Motion Editing Techniques for Realistic

Human Motion Generation

Submitted by

Jae Woo Kim

12th December 2000

To
Dr. Jaems K. Hahn

In partial fulfillment of the course
Advanced Topics in Computer Graphics
CSCI 368
Table of Contents

1. Introduction

1.1 Limitation of the Traditional Animation

1.2 Motion Capture for Realistic Animation

1.3 Motion Editing

2. Interactive Motion Editing with Constraints

2.1 Spacetime Constraints

2.2 Retargeting Motion to New Characters

2.3 A Hierarchical Approach to Interactive Motion Editing

2.4 Physically Based Motion Transformation

3. Motion Editing with Global Controls

3.1 Signal Processing Techniques

3.2 Fourier Principles for Emotion-based Human Figure Animation

3.3 Motion Warping and Blending

3.4 Multidimensional Motion Interpolation

3.5 Stylistic Motion Synthesis

4. Discussion and Conclusion

References
1. Introduction

1.1 Limitation of the Traditional Animation

Motion control of articulated figures such as humans has been a challenging task in computer animation. Using traditional key-framing, it is relatively straightforward to define and modify the motion of rigid objects through translational and rotational trajectory curves. However, manipulating and coordinating the limbs of an articulated figure via keyframes or the spline curves they define is a complex task that draws on highly developed human skills. More general, global control of the character of an animated motion would be useful in fine-tuning keyframed sequences.

Much of recent research in motion control of articulated figures has been directed towards reducing the amount of motion specification to simplify the task of the animator. The idea is to build some knowledge about motion and the articulated structure into the system so that it can execute certain aspects of movement autonomously. This has lead to the development of higher level control schemes where the knowledge is frequently specified in terms of rules, and physics-based modeling techniques in which knowledge is embedded in the equations of motion, constraints and possibly an optimization expression. Both approaches often suffer from lack of interactivity: they don’t always produce the motion which the animator had in mind, and complex models have a slow interactive cycle. To increase the expressive power of such models, more control parameters can be introduced. [2][4][7][11]
1.2 Motion Capture for Realistic Animation

An alternative method to obtain movements of articulated figures is motion capture where the motion is captured from live subjects. Systems for real-time 3D motion capture have recently become commercially available. These systems hold promise as a means of producing highly realistic human figure animation with more ease and efficiency than traditional techniques afford. Motion capture can be used to create custom animation, or to create libraries of reusable clip-motion. Clip-motion libraries could facilitate conventional animation, or serve as databases for on-the-fly assembly of animation in interactive systems.

Although a variety of technologies have been developed to fairly reliably measure motion capture data, the computer graphics literature makes scant mention of editing techniques for recorded motion. In the absence of effective editing tools, a recorded movement that is not quite “right” requires the whole data capture process to be repeated. So, it is desirable to develop tools that make it easy to reuse and adapt existing motion data. [2][4]

1.3 Motion Editing

The ability to edit captured motion is vitally important. Custom animation must be tweaked or adjusted to eliminate artifacts, to achieve an accurate spatial and temporal match to the computer generated environment, or to overcome the spatial constraints of motion capture studios. To reuse clip motion we must to be able to freely alter the geometry and the timing. To be useful, editing should be much easier than animation from scratch, and should preserve the quality and naturalness of the original motion.
Much of the recent research in motion control has been devoted to developing various kinds of editing tools to produce a convincing motion from prerecorded motion clips. To reuse motion-captured data, animators often adapt them to a different character, i.e., retargeting a motion from one character to another[5], or to a different environment to compensate for geometric variations[2][11]. Animators also combine two motion clips in such a way that the end of one motion is seamlessly connected to the start of the other.

There also have been the research about how to represent and apply the emotion or the style of the original motion to another [8].

