A Framework for Motion Based Bodily Enaction with Virtual Characters

Roberto Pugliese and Klaus Lehtonen

Department of Media Technology,
School of Science,
Aalto University,
Espoo, Finland
{roberto.pugliese,klaus.lehtonen}@tkk.fi

Abstract. We propose a novel methodology for authoring interactive behaviors of virtual characters. Our approach is based on enaction, which means a continuous two-directional loop of bodily interaction. We have implemented the case of two characters, one human and one virtual, who are separated by a glass wall and can interact only through bodily motions. Animations for the virtual character are based on captured motion segments and descriptors for the style of motions that are automatically calculated from the motion data. We also present a rule authoring system that is used for generating behaviors for the virtual character. Preliminary results of an enaction experiment with an interview show that the participants could experience the different interaction rules as different behaviors or attitudes of the virtual character.

Keywords: enaction, motion capture, bodily interaction, authoring behaviors.

1 Introduction

Authoring believable behaviors for virtual characters is a crucial step towards the creation of immersive gaming experiences. In social encounters behaviors emerge as humans react to actions of others in a continuous feedback loop. This process is sustained by bodily interaction among the different parties. Human-computer bodily interaction is possible even at consumer level with latest sensor technology.

We are interested in behaviors that can be observed in and activated by bodily motion and how to use this as a medium of interaction with a virtual character. Our interests are not in traditional goal-oriented interaction or in symbolic language. For these reasons an enactive loop, where both parties can continuously affect the other through actions and the style of motions, was chosen as the model of interaction instead of using discrete gestures.

We present a framework that allows bodily interaction between a human and a virtual character in an enactive loop. The implementation takes a long motion capture sequence as input and automatically segments it into a motion library, indexed by motion styles, that is used to animate the virtual character. We also present a rule authoring system that is used for generating behaviors for the virtual character.

2 Related Works

In this section we explore earlier works related to enaction and to techniques that enable interaction through motion with animated characters.

2.1 Enaction

Enactive Media is an approach to design modalities of human-machine interaction. While traditionally interactivity has been approached with theories and tools for goal-oriented tasks, the enactive paradigm focuses on a tight coupling between machine and the user, here a participant or enactor. The process is a feedback loop: the actions performed by the enactor affect the medium that in turn affects the following actions of the enactor. The coupling is sustained by means of bodily and spatial involvement, or enactment [1]. An enactive system may involve even a community of agents in participatory sense-making [2].

We want to create a process where the participant will be able to notice different behaviors in the virtual character as a response to his or her own behavior. The rules governing the interaction do not need to be explicit but they can be learned by interacting, in accordance with the original definition of enaction from Bruner [3], that is to learn by doing. While an enactive account for human-computer interaction has been provided in other fields such as facial expressions of virtual characters and movie creation [4], based on psychophysiological input, an implementation of enaction with a virtual character based on bodily motion is yet absent.

In our enactive setting no assumption about the meaning of gesture is done a priori but meaning is actively constructed by the participant and emerges from the enactive loop. This calls for representing the quality of the interaction and the motion style in an objective and non-hierarchical way. We borrow the spatial ontology (ontospace) approach by Kaipainen et al. [5] as a solution. An ontospace is defined by ontological dimensions (ontodimensions) that correspond to descriptive properties of the content repertoire, which in our case are motion clips.

2.2 Interaction through Motion with Animated Characters

Animating characters is possible with motion graphs that contain captured motion segments and a list of allowed transitions between the segments [6]. A motion graph can be constructed automatically from a large corpus of motions and can be used to produce arbitrarily long continuous motions [7].

In a previous work, full-body interaction with a virtual creature meant mainly giving commands and instructions to virtual creatures and the set up did not allow symmetrical interaction [8]. Similarly, Improv [9] allows interaction with virtual characters. It allows creating scripted sequences of animation and interaction by means of if-statements based on the properties of the characters. These earlier systems concentrated mainly on goal-oriented actions.

