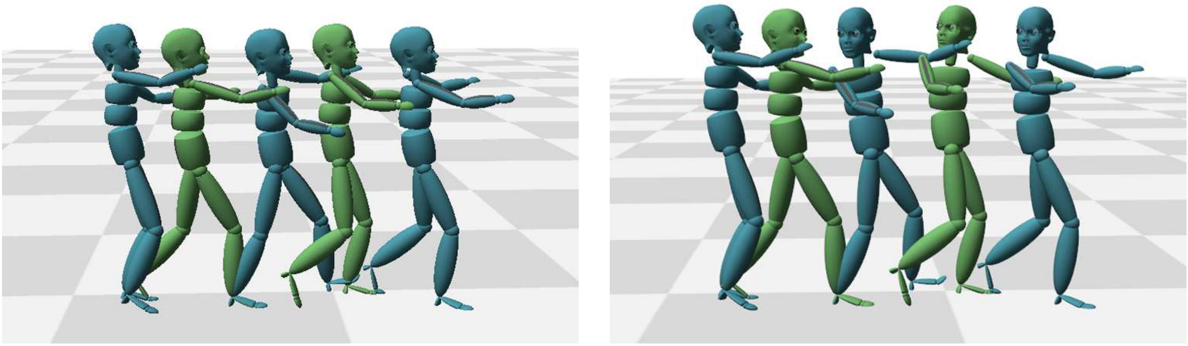
Chapter 5

Gaze Control

Because humans depend heavily on their sense of sight, it is natural for a person to look at interesting objects in the environment or down the direction of travel. To an observer, shifts in the location or direction where a character is looking might not only indicate a shift in attention but can also convey a person's goal before they act on it. For instance when faced with a choice between five different colored cups lined up on a table, a person will fixate on the cup they choose before reaching their hand out to grasp the cup. I call these natural cues the *gaze* of a human.

Currently, methods for motion clip synthesis fail to provide a way to synthesize high-quality motion of a character adjusting their gaze independently of overall body motion. Because of this, interactive characters in video games and training simulations most commonly lack gaze cues altogether. This lack of gaze cues leads to characters who look as if they are disconnected from the environment they are placed in. A human in a real environment needs to adjust their gaze many times in order to effectively take in their surroundings.

In this chapter, I present a new method for decoupling the gaze of a character from his full body motion in a way that allows high-quality, controllable motion clips to be synthesized at runtime. This model takes the form of a *parametric gaze map* that maps a requested gaze change to a low-dimensional representation of the way a human adjusts their gaze. This parametric gaze map can be applied to a base full body motion in order to adjust the gaze direction of that motion realistically and by an accurate amount. This model combines knowledge from the biological and psychological sciences with examples captured in a motion capture studio in order to achieve these realistic-looking results using little storage space. The method produces results that retain the correlations originally exhibited in each base motion while effectively adjusting the motion's gaze



a **b**

Figure 5.1 Gaze adjustment for a zombie motion. (a) A motion captured example motion of a person walking like a zombie. (b) A motion generated by adjusting the gaze of the zombie walking motion so that the gaze is directed over the right shoulder.

direction. Figure 5.1 shows the result of the technique applied to a motion of a person acting like a zombie. The gaze adjusted motion still clearly shows the character acting like a zombie but the overall configuration of the upper body has been adjusted in a natural way in order to achieve the desired gaze direction.

The rest of this chapter is organized as follows. Section 5.1 describes in detail the example-based, biologically and psychologically inspired model used to control gaze. Next, Section 5.2 describes how to construct and use a parametric gaze map to adjust character gaze. This section includes a detailed explanation of the capture procedure for example motions of a person adjusting their gaze in a motion capture studio. Section 5.3 then presents some results of the technique. Finally, Section 5.4 concludes with a general discussion of the method's uses.

5.1 Gaze Motion Model

In an interactive application, there are many instances when character gaze could be important. A character might look at objects that they are interacting with or gaze at key people or scenery as they walk by. These types of environmentally driven gaze cues can cause the character to appear more connected to the environment. Gaze cues might also be used more directly by a user. In particular, for interactive environments where many different virtual characters are being controlled by different people, such as in a massively-multiplayer online video game, a user could directly adjust the gaze of a character in order to point out interesting landmarks. As it is already commonplace in video games to control the global orientation of a character independently of the viewing direction, it would be possible to use these same controls to adjust the gaze direction of a character in order to provide better cues to other players.

A small number of recent video games have exhibited gaze cues, indicating an interest in gaze, but typically the methods employed consist of little more than twisting the head in order to point the eyes in the gaze direction[1]. But studies originating in the biological and psychological sciences indicate that gaze control is not a simple function but one that involves coordination between the

---

[1] One notable exception is the game Half-Life 2 where many man hours of work produced an intricate facial muscle model and eye focus controller, creating NPCs with realistic eye motion. However, the focus in Half-Life 2 is on the face and eyes, not necessarily on the motion of the body that achieves an overall gaze pose.

eyes, head, and spinal joints. In fact, these studies have found that head rotation actually factors little into overall gaze change [GV87, FS00].

Interactive applications are in need of a realistic method for adjusting the gaze of a character. While a character is performing some other action or actions with their full body, such as walking, running, shooting a gun, or holding a suitcase, it would be desirable to be able to redirect the character's gaze in response to environmental or user-directed gaze goals without unduly affecting the overall body motion.

The goal of gaze control in this chapter is to be able to synthesize a new *clip* of motion where a character performing some independent full body motion, called the *base motion*, adjusts his pose in a natural way such that his eyes point in a specified direction in relation to where they would point without the adjustment. Gaze control allows gaze to be directed independently of full body motion.

Because of the importance of both efficiency and realism, my approach to gaze control is to use motion-captured examples of a person adjusting their gaze to build a low-dimensional model for motion adjustment that effectively captures important aspects of gaze. In the rest of this section, I review work in the biological and psychological sciences that study how real humans adjust their gaze. I then describe in detail my example-driven model for gaze control, called a *parametric gaze map*, that is informed by these observations.

5.1.1 Biological and Psychological Observations

For many decades, the biological and psychological communities have been interested in understanding how humans (and other primates) unconsciously adjust their gaze. These studies have observed a number of interesting facts about human gaze that can be used to support and inform a model for virtual gaze control.

Many of the studies on how humans adjust their gaze focus only on how the head and eyes contribute to overall gaze by restricting the motion of the shoulders [KGM07, GV87, FS00], but more recently researchers have begun to focus on the importance of the spine for adjusting gaze as well. In particular, McCluskeyl and Cullen's [MC07] recent study of the way unrestrained rhesus monkeys adjust gaze shows that the spinal joints can have a significant effect on gaze direction, especially for gaze shifts of more than 40°.

Psychological literature observes that the proportional motion of the spinal joints, head, and eyes used to adjust gaze to a desired direction is primarily guided by complex psychological processes in the brain rather than by the mechanics of the body [GV87, FS00]. And this work further shows that these proportions change depending on the subject as well as the size of the gaze adjustment. These observations argue against models of human gaze that attempt to achieve a desired gaze goal by minimizing the amount of energy expended.

