

# Gestural Coupling Between Humans and Virtual Characters in an Artistic Context of Imitation

Elisabetta Bevacqua<sup>(✉)</sup>, Céline Jost, Alexis Nédélec, and Pierre De Loor

UEB, Lab-STICC, ENIB, Brest, France  
{bevacqua,jost,nedelec,deloor}@enib.fr

**Abstract.** We present a human-agent interaction based on a theatrical mirroring game. The user and the agent imitate each other's body movements and introduce, from time to time, changes by proposing a new movement. The agent responses are linked to the game scenario but also to the user's behavior, which makes each game session unique.

**Keywords:** Human-agent body interaction · Coupling

## 1 Context

This demonstration has been realized within the collaborative French project called Incredible. The fundamental goal of this project is to reproduce the mutual influence that is intrinsic to human-human interaction for a human-virtual agent interaction. The Incredible project focuses, especially, on the gestural behavior and gestural expressive quality shown by both the virtual agent and the user while interacting. This type of interactions contains some spontaneous and emergent behaviors which are hard to tackle for virtual characters because the evolution of the interaction is shared between the protagonists. A first attempt was proposed by [5] but this work does not show the evolution of the decision relative to the quality of the interaction. Previous works on human-human interaction have studied the dynamical evolution of communication. In “Alive Communication” [3], Fogel and Garvey showed that, while interacting, people coordinate their behaviors and mutually influence their actions and intentions. Ordinary variability, that is slight modifications in the behavior due, for example, to personal style, emerges naturally without braking the bounds of the interaction. From time to time, extraordinary variability (called *innovation*) can appear, obliging interactants to make an effort to integrate the novelty in order to keep the communication going. This theory highlights the dynamical evolution of human-human interactions, the mutual influence and the capability to resist and to adapt to changes. Regular and unexpected behaviors appear naturally and previous works showed that reproducing the equilibrium between them in human-agent interactions is fundamental to improve agent believability and user's engagement [1]. Within the Incredible project we aim to reproduce this evolving equilibrium between regularity and surprise, that we call *coupling* [2]. This demonstration shows our last progress in this direction.

## 2 The Incredible Project Framework

To reach our goal we propose a framework composed of five modules: the **Capture module** retrieves data from tracking devices and generates a unified 15-joints skeleton. Skeletons are sent to the **Analysis module** which tries to recognize the current gesture and its expressivity. The **Decision module** determines the agent's response depending on the user's behavior and the interaction scenario. The **Synthesis module** computes the final skeleton animation that is displayed by the **Rendering module**, implemented in Unity3D.

The Incredible project is still an ongoing project, and this demonstration focuses mainly on the Analysis and Decision modules. The Analysis module collects, in a continuous flow, data from devices enabling skeletal tracking of a single person. The recognition process is based on skeleton analysis and motion features computation. Skeleton joints are used to compute features which are stored in a reference database. A Principal Component Analysis (PCA) is computed to select the most important features, useful in discriminating gestures. During real-time recognition, using distance measures, real-time selected features are compared to the reference database to find the most similar gesture. More details can be found in [4]. The objective of the Decision module, shown in Fig. 1(a), is to generate an appropriate action of the virtual agent according to the user's non verbal behavior. An action can be an animation to play, a joint to move, or an expressivity to change (for example, moving faster). The decision is based on the level of coupling between the human and the virtual agent, which shows mutual influence, engagement, and willing to keep on interacting together. A high level of coupling indicates that, both human and virtual agent, are engaged in the interaction, while a low level of coupling results in a bad feeling about the interaction. Analyzing this coupling, the main objective is to maintain the interaction. According to Fogel [3], three strategies are possible: (i) to maintain the current coupling level (co-regulation) if everything is going well, (ii) to induce a small variation of coupling level (ordinary variability) for example to avoid human's boredom, or (iii) to change dramatically the coupling level (innovation) for example to reengage the human. To take such a decision, the Decision module refers to a scenario which describes its expectations (what human is supposed to do), and its goals (what virtual agent is supposed to do). The satisfaction of expectations implies that the level of coupling increases, otherwise it decreases. The first implementation of this module is scenario-dependent and rule-based. A scenario is written in XML and describes actions which can be done by the virtual agent, a sequence of possible events and associated expectations. Rules, also written in XML, describe relations between the level of coupling and the goal to reach. This is only a first attempt to reproduce the mutual influence of human-human interaction. Our next objective is to make the decision module more generic to decrease the importance of a predefined scripted scenario. Moreover, a learning process will be added to automatically define the rules which conducted to the best experience between the human and the virtual agent.

