Eurographics/ ACM SIGGRAPH Symposium on Computer Animation (2012)
P. Kry and J. Lee (Editors)

Finger Walking: Motion Editing with Contact-Based Hand Performance

Noah Lockwood[1,2] and Karan Singh[1]

[1]University of Toronto
[2]Industrial Light & Magic

Abstract

We present a system for generating full-body animations from the performance on a touch-sensitive tabletop of "finger walking", where two fingers are used to pantomime leg movements. A user study was conducted to explore how users can communicate full-body motion using their hands, which concluded that finger walking is a naturally-chosen and comfortable performance method. Based on contact data recorded during this study, the properties of a variety of performed locomotion types were analyzed to determine which motion parameters are most reliable and expressive for the purpose of generating corresponding full-body animations. Based on this analysis, a compact set of motion features was developed for classifying the locomotion type of a finger performance. A prototype interactive animation system was implemented to generate full-body animations of a known locomotion type from finger walking by estimating the motion path of a finger performance, and editing the path of a corresponding animation to match. The classification accuracy and output animation quality of this system was evaluated in a second user study, demonstrating that satisfying full-body animations can be reliably generated from finger performances.

Categories and Subject Descriptors (according to ACM CCS): I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Animation

1. Introduction

Creating full-body animations can be an imposing task for novice animators, since animation software often require a large number of parameters such as positions and joint angles to be specified at precise times. One approach to address this problem is *performance interfaces*, which utilize the timing of a user's actions. Performance interfaces have the potential to be particularly accessible, since realistic timing of motions can be generated not only from the intentional timing of a performance, but also implicitly by the physical constraints on a user's motion.

We conducted an exploratory user study to examine how motion can be communicated through hand motions, by having users perform a number of motion types in whatever manner they preferred. A majority of users chose to use "finger walking", where the middle and index fingers of the dominant hand are used to pantomime the leg motion of a full body (Figure 1). However, analysis of the touch-sensitive tabletop data collected during this study has shown that seemingly-precise and expressive finger motions can not only possess inconsistent motion parameters, but can differ significantly from the corresponding full-body motions.

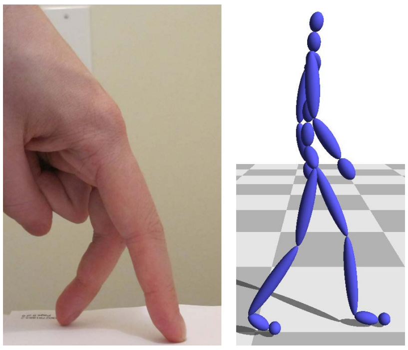

Figure 1: *Finger walking (left) is a natural and expressive way to communicate full-body locomotion (right).*

DOI: 10.2312/SCA/SCA12/043-052

Based on this first study, our goal was to create an interactive animation system that is controlled by finger walking, but is tailored to the particularities of how users perform finger walks. We have focused on using contact data as user input, since interaction with a touch-sensitive tabletop is entirely passive and does not impede hand motion, as an instrumented glove or motion capture markers would. However, directly mapping the contact positions and timing from a finger walking performance to an animated character could result in unrealistic movement due to the inconsistencies made clear by our study analysis. Therefore, we take a *gestural* approach instead, where our system modifies preexisting full-body animations to match the general spatial parameters of a finger motion, without matching the performance so closely that the results appear unrealistic.

Our system can operate on a user's performance in two ways. The motion type of a user's finger performance can be determined by summarizing the performance in a compact feature vector, and comparing it to the feature vectors of other performances with known types. Given a particular locomotion type, our system estimates the central path of the motion from its "footprints" and edits a full-body animation of the corresponding type to match this path.

We conducted a second study to evaluate the accuracy of our motion type classification and user satisfaction with the motions generated from their finger performances. We concluded that our system can reliably classify the locomotion type of a new user's input, as well as generate animations matching the intent of the user's performance.

Our work makes the following contributions:

- An analysis of how users express full-body motion using finger walking.
- A finger walking animation performance interface, using feature-based motion classification and path-based editing to generate full-body animations matching the user's gestural input.
- A novel kinematic path estimation algorithm which determines a motion path based on only a set of ordered footprint positions.

2. Related Work

Performance interfaces utilize a user's physical interactions over time to generate character animation. These interfaces are particular appealing for novice animators who may not be able to specify their intentions, but can perform them. Directly mapping partial or full-body motion to a character can be referred to as "computer puppetry"; Sturman [Stu98] provided a history of the technique, including contemporary collaborative approaches. To reduce the intrusiveness of measuring devices, some recent performance techniques measure only selected parameters of full-body motion and reconstruct entire poses based on examples [CH05, LWC*11]. Less direct control is also possible, by mapping only the important parameters of the user's pose onto the character [SLSG01], or by augmenting the user's motion with physical simulation of environmental or object interaction [IWZL09, NWB*10].

Hand performance can be an effective and simple means of communicating full-body motion, and is used in both amateur performances as well as advertisements. Hand performance has also been applied to animation generation, by mapping finger joint angles to joints in a character skeleton [SZ93, LZK04], or by using fingertip positions as constraints on foot positions [WP09]. These previous approaches have directly mapped finger joint angles or positions to a character's pose, which can result in inconsistent animation; our technique treats user performances as suggestive instead, to ensure smooth and realistic animation.