But the biological and psychological literature also observes that gaze adjustments follow a definable pattern. One pattern observed in gaze adjustment motions is that the onset of movement used to adjust gaze cascades from the eyes down the spine [KGM07, GV87, FS00, MC07]; the eyes begin to move before the head, and the head begins to move before the upper back, and so on. The study performed by Kim et al. [KGM07] drew attention to another pattern of gaze adjustment motions. In this study, when presented with a target to look at, a subject would quickly adjust their gaze direction in a gross way, usually overshooting their target, and then slowly readjusting in order to correct for this gross motion. However, Kim et al. also observed that the details of how a person adjusts their gaze differs from one subject to another. The speed of the motions, the point at which each joint begins to move, and the amount by which subjects overshoot their target differ between subjects.

These findings in the biological and psychological communities support my approach of using an example-based, low-dimensional gaze model to adjust character gaze in two ways. First, the findings show that individuals adjust gaze differently depending on many factors, including personal characteristics, such as flexibility. This individuality motivates an example-oriented approach that would allow the gaze of a virtual character to be adjusted in accordance with an individualized model. Second, biologists and psychologists observe that even though different people adjust their gaze differently, each of these individuals follows patterns that are consistent across all gaze adjustment motions. These similarities can be used to design a low-dimensional model for the way an individual gaze adjustment is made.

5.1.2 A Biologically and Psychologically Inspired Model for Gaze

The biological and psychological literature reviewed in Section 5.1.1 provides a pattern for gaze control. In particular, it observes that:

1. the spinal joints, head, and eyes contribute to the adjustment of gaze
2. the initiation of the movement of these different joints follows a cascading effect, and
3. during the gaze adjustment, gaze often overshoots its goal orientation, resulting in a slow readjustment period.

But this general pattern does not provide the quantitative details necessary to adjust gaze. These missing details include how much the gaze overshoots, the time at which this *overshoot peak* is reached, the time delay between the movement initiation of each of the joints, and the proportional amount that each joint contributes to the overall gaze adjustment. In fact, the literature suggests that there is not a single correct value for these terms; instead, details differ depending on the size of the gaze adjustment as well as the individual who performs the adjustment.

My method for decoupling character gaze from overall body motion is strongly guided by these observations. To represent the adjustments made to a base motion in order to adjust gaze, I have developed a low-dimensional model that explicitly accounts for the patterns observed in real human motion. Quantitative values associated with the model are filled in using example motions of a human subject adjusting their gaze.

My model is called a *parametric gaze map*. A parametric gaze map maps the amount by which a character's gaze needs to be adjusted to a motion of the spinal joints, head, and eyes that in a short amount of time will achieve the desired gaze adjustment when added to the base motion[2]. I can parameterize the amount of a gaze adjustment as the change in yaw, r_{yaw}, and pitch, r_{pitch}, of the eyes in relation to the root. This parameterization is equivalent to a latitude/longitude parameterization of the unit sphere. Biologically, it is possible to adjust the gaze of a person looking straight forward by an amount that corresponds to any point on the unit sphere, (r_{yaw}, r_{pitch}) where

---

[2] This mapping assumes that the way that a person adjusts their gaze does not depend on their starting configuration, but within the psychological community there is a debate as to whether this is true or not [KGM07].

$r_{yaw} \rightarrow \{-180°, 180°\}$ and $r_{pitch} \rightarrow \{-90°, 90°\}$. In fact, [KGM07] observes that horizontal gaze in the range $\{-180°, 180°\}$ is possible *even* when the shoulders are rigidly held.

For each (r_{yaw}, r_{pitch}) pair, the parametric gaze map supplies the spine, head, and eye adjustments needed to achieve the desired gaze change. I call these additive spinal motions *gaze displacement motions*. A gaze displacement motion consists of two parts that explicitly encode the patterns observed in real human motion. A gaze displacement motion consists of two parts. The first explicitly represents the change in a character's skeletal *pose* that is used to achieve a desired gaze orientation. The second focuses on the pattern of the *motion* used to achieve this final pose.

I represent the final pose adjustment for a gaze displacement motion, $\mathbf{P}^{+}$, as a displacement of the spinal joints, head, and eyes from the base pose[3]. For my skeletal model, this can be stored as a 5-tuple of quaternion rotational offsets:

$$\mathbf{P}^{+} = \{\mathbf{q}^{+}_{Pelvis}, \mathbf{q}^{+}_{LowerSpine}, \mathbf{q}^{+}_{UpperSpine}, \mathbf{q}^{+}_{Neck}, \mathbf{q}^{+}_{Eyes}\} \qquad [5.1]$$

The motion portion of a gaze displacement motion uses a 7-tuple of floating point numbers. This includes:

- l: The length of time in seconds that it takes to adjust gaze, starting from the point when the adjustment begins and ending when the full gaze adjustment, $\mathbf{P}^{+}$, has been reached.
- $\mathbf{T} = \{t_{Pelvis}, t_{LowerSpine}, t_{UpperSpine}, t_{Neck}\}$: The time at which each joint starts moving, other than the eyes, which are assumed to start moving immediately. For convenience, this time is normalized to the range 0 to 1 with respect to the length of the motion.
- t_{β}: The time at which the motion reaches its overshoot peak, also normalized to the range 0 to 1 with respect to the length of the motion.
- β: And, finally, the amount by which the motion overshoots, stored as a multiplier associated with the final pose displacement.

---

[3] Note that because eyes are not a part of the skeletal hierarchy presented in Section 3.1 but are an important part of gaze, I augment my skeleton hierarchy with the addition of eyes. For data without eye orientation information, the eyes are assumed to be looking straight forward from the head.

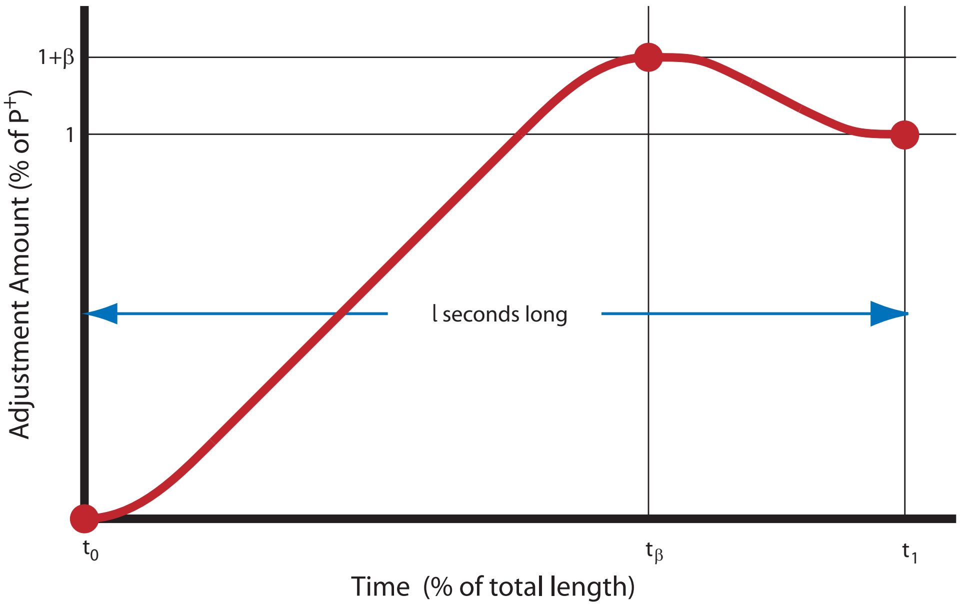


Figure 5.2 A pictorial depiction of a gaze displacement motion.

