# Human Computing, Virtual Humans and Artificial Imperfection

Z.M. Ruttkay, D. Reidsma, A. Nijholt
University of Twente, PO Box 217
7500 AE Enschede, the Netherlands
{zsofi,dennisr,anijholt}@cs.utwente.nl

## ABSTRACT

In this paper we raise the issue whether imperfections, characteristic of human-human communication, should be taken into account when developing virtual humans. We argue that endowing virtual humans with the imperfections of humans can help making them more 'comfortable' to interact with. That is, the natural communication of a virtual human should not be restricted to multimodal utterances that are always perfect, both in the sense of form and of content. We illustrate our views with examples from two own applications that we have worked on: the Virtual Dancer, and the Virtual Trainer. In both applications imperfectness helps in keeping the interaction engaging and entertaining.

## Categories and Subject Descriptors

H.1.2 [**Models and Principles**]: User/Machine Systems - *Human information processing*; H.5.1 [**Information Interfaces and Presentation**]: Multimedia Information Systems - *Animations, Artificial, augmented, and virtual reality* H.5.2 [**Information Interfaces and Presentation**]: User Interfaces – *Evaluation / methodology, User-centered design, User interface management systems*

## General Terms

Design, Human Factors, Theory

## Keywords

Virtual Humans, Embodied Conversational Agents, Human Computing, Imperfections

## 1. INTRODUCTION

The concept of *Human Computing* unifies several objectives: to make the usage of computers easy and natural, allowing such very 'human' behaviors like being emotional or bored, endow computer systems with adaptive and empathic response, facilitating applications where joy and engagement become more important than the problem-solving oriented 'categorical computing' practice.



Virtual Humans (VHs) can play an interesting role in approaches to the Human Computing objectives. A VH is an animated agent (also referred to as embodied conversational agent) [3] residing on the screen of a PC (or projected from there to a wall or placed in a Virtual Environment), resembling in appearance and communication capabilities a living, human creature. Depending on the application, humans treat computers as social actors [18]. Human-like appearance and behavior of the interface, in particular when the computer is hidden between a virtual human, strengthen this acceptance of the computer as a social actor [6,13] or even a persuasive social actor [5] that changes our attitudes and behavior. Virtual Humans, more than regular graphical user interfaces, invite natural human interaction behavior and therefore it is useful to be able to capture and understand this behavior in order to let it play its natural role in an interaction [14].

Assuming that we are able to invite, capture and understand natural behavior the question arises whether these VHs are able to maintain this naturalness in their communication with their human partners. While the resources of increasing computer power have been used to improve the technical aspects of virtual humans (vision, synthetic speech, smooth animations, high-polygon models), some essential human characteristics have hardly been addressed. These characteristics deal with subtleties and imperfections in human-human communications that generally are not considered to be interesting from the viewpoint of efficient and correct human-computer interaction. We will look at these characteristics from the viewpoint of human-virtual human interaction.

### 1.1 'Machine Humans'

Traditionally, there are different qualities associated with machines (including computers) and humans. Machines, and machine-made products are praised for being reliably identical and precise, irrespective of time and conditions of production, exhaustive, fast, deterministic in handling huge amounts of information or products, slavishly programmable… as opposed to humans being non-repetitive and less predictable, error-prone and non-deterministic in their 'functioning'. On the other hand, when it comes to flexibility, adaptation, error recovery, engagement… humans are valued higher. In the old and simplified view, machines are to perform dull, repetitive tasks, leaving more space for the human for the creative design, experimentation.

The above dichotomy evokes re-thinking from the point of view of the computers too: Having VHs as 'machine humans', the question arises how to unify the seemingly conflicting qualities of machines and humans? Do we want a VH to be hesitant or even make errors, to be non-deterministic, or, just the opposite, are we eager to use the qualities made possible by the machine? Do we need the fallible human qualities to increase the naturalness and believability of human-looking software entities? Are there additional, practical values too of human imperfections? Is there a third choice, namely VHs which unify the functionally useful capabilities of machines and humans, and thus are not, in principle, mere replicas of real humans?