Handheld input devices can also be manipulated in real-time to form performance interfaces. The two-dimensional motion over time of a mouse can be used to re-time animations [TM04] or control particular aspects of a character's pose [LvdPF00, IMH05b]. Objects with more degrees of freedom can also be used for expressive performances, such as paper cutouts which can be positioned and rotated [BJS*08], a tilt-sensitive stylus [Osh04], or a 3D-tracked object [DYP03]. Paired objects can also be manipulated simultaneously to control particular bones in a character's skeleton [OTH02] or foot positions [KPL05].

Multitouch devices provide another simple and effective way to control animation parameters in real time, by dragging multiple portions of a deformable character [IMH05a] or by utilizing specialized interfaces with asymmetrical controls for each hand [KN10]. Like our approach, Sugiura et al. [SKW*09] utilized a multitouch interface not for direct manipulation, but for finger walking; however, their approach was specialized to pre-programmed gestures on a handheld device, while our locomotion type classification is example-based and utilizes a larger multitouch tabletop to allow user control over a motion's path.

Utilizing a touch-sensitive device for finger walking requires considering finger contacts as foot contacts, which have been used in a variety of ways for animation purposes. Footstep positioning is important for planning realistic character motion through an environment [CLS03] as well as simulating robust walking behavior [CBYvdP08]. User-specified footprint positions can also be used to generate full animation of both bipeds [vdP97] and quadrupeds [TvdP98]. Yin and Pai [YP03] used a pressure-sensitive device to allow interactive foot contact-based control over a character's pose and motion; while we use feature-based classification and motion editing in a similar way, our approach compensates for the inaccuracies of finger walking and is applied to locomotion rather than stationary motions.

3. Exploratory Study

We conducted an initial exploratory study to identify ways that users would choose to communicate full-body motion using their hands. Our goal was to identify a preferred and comfortable method of interaction, to inform the design of an animation interface.

3.1. Methodology

The study was designed to passively measure hand performances with minimal intrusion to the participants, in order to maintain all possible performance strategies or techniques. To accomplish this, a Microsoft Surface touch-sensitive table was used to record any surface contact that occurred, and a video camera recorded above-surface motions and gestures (Figure 2). As the Surface display was inactive to avoid encouraging surface-based interaction, a secondary monitor was placed beyond the Surface to display instructions. Participants were presented with a scripted introduction which explained the purpose of the study and the equipment being used. They were also informed that they could express motion in any manner they chose within the "capture volume" above the Surface, from any direction, as there was enough space to interact with the Surface from any side.

The study consisted of two stages. In the first "freeform" stage, names and brief descriptions of locomotion types (such as walking or running) were presented, and the participants were asked to perform these motions without any spatial instructions, such as directionality. In the second "mimicked" stage, short motion-captured animations varying in locomotion type and direction (such as walking and turning or running forward) were played, instead of presenting written descriptions. The participants were asked to communicate the general type and direction of the animation, without specifically trying to capture particular details, such as the number of steps. Participants could restart their performances and replay animation clips at any time. After deciding when a performance was complete, participants rated their satisfaction with their performance on an integer scale of 1 ("very unsatisfied") to 5 ("very satisfied"). After the final performance, an image of a participant's hands placed palm-down was captured using the Surface, and a short survey of follow-up questions was administered.

3.2. Results and Observations

The study was performed by 12 volunteer participants (10 male, 2 female), each performing 28 motions (10 in the first stage, 18 in the second stage). The methods of performance were as follows:

- 73 percent employed "classic" finger walking on the Surface, with the middle and index fingers of the dominant hand pantomiming the leg movement of the corresponding full-body motion.
- 12 percent were by pantomiming upper-body motion above the Surface, such as hands swinging back and forth during a walk.
- 10 percent used both hands to pantomime foot movement, with either fists or open palms alternately "stepping" on or above the Surface.
- 5 percent used a single abstraction of solely full-body position, such as a single finger or clenched fist, on the Surface.

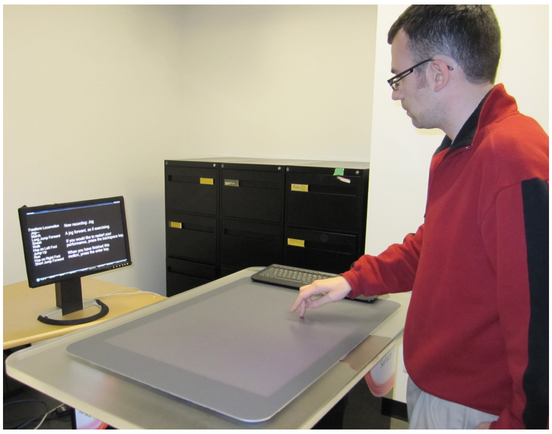

Figure 2: *Equipment used to study hand performances. Directed by instructions on the external monitor, participants were free to move around the Microsoft Surface, which captured contact data, while their above-surface motions were video recorded.*

While some participants experimented with multiple techniques, 11 out of 12 used finger walking at least once. Overall, participants were more satisfied with the finger walking performances; the average rating of finger walking motions was 3.7 out of 5, compared to 3.3 out of 5 for all other types of performance. During follow-up questions, most users indicated that their choice of finger walking was because motion of the lower body was more important to their performance, and they did not need to incorporate upper body motion to communicate locomotion type.

Therefore, finger walking seems to be both a natural and comfortable choice for communicating full-body motion through hand performance, and we closely examined and analyzed the contact data gathered during only the finger walking performances.

