Multi-modal Behavioral Cues from Bodily Interaction in Ambient Entertainment Applications

Anton Nijholt, Betsy van Dijk, and Dirk Heylen

University of Twente, Human Media Interaction (HMI), PO Box 217, 7500 AE Enschede, the Netherlands, a.nijholt@ewi.utwente.nl

Abstract

Exertion interfaces require bodily activity. Users have to perform exercises, they have to dance, they have to golf or football, and they have to train particular bodily skills. Unlike game interfaces where we can observe lots of research activity recently trying to define, interpret and evaluate issues such as 'flow' and 'engagement', in exertion interfaces these concepts need to be reconsidered and new ways of evaluation have to be defined. Here we embed exertion interface research in ambient intelligence and entertainment computing research. Examples are discussed and views on evaluation are expressed.

Introduction

In previous years exertion interfaces have been introduced [1]. In game or entertainment environments the 'user' may take part in events that require bodily interaction with sensor-equipped environments. E.g., in an urban game, mobile devices may be used to inform the users about activities they have to perform or about activities of their partners or opponents in the game. The game can require the gamer to walk, run, or perform other activities, in order to compete or cooperate with others involved in the game. Exertion interfaces have also been introduced in home and office environments that offer, elicit and stimulate bodily activity for recreational and health purposes.

For example, in a smart, sensor-equipped, home environment bodily activity can be employed to control devices, or the smart home environment might anticipate our activities and behave in a pro-active and anticipatory supporting way. Although in home environments there is freedom when and how to perform tasks, there nevertheless are regular patterns of bodily activity and therefore activities can be predicted and anomalies can be detected. In task-oriented environments, e.g. an office environment, people probably have more well-defined tasks where efficiency plays an important role. Smart office furniture can provide context and task aware support to a moving office worker.

Exertion and Entertainment Interfaces

In order to design and implement successful exertion interfaces we need an environment that can detect, measure, and interpret physical activity. One of the best known exertion interfaces is 'sports over a distance', where players from different sites have to hit a wall with a ball [1]. The position on the wall and the force with which the ball hits the wall are important for winning or losing the game. In this particular exertion interface there is no direct sensing of body movements or physiological information. Only the result of the exertion is measured and mediated. There also exist exertion interfaces with direct sensing of bodily activity (body movements, gestures, bodily and facial expressions, dynamic aspects of expression, etc.) and of speech activity that accompanies bodily activity (effort and pain utterances, laughs, prosodic aspects of speech utterances ...). Cameras and microphones allow visual and audio processing of a user's activity and any other sensors are available. There are sensors that provide information about location changes (tracking bodies and faces of individuals), frequency and expressiveness of movements, effort measuring, etc. One step further is to take into account physiological information obtained from the user. This information can be used both to guide the interaction and to measure the user experience. In particular brain-computer interfacing (BCI) is a source of information in our research to learn about the way the user experiences the interaction besides offering the user control.

Multimodal, Joint, and Coordinated Activity in Embodied Interaction

Exertion interfaces emphasize the conscious use of bodily activity (jogging, dancing, playing music, sports, physical exercises, fitness, etc.) in coordination and sometimes in competition with other human users (friends, community or team members, accidental passers-by, opponents, etc.). Real-time coordinated interaction between human partners or between humans and virtual or robotic partners makes exertion interfaces exciting. Coordination may be required by the rules of the game or the exercise or the tasks ask for it, but most of all people engage in coordinated interaction because it brings satisfaction and enjoyment. For users of exertion interfaces the interaction supporting feedback and the interaction experience are important [2].

We take inspiration from Clark's work on joint activity in many of our studies. [3]: "A joint action is one that is carried out by an ensemble of people acting in coordination with each other. As some simple examples, think of two people waltzing, paddling a canoe, playing a piano duet, or making love."