Before embarking on these questions, we want to relate them to the present state of VH research. It is an enormous challenge in itself to make a VH (a 'machine') talk, smile, gesture and move 'perfectly' as humans do. These efforts are by all means essential. The imperfections we are going to address are additions, usually involving several modalities, often on the cognitive level. Also, imperfections of the technology are not what we consider desirable, yet acceptable. E.g. the imperfection of most of the current TTS (Text-to-Speech) systems is a major bottleneck in developing engaging conversational agents, as the robot-like speech quality is difficult to understand and tiring to listen to. Yet the language spoken by the VH should not consist of only complete and correctly formulated sentences. When hesitating, thinking or recalling an episode, people may use partial or ill-formed utterances, which they may correct afterwards. The modeling and generation of such a phenomenon is, from a computer science point of view, a challenge similar to the generation of 'correct' utterances.

## 1.2 About this Paper

In this paper we focus on those qualities of real humans that are characteristically present in every-day life, but are hardly covered by the efforts on attaining higher-level cognitive model-based behavior of VHs. We will look at subtleties and 'imperfections' inherent in human-human communication, and investigate the function and merits of them. By *subtleties*, we refer to the rich variations we can employ in verbal and nonverbal communication to convey our message and the many ways we can draw attention to our intentions to convey a message. By *imperfections* we refer to phenomena which are considered to be incorrect, imperfect according to the normative rules of human-human communication. Both enrich the model of communication, the first one by taking into account more aspects and details (e.g. emotions and personality of the speaker), the second one by being more permissive about what is to be covered. We consider imperfections as those natural phenomena of every-day, improvised language usage which are not considered to be correct and thus, are not 'allowed' by some descriptive rules of proper language usage, put together by an academic body. For instance, re-starting a sentence after an incomplete and may be grammatically incorrect fraction is such an imperfection. We are aware though of the continuous change of norms for a language, pushed by actual usage. The space here does not allow dwelling on the relationship between intelligence and the capability of error recovery and robust and reactive behavior, in general. Here we present the issue from the point of view of communication with virtual humans and from the point of view of perception of VHs. In the next section we touch upon some related work on human-like qualities in VH technology and present a few examples from own work in different application domains. The two further sections address certain subtleties and imperfections in VHs. In the closing section, we come back to the initial question, that is, why allow subtleties and imperfections in the communication with virtual humans, but now from a perceptional point of view.

# 2. VIRTUAL HUMANS: ENGAGEMENT AND ENJOYMENT

## 2.1 Introduction

The main motivation to develop VHs is in applications where human communication with (real) humans is the norm such as sales, news reading, tutoring, and in applications requiring human 'players', such as training by simulation. In these applications, it is essential that the (potential) users like the application, find the VH believable, get engaged… and perform the task required. In the field of VH research, several subtle human qualities have been proven useful for VHs besides the straightforward usage of natural communicational modalities. It has turned out that small talk, showing interest and providing empathic feedback, getting emotional, exhibiting a coherent personality are human qualities which play an important role in both the subjective judgment of a (virtual) human partner and the objective influence of him. It has also been addressed that virtual humans should be individuals, in their behavior, verbal and non-verbal style, body, face and outfit [21]. Some subtle aspects of interaction, like the style of the VH [20] and his/her attitude towards the user have proven to be important in judging them. Comprehensive surveys of technical and cultural issues in the design of VHs can be found in [17] and [15], respectively.

In this section, we present two applications currently being developed at the HMI (Human Media Interaction) research group: the Virtual Dancer [19], and the Virtual Trainer [22]. These two novel applications are summarized in preparation to our general discussion on subtleties of VHs, where we will use illustrative examples from these two applications. Both applications require virtual humans with capabilities beyond the ones in more restricted or traditional functions such as providing information or tutoring. These two seemingly very different applications share some basic features, and have actually been developed relying on a similar framework. In both applications, the VH:

- has visual and acoustic perception capabilities,
- has to monitor and react to the user continuously,
- has to use subtle variants of a motion repertoire generated on the fly, and
- uses both acoustic (music, speech) and nonverbal modalities in a balanced and strongly interwoven manner.

