Perceptual evaluation of footskate cleanup

Martin Pražák, Ludovic Hoyet and Carol O’Sullivan

Graphics, Vision and Visualisation Group, Trinity College Dublin.

Abstract

When animating virtual humans for real-time applications such as games and virtual reality, animation systems often have to edit motions in order to be responsive. In many cases, contacts between the feet and the ground are not (or cannot be) properly enforced, resulting in a disturbing artifact known as footsliding or footskate. In this paper, we explore the perceptibility of this error and show that participants can perceive even very low levels of footsliding (<21mm in most conditions). We then explore the visual fidelity of animations where footskate has been cleaned up using two different methods. We found that corrected animations were always preferred to those with footsliding, irrespective of the extent of the correction required. We also determined that a simple approach of lengthening limbs was preferred to a more complex approach using IK fixes and trajectory smoothing.

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

1. Introduction

Animated virtual characters are an essential part of virtual reality (VR) applications and computer games. While their appearance has reached impressive levels of realism, producing natural animations has proven to be a tougher problem. Inevitably, practically all marketed applications contain a certain amount of animation artifacts, with one of the most common being footsliding, i.e., incorrect contacts between the feet and the ground. While its causes can be diverse, ranging from noise gathered in the motion capture pipeline, through errors introduced during motion editing [BW95, WP95, Gle98], to shortcuts used to increase characters’ responsiveness [MP07], the visible result is always very similar—movement of the feet that is inconsistent with expected foot-ground interactions. As a consequence, footsliding is a very common problem in computer animation, especially in games and interactive applications where motions are edited on the fly. For this reason, the detection and correction of this problem have been addressed in previous work. However, the perceptual saliency of the artifact itself, or of the changes made to correct it, has not been explored to date.

In this paper, we aim to fill this important gap. In our first experiment (Section 4), we determined minimal perceivable footsliding thresholds, assuming that the observer...
is aware of its presence. We showed that participants can perceive even very low levels of footsliding (<21mm in most conditions), especially when environment cues highlight the artifacts. In the second experiment (Section 5), we introduced footsliding at levels that are clearly perceivable, and corrected them using two methods (Kovar et al.’s method [KSG02] and lengthening of body segments alone, as in Harrison et al. [HRP04]). As each of these methods alter the original animation in certain ways, they can introduce side effects and artifacts into the original motion. By asking users to compare the corrected motions with uncorrected ones, we showed that corrected animations are always preferred to animations with footsliding, independently of the correction level required. Results also showed that participants considered animations corrected with the body segment lengthening method to be of higher quality than those edited using Kovar’s method. Our results provide valuable insights for developers of games and VR applications by providing thresholds for the perception of footsliding artifacts, as well as a visually effective method for correcting this disturbing artifact.

2. Background

Constraint-based motion synthesis and motion editing form a large group of methods for altering or generating realistic animations. In these methods, the most common constraint is a footstep – a constraint ensuring the spatiotemporal relation between a foot and the ground during the contact phase. The solution to this constraint usually involves a combination of inverse kinematic and root displacements [LS99, KMA05].

One of the most comprehensive methods to-date was presented by Kovar et al. [KSG02], where footskate cleanup was conducted by successive steps of character root displacement, root trajectory smoothing, leg inverse kinematics and leg segment lengthening. A necessary prerequisite for any cleanup method is the knowledge of the footsteps’ position and timing. While explicitly known for procedural animation methods, data-driven approaches require either manual specification of this information or a constraint detection step [LB06, GBT06, IAP06].

However, editing of animations can create artifacts which can affect the perception of its naturalness. Perceptual properties of these artifacts depend on many parameters, such as the character’s representation [HOT98, CHK07, MJH+07], the number of character models or animations [MLD+08], or the artifact’s type [RAP08]. The perceptual effects of a variety of anomalies in human animation was also recently studied by Hodgins et al. [HJO+10].

Closely related to constraints correction and the perception of motions artifacts, Harrison et al. [HRP04] studied to what extent the lengths of a virtual character’s links can be changed without the viewer being aware of this change. They found that length changes of over 20% can go unnoticed when the attention of the ser is not focused on the affected link. This is a finding that could be very useful for footskate cleanup.