3.3. Data Analysis

The first method of examining finger walking contact data was to statically visualize all of the contacts which occurred during a single performance. Contact information is provided by the Surface as ellipses with associated times; a single contact consists of a sequence of ellipses, representing contact shape from initial contact to lift-off. Since these contacts can be sampled at irregular rates, for the purposes of visualization, linear interpolation was used to re-sample all ellipse parameters (position, rotation, major and minor axes) at 30Hz; an example of this visualization for a finger walk, with a detail of a single contact, is shown in Figure 3.

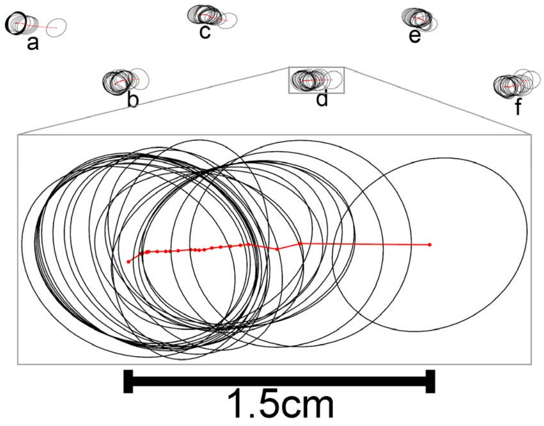

Figure 3: *Contact shapes for a forward finger walk (top), sampled at 30Hz; steps occurred in order a-f. An enlarged view of contact d (bottom), with ellipse center path in red, shows how contacts generally begin with a stomp, followed by an accelerating roll forward.*

Most finger walking contacts during forward motion do not consist of a smooth and constant roll in the direction of the motion, but instead generally begin with a "stomp" down with low forward velocity, and roll forwards while accelerating. It was not unusual to observe finger contacts which slip just before liftoff, most likely due to reduced friction when a user's fingernail is the predominant contact. Finger contact shape also appears to be unreliable and does not have a consistent orientation, such as elongation relative to the direction of travel, and the path of each contact's center is generally forward, but noisy.

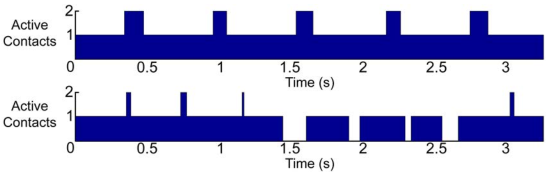

Figure 4: *The number of active contacts over time during a full-body walk (top) is consistent, with regular, brief periods of double-contact. The number of active contacts during a finger walk (bottom), however, can be significantly inconsistent while the motion may still appear appropriate.*

There are other expressive parameters of finger motions which can be compared to the equivalent parameters of corresponding full-body motions. For example, the number of active contacts - i.e., fingers or feet in contact with the ground - can be examined as it changes over time. During a full-body walk, the majority of time is spent on one foot, with regular brief periods of double-stance after the swing foot lands but before the stance foot lifts off. However, since a finger walking hand is not propelled by ground contact, but rather by the movement of the arm or upper body of the performer, the number of active contacts over time can be significantly inconsistent even for finger motions which appear "correct". Figure 4 compares the active contacts over time of a motion-captured full-body walk and a freeform finger walk which was rated 5 out of 5 by its performer.

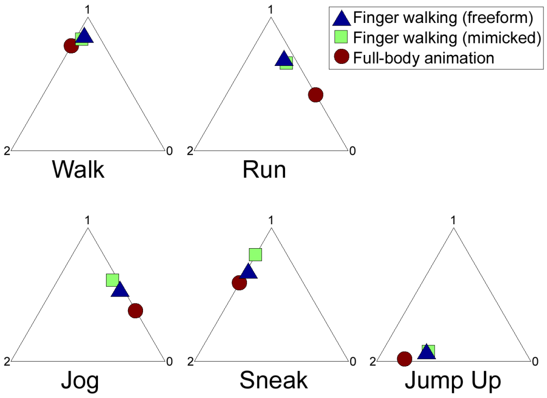

Figure 5: *Average proportion of time spent during a motion with 0,1, or 2 active contacts, plotted by barycentric coordinates, for a variety of locomotion types. The relationship between freeform and mimicked finger walking and a corresponding full-body motion differs depending on locomotion type.*

Examining the number of active contacts *overall* throughout an entire motion, rather than moment-to-moment, can be more robust to temporary variations. Figure 5 shows ternary plots of the proportion of time spent during a number of locomotion types with zero, one, or two active contacts, for the freeform performances, mimicked performances, and the example motion capture clips. The full-body motions almost entirely consist of two contact states, for example, either zero or one contact for running, and one or two contacts for walking. The finger motions, however, usually contain a more significant proportion of the third contact state, such as zero contacts during walking. It also appears that finger motions are biased towards single-contact states relative to the full-body motions; however, the magnitude of this bias seems to be locomotion type-dependent.

There are many other parameters which can be used to compare finger motions to full-body motion. Figure 6 shows the average contact frequency for a number of locomotion types, for the performed and example motions. For some

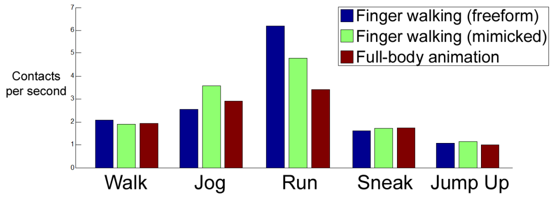

Figure 6: *Contact frequency for a variety of motion types. Again, the relationship between freeform and mimicked finger walking and a corresponding animation differes depending on locomotion type.*

types, the contact frequency is similar. However, running contacts were performed significantly more rapidly than full-body running, even during the mimicked motions where an example had been played. The frequency of jogging contacts was initially under-performed and then over-performed during the mimicked motions.