In this paper, we will survey the motion editing techniques developed so far. This paper categorizes the motion technique techniques as follows:

- Interactive motion editing with constraints
- Motion editing with global controls

In section 2, we will review interactive animation tools for constraints-based motion editing and in section 3, we will review an alternative approach is to provide more global animation controls.
2. Interactive Motion Editing with Constraints

Much recent effort has addressed the problem of editing and reuse of existing animation. A common approach is to provide interactive animation tools for motion editing, with the goal of capturing the style of the existing motion, while editing the content. Gleicher[4] provides a low-level interactive motion editing tool that searches for a new motion that meets some new constraints while minimizing the distance to the old motion. A related optimization method is also used to adapt a motion to new characters[5]. Lee et al. provide an interactive multiresolution motion editor for fast, fine-scale control of the motion[6]. Most editing systems produce results that may violate the laws of mechanics; Popovic and Witkin[7] describe a method for editing motion in a reduced-dimensionality space in order to edit motions while maintaining physical validity. Such a method would be a useful complement to the techniques presented here.

2.1 Spacetime Constraints

Gleicher[4] suggested a method for editing a pre-existing motion such that it meets new needs yet preserves as much of the original quality as possible. Their approach enables the user to interactively position characters using direct manipulation. They used spacetime constraints which consider the entire motion simultaneously. These methods enable the user to specify constraints over the whole motion and use a solver to compute the “best” motion that meets these requirements. Like more traditional keyframe and inverse kinematics methods, the user makes adjustments to an animated character with direct manipulation, for example pulling on a
character’s hand to reposition it. But, to achieve these new positions, the animation system makes adjustments that attempt to preserve the original motion. For example, in figure 1, a user pulled the hand at frame 20, the pelvis at frame 24, and the knee at frame 37. At frame 20, the hand constraint is maintained while all the other parts of the body moved to satisfy the constraints.

Like other spacetime constraint methods, their system considers the entire motion in making changes. Unlike the previous spacetime systems, they solve the numerical constraint problems fast enough to provide direct manipulation dragging. To achieve this new style of motion editing, they devised a constraint formulation that is rich enough to be effective, yet simple enough to permit rapid solution and they introduced fast methods for solving these constraint problems.

Using this method, it is very difficult to encode the style of motion mathematically, in other words it is almost impossible to specify meaningful attributes such as “gracefully”, “as if you were scared,” or even “like Charlie Chaplin.” A more serious omission is the
lack of Newton’s laws in their system. The physics constraints have structure that appears to make solution of the optimization problems much more difficult. There is a tradeoff between having constraints and objectives that are better able to compute better edits and having a simple enough problem to solve at interactive rates. They have chosen the latter approach.

2.2 Retargetting Motion to New Characters

Gleicher[5] suggested a technique for retargeting motion that is the problem of adapting an animated motion from one character to another. Their focus is on adapting the motion of one articulated figure to another figure with identical structure but different segment lengths. Their method creates adaptations that preserve desirable qualities of the original motion. They identify specific features of the motion as constraints that must be maintained. A spacetime constraints solver computes an adapted motion that re-establishes these constraints while preserving the frequency characteristics of the original signal.

Constraints are the primary tool used to identify features of the original motion that must be present in the retargeted result. Specification of these constraints typically involves only a small amount of work in comparison with the tasks of creating the characters and motions. Constraints are generally defined once for each motion, and this one set of constraints is used for any retargettings. Since there are typically many possible motions that satisfy the constraints, they use an objective function to select the best choice. A simple objective could be “minimize the amount of noticeable change.”
Figure 2. Example of retargetting motion to new characters

In figure 2, differently sized characters pick up an object. Their positions are determined by the position of the object. The left shows the original actress, the center shows a figure 60% as large and the right shows a figure with extremely short legs and arms and an extremely long body. The cones represent footplant positions.

Their approach makes many sacrifices to achieve practicality. They tell their solver little about the original motion or general motion properties, and their choice of the mathematical problem is heavily influenced by what can be solved efficiently. While their pragmatism pays off in the practicality of the method, they sometimes pay a cost in the quality of the resulting motions. Some of the problems they see are artifacts of the specific simple objective they have chosen and their reliance on simple frequency limits on the adaptations. Other problems occur because they have no guarantees on the many properties they do not explicitly model in their constraints and objective. For instance, their lack of physics constraints can lead to unrealistic situations.
2.3 A Hierarchical Approach to Interactive Motion Editing

Lee et al.[6] suggested a technique for adapting existing motion of a human-like character to have the desired features that are specified by a set of constraints. This problem can be typically formulated as a spacetime constraint problem. Their approach combines a hierarchical curve fitting technique with a new inverse kinematics solver. They employed the multilevel B-spline fitting technique. They also present an efficient inverse kinematics algorithm which is used in conjunction with the fitting technique.

