Chapter 1

Animation from Observation — DRAFT

Excerpted from Motion Capture and Motion Editing: From Observation to Animation, *by Michael Gleicher. This book will be published sometime after its finished by AK Peters. I have no idea when I will have it finished.*

This draft is current as of January 24, 2002.

This book is about producing computer animation from observations of the real world. Or to be more specific, it is about animating characters in computer animation based on observations made by sensing moving objects in the real world. Or, to be more true to our focus, it is about obtaining the motion for animated computer characters through a process that includes using sensors to record the movement of objects in the real world. Or really, it is about how to create the movements of computer characters by re-using the movement created by some other character, either by observing it in the real world, or from some library of movements. Or, ...

The circumlocutions of the previous paragraph indicate that its hard to concisely and precisely define what this book is about. I was specifically trying to avoid using the common term *motion capture,* that most (including the publisher, my co-authors, most readers, . . .) were probably expecting. Part of this stems from a desire to avoid using a technical term (or worse, a semi-technical term) before it is properly defined. In the case of the term motion capture, this is especially problematic since the term tends to mean different things to different people. More importantly, this book really isn't just about motion capture by just about anyone's use of the term. In fact, one of our goals in this book is to convey our notion that motion capture by itself is not the same as the entire process required to produce "animation" using motion capture.

Our first task in this book is to define what it is that this book is about, what we mean by motion capture, and how the other pieces of the animation process must be fit together with motion capture in order for successful results to be created. This task is the primary goal of this chapter. Making a quick survey of the entire process not only allows us to make our point that motion capture is only one step of a larger process, but it also allows us to introduce terminology and provide quick tour of what will come later in the book.

This last aspect is particularly important. The set of talents involved in the computer animation process is quite diverse, ranging from highly technical and mathematical to artistic. While most individuals will probably focus on certain aspects of the process, an understanding of the entire process is key to success, whether you're a software developer inventing new algorithms, an artist trying to craft an animation, or a producer trying to budget a project. It is critical that everyone has a common understanding and terminology. Therefore, we have tried to create a book that will service this diverse set of needs. This means that not all of the book will be relevant to all people. I do recommend that everyone start with this introductory chapter, as it spells out our philosophy, goals, and terminology, and can help guide you to which other parts of the book are relevant to your needs.

This chapter begins by making an explicit statement of our philosophy on the subject. We will then proceed to define what the subject is by examining what is meant by computer animation, and exploring the ways in which the motion for computer animation can be created. This will allow us to examine how motion capture compares with alternative methods, and to see how it fits in with the larger puzzle of animation production.

1.1 Caveats

As you might have already surmised, we[1] have an agenda in this book. There is a definite philosophy that we would like to convey. Rather than be subversive, we prefer to make this agenda explicit up front, despite the fact that we will not justify it, or even define the terms, until later in the book:

1. Motion capture can work to produce computer animation.
2. Motion capture is just part of the process.
3. Motion capture is not the right tool for everything, or every project.
4. Planning is at least as important as any other phase of the project.
5. Understanding how motion capture fits into the bigger picture, both in terms of the overall animation process as well as with other tools.
6. That there is no competition between methods: to be truly successful, motion capture must be well integrated with other approaches such as keyframing.
7. The core issues of motion capture stem from the need to build and use representations that are abstractions of motion. These issues transcend any particular technology.
8. Motion editing, the act of changing a motion to fit ones' needs, is as much a part of the production as capturing the motion in the first place.

[1]The "we" here is different than most of the "we" in this book. For this collaboration, I have explicitly chosen to work with people who share my views.

Generally, we will try to use the first person singular to say something that is an opinion of the author (e.g. I believe), and use the first person plural to refer to the author in conjunction with the community. I take responsibility for my opinions, but what I achieve, I achieve only by building on the work of others.

The history of using motion capture for animation is full of successes and failures. There is certainly no single recipe for failure, and (I believe) that there is no single recipe for success either. This book is put together on the belief that there are a few common themes in any successful strategy:

- Planning is important.
- Solve problems sooner, rather than later.
- Remember that it is about performance (as in an actor's performance, not as in speed of computers).
- Have realistic expectations.
- Begin with the ends in mind, and keep the end goals in mind all along.