Putting all of these numbers together, a gaze displacement motion, $\mathbf{G}$, can be defined as:

$$\mathbf{G} = \{\mathbf{P}^{+}, l, \mathbf{T}, t_{+}, \beta\} \tag{5.2}$$

This representation of the motion of a person adjusting their gaze is a direct mapping of the three pattern observations from the biological and psychological sciences described at the beginning of this section. The values for any particular gaze displacement motion can be filled in using an example gaze adjustment motion (see Section 5.2.2.2). See Figure 5.2 for a pictorial representation of a gaze displacement motion.

I construct a parametric gaze map from a set of example gaze displacement motions using a form of blending-based parametric synthesis. The gaze displacement motion, G, returned for any (r_{yaw}, r_{pitch}) can be layered onto a base motion to adjust the character's gaze. By representing an entire example motion with just a few parameters, my method not only reduces the amount of storage needed for a gaze displacement motion but also allows gaze to be applied efficiently at runtime.

5.2 Parametric Gaze Maps

This section describes my method for capturing example motions of a person adjusting their gaze in a motion capture studio, converting the raw motion capture data collected at the studio into a gaze displacement motion, building a parametric gaze map from these gaze displacement motions, and applying the parametric gaze map to a base motion to adjust gaze.

5.2.1 Capturing Gaze

For this dissertation, I sought to capture example motions for a parametric gaze map that show how a human adjusts their gaze when asked to look at different locations in 3D space. Unfortunately, because of the confines of the capturing environment, it was necessary to limit vertical gaze to a much smaller range - $r_{pitch} \rightarrow \{-12°, 12°\}$ for all of the experiments presented in this dissertation; the horizontal gaze range of the subject was captured in full. Because of my model's dependence on a specific type of gaze motion, this section describes in detail how these example motions were captured for this dissertation.

5.2.1.1 Motion Capture Preparation

To capture example motions of a subject adjusting their gaze, it is useful to have objects in the motion capture environment that can be used as gaze targets. Acting as these points of interest, 33 signs labeled with letters were placed around the motion capture environment. With a human subject at the center, signs were placed at 36° increments along the horizontal axis and at 12° increments along the vertical axis, starting from the point $(0, 0)$, or the point where the subject's gaze is directed when at rest. Figure 5.3 diagrams the setup of the signs along these axes.

Markers were then attached to the subject's upper body along the spine, on the head, and on the arms. For the subject's spine, I marked three landmarks with an equilateral triangle of markers: the lower back, the upper back, and the base of the neck, between the shoulder blades. By using a triangle of markers, I was able to guarantee that I could construct a coordinate system at each of these points along the subject's spine during processing (see Section 5.2.2.2). I also placed a marker at the base of the spine as a reference point for the motion. The subject's head was outfitted with a hat with four markers attached to the sides in a square and one marker attached to the top, in the middle. A small additional marker was placed between the subject's eyes. I used several extra markers for visualization purposes only: one on each of the subject's shoulders, elbows, and wrists. These markers were observed and recorded by a standard optical motion capture system. Figure 5.4 shows the arrangement of these markers in 3D space.

My marker setup does not support capturing the motion of the eyes, but as stated previously, the eyes play an important role in gaze. As described later in Section 5.2.2.1, the orientation of the eyes in relation to the head is inferred from the overall orientation of the upper body in conjunction with the known target gaze direction.

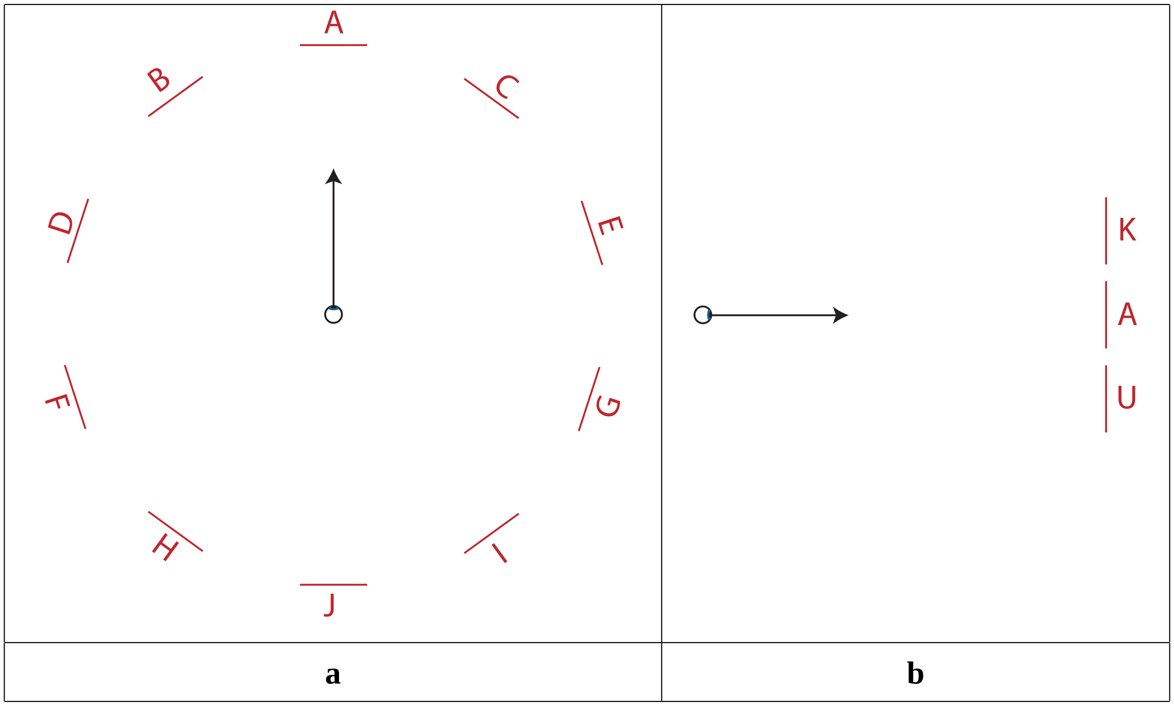


Figure 5.3 Sign placement. Signs with letters on them were arranged within the motion capture environment to act as targets for gaze capture. In all, there were 33 signs arranged in 3 latitudinal lines and 11 longitudinal lines. (a) shows an overhead view of sign placement along the middle latitudinal line. The eye and arrow in the center depict the location and orientation of the subject at rest. (b) shows a side view of sign placement along the front longitudinal line. Again, the eye and arrow depict the location and orientation of the subject at rest.