Using the kinematics solver, they can adjust the configuration of an articulated figure to meet the constraints in each frame. Through the fitting technique, the motion displacement of every joint at each constrained frame is interpolated and thus smoothly propagated to frames. They are able to adaptively add motion details to satisfy the constraints within a specified tolerance by adopting a multilevel B-spline representation.

The performance of their system is further enhanced by the new inverse kinematics solver. They devised a closed form-form solution to compute the joint angles of a limb linkage. This analytical method greatly reduces the burden of a numerical optimization to find the solutions for full degrees of freedom of a human-like articulated figure.

Their approach is distinguished from the work of Gleicher who addressed the same problem. He provided a unified approach to fuse both relationships into a very large non-linear optimization problem, which is cumbersome to handle. Instead, decouple the problem into manageable subproblems each of which can be solved very efficiently.
2.4 Physically Based Motion Transformation

Popovic and Witkin[7] introduced a novel algorithm for transforming character animation sequences that preserves essential physical properties of the motion. By using the spacetime constraints dynamics formulation their algorithm maintains realism of the original motion sequence without sacrificing full user control of the editing process.

In contrast to most physically based animation techniques that synthesize motion from scratch, they take the approach of motion transformation as the underlying paradigm for generating computer animations. In doing so, they combine the expressive richness of an input animation sequence with the controllability of spacetime optimization to create a wide range of realistic character animations. The spacetime dynamics formulation also allows editing of intuitive, high-level motion concepts such as the time and placement of footprints, length and mass of various extremities, number of body joints and gravity.

At its core their algorithm uses spacetime optimization because the spacetime formulation maintains the dynamic integrity of motion and provides intuitive motion control. Because such methods have not been shown to be feasible for human motion models, they must also find a way to simplify the character model.

The entire transformation process breaks down to four main stages as follows:

- **Character Simplification.** Create an abstract character model containing the minimal number of degrees of freedom necessary to capture the essence of the input motion.
  Map the input motion onto the simplified model.

- **Spacetime Motion Fitting.** Find the spacetime optimization problem whose solution closely matches the simplified character motion.
• **Spacetime Edit.** Change spacetime motion parameters, introduce new pose constraints, change the character kinematics, objective function, etc.

• **Motion Reconstruction.** Remap the change in motion introduced by the spacetime edit onto the original motion to produce the final animation.

The main shortcoming of their approach is that large portions of the motion fitting algorithm stage are performed manually. They have found the simplification process quite intuitive the simplification is performed only once per input motion sequence, so the effort spent by the motion library creator is amortized over the large number of possible transformed animation sequences. Nevertheless, automating this manual decision-making process would enable on-the-fly construction of a physically based spacetime formulation from an input animation.

3. **Motion Editing with Global Controls**

Researchers tried to edit motion sequences with global controls. Some of them used signal processing techniques or frequency domain techniques to represent and apply the emotion or the style of the original motion. Signal processing systems, such as described by Bruderlin and Williams[2] and Unuma et al.[10], provide frequency-domin controls for editing the style of a motion. Witkin and Popovic[11] blend between existing motions to provide a combination of motion styles. Rose et al.[9] use radial basis functions to interpolate between and extrapolate around a set of aligned and labeled example motions (e.g., happy/sad and young/old walk cycles), then use kinematic solvers to smoothly string together these motions. Similar functionality falls out of our framework. Brand and
Hertzmann[3] suggested style machines which generates stylistic motion by learning motion patterns from a highly varied set of motion capture sequences.