We will try to avoid the debate about technology choices. If for no other reason than the technology is a rapidly moving target, so the hardware in three years will probably have different tradeoffs than what we use now. More importantly, I believe that success does not come from having the latest and greatest technology: it comes from the well executed use of available technology. Also, all technologies have related problems. The details of these issues change, but the general strategies remain the same.

1.2 Computer Animation

Explaining the title of this book is problematic since we wanted to avoid defining motion capture. However, even the term animation is now problematic.

Literally, the term animation means "to bring to life." Until recently, this traditional notion of animation really applied to computer animation as well. The creation of a sequence of images with enough quality to "bring [something] to life" needed to be done in an off-line fashion, where the movements, and typically even the images, were created in advance. Computer animation meant something similar to traditional animation, a "film" where each image was "drawn." The difference was that in computer animation, the computer that did the drawing.

More recently, this meaning of computer animation has been stretched in several ways, many of which are particularly relevant to our subject:

1. The ability to create high-quality imagery quickly means that animation is now practical in interactive applications.

   As computing hardware improves, the quality of the imagery possible in interactive setting also improves. The current state of the art for character animation in interactive applications is improving rapidly. While the characters in computer games and virtual environments may not yet be at the standard of what we see in feature films, the distinction is rapidly decreasing. Current generation games already feature human characters that move in realistic ways.

2. New display venues, such as internet broadcasting, make new formats practical, or even common. This means that the range of "animation" projects includes low budget, short projects, as well as large-scale feature film projects. These new venues often have different resource availabilities, in terms of time, budget, and talent, than more traditional venues.

3. Increasingly, animation is merged with "live action" footage, to create animated characters that interact with a real world or to modify the real world. This realm of "visual effects" is becoming increasingly important to the film industry, where computer imagery is being used to create impossible (or at least impractical) worlds that appear real.

   Visual effects place an interesting set of demands on the motion of characters. Often, real action must be mimicked or a real environment must be accommodated.

4. New applications in medicine and training place new demands on realism. If the goal of an animation is to realistically convey movement to help someone in a real world setting, there are different demands on this motion than in an entertainment setting.

We will explore these applications of computer animation in general, and of animation from observation specifically, in Chapter **??**. Our point here is to say that we need a very general notion of what animation is.

1.3 The Challenges of Animation

While the definition of computer animation from the last section is non-specific, there is a common thread: animation always involves something that is animated, that is, brought to life through movement. This movement is what provides the unique challenge of animation. While animation is not unique in being a visual media, or even in being a spatio-temporal art form, it is the only art form in which movement is specified and controlled as an independent entity.

Animation is a uniquely expressive art form: it provides the creator with control over both the appearance and the movement of characters and objects. This gives artists tremendous freedom, which when well used, can create works with tremendous impact. This freedom, however, also becomes a curse: while everything can be controlled, everything must be controlled. Control over the movement of objects is a difficult task, requiring skill and labor.

The need to specify movement creates unique challenges for animation. Because specifying motion is so uncommon, few people have developed the skill to do it, there is much less understanding of it, and there is even less of a vocabulary to describe it. While most people might pick up a crayon and start drawing at a very young age, few ever specify motion until they become animators. And while our language provides precise, detailed mechanisms for specifying geometry and form, commonly used language lacks equivalent terms for motion.

What this means is that motion is very difficult to describe in a concise way. While geometry leads to descriptions that are easily recreated (e.g. a 3 inch green square), motion rarely can be described in such detail. Typically motion must be described in terms of results, metaphors, and moods. We talk about a happy or sad motion, a motion that lifts a box to the shelf, a regal march, etc.

What makes the problem of describing motion even more difficult is that motion is incredibly rich, expressive, complex, and communicative. Our ability to see intent and mood - or even to distinguish performers - by motion is well documented. Psychology studies using moving light displays (cite) show that from motion alone we are capable of recognizing activities and people. We aren't able to say what about a motion makes it a walk or strut, happy or said, Fred Astaire or Michael Jordan or our best friend, but we know it when we see it.