In human-computer bodily interaction one needs to extract motion cues able to describe the human motion in a machine friendly way. We follow a methodology based on previous work in the field of analysis of expressive gesture in music and dance performances [10]. The process has camera-based tracking and calculation of motion features that serve as descriptors for motions. Those descriptors include amount of movement and body contraction/expansion.

3 Framework

The system that we have built simulates a situation where two persons are separated by a glass wall and are able to interact only through bodily motions. This creates an enactive loop and allows replacing one or two of the persons with a virtual character (Fig. 1), in our system rendered as a stick figure. The enactive system reacts to human motion (input) by triggering a recorded motion clip (output).

3.1 Enactive Loop

Any motion clip (either recorded or realtime captured) can be associated to a point in an ontospace based on its values of the motion descriptors as coordinates. Before the enaction (Fig. 2) can start the output ontospace is filled with acted motions (Z). The loop starts by mapping the human motion into the input ontospace with the descriptors (A). Then a rule system determines the desired position of the virtual character in the output ontospace (B). Next the animation engine searches for the closest motion to the desired position from the acted motions (C). The virtual character then proceeds to play the motion (D). As the last step the human observes the motion of the virtual character (E), which affects the motion of the human, etc.

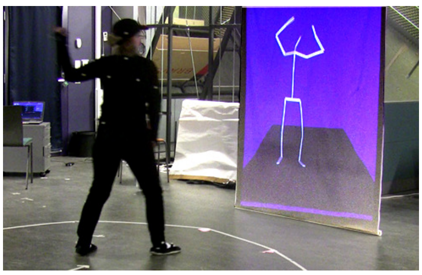

Fig. 1. Live enaction.

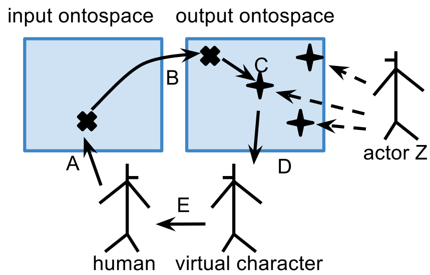

Fig. 2. Enactive loop sustained by the human and the virtual character.

3.2 Motion Analysis and Generation Using Descriptors

For the enactive loop we needed a virtual character that moves with varying motion styles and reacts to the motion style of the human. We use Quantity of Motion (QoM) and Distance descriptors for motions. The former is used as an estimation of the energy and the motion style while the latter characterizes the interaction between the social spaces of the human and the virtual character.

Our definitions are the following: Quantity of Motion (QoM), sum of the frame to frame displacements of all the joints in the body of the character divided by the number of frames in the motion, minus the minimum amount of motion required to move from the starting position to the end position; Distance, the distance of the center of the body from the wall separating the characters.

These descriptors allow placing every possible motion in a two-dimensional ontospace (Fig. 3), constituting a simple case that still allows authoring behaviors. To normalize the descriptors values Distance was scaled linearly, but for QoM we used a log-like function. This takes into account that humans perceive very small changes in the amount of motion if the overall speed is low, but for high speeds the change needs to be much larger to be noticed [11].

Our virtual character is a program that takes desired descriptor values as input and then generates an animated motion sequence that fits to the desired values. To be able to do this we created a motion library containing idle standing (consentration of dots in Fig. 3), walking and running (extremes of Distance in Fig. 3) and jumping actions (high QoM in Fig. 3). These actions were acted with varying styles to evenly populate the ontospace with motion segments. Total of six minutes of motion was automatically segmented to create a motion graph with approximately one second long clips that allow smooth transitions to many other clips. The segmentation was based on finding frames of motion that have a similar pose and speed. After playing a clip the number of alternative following clips ranged with our motion library from 2 to 240.

