Graceful Anytime Interruptibility for Virtual Agents

Klaus Brügmann
Friedrich-Alexander-Univ. Erlangen-Nürnberg
Computer Graphics Group
Am Weichselgarten 9, 91058 Erlangen, GER
d3lphi@gmx.de

Helmut Prendinger
National Institute of Informatics
2-1-2 Hitotsubashi, Chiyoda-ku
Tokyo 101-8430, JP
helmut@nii.ac.jp

Marc Stamminger
Friedrich-Alexander-Univ. Erlangen-Nürnberg
Computer Graphics Group
Am Weichselgarten 9, 91058 Erlangen, GER
stamminger@cs.fau.de

Mitsuru Ishizuka
Graduate School of Information Science
and Technology, Univ. of Tokyo
7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, JP
ishizuka@i.u-tokyo.ac.jp

ABSTRACT

The demo will show two highly expressive virtual agents introducing the audience to a traditional Japanese room. The agent models and gestures have been created by a professional Japanese character designer. By adapting and extending an influential technique ("Verbs and Adverbs"), we added parameterized emotional expression and gesture interruptibility as new behaviors to increase agent believability, interactivity, and responsiveness. In the demo, users will be provided a simple interface to test the new features.

Categories and Subject Descriptors

H5.1 [**Information Interfaces and Presentation (e.g., HCI)**]: Multimedia Information Systems—*animations*

General Terms

Design

Keywords

Character animation, Verbs and Adverbs technique

1. INTRODUCTION

The deployment of life-like interface agents is gaining popularity in various application domains, ranging from e-learning and computer-mediated interaction to experimental multi-agent interaction environments, games, and interactive storytelling [5]. It has been shown that virtual agents may improve human–computer interaction for the dimensions of naturalness, effectiveness, and entertainment value. However, as pointed out in [2, p. 54] "[...] because an agent looks like a human, people expect it to behave like a one as well and will be disturbed by, or misinterpret, discrepancies from human norms." The expression of emotion and personality is often seen as a core requirement for life-like and believable behavior. An influential approach is described in [1], where a subset of the Laban Movement Analysis (LMA) method is used to parameterize agent behaviors according to emotional state.

ACE'07, June 13–15, 2007, Salzburg, Austria.
ACM 978-1-59593-640-0/07/0006.

The aim of our work on agent animation is twofold. *First*, we intend to increase the naturalness of emotional expressivity of 3D agents by gesture parametrization (intensity and speed) and smooth transitions between affect-conveying gestures. Hereby, we adapted a previously developed method for motion parametrization and interpolation, the so-called "Verbs and Adverbs" (V&Adv) technique [6]. While V&Adv focused on walking motions of simplistic (stick-like) figures, our interest is in gesture behavior of highly realistic 3D agents. Our *second* aim is to increase the level of interactivity in virtual agent performances. The agent system and authoring language MPML3D described in [3] allow for anytime interruption of verbal utterances. We extended this concept to also cover interruptibility of gestures.

2. GESTURE PARAMETRIZATION

For gesture parametrization, the "Verbs and Adverbs" technique was applied [6]. V&Adv is a method to generate new motions for virtual characters from existing ones on a frame-by-frame interactive basis. The V&Adv concept introduces a graph which depicts available basic motions, called "verbs", and legal transitions between them. For each basic motion, several variations (motion samples) representing nuances of motion styles exist, which are called "adverbs". Adverbs span a multidimensional expression space that is populated by the available sample motions. In addition to smooth transitions between verbs, the technique offers a way to produce interpolated motion output based on the adverb space. The V&Adv technique requires that motion variants are similar in anatomy. However, it allows for samples with differing keyframe timing as well as differing overall duration. Unlike other approaches, it does not rely on the frequency domain. It is based directly on keyframes in the time-domain, and thus integrates processing of non-periodic motions in a straight-forward way.

2.1 Parameterized Emotion Expression

The V&Adv technique overcomes the limitations of discrete gesture scheduling approaches and thus supports different intensity levels in the expression of emotions. Here we briefly describe how V&Adv was adapted for the purpose of emotion-loaded gestures. As a basis, the character designer prepared animation samples for several gestures types (iconic, metaphoric, etc.), whereby each gesture is provided

Figure 1: Example of iconic gesture: "A typical byoubu panel is [this] wide."

at different levels of speed and spatial expansion. For instance, Fig. 1 illustrates one sample of an iconic gesture for showing size horizontally. Using that pose, the size shown can be defined and adjusted continuously and smoothly at runtime while keeping the actual pose. Other gestures will convey a calm or hectic emotional state, depending on performance speed. Internal state parameters are used to interpolate the actual motion on a frame-by-frame basis.