Because creating motion is so difficult, there have emerged a few basic strategies for obtaining it for computer animation. While traditionally, computer animation has looked at the first three of these (see, for example Hodgins [**?**, **?**]), the fourth is emerging as a method in its own right:

1. We can rely on talented animators to translate the high-level desires for the motion into a convenient, low-level representation for the computer. This process most closely mimics traditional animation.

2. We can attempt to define mathematical models that recreate the motions, typically, by trying to describe motions by physical principles or algorithmic procedures.

3. We can create the movement by demonstration. That is, we can record the movement from some real-world source, allowing us to provide an example to be recreated.

4. We can create new movements by finding previously created examples that are as close as possible to what we need and adapting them to meet our desires. While this method relies on one of the other methods to create the initial motion, as the collection of examples to draw from grows large enough, adaptation emerges as a creation method in its own right.

The first method is commonly called key-framing, because the most common tools for animators to specify motion requires them to specify key frames that are interpolated. The second category includes simulation and procedural synthesis. The third category is commonly referred to as motion capture. The final category we typically refer to as motion editing.

The methods are never truly distinct, nor independent. An animator often will look at an example or demonstration of a movement in creating it, procedural methods often use hand-generated controls, editing often requires manual adjustment, etc. While pigeon-holing techniques into concise categories helps us understand the methods, and provide us with focus in studying them, it also has the potential to invite separation, comparison, or even competition, none of which are useful in getting the job done.

In practice, all of these methods work together to form the toolbox with which we create animation. While our focus in this book is on motion capture and motion editing, one of our central tenets is that these methods are only an addition to the toolbox. All of the tools have their strengths and weaknesses, and all have the most power when integrated with all of the others. The best way to achieve the best possible results is to have fluency with all of the tools in the tool box and to be able to select among them, combine them, and integrate them as needed to achieve desired results.

Within the animation community, there is historically a tension between “animators” (those artists who specialize in creating motion by hand) and motion capture technicians/users [Cameron]. This tension comes from many factors, some of them real and some of them perceived. The two main sources of this tension are:

1. unrealistic expectations about what motion capture can do (that it can automatically produce motion that displaces animators);
2. motion capture technology development has not considered the use of the data, leaving animators with data that is difficult to deal with.

This tension is a silly and unfortunate piece of history.

1.4 Motion Capture

We must now take on the challenge of defining what we mean by the phrases *motion capture* and *animation from observation*. While making such definitions is an academic exercise, these difficulties are in many ways instructive.

The most obvious way to define motion capture would be as a mechanism for recording the motion of something that moves in the world. By this definition, videotaping would qualify. Pointing a video camera at a person “captures” their motion. We can play it back and see what they did. For some reason, this is not what we commonly mean by motion capture.

What truly distinguishes motion capture is not only that it records the movements, but that it abstracts away geometry in order to provide a manipulatable representation. The distinction (for me at least) is that motion capture creates a representation that distills the motion from the appearance; that it encodes the motion in a form that is suitable for the kinds of processing or analysis that we need to perform. This definition of motion capture is dependent on what we are going to do with the result.

Motion capture is the recording of a manipulable representation of a motion from sensing that motion. This is still quite broad: we have specified neither what kinds of representations, nor what kinds of sensors. Most commonly, motion capture is used to mean the creation of moving 3D geometry based on the output of sensors such as magnetic trackers or video analysis.

The result of motion capture is not animation! It is some data that represents what happened in the real world over a period of time. We might use this data to create animation, or we might use this data to perform some analysis.

Motion capture is different from the process of creating animation from observations. For one, motion capture may be done for a variety of reasons besides animation, such as biomedical analysis, surveillance, sports performance analysis, or as an input mechanism for human-computer interaction. Each of these tasks has similarities and differences with the problems of creating animation. At the first stage of each, there is a need to create the observations that are then interpreted, e.g. capture the motions. Many of the methods used in animation have their roots in the bio-mechanical or medical domains. A discussion of such applications is beyond the scope of this book.

