Proceedings of ITI 08 International Conference of Information Technology Interfaces, Cavtat/Dubrovnik, Croatia. IEEE Xplore.

Modeling Prehensile Actions for the Evaluation of Tangible User Interfaces

Georgios Christou
European University Cyprus
6 Diogenes St., Nicosia, Cyprus
gchristou@acm.org

Frank E. Ritter
College of IST
Penn State University
University Park, PA 16802, USA
frank.ritter@psu.edu

Robert J. K. Jacob
Tufts University
161 College Ave., Medford, MA 02155, USA
jacob@cs.tufts.edu

Abstract. *Prehension, or reaching-to-grasp is a very common movement performed by users of Tangible User Interfaces (TUIs) because through this movement users can manipulate tangible artifacts. Here we present an experiment that provides evidence towards the hypothesis that prehensile movements can be modeled based on the amplitude of the prehension movement. We then explore consequences of this evidence on the modeling and evaluation of TUIs using tools that can predict task completion times, such as Goals, Operators, Methods and Selection Rules (GOMS), as well as the implications for TUIs compared to Direct Manipulation Interfaces.*

Keywords. Tangible User Interfaces, Prehension, Fitts' Law, Reaching-To-Grasp.

1. Introduction

Leaps in technology recently have made it easier to interact with data in a tangible way. Tangible User Interface (TUI) [9, 22] research has provided many examples of interesting data representations as tangible artifacts [10, 11, 24, 26], and system models that allow the design of tangible interfaces [20, 22, 23]. TUIs are interfaces whose goal is to provide physical artifacts that are familiar to the user, but that are augmented with digital data. For example, URP [26] is a tangible interface of an urban planning simulator that provides mockups of buildings on a flat surface, and through illumination effects, the buildings' shadows are shown to the user. The user can change the placement of the buildings, as well as the light sources to simulate the changing position of the sun in the sky, to find the optimal placement of the buildings so that the shadow of each will not fall on the others. Thus, through the manipulation of the physical artifacts, the user manipulates the digital environment as well.

But the evaluation of TUIs has been based on experiments with specific interfaces, because there are many differences between TUIs and contemporary interaction styles, such as Direct Manipulation interfaces (DM) [8, 21], for which some of the existing evaluation methods have been created. In TUIs, users interact with the computer by actions that may not be well represented in contemporary evaluation methods. Only recently have frameworks that allow general evaluations of RBI interaction styles been presented [4].

Here we present a model for prehension, or reaching-to-grasp, a movement that is very common in TUIs. We propose a model that follows the logic of Fitts' Law [5, 6], using movement's amplitude to predict the movement's completion time. Obviously, the model is different from Fitts' Law, because, as is explained in the next section, pointing is different from prehension. Using this model, evaluators of TUIs will be able to predict the time required for reaching and grasping the artifacts that are required for the interaction with the TUI. Thus, task completion times that include prehensile actions may be calculated.

In a TUI, prehension is one of the most common actions performed, because it is through this action that a human will grasp an artifact to manipulate it. The same holds true for Virtual Reality (VR) [7], Augmented Reality (AR) [1], and Ubiquitous Computing (Ubicomp) [25]. In these interaction styles prehension is one of the most commonly used actions. For example, in VR, prehension may be performed by point-and-click, but it may also be performed through the use of virtual gloves. In AR and Ubicomp this action is more pronounced, because the user interacts not only with virtual artifacts, but with real world ones as well, which can only be manipulated through prehension.

Modeling prehension then, can be tied directly to modeling and evaluating Reality Based Interfaces (RBIs) [12, 13], of which TUI is a sub-class. Direct Manipulation interfaces (DM) [8, 21] can be modeled and evaluated very accurately by using many different methods, such as GOMS [3], etc. However, it seems that these established methods of interface evaluation need revising to work on emerging interaction styles [4], because, especially predictive evaluation methods, do not include any means of modeling prehension, and probably other allowable actions that do not occur in the DM interaction style. Thus, to revise these methods, we first need to study how to model the allowable actions in new interaction styles that do not appear in contemporary or previous generation ones. This was our first motivation on studying how completion time of the whole movement can be predicted by the amplitude of the movement.