3.3 Authoring Rules

In our methodology, authoring the rules corresponds to finding a meaningful transformation of the input ontospace, the one of the human, into the output ontospace, the one of the virtual character. The transformation is a mapping defined by example point-pairs in the input and output ontospace. To make the mapping work for inputs in between the example points, we search for the k-nearest neighbors in the input space and determine the output with a weighted interpolation of the corresponding points in the output space. For this, we made a GUI for creating mappings between the ontospaces by specifying examples of corresponding point pairs (A, B, C and D in Fig. 4). In the case of a two-dimensional ontospace this means clicking a point in the input space and then clicking the desired output in the output space. An obvious mapping is the identity transformation which makes the virtual character imitate the motion style of the human. Once a rule is defined, it can be used with a larger motion library without any extra manual work.

In the case of a large number of descriptors, authoring the rules with a mouse can become a tedious and difficult task. A promising alternative approach could be to first record actions with motion capture and then to record the responses to those actions. With this motion data, it should be possible to populate the input and output spaces and obtain behaviour rule.

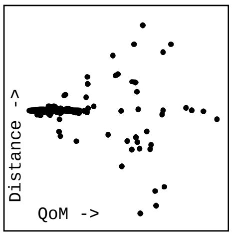

Fig. 3. Ontospace populated with motion segments (dots). The coordinates of the dots are the QoM and the Distance.

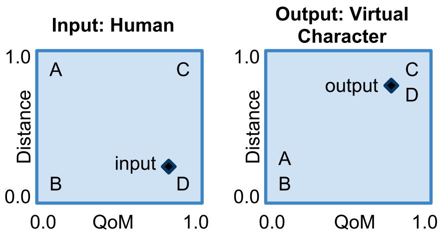

Fig. 4. An example of authoring the behaviour rules with point pairs A, B, C and D.

4 Enaction Tests and Interview with the Participants

In order to validate our methodology and evaluate the effectiveness of our implementation, we conducted enaction tests where a participant had to bodily interact with a virtual character projected in front of them. The participants were 7 unpaid volunteers, 5 male and 2 females of age from 25 to 55.

At the beginning of the experiments the participant was inside a motion capture room and informed about the area where he or she could move and the fact that the virtual character was a stick figure which is able to see the human as a stick figure. No explicit goal of the experiment was stated besides the suggestion to freely explore the bodily interaction with the virtual character.

Six conditions were presented. The first one was always a plain imitation rule used to familiarize the participant with the setting and the interaction paradigm. The next 6 conditions were other rules in randomized order. Those rules were created from a mathematical point of view to try different mappings of descriptors and restricting the virtual character to a limited area of the ontospace.

Condition A was a plain imitation rule. Conditions B and C were imitations with QoM of the virtual character limited to low values in condition B and to high values in condition C. Condition D mapped the QoM of the human to the Distance of the virtual character, as in Fig. 4. This causes the virtual character to back off when the human does motions with high QoM. Condition E inverted the QoM of the human for the virtual character. This makes the virtual character have high QoM, for example, by jumping and waving hands when the human is standing still. In condition F the virtual character played random motion clips without being affected by the human.

In each condition the participant was free to experiment with that rule for 2 minutes. After the experiment we interviewed all the participants to get detailed information about their experiences.

5 Interview and Discussion

Evaluating bodily motions during enaction in an objective manner is a more difficult task than evaluating pre-recorded videos of motions. The main reason is that the conditions are not fully controllable and repeatable because by definition the outcome strongly depends on what the participant does.

On questions related to quality of the interaction with the virtual character we found out that all the participants felt there was interaction in some conditions and also that they could identify different behaviors. There was a general agreement about a character that showed a recognizable scared behavior. This character belonged to the condition D which causes the virtual character to back off when the human does motions with high QoM. Another often mentioned behavior was aggressiveness. This was probably caused by the motions with high QoM such as jumping and waving hands.

The participants said that in some conditions it was hard to understand what made the character react. Besides the condition F with a randomly acting character, this could be explained by that the reaction time of the character could become too slow if the character was playing a long motion. We are aware that two descriptors are not enough to properly describe human actions. We realized that a too simple system makes the participant focus mainly on discovering the rules and the descriptors rather than being in the flow of enaction.