## 2.2 The Virtual Dancer

In a recent application built at HMI, a virtual human – the Virtual Dancer – invites a real partner to dance with her [19] (see Figure 1). The Virtual Dancer dances together with a human 'user', aligning its motion to the beat in the music input and responding to whatever the human user is doing. The system observes the movements of the human partner by using a dance pad to register feet activity and the computer vision system ParleVision[1] to gain information about arm and body movements. Using several robust processors, the system extracts global characteristics about the movements of the human dancer like 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

[1] http://hmi.ewi.utwente.nl/showcases/parlevision

that are in some way ‘appropriate’ to the dancing style of the human dancer.

Figure 1: Interacting with the Virtual Dancer

There is a mapping from the characteristics of the observed dance moves to desirable dance moves of the Virtual Dancer. The interaction model reflects the intelligence of the Virtual Dancer. By alternating patterns of following the user or 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 being addressed in this research.

Clearly, the domain of dancing is interesting for animation technology. Previous research in this area focused on the regeneration of new dance sequences from captured dancers [9,12,23] or on animation specification and execution [11,16].

In our research we focus on the interaction between human and virtual dancer. The interaction needs to be engaging, that is, interesting and entertaining. Efficiency and correctness are not the right issues to focus on. In this interaction perfectness can become boring and demotivating.

## 2.3 The Virtual Trainer

The *Virtual Trainer* (VT) application framework is currently under development [22] and involves a virtual human on a PC, who presents physical exercises that are to be performed by a user, monitors the user’s performance, and provides feedback accordingly at different levels. Hence, our VT should fulfill most of the functions of a real trainer: it not only demonstrates the exercises to be followed, it should also provide professionally and psychologically sound, human-like coaching. Depending on the motivation and the application context, the exercises may be general fitness exercises that improve the user’s physical condition, special exercises to be performed from time to time during work to prevent for example RSI (Repetitive Strain Injury), or physiotherapy exercises with medical indications. The focus is on the reactivity of the VT, manifested in natural language comments relating to readjusting the tempo, pointing out mistakes or rescheduling the exercises. When choosing how to react, the static and dynamic characteristics of the user and the objectives to be achieved are to be taken into account and evaluated with respect to biomechanical knowledge and psychological considerations of real experts. For example, if the user is just slowing down, the VT will urge him in a friendly way to keep up with the tempo, acknowledge with cheerful feedback good performance and engage in a small talk every now and then to keep the user motivated.

Related work on VTs can be found in, among others, [4], where a physiotherapist is described with similar functionality as ours, [2] with an interesting Tai Chi application, and [1], reporting about work on an aerobics trainer.

# 3. SOME SUBTLETIES FOR VHs

## 3.1 Who is in Control?

VHs enter domains where they are not to be the dominant partner, and thus the control scenario is not well established. On one hand, the VH should not be completely dominant and thus needs to be able to follow the initiative of the human partner. On the other hand, VHs that show no initiative of their own do not seem to be human-like conversational partners but more like responsive machines that are to be controlled by the user using multimodal commands.

In the Virtual Dancer application, this issue of control is one of the important points of the application. The control between human and VH is both mixed-initiative and implicit: the VH alternates phases of ‘following dancing behavior’ where it incorporates elements of the human’s dance in its own dance with phases of ‘leading dancing behavior’ where it introduces new elements which the human will hopefully pick up.

In the VT scenario, the issue of control is essential, and subtle. Basically, the user is to perform the instructions given by the VT. However, this is not what happens all the time. The reaction for the VT depends on the assessment of the situation, including past performance and knowledge of the user. When the VT concludes that the user has just lost tempo, or is getting a little lazy, than the VT reinforces his/her own tempo to the user. But if the user looks very exhausted, or (s)he has a ‘bad day’ of decreased performance, then the VT may ‘give in’ to the user and slow down his/her tempo to comfort the user. Hence the VT is attentive and reactive, instead of imposing a predefined scenario on the user. The decision concerning when and how ‘to give in’ to the user must be based on detailed domain-specific and psychological knowledge.