The proposed model is presented with an experiment that centers around prehension. We derive a model for prehension that describes the action based on the amplitude of the movement. We then discuss implications of the findings on user performance and modeling in TUIs, but not for the other classes of RBIs, as generalization over all RBIs is out of the scope of this paper.

2. Background and Motivation

There are many experiments that have been performed to study prehensile movements, from looking at the velocity of the hand during the action, to looking at the shape that the fingers take during the hand's transport towards the target, to how the shape of the hand changes during the whole movement. It is beyond the scope of this paper to provide a survey of the literature for these investigations, but the interested reader is directed to Jones and Lederman [14] and Mackenzie and Iberall [17].

Through the aforementioned investigations, it has been shown that prehension can be described as "...consisting of three phases: moving the arm next to the object, shaping the hand in expectation for grasping, and actually grasping the object" [14]. As can be surmised, the pointing motion that is described by Fitts' Law [5, 6] is one part of the prehensile motion, and for that reason prehension must be slower than pointing.

The transport phase of the hand is comprised of an accelerating motion and a decelarating motion, but these two phases are not equal in the time that each takes, during the transport phase of the hand. In fact, Marteniuk et al. state that "the velocity curve of the movement becomes more asymmetric as the target becomes more fragile or smaller" [18]. Therefore, as mentioned above, the times of acceleration and deceleration vary according to the target's properties. In the presented experiment, we did not take into account all the potential properties of the target that may effect the movement, such as fragility, size, or weight. Rather, we focused on finding a model for the general case of an artifact that is not very fragile, and that accomodates easy grasping, in that it is large enough to fit comfortably in the hand.

The work that is most similar to ours is that of Bootsma et al. [2], who studied how prehensile movements are affected by movement amplitude, target size, and target width. The result of their work however, is not a model that describes the prehensile motion completely. Rather, they studied how the three attributes can be combined to predict the completion time of the transport component and how these same attributes affect the peak hand apperture. They do not, however, produce a model that is able to predict the completion time of the whole prehensile movement.

Another model comes from Mackenzie and Iberall [17], who propose a very detailed model that eventually can be used to create robotic systems that perform prehensile motions. Their model however, is one that tries to explain how the brain (or a neural network), can control the hand, so that prehensile behavior is exhibited. As such, this model takes into account variables whose values are not readily available for evaluation of interactive systems, but rather come from engineering practices, when one tries to simulate prehensile behavior in, for example, robotic systems. Thus, such a model is too complex to be used as an evaluation tool.

Therefore, the reason that we performed this experiment and that we provide the subsequent model, is that even though there have been many studies of the constituent motions of prehension, as well as models and theories of grasping, there seems to be a lack of studies on prehension to provide a simple predictive model that may be used in the evaluation of RBIs. This is the motivation of the current study, which focuses on modeling prehension for the predictive evaluation of the movement in TUIs.

3. Experiment

The described experiment was performed to calculate the completion time needed to perform a prehensile movement. The data gathered was then used to investigate whether a model of prehension can be created that predicts the completion time of the movement from its amplitude.

3.1. Participants

All of the participants in the experiment were students at the Department of Computer Science and Engineering of the European University Cyprus. They were all between the ages of 20 and 30 years old. The participants were enrolled in various courses in Computer Science, and they were compensated for their participation in the experiment with extra credit in their individual courses. 15 participants took part in the experiment, 12 males and 3 females.

3.3 Materials and Design

The participants were seated in front of a table, which was measured on its width. We placed a scale across the table's width, with markings from 5cm to 70cm, at every 5cm, as shown in Fig. 1.

For every trial, the cup used measured 9.5cm diameter and 11.8cm height. It was placed on its base on one of the markings. The placing of the cup on each of the markings was random. For example, the cup would be placed at the 5cm marking, then at the 25cm marking, and so on. The reason we used a plastic cup for this experiment is that we wanted an object that was common enough so that our participants had experience in interacting with it, and we also needed an object that was designed to be grasped in the first place. We believe that a plastic cup fulfills these two requirements.

Each participant performed 5 trials at each marking from 5cm to 70cm, for a total of 14 markings.