The participants said that their own behavior was affected by the behavior of the virtual character and many of the participants said that they started to mimic the gestures seen in the character. Most of the time, participants moved more when the virtual character was active and less when the character was passive. These facts indicate that the interaction we designed is effectively a case of enactive loop, where both parties affect each other.

6 Conclusions and Future Work

We have presented a framework to design bodily interaction with virtual characters based on the concept of enaction and an authoring tool to specify different behaviors for them that gradually emerge during and due to the interaction. Behaviors are created by mapping the input ontospace of the human, described by motion descriptors, into the output ontospace of the virtual character, populated with automatically evaluated motions. Preliminary tests with participants showed that experiencing different interaction rules as different behaviors or attitudes of the virtual character is possible even in the simplest case of a two-dimensional motion descriptor space.

Defining motion styles with motion descriptors allows using large amount of captured motion without adding more work as no manual annotation is required.

In the future, we plan to add new motion descriptors and differentiate different parts of the body. The manual process of authoring behaviors could be replaced by acting them out in the case of a large number of motion descriptors. Furthermore, we intend to use interpolation among different rules to create virtual characters changing their behaviors during the enaction.

Acknowledgments. This work is part of the Enactive Media project, funded by the Academy of Finland, decision number 128132. We want to thank the Enactive Media team Mauri Kaipainen, Niklas Ravaja, Tapio Takala, Pia Tikka and Rasmus Vuori for the fruitful discussions and intellectual contribution to this paper.

References

1. Varela, F., Thompson, E. and Rosch, E. Embodied Mind: Cognitive Science 2. and Human Experience. MIT Press, Cambridge, MA, (1991)
2. De Jaegher, H., & Di Paolo, E.A. Participatory sense-making: An enactive approach to social cognition. Phenomenology and the Cognitive Sciences, 6(4), 485-507, (2007).
3. Bruner, J. Toward a theory of instruction. Belknap Press of Harvard University Press, Cambridge, MA, (1966)
4. Kaipainen, M.; Ravaja, N.; Tikka, P.; Vuori, R.; Pugliese, R-; Rapino, M.; Takala, T.: Enactive Systems and Enactive Media. Embodied human machine coupling beyond interfaces . Leonardo, (2011) (in press)
5. Kaipainen, M., Normak, P., Niglas, K., Kippar, J. and Laanpere, M. Soft ontologies, spatial representations and multi-perspective explorability, Expert Systems 25, 5, 474–483 (2008)
6. Kovar, L., Gleicher, M., Pighin, F.: Motion graphs. In Proc. of SIGGRAPH '02. ACM, New York, NY, USA, 473-482. (2002)
7. Zhao, L., Normoyle, A., Khanna, S., Safonova, A.: Automatic construction of a minimum size motion graph. In Proc. of the 2009 ACM SIGGRAPH/Eurographics Symposium on Computer Animation (SCA '09), Dieter Fellner and Stephen Spencer (Eds.). ACM, New York, NY, USA, 27-35. (2009)
8. Blumberg, B., Galyean, T.: Multi-level direction of autonomous creatures for real-time virtual environments. In Proc. of SIGGRAPH '95, Susan G. Mair and Robert Cook (Eds.). ACM, New York, NY, USA, 47-54. (1995)
9. Perlin, K., Goldberg, A.: Improv: a system for scripting interactive actors in virtual worlds. In Proc. of SIGGRAPH '96. ACM, New York, NY, USA, pp. 205-216. (1996)
10. Camurri, A., Mazzarino, B., Ricchetti M., Timmers, R., Volpe, G.: Multimodal Analysis of Expressive Gesture in Music and Dance Performances, Gesture-Based Communication in Human-Computer Interaction, Lecture Notes in Computer Science, vol. 2915/2004, pp. 357-358, (2004)
11. Levine, S., Krhenbhl, P., Thrun, S. and Koltun, V.: Gesture controllers. In Proc. of SIGGRAPH 2010, Hugues Hoppe (Ed.). ACM, New York, NY, USA, Article 124, 11 pages. (2010)