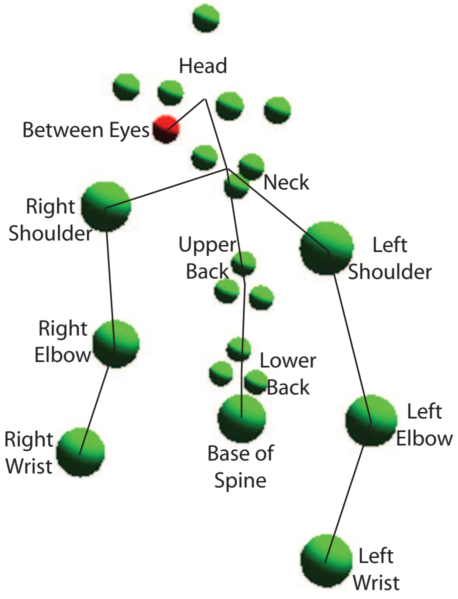


Figure 5.4 Marker placement for gaze capture.

5.2.1.2 Capturing Each Example Motion

The goal of the motion capture session was to capture the motion of the subject adjusting his gaze from a starting position to each of the target locations indicated by the lettered signs. For each captured motion, an assistant and I would:

1. point a laser pointer at the target sign, and
2. play a tone twice, in a short, long pattern, using a tone making device positioned behind the target sign. The tone making device was capable of playing 3 tones. The highest tone was always played for targets located along the highest latitude line, the middle tone was always played for targets located along the middle latitude line, and the lowest tone was always played for targets located along the lowest latitude line.

Because I wanted to capture the natural way that the subject adjusts his gaze, it was important that the subject not consciously plan his motions. Thus, the 33 motions were captured in a random order. The subject was told to do the following for each of these trials:

1. Look straight ahead at the letter **A**.
2. Close your eyes.
3. When you hear the first, shorter tone, open your eyes.
4. When you hear the second, longer tone, look at the target sign, as indicated by the tone and laser pointer.
5. Read what is written on the target sign.

Each of the motions were captured at $60hz$ and stored as raw motion capture data (i.e., the 3D locations of each of the markers on each frame). The next section describes how I use this raw data to create a parametric gaze map.

5.2.2 From Raw Motion Capture Data to a Parametric Gaze Map

The motions captured using the method described in Section 5.2.1 provide examples of the way the subject adjusts his gaze over a large range of possible gaze changes. In this section, I describe how to transform these motion-captured observations from moving 3D positions in space to a parametric gaze map that can be effectively applied at runtime to adjust a character's gaze.

5.2.2.1 Cleaning the Data

When motion is captured in a motion capture studio, the raw data that is produced often needs to be cleaned to conform to the standards needed for the example motions. For the case of the gaze example motions captured for gaze control, it is important that the 3D locations of each of the markers appear in every frame of motion, that the upper body motion is stabilized in relation to any extraneous lower body motion, and that the motion is trimmed to only include the region of time that is interesting for gaze control.

When a motion is captured using the sensors at a motion capture studio, it is possible that one marker might become "invisible" to the sensors for a brief period in time. For instance, for the optical motion capture system used to capture all of my example motions, it is possible for markers to become blocked from the cameras by objects in the environment or self-occlusion of the subject. If they are well-placed, it is unlikely for the markers to disappear for long. Thus the first step for cleaning the captured gaze examples is to fill in any missing marker data points.

For each motion-captured gaze example, missing marker locations are inferred using linear interpolation of existing data in temporally nearby times. In particular, a missing marker data point at time t is calculated as the weighted average of the marker's last known location, prior to time t, and next known location, after time t, weighted according to the amount of time between the observations.

After filling in missing marker data, the motion examples can be stabilized. My gaze model presented in Section 5.1 considers only the motion of the spine, head, and eyes. Yet, slight changes in the global position of the base of the subject's spine affects the velocities of the markers on the rest of the upper body. Thus, the second step needed to clean the gaze example data captured at the motion capture studio is to stabilize the root such that it does not move, while maintaining the relative distance between each of the markers in 3D space. This can be done efficiently by shifting the entire captured motion on a frame-by-frame basis so that the marker at the base of the spine is always at the origin.

Finally, because the motion capture equipment begins recording each motion before the subject starts to move and continues to record after the subject has completed the task of changing his gaze, the raw data for each motion-captured example contains additional frames of motion before and after the needed example motion. Ideally, each motion would begin at the moment when the subject begins to move and end after the subject has finished reading the letter printed on the target sign (see Section 5.2.1 to review the sequence of steps used to capture the gaze example motions in a motion capture studio). This region of motion can be extracted from the adjusted motion by locating periods of time when the subject is barely moving, henceforth referred to as *periods of rest*. To do this, my system uses the following simple, automated technique:

1. For each frame of motion, t_i, it calculates the average velocity magnitude of the markers, v_i, using finite differences.

2. Next, the system labels a frame i as “still” if v_i is below a tunable threshold, V. For processing all of the data in this dissertation, $V = 6mm/s$. A sequence of “still” frames are then grouped together to form a period of rest.

3. To ward against the effects of noise, the system combines periods of rest that are separated by a short amount of time ($1/6$ of a second for my data) to form a single larger period of rest.

4. I can assume that the first period of rest temporally corresponds to the period of time when the subject is waiting for the cue to change his gaze. Similarly, the second period of rest most likely corresponds to the time when the subject has settled his gaze on the target sign and is reading the writing. Thus, the motion data is automatically cropped by the system such that it only contains the frames of motion starting from the end of the first period of rest to the end of the second period of rest.

In most cases, this simple technique effectively crops the example motions. The notable special case is the gaze motion whose target is $(0, 0)$. Since a change of gaze from $(0, 0)$ to $(0, 0)$ does not actually include a period of motion, the example must be cropped manually. All of the other example motions for this chapter where cropped automatically.

5.2.2.2 Building a Parametric Gaze Space

After the captured example motions have been processed, they can be quickly and automatically constructed into a parametric gaze map that allows fast and accurate synthesis of motions with gaze adjustments. First, each of these motions are translated into a gaze displacement motion (see Section 5.1). These gaze displacement motions are then mapped into a parametric motion space of gaze motion, forming the parametric gaze map. This parametric gaze map can be applied at runtime to adjust gaze as described later in Section 5.2.3.

To calculate gaze displacement motions for the processed motion data, the system analyzes the orientations of the joints throughout the example motion. To do this, it is necessary to be able to calculate the orientations of each joint in the captured motion. For the joints along the spine that were marked with a triangle of points, the orientation can be defined by constructing the local coordinate system defined by the triangle (or the square, for the head). When placed in matrix form, this coordinate system represents the orientation of the landmark in relation to the origin. This matrix can be converted to a quaternion representation using the standard conversion equation [Sho85].

Given this method for computing the global orientation of a joint from the raw motion capture data, my system calculates the pose portion of the gaze displacement motion, $\mathbf{P}^{+}$, that coorresponds to a processed, example motion using the following method:

1. The final orientations of the joints are translated into a displacement from the initial pose by subtracting the orientation of each joint in the first frame from the corresponding joint in the last frame.

2. These global displacement orientations are then converted into hierarchical displacement orientations, or into a local orientation displacement from its parent joint.

3. Finally, because each gaze displacement motion was captured with a known, calibrated gaze goal, (r_{yaw}, r_{pitch}), the local orientation of the eyes in relation to the head can be inferred as the orientation needed in order to achieve a global displacement of the eyes equal to (r_{yaw}, r_{pitch}).