However, little is known about the perception of animations with footsliding or of animations where this artifact has been corrected. In this paper, we aim to study the perceptibility of footsliding in an animation, and to compare the perception of the quality of the resulting animation after correcting the artifact using limb lengthening [HRP04] or the method presented by Kovar et al. [KSG02].

3. Stimuli

Because our goal is to ask participants to judge the quality, naturalness and artifact levels of motions, we need to ensure that the original motions (before alterations based on the experiment objective) are of high quality, natural and artifact-free. Motion capture technology can provide such motions. For our experiments, we selected a set of motions from our database (83 actors, with a variety 100+ motion types per actor). First, based on the biometric data, we selected 10 actors (5 male and 5 female) with normal body builds. Second, for each actor we selected a straight walk with speed close to the actor’s average walking speed. Finally, we carefully checked that the selected motions did not contain any visible artifacts.

A classical representation of a motion clip relates motion data to the character’s skeleton, which represents his/her body structure. While real-world applications would use more realistic skinned human models, which have been shown to convey more information than simpler models [HOT98], one of the sources of a diversity of animation artifacts is motion retargeting, i.e. adapting the motion to a new skeleton (e.g., a skeleton of a displayable character) [KSG02]. In order to avoid these alterations, we chose to display the motions on virtual mannequin figures, generated to accurately reflect the morphology of the original actors (Figure 2). These mannequins offer a better representation than stick figures, as they convey spatial information, while avoiding the problem of retargeting inherent in realistic skinned characters.

One of the disadvantages of motion capture technology is, due to its physically-based nature, that a certain amount of artifacts will be always present in the captured motions (Figure 1, top), caused by measurement noise and the capturing/processing pipeline. To ensure that we started with perfect animations, we first detected footstep constraints using a method based on the work of LeCallennc and Boulic [LB06] and then cleared any residual foot motion using the method of Kovar et al. [KSG02]. As we did not perform any retargeting step, the footskate correction on captured motions caused only minimal changes (Figure 1, middle). The parameters of the detection algorithm were manually adjusted for each animation, and the resulting constraints checked to ensure their accuracy.
4. Baseline Experiment

In the first experiment, our aim was to determine the saliency of footsliding for worst-case scenarios, when participants are aware of its presence or when it is emphasized by the properties of the environment. We hypothesized that participants would be able to perceive even a small amount of footsliding and that environment cues can significantly highlight these artifacts.

4.1. Footsliding

In order to allow precise control of the footsliding, we modelled its effects by altering the translational component of the root trajectory. Our model consists of two distinct components relative to a character's heading direction – the root trajectory. Our model consists of two distinct components relative to a character's heading direction – the root trajectory. Our model consists of two distinct components relative to a character's heading direction – the root trajectory. Our model consists of two distinct components relative to a character's heading direction – the root trajectory. Our model consists of two distinct components relative to a character's heading direction – the root trajectory.

Front footskate was introduced by multiplying the forward directional movement of the root by a speed coefficient. This resulted in speeding-up (i.e., gliding) for values above one, or slowing-down (i.e., moonwalking) for values below one (with no change when the value was exactly one). Over all the motions used for this experiment, an increase or decrease of 10% of the root velocity corresponded to an average foot position change of 74.2±1.3mm per foot step. The amount of footsliding in mm is linearly related to the speed coefficient.

Side footskate was introduced by adding a perpendicular displacement proportional to a sinus function (with period consistent with that of the animation and phase determined by the midpoint of the left foot constraint). The neutral element in this case was zero, while positive values led to “outwards” footskating (i.e., the constrained foot was moving away from the character’s root projection on the ground) and negative values led to “inwards” footskating.

This approach ensured that only smooth changes were introduced into the original motions while providing a controllable way of introducing footsliding artifacts similar to the ones seen in VR applications (speeding up, slowing down, or blending between straight and curved locomotions).

