learn one constraint. One of the big advantages however of using a HMM is that there are straightforward ways to learn one constraint. One of the big advantages however of using a HMM is that there are straightforward ways to learn one constraint. One of the big advantages however of using a HMM is that there are straightforward ways to learn one constraint.

VI. CONCLUSION AND FUTURE DEVELOPMENTS

Our work presents a computational model illustrating control based interpretation of the direct matching hypothesis. We transposed the model to humanoid robots and translated the “motor representation” into robot’s motion controller. To simulate an imitation scheme where a robot observes another robot, we went through a simple state observer based implementation of this approach, applied to the case of walking and stepping motions. When the robots are equipped by the same motion controller, our main result has been to observe that anticipation was realized easily and efficiently.

Our result involved two robots; the next step now is to test the robustness of this approach between different systems. It has been already shown that humans use their mirror neuron system to interpret robots action despite the big mismatch between the systems [3]. In order to achieve the inverse process we have equipped a person with a wireless network of sensors and this person is instructed to perform random steps following a known rhythm given by a metronome. We are now processing the human walking signals, and experiments with a real HRP-2 robot should follow soon.

VII. ACKNOWLEDGMENTS

This research is partially supported by the French Agence Nationale de la Recherche, grant reference ANR-08-JCJC-0075-01 and the French Ministry of Economy, Finances and Industry, grant reference DGE-08-2930720. For this research P.-B. Wieber is supported in part by a Marie Curie International Outgoing Fellowship within the 7th European Community Framework Programme; A. Kheddar is supported in part by the EC FP7 VERE IP project.

REFERENCES