Next, to fill in the motion portion of the gaze motion, it is necessary to locate the point in time when the motion reaches its overshoot peak, as well as the points in the example motion where each joint begins to move.

Locating the Overshoot Point:

1. The method starts by calculating the velocity of the global orientation of the head in each frame, i. This velocity can be calculated using finite differences as the angle between the orientation of the head in frame $i - k$ and and the orientation of the head in frame $i + k$, where k is a tunable parameter used for finite differencing.
2. A frame j is identified as a possible overshoot peak if the velocity at j has a different sign from the velocity of $j - 1$.
3. The overshoot point is identified as the first flagged frame where the average velocity of the frames $\{j - m, j - 1\}$ and the average velocity of the frames $\{j + 1, j + m\}$ have different signs. Again, m is a tunable parameter used to define a local window size around the frame in question.

Identifying the Start Times for Each Joint: To identify the start time of each joint, the algorithm calculates the velocity of the local orientation of the joint in question using the same method used to locate the overshoot point. The first frame at which the magnitude of this velocity surpasses a tunable threshold is defined as the start point for that joint. If no start point is found, the algorithm assumes that the joint begins moving slowly at the start of the gaze change. For all of the examples in this chapter, I set the tunable velocity threshold to $18°/s$.

As described in Section 5.1, the motion of a person changing their gaze is strongly dependent on the gaze parameters (r_{yaw}, r_{pitch}). The gaze example motions act as a sampling of the space of all possible gaze changes. Following the blending-based parametric synthesis work presented in Section 3.5, I can represent the way a person adjusts their gaze as a parametric gaze motion space, parameterized on the orientation of the gaze, (r_{yaw}, r_{pitch}). Each motion in this space is defined as a blend of the example gaze motions. My system directly uses the parametric synthesis method of Kovar and Gleicher [KG04] presented in Section 3.5 to construct this space of gaze motions. However, since the motions in this motion space are represented using the gaze motion model presented in Section 5.1, it is necessary to be able to blend motions in this representation.

The form of the representation of a gaze motion makes blending two gaze motions together straightforward. Because the pose of a gaze motion is represented as a displacement from the base pose, blending the poses of two gaze motions does not require the motions to be spatially aligned. Thus, to blend the pose portions of two gaze motions, the 5-tuples of quaternion offsets can be linearly interpolated without modification. Similarly, the motion portion of a gaze motion does not require temporal alignment since the key events are already explicitly marked in time (e.g., the gaze motion representation explicitly stores the time at which the head begins to move). So blending the motion portion of two gaze motions is as easy as linearly interpolating each of the 7 floating point numbers used to describe the shape of the motion (see Section 5.1).

5.2.3 Applying the Parametric Gaze Map

A parametric gaze map can be used to efficiently adjust the gaze of an existing base motion clip. A user or application simply supplies the desired change in gaze, (r_{yaw}, r_{pitch}). These parameters are used by the parametric motion space of gaze motions represented by the parametric gaze map. Using the blending-based synthesis method discussed in Section 3.5 combined with the gaze motion blending algorithm presented in Section 5.2.2.2, the parametric gaze map can supply a gaze displacement motion that accurately describes how to adjust the character's gaze by the desired amount.

This new target gaze motion can then be applied to a base skeletal motion by adding in the appropriate spinal, head, and eye orientation displacements on each frame. Using the definition of a gaze motion from Section 5.1, the orientation displacement of a joint, j, at time t, is:

$$\mathbf{q}_j^+(t) = \left\{ \begin{array}{ll} I & \text{if } t < t_j \\ (\frac{t-t_j}{t_\beta - t_j} * (1+\beta)) * \mathbf{q}_j^+ & \text{if } t_j \leq t \leq t_\beta \\ (1 + (1 - \frac{t-t_\beta}{1-t_\beta}) * \beta) * \mathbf{q}_j^+ & \text{if } t_\beta < t \end{array} \right\}$$

where I represents the identity rotation. In other words, there is no local orientation displacement of a joint, j, from its parent until time t_j. At time t_j, the joint's displacement begins changing linearly until it reaches its overshoot orientation, $(1+\beta) * \mathbf{q}_j^+$, at time t_β. Finally, the joint's displacement changes linearly from this overshoot point until it reaches its goal orientation, $\mathbf{q}_j^+$, when time equals 1.

In practice, the system produces motions with $c(1)$-continuity by replacing the linear interpolation terms $\frac{t-t_j}{t_\beta - t_j}$ and $\frac{t-t_\beta}{1-t_\beta}$ with an ease-in/ease-out function that was originally presented in [HWG07]. This function introduces acceleration and deceleration periods to the interpolation, providing better continuity.

By only applying this smooth displacement map to the base skeletal motion, the system guarantees that it will not introduce discontinuities that might be striking to the eye. Additionally, the method avoids producing motions that appear too smooth by layering these low-frequency gaze changes on top of the motion that already exists. The high-frequency details that exist in the base motion often contain correlating submotions, such as the slight bobbing of the head in time with footsteps, that are necessary to make a motion appear natural. All of these correlating details are retained when a parametric gaze map is used to adjust a character's gaze.

5.3 Results

To test the methods presented in this chapter, I applied the parametric gaze map I developed to a number of different base motions.

Figure 1.4 shows the result of adjusting the gaze of a character stepping up onto a platform so that he looks over his left shoulder. The motion retains the correlating details of the original base motion, such as a dramatic nod of the head as the character steps up onto the platform, while the overall gaze of the character is adjusted realistically to meet the requested gaze change. The motion of the character also exhibits overshoot of the gaze goal, a characteristic of real human motion [KGM07]. Figure 5.1 shows another example of a gaze change applied to the motion of a person walking like a zombie. Note that the resulting adjusted motion is still recognizable as a zombie motion, but the pose of the character's upper body has been adjusted to change gaze.

