A Tangible Interface for High-Level Direction of Multiple Animated Characters

Ronald A. Metoyer  Lanyue Xu  Madhusudhanan Srinivasan
School of Electrical Engineering and Computer Science  Oregon State University

Abstract
Many training, education, and visualization environments would benefit from realistic animated characters. Unfortunately, interfaces for character motion specification are often complex and ill-suited to non-experts. We present a tangible interface for basic character manipulation on planar surfaces. In particular, we focus on interface aspects specific to 2D gross character animation such as path and timing specification. Our approach allows for character manipulation and high-level motion specification through a natural metaphor – the figurine. We present an example interface for designing and visualizing strategy in the sport of American football and discuss usability studies of this interface.

Key words: tangible interfaces, animation

1 Introduction
Advances in computer animation technology, specifically motion capture, have led to natural looking 3D characters that can be controlled at a high level. These algorithms produce strikingly realistic motion while maintaining high-level constraints such as desired paths, poses, or tasks. We are interested in leveraging these algorithms for natural character motion and combining them with an intuitive interface for naive users to quickly generate animated content. To do so, naive users must be able to provide high-level character direction in the form of desired paths, tasks, and timing. This paper addresses the problem of specifying character motion on planar surfaces. This problem is of particular interest in several areas including sports training, military training, architecture, and urban planning.

Coaches in American football, for example, often strategize by designing plays on paper. Recent software packages enable coaches to design plays using a mouse based interface[33]. The ability to create animated content and visualize plays in 2D and 3D would be advantageous because the coaches and players could be presented with several views of the play as it progressed in time rather than being limited to a 2D static top-view drawing. Animated motion could also be used in training simulators for positions such as the quarterback in American football. Our system allows users to design dynamic content for training and visualization of football plays using a familiar interface – physical icons or figurines (Figure 1).

Figure 1. A user specifying passing routes with the coach's table interface.

Other applications could also benefit from simple interfaces for generating animated content. Sports broadcast companies are beginning to use animated visualizations to present strategy to the television audience. The military generates content of battlefield scenarios for use in training and visualization. Likewise, architects and urban planners could benefit from visualizing pedestrian motion within a building or in emergency situations.

2 Figurines and Affordances
One of the primary factors contributing to the effectiveness of tangible interfaces is the concept of “affordances”. When an object clearly indicates the kind of interaction that it is meant to support, that object is said to “afford” its behavior[28]. Figurines, or more generally, physical objects that represent characters, afford the specification of gross character motion, primarily because most people have manipulated figurines to “animate” them for play purposes.
Figure 2. Electric football field. The inset image shows a single player figurine on a typical plastic base.

In the particular case of football, the game of electric football presents an inspiring interface. The user positions figurines on a miniature electric football field. The figurines have simple handles for pre-specifying their desired direction of motion. When switched on, the field vibrates, causing the players to move across the surface. Offensive and defensive players collide and move forward until progress of the ball is stopped and the play is finished (Figure 2). Another example of a childhood toy that reinforces our experience with figurine manipulation is “the toy town” (Figure 3). One could imagine using such figurines to populate and animate pedestrian simulation scenes.

We aim to take advantage of the natural interaction afforded by figurines for digital animation purposes. Our goal in this paper is to begin to explore user interaction methods for specifying animated content in a tangible setting. While most previous work in tangible interfaces has focused on representing and interacting with static content, we are interested in character motion. Interacting with dynamic characters will require that we address new concerns such as motion, task, and timing specification within the tangible setting.

3 Background

The idea behind tangible interfaces is that we live in two worlds, physical and digital, and we are adept at working in the physical world but need better ways of interacting in the digital world. The tangible interface approach blurs the line between the two worlds by designing physical interfaces for manipulating and interacting with digital content.

Some of the earliest work in tangible interfaces focused on using physical objects to aid in the creation of 3D digital geometry[1,2,16,17,18,19]. Since this early work, others have revisited this problem[3,4,5,9,31]. Physical interfaces have also been used for intuitive manipulation of existing 3D digital objects such as medical models or virtual environment props[20,30] and as an intuitive means for navigating complex information spaces[11,21].