3.4. Discussion

Based on the analysis of finger motion contact data, it appears that even when performing motions to their satisfaction, precise parameters of user performances such as contact shape, position, and state can be imprecise and unreliable. Examining more robust overall characteristics yields greater consistency - however, the correspondence between finger and full-body motions appears to be highly dependent on the motion type.

Therefore, a method of creating full-body animation by directly mapping finger motion parameters would need to handle this varying abstraction in order to generate plausible results. Instead, given the inconsistencies in user inputs, we hypothesize that users treat this type of performance as illustrative - demonstrating general characteristics such as motion type and overall direction, without concentrating on either the consistency or realism of specific motion and contact parameters, and aim to develop a finger walking interface to accommodate this.

4. Implementation

To accommodate illustrative input by the user, we propose a *gestural* interface to generate animation which match high-level characteristics of a finger motion, rather than from the finger motion directly. This system consists of two components that operate on a user's finger motion. The locomotion type of a new user's performance can be automatically determined, based on the previous users' motions of known types. Given a particular locomotion type, an appropriate full-body animation can be selected and edited (Figure 7), without replicating finger motion inconsistencies which would make it look unrealistic. This approach of classification and gestural motion editing from a single input is similar to the Motion Doodles work of Thorne et al. [TBvdP04], which used continuous sketched paths from a "side view" instead of finger contacts.

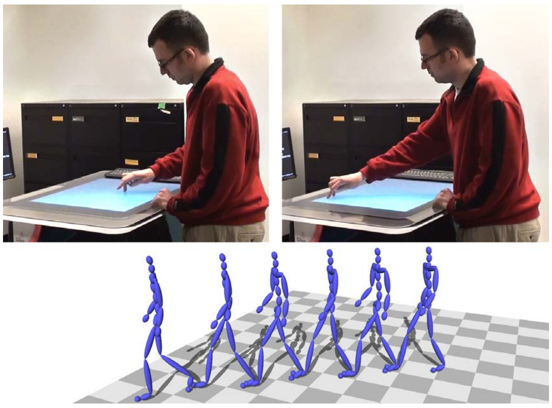

Figure 7: *A finger motion performed on our prototype system (top) and the edited full-body animation which was automatically generated as output (bottom).*

4.1. Feature-Based Classification

Since finger motions can vary in their similarity to full-body motion depending on the locomotion type, attempting to classify the locomotion type of a new finger motion by direct comparison to full-body motions could prove unreliable. Therefore, we have chosen an example-driven approach to classify new finger motions based solely on previously-recorded finger motions. Given a finger motion of unknown type, a feature vector is determined, which can be compared to feature vectors from existing finger performances of known types, from a variety of users.

Every feature must be valid for all locomotion types, and should produce similar results for motions of the same type which differ in the number of steps. Furthermore, these features should be robust to the types of potential irregularities examined in Section 3.3, such as inconsistent contact state or finger slips, or even missing steps entirely.

Our feature vector consists of six features of a finger motion:

- The contact time proportions for zero, one, and 2 active contacts,
- Average velocity,
- Average stride frequency,
- Average contact-to-contact distance,

where distance between contacts is measured from the more reliable first point of contact.

The accuracy of a number of standard classification techniques were tested using this feature vector, with accuracy ranging between 65 and 80 percent. A full analysis of classifier accuracy on a larger dataset is included with the results of our second study in Section 5.2.

4.2. Data Normalization

The animation system should accommodate not only a variety of locomotion types, but a variety of users as well. While the feature vector is based on contact positions and timing, which are not explicitly user-dependent, there is one very important way in which differences among users can potentially affect the results: finger length.

Compensating for subject size is a common consideration in biomechanics applications such as gait analysis, where the consequences of different heights can be significant; for example, natural walks for two people of different heights will vary in velocity and stride length. To address this, gait velocity in particular can be rendered dimensionless - effectively, normalized - using the Froude number [VO05], a nonlinear quantity which relates relative stride length (stride length divided by leg length) to velocity by a nonlinear equation. This is appropriate because morphologically similar humans of different sizes are affected by a gravity force proportional to their mass, which does not vary linearly with height.

We are inspired to take a similar approach by normalizing our distances as well, using finger length instead of leg length. While free-standing humans require nonlinear normalization, we hypothesize that linear normalization is sufficient. This has the effect of modifying the spatial units of measurement so that, for example, average finger motion velocity is measured not in centimeters per second, but finger lengths per second.

The Surface can be used to quickly and automatically measure finger lengths. The intensity images captured by the Surface (Figure 8, left) are thresholded at a standard value to produce a binary image (Figure 8, middle). Fingertips on the largest blobs in the image are detected using the technique of Malik et. al [MRB05], and the clockwise ordering of the fingers relative to the thumb identifies the dominant hand. While the lengths of the index and middle fingers are different (Figure 8, right), applying the lengths appropriately to different "steps" would require identifying the finger of a particular contact, which could be complex and/or unreliable. Therefore, all distances are normalized by the average finger length.

Gait analysis also calculates step-to-step distance in a different way; rather than measuring the direct displacement between step positions, only the "forward" distance is considered. To approximate this, a novel path estimation technique (described in Appendix A) is used to determine a central path for a finger motion based on the ordered contact positions. The contact-to-contact distance can then be measured in path arc-length between the contacts' projected positions along the path.