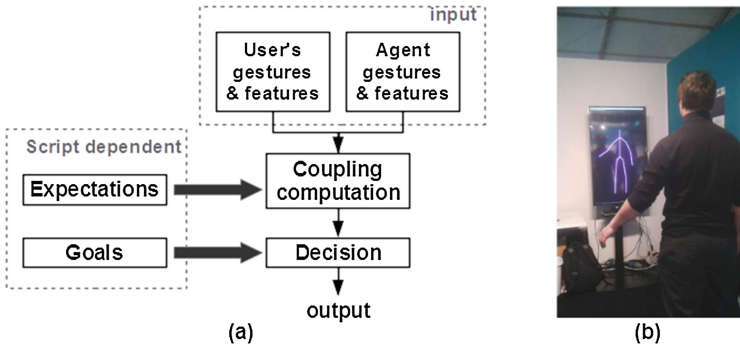


Fig. 1. (a) Decision module architecture. (b) Theatrical mirroring game.

### 3 Theatrical Mirroring Game Scenario

As explained in Sect. 1, the Incredible project focuses on gestural behavior, particularly in the artistic domain. For such a reason, we collaborate with a theater company to define scenarios in which the whole interaction is based solely on body movements (neither speech nor facial expressions). The mirroring game is a good example of the researched interactions: two players imitate each other's movements and introduce, from time to time, changes by proposing a new movement or by modifying the expressivity, that is the manner in which the movement is performed. A very interesting characteristic of this game is that at any moment both players can be the leader and the follower, for example a person can control the movement of left hand while the other is leading that of the right hand. If the game is correctly realized, it is impossible for an external observer to guess who is in control of which part of the body and just the players can feel it. They are strongly aware of how their behavior is both influenced and influential. To play successfully this type of game, coupling between the players is fundamental; for such a reason, the mirroring game is the scenario we chose in this demonstration. It evolves as follows: the virtual agent performs idle movements in its virtual environment, patiently waiting for a user who would like to play with it. When the user appears in front of the agent and the Kinect tracks him, to establish the interaction, the agent and the user must communicate. Since neither verbal behavior nor facial expressions are allowed, both interactants have nothing but gestures to express themselves. So the agent greets the user waving its hand (or greets back if the user waves his hand first). From this moment, to start the game, both virtual and human players can express their intention to play by performing a bow. It does not matter who bows first, what is important is that, by bowing, both show that they agree to play the game together. Figure 1(b) shows the demo set-up. During the game the agent expects that the user imitates its gestures and their expressivity and that from time to time the user proposes something new. So, when the Analysis module informs the Decision module about the user's behavior, this module compares the input data with

the agent expectations and updates the level of coupling. In this scenario, there is just one expectation: “the human and the virtual agent are doing the same thing”. To measure this coupling, human position and features are compared to virtual agent position and features. The coupling is inversely proportional to the resulting distance. For example, if the human and the virtual agent are doing the same thing, distance is minimum, therefore the coupling is maximum. In this context, the virtual agent has two possible goals: to imitate human (goal 1) or to introduce new movement (goal 2). In relation to “Alive Communication” theory, we consider that co-regulation is goal 1, that ordinary variability is goal 2, and that innovation is stopping the game. During the game, virtual agent alternates these goals according to rules until the end of the game. In this first implementation, rules are minimalist and describe the relation between coupling and goals:

```

if state == goal1 then
  if coupling == too low then stop the game
  else if coupling == high then state ← goal2
  else coregulation
  end if
else if state == goal2 then
  if coupling == too low then state ← goal1
  else coregulation
  end if
end if

```

At any moment, both the agent and the user can stop the game showing a specific gesture. The agent could stop the game for several reasons, for example if the game has lasted long enough or if the user is not a “good” player and the coupling level remains low (the user rarely follows new movement propositions, that is he rarely adapts his behavior to that shown by the agent). When the interaction is over, the virtual character greets the human player and goes back to its idle motion waiting again for a new player.

**Acknowledgments.** This work was funded by the ANR INGREDIBLE project: ANR-12-CORD-001 (<http://www.ingredible.fr>).

## References

1. Bevacqua, E., Stanković, I., Maatallaoui, A., Nédélec, A., De Loor, P.: Effects of coupling in human-virtual agent body interaction. In: Bickmore, T., Marsella, S., Sidner, C. (eds.) IVA 2014. LNCS, vol. 8637, pp. 54–63. Springer, Heidelberg (2014)
2. De Loor, P., Bevacqua, E., Stanković, I., Maatallaoui, A., Nédélec, A., Buche, C.: Utilisation de la notion de couplage pour la modélisation d’agents virtuels interactifs socialement présents. In: Conférence III. Oxford University Press, France (2014)
3. Fogel, A., Garvey, A.: Alive communication. *Infant Behav. Dev.* **30**(2), 251–257 (2007)
4. Jost, C., Stanković, I., De Loor, P., Nédélec, A., Bevacqua, E.: Real-time gesture recognition based on motion quality analysis. In: INTETAIN (2015)
5. Pugliese, R., Lehtonen, K.: A framework for motion based bodily enaction with virtual characters. In: Vilhjálmsón, H.H., Kopp, S., Marsella, S., Thórisson, K.R. (eds.) IVA 2011. LNCS, vol. 6895, pp. 162–168. Springer, Heidelberg (2011)