Chapter 7

Discussion

The problem of synthesizing quality human motion is a difficult one not only because of the intricacies of the human body but also because humans are good at being able to recognize when a motion is flawed. For interactive applications, the problem becomes even harder; motions must necessarily be generated under much tighter processing and storage requirements while at the same time accurately meeting the ever-changing requests of a user. Example-based methods for human motion synthesis have aided the goal of producing natural looking, accurate motions; these techniques use nuanced example motions captured in a motion capture studio to synthesize new motions that are as nuanced as the originals. However, most of these example-based synthesis methods are not used in practice due to two drawbacks:

1. the exponentially large number of example motions needed to control many different motion parameters simultaneously

2. the processing time and latency necessary for synthesizing a stream of controllable motion

This dissertation has addressed both of these problems by decreasing the number of example motions needed for motion clip synthesis through parameter decoupling and by increasing synthesis and response time through motion clip organization. Taking a practical approach, I focused on developing methods for decoupling commonly used motion parameters and automatically constructing highly structured control mechanisms for realtime character control. Specifically, I made the following technical contributions:

1. **Runtime Method for Splicing Upper-Body Actions with Locomotion.** Chapter 4 introduced a new method for splicing the upper body action of one motion with the lower

body locomotion of another. Unlike prior techniques for performing upper-body, lower-body splicing, my method explicitly considers the natural correlations between the upper and lower body. My approach of using temporal and spatial alignment to retain correlations produces high-quality results that exhibit physical as well as stylistic characteristics seen in the original two example motions. And because the method decouples motion parameters, it considerably reduces the storage requirements for interactive applications that need many different action/locomotion combinations at runtime. Since upper-body/lower-body splicing is already commonplace in the video game industry using naïve methods, my technique has potential to influence real world motion synthesis in interactive applications.

2. **Runtime Method for Adjusting the Gaze Direction of a Character.** In Chapter 5, I presented another decoupling method. This method decouples gaze direction from overall body motion. By basing my method on studies in the biological and psychological sciences, I have been able to produce a simple, efficient model for human gaze that uses little storage space and can be easily applied to any base motion at runtime. The ability to control gaze at runtime in a realistic way not only has the potential to increase the feeling of connectedness between a character and its environment, but it also has the potential for increasing character responsiveness through use of subtle gaze cues.

3. **Method for Automated Authoring of Parametric Motion Graphs.** My method for constructing a parametric motion graph as presented in Chapter 6 is highly-automated. The technique tackles the difficult problem of defining good transitions between entire spaces of motions using sampling. Without using an automated graph construction method, an author would be hard pressed to manually define these transitions without carefully hand-tweaking the motion data itself.

4. **Runtime Method for Using Parametric Motion Graphs.** Chapter 6 also shows how the structure of parametric motion graphs makes them a natural mechanism for controlling interactive characters in realtime. The synthesized motion streams exhibit quality transitions without pauses due to processing. In fact, processing times are extraordinarily quick, even

using my unoptimized, prototype system. This ability to synthesize controllable motion streams in realtime for interactive characters makes it possible for even low-budget applications to explore higher-quality methods for motion synthesis.

With these specific technical contributions, I have shown that example-based motion synthesis methods can be used to effectively synthesize quality human motion that accurately meets user requests in an interactive application.

While the methods presented in this dissertation either tackle the problem of motion clip synthesis through decoupling or focus on the problem of motion stream synthesis using highly structured control mechanisms, these methods could be combined to gain the advantages of both types of techniques. One possibility with great potential is to use a combination of motion clip synthesis methods within each node of a parametric motion graph (Chapter 6) in order to decouple motion parameters during motion clip generation. For instance, one node in a graph might represent all motions where a person walks at a particular curvature and looks in a particular direction. Internally, this node could use blending-based parametric synthesis to synthesize a base walking motion and then use the technique presented in Chapter 5 to layer the gaze direction of the character onto the base motion. The algorithms presented in this dissertation would not change even though the clips produced by each node of the parametric motion graph might be synthesized using a layering of decoupled methods; any smooth parametric motion space can be sampled no matter how the synthesized motions are produced.