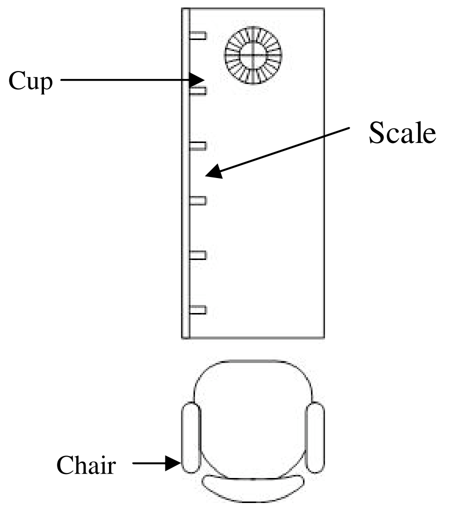

Figure 1 The layout of the experiment

3.2. Procedure

The experiment was performed by using a measured scale on top of a table, with the participants sitting in front of the table, at a distance of about 35cm from the edge of the table to the body of the participants. The scale was created using 5cm intervals from the edge of the table and it ended at 70cm.

The participants were asked to reach and grab a plastic cup that was placed on the table scale. Each participant performed 5 trials over 14 positions of the cup (from 5cm to 70cm at 5cm intervals), for a total of 15 participants x 5 trials x 14 positions = 1050 trials, and where the position of the cup for every trial was decided at random.

For each trial, each participant was asked to place their dominant hand on the edge of the table. The experimenter gave a verbal cue to the participant, who would proceed to reach the plastic cup and grab it, with the instruction to be as fast and accurate as possible. The session was videotaped with a Sharp VL-AH50 video camera and the time to reach and grab the cup was found by analyzing the video transcript. If the participant could not hold on to the plastic cup, or did not manage to grab it, the trial was considered an error, and was repeated.

The video of the session was analyzed using VirtualDub [16], a program that allows frame by frame analysis of digital video files.

3.4. Results

The results, with the relationship between the amplitude of the movement and the time taken to perform the movement at each distance, and with the fitted regressed line, are shown in Fig. 2.

The regression line was calculated using SPSS, and it is described by the equation: $Time = 241 * e^{0.0132*Amplitude}$ ($R^2 = 0.998$). The regression analysis provides strong evidence to support the hypothesis that prehensile movements may be described by an exponential model. Compared to the Fitts' Law model that describes pointing, it is evident that pointing is faster than prehension, as been has described in the existing literature.

4. Discussion

We believe that the existing models of prehension are too involved to be used in the evaluation of TUIs, as mentioned in a previous section. Thus, we propose this model as an approximation to the more involved models, and as a suitable model for the evaluation of interfaces that require the use of prehension motions.

The experiment provides strong evidence to support the hypothesis that prehension actions may be modeled by an exponential model. Such a model describes a motion that is slow compared to other arm motions used in interacting with computers, such as pointing.

We believe that this model provides an accurate description of the increasing difficulty of reaching and grabbing objects that are not in the immediate reach of the user. There are several reasons why the model is slow, as described in the existing literature [14, 17].

One reason is that the model starts off by describing reaching and grasping for artifacts that are inside the area where the user can reach and grab things without making any other body movements. These items can be grabbed relatively fast. But as the items are placed further away from the actor, more of the body is required to be moved, such as when reaching to grab something that is outside the aforementioned area. This results in making the reaching and grabbing action exponentially slower.

Another reason is that, unlike pointing movements, the reach-and-grasp movement consists of both gross and fine movements. Gross movements happen during the transport phase, when the arm moves the hand towards the target artifact, but then the deceleration phase begins, where the hand is shaped into the appropriate way to accommodate the grasping of the target artifact. Because of this deceleration phase, and the accompanying movement of shaping the hand and eventually grabbing the artifact, the speed with which we reach-and-grasp, compared to the speed that we point, is slower.

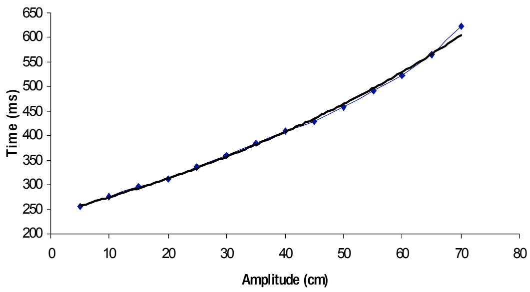

Figure 2 The graph shows the relationship between distance and completion time. Also shown is the regressed line.