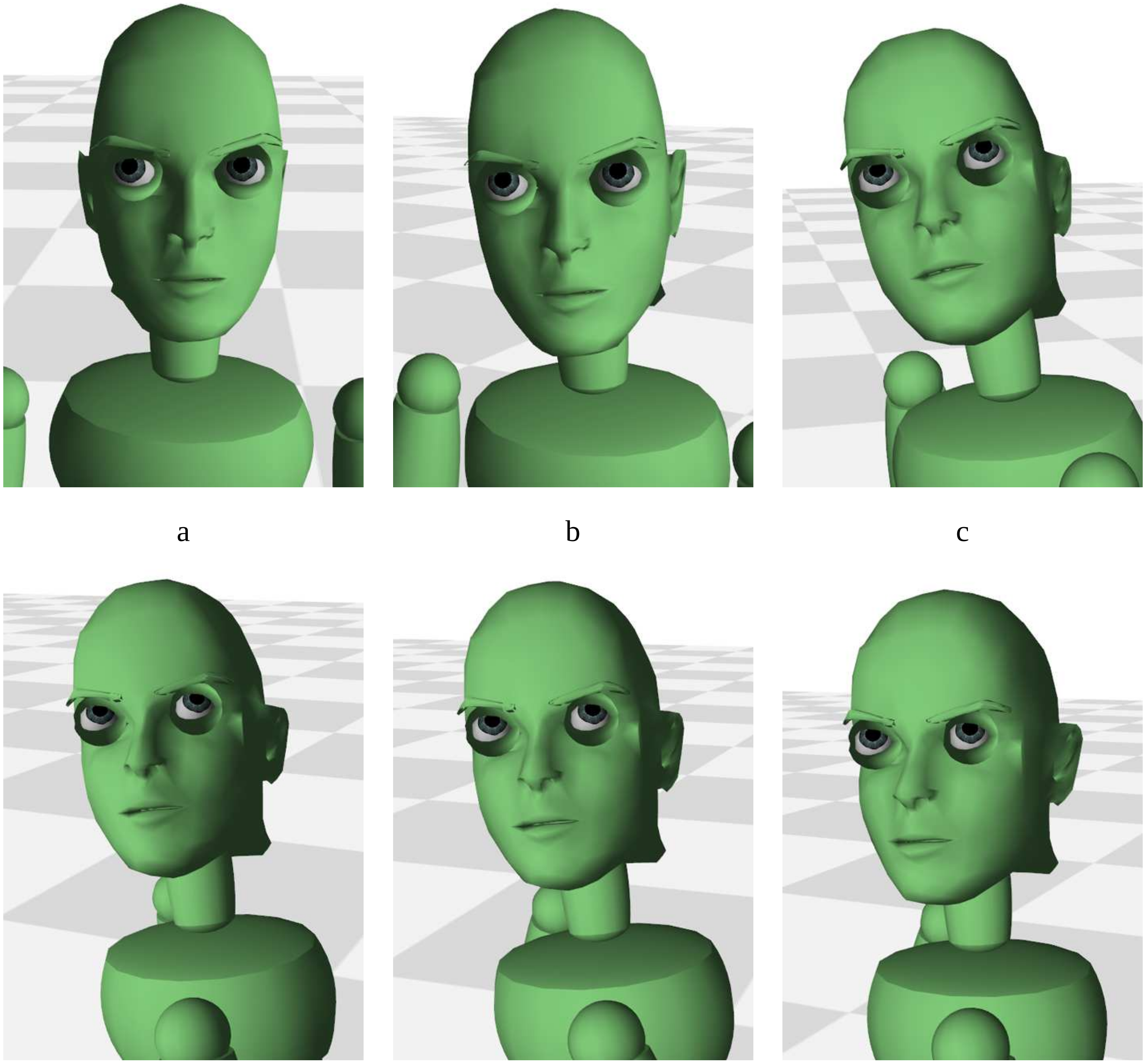
Figure 5.5 more effectively illustrates the overshoot phenomenon. This figure shows the results of adjusting a walking character's gaze such that it is pointed towards the ceiling over the left shoulder. Before the gaze adjustment is made, the eyes peer straight forward in relation to the root (Figure 5.5a). As the character adjusts his upper body orientation in order to meet the requested gaze, the eyes pass their intended goal (Figure 5.5d). The character then follows up by slowly correcting for this overshoot until he achieves the desired gaze adjustment (Figure 5.5f).

One of the strengths of using an example-based method for controlling gaze is that it captures nuances of gaze adjustments that are particular to specific subjects. For instance, the subject whose gaze was captured for the experiments in this chapter does not adjust his gaze in a symmetrical way; through evaluation of the captured motion data, it is clear that the subject's spine is more easily twisted towards the right. For all of the captured example motions where the subject turned towards the left, the subject's eyes and head contributed in a much greater proportion to the overall gaze change than when the subject turned towards the right. This asymmetrical gaze characteristic is transferred using the parametric gaze map. Figure 5.6 shows the final skeletal configuration after applying the parametric gaze map to a motion consisting only of the dress pose[4]. This figure illustrates how motions that represent a gaze change to the left result in a pose where the spinal joints contribute relatively little to the overall orientation when compared with the same gaze change to the right.

Figures 5.7 and 5.8 show additional examples of a character's gaze being adjusted using a parametric gaze map. In the first is a detailed look at the motion synthesized by adjusting the gaze of a person carrying a heavy box along a curved path so that the character looks more intently at

---

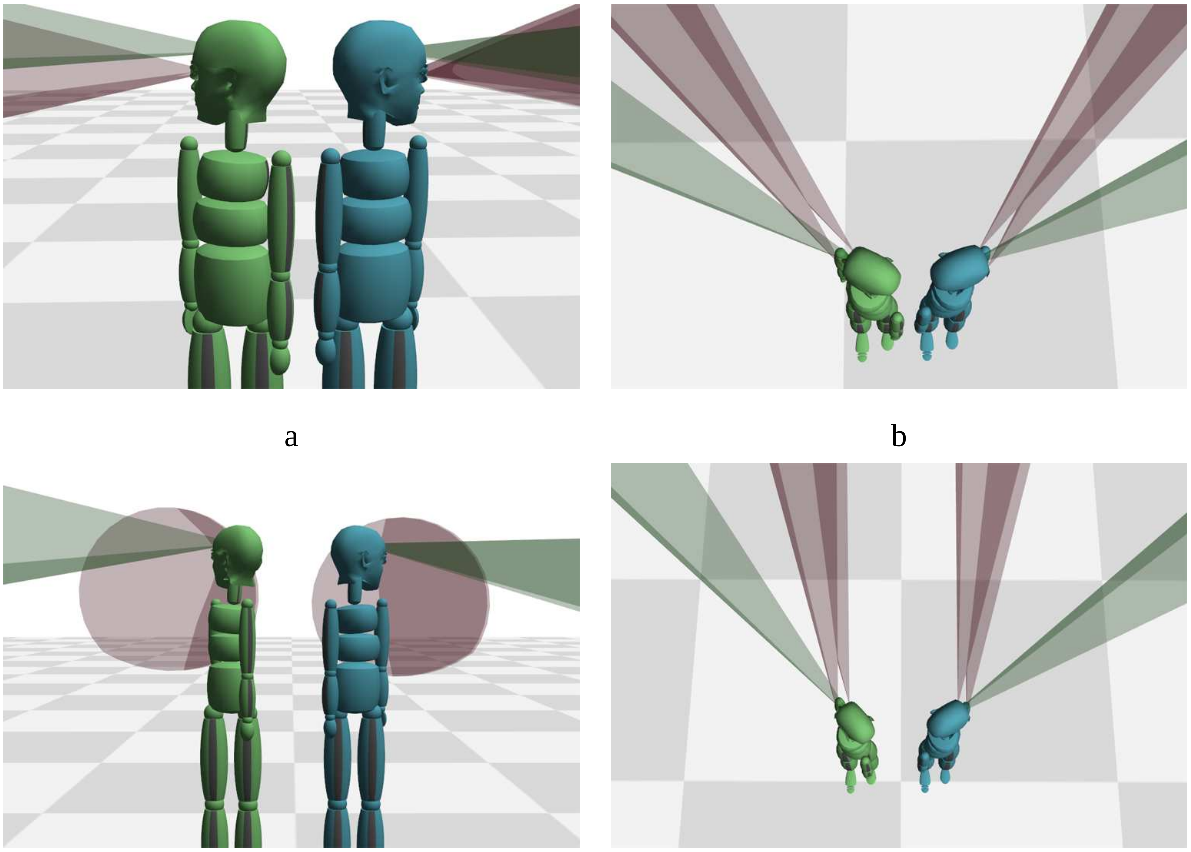
[4] The dress pose of a skeleton is the pose of the skeleton where all of the local joint orientations are set to 0



a b c

d e f

Figure 5.5 A motion of a person adjusting their gaze toward the ceiling over their left shoulder. This motion was synthesized by applying a parametric gaze map. The character starts by looking straight forward in (a). Notice how in (d) the character has overshot his gaze goal (focus on the rotation of the eyes, as this is the easiest way to notice the overshoot in a series of still images) before coming to rest at his goal adjustment in (f).