## 3.2 Coordination of Modalities

There are several reasons why coordination of modalities in our applications needs a level of subtlety and sophistication beyond what is present in much of the current work.

For example, in present VH applications it is usually assumed that speech is the leading modality, and accompanying gestures and facial expressions should be timed accordingly. Dropping this assumption influences the design of a balanced multi-modal behavior that does not place natural language in a privileged role. It also influences the planning algorithms used to generate the behavior. We know of one work only where the instruction utterances’ tempo is scaled to the duration of the hand movements in explaining an assembly task [10]). This clearly indicates how rarely this issue is attended by current research efforts.

Another challenging characteristic for both applications is the need for alignment of multimodal behavior to external channels or events. Such feature is essential in any application where the VH does not converse by switching between pre-programmed active speaking and passive listening but reacts according to a model of bidirectional communication behavior. 'Outside events' are usually not covered in the framework of modality coordination, but as an issue of planning specific gestures (e.g. pointing at or reaching for moving objects). However, the fact that behavior in any of the modalities may be constrained in its timing by sources outside the VHs' influence calls for subtle and strongly adaptive planning and re-planning.

In the Virtual Dancer application the issue of coordination focuses wholly on the alignment of the dance behavior to the music. There is also a relation between the performed dance 'moves' and the dance of the user, but as yet, there is no tight timing relation between those two.

In the VT scenario, the issue of coordinating speech, motion and a given piece of music has turned out to be central, relating to two aspects mentioned above. Namely, a virtual trainer explains postures and basic movements and conducts rhythmic exercises. The exercises may be performed in different tempi. The speech like counting and providing comments on posture, which is often of secondary importance, should be aligned to the motion. The alignment should be subtle, in one case making sure that the emphasized syllabus of the counting coincides with the end of the stroke of the movement, while in another case the counting starts together with a repetitive motion unit, or an expressions is uttered in an elongated duration, during the finishing movement of a series. This requires real-time reactive planning of speech, occasionally resulting in what would usually be seen as 'unnaturally slow' utterances which are justified by the elongated duration of the corresponding motion. TTS systems may not even be prepared to generate such slowed down, unnatural speech. On the other hand, the timing of the exercise presented by the VT may be driven by external audiovisual cues of varying tempo. An exercise author may specify the tempo of an exercise by clapping or tapping, or as being aligned to music beats. However, this is often not enough in the case of VTs. For example, a VT may need to align its performance to the user doing the same exercise while counting along. This again requires a subtle real-time planning, affecting duration and alignment of sub-segments of a motion.

# 4. BEYOND PERFECTION

In the introduction, we raised the question whether we should aim for 'perfect' VHs or whether there are practical values to human imperfections too. As with everything, there are a multitude of different viewpoints on classes of imperfection. Imperfections in human communicative behavior can comprise imperfections in the *execution of chosen communicative acts* (e.g. mumbling, sloppy gestures …), the *information transfer* (e.g. ambiguous wording, redundant or incomplete content …), the *knowledge or decisions*, etc. In this section we discuss some examples where such imperfections may be desirable rather than something to be avoided.

## 4.1 User Adaptation

At first sight one might require that a VH is as perfect as can be, in its cognitive decisions, conversational and motion skills. One reason for lifting this requirement may be user adaptation. As a VH may come across users of different intellectual and communicational level, the capability to recognize such a situation and scale down the VHs' own functioning is more beneficial than overwhelming the user with a perfect and too demanding performance. This is already happening in the case of some TTS services, where the generated English pronunciation is not the 'official nicest' one, but one 'tortured' according to the practice of users whose mother tongue, like many Asiatic languages, has a very different acoustic scheme. Adapting to the level of a user by hiding some of the cognitive capabilities of the VH is already a usual practice in gaming and in tutoring situations. Lastly, the VT scenario exemplifies a possible reason for adapting the motion skills to those of the user by showing up in the embodiment the most appropriate for the given user. His/her gender, age and motion characteristics may be similar to the user's, in order to avoid having a too huge gap between or example a 'fit and young' trainer and the 'fattish, somewhat rigid-moving' user. On the other hand, deviations from usual trainers in the other (superman) direction may have positive effects too. Imagine the VT making every now and then extreme, beyond-realistic-capabilities jumps to cheer up the user or grab attention.