4.2. Experiment Design

The experiment was divided into two main blocks. In the first block (grid environment), we tested a worst-case scenario where a grid texture on the ground provided guides for accurate footskate detection (Figure 3, top row). The grid was white on a blue background with interline distances of 10cm. The second block (neutral environment) used a plain blue ground without any texture (Figure 3, bottom row). Each block was further divided into two parts (Part 1 and Part 2). Part 1 tested frontwise footsliding with motions viewed from the side (Figure 3, right column), while Part 2 tested side-wise footsliding with motions viewed from the front (Figure 3, left column). This choice of viewport/footsliding combinations ensured that we tested the worst-case scenarios (where the footsliding direction was always perpendicular to the view direction), because the viewing direction can play an important role in perceived levels of footsliding.

To accurately determine the perceptual threshold for footsliding, we used an adaptive double staircase experiment design [Cor62]. Our design was a standard Yes-No (a variant of 2-alternative forced choice) with fixed up and down steps, as described by Garcia-Perez [GP01]. We used a down/up step ratio of 0.871, converging at 52.38% of the psychometric curve, and set the stopping condition to 20 reversals.

In each trial, participants observed one of the ten motions, randomly selected, with the introduced amount of footsliding corresponding to the current state of the staircase experiment. They were asked to determine if the character is footsliding, and to answer by pressing the “yes” or “no” button, clearly marked on the computer’s keyboard. The description of what footsliding was is provided both formally on the instruction sheet and informally by the experimenter. There was no explicit time limit and the adaptive nature of the method used did not allow to determine the length of the experiment precisely beforehand, but none of the participants took more than 20 minutes. Because of the complexity of the task, each participant performed at most two parts and received a book voucher for their efforts.

Nine participants took part in the grid environment block of this experiment (2F, 7M), with five participants for grid Part 1 and four participants for grid Part 2. Fourteen participants took part in the neutral environment block (5F, 9M), with eight participants for neutral Part 1 and six participants for neutral Part 2. For this and all subsequent experiments, all participants were naïve to the purpose of the experiment and came from various educational backgrounds.

4.3. Results

For each experiment block/part, we used Matlab psignifit toolbox [WH01] to fit a logistic psychometric curve into the
data, both to each participant and to the overall merged results. A psychometric function represents how the participant’s response to stimuli varies depending on the variation of these stimuli. The results of the evaluation are presented in Figure 4, which shows both ratings per-participant and overall psychometric curve. The overall point of subjective equality (PSE, point where participants are equally likely to find the stimulus acceptable or otherwise) and just noticeable difference (75% JND, minimum increase of the PSE stimulus value needed to be detectable 75% of the time) are summarized in Table 1. As expected, the results show that the accuracy of footsliding detection is greatly improved due to the visual cues present in the grid environment. Furthermore, the levels of footsliding perceived in the grid environment (especially for the side case) suggest, that the limit is approaching the display resolution of the screen. Our interpretation is that, effectively, any footsliding, no matter how small, will be perceived in an environment with visual cues.

However, results are different for the neutral environment. In the case of front footsliding, the results showed that participants are more disposed to perceive footsliding caused by slowing-down rather than speeding-up, with a speed co-
efficient PSE of 0.971 and 1.115 respectively. These values corresponded to respective errors of approximately 21mm and 84mm. In the case of side footsliding, a lower threshold with PSE < 20mm was found.

Table 1: PSE and JND for baseline experiment. Absolute values are given in metres; relative values as a ratio of the motion speed-up/slow-down.

<table>
<thead>
<tr>
<th>Case</th>
<th>Grid environment</th>
<th>Neutral environment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Speed coef.</td>
<td>Footsliding amount [m]</td>
</tr>
<tr>
<td>gliding</td>
<td>0.984</td>
<td>0.007</td>
</tr>
<tr>
<td>moonwalking</td>
<td>0.021</td>
<td>0.010</td>
</tr>
<tr>
<td>sliding outwards</td>
<td>0.971</td>
<td>0.034</td>
</tr>
<tr>
<td>sliding inwards</td>
<td>-</td>
<td>-0.005</td>
</tr>
</tbody>
</table>

5. Footskate Cleanup Experiment

Our previous experiment showed that the perceptual threshold for footsliding is relatively low, especially when environment cues are present to highlight its effect. However, it is still unclear whether footsliding should be corrected in all cases, or if sometimes remedial action can actually deteriorate the quality of the motion even further. For this reason, in our second experiment we compared animations with footsliding to animations that had been corrected. We hypothesized that possible artifacts might play an important role in a participant’s responses to the corrected motions, and might even lead to them preferring the original animation despite footsliding being present.