Planar desktop physical interfaces have also been an area of great interest. An early desktop interface that merged real and digital paper and content on a physical desk was presented by Wellner in 1993[35]. In 1995, Fitzmaurice introduced the “bricks” concept for manipulating digital content on a planar surface using physical handles[14]. Ullmer and Ishii presented models for tangible user interfaces in the metaDESK project. The metaDESK explored the mapping of traditional graphical user interface elements such as windows, icons, menus, etc. to tangible user interface elements such as the lens, phicons, trays, etc.[32]. Our approach builds upon the “bricks” concept by representing characters with bricks such as physical blocks. Particular attention in desktop physical interfaces has been paid to spatial applications such as urban planning and factory simulation which are especially suited to planar interaction [6,7,10,15,23].

We are particularly interested in designing animated motion using a physical interface. Early tangible interfaces for 3D articulated character animation can be traced to the Monkey device[12]. The Monkey is essentially a small scale skeleton of a human that is instrumented with rotation sensors that provide joint angles to an animation package. Anderson and his colleagues present techniques for modeling articulated characters using physical clay. The character is then scanned, matched to template of character models, and
Figure 4. An example of a football play. Various shapes have been used for each position for clarification only.

Figure 5: The Coach’s Table system.

animated using physical simulation[4]. More recently, Oore and colleagues presented a desktop device for creating performance animations of an articulated character[29]. Johnson et al. introduce sympathetic interfaces to describe an instrumented plush toy that is manipulated to control an interactive story character[24]. Our work differs from previous work in the animation area in that we are not concerned with articulated character motion, but with using physical icons for designing the gross motion of multiple characters that potentially interact with one another on a planar surface.

4 Football Strategy Domain

In this paper, we address the particular problem of specifying motion for the sport of American football. In American football, there are 22 players on the field at any one time: 11 offensive players and 11 defensive players. Defensive motion is typically reactive to the offensive strategy while offensive motion is typically well choreographed by a coach. We will focus on offensive motion and strategy specification.

A coach typically designs a playbook using symbols to represent players, paths and tasks. The four primary player positions are lineman, running back, receiver, and quarterback. Each player performs one of three general tasks based on his position: run, catch, or block. Linemen are given blocking assignments. Blocking is the act of obstructing the forward progress of a defensive player. Running backs are typically given blocking assignments or running assignments. A running assignment is a specification of where to run on the field. A receiver is typically given a pass route assignment. A pass route defines where the receiver should run and what path the receiver should take in order to get to that location and receive the ball thrown from the quarterback. The quarterback plays the most pivotal position. The quarterback initiates each play and either hands the ball to a running back, runs himself, or throws the ball to a receiver (Figure 4).

Although several software packages exist for designing football plays, to our knowledge, none of them have incorporated animated motion or 3D views. Furthermore, none of them have focused on interfaces beyond the mouse that can facilitate natural two-handed or multiple person interaction. This system is designed not only for specifying static play descriptions, but dynamic player motion. Our system takes advantage of a user’s experience with figurines by providing a tangible interface for specifying paths, tasks, and important timing information including game events, task duration, and relative timing of interactions. The system, the ‘coach’s table’ is composed of three parts: vision/display system, figurines and buttons interface, and 2D/3D motion playback.

5 Vision/Display System

We use an infrared vision system similar to that of the metaDESK to track objects on the table[32]. An infrared (IR) lamp is positioned above the table and provides infrared light coverage for the table surface. Objects on the surface are marked with highly reflective tape in specific geometric shapes. A monochrome camera with an IR pass filter produces a clear image of the object markers. The illuminating light of the IR does not interfere with the projected image and most importantly, the projected image does not interfere with the object tracking (Figure 5).
The vision software is built on top of the Intel OpenCV library[22]. The images are captured at 30hz using a standard video capture card. Each image is processed to produce a binary image of the reflected geometric shapes placed on the bricks. The vision system is calibrated by placing a set of three reflective bricks on three known points on the field. These three points are used to solve a linear system of six equations for the affine transformation from the camera image to the projected image.

Offensive players are represented with wooden blocks marked with reflective circles. When a player is selected and moved, the system identifies which object was moved, tracks the movement of the centroid of the object, and updates the path information for that player. Aside from objects on the playing field, the system must recognize interaction with the application interface buttons. A button is represented by a projected rectangular button area marked with a label. This implementation is similar to the trays approach[14,32]. To select a button, the user simply places a brick in the area of the projected button.

A projector is mounted horizontally in the ceiling and a mirror redirects the image to the table. For the football application, the projector produces an image of the football field, player representations, and interface elements. The display and vision systems run as two separate processes on two separate machines. The two processes communicate via a TCP/IP socket connection.

