

SmartBody: Behavior Realization for Embodied Conversational Agents

Marcus Thiebaut¹, Andrew N. Marshall¹, Stacy Marsella¹, Edward Fast²,
Aaron Hill², Marcelo Kallmann³, Patrick Kenny², and Jina Lee¹

¹ USC Information Sciences Institute
{thiebaut, amarshal, marsella, jinal}@isi.edu

² USC Institute for Creative Technologies
{fast, hilla, kenny}@ict.usc.edu

³ University of California, Merced
mkallmann@ucmerced.edu

Virtual humans should be *responsive* to unexpected events, *life-like* in their behavior and *interpretable* so users can interpret their response to situations, including their cognitive and emotional state. To address such demands, a variety of approaches have been used to animate virtual humans. While procedural animation is flexible to respond to events, hand-animated (or motion captured) sequences are better suited to display life-like expressive behavior.

We propose that a combination of animation approaches may be required to achieve responsive, life-like and interpretable behavior. Our motion control architecture [1] is based on controllers that can be hierarchically interconnected in real-time in order to achieve continuous motion. This approach has evolved into a general character animation system called SmartBody, an open source modular framework for realizing embodied characters. Controllers in SmartBody can employ arbitrary animation algorithms such as keyframe interpolation, motion capture or procedural animation. Controllers can also schedule or blend other controllers. Composable controllers provide flexibility and a wide range of potential behaviors from a smaller set of simpler components.

In creating SmartBody, we seek a platform that supports research and fosters collaboration in virtual human animation techniques as well as reduces the barrier of entries to crafting virtual human applications.

1 Overview of Architecture

The problem of synthesizing realistic behaviors for conversational characters has been addressed by several researchers (e.g., [2–4]). The main approach taken has been to develop an animation architecture to fine tune it to meet with the behavioral requirements at hand. Our main contribution is the design, development and experimentation of a generic system able to address wide requirements. Our approach is based on three main concepts: 1) development of primitive behavioral motion controllers, 2) online hierarchical composition of controllers, and 3) a behavioral and schedule manager able to assemble and orchestrate controllers following input BML specifications.

Smartbody implements the behavior realization component of the SAIBA embodied conversational agent framework [5]. That is, it transforms BML performance descriptions into character animation and synchronized audio.

The behavior processing of Smartbody has two major components. The first is the behavior & schedule manager, which parses incoming BML. In our applications, this BML comes from a non-verbal behavior generator (NVBG) [6] that generates BML from the agent's communicative functional markup language (FML) request. The behavior & schedule manager maps each BML request to a set of skeleton-driving motion controllers, feedback events, and schedule edits.

The second major component is the motion controller engine that maintains and evaluates a hierarchical schedule of immediate and future motions for each character. The behavior & schedule manager makes necessary edits to this schedule, usually comprised of inserting new motions. For each evaluation cycle, the controller engine executes and combines the immediate motion controls, and generates a set of skeletal joint rotations and translations.

The results are sent over a fast network protocol to the system's rendering engine. The renderer applies the skeletal motion to the character's deformable mesh, displayed within the application scenario. This separation from the application rendering system keeps SmartBody portable to multiple rendering engines and allows us to distribute the animation processing of multiple characters.

Smartbody has been applied successfully to six virtual human projects, including SASO-ST, ELECT, Virtual Rapport and Virtual Patient. The projects range from research platforms to training applications (Figure 1). Overall, SmartBody has proven to be efficient and flexible, greatly reducing time and cost.

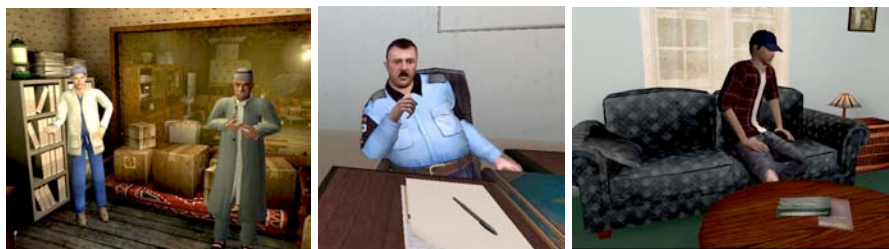


Fig. 1. SASO-ST, ELECT & Virtual Patient.

References

1. Kallmann, M., Marsella, S.: Hierarchical motion controllers for real-time autonomous virtual humans. In: Proc. of Intell. Virt. Agents (IVA'05), Kos, Greece (Sept. 2005) 243–265
2. Kopp, S., Wachsmuth, I.: Synthesizing multimodal utterances for conversational agents. *Computer Animation and Virtual Worlds* **15**(1) (2004) 39–52
3. Cassell, J., Vilhjálmsón, H.H., Bickmore, T.W.: Beat: the behavior expression animation toolkit. In: Proceedings of SIGGRAPH. (2001) 477–486
4. Carolis, B.D., Pelachaud, C., Poggi, I., de Rosis, F.: Behavior planning for a reflexive agent. In: Proc. of the Int. Joint Conf. on Art. Intell. (IJCAI'01), Seattle (Sept 2001)
5. Kopp, S., Krenn, B., Marsella, S., Marshall, A., Pelachaud, C., Pirker, H., Thorisson, K., Vilhjálmsón, H.: Towards a common framework for multimodal generation: The behavior markup language. In: Proc. of Intell. Virt. Agents (IVA'06). (2006) 105–111
6. Lee, J., Marsella, S.: Nonverbal behavior generator for embodied conversational agents. In: Proc. of Intel. Virt. Agents (IVA), Marina del Rey, CA (2006) 243–255