5.1. Footskate Cleanup

Considering the importance and frequency of footsliding artifacts, surprisingly low number of methods have been proposed to correct them. For our experiment, we selected two basic methods commonly employed in real-time systems. The first and more comprehensive method was introduced by Kovar et al. [KSG02]. In this method, footskating is cleaned up using successive steps of character’s root displacement, root trajectory smoothing, leg inverse kinematics and leg segment lengthening. The second and simpler method employs only lengthening of the character’s limbs to reach the desired end-effector position (which corresponds to the last step of Kovar’s algorithm). While the first method is much more sophisticated, and therefore is expected to perform significantly better, simple limb lengthening might provide acceptable results, as hinted by work of Harrison et al. [HRP04] (under certain conditions, limb lengthening of up to 20% might go unnoticed by the observer). In the experiment, these two methods will be compared both against each other and with uncorrected animations. In further text, these three options will be described as Kovar’s correction (K), Lengthening correction (L) and uncorrected (U).

5.2. Experiment Design

In VR applications, footsliding in the direction of the character’s movement is a necessary price for responsiveness, and therefore is more common than side footsliding. For this reason, in this experiment we focused on this aspect alone. Our results should generalize to side footskate as well, but that needs to be confirmed in future work.

Twelve participants volunteered for in this experiment (3F, 9M). Seventy-two animations were presented in randomized order, with factors: 2 directions (gliding and moonwalking) \times 3 footsliding levels (50%, 75%, 99%) \times 3 comparisons (U-K, U-L, K-L) \times 4 repetitions. The 3 footsliding levels, selected according to the results of the Baseline experiment, were 50%, 75% and 99%, which represent an average footsliding amount of 84mm, 111mm and 196mm respectively for gliding and 21mm, 30mm and 60mm respectively for moonwalking.

In each trial, participants viewed two characters display-
ing the same motion with the same level of introduced footsliding, differing only in the correction method (uncorrected, Kovar’s or lengthening). The character displayed on the left was always blue and the right one always orange, matching two coloured keys on the keyboard. In each displayed pair, a random selection was made to decide which of the characters will display which motion. Participants were instructed to select the character displaying the most natural motion, without any additional information about the bodypart or artifact they should focus on.

In the previous experiment, the perception of front footsliding was displayed from the side view, as it gave the most information about the front sliding motion of the feet. In this experiment, though, the artifacts are present both in the feet and in the character’s overall movement. Because camera angle can influence the perception of both aspects differently, we presented a canonical viewpoint, giving as much information as possible about both the overall quality of the motions and the footsliding (Figure 5).

5.3. Results

The design of our experiment inherently contains two distinct sets of data that cannot be analysed in a single step. Our analysis reflects this by providing separate results and interpretation for each case. The first data set includes a comparison of corrected motions displayed together with the uncorrected ones (cases U-K and U-L). The second is used to directly compare the two correction methods when displayed simultaneously (K-L).

To correct or not to correct? To evaluate participants’ preference between corrected and uncorrected motions, we extracted a subset of our data that contained the answers from trials when uncorrected motions (U) were displayed simultaneously with motions corrected using either of the correction methods (K or L).

We performed a three-way repeated measures ANOVA with within-subjects factors comparison (U-K and U-L), direction (moonwalking and gliding) and footsliding level (50%, 75% and 99%). In this evaluation (and in all subsequent evaluations), the post-hoc analysis was performed using standard Newman-Keuls test for comparison of means. The results showed a main effect of comparison \( (F_{1,11} = 7.376, p < 0.05) \), where footsliding animations corrected with the lengthening method were preferred to uncorrected animation more often than animations corrected with Kovar’s method (Figure 6). The results also showed a main effect of footsliding level \( (F_{2,22} = 3.737, p < 0.05) \), where corrected animations were more likely to be preferred with higher levels of the footsliding (Figure 7). There was no effect of direction.