Second, motion capture is just part of the process of producing animation using motion capture. Over the course of the book, we will follow our way through this complete process. A briefer survey is provided later in Section 1.9.

1.5 Animation from Observation

Since the earliest days of the art form [**?**], animators have observed the movement of real creatures in order to create animated motion. Sometimes this simply takes the form of an artist carefully observing nature for inspiration. Another process is to transfer the movement from a recording of the movement to the animated objects. The earliest mechanism for doing this was the Rotoscope, a device that projected frames of film onto the animator's workspace, providing the animator with a guide for their drawings.

Computer animation brings the potential for automating the process of creating animated motion from observations of real moving objects. Optical, mechanical, or magnetic sensors record the movements that can then be transferred to animated characters. This process is commonly referred to as motion capture, although the act of ”capturing the motion” is only one aspect of creating animation from observations of real motion.

Capturing the motion is only part of the problem of using this data to create animation. Commonly, the term motion capture is used to describe the whole process. This has the problem that it neglects other aspects of the task, and sets up some unreasonable expectations about how much work needs to be done to move from the sensor data to animation.

Motion capture for animation implies that we will somehow be changing something about what we have recorded-if we did not intend to change something, we could have simply replayed a video. Almost always, we will at least change the character to which the motion is applied from a real person to some graphical model. By definition, to animate means to bring to life, so technically it is the act of making a lifeless object (a graphics model) move that makes what we're doing animation.

While we have said this before (and will say it again), motion capture is only one part of the process for creating animation from observation.

All that said, I will often be guilty of abusing the term “motion capture” in this book, in this very same way.

1.6 Why is this hard?

The fact that you've read this far implies that you already appreciate that animation from observation is challenging. However, to better motivate the discussions that follow, it is useful to consider why it is so challenging, and why it requires a book's worth of information (or more) to do it well.

The challenges of motion capture may not be so obvious. At a philosophical level, most of us can watch another person make some movements and understand what is happening. Vendors of hardware for helping computers accomplish similar tasks are continually advertising their new system as making the process easier, almost implying that the most difficult part of using motion capture is financing their system. Trade show exhibits show some energetic dancer (almost always someone who looks good in tight lycra) having their movements recorded and translated into wonderful animation without complication. Motion capture service bureaus tell us that they have the equipment and expertise required to really make their products useful to anyone. From such marketing, one might believe that the only problem with motion capture for animation is unemployment of traditional animators!

The evidence is clearly to the contrary. We still hear horror stories of motion capture projects gone wrong. Motion capture hardware developers are continually refining their products and better educating their customers as to their most effective use. Capture services still need to educate their customers [2] and are still honing their crafts. Tool developers are continually refining their products.

Given the investment in the development of motion capture technology, the sophisticated tools available, and the extensive experience of motion capture practitioners, shouldn't we expect animation from observation to be a solved problem?

Probably not. Motion capture and animation from observation are fundamentally hard problems that are unlikely to be "solved" completely any time soon. The fundamental difficulties fall into a few major categories:

1. Human motion is incredibly rich, complex, detailed, nuanced and unpredictable. To make matters worse, the motions that we most want to capture tend to be those "exceptional" performances that are not run-of-the mill and predictable;

2. Human motion is difficult to sense. While technology is increasingly addressing this issue, however, fundamental challenges may preclude an "ultimate" technology that is non-invasive and yet not hampered by challenges such as obscurances and variations of scale.

3. For animation, we typically do not need all of the detail of human motion, rather we need to fit the observations to a simpler model that corresponds to the character that we plan to animate.

4. What is actually performed is rarely exactly what we want. At best, what is captured is an accurate representation of what a particular person may have performed - not necessarily what our target character would have done. Effectively we must find ways to transfer the "essence" of the performers motion to the new setting required by the animation.