4.1 TUI vs. DM Performance

The exponential form of the model also suggests that TUIs may not be as fast as Direct Manipulation (DM) interfaces [8, 19], in terms of motor performance by their users. DM interfaces use pointing, and it has been shown that pointing is faster than reach-and-grasp [14, 17]. Because TUIs rely heavily on prehension actions, this potentially makes task performance slower than in DM.

However, a counterargument to this is that if the artifacts that require the reaching and grabbing motions are near the user, such that no significant body movement is required other than arm movements, then those TUIs may be as fast, or even faster, than their DM counterparts. Fig. 3 shows an example of how this could occur, although the figure is not based on our data. As is demonstrated from the figure, a situation may arise where the prehensile movement is faster for certain distances, than the pointing motion.

We believe though, that this situation is not the norm, but the exception. The reality is that the pointing will probably be faster than reaching-and-grasping in most situations, and therefore a DM interface will allow for faster

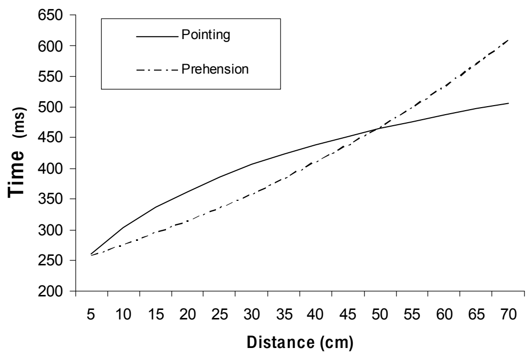

Figure 3 Example of comparison of the prehension and pointing motions

motions. This will result in faster execution of tasks that involve pointing vs. prehension.

4.2 Limitations of the Model

The model that we present here has certain limitations in that the only tested property of the artifact to be grabbed is its distance from the actor of the movement. There are several other properties that may impact the model, such as the width [2] and the weight of the artifact, whether it is a container that contains a liquid that may be spilled, the target's fragility, and others. While we believe that these factors may play a role in the model's behavior, we also believe that most of the artifacts that are presented in TUIs are well represented by our choice of the tested artifact.

5. Conclusions and Future Work

In this paper we have presented an experiment that investigates whether a prehensile motion's completion time can be predicted through the movement's amplitude. The experimental results provide evidence to support the hypothesis that prehensile motions can be described by an exponential model of the form Completion Time = $a*e^{b*Amplitude}$, where a and b are experimentally determined constants.

According to this model, prehensile motions are relatively slow, especially when compared with pointing motions, which can be described using a logarithmic model. We have identified some reasons that make prehension slower than pointing. One is that prehension is composed of a pointing action and a grip-shaping action. This second action is what adds most of the overhead to the reaching-to-grasp motion.

Second, the artifacts may not always be in the area where the user can reach comfortably, without any added body motions. The added body motions to reach these artifacts also require time to be completed, whereas pointing does not require the equivalent body movement.

Thus, we conclude that TUIs that use prehensile motions may be slower than their DM interface counterparts that use pointing motions.

We have also described limitations of the model, in that it does not take into account all properties of the artifact to be grabbed. Rather, it relies only on the motion's amplitude. However, we believe that the model is still valid, because the artifact used in the experiment sufficiently represents artifacts found in TUIs.

We continue to refine the model, with two goals for its development. First we will create a more generic model that will take into account more attributes. However, we do not want to create a model that is as complex as other existing models. Rather we want to include as much information in the model as necessary for precise prediction, but also keep it simple so that it can be used in quick evaluations. Second, we are working towards creating operators that can be integrated into evaluation methods such as GOMSL [15], to allow these methods to model TUIs as well.

5. References

[1] Azuma, R.T., A Survey of Augmented Reality. Presence: Teleoperators and Virtual Environments, 1997; 6: 355-385.

[2] Bootsma, R.J., et al., The speed-accuracy trade-off in manual prehension: effects of movement amplitude, object size and object width on kinematic characteristics. Experimental Brain Research, 1994; 98(3): 535-541.

[3] Card, S.K., T.P. Moran, and A. Newell. The Psychology of Human Computer Interaction. Hillsdale, NJ: Erlbaum; 1983.

