Motion Editing: A Comparative Approach

Paul S A Reitsma

April 17, 2000

Committee: Nancy Pollard, John Hughes, David Laidlaw

Advisor: Nancy Pollard

1 Abstract

Motion editing is increasingly being accepted as a means for expanding the range of available motions with far less expenditure of time and effort than the original motion creation method. Adding flexibility and breadth to motion capture data is particularly valuable, provided the new motions are sufficiently "good." This research will attempt to not only explore new ways of deriving these "good" motions, but also to address the question of what it means for a motion to be "good" in different contexts. In particular, it will address the oft-neglected area of transitions between motions and the effects of preserving high-frequency information and/or blending different joint parameters on the character's balance through the generated composite motion.

2 Introduction

There are several widely-used methods for generating high-quality animated motions, with the most common being motion capture, controller-based technologies, and simple keyframing. Each of these has certain strengths and limitations; motion capture, for example, is excellent at generating realistic human motion, since real human motion is captured, but is limited to motions that can be performed within the confines of a motion capture studio. Additionally, all of these methods are both complicated and time-consuming, making them fairly inflexible to the needs of animators.

The idea behind motion editing is to allow animators to alter the motions generated by these techniques to the requirements of each specific animation. For it to be used, the editing must be significantly simpler and easier than simply generating appropriate new motions via the original technique. Additionally, the edited motions must be as high-quality as the originals, and maintain any particular characteristics of the original that the animator might want; the simplest system in the world will not be used if it does not give good enough results.

As well, motion editing offers not only the flexibility to derive new motions from existing ones more easily than creating those motions from scratch, it also holds promise for generating motions that could not be created by the original method at all. Motion capture data, for example, is limited to what an available actor can do in the confines of a motion capture studio; any action that will not fit inside a studio or that a human can not do simply can not be captured, but perhaps can be created by editing motions that can be captured.

Similarly, motion editing may allow different generative techniques to be combined. The work of Unuma et al[10] suggests that the characteristic style of a motion can potentially be isolated and edited. This could allow the style of a motion captured actor to be altered to subtly change the motion or to make the motion fit more closely with other motions in the animation. As well, that style could perhaps be copied from motion capture data onto simulated motions, allowing, for example, keyframed motions to be used with captured motions in the same animation of a character without jarring differences between the parts of the animation which draw from different sources. Since humans are so skilled at interpreting human motion that it is possible for a particular person to be recognized from motion capture data being performed by a simple skeleton after only a short time, it may be that these subtle characteristics will be necessary to take into consideration.

Motion editing can be thought of as a way to use generated motions as building blocks to create animations. The main operations available with these blocks are blends, overlays, and transitions. Blending two or more motions allows interpolation between them, such as producing a medium kick from a low kick and a high kick, or extrapolation, such as producing an even higher kick. More generally, alteration of a single motion, such as forcing it to interpolate certain constraints, could also be classified as blending.

Overlaying takes two motions and causes the character to do some or all of each, resulting in a more complicated motion; for example, a kick and a martial arts block could be overlaid to create a motion where the character kicks while blocking without sacrificing the characteristics of either.

Finally, transitioning allows these motions to be strung together into arbitrary sequences of arbitrary length, changing a database of motion clips into an actual animation. It is worth noting that two main types of transitions exist; those between disjoint motions which are separated by some non-negative amount of time in the animation, and those between overlapping animations, where the character changes smoothly from executing one motion to executing the other.

These editing primitives offer tremendous flexibility in using motion clips to create animations or new motion clips. In theory, continued application of these primitives should allow a wide range of motions to be generated from a small starting set.

This raises the difficult question of how far motion editing can push the underlying motions and generation techniques while still producing good results. Unfortunately, this question has been essentially ignored to date and is far too large to answer in its entirety within the scope of this research; however, this research will take an important first step by examining how several different motion editing techniques and aspects influence the subproblem of character balance in generated transitions.

3 Background

Motion editing as a whole has received a modest amount of attention. Michael Gleicher has used spacetime constraints, a local optimization method that applies results over the entire motion, to retarget motions to new characters and to apply constraints to those motions[5][6], although specifying high-level constraints in the complex mathematical methods he uses can be challenging. Zoran Popovic has worked with spacetime constraints and energy minimization to force motions to comply with constraints, such as alternative character or environmental parameters[3][4]. Victor Zordan has approached the problem using controllers which track and attempt to produce the original motion, subject to certain requirements that the controllers must fulfill[7]. Lee and Shin use hierarchical B-splines to preserve high-frequency information from the original motion while using spacetime constraints and joint-curve warping to alter the original motions[8]. Unuma et al use fourier analysis to preserve the high-frequency data while isolating and modifying motion parameters such as gait[10].