5. Getting the desired performance is difficult, yet it is crucial to our success at motion capture (since the performance is what is captured). Given that the tools are ineffective for changing the essence of motion, it is critical that the original performance has whatever "essence" that we desire since it will be extremely difficult to add it in later.

We will explore these challenges in more detail in Sections **??** through 1.7.

Before making the situation sound impossible, we want to stress that these issues *can* be addressed with a combination of thought, planning, technology, and hard work. The key elements of this are:

[2] A production coordinator at a large studio reacted to the proposal for this book by asking if they could require all of their customers to read it first.

1. only choose motion capture where it is appropriate. The developers of motion capture will not be hurt if you choose to keyframe a motion that is better created that way. They will be hurt if your project fails because you have tried to use motion capture where it is inappropriate.

2. be realistic in your expectations of what can be achieved;

3. plan motions such that they are as easy as possible to sense and to post-process later;

4. make sure that the performance is as close to the desired results as practical, or differs in ways that the technology can address;

5. solve problems sooner rather than later;

6. acknowledge that processing to adapt resulting data is a vital part of the application of motion capture data.

Hopefully, it is obvious from this list that planning and processing are two key elements of creating animation from observation and are key things to learn about if you want to address the difficult issues in applying motion capture methods to the production of animation.

Before exploring how we will address the fundamental challenges of using motion capture to create animation from observation.

1.6.1 Non-Fundamental Challenges

Some issues in creating animation from observation are fundamental. That is, they stem from the very nature of the problem. Other issues are more a statement about the particular state of the art or technologies chosen.

We do not mean to belittle these "non-fundamental" challenges. Such things are just as likely to make a particular project a success or failure. In fact, being able to address these challenges is quite critical, quite simply because they can often be addressed. However, it is more difficult to discuss them in a general way.

One large source of non-fundamental challenges is the particular tracking technology used. Tracking human motion *is* a fundamentally hard problem, as we will explore further in Section 1.6.3. To address this challenge, we typically seek engineering solutions, that is, we make comprimises to make the problem easy enough that we can solve it through good engineering. The current capture technologies, as well as some future prospects will be discussed in Section **??**.

For example, computer vision and video processing technologies are still a long way off from being able to look at "normal" video of a person wearing normal clothing and determine their movement to the precision needed for animation. To track visually, current systems use markers that are easy for the tracking system to see (typically using special cameras), as will be discussed in Section **??**. At any given time, the state of the art with these systems has particular limits: the cameras are only so sensitive, the markers need to be at least a certain size, the number of markers that can be tracked is bounded, the accuracy that markers can be resolved to is limited, the amount of range that the performer can be seen in is bounded, etc. Next year, someone will produce a better camera or more clever processing algorithm and some of these "limits" will change.

This every changing technology makes it hard to be specific about non-fundamental limitations. These limitations are very real: if the current state of the art precludes what you are trying to do, you cannot do it (unless you wait for technology to catch up). However, what these limits are will undoubtedly change by the time you read this[3].

1.6.2 Challenge: The Richness of Human Motion

Part of the challenge of motion capture comes from the fact that what we want to capture is complicated. Motion is complex, subtle, unpredictable, highly varied, and expressive. To make matters worse, people are remarkably good at looking at it. The psychology literature is full of examples that show that people can make amazing amounts of understanding of what is happening from a small amount of motion information. For example, from a very shadowy image, or the movement of a few lights tied to a performer, we can identify what they are doing, and sometimes even who it is.

This complexity is, of course, the reason that we need to do this in the first place. If motion were simple, we could probably find ways to synthesize it computationally, or it wouldn't be so challenging to create by hand.

To make matters worse, the reason we need to do motion capture is often that it is precisely these subtleties and unpredictablenesses that we want to obtain. If we just wanted "walking," we could use some generic motion and not need to worry about getting high-quality motion capture data. However, what we often desire is a particular person's walk, or a particular style of walk, or a particular mood.

To see why this is difficult, consider specifying these particular properties. We want Joe's walk, or a regal yet happy walk. What makes a walk have these properties? While we might be able to identify some details that are part of these properties, in general, it is hard to pinpoint most. We know it when we see it, but we cannot describe it mathematically.