Figure 6. The interface consists of five areas with the bulk of the interface occupied by the field of play. The physical interface is approximately 46 in. x 36 in.

Figure 7. Top-view diagram of the five areas of the Coach's Table interface.

6 Interface Elements

There are several key tasks a user should be able to perform for the creation of animated football content.

- Player formation–initial positions
- Player selection
- Motion path and task specification
- Time duration specification
- Timed event specification
- Animation playback

In the following sections, we will discuss the interface design for facilitating these interactions.

6.1 Button Layout

The interface is divided into five main areas (Figures 6 and 7). To the right side of the field are the interface buttons for playback and editing. These elements are not specific to football, but are necessary for general animated content creation. To the left of the field are the football specific interface buttons divided into two areas: tasks and timing. The task area is reserved for tasks that may be assigned to any particular player, such as blocking. The timing area is reserved for various forms of timing specification including player speed, time events, and time durations. The bulk of the display is occupied by the playing field.

In the button areas, a brick’s position is not tracked until it is stationary for a fraction of a second, allowing users to move bricks within view of the camera without accidental tracking. On the field of play, the same is true until a player has been selected. Once selected, all motion is tracked until the player is deselected.
6.2 Players and Motion
The character interface consists of wooden blocks representing players and serving as our bricks. The bricks can be used to position players on the field, select players, deselect players, move players, and assign player tasks and timing information through the interface buttons.

To put offensive players on the field, the user first selects the “offense” button. The bricks then behave like rubber stamps, depositing players at locations on the field where the bricks are momentarily placed. Players are assumed to be oriented in the forward direction initially and oriented along the path of motion as the play progresses.

Play specification involves selecting players and assigning paths and tasks to each of them. Users select a player by placing a brick at the player’s location on the field. When selected, the player is highlighted with a color change. The user specifies player paths by moving the bricks across the field of play as shown in Figure 1. The path is projected onto the field as a dotted line. To deselect a player, the user must cover the brick and remove it from the field of play.

6.3 Timing Specification
Animation requires that the user be able to specify timing information. There are four types of timing specifications in our system – time durations, events, relative timing, and speed. The desired time duration can be specified for a particular task using a time slider. The slider icon consists of a projected slider with values from 0 to 5 seconds. A typical football play lasts on average 4.5 seconds. The user specifies a time duration by placing a brick on the desired value. The selected duration is highlighted with an indicator bar (Figure 8).

We chose the slider approach as opposed to an absolute time clock such as that used in Urp[23] or a keyframe approach because football coaches are more accustomed to specifying durations. For example, a coach might tell a player to block for two seconds, then run a particular pass route.

Another important timing specification involves timed events. In football, there is one key time event: the ball “snap”, or “hike”, in which the center lineman hands the ball to the quarterback. Limited motion can occur before the ball is hiked and the remainder of the motion must not begin until after the ball is hiked. In this system, the football hike is noted by placing a brick on the “hike” button. Any motion specified before this action is assumed to take place before the hike. Any motion specified after this action is assumed to take place after the hike.

The ability to specify events such as the hike is particularly important because it allows the user to create interactive content. If only time durations are allowed, essentially all motion is pre-scripted in time. If, on the other hand, the system can recognize and annotate events, then we can later allow events to trigger motion in an interactive environment. For example, imagine the play content being used in a 3D immersive quarterback training environment. The quarterback could use a voice command to represent the “hike”, providing the event and triggering motion of the players.

Coaches must also be able to specify that certain motions happen before others. For example, players sometimes cross paths on the field of play and the coach must be able to specify which player crosses first. In our system, when there is a potential player contention in space, motions are assumed to happen in the order in which they were specified. If two players cross paths, the player to first reach the point of contention during path specification will be assumed to reach it first during motion generation. This allows for a very natural specification of crossing paths where the user simply moves the two players about the field exactly as he would expect them to move making sure that the correct player passes the contention point before the other player.

Finally, users must be able to specify a player’s speed. One could specify velocity using a keyframing approach, but coaches and naive users in general are not
accustomed to this type of speed specification. In football plays, motion is typically specified to be full (fast) speed or half (slow) speed. The system therefore allows the user to tag player paths with fast or slow speeds using an interface button. The default speed is fast.

