Spotlight on Sergey Levine: Automatic Generation of Avatar Gestures

Karis Tang-Quan

Second Life. Keneva. Cybertown. Moove. The list of virtual worlds continues to grow as more and more users subscribe to the relatively new pastime of virtual reality. The increasing number of subscriptions means the creation of more three-dimensional avatars, or a computer user’s alter ego in the virtual world. This popularity of avatars calls for better character animation, including body language animation. Sergey Levin is answering this call.

Working in the Department of Computer Science, Sergey has completed a project on the automatic generation of gestures for avatars in virtual worlds. Since gestures and body motions are an important part of human communication, Sergey felt it was vital to include these interactions in the virtual environment as well.

The result is that avatar interactions would be closer to the actual interactions that go on between human beings. As is expected with computer science projects, Sergey’s research methodology deviated from the typical natural science experimental design or humanities readings. In the end, Sergey’s work produced results in the real world for the virtual world.

First, Sergey, with the help of his research group, recorded a body of motion capture data. The data were segmented in a way to generate relationships between speech and the occurrence of common gestures. A probabilistic model was then trained to exploit the found relationships. By using the common relationships, this probabilistic model selects the appropriate motion to match a new speech signal. The system of the probabilistic model was designed to flow accurately and smoothly from gesture to gesture based on the current speech.

Real people evaluated the model by comparing the synthesized motions to the motions of real human beings. Using a hidden Markov model, Sergey was able to create a well-trained probabilistic model for his project.

With this probabilistic model, Sergey then wanted to create a human character that had body gestures to match his or her spoken words. Not only would the gestures match the speech, but also the system would run in real-time with no knowledge of future phrases. The probabilistic model would be able to predict the appropriate future body movements based on its inferred data from the current speech. The timing and rhythm of the body gestures would correspond with the character’s speech, bringing life and reality to animated characters.

Giving life to animated characters is not an easy task. Sergey not only became an expert in his computer science field of study but also trained to understand linguistics, psychology, artificial intelligence, and animation. Since gesture synthesis is so interdisciplinary and Sergey began with a background in computer graphics, he accumulated a vast amount of knowledge through the course of one research project. Though it has been a challenge for him, it also provides the greatest reward: the understanding of unfamiliar subjects and original research.

The challenge of research has taught Sergey that an undergraduate research experience is, in his own words, “humbling.” Despite having a solid background in computer science and having taken numerous Stanford courses, Sergey realized how much he still has to learn. Research is not all about the end product, but what discoveries were made along the way.

Through his research experience, Sergey has realized the importance of reading the right articles and books and having those perfect conversations with key people that can stimulate ideas. Creativity and problem-solving skills provided the foundation for starting research, but the people around him made it possible to complete the project.

Sergey found support for his project in the Virtual Worlds Group in the Computer Science Department. The head of the group Professor Vladlen Koltun realized Sergey’s interest and ability for research in a class on Virtual Worlds. It was through this class that Sergey saw the research potential of networked virtual environments. Naturally Professor
Koltun, knowing that Sergey can meet the challenge of a broad, interdisciplinary subject, suggested Sergey join the Body Language project. It was a perfect match for Sergey to be involved in the development of new technologies that are currently being applied.

While receiving guidance from graduate student Jerry Talton in the Virtual Worlds Group, Sergey began working on the growing project. As his project expands, Sergey incorporates new people in his research. In addition to direct contact with Professor Koltun, Sergey works with Visiting Assistant Professor Christian Theobalt in the Artificial Intelligence Lab, trying to incorporate more sophisticated AI techniques in his project. Previously, Sergey has worked with the Stanford Biomotion Lab and Animotion Inc. to integrate an interdisciplinary approach to his research.

While the number of virtual world users growing, Sergey continues his research on generating the appropriate body gestures for human characters. He hopes that his work will be applied to virtual environments that lack the multi-modal aspect of communication that happens in the real world.

Recently, Sergey submitted his research work for publication and looks forward to hearing back. His current and future work includes an extension of the current system that trains the model. This improved system would reveal further insight to the correlation between motion and speech. These revelations are necessary for providing realistic character animation in the virtual world and vital information on behavioral interaction in the real world.

Through his humbling experience with interdisciplinary research, Sergey has learned an immense amount on human interactions through body motions. With this understanding, Sergey does not need a Second Life to give life to animated human characters.