## 4.2 Clarification and Commitment

Certain types of imperfections in communication need 'repair'. A (virtual) human who mumbles needs to be asked to repeat itself. A (virtual) human using ambiguous language, or serious disfluencies, may be asked for clarification, literally, or through nonverbal mechanisms such as a lifted eyebrow or a puzzled expression. The traditional judgment of such repair in Human-Computer interaction is as an undesirable, necessary evil, hence the term 'repair'. However, as Klein et al. state, one of the main criteria for entering into and maintaining a successful joint activity is "Commitment to an intention to generate a multi-party product" [8]. This commitment needs not only be present, but must also be communicated to the conversational partner. Imperfections in the communications of a VH, and the subsequent so-called 'repair dialogues', could offer the perfect occasion for both VH and human to signal their commitment to the interaction. One must be really committed to an interaction if one is going through the trouble of requesting and/or giving clarifications, repeating oneself, etc...

There are two view angles to this commitment issue. The first relates to imperfections at the side of the human user. When we assume that entering into clarification and repair dialogues is a strong signal of commitment to the conversation, we see a clear reason to invest in the development of techniques for clarification dialogues beyond what is needed to let the conversation reach its intended goal. We even see a reason to make the VH initiate clarification dialogues when they are not absolutely necessary for reaching the goal.

The second relates to imperfections at the side of the VH. If the human is forced to ask the VH for clarification all the time, this will be considered a drawback in the interaction capabilities of the system. However, there may be certain subtle advantages to a VH that uses ambiguous, disfluent or otherwise imperfect expressions. Judicious use of ambiguities at non-critical points in the conversation, at a point where it is likely that the user will ask for clarification (explicitly or nonverbally), gives the VH a chance to show off its flexibility and its willingness to adapt to the user. This gives the human user a feeling of 'being understood' and of commitment from the side of the VH. Again, such an approach would need a lot of investment in repair and clarification capabilities of the VH.

Finally, one must remember that in many existing cases machine provided statements such as 'give the security code of your credit card', which should be clear and unambiguous, are not. If 'space' for clarification is not provided, the user will not only get frustrated but the task performance will decay.

## 4.3 Signaling Mental State and Attitude

In other situations imperfections express an explicit part of the content of the conversation. Imperfections in the multimodal generation process, such as hesitation, stuttering, mumbling and disfluencies can signal indications of the VH's cognitive state (e.g. 'currently thinking'), conversational state ('ready to talk'), attitude towards the conversation partner or the content (belief, certainty, relevance, being apologetic).

Cognitive (e.g. thinking ) or conversational (e.g. 'ready to talk') state is often reflected by gaze and body postures of the VHs. However, the usage of (non-speech) vocal elements in these and other situations has not been addressed widely yet. For example, by analyzing a multi-party real-life conversation, we found that non-speech elements abundantly interwove with the 'meaningful, articulated' utterances. While some of the erroneous and non-verbal utterances reflect the 'processing deficiencies' of the speaker (e.g. difficulty in formulating a statement in correct format), others have important function in regulating the dialogue (e.g. indicating request for turn taking by making some sound) or in qualifying the 'verbatim' content (e.g. a hesitant pause before making a statement indicates that the information to be conveyed may not be correct). In our analysis of multi-party real-life conversation we could identify speech situations as well as personality, emotional and cognitive state of the speaker as indicators of the frequency and type of non-speech elements used.

## 4.4 The Importance of Imperfections

The above subsections presented a number of considerations for turning VHs 'imperfect'. Before ending the paper with a short discussion we will touch lightly upon a few of the countless other themes that in some way also can involve imperfections.

Disfluencies and other imperfections are a major characteristic of spontaneous speech. If a VH talks without disfluencies it may come across as stilted, not spontaneous enough. In the ongoing work mentioned earlier we do encounter a common phenomenon in the language usage itself: speakers do not express themselves in perfect sentences, especially if they are answering an unexpected, unusual question or contribute to a discussion. Often, they abandon an erroneous start of a sentence and correct, or repeat the start in another form.