### 6.4 Task Specification

Users must also be able to specify tasks such as blocking. Users specify the blocking task by placing a brick on the “block” label. Any task selection is applied to the currently selected player at the current time in the motion specification.

All football specific interface events such as path start, path stop, block, time durations, path contentions, hike, fast, slow, etc. generate play events that are stored in a list for the currently active player. To generate motion, the system executes events on each player’s list. Certain events, such as the hike, are waiting events and require that a player wait until all other players reach their hike event before it can continue. This approach allows us to generate motion as a series of high-level discrete events where the continuous low-level timing and motion details are determined by the underlying motion simulation. Currently, the underlying motion simulation for each player is a simple point mass simulation. In section 9 we discuss other options for the low-level motion simulation.

### 7 2D/3D Animation Playback

Once the user has completed the design of the play, he can then review the final motion using the play review buttons, “Play” and “Rewind”. Motion can be viewed in either 2D or 3D. The mode is chosen by placing a brick on the appropriate label. In 2D mode, the field is shown directly from above and players are represented as circles. In 3D mode, players are represented as six foot tall cylinders and the field is viewed from a position above and behind the quarterback (Figure 9).

### Table 1: General questions posed to the seven subjects. The subjects responded on a 5 point scale with 1 being “Very” and 5 being “Not Very”.

<table>
<thead>
<tr>
<th>General Questions</th>
<th>Ave. Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>How well do you understand the game of football?</td>
<td>2.57</td>
</tr>
<tr>
<td>How easy was it for you to learn to use this system?</td>
<td>2.14</td>
</tr>
<tr>
<td>How natural was the interface?</td>
<td>2.14</td>
</tr>
</tbody>
</table>

### Table 2: Average scores for several questions about the football specific components of the interface.

<table>
<thead>
<tr>
<th>Football Specific Task Questions</th>
<th>Ave. Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>How easy was it to place players on the field?</td>
<td>1.43</td>
</tr>
<tr>
<td>How easy was it to select players?</td>
<td>1.86</td>
</tr>
<tr>
<td>How easy was it to specify paths?</td>
<td>1.29</td>
</tr>
<tr>
<td>How easy was it to specify tasks, such as block?</td>
<td>2.71</td>
</tr>
<tr>
<td>How easy was it to specify time durations?</td>
<td>2.71</td>
</tr>
<tr>
<td>How easy was it to specify time events such as hike?</td>
<td>2.43</td>
</tr>
</tbody>
</table>

### Table 3: Average scores for several questions about the playback and editing specific components of the interface.

<table>
<thead>
<tr>
<th>Playback and Edit Questions</th>
<th>Ave. Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>How easy was it to fix mistakes or modify a play?</td>
<td>2.85</td>
</tr>
<tr>
<td>How easy was it to review the play?</td>
<td>1.29</td>
</tr>
<tr>
<td>How easy was it to view the play in 3D?</td>
<td>1.0</td>
</tr>
</tbody>
</table>

### 8 Usability Studies and Discussion

We have performed studies to determine the usability of the system and to help us identify and resolve flaws in the design. Research suggests that five subjects are sufficient to identify shortcomings in an interface and to show that an interface is usable[34]. We observed a total of seven randomly chosen subjects as they completed seven tasks with the interface. The experiments exercised the various tasks a user or coach might need to perform in order to define and review an interactive football play. The final task included designing a complete play for 11 offensive players.

The very first task was designed to determine if the block figurines afforded character motion specification. Before any introduction to the system, the users were shown the interface with a single block on the field of play. The user was then shown a diagram of a receiver pass route and asked the following question: “The block on the table represents a football player. How would you tell the player to run this route?” All seven subjects grasped the block and moved it about the field of...
play to trace out the passing route. This result reinforced our belief that the figurine metaphor naturally affords character animation.

The subjects were then given a short tutorial and allowed five minutes to practice with the interface. They were then given six more tasks to complete. All seven subjects completed all seven tasks and a post experiment questionnaire. Overall, the subjects seemed comfortable with the interface and several commented on how easily they were able to create fairly complex timed motion (Table 1). Table 2 shows responses for several questions about the football specific tasks that the users were asked to complete. Table 3 shows the responses for questions regarding the playback and edit capabilities of the interface. All responses were on a 5 point scale with 1 being “Very” and 5 being “Not Very”.