1.6.3 Challenge: Sensing Human Motion

The fact that people are so good at seeing motion makes it ironic that machines are so bad at it. I believe that there are a few general reasons for this apparrent paradox:

1. People see the overall big picture in the motion, not the details. While a person can say "that is a person walking," they cannot say "there arm is 27.3 degrees," or "their toe is 2.5cm off the ground," or even "the are moving northwest at 1.6 meters per second."

2. People don't actually see the motion: they figure out what is happening by piecing together details that they do observe. For example, you can identify a woman wearing a long dress walking despite not being able to see her legs. We clearly use large amounts of experience and knowledge in determining what we see.

[3] Especially given how slowly I am writing!

3. Biological vision systems are amazingly highly evolved. While the obvious examples of this are the amazing feats that can be accomplished by animals, such as birds of prey seeing their targets at amazing distances, even more mundane human vision is exceptional. For example, the dynamic range that our eyes are capable of is truly astounding.

Computer Vision, the study of teaching computers to see and understand what they are seeing, has been an active area of research for at least 30 years. While the field has (and continues to) make impressive progress on a number of fronts, it is unlikely that watching "natural" video of people will be likely to produce high quality motion data for a long time to come. We will discuss this in more detail in in Section **??**.

At a much more basic level, the problem of watching human (or animal) motion is problematic. The shapes are irregular, different pieces will often obscure and occlude other pieces, few specific points are distinguished, pieces are often the same color, etc.

To make the sensing problem solvable, motion capture techniques rely on simplifying the problems, or giving computers artificial senses that do not have these problems. Systems based on making the vision problem easier, called *Optical Motion Capture Systems*, rely on using special markers and controlled environments to make seeing individual parts easier. Such systems will be discussed in Section **??**.

Vision is the primary way that we sense the motion of others. To a limited degree, we can use touch (such as having someone guide our hand through a complex manuerver, or leading a partner while dancing). Touch is a very practical mechanism for small scale motion detection - indeed, it is our primary modality for computer input (mice, keyboards, ...). Motion capture systems based on "touch" (or mechanical connection to the performer) have problems in that they are more cumbersome and intrusive (since they have to touch the performer). Such systems will be discussed in Section **??**.

Designers of computing systems are not limited to mimicking the senses that are used by biological systems. We can engineer sensors for things such as radio frequencies, radiation and magnetic resonance. The medical community has made extensive use of such sensing to "see" internal structures of the body. While the techniques used in medicine such as MRI, X-Ray, and CT scanning are most likely impractical for animation production, they do suggest that the potential for more advanced sensing technology exists.

One synthetic sense that has been successfully applied to motion capture is the sensing of magnetic fields. As will be discussed in Section **??**, using magnetic tracking technology leads to motion capture solutions with a very different set of tradeoffs from optical ones. For example, unlike light, magnetic fields are not obscured by humans. However, they are more limited in range, and require more sophisticated sensors.

1.6.4 Challenge: Modeling Humans for Animation

The success and rapid advance of medical imagine technologies implies that it will be possible (albeit possibly at an extreme cost) to sense humans with amazing precision. Even with the availability of such data, the animation problem is not solved. The most significant issue is that the observations are rarely in a form that is what is needed for animation. It is very easy to have too much data.

For example, suppose we had a tracking technology that could tell us the exact configuration of all 206 bones of a human being. This level of detail is clearly not needed for many animation applications. Simpler representations are often preferable.

For animation, we typically have other concerns in how we model characters than biological precision. We want our characters to be simple enough that we can draw them efficiently and control their movements. We do not want to expend our energy and effort modeling detail that will not be needed.

For example, consider a human knee. When a surgeon looks at an X-Ray and MRI scan, they can see a complex system of bones, cartiledge, muscles, tendons, fat, and other tissues. If they are performing an operation, they need to see all of this detail such that they can effectively perform their surgery. Many of these pieces are involved in the normal motion of the knee. If you observe the movement of the knee up close (as a physical therapist might), the complexities of the motion caused by this system of many parts is apparent.