The problem of transitioning between two motions has received less attention. Rose et al use spacetime constrains and energy minimization to fill in the part of the animation between two disjoint motions in a manner somewhat similar to pose interpolation[9], while Lee and Shin also make some use of this technique. Unuma et al transition between motions like walking and running by interpolating between the motion parameters, gradually changing the walk to a run. Finally, Witkin and Popovic do simple joint angle blending[11], which Zordan also makes some use of.

4 Metrics

Unfortunately, the previous work uses few or no metrics to measure the success of the editing method; energy minimization is almost uniformly the only metric employed beyond human intuition. While human intuition is a very powerful tool for evaluating motion that humans are familiar with, such reliance on it is limiting. Automatic animation generation from motion clips, for example, requires other metrics, simply because a human is not present. Additionally, it may at times be useful to have a more precise or quantitative measure of certain characteristics than human observation can provide.

Additionally, the wide range of potentially useful metrics can tax the capabilities of human observers to consider them all, even should they have the necessary expertise to do so. Physical metrics, such as musculoskeletal limitations or character balance, are perhaps the easiest to see since they directly affect the realism of the motion. Actor metrics are likely harder; these come from the actor's intent for the motion, such as speed, power, snap, accuracy, and so on for a martial arts kick. Finally, style metrics may be subtle enough that explicit human observation and evaluation is difficult. These metrics include motion appropriateness, which takes into account how a human catching a basketball will move differently than a human catching a bowling ball. Even if the basic motions are the same, the greater power requirements of the latter will move the action more into the large trunk muscles, producing subtle changes that will give cues about the weight of the object.

These considerations may be crucial for highly expressive or accurate animations. Teaching animations, for example, should usually be highly realistic so as not to mislead those trying to learn the skill; similarly, a system with sufficiently robust realism and actor metrics may be able to offer suggestions for optimizing those metrics that the human from which the motion was generated can then try to adopt to improve his or her performance.

Additionally, metrics would allow comparison of the results of different motion editing techniques, characterizing the strengths and weaknesses of each approach and the degree to which a method can modify original motions while still producing output of sufficient quality for the application at hand.

5 Approach

The core system this research will use employs signal-processing techniques based on those of Bruderlin and Williams[1]; in short, an expanding Gaussian kernel is progressively applied to obtain lower frequency pyramids, with each band simply being the difference between adjacent pyramids. This is used to decompose the joint signals into frequency bands, blend those bands, and reconstitute the modified signal with data preserved at all frequency levels. Simple numerical techniques are used to obtain the derivative curves from the joint angle data, and accurate bone lengths combined with approximate segment mass figures will be used to obtain the modest amount of physics of the system which will be required.

All transitioning methods will be done within this system, and metric-evaluation modules will be added as needed. In particular, the following methods, metrics, and tests are planned:

- Transitioning methods

    1. Joint angle blend

A naive ease-in/ease-out blending approach works reasonably well for joint angles; both Zordan as well as Witkin and Popovic have used this type of blend before.

2. Joint velocity blend

   Transitioning between overlapping motions makes derivative blend usage a more difficult problem than use of joint position. A naive blend between the velocity curves will, just as with a joint angle blend, result in a new, displaced curve. When velocity is integrated to find the joint angle curves that drive the animation, however, any changes at time t are also propagated to all later times, causing very strange effects.

   With the particular problem domain used in this research, lowering the leg at the end of one kick will cancel some of the velocity of raising the leg at the beginning of the next, leading to non-physical effects like the leg being lowered too far at the end of the kick sequence; the "leg up/leg down" rhythm of kicking is not preserved by the naive approach of directly blending velocity curves and then integrating to determine joint position.

   To address this problem with propagating error, one possible approach is to add in an appropriate integration constant to cause the figure to end in the correct location. Since we have data of kicks where the character returns to almost exactly the original location and kicks in which it does not, this hypothesis can be tested over a range of easy (final location known and same as starting location) and difficult (final location significantly different) cases. Other approaches will be evaluated as the utility of this one becomes apparent.

3. Joint acceleration blend

   Since acceleration must be integrated twice to generate the joint angle curves that drive the animation, all of the error propagation problems which are present in naive velocity-based transitions are much more prevalent in acceleration-based transitions.

- Metrics

  1. Character balance

     Being a physically-based metric, character balance will first require using bone and mass data to approximate the physics of the system. We will use an approach inspired by Winter et al[12] which involves tracking the location of the centre of pressure within the support base defined by the ground contact points, and appropriate consideration of bodily momentum will also be crucial. The metric value will depend on a combination of the centre of pressure location with respect to the support base, the movement of the centre of pressure, and the movement of the body itself if this is not sufficiently captured by the movement of the centre of pressure.