The results also show that, overall, the corrected motions were considered as better quality in more than 70% of the trials, independent of the level of introduced footsliding. To explore this fact further, we carried out single t-tests on collapsed data to determine if there was a real preference of correction methods over non-correction (i.e., preference significantly different from 50%). The data was collapsed over direction and level factors, as we were interested in the overall preference of each correction method against non-correction. The results showed that both correction methods were preferred to non-corrections in significantly more than 50% of the cases \( t(11) = 2.946, p < 0.05 \) for Kovar’s corrections and \( t(11) = 8.666, p < 0.00001 \) for Lengthening ones. From this result we can conclude, that correcting footsliding should always be preferred over not correcting it, even if it implies large changes of the motion.

Figure 6: The main effect of comparison factor on trials where corrected motions (K or L) were presented together with uncorrected ones (U).

Figure 7: The main effect of footsliding level factor on trials where corrected motions (K or L) were presented together with uncorrected ones (U).
Preferred correction method? The analysis above shows that corrected animations are preferred to uncorrected ones, with the Lengthening method ranked on average higher than Kovar’s method. The second subset of our data allows for a more direct comparison between the two correction methods when displayed side-by-side.

We performed a two-way repeated measures ANOVA with within-subjects factors direction and footsliding level on the preference of L over K. Results showed no significant effect of direction or footsliding level. We then averaged participant ratings over these two factors and carried out a single t-test to determine if one of the correction methods was significantly preferred to the other (i.e., preference significantly different from 50%). The results showed that the body segments lengthening method was ranked higher than Kovar’s approach 72% of the time ($t(11) = 5.693, p < 0.0005$, Figure 8).

6. Discussion and Future Work

In this paper we have presented a set of experiments addressing the perceptual effects of footsliding in VR applications. Our results have shown that participants can perceive even very small levels of this artifact, with a trend towards accepting higher levels of footsliding in the direction of character’s movement. Moreover, we have shown that the ability to see the footsliding is increased even further by visual cues of the environment, such as a grid texture on the ground, up to the point bordering with the screen resolution.

Out of the two tested correction methods, users preferred the simpler lengthening method. This can imply, that the perceptual threshold for limb lengthening is relatively high (a finding consistent with the work of Harrison et al. [HRP04]), and that the artifacts introduced by Kovar’s method are perceptually more salient. However, attention would also affect the user’s preference, as Harrison et al. [HRP04] demonstrated that limb length changes of over 20% can go unnoticed when the user’s attention is not focused on the extending limb.

In our experiments, the stimuli character was a wooden mannequin figure. While a well-justified choice for our scenario, future work include to evaluate how these results would scale on more realistic characters, as they can prove to be even less tolerant towards both footsliding and correction artifacts. Also, a general perceptual evaluation of limb lengthening on highly realistic humanoid characters is yet to be performed.

The viewpoint and camera angle can also play a significant role in footsliding perception. In our experiments, we have chosen two approaches – in the first experiment, we used the worst-case scenario, placing the camera in direction perpendicular to the introduced footsliding. While a valid choice for our goal, a different angle might prove more forgiving. In the second experiment, we have used the “canonical” viewpoint. This choice was motivated by the fact that in this experiment it is not only footsliding that we are evaluating, but also overall artifacts introduced by correcting it. A different camera angle might mask either of these artifacts, shifting the user’s preference towards one or another.

To conclude, future research on the perception of motion artifacts (with one being the footsliding) will involve high quality human characters, along with the influence of environment, camera angle, camera motion, and the effects of the user’s attention.

Acknowledgments

We wish to thank the reviewers for their comments, and the participants in our experiments. This work was sponsored by Science Foundation Ireland as part of NaturalMovers and Metropolis projects.

References


© The Eurographics Association 2011.
M. Pražák & L. Hoyet & C. O’Sullivan / Perceptual evaluation of footskate cleanup

[GP01] Garcia-Pérez M. A.: Yes-no staircases with fixed step sizes: psychometric properties and optimal setup. Optometry and vision science official publication of the American Academy of Optometry 78, 1 (2001), 56–64. 3


[IAF06] Ikemoto L., Arikan O., Forsyth D.: Knowing when to put your foot down. Proceedings of the 2006 symposium on Interactive 3D graphics and games SIG3D 06 1, March (2006), 49. 2


© The Eurographics Association 2011.