Ambiguous wordings and underspecification will help to give the user a sense of 'freedom' in the dialogue, to feel less constrained [7]. (Note though that we then run the risk of the user indulging in the same kind of ambiguous language use to an extent that the VH cannot handle it). Hesitations, mumbling and disfluencies are also mechanisms that play an important role in reducing the amount of threat potentially perceived by the human user.

Making mistakes may have the positive side-effects of users perceiving the VH as more believable and more individual. Humans are not perfect, and usually there is not a single 'perfect' way of action. When researchers create a model of, for example, turn taking behavior as a general norm, and we then implement this model in a VH, the fact that such a model is an abstraction by nature makes certain that the VH will behave in a way that no *real* human would do. Making mistakes also gives a VH unrivalled opportunity for exhibiting an individual style. After all, as humans we can probably distinguish ourselves from others through the mistakes that we make as much as through the things we do perfect.

# 5. DISCUSSION

We have argued that carefully 'designed' imperfection in the communication may increase not only the believability but also the scope of applicability and effectiveness of the VH in question. People are not uniformly perfect, but have different capabilities, and have means to recognize and cope with errors and limitations. Endowing VHs with the imperfections of humans can help making them more 'comfortable' to interact with. The natural communication of a VH should not be restricted to multimodal utterances that are always perfect, both in the sense of form and of content.

This scenario aims at hiding the 'machine' nature of the VH. This may be very much what we want on the communication level, but how about the cognitive level? For instance, if a VH is to find an item from a huge database, or make a move as a chess opponent player, should the imperfect – that is, slower, error prone – human behavior be mocked up, even if the computer is ready with the perfect answer immediately? In general, a VH's amount and processing of knowledge should 'resemble' the capabilities of humans, in order to make the VH believable and life-like, as opposed to some omnipotent, super creature.

On the other hand, such a scenario may sound completely irrational, as it makes no use of the power of the computer. Given the fact that most humans are indoctrinated from birth with the adagio that 'computers are fast in calculating', hiding this capability behind artificial imperfection might even be perceived as unrealistic. Except maybe when the user is given to understand that (s)he interacts with the computer *through the intermediation of* the VH rather than with the computer *embodied by* the VH.

A fascinating, and not yet explored question is if VHs should converge to a specific type, who do communicate with humans in an effective and natural way, but have additional, 'beyond human' mental and communicational capabilities, thanks to the computing power behind them. We argue that such VHs may indicate that they are not 'real humans', by a non-photorealistic look, and eventually, by having a human-like communicative repertoire extended with beyond-realistic features known from cartoon or science-fiction [24]. But such VHs too should be, all the same, individual, fallible and creatures to be liked by humans, not a population of all-alike robots.

# 6. REFERENCES


[1] Babu, S., Zanbaka, C., Jackson, J., Chung, T., Lok, B. Shin, M.C. and Hodges, L.F. Virtual Human Physiotherapist Framework for Personalized Training and Rehabilitation. In *Proc. Graphics Interface 2005*, Victoria, Canada, 2005.

[2] Chao, S-P., Chiu, C-Y, Yang, S-N, and Lin, T-G. Tai Chi synthesizer: a motion synthesis framework based on key-postures and motion instructions. *Computer Animation and Virtual Worlds*, Vol. 15, 2004, 259-268.

[3] Cassell, J., Sullivan, J., Prevost, S., and Churchill, E. (Eds.). *Embodied Conversational Agents*. MIT Press, 2000.

[4] Davis, J.W. and Bobick, A.F. Virtual PAT: A Virtual Personal Aerobics Trainer. MIT Media Laboratory, TR 436, 1998.

[5] Fogg, B.J.. *Persuasive Technology. Using Computers to Change What We Think and Do*. Morgan Kaufmann Publishers, San Francisco, CA, 2003.

[6] Friedman, B. (Ed.). *Human Values and the Design of Computer Technology*. CSLI Publications, Cambridge University Press, 1997.