Figure 8: *While the Surface can provide approximate depth images (left), appropriate thresholding yields a hand shape (middle) which can be automatically analyzed to identify and measure fingers (right).*

A comparison of classification accuracy using all combinations of these variations - absolute versus relative units, and direct versus path step distances - is presented in Section 5.2.

4.3. Full-Body Motion Editing

Given a particular locomotion type, the goal of our system is to generate a corresponding full-body animation matching some aspects of the user's performance. This is accomplished by editing "canonical" animations: short animation clips representing each locomotion type, which can be automatically looped to generate a smooth animation of any number of steps. However, editing the canonical animations to closely match particular characteristics of a finger motion could result in unrealistic animation, given the inconsistencies observed in finger motions (Section 3.3). Therefore, user input is treated as gestural - essentially, as instructions of the form of *"do this, there"* - and the canonical animation is edited to match the broader spatial parameters of a finger motion, in the form of its motion path.

Canonical animations are edited using the path-based editing technique of Lockwood and Singh [LS11], which identifies a sparse set of editing handles along the path of an animation. Specifying new handle positions automatically modifies the path and poses of the entire animation to match, while maintaining similar motion timing. After identifying the motion path of a finger motion using our path estimation technique (Appendix A), the editing handles of a straight-ahead canonical animation can be automatically placed along the finger motion path, resulting in a new animation with a very similar path to that of the user's performance.

There are two remaining degrees of freedom in this process: the *animation scale* and the *number of cycles* for the canonical animation to be edited. One possible method for determining scale is to attempt to match the average step size of the input finger motion and the animation. Unfortunately, this could result in animated characters of significantly different sizes being generated even for a single user. Instead, based on feedback from participants during our exploratory study, a single scale is determined per user such that the leg

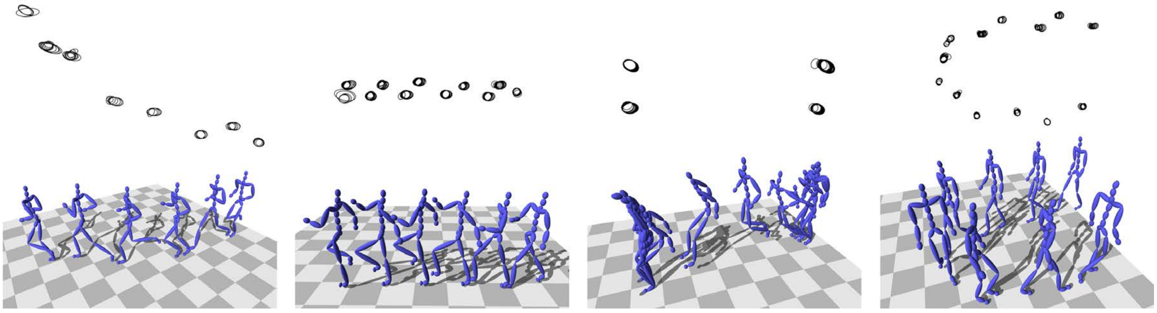

Figure 9: *Sampled finger contacts (top) and resulting animations (bottom) from separate participants in the performance study. Left to right: running and turning, marching, jumping forward, and a walking u-turn.*

length of the animated character is equal to the user's finger length (shown in Figure 1), which results in consistent character size across locomotion types and performances.

To determine the appropriate number of cycles of a scaled animation, the layout of the editing handles along the path must be considered. With a fixed scale, arranging the handles to match aspects of the performance could result in erratic motion or unrealistic gait parameters, such as unnaturally long strides. Instead, in keeping with treatment of user input as gestural, the canonical motion is cycled a sufficient number of times to lay out the editing handles along the entire path while maintaining the original distance between each pair of handles, i.e., "bending" the animation without "stretching" it. This will very nearly maintain the consistency of stride length in the output animation; however, one potential downside of this method is that the animation's timing may be significantly different from the user's performance.

5. Performance Study

To evaluate the finger walking animation system, a second study was conducted with a new set of participants, who performed finger motions and rated the resulting animations.

5.1. Methodology

The equipment for this study was identical to the first study (Section 3.1). However, one practical difficulty of this study was the need to display an animation to the user which corresponded to their finger motion, and we felt that displaying this motion solely on a secondary monitor would make this correspondence difficult for users to evaluate. While an augmented reality system to display the animated character in the space above the surface would have been ideal, it was not practical. Instead, we opted for a multi-view approach: after each performance, the resulting animation was played from a pre-set 3/4 view on the secondary monitor, as well as from an orthographic top-down view on the Surface itself, in spatial correspondence with the performance.

Participants were instructed to use only the middle and index fingers of their dominant hand to pantomime motions on the Surface. After scanning their hands to determine finger length, a first "freeform" stage of trials instructed the participants to perform different locomotion types based strictly on description. The second "directed" stage used pre-set lanes displayed on the Surface combined with locomotion type descriptions on the secondary monitor, to generate motions with particular path shapes. After each finger motion was completed, the output animation was played once automatically, and could be replayed by the participant any number of times before they rated their satisfaction with how well the animation represented their intended motion, on a scale from 1 ("very unsatisfied") to 5 ("very satisfied"). There were 5 cyclic locomotion types (walk, job, run, sneak, march) and 3 non-cyclic (jump up, short jump, long jump), with motion path shapes which were either straight ahead or with a single turn of 45, 90, or 180 degrees.