Through these studies, we were able to identify the key shortcomings of the interface. First, users were not satisfied with the tools for modifying plays. Our system currently only allows the user to undo a player completely and restart that player’s motion from scratch. Other possible options include the ability to undo the last motion specification (path, task, time, etc.) as well as tools for modifying a player’s path – similar to drawing tools for specifying curves.

Some users also found it difficult to remember the correct order of operations. For example, to assign a blocking task to a player, the user must first select a player and then select the task to apply to that player. Several users executed these subtasks in the opposite order, quickly realizing their mistake and correcting the action.

Another common problem was trying to move players without selecting them first. Other problems were due to the vision system. Users sometimes inadvertently covered the bricks while specifying motion, interfering with the tracking. Users also sometimes forgot to cover the bricks to deselect a player before removing it from the field. These problems suggest that a more robust tracking system or an alternative tracking system, such as a multitouch surface, might be more usable.

One task produced results that were particularly interesting. The subjects were asked to specify the passing routes for two players in which the routes crossed paths. One player was specified as crossing before the other. We observed that several users grasped the two players simultaneously and specified their motion in parallel, resolving the possible collision by sliding the bricks across each other in the desired order. This natural use of both hands for specifying interactions between multiple players is promising and illustrates one advantage of tangible interfaces for multiple character motion specification.

9 Future Work

Usability studies indicate that naive users can quickly design interactive football plays and visualize them using our interface. Although we have experienced success with our initial implementation, there are still several areas for improvement and further study.

Currently, the system is designed for basic play specification and simple visualization. Ultimately, we would like this tool to be useful for creating high quality 3D animated character motion. For example, a sports broadcaster may want to create a 3D play for demonstration to the television audience. To create compelling content, she will need to specify paths and routes for players as well as behaviors such as blocking, jumping, tackling, etc. The system should then generate compelling 3D character motion. We can currently generate 3D motion by matching the desired player path to the paths of several 3D motion capture sequences and choosing the nearest sequence (Figure 10). A more general solution that involves the application of recent advances in motion graph algorithms are under development[8,25,26,27]. This approach would allow the user to specify animation constraints such as key poses, time durations, paths, and behaviors. These constraints would guide the selection of a series of motion clips that most closely fulfill the user’s constraints. An im-
portant aspect of 3D content creation that we have not addressed is flexible camera motion. To create compelling 3D content for training and visualization, the user should be allowed to specify general camera views and possibly camera motion paths. One option for specifying camera views is the active lens approach [32].

Our descriptions to this point have involved single users. Single users can use both hands to specify player positions and motion well as manipulate the interface buttons. This implies that multiple users could be seated at the coach’s table to specify player motion. One possible problem is contention for the interface buttons. For example, if one user wanted to apply a block to the currently selected player, then the other user would have to deselect his players so that the blocking task would not be applied to them. One solution to this problem is to give each user their own unique access to the buttons either through separate button trays or through uniquely identified bricks for selecting buttons. A more elegant solution would be to embed interface components within the brick attached to a player using an approach similar to the flipbricks described by Fitzmaurice [13]. For example, each player could be represented by a block where each face of the block represented a different player task such as block, jump, or dive. To apply a task to a player, the user would simply flip the brick to show the correct task on top. This approach would of course be limited by the number of “sides” on the brick. Another option would be to use voice input combined with identification. For example, “offense, block” would denote the blocking task for the selected offensive player.

We are currently using blocks as a metaphor for a character figurine. In one sense, this is a reasonable abstraction that users are able to understand as is evidenced by the first task presented in our usability studies. On the other hand, a figurine with sensing capabilities might provide more expressive input modes. For example, imagine a figurine with articulated arms and legs and a system that could recognize various poses. The user could put the character in a blocking pose to specify the blocking task. This approach would only be useful if the poses were very easy to specify.

We are also interested in getting feedback from experts in the football domain. The football team at our university has expressed interest in our goals of providing tools for naive users to create 3D content. After completing several modifications based on our usability studies we hope to test the system on members of the coaching staff at the university.

In conclusion, we have presented a table-top tangible interface for specifying motion for multiple characters on a football playing field. The interface is coupled with a 2D simulation for visualization of the final scene. We believe that such interfaces will prove extremely useful in bringing animated content creation to the naive user and in facilitating the collaborative design of simulation and training content.

Acknowledgements
We would like to thank Electronic Arts for supplying us with motion capture data.

References


[27] Metoyer, R., Ph.D., Building Behaviors With Examples, Georgia Institute of Technology, 2002