[7] Gaver, W., Beaver, J., and Benford, S. Ambiguity as a resource for design. In *Proceedings of CHI 2003*, 233-240.

[8] Klein, G., Feltovich, P.J., Bradshaw, J.M., and Woods, D.D. Common Ground and Coordination in Joint Activity. In W. B. Rouse and K. R. Boff (Eds.), *Organizational Simulation*. John Wiley, 2004, 139-184.

[9] Kim, T., Il Park, S., Yong Shin, S.: Rhythmic-motion synthesis based on motion beat analysis. *ACM Transactions on Graphics* **22**(3) (2003) 392–401.

[10] Kopp, S., and Wachsmuth, I. Model-based animation of co-verbal gesture. In *Proceedings of Computer Animation*, Geneva, June 2002, 252-257.

[11] Mataric, M., Zordan, V., Williamson, M.: Making complex articulated agents dance. *Autonomous Agents and Multi-Agent Systems* **2**(1) (1999) 23–43.

[12] Nakazawa, A., Nakaoka, S., Kudoh, S., Ikeuchi, K.: Digital archive of human dance motions. *Proceedings of the International Conference on Virtual Systems and Multimedia (VSMM2002).* (2002) 180–188.

[13] Nass, C., Isbister, K., and Eun-Ju Lee. Truth is Beauty: Researching Embodied Conversational Agents. Chapter 13 in [3], 374-402.

[14] Pantic, M., Pentland, A., Nijholt, A., and T. Huang. Human Computing and Machine Understanding of Human behavior: A Survey. In *Proceedings Eighth International Conference on Multimodal Interfaces (ACM ICMI 2006)*, Banff, Canada, November 2006 (these proceedings).

[15] Payr, S. and Trappl, R. (Eds.). *Agent Culture. Human-Agent Interaction in a Multicultural World*. Lawrence Erlbaum Associates, Mahwah, NJ, 2004.

[16] Perlin, K.: Real time responsive animation with personality. *IEEE Transactions on Visualization and Computer Graphics* 1(1) (1995) 5–15.

[17] Prendinger, H. and M. Ishizuka, M. (Eds.). *Life-Like Characters: Tools, Affective Functions and Applications.* Springer-Verlag, Berlin, Heidelberg, 2004.

[18] Reeves, B. and Nass, C. *The Media Equation: how people treat computers, televisions and new media like real people and places.* Cambridge University Press, 1996.

[19] Reidsma, D., Welbergen, H. van, Poppe, R., Bos, P., and Nijholt, A. Towards Bi-directional Dancing Interaction. In: R. Harper, M. Rauterberg, M. Combetto (Eds.), *Proceedings 5th International Conference on Entertainment Computing,* LNCS 4161, 2006, 1-12.

[20] Ruttkay, Z.M. and Pelachaud, C. Exercises of style for virtual humans. In *Proceedings of Animating Expressive Characters for Social Interaction Symposium*, Imperial College, London, April 2002, 85-90.

[21] Ruttkay, Z.M. and Pelachaud, C. (Eds.). *From Brows to Trust. Evaluating Embodied Conversational Agents*. Kluwer's Human-Computer Interaction Series – volume 7, 2004.

[22] Ruttkay, Z.M., Zwiers, J., Welbergen, H van, and Reidsma, D. Towards a Reactive Virtual Trainer. In *Proceedings of the 6th International Conference on Intelligent Virtual Agents, IVA 2006*, Marina del Rey, CA, USA, Springer Verlag, LNAI 4133, 2006, 292-303.

[23] Shiratori, T., A. Nakazawa A. & Ikeuchi, K. Dancing-to-Music Character Animation. *EUROGRAPHICS 2006*, E. Gröller and L. Szirmay-Kalos (Guest Editors), Volume 25 (2006), Number 3, to appear.

[24] Wang, J., Drucker, S., Agrawala, M., and Cohen, M. The Cartoon Animation Filter. *ACM Trans. on Graphics (Proc. of SIGGAPH2006)*, 2006, to appear.