TouchOver: Decoupling Positioning from Selection on Touch-based Handheld Devices

Adriano Scoditti, Thomas Vincent,
Joelle Coutaz, Renaud Blanch
UJF-Grenoble1 ; CNRS UMR 5217
Laboratoire d'Informatique de Grenoble (LIG)
Grenoble, France
Firstname.*Lastname*@imag.fr

Nadine Mandran
CNRS UMR 5217
Laboratoire d'Informatique de Grenoble (LIG)
Grenoble, France
Firstname.*Lastname*@imag.fr

RESUME

Contrairement à la souris, les écrans tactiles des dispositifs mobiles n'ont pas d'état *mouseover* pour fournir à l'utilisateur des informations dynamiques pro-actives. De plus, sur écran tactile, la détection des actions "appuyer" et "relâcher" du doigt rend difficiles les sélections requérant une grande précision.

En réponse à ces limitations, nous proposons *TouchOver*, une technique multimodale pour dispositif mobile qui tire partie de l'écran tactile et des accéléromètres : le positionnement est effectué avec le doigt sur la surface tactile et la sélection par inclinaison du dispositif vers l'avant. Ainsi, *TouchOver* introduit un état *mouseover* et améliore la précision de la sélection tout en restant compatible avec les techniques d'interaction existantes. Dans une étude formelle, nous comparons *TouchOver* à deux autres techniques de sélection. Les résultats montrent une amélioration significative de la précision ainsi qu'un bon compromis entre vitesse d'exécution et précision.

MOTS CLES

Interaction tactile, gestuelle, mutlimodale, dispositifs mobiles.

ABSTRACT

When compared to conventional desktop mouse input, touch input on handheld devices suffers from the lack of a main feature: that of a *mouseover* state that can provide users with dynamic pro-active information. In addition, with touch screens, selection precision is limited by undesired extra finger tracking during finger press and lift movements.

We propose *TouchOver*, a multi-modal input technique for touch-screen accelerometers-enabled handheld devices where positioning is performed with a finger on the touch surface, while selection is triggered by a gentle "tilt forward" of the device. By doing so, *TouchOver* adds a *mouseover*-like state and improves selection precision while remaining compatible with existing interaction techniques such as Shift [10] devised to improve precision. Our formal user study shows a significant precision improvement over two other selection techniques as well as a good tradeoff between speed and accuracy.

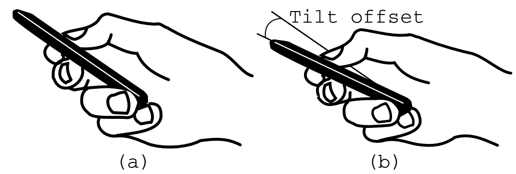

Figure 1: *TouchOver* users switch between two interaction states (a) and (b) by tilting the device gently while still interacting with their finger. This makes it possible hovering, dragging, and feedforward-enabled interaction techniques, as well as visual and eyes-free interface exploration, while improving selection precision.

Categories and Subject Descriptors

H.5.2 [**User Interfaces**]: Input Devices and Strategies

General Terms

Design, experimentation, human factors.

Keywords

Touch, gestures, mutli-modal input, mobile devices.

1. INTRODUCTION

Touch-enabled handheld devices provide interactions that are often described as more "natural" for the user. However, as demonstrated by Buxton [2], touch-screens support only a limited set of the interactions made possible by graphics tablets with stylus or even by desktops with the mouse. For example, it is currently not possible with touch-screens to perform one-finger concurrent pointing, scrolling and dragging, or to obtain feedforward information such as *on-over* preview, without triggering system interpretation (e.g., object or command selection). Either the finger is tracked as it comes in contact with the touch surface and triggers a system action as soon as it leaves the surface, or the finger is *out-of-range* with no input available to the system, thus with no effect on the system. There is no intermediate state where a position is given to the system without triggering system interpretation. As a result, interaction designers for touch-enabled handheld devices are required to introduce extra modes and ad-hoc solutions to provide additional features.

For example, to display a web page link address, web browsers on smart phones use a *touch and hold* interaction and a pop-up window, whereas web browsers on desktops use the *mouseover* state. For the iPhone home screen, users need to *touch and hold* an application icon to switch from the nominal pointing and scrolling mode to the edit mode. Although attractive at first sight, these techniques have drawbacks and break the interaction flow. Alternatives have been proposed [1, 5].

In addition to the absence of an intermediary *tracking* state, touch-screen inputs