5.2. Results and Discussion

8 new participants (5 male, 3 female, all right-handed) volunteered for this study, and each performed a total of 21 motions (8 freeform, 13 directed). Some performances and resulting animations are shown in Figure 9. Classification accuracy of locomotion type was evaluated afterward, using leave-one-subject-out cross-validation on the feature vectors of the finger walking motions from both studies. A number of standard classification techniques were used: k-Nearest Neighbor (with $k = 1,3,5,7$), Mahalanobis distance (used for motion classification by Yin et al. [YP03]), and Support Vector Machines using both linear and radial basis function kernels. Figure 10 shows the accuracy of classifiers using feature vectors calculated with either absolute or relative units, and direct or path-based contact distances (Section 4.2).

Classification accuracy ranged from 62 to 74 percent. Within a particular classifier, accuracy was generally improved by using path-based instead of direct contact-to-contact distances, and relative instead of absolute units, but

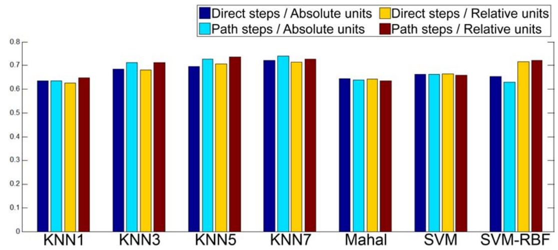

Figure 10: *Locomotion type classification accuracy for all classifiers and varieties of feature vectors.*

not by a large margin (typically 1-2%). We believe that this demonstrates the usefulness of our selected motion features and data normalization methods for reliable classification, which can only be improved by the application or development of more specialized classification algorithms, which we leave to future work.

Locomotion type also had an effect on classification accuracy. Figure 11 shows the average and standard deviation of accuracy for each locomotion type across all classifiers and feature vectors. The accuracy is significantly lower for the more "vaguely-named" locomotion types (jog, march, sneak), where greater variation in performance parameters can cause mis-classifications; for example, one user's performance of a march may be very similar to another's walk.

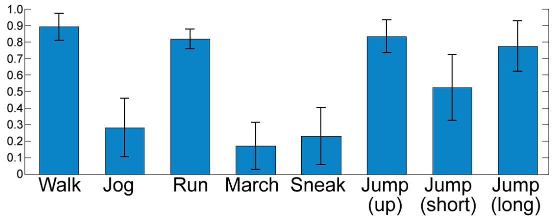

Figure 11: *Classification accuracy for individual locomotion types, combined from all classifiers and varieties of feature vectors.*

Participants rated the generated animations very high in satisfaction, with an average of 4.67 out of 5. Most participants did not comment about any discrepancies between their performances and the animations, indicating that the treatment of the input as gestural was appropriate. However, after the study was completed, one participant remarked that the animations seemed generally faster than their performances, while one other participant said the opposite: that animations seemed generally slower than the performances. To examine the potential discrepancies in timing between the performances and the animations, the ratios of the animation and performance durations and contact frequencies were examined, as shown in Figure 12.

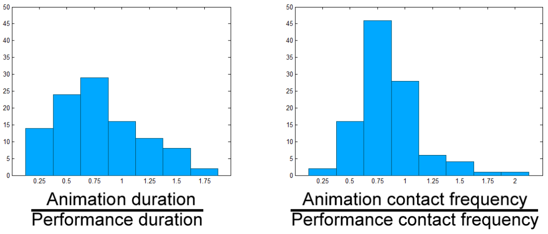

Figure 12: *Histograms of the ratio between duration (left) and contact frequency (right) for each generated animation and its corresponding performed finger motion.*

Overall, animations were slightly shorter in duration than performances, which could indicate that character scale was overestimated, since a smaller character moving at the same relative speed would take longer. The contact rate of the animations was also slightly slower. However, given the high user satisfaction ratings, we are cautious about altering the animation timing to more closely match; this disparity is mostly accounted for by the tendency to over-perform the contact frequency of running to an unrealistic degree, as observed in our first study (Figure 6).

6. Conclusions

We have presented a user-centered approach to the development of an interactive animation system based on finger walking on a touch-sensitive tabletop. An exploratory user study was conducted which identified finger walking as a natural and comfortable method of communicating full-body motion through hand performance. Analysis of this data and its characteristics led to the development of a gestural performance interface using feature-based classification of finger motions and path-based editing of corresponding full-body motions. This system was evaluated in a second user study, which found that motions can be reliably classified using the feature vector, and satisfying animations can be generated to match user performance.

Our current system has certain limitations which could be explored in future work. Our focus on locomotion allows sufficient expression through surface contact, but additional information such as hand position, orientation, and pose could be very useful for expanding the types or styles of motions which could be recognized.

A system controlled by finger walking could potentially generate motion in a number of ways different from our canonical motion editing approach. For example, motion class specification and a motion path can be used to splice new motions together using motion graphs [SH07]. Online motion generation is also possible, which would allow the user to adjust their technique during their performance.

Finger walking interfaces could also be extended beyond our method. Different types of surface interaction could be studied, to examine, for example, whether finger walking in a "treadmill" style on a smartphone screen [SKW*09] is similar to freeform finger walking on a touch-sensitive tabletop. Including additional features like a second hand (perhaps to indicate upper body motion, or additional legs for non-bipeds) or useful props, such as the finger shoes favored by amateur finger performers online, could extend this interface even further.

