Using motion capture data to generate and evaluate motion models for real-time computer animation

H. van Welbergen[1]

[1] *Human Media Interaction, University of Twente, Enschede, the Netherlands. welberge@ewi.utwente.nl*

Introduction

In the field of computer animation, we are interested in the creation of movement models that make a virtual human (VH) move in a natural way. Ideally, one should not be able to distinguish the movement of such a VH from that of a real human. Furthermore, we want to be able to exert control over such motion in real-time, so that the motion can be adapted or fully generated during interaction with VHs. Obtaining such control in real-time typically comes at the cost of naturalness. That is why VHs that look and move in a very natural way are seen in movies, where all behavior is predefined, while we can not interact with such natural looking and moving humans in real-time.

Our approach to motion generation is bottom up. We start out with motion capture data and replace the motion on a part of the body motion by a movement model. For example, we could replace the lower body movement by a balancing movement model, or we could replace movement of the head and eyes by a gaze model, while the rest of the body is still moved by motion capture. This way, a movement model can be evaluated in isolation in a user test, by comparing it with motion capture data.

Creating motion models

Motion editing techniques use the recorded motion directly in the motion model. The motion is generated using a combination of existing motion recordings. Recordings that are 'close enough' can be concatenated to generate new motion [1], or several motions can be blended to form a desired motion [2]. Control in motion editing techniques is about finding the right motions to combine and finding blend weights that produce a desired motion.

Physical models steer the motion of the body by applying forces in joints. In real-time physical animation, these forces are calculated by models from control theory: a desired state of the body is defined, and the forces are steered so that the body gets closer to this desired state [6]. For example, our physical balance model steers forces in the hips, knees and ankles, in such a way that the body's center of mass moves closer to a desired position in which the body does not fall over. Control in such models is about finding the right control model for a certain task, and about setting parameters in the desired state of the body (such as the height of the hips above the ground and the position of the balance point in our balance model).

Procedural simulation defines mathematical formulas that control motion, given motion time and a set of movement parameters. This can be used to directly control the rotation of joints [3]. A typical application is at a slightly higher level: the movement path of hands through space is defined mathematically to generate gestures [4].

Procedural models and physical models are typically created on the basis of models from biomechanics or behavior science, rather than directly basing them off motion capture. The parameters that steer these models are designed to be intuitive for motion authors, but are often related. Motion capture can serve as a way to find dependencies between these parameters. For example, we have shown that the movement path of the hand decrease linearly with the tempo in a clapping task [5]. A change of one parameter then changes all parameters that are related to it. If more than one parameter is specified, conflicts might arise. These conflicts can be solved in several ways, for example by finding some kind of 'best fit' of parameters values, weighted by their importance.

Evaluating motion models

VHs usually do not have a photo-realistic embodiment. Therefore, if the naturalness of VH animation is evaluated by directly comparing moving humans with a moving VH, the embodiment could bias the judgment. To remedy this, motion captured human movement can be casted onto the same embodiment as the VH. This motion is then compared with generated animation. Typically this is done in an informal way. A motion Turing Test [6] could be used to do this more formally. In such a test, subjects are shown generated movement and similar motion captured movement, displayed on the same VH. Then they are asked to judge whether this was a 'machine' moving or a real human. However, such a human judgment is not sufficient to measure the naturalness of motion. Even if a certain movement is judged as natural, an invalidation of naturalness that is not noticed consciously can still have a social impact [7]. Unnatural moving characters can be evaluated as less interesting, less pleasant, less influential, more agitated and less successful in their delivery. So, while a VH Turing test is a good first measure of naturalness (at least it looked human-like), further evaluation should determine if certain intended aspects of the motion are delivered. Such aspects could include showing emotion, enhancement of the clearness of a spoken message using gesture, showing personality, etc.

We use a movement model that steers a part of the body, and steer the rest of the body using motion capture. We can then compare the motion generated by the movement model combined with motion capture with the same motion generated solely by motion capture in a motion Turing test. In a similar way, we can test if a certain aspect of motion is important for naturalness, by using a model that either removes this aspect, or replaces it by noise. Our method does not only provide us the means to test motion models in isolation, but it also provides meaningful technology to combine motion models with kinematic motion. In a later stage, we plan to use this approach to test combinations of one or motion models that were evaluated to work well in isolation.

Acknowledgements

References

[1] Kovar L., Gleicher M. Pighin F. H. (2002). Motion graphs. *ACM Transactions on Graphics* **21(3)**, 473-482.

[2] Wiley D. J., Hahn J. K.(1997). Interpolation Synthesis of Articulated Figure Motion. *IEEE Computer Graphics Applications* **17(6)**, 39-45.

[3] Perlin, K. (1995). Real Time Responsive Animation with Personality. *IEEE Transactions on Visualization and Computer Graphics* **1(1)**, 5-15.

[4] Chi D. M., Costa M., Zhao L., Badler N. I. (2000). The EMOTE model for effort and shape. *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, 173–182.

[5] van Welbergen H., Ruttkay Z.M. (To appear). On the parameterization of clapping. *Proceedings of the 7th International Workshop on Gesture in Human-Computer Interaction and Simulation (Lisbon, May 26 2007)*

[6] Hodgins J. K., Wooten W. L., Brogan D. C., O'Brien J. F. (1995). Animating human athletics. *Proceedings of the 22nd annual conference on Computer graphics and interactive techniques*, 71-78.

[7] Reeves B., Nass, C. (1996). The media equation: how people treat computers, television, and new media like real people and places. Cambridge University Press.