a b

c d

Figure 5.6 Asymmetric gaze changes captured by a parametric gaze map. (a) and (b) show front and top views, respectively, of the final pose configuration of a skeleton after applying a gaze change of approximately 140° to the right and left. The skeleton was initially configured in a dress pose, where each of its local joint orientations were set to 0. (c) and (d) show a similar application of the parametric gaze map for rotations of approximately 175° to the right and left. Notice how in both instances, the righthand gaze change is not symmetrical with the lefthand gaze change. Instead the eyes, whose orientation is clearly indicated by the red cones protruding from the head, contribute more to the overall orientation when the character turns towards the left. Compare the angles between the red eye cones and the green head orientation cones.

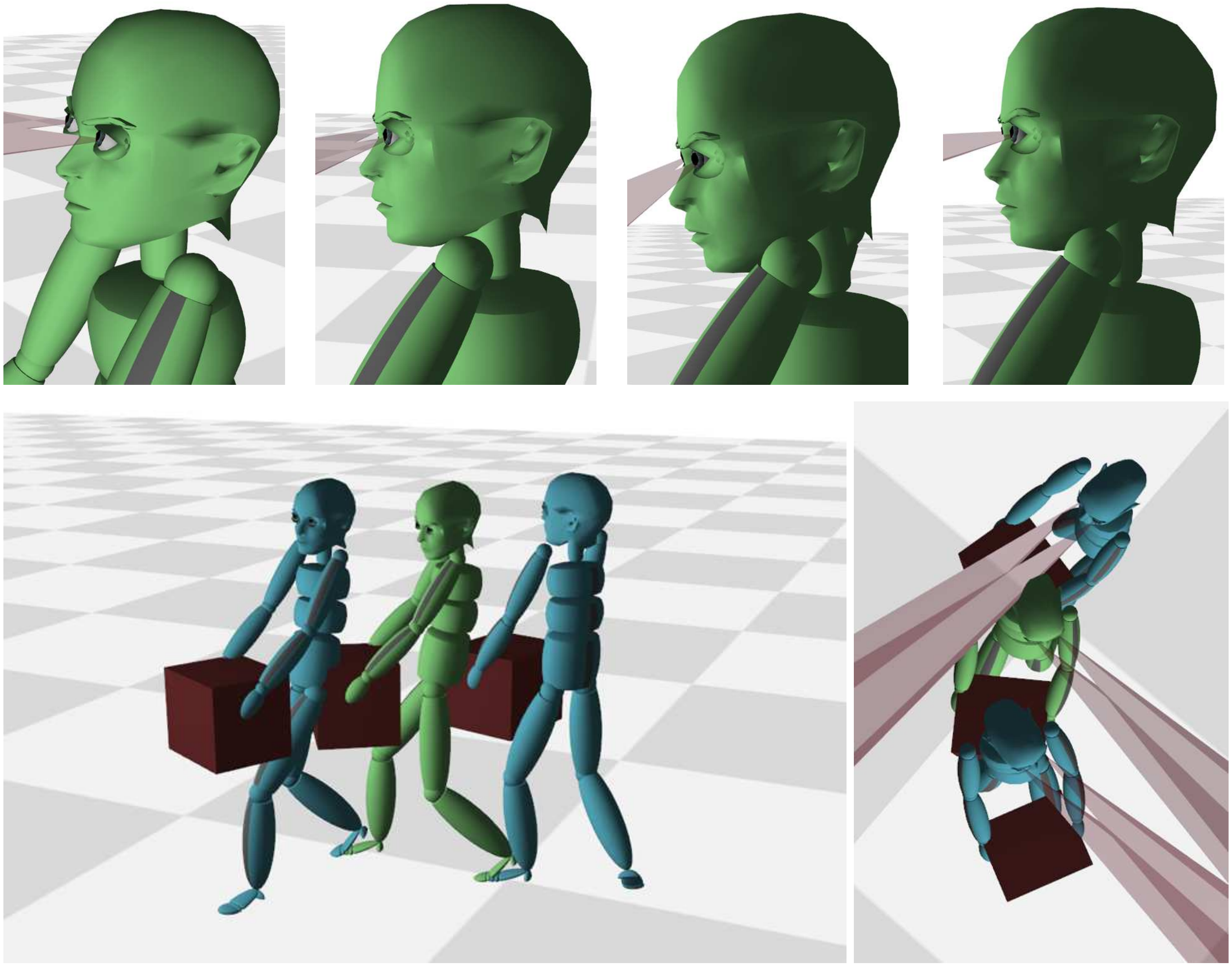


Figure 5.7 A motion synthesized by adjusting the gaze of a character walking along a curved path such that the character looks at the ground along his intended direction of travel. The images in the top row focus on the motion of the head and eyes over time, while the images in the bottom row provide an overview of the entire motion from the side and the top.

the ground along his direction of travel. Figure 5.8 compares the results of different gaze changes applied to the same base motion.

5.3.1 Algorithm Performance

In this section, I describe how my biologically and psychologically inspired method for adjusting the gaze of a character at runtime performs in each of the six categories described in Section 1.1.

Efficient Synthesis The examples in this paper were computed on a laptop computer with a $1.75GHz$ Pentium M Processor, $1GB$ of RAM, and an ATI Mobility Radeon X300 graphics card. All of the generated motions were sampled at $30Hz$. Because applying a gaze change to an existing motion clip is a fast, constant time operation, it is possible to apply gaze changes quickly at runtime. The only part of applying a parametric gaze map that is not necessarily constant is constructing the blended gaze motion in the parametric motion space of gaze motions. The time it takes to construct a blended gaze motion is dependent on the number of example motions being blended together. But because Kovar and Gleicher's method for blending-based parametric synthesis [KG04] limits the number of motions that can be blended together at each of the sample points in a parametric motion space, this time is effectively bounded.

Efficient Data Storage The storage requirements for a parametric gaze map are low. Since the example motions are stored using the compact, biologically and psychologically inspired gaze motion model, it is possible to store many example motions in a small amount of space. Additionally, since my algorithm effectively decouples gaze from full body motion, it reduces the number of example motions that need to be stored for quality control of more than one motion parameter simultaneously.

Low Latency or Response Time Since the method for adjusting the gaze of a character is designed for motion clip generation and not motion stream generation, this characteristic is not applicable.

