The INGREDIBLE Database:
A First Step toward Dynamic Coupling in Human-Virtual Agent Body Interaction

Igor Stanković¹, Pierre De Loor¹, Virginie Demulier²,
Alexis Nédélec¹, and Elisabetta Bevacqua¹

¹ UEB, Lab-STICC, ENIB
² LIMSI-CNRS, France
{stankovic,deloor,nedelec,bevacqua}@enib.fr
demulier@limsi.fr

1 Introduction

This paper describes a human-avatar interaction database, built from the recorded interactions of two actresses wearing motion capture (MoCap) suits. They were located in separate rooms, and communicated only via each other’s avatar. This form of interaction is noteworthy for two reasons: firstly, due to the use of avatars, no facial expressions nor gaze were visible, ensuring that the only interaction cues were body movements; secondly, from a psychological point of view, this database can tell us how people communicate with virtual representations of other individuals.

To our knowledge, none of the existing multi-modal expressive behaviour databases meet our requirements. Most databases, such as [1], do not focus on interaction or have been collected for different reasons. Other databases, for example [2], contain recordings of human-human interactions and focus mainly on verbal communication limited to the upper part of the body (i.e. face, head, and hands).

The database presented here is part of the INGREDIBLE project, which aims to develop a virtual agent capable of maintaining a gestural coupling [3] with a human. The assumption is that the increased persistence of coupling – with some evolving rules yet to be defined – will heighten the realism of a virtual character’s presence [4] and benefit their use in applications such as personal assistants and intelligent virtual tutors, and provide more realistic interactive behaviours in video games.

2 Technical Description and Collected Data

MoCap recordings for the database were carried out with the aid of two professional actresses from the theatre company Derezo¹. The actresses were located in different rooms, and only able to see each other’s avatars. In order to make the

¹ http://www.derezo.com
INGREDIBLE database more widely usable, the database utilizes two different MoCap suits and systems to collect data: Art-Human and Moven.

We have two main reasons for requiring a MoCap database: firstly, the database will be used to develop feature analysis tools able to recognize users’ gestures; secondly, MoCap recordings are necessary for the animation of a virtual agent. We also recorded synchronised videos of the two actresses while interacting in order to annotate their movements and find cues of dynamic coupling.

There are several types of recordings in the database. Recordings were either non-interactive or interactive. Non-interactive means that the actresses did not interact, and so no avatar was displayed in front of them. Instead, they were asked to perform a series of predefined gestures, repeating each one with variations in three dimensions. The modification criteria were amplitude (narrow, medium, wide), speed (slow, medium, fast), and fluidity (staccato, medium, fluid). In the interactive approach, the actresses communicated with each other through their human-size avatars, which were displayed on a screen in front of them. They were introduced to this environment by being encouraged to interact freely for as long as they wished. These first recordings often provided us with very interesting and spontaneous data, but the actresses rapidly grew bored without a prescribed task to perform. To add artistic, gestural, and expressive details to the interactions (according to the requirements of the project), we defined two interacting situations: 1.) imitation and 2.) bodily emotional dialogue.

The resulting dataset consists of 114 different recordings, 57 captured by each suit, with more than 150 hours of recording and 27 GB of data. The database stores recorded Art-Human and Moven data converted to the .bvh format. It also holds Art-Human raw data in .txt files and Moven .mvn data.

3  Future Work

The database contains some limitations (e.g. only two participants, both participants are female, both participants are actresses), so as a part of our future work, more recordings would be necessary. At present, the videos are being annotated by a team of psychologists aiming to extract cues of dynamic coupling.

Acknowledgement. This work has been funded by the ANR INGREDIBLE project: ANR-12-CORD-001 (http://www.ingredible.fr)

References