

At the Virtual Frontier: Introducing Gunslinger, a Multi-Character, Mixed-Reality, Story-Driven Experience

Arno Hartholt¹, Jonathan Gratch¹, Lori Weiss¹, the Gunslinger Team²

University of Southern California, Institute for Creative Technologies, 13274 Fiji Way,
Marina del Rey, CA, USA

¹{hartholt, gratch, weiss} @ ict.usc.edu

1 Introduction

Gunslinger is an interactive-entertainment application of virtual humans that transforms an iconic Wild West movie scene into a vivid semblance of reality. The project combines virtual humans technology with Hollywood storytelling and set building into an engaging, mixed-reality, story-driven experience, where a single participant can interact verbally and non-verbally with multiple virtual characters that are imbedded in a real, physical saloon. The Gunslinger project pushes the frontier of virtual humans research by combining question-answering dialogue techniques with explicit story representation. It incorporates speech recognition techniques and visual sensing to recognize multimodal user input. It further extends existing behavior generation methods such as SmartBody to drive tightly coupled dialogue amongst characters. These capabilities strive to seek a balance between open ended dialogue interaction and carefully crafted narrative.



Fig 1. The Gunslinger set.

2 System

Gunslinger is based on the Institute of Creative Technologies virtual humans architecture [1]. A speech recognizer is used to give hypothesized transcriptions of a raw speech signal [2]. A visual recognition module detects and tracks the participant's motion in the physical space over time while the visual understanding module infers higher level behaviors such as drawing a gun, facing a character, and movement with respect to a character. The NPCEditor is used for natural language processing and dialogue management [3]. It uses a statistical text classification algorithm that selects the character's responses based on the user's utterances. Gunslinger uses the Non-Verbal Behavior Generator (NVBG) [4] and SmartBody [5] to plan and realize the

story's three characters, respectively. NVBG analyzes the speech text to propose non-verbal behaviors that are passed to SmartBody using the Behavior Markup Language. Sound is realized using ARIA, a real-time audio processing engine which leverages commodity computing hardware for high-quality audio processing [6].

3 Conclusions and Future Work

The Gunslinger project allowed us to advance the virtual humans work in such areas of multi-modal input, multi-party dialogue and story-driven interaction. The work advanced the state of these individual components and resulted in an architecture that could exploit synergies between them. We were able to advance the work being done in visual recognition and understanding by adapting efficient and reliable visual-sensing approaches to the problem of recognizing human nonverbal behaviors. The NPCEditor has been extended substantially with multiple characters support, scripting capabilities, story phases, vision integration and interruption. In the future we plan to extend Gunslinger in several directions. We will explore ways to increase the sensory fidelity of the system through the use of photorealistic characters, 3D motion parallax, and "4D" effects such as projected bullet holes. We are performing extensive user testing with the aim of extending the basic story line to allow a much richer range of user preferences. In summary, Gunslinger provides an exciting research sandbox to explore the limits of dramatic human-to-virtual-human interaction.

² Anton Leuski, Louis-Philippe Morency, Matt Liewer, Marcus Thieboux, Stacy Marsella, Prathibha Doraiswamy, Andreas Tsiartas, Kim LeMasters, Ed Fast, Ramy Sadek, Andrew Marshall, Jina Lee, Lance Pickens.

References

1. Kenny, P., Hartholt, A., Gratch, J., Swartout, W., Traum, D., Marsella, S., Piepol, D.: *Building interactive virtual humans for training environments*. In: Proceeding of I/ITSEC. (2007)
2. Sethy A., Narayanan S., and Ramabhadran B.. *Data-Driven Approach for Language Model Adaptation using Stepwise Relative Entropy Minimization*. In Proc. of ICASSP, volume 4, pages 177-180, 2007.
3. Leuski, A. and D. Traum, *A statistical approach for text processing in virtual humans*, in *26th Army Science Conference*. 2008: Orlando, Florida.
4. Lee, J. and S. Marsella. *Nonverbal Behavior Generator for Embodied Conversational Agents* in 6th International conference on Intelligent Virtual Agents, Marina del Rey, CA, 2006
5. Marcus Thieboux et al, *SmartBody: Behavior Realization for Embodied Conversational Agents*, in Proceedings of Autonomous Agents and Multi-Agent Systems, 2008.
6. Sadek, Ramy *A Host-Based Real-Time Multichannel Immersive Sound Playback and Processing System*, Proceedings of the Audio Engineering Society 117th Convention, New York, Oct 2004