From a distance, however, the knee can be viewed as a hinge between the two rigid segments of the leg. Especially if the details of the motion are obscured by clothing. This simplification does lose details: the knee is not a hinge, and the complexity of its motion may cause some of the subtleties in motions, such as it absorbing the impact of a footplant while running.

For computer animation, we very often model the knee as a hinge between two rigid segments. This gives us a small number of parts to be concerned with. Also, a hinge is very easy to characterize mathematically (it is what we call a *kinematic pair* where the two pieces that are connected move in a simple relation to one another, in this case rotation).

For any given application, there is an appropriate level of abstraction that trades off detail for pragmatic concerns. Even within animation this varies. For example, when doing animation for film, it may be appropriate to model the curvature of the back to get natural looking bending motions, whereas in doing a real-time simulation or game, the back might be treated as a small number of rigid elements in order to achieve performance goals.

No matter what level of abstraction is chosen to represent the character and its motion, it most certainly will not be as complex as the original performer. This means that there will be a translation problem: we must somehow map between observations of the complex performer to the simplified character. Something will neccessarily be lost. The challenge is to do the translation in a way that as much of what is important in the motion is preserved, and only unimportant things are lost.

1.6.5 Challenge: Retrieving the Essence of Motions

The last sentence of the preceeding section brings up an issue that is at odds with what was discussed in Section 1.6.2. On one hand, we want to be able to preserve what is important in motions as we capture and process them. However, the things that are important are often subtle and difficult to characterize. This is the source of the fundamental challenge of motion editing, and is a challenge for other aspects of producting animation from observation.

At a more concrete level, what we measure during capture, and store during animation, are specific details of a motion. The position of something, an angle. A huge sequence of the details makes up a motion. What is important are typically high-level properties.

This gap between what we sense, store, and manipulate with animation systems and the high-level properties that are important in motions for creating the desired animation is one of the most significant sources of challenges for all aspects of animation production. This challenge will affect all stages of the animation process, and will impact all aspects of our methodologies. For example, we will need to plan to make sure we capture sufficient details to achieve the high-level qualities we need and capture with sufficient fidelity such that these qualities are not lost. The captured observations must be processed in a way such that the details are not lost in a way that disturbs the high-level properties, and edited in ways that preserve or control these high level properties, despite the fact that what the method must work with are the myriad of details.

The ability to have computers manipulate animation at higher levels of abstraction, rather than directly managing the details, has been a holy grail of computer animation research.

1.7 Challenge: Getting the Right Performance

The pervious challenges imply that, at best, what we capture is the actual performance. This means that it is critical that the performance we capture is the performance that we want. Put another way, since it is so difficult to manipulate the "essence" of a motion, we need to make sure that the initial performance has the desired "essence" since we cannot add it in later (cannot is too strong a word - see Section 1.9.1.

Getting the proper performance is a combination of planning, acting, and directing. The art of using these three phases to achieve a desired performance is not unique to motion capture: it is quite similar to what is needed for more traditional media such as theater and film. Acting, dancing, directing, and stunt acting are well developed art forms. This implies both that creating a performance is challenging (since it requires such work to do in the traditional media), but that there are people who are good at doing it.

Performing for motion capture is different than performance for more traditional media. On one hand, since we usually only capture the gross motion, the smaller details are unimportant. This makes things easier since the performers do not need to worry about the details that are not captured. On the other hand, because the performance must happen in the "vaccuum" of the motion capture studio (an empty room, rather than a populated set), often disconnected from the final result, and with complex equipment to deal with, new complications are added.

Experience with the performance arts (particularly acting and directing) is therefore a critical skill for the motion capture practioner. As evidenced by all the effort that goes into the performing arts, creating a performance is challenging (which is why I consider getting the right performance a fundamental challenge of creating animation from observation).

Motion capture is often plagued by the "Joe from Accounting" syndrome - when insufficient attention is paid to getting the proper performers, and the needed performance, the final results will suffer. Even in the event that "programmer motion" is what is desired, coaxing a "normal" and "natural" performance out of someone is an art. One of the first concepts of acting is that there is no such thing as a generic walk.

