

# Establishing Rapport with a Virtual Dancer

Anton Nijholt, Dennis Reidsma, Herwin van Welbergen  
Human Media Interaction Research  
University of Twente  
PO Box 217, 7500 AE Enschede, the Netherlands  
[anijholt@cs.utwente.nl](mailto:anijholt@cs.utwente.nl)

*Abstract. We discuss an embodied agent that acts as a dancer and invites human partners to dance with her. The dancer has a repertoire of gestures and moves obtained from inverse kinematics and motion capturing that can be combined in order to dance both on the beat of the music that is provided to the dancer and sensor input (visual and dance pad) from a human partner made available to the virtual dancer. The interaction between virtual dancer and human dancer allows alternating 'lead' and 'follow' behavior, both from the point of view of the virtual and the human dancer.*

## 1. Introduction

Human-computer interaction research has moved from traditional mouse, keyboard and graphical desktop user interface to (invisible) ambient intelligence user interfaces. Ambient intelligence has been defined as ubiquitous computing + social and intelligent interfaces. In the ambient intelligence point of view of the future we have sensors in our personal and public environments (at home, in our car, during shopping, et cetera) and these sensors allow us to enter conversation, entertainment, education, health and training environments in which we can interact with virtual humans that play a role of guide, coach, teacher, buddy, personal assistant, or adviser. We are interested in interfaces where we have whole-body interaction. Embodied agents can play useful roles in those interfaces, for example, to encourage physical movements when playing the role of a fitness trainer, to demonstrate and advise on aerobic exercises, or to help in revalidation. Art, games and entertainment are other application areas. Playing with (or against) a robot, performing with a virtual orchestra, take part in a play with real and virtual actors, act in an immersive virtual environment are other examples where the smart environment rather than the desktop invites, allows and supports natural human activities and associated body movements. These smart environments can be part of our home environment where they can be made to ensure privacy, but they can as well be part of (distributed) virtual collaborative environments and of public entertainment environments.

## 2. Multimodal Input for the The Virtual Dancer

We introduced the Virtual Dancer [1], an interactive dancing agent that dances together with the user to the beat of the music (Figure 1). It will adapt its performance to whatever the human user is doing, who is observed using real time computer vision.

A prerequisite for a virtual dancer is the ability to interpret the music to find the beats to which the dance should be aligned. For this we implemented the beat detection algorithm described in the publications of Anssi Klapuri. This algorithm detects the tempo and beat in the music played for the dancer. Once the beat is known, the virtual dancer should be able to dance along. For this we constructed a database of many different dance moves, collected using motion capturing or created manually, using parameterized animation. The virtual dancer selects



Figure 1: Interacting with the Virtual Dancer

the most appropriate dance moves from the database, given the observations of the movement characteristics of the human dancer. These moves are then timed to the beat, by locally warping its timing so that the beat positions in the move match with those in the music. For example, in a complex clapping animation, the clap-points are aligned exactly to the predicted beat times, so the dancer will clap to the beat of the music. The transition from one move to the other is made using an IK (Inverse Kinematics)-generated stepping motion and interpolation techniques.

Other input for the Virtual Dancer comes from computer vision. The system observes the movements of the human dance partner using the computer vision system ParleVision. Using several robust processors, the system extracts global characteristics about the movements of the human dancer, such how much (s)he moves around or how much (s)he waves with the arms. Such characteristics can then be used to select moves from the database that are in some

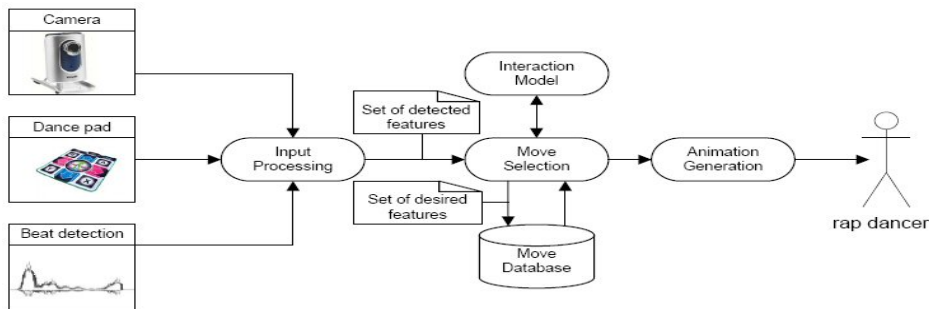


Figure 2: Architecture of the Virtual Dancer

way “appropriate” to the dancing style of the human dancer. Finally, there is a dance pad that registers feet activity.

### 3. Interactive Dancing

The architecture of the system is presented in Figure 2. There is a mapping from the characteristics of the observed dance moves to desirable dance moves of the Virtual Dancer. As described above, to realize this mapping, the database with moves is searched for moves that satisfy as much as possible the desired features. The interaction model (see Figure 2) reflects the intelligence of the Virtual Dancer. By alternating patterns of following the user with taking the lead with new moves, the system attempts to achieve a mutual dancing interaction where both human and virtual dancer influence each other. Finding the appropriate nonverbal interaction patterns that allow us to have a system that establishes rapport with its visitors is one of the longer term issues that is being addressed in this research.

### 4. Dance Events and Dancers

We have presented our virtual dancer installation in April 2006 at a large Dutch ICT event in Amsterdam and at the ACM CHI conference in Montréal. The installation turned out to be very robust and, clearly, visitors had a lot of fun (Figure 3), although they often wondered what had to be done (fulfill a task, play a game, enter a competition), while all we had in mind was to provide users with task less but enjoyable interaction. An other observation we made was that because of the dancer’s movements and feedback people attributed more intelligence to the virtual dancer than it had. This is of course in line with observations from Joseph Weizenbaum long time ago on his Eliza chatbot and it is also in line with observations made on Jeremiah [1], a virtual face displayed on a wall in a room that shows emotions in response to the actions of people visiting the room. That is, “people want to believe that he is more than he is and as such read far more into his direct response to their activities”. We recorded the interactions with the visitors to our installation and these recordings are used to identify and analyze the interaction patterns and will lead us to future research.



Figure 3: Fun interactions with the Virtual Dancer

### References

- [1] D. Reidsma, A. Nijholt, R.W. Poppe, R.J. Rienks, & G.H.W. Hondorp. Virtual Rap Dancer: Invitation to Dance. *Proceedings CHI2006, Conference on Human Factors in Computing Systems*, ACM, Montréal, Québec, Canada, April 2006, 263-266.
- [2] R. Bowden, P. Kaewtrakulpong, & M. Lewin. Jeremiah: the face of computer vision. *Proceedings SMARTGRAPH '02, 2nd international symposium on Smart Graphics*, ACM Press, New York, NY, USA, 2002, 124–128.