To develop the algorithms presented in this dissertation, I built a custom testbed where motion clips and motion steams could be synthesized both online and offline. My system allows the upper-body action of one motion to be spliced onto the lower-body locomotion of another, allows a parametric gaze map to be applied to any base motion, and supports the interactive control of virtual human characters using parametric motion graphs. Because of the dependence on example motions, the effective use of these techniques in real world situations would not only require more optimized implementations of these algorithms but also tools for analyzing and browsing a motion database. Additional artist-oriented tools would also need to be developed to allow the authoring and tweaking of parametric gaze maps (Chapter 5) and parametric motion graphs (Chapter 6).

The rest of this chapter describes the applications for the methods I have introduced in this dissertation, the limitations of the overall approach of my work, and some potential directions for future research related to human motion synthesis for interactive applications.

7.1 Applications

The main focus of my work is on synthesizing human motion for interactive applications. The industry that often drives technology development in this realm is the video game industry. And the needs of the video game industry have strongly inspired the algorithms presented in this dissertation. My methods and insights have the potential to improve the quality and accuracy of human motions in a video game. Video games need the ability to decouple upper-body action from lower-body locomotion and currently fulfill this need using a simple naïve DOF replacement algorithm. My technique could be used as a direct replacement for these methods in order to improve splices in next-generation video games. And now gaze control is becoming a part of video games, again using simple, yet efficient methods that sacrifice motion quality. However, my gaze control model could be applied at runtime in order to perform more realistic looking gaze adjustments. Finally, the idea of combining blending-based parametric synthesis and synthesis-by-concatenation using sampling can increase the accuracy and quality of the motion streams synthesized in modern video games and greatly reduce the amount of time it takes to author controllable human characters.

My work also applies well to the area of online or offline crowd simulation. Crowd simulation is also restricted to using efficient methods for motion synthesis because of the sheer number of motions that must be synthesized. My decoupling methods can be used to efficiently synthesize a large variation of different motions using the same small set of example motions. And parametric motion graphs can allow both high-quality random motion synthesis (see Section 6.2.2.1) as well as motion synthesis that is guided by high-level goals (see Section 6.2.2.2), like those that an agent in a crowd might be required to meet. Using my techniques, a crowd might exhibit greatly variability as well as more accurate interactions with the environment.

Finally, my work applies well to synthesizing motions for training simulations that can aid in preparing workers for their jobs or teaching students about new subjects. The ability to control gaze in an interactive application could lend a greater sense of believability to the actors in a simulation, making it easier to connect the lessons learned with the real world. Furthermore, as with video games, the ability to divide upper-body action from locomotion can increase flexibility in the motions that can be mixed-and-matched at runtime. Finally, because of the ease with which controllable characters can be produced using parametric motion graphs, even low-budget training simulations can begin to experiment with how motion quality affects the transference of the lessons being taught within the simulation. The impact of my work on training simulations could be significant in increasing the realism of these simulations while reducing the time and effort needed to author believable human characters.

7.2 Limitations

While the work presented in this dissertation meets my goal of increasing the utility of example-based motion generation methods for interactive applications by supplying insight into techniques that provide efficiency, low latency, high accuracy, quality, and automated authoring, there are still limitations. This section describes some of the most important limitations to consider when applying my approach to real applications.

7.2.1 Data

The effectiveness of any example-based human motion synthesis technique is necessarily tied to the quality and consistency of the motion database. It is important that capture motions are of high-quality as example-based techniques make this assumption. Furthermore, example motions must be captured in such a way as not to exhibit motion variations that are not intended. For instance, it is easy to capture 30 example motions of a person walking at various curvatures; it is much harder to ensure that the actor does not begin to exhibit subtle motion characteristics associated with fatigue towards the end of the motion capture shoot. It is usually not desirable to produce an interactive character that appears “tired” whenever he curves sharply to the right, for example.

All of the methods presented in this dissertation assume that the motions being blended or spliced are compatible with one another. The motions cannot come from two different actors with drastically different proportions or be mapped to different skeletal hierarchies. Some of the problems associated with data set compatibility can be overcome using motion retargeting methods, like those in [Gle98, SLGS01], but these types of techniques cannot address more stylistic compatibility issues; for instance, it might be possible to splice the upper body of a drunk person walking with loose joint control and the lower body of another person walking with a rigid, straight-legged waddle, but these motions are not necessarily compatible with one another.