According to the A&Adv approach, for each DOF (degree of freedom) a motion function is evaluated at every frame. This function is not represented explicitly, but evolves at runtime, depending on the current parameter settings that constitute a point in the adverb space. To generate the actual motion, the sample motions of the current verb will be interpolated, weighted according to the distance of their own location to the current adverb setting, using radial basis functions. To generate consistent blends, sample motions have to be similar in structure, i.e. consist of the same set of movement phases. However, those do not have to coincide temporally, thus allowing for individual phase lengths of sample motions. Before interpolating, time warping is applied to align the phases of the involved sample motions. As the interpolated motion function is computed at each frame, it is also guaranteed that expression parameters can be adjusted anytime.

2.2 Non-Repetitiveness

Besides emotion expression with varying intensity levels, gesture parametrization also achieves a reduction of unnatural repetition of movements. This is important since human actors will always have some variation in their gestures, even if they intend to perform the very same gesture. Moreover, we can reduce repetitiveness by applying some randomization to various so-called 'idle' gestures, including "swaying" and "breathing". The latter are keyframed movements that are blended with actual action gestures to promote movement variation control. Randomization is achieved by integrating noise functions into gesture animation [4].

3. GESTURE INTERRUPTION AND TRANSITION

A useful feature of A&Adv is that it can be used for high-level animation control of non-periodic motions, as required for real-time behavior planning and real-time interruptibility. To generate smooth and realistic transitions from one movement to another, V&Adv specifies a synchronization interval for each pair of movements where a transition is legal. In our case, however, transitions must start immediately to produce realistic animation. Therefore our animation engine does not require animations (gestures) to reach a predefined synchronization frame before transiting.

When a character prematurely stops the current gesture because it is interrupted by another, and continues to perform a different gesture, a transitional curve will be applied that forms a transition between the instant of interruption and the entering frame of the gesture to be performed next. For each DOF of the relevant parts of the mesh, a Bezier curve of degree three will be generated to preserve the trajectory's C^2-continuity. Depending on the current speed of motion, the starting time of the succeeding motion is scheduled, and the curve is re-parameterized to ensure an appropriate transition period. As in V&Adv, the transition curve will be computed for each frame (if necessary), to account for parameter changes affecting the succeeding gesture. In this way, we can produce smooth graceful transitions that may occur anytime during the performance of a gesture.

Acknowledgements

The research was supported by the Research Grant (FY1999–FY2003) for the Future Program of the Japan Society for the Promotion of Science (JSPS), by a JSPS Encouragement of Young Scientists Grant (FY2005–FY2007), and an NII Joint Research Grant with the Univ. of Tokyo (FY2006). The first author was partly supported by a German Academic Exchange Service (DAAD) Scholarship.

4. REFERENCES

[1] N. I. Badler, J. Allbeck, L. Zhau, and M. Byun. Representing and parameterizing agent behaviors. In *Proceedings of Computer Animation Conference (CA-02)*, pp. 133–143. IEEE Computer Society, 2002.

[2] J. Gratch, J. Rickel, E. André, J. Cassell, E. Petajan, and N. Badler. Creating interactive virtual humans: Some assembly required. *IEEE Intelligent Systems*, pp. 54–63, 2002.

[3] M. Nischt, H. Prendinger, E. André, and M. Ishizuka. MPML3D: a reactive framework for the Multimodal Presentation Markup Language. In *Proceedings 6th International Conference on Intelligent Virtual Agents (IVA-06)*, Springer, pp. 218–229, 2006.

[4] K. Perlin. An image synthesizer. In *Proceedings of SIGGRAPH'85*, pp. 287–296. ACM Press, 1985.

[5] H. Prendinger and M. Ishizuka, editors. *Life-Like Characters. Tools, Affective Functions, and Applications*. Cognitive Technologies. Springer Verlag, Berlin Heidelberg, 2004.

[6] C. Rose, B. Bodenheimer, and M. F. Cohen. Verbs and adverbs: multidimensional motion interpolation. *IEEE Computer Graphics and Applications*, 18(5):32–40, 1998.