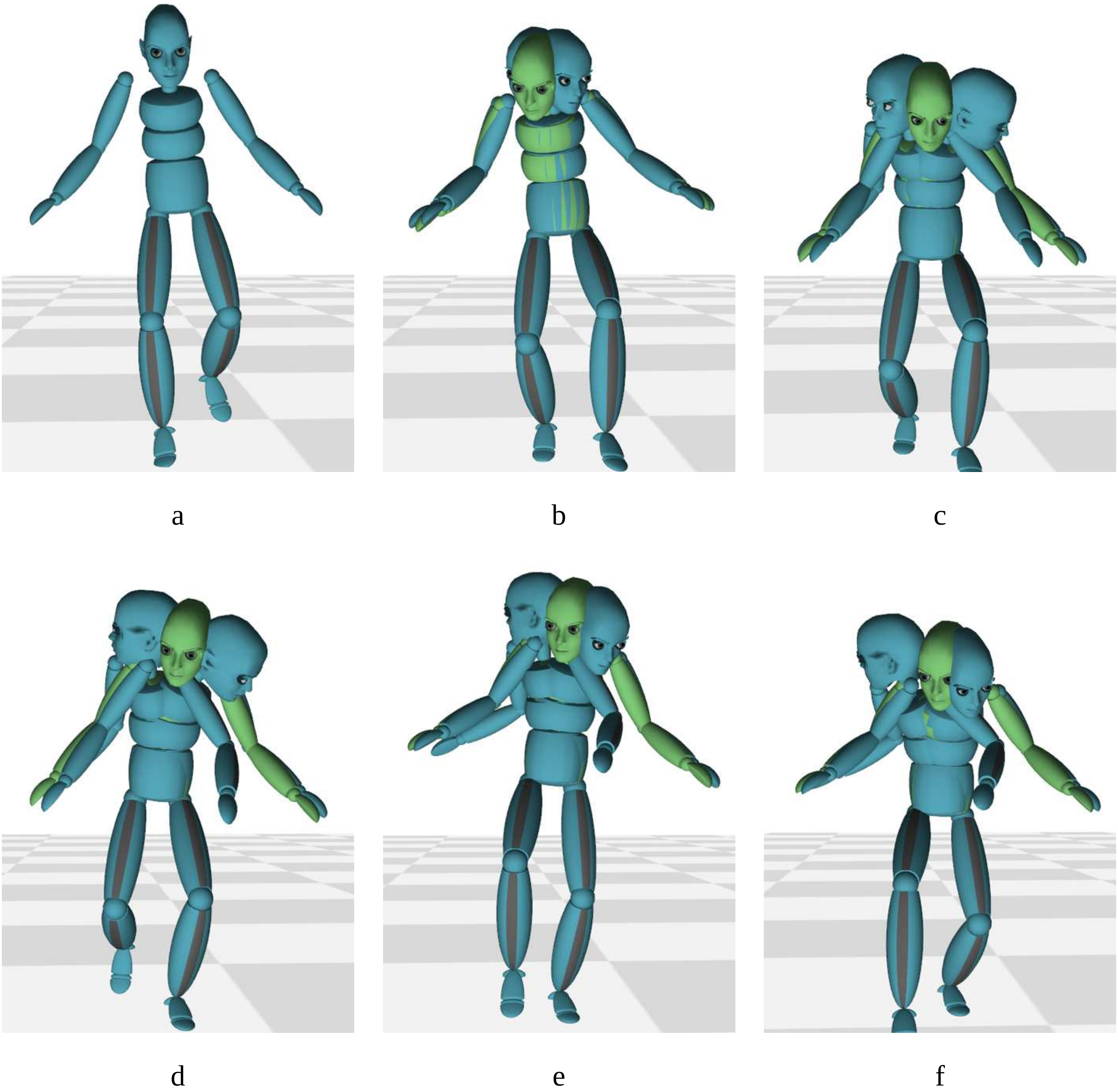


Figure 5.8 Variations on a motion: Three motions synthesized by applying a parametric gaze map to a tip-toeing motion using different gaze goals. One character is turning to look high over his right shoulder, another is looking down towards his feet, and the last is turning towards the left by a small amount.

Accurate Motion Generation Because each gaze displacement motion is computed from the example motions using blending-based parametric synthesis methods, it is possible to identify a motion that should accurately adjust a character's gaze. But there are limitations to the model that could compromise the accuracy of the method. It is necessary that the base motion that a parametric gaze map is applied to not already contain any gaze changes. A low-frequency gaze direction change in the base motion would require a new gaze change goal. Furthermore, the small details in the base motion that are retained by only applying a smooth displacement map during gaze changes might cause the gaze direction of the character to "bob." This bobbing can be naturally corrected by adjusting the eye and/or head orientation slightly in order to compensate for these small submotions.

Visual Quality Again, the quality of a virtual character motion is a partially subjective characteristic. By limiting the changes made to the base motion (see Section 5.1), deviations are small relative to the original example motion, which is assumed to be of high-quality. In addition, all of the changes made to adjust the spine, head, and eyes are smooth and c2-continuous. As the gaze model is informed by studies in the biological and psychological communities, it explicitly captures characteristic features of gaze change motions. However, the gaze motion model is lossy. In the gaze motion representation, the motion of each joint from starting pose to overshoot pose and from overshoot pose to final pose are assumed to be simple functions. Any deviation of the actual motion from this simple representation will be lost by the model.

Automated Authoring The algorithm for building the parametric gaze map for gaze control is highly automated, using only a small number of user-defined parameters, with the notable exception that a human is needed to capture the original example motions. It is also necessary for a user to supervise the automatic processing of the raw motion capture data in order to intervene when automated cleanup fails.

5.4 Discussion

This chapter presented a new method for decoupling gaze from overall body motion, greatly reducing the number of example motions needed to control gaze on top of other parameters simultaneously. Because the method is inspired by biological and psychological research, it treats motions of a person changing their gaze in a natural way, explicitly capturing many characteristics of these motions, such as overshoot and cascading joint motion. And, because the model uses captured example motions of a person adjusting their gaze, it is capable of synthesizing motions with stylistic properties associated with the way an individual adjusts their gaze, such as asymmetry.

For the examples in this chapter, only motion-captured example motions and base motions were used in my experiments, but procedural, keyframed, or edited motions would serve just as well. For instance, it should be possible to synthesize a stream of walking motion using one of the techniques discussed in Section 2.2.4 and then to apply the gaze model to these synthesized motions in order to have the character look at interesting objects in the environment. An artist could also keyframe (see Section 2.1.1) a set of example motions of a character changing their gaze. For instance, an artist could produce a small number of example motions where a character whose torso has been injured changes his gaze. These example motions could then be used to build a parametric gaze map for the way the character adjusts their gaze when injured.

While the work presented in this chapter provides a reliable method for adjusting the gaze of a character when the target is known, it does not tackle the equally important question of where a character's gaze should be directed at any point in time. Yet my reliance on psychological literature to develop a model for gaze could be extended to look at models for realistically determining when humans direct their gaze towards specific types of targets. By providing a method for decoupling gaze from full body motion, I hope that my work will inspire others to tackle the problems associated with gaze control.