Acknowledgements

We thank the members and alumni of the Dynamic Graphics Project lab, especially Prof. Khai Truong, Dustin Freeman, John Hancock, and Martin de Lasa, for essential discussion and feedback. Many thanks as well to Isabelle Ortega, for invaluable and tireless support. We are also grateful to the anonymous reviewers, and to our study participants. All motion capture clips were from the Carnegie Mellon University motion capture database.

Appendix A: Closest Points of Approach Path Estimation

An approximation of a central motion path can be useful for a variety of purposes, such as motion analysis and categorization, and path-based motion editing. Our method approximates the motion path from a sequence of contact positions by attempting to project the contact positions onto the path, which are then interpolated by a C2 natural cubic spline.

The projected positions are determined by combining a series of local solutions of the "Closest Points of Approach" problem: given two contact positions p and q and their approximate lateral directions u and v, the projected positions are the closest points along the lines $p(t) = p + t \cdot u$ and $q(t) = q + t \cdot v$ (Figure 13).

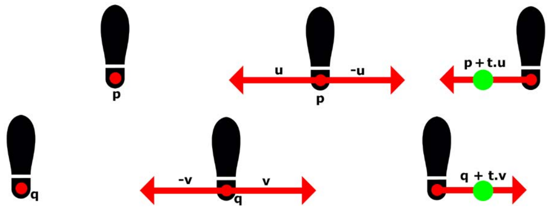

Figure 13: *Solving for the Closest Points of Approach using contact positions (left) and their corresponding lateral directions (middle) allows the projection of the footprints onto the motion path to be estimated (right).*

To apply this to contact position projection, we assume that the two steps are laterally equidistant from the central path, i.e., that $||u|| = ||v||$. One lateral direction is negated if necessary, to ensure that the closest points of approach are "between" the contact positions (Figure 13). The full process is described in Algorithm A.

Algorithm 1 CPA(p, q, u, v)

```
Require: $||u|| = ||v||$
if $sgn(u \cdot (p - q)) = sgn(v \cdot (p - q))$ then
    $u \leftarrow -u$
end if
$t \leftarrow \frac{(p-q)\cdot(u-v)}{||u-v||^2}$
return $(p + t \cdot u, q + t \cdot v)$
```

The CPA process in Algorithm A is applied to each pair of subsequent contact positions, yielding one estimate for the projected positions of the first and last contacts; for all interior contacts, the projected position is calculated as the average of its two corresponding estimates. Once the positions are determined, a C2 natural cubic interpolating spline can be fit to the projected positions, approximating the motion path. The lateral directions at each contact position are determined from the normal vector at the nearest point along the spline path; the initial lateral directions are determined from a temporary spline which interpolates the midpoints between pairs of subsequent contacts.

This path can be iteratively refined by repeating the process, using more accurate lateral directions from the current path to determine a new path. In all of our uses, this procedure is stable and converges. We have found a useful termination criteria to be when the maximum change in lateral direction between iterations drops below a certain threshold; we used a value of 1 degree. For forward paths, the process usually converges after 1 iteration; for more complex paths, 3 iterations is the usual maximum. Some results for finger motion paths are shown in Figure 14.

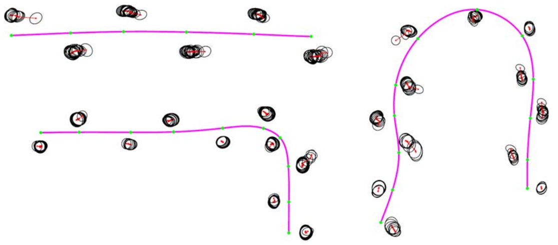

Figure 14: *Our Closest Points of Approach algorithm can quickly determine a path from footprints for a variety of motion types and path shapes, such as forward walking (top left), sharp turns (bottom left), and 180-degree turns (right).*

One assumption of this algorithm is that contacts occur in a generally "forward" direction. While it has not been extensively tested on more erratic motions with significant back-and-forth components such as dancing, this presents a problem even during forward motions when double stances occur (Figure 15, left). A simple solution to this problem is to "merge" contacts which overlap significantly in time into a single contact position at their midpoint, and to constrain the corresponding path position to this point (Figure 15, right). We have found that merging contacts when the duration of the overlap is at least 50% of the duration of either contact is a suitable criteria for generating appropriate paths.

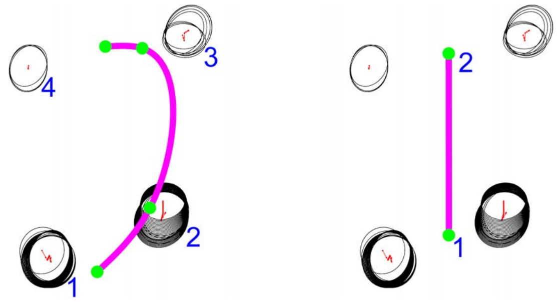

Figure 15: *Imposing a strict ordering on contacts can generate inappropriate paths for double-stances (left), which is addressed by merging contacts which overlap significantly in time (right).*

References

[BJS*08] Barnes C., Jacobs D. E., Sanders J., Goldman D. B., Rusinkiewicz S., Finkelstein A., Agrawala M.: Video puppetry: a performative interface for cutout animation. In *ACM SIGGRAPH Asia 2008 papers* (2008). 2