And, as with many other example-based motion synthesis methods, the techniques presented in this dissertation depend on having enough example motions to adequately cover all possible motions that might be synthesized. My techniques are incapable of producing motion that is considerably different from the motions in the database. This includes transitions between motions. My reliance on linear blend transitioning rather than a more complex method means that I limit the motions that can be appended together to those that already appear similar. However, the problem of generating a quality transition between any two motion clips is an open problem that may be intractable.

7.2.2 Generalization

With the methods presented in this dissertation, I have shown the utility of decoupling and structure by developing methods that directly address some specific problems in motion synthesis for interactive character control. While these methods illustrate the advantages of my general approach, the methods themselves do not directly generalize. In particular, the methods for decoupling motion parameters during motion synthesis do not provide a method for decoupling any general parameter from any other.

Furthermore, all of the methods presented are primarily designed for the control of full body motion. Techniques for better control of facial features or hands, either independent of or in conjunction with full body motion, are not addressed.

7.2.3 User Input

While a goal of my work is to automate the process of authoring controllable human characters for interactive applications, my methods do depend on goals or hints supplied by a user. For instance, while my technique for building a parametric motion graph automates the process of constructing transition mappings from one parametric motion space to another, it still depends on user input to design the global shape of the graph (i.e., which nodes connect to which other nodes). The process of building an edge between parametric motion spaces also requires that a user specify a small number of tunable parameters that effectively define the tradeoff between motion quality and accuracy. Similarly, the automated methods for processing example motions of a person adjusting their gaze require that a user specify values for several tunable parameters; these parameters allow the algorithms to be adjusted for different motion capture environments by depending on the knowledge of the user.

Additionally, all of the methods presented depend on a user to supply the high-level information about what motion to synthesize. In truth, this ability to supply a high-level goal is the point of interactive motion synthesis, but in some cases, such as for gaze control, there may be automated methods that could help identify a plausible goal given the configuration of the environment and the current motion of the character.

7.3 Future Work

Considering the current trend towards using video games and training simulations to supplement classroom learning and on-the-job training, the need for quality, interactive human motion synthesis will only grow. And yet, there are still many open problems in the area of human motion synthesis:

Better Understanding of Human Perception: One area of human motion research that is only just starting to be explored is the area of human perception. When synthesizing human motion, techniques often introduce artifacts, such as foot-sliding, knee-popping, and bone-stretching. The goal of many researchers is to limit the introduced artifacts to ones that are not offensive. Yet few papers like [HRvdP04], which studies the perceptual effects of bone-stretching, have appeared. More work is needed to determine which artifacts humans can perceive and to what extent. Using careful experimental design, it may be possible to study the perception of human motion to help inform motion synthesis techniques.

Better Motion Quality Metrics: One important characteristic of a virtual character motion is its quality. In this dissertation, I have sought to provide only high quality motion synthesis methods. However, as stated throughout this document, quality is a partially subjective characteristic and there is no standard way to measure motion quality directly. My work focuses on ensuring only some quality characteristics, such as artifact appearance and continuity. Yet, in the future, better metrics and processes for measuring the quality of a human motion could greatly impact not only my own work but also the work of others in the motion synthesis community.

Better Physical Models: One downside to example-based human motion synthesis is that the motions generated using these techniques might break some of the laws of physics. Some researchers have begun studying physically based models for motion generation, but these models fail to recognize the importance of characteristics, such as personality, that a purely physical approach cannot capture. Future work on ways to combine example-based approaches with physically based approaches in order to create realistic motions with desirable physical as well as stylistic properties could have great potential to the interactive application community.

Motion Browsing: The proliferation of example-based motion synthesis techniques means that the number of motions in a typical motion database is bound to grow. Large numbers of motions not only present a problem for motion synthesis techniques, but large number of motions also make it difficult to browse a motion database. The community is in need of better ways to present the data in a motion collection so that a human can browse it rapidly. Solutions to the problem of motion browsing might include better methods for automatically identifying motion characteristics as well as better methods for “summarizing” a motion in a single thumbnail image.

I hope that my work on efficiently generating accurate, quality motion using an example-based approach will increase the quality and accuracy of human motion in real world applications and inspire others to develop practical yet quality-oriented algorithms for motion synthesis in interactive applications.