### 3.1 Signal Processing Techniques

Bruderlin and Williams[2] applied techniques from image and signal processing domain to designing, modifying, and adapting animated motion. Using their method, existing motions can be modified and combined interactively at a higher level of abstraction. The techniques represent a pragmatic approach to signal processing by providing analytic solutions at interactive speeds, and lend themselves to higher level control by acting on several or all degrees of freedom of an articulated figure at the same time. They applied the principles of image multiresolution filtering to motion parameters of an articulated figure, motivated by the following intuition: low frequencies contain general, gross motion patterns, whereas high frequencies contain detail, subtleties, and most of the noise.

An application of motion multiresolution filtering is illustrated in figure 3. Displayed like an equalizer in an audio amplifier, this is a kind of graphic equalizer for motion, where the amplitude (gain) of each frequency band can be individually adjusted via a slider before summing all the bands together again to obtain the final motion.

There method can be applied to

- Multitarget motion interpolation: To blend two or more different motions (e.g., happy walk, sad walk, angry walk, etc.)

- Waveshaping: To limit the joint angles for a motion sequence of an articulated figure waving or to map the shape of input motions to a “characteristic” function.
• Motion displacement mapping: To provide a means to change the shape of a signal locally through a displacement map while maintaining continuity and preserving the global shape of the signal.

3.2 Fourier Principles for Emotion-based Human Figure Animation

Unuma et al.[10] described the method for modeling human figure locomotions with emotions. They used Fourier series expansions of experimental data of actual human behaviors to interpolate or extrapolate the human locomotions. For example, the transition from a walk to a run is smoothly and realistically performed by the method. Moreover an individual’s character or mood, appearing during the human behaviors, is also extracted by the method. For example, the method gets “briskness” from the experimental data for a “normal” walk and a “brisk” walk. The “brisk” run is generated by the method, using another Fourier expansion of the measured data of running. The superposition of these human behaviors is shown as an efficient technique for generating
rich variations of human locomotions. In addition, step-length, speed, and hip position during the locomotions are also modeled, and then interactively controlled to get a desired animation.

Based on the Fourier series expansion of the original measured data, a functional model is defined for generating a rich variation of movements, far from the original. In making a human figure animation, however, they must treat not only the emotional aspect but also the kinematic aspect. Therefore the functional model of their method is further extended to provide intuitive parameters for simultaneously controlling emotional and kinematic human locomotions, where the kinematic control, for example, prescribe speed and step-length of the human figure model. In addition real-time and interactive control is performed with the functional model and this consequently provides wider variations of human figure animations than previous approaches.

However, their method is not invertible. This means that the transition from running to walking by the method is unnatural, while the realistic transition from walking to running is made by the method. This problem must be addressed. The very limitation of the superposition technique will be clarified under more explicit formulation.

### 3.3 Motion Warping and Blending

Witkin and Popovic[11] introduced a simple technique for editing captured or keyframed animation based on warping of the motion parameter curves. The animator interactively defines a set of keyframe-like constraints which are used to derive a smooth deformation that preserves the fine structure of the original motion. Motion clips are combined by overlapping and blending of the parameter curves. They showed that whole families of
realistic motions can be derived from a single captured motion sequence using only a few keyframes to specify the motion warp.

Their method is similar to that of conventional keyframing, in that the animator interactively modifies the pose at selected frames. The main contribution of their research is to introduce motion warping as a means of editing captured motion and to demonstrate that even very complex motions such as a human walk or a tennis swing can be radically reshaped using just a few keyframes without losing their realistic appearance. In figure 4, we can see a frame from a captured motion sequence of a tennis forehand shot, and the corresponding frames from two warped sequences. Only a single keyframe at the moment of impact was required to produce the warped sequences.

To create transitions between clips, they perform motion blends using a technique in which the motion to be joined are overlapped, with one or more critical correspondence
points identified. The combined motion is generated by time warping the constituent motions to align the correspondence points, then blending using time-dependent weights. A key advantage of motion warping is that it fits well into the familiar keyframe animation paradigm, allowing a wide range of existing tools, techniques, and skills to be brought to bear. On the other hand, motion warping shares some limitations of standard keyframing, for example the difficulty of enforcing geometric constraints between keys. Constraints techniques applicable to conventional keyframing can be applied to motion warping as well.