[4] Christou, G. Towards a New Method of Evaluation for Reality-Based Interaction Styles. In: Extended Abstracts of CHI 07 Conference on Human Factor in Computing Systems; 2007; San Jose, CA; p.

[5] Fitts, P.M., The Information Capacity of the Human Motor System in Controlling the Amplitude of Movement. Journal of Experimental Psychology, 1954; 47(6): 381-391.

[6] Fitts, P.M. and J.R. Peterson, Information Capacity of Discrete Motor Responses. Journal of Experimental Psychology, 1964; 67(2): 103-112.

[7] Foley, J.D., Interfaces for Advanced Computing. Scientific American, 1987; 257(4): 127-135.

[8] Hutchins, E., J. Hollan, and D. Norman, Direct Manipulation Interfaces, in

User Centered System Design: New Perspectives in Human-Computer Interaction, D.A. Norman and S.W. Draper, Editors. 1986, Erlbaum: Hillsdale, NJ. p. 87-124.

[9] Ishii, H. and B. Ullmer. Tangible Bits: Towards Seamless Interfaces between People, Bits and Atoms. In: CHI Conference on Human Factors in Computing Systems; 1997; Atlanta, GA, USA; p.

[10] Ishii, H., A. Mazalek, and J. Lee. Bottles as a Minimal Interface to Access Digital Information. In: CHI '01 Extended Abstracts on Human Factors in Computing Systems; 2001; New York, NY; p. 187-188.

[11] Jacob, R.J.K., et al. A Tangible Interface for Organizing Information Using a Grid. In: CHI 02 Conference on Human Factors in Computing Systems; 2002; Minneapolis, MS; p. 339-346.

[12] Jacob, R.J.K., et al. Reality-Based Interaction: Unifying the New Generation of Interaction Styles. In: Extended Abstracts of ACM CHI 07 Conference on Human Factors in Computing Systems; 2007; San Jose, CA; p. 2465-2470.

[13] Jacob, R.J.K., et al. Reality-Based Interaction: A Framework for Post-WIMP Interfaces. In: CHI 08 Conference on Human Factors in Computing Systems; 2008; Florence, Italy; p.

[14] Jones, L.A. and S.J. Lederman. Human Hand Function. New York, NY, USA: Oxford University Press; 2006.

[15] Kieras, D. A Guide to GOMS Model Usability Evaluation using GOMSL and GLEAN4. 2006 [cited September 20th, 2006]; Available from: http://citeseer.ist.psu.edu/kieras99guide.html.

[16] Lee, A., *Virtual Dub*, 24th January 2008: http://www.virtualdub.org/.

[17] Mackenzie, C.L. and T. Iberall. Advances in Psychology: The Grasping Hand. Amsterdam: Elsevier Science BV; 1994.

[18] Marteniuk, R.G., et al., Constraints on Human Arm Movement Trajectories. Canadian Journal of Psychology, 1987; 41: 149-176.

[19] Schneiderman, B., Direct Manipulation: A Step Beyond Programming Languages. IEEE Computer, 1983; 16(8).

[20] Shaer, O., et al., The TAC Paradigm: Specifying Tangible User Interfaces. Personal and Ubiquitous Computing, 2004; 8(5): 359-369.

[21] Shneiderman, B., Direct Manipulation: A Step Beyond Programming Languages. IEEE Computer, 1983; 16(8).

[22] Ullmer, B. and H. Ishii, Emerging Frameworks for Tangible User Interfaces. IBM Systems Journal, 2000; 39(3&4): 915-931.

[23] Ullmer, B., H. Ishii, and R.J.K. Jacob, Token+Constraint Systems for Tangible Interaction with Digital Information. ACM Transactions on Computer-Human Interaction, 2005; 12(1): 81-118.

[24] Underkoffler, J. and H. Ishii. Urp: A Luminous-Tangible Workbench for Urban Planning and Design. In: CHI 99 Conference on Human Factors in Computing Systems; 1999; Pittsburgh, PA; p. 386-393.

[25] Weiser, M., *The Computer for the 21st Century*, in *Scientific American*. 1991. p. 94-104.

[26] Zigelbaum, J., et al., *Tangible Video Editor: Designing for Collaboration, Exploration, and Engagement*. 2005, Department of Computer Science, Tufts University: Medford, MA.