2. Physical intuition
   As mentioned, human intuition about human movement is well-developed and very powerful, particularly for motions with which the observer is very familiar. Subjective human evaluation of the results of the various methods should provide not only an implementation guide, but also a useful comparison to any more quantitative metrics employed. Since this type of evaluation is much more comprehensive than any one quantitative metric, we will take advantage of it for early guidance on how the methods should operate. Additionally, comparisons between transitions in which human intuition and quantitative metrics agree or, especially, disagree, may illuminate interesting areas for future research.

- Tests
  The initial tests employing human intuition will be used to tune the system into a reasonable state from which comparisons between what an observer considers good and what the metrics consider good can be easily done. Additionally, these tests can be done before any metrics have been implemented, allowing the testing scheme to itself be evaluated for suitability and coverage. The results of these initial tests may suggest alterations to the set of metric tests to be run. All metric-based tests will examine the same questions as the corresponding intuition-based tests.

  1. Position vs. acceleration blends; physical intuition metric.
     This test will examine whether an appreciable difference appears between transitions created with simple position blends and those which are presumably closer to the physical situation, having been based on joint acceleration which is closely related to joint torque.

2. Position blends with and without frequency banding; physical intuition metric.
   Does preservation of high-frequency information make a noticeable difference even at the simplest level?
3. Acceleration blends with and without frequency banding; physical intuition metric.
   Due to the doubly-integrated nature of animations resulting from acceleration-based transitioning, the benefits of frequency banding may be more important in such a situation.
4. Blends with and without hierarchical frequency banding; physical intuition metric.
   The system allows each frequency band to be blended with a different length of blend interval, although the centre of the interval is uniform. Flexibility of this sort helps prevent ghosting effects in image blending, and may have similar utility in motion blending.
5. Position vs. acceleration blends; balance metric.
6. Position blends with and without frequency banding; balance metric.
7. Acceleration blends with and without frequency banding; balance metric.
8. Blends with and without hierarchical frequency banding; balance metric.

6 Implementation

The basic system is in place and does frequency banding, joint angle blending, early versions of joint velocity and joint acceleration blending, and import and export to and from Maya. The other functionality remains to be added, as well as removing the remaining incompatibilities between the system and the format of the motion capture data.

7 Timeline

Apr Investigate and present research problem.

May Refine problem; retool current editing system for this problem's requirements.

Jun Continue implementing derivative blends.

1. Velocity with appropriate constant of integration.
2. Acceleration with appropriate function of integration.

Jul Finish derivative blends and system retooling.

Aug Begin implementing balance metric.

1. Back out physics.
2. Determine centre of pressure.
3. Determine support base.
4. Determine appropriate metric based on location of centre of pressure in support base and on body momentum.

Sep Balance metric.

Oct Finish any last issues with derivative blends and balance metric.

Nov Resolve any issues that remain outstanding.

Dec Compare methods.

1. Position vs. acceleration blends; physical intuition metric.
2. Position blends with and without frequency banding; physical intuition metric.
3. Acceleration blends with and without frequency banding; physical intuition metric.
4. Blends with and without hierarchical frequency banding; physical intuition metric.
5. Position vs. acceleration blends; balance metric.
6. Position blends with and without frequency banding; balance metric.
7. Acceleration blends with and without frequency banding; balance metric.

8. Blends with and without hierarchical frequency banding; balance metric.

Jan Continue comparisons.

Feb Finish comparisons; analyze results.

Mar Presentation.

Apr Final analysis and writeup.

8 References

1. *Motion Signal Processing*, Armin Bruderlin and Lance Williams, Siggraph 1995.

2. *Optimal Control of Reaching Movements*, Y. Uno and M. Kawato.

3. *Physically Based Motion Transformation*, Zoran Popovic and Andrew Witkin, Siggraph 1999.

4. *Character Animation by Spacetime Optimization*, Zoran Popovic.

5. *Retargetting Motion to New Characters*, Michael Gleicher, Siggraph 1998.

6. *Motion Editing with Spacetime Constraints*, Michael Gleicher, 1997 Symposium on Interactive 3D Graphics.

7. *Tracking and Modifying Upper-Body Human Motion Data with Dynamic Simulation*, Victor Zordan and Jessica Hodgins.

8. *A Heirarchical Approach to Interactive Motion Editing for Human-like Figures*, Jehee Lee and Sung Yong Shin.

9. *Verbs and Adverbs: Multidimensional Motion Interpolation*, Charles Rose, Michael F. Cohen, and Bobby Bodenheimer, IEEE Computer Graphics and Applications, Vol. 18, No. 5, September/October 1998.

10. *Fourier Principles for Emotion-based Human Figure Animation*, Munetoshi Unuma, Ken Anjyo, and Ryozo Takeuchi, Computer Graphics Proceedings 1995.

11. *Motion Warping*, Andrew Witkin and Zoran Popovic, Siggraph 1995.

12. *Stiffness Control of Balance in Quiet Standing*, D. Winter et al, J. Neurophysiology, 80(3), pp 1211-1221, Sep. 1998.