A further limitation is that motion warping is a purly geometric technique, not based on any deep understand of the motion’s structure. Consequently, as with analogous image morphing techniques, extreme warps are prone to look distorted and unnatural. A physically based technique might overcome this limitation.

3.4 Multidimensional Motion Interpolation

Rose et al. [9] introduced methods and data structures used to leverage motion sequences of complex linked figures. They developed a technique for interpolating between example motions derived from motion capture or produced through traditional animation tools. These motions can be characterized by emotional expressiveness or control behaviors such as turning or going uphill or downhill. They call such parameterized motions “verbs” and the parameters that control them “adverbs.” Verbs can be combined with other verbs to form a “verb graph” with smooth transitions between them, allowing an animated figure to exhibit a substantial repertoire of expressive behaviors.
A combination of radial basis functions and low order polynomials is used to create the interpolation space between example motions. Inverse kinematic constraints are used to augment the interpolations in order to avoid, for example, the feet slipping on the floor during a support phase of a walk cycle. Once the verbs and verb graph have been constructed, adverbs can be modified in real-time providing interactive or programmatic control over the characters’ actions. This allows the creation of autonomous characters in a virtual environment that exhibit complex and subtle behavior.

3.5 Stylistic Motion Synthesis

Brand and Herzmann[3] approached the problem of stylistic motion synthesis by learning motion patterns from a highly varied set of motion capture sequences. Each sequence may have a distinct choreography, performed in a distinct style. Learning identifies common choreographic elements across sequences, the different styles in which each element is performed, and a small number of stylistic degrees of freedom which span the many variations in the dataset. The learned model can synthesize novel motion data in any interpolation or extrapolation of styles. For example, it can convert novice ballet motions into the more graceful modern dance of an expert. The model can also be driven by video, by script, or even by noise to generate new choreography and synthesize virtual motion-capture in many styles.

For the purpose, they introduced the style machine – a statistical model that can generate new motion sequences in a broad range of styles, just by adjusting a small number of stylistic knobs (parameters). They seek a model of human motion from which they can
generate novel choreography in a variety of styles. Rather than attempt to engineer such a model, they attempted to learn it by using unsupervised learning technique. They used the Hidden Markov Model (HMM) to learn the choreography and the style of motion data. A HMM is a probabilistic finite-state machine consisting of a set of discrete states, state-to-state transition probabilities, and state-to-signal emission probabilities to learn. They added to the HMM a multidimensional style variable $v$ that can be used to vary its parameters, and call the result a stylistic HMM (SHMM). Using SHMM, they separated structure, style and accidental properties in a dataset by minimizing entropies in the SHMM and finally they can generate new choreographies with various styles.
4. Discussion and Conclusion

Much recent effort has addressed the problem of editing and reuse of existing animation. In this paper, we surveyed some of recent motion editing researches. We divided the approaches into two categories, one is interactive motion editing with constraints and the other is motion editing with global controls.

The researches of interactive motion editing with constraints provide interactive animation tools for motion editing, with the goal of keep the characteristics of the original motion, while editing the content. They provides a low-level interactive motion editing tool that searches for a new motion that meets some new constraints while minimizing the distance to the old motion.

An alternative approach, researches on the motion editing with global control is to provide more global animation controls. Signal processing systems and frequency domain techniques are used in editing the characteristics or styles of a motion. They are also used to blend between existing motions to provide a combination of motion styles. There was a research on a stylistic motion synthesis by learning motion patterns from a highly varied set of motion capture sequences. They developed Stylistic Hidden Markov to represent the structure and the style of the motion capture data.

One of the main purpose of motion control and motion editing researches is to reduce the specifications of the motion which the user has to give, while not losing the control of the motion. There exists a trade-off between the reduction of the specification and the preservation of the control of motion. In order to achieve this purpose the system must have knowledge about the motion and a kind of intelligence which can replace some of
the user’s role. The artificial intelligence approaches, such as style machines, must be useful to achieve this purpose.
References