Communication, like dancing, includes coordinated nonverbal activity. We have studied face to face conversations, multi-party interaction, interactions between a virtual and a human dancer [4], a virtual conductor and a human orchestra [5], and a physiotherapist and her student [6]. Underlying joint activities are rules and scripts. To learn these and to put them into practice requires social intelligence, guided by empathy, moods and emotions. Despite many research results from social and behavioral sciences, computational models of joint activities are hardly available. This makes it difficult to design interfaces that aim at providing a similar interactional experience between real humans and virtual humans or robots, as is provided in a real-life human-human exertion activity, as in dancing, paddling, playing quatremains, and making love. Endowing the computer with a human-like appearance strengthens the expectation the computer will take part in joint activities in human-like ways. Hence, there is a need for computational modeling of human joint activities. We replace one of the human partners in a joint exertion activity by a computer (i.e., a robot or a virtual human). Hence, we need to model the exertion interaction in order to have the computer behave in a natural and engaging way.

In addition to rules that underlie joint activity there can be a need to align the interaction to external events over which the interaction partners do not necessarily have control. E.g., if we have a human and a virtual dancer then their moves have to be aligned with the music. Similarly, a virtual conductor and his human orchestra follow the score; a virtual aerobics trainer

interaction partners do not necessarily have control. E.g., if we have a human and a virtual dancer then their moves have to be aligned with the music. Similarly, a virtual conductor and his human orchestra follow the score; a virtual aerobics trainer and human student have to align their movements to the supporting music.

In our research we look at

Measuring activity to improve the interaction (using off-line information: history, personality, …) and adapt the system to the user's history and current activities

Measuring activity in order to know about the involvement of the user (flow, pleasure, pain, effort, tiredness, ...) and adapt the system in order to improve engagement

Earlier [7] we argued that for applications that have the interaction itself as their goal, the interaction and the user experience need to be evaluated, rather than efficiency. In our present research we investigate ways to measure engagement by looking at the degree of coordination between the activities of a human and a virtual partner in exertion and other entertainment interfaces [8]. In this research, supported by [9,10] we investigate how to make entertainment interactions more enjoyable by looking at interaction synchrony, where, on the one hand we aim at disturbing this synchrony in order to introduce a new challenge, and on the other hand we aim at convergence towards coordinated anticipatory multi-modal interaction between human and artificial partner and environment.

References

1. Mueller, F., et al. (2003). Exertion Interfaces: Sports over a Distance for Social Bonding and Fun. In: *Proc. CHI 2003*. ACM Press, USA, 561-568.
2. Bianchi-Berthouze, N., et al. (2007). Does body movement engage you more in digital game play? And Why? In: *Affective Computing and Intelligent Interaction.* LNCS 4738, Springer, 102-113.
3. Clark, H (1996). Using Language. Cambridge University Press.
4. Reidsma, D., et al. (2006). Towards Bi-directional Dancing Interaction. *5th Intern. Conf. on Entertainment Computing (ICEC 2006)*, R. Harper et al. LNCS 4161, Springer, Berlin, 1–12.
5. Maat, M. ter, et al. (2008). Beyond the Beat: Modeling Intentions in a Virtual Conductor. *2nd Intern. Conf. on INtelligent TEchnologies for interactive enterTAINment (INTETAIN).*
6. Ruttkay, Z.M. & Welbergen, H. van (2006). On the timing of gestures of a Virtual Physiotherapist, *3rd Central European MM & VR Conf.* C.S. Lanyi (Eds), Pannonian Univ. Press, Hungary, 219-224.
7. Poppe, R.W. et al. (2007). Evaluating the Future of HCI: Challenges for the Evaluation of Emerging Applications. In: *AI for Human Computing,* T. Huang et al. (eds.), Springer, 234-250.
8. Nijholt, A. et al. (2008). Mutually Coordinated Anticipatory Multimodal Interaction. In: Nonverbal Features of Human-Human and Human-Machine Interaction. LNCS, Springer, to appear.
9. F. Tanaka et al. (2004). Dance Interaction with QRIO. International Workshop on Robots and Human Interactive Communication, IEEE.
10. Tomida, T. et al. (2007) miXer: The Communication Entertainment Content by using "Entrainment Phenomenon" and "Bio-feedback". *ACE 2007*, ACM, 286-287.