1.8 Abstractions of Human Motion

We have defined motion capture to be the construction of a representation of a motion from observations. In this section, we briefly consider the problem of what that means. For this discussion, we consider human body movements, although identical issues arrise in observing other creatures, or capturing facial motions.

A first consideration is how much information we need to record in order to represent a human motion with sufficient detail. The answer to this is "it depends." Clearly, there are cases where an immense amount of detail is crucial, and other cases where very little is needed. For example, if we are planning a very sensitive surgery, we may very much care about all of the minute details. At the extreme, we might consider wanting to know about every sub-atomic particle that makes up a particular person as they move. While the motives for such immense details are unclear, its impracticality makes an important point: more detail is not necessarily better. Additional detail comes with a cost. Each additional piece of detail means that there is more information to obtain, store, validate, manipulate, analyze and visualize. The cost of too-little detail is insufficient fidelity to record potentially important features of a motion.

discussion of various abstractions that we use in modeling human motion, including skeletons, and a discussion of why these do not fit a real person precisely

1.9 The Process

1.9.1 Post-Processing

1.10 Roadmap: How to Use this Book

Hopefully, from this introductory chapter, you have a reasonable expectation as to what this book is about. However, the nature of this book makes a reader's guide useful.

1.10.1 Audience

The endeavor of computer animation involves a wide array of people. Animators, writers, producers, technicians, programmers, marketers, writers, and even avid fans all have interests in the process by which animations are made. All of these different participants are likely to have some interest in understanding how motion capture fits into the process of animation.

This wide-ranging audience provides a dilemma for an author. On one hand, it is important that all of these different audiences are able to communicate with one another about the subject through the use of a common vocabulary. This suggests that a single book for everyone is essential. On the other hand, everyone has far different needs. The equations that a software developer wants to help them implement tools often scare away many of the less-technical readers. In the end, I think the need for common understanding prevails, and therefore hope this book can serve everyone's needs.

If you are a non-mathematically inclined reader, please do not be scared away by the chapters containing equations. We have intentionally made these pieces self-contained, and these parts can be omitted if you do not want to understand this level of details.

Flow-chart of book chapters.

Paragraph about what an artist/animator would want from this book.

Paragraph about what a producer would want.

Paragraph about what a developer would want. If you're the type who is wondering "where's the code," you may be a little disappointed if you flip through the book, . Almost all of the problems in working with motion capture defy the simple solutions that can be written into small code snippets that can be placed in a book. Some of the basic pieces, which can be small, have already become part of most standard libraries. Instead, we hope that a book like this one will help give you insights and ideas that you can use to develop your own tools and systems.

Paragraph about what a researcher would want.

1.10.2 What This Book Is Not

We have explicitly chosen to keep the topic of this book somewhat narrow.

Because our focus is on animation from observation, not on motion capture, we do not consider other applications of human tracking. Many of the technologies used to record movements for animation have other applications. Optical tracking, for example, was originally developed for biomedical analysis applications, and this still provides a large portion of the usages of such systems. While topics in digitally recording human motion for applications in medicine, sports performance understanding, and surveilance are quite interesting, they are beyond the scope of this book.

Our focus is on animation from observation for figure animation, as opposed to facial animation. While many of the issues and techniques for creating facial animation using motion capture techniques are similar to those of full-body capture, there are also many differences. At a high level, figure motion capture is mainly concerned with the gross movement of rigid pieces, while facial animation is more concerned with the fine details of soft-tissue movement. Facial motion capture and animation is an interesting topic in its own right, and deserves a better treatment than we could provide in this book.

Our focus in this book is on "direct" mappings from the movements to the animated character, rather than puppetry. With puppetry, movements of the performer are mapped to quite different movements of the target character. Just as a real puppeteer might pull on the strings of a marionette to make it walk, there are styles of animation where the performer turns knobs to move a character, or uses their hands to control a character's facial expressions.