[CBYvdP08] Coros S., Beaudoin P., Yin K. K., van de Pann M.: Synthesis of constrained walking skills. In *ACM SIGGRAPH Asia 2008 papers* (2008). 2

[CH05] Chai J., Hodgins J. K.: Performance animation from low-dimensional control signals. In *ACM SIGGRAPH 2005 Papers* (2005). 2

[CLS03] Choi M. G., Lee J., Shin S. Y.: Planning biped locomotion using motion capture data and probabilistic roadmaps. *ACM Trans. Graph. 22*, 2 (2003). 2

[DYP03] Dontcheva M., Yngve G., Popović Z.: Layered acting for character animation. In *ACM SIGGRAPH 2003 Papers* (2003). 2

[IMH05a] Igarashi T., Moscovich T., Hughes J. F.: As-rigid-as-possible shape manipulation. In *ACM SIGGRAPH 2005 Papers* (2005). 2

[IMH05b] Igarashi T., Moscovich T., Hughes J. F.: Spatial keyframing for performance-driven animation. In *SCA '05: Proceedings of the ACM SIGGRAPH/Eurographics symposium on Computer Animation* (2005). 2

[IWZL09] Ishigaki S., White T., Zordan V. B., Liu C. K.: Performance-based control interface for character animation. In *ACM SIGGRAPH 2009 papers* (2009). 2

[KN10] Kipp M., Nguyen Q.: Multitouch puppetry: creating coordinated 3d motion for an articulated arm. In *ACM International Conference on Interactive Tabletops and Surfaces* (2010). 2

[KPL05] Kolhoff P., Preuss J., Loviscach J.: Walking with pens. In *Eurographics 2005 Short Presentations* (2005). 2

[LS11] Lockwood N., Singh K.: Biomechanically-inspired motion path editing. In *SCA '11: Proceedings of the ACM SIGGRAPH/Eurographics Symposium on Computer Animation* (2011). 6

[LvdPF00] Laszlo J., van de Panne M., Fiume E.: Interactive control for physically-based animation. In *ACM SIGGRAPH 2000 papers* (2000). 2

[LWC*11] Liu H., Wei X., Chai J., Ha I., Rhee T.: Realtime human motion control with a small number of inertial sensors. In *I3D '11: Proceedings of the symposium on Interactive 3D graphics and games* (2011). 2

[LZK04] Lam W.-C., Zou F., Komura T.: Motion editing with data glove. In *ACE '04: Proceedings of the ACM SIGCHI International Conference on Advances in computer entertainment technology* (2004). 2

[MRB05] Malik S., Ranjan A., Balakrishnan R.: Interacting with large displays from a distance with vision-tracked multi-finger gestural input. In *UIST '05: Proceedings of the ACM symposium on User interface software and technology* (2005). 6

[NWB*10] Nguyen N., Wheatland N., Brown D., Parise B., Liu C. K., Zordan V.: Performance capture with physical interaction. In *SCA '10: Proceedings of the ACM SIGGRAPH/Eurographics Symposium on Computer Animation* (2010). 2

[Osh04] Oshita M.: Pen-to-mime: A pen-based interface for interactive control of a human figure. In *Eurographics Workshop on Sketch-Based Interfaces and Modeling* (2004). 2

[OTH02] Oore S., Terzopoulos D., Hinton G.: A desktop input device and interface for interactive 3d character animation. In *Graphics Interface* (2002). 2

[SH07] Safonova A., Hodgins J. K.: Construction and optimal search of interpolated motion graphs. In *ACM SIGGRAPH 2007 papers* (2007). 8

[SKW*09] Sugiura Y., Kakehi G., Withana A., Fernando C., Sakamoto D., Inami M., Igarashi T.: Walky: An operating method for a bipedal walking robot for entertainment. In *ACM SIGGRAPH Asia 2009 Emerging Technologies* (2009). 2, 9

[SLSG01] Shin H. J., Lee J., Shin S. Y., Gleicher M.: Computer puppetry: An importance-based approach. *ACM Trans. Graph. 20*, 2 (2001). 2

[Stu98] Sturman D.: Computer puppetry. *IEEE Computer Graphics and Applications 18*, 1 (1998). 2

[SZ93] Sturman D. J., Zeltzer D.: A design method for "whole-hand" human-computer interaction. *ACM Trans. Inf. Syst. 11*, 3 (1993). 2

[TBvdP04] Thorne M., Burke D., van de Panne M.: Motion doodles: an interface for sketching character motion. In *ACM SIGGRAPH 2004 Papers* (2004). 5

[TM04] Terra S. C. L., Metoyer R. A.: Performance timing for keyframe animation. In *SCA '04: Proceedings of the ACM SIGGRAPH/Eurographics symposium on Computer animation* (2004). 2

[TvdP98] Torkos N., van de Panne M.: Footprint-based quadruped motion synthesis. In *Graphics Interface* (1998). 2

[vdP97] van de Panne M.: From footprints to animation. *Computer Graphics Forum 16*, 4 (1997). 2

[VO05] Vaughan C., O'Malley M.: Froude and the contribution of naval architecture to our understanding of bipedal locomotion. *Gait & Posture 21*, 3 (2005), 350–362. 6

[WP09] Wang R. Y., Popović J.: Real-time hand-tracking with a color glove. In *ACM SIGGRAPH 2009 papers* (2009). 2

[YP03] Yin K., Pai D. K.: Footsee: an interactive animation system. In *SCA '03: Proceedings of the ACM SIGGRAPH/Eurographics symposium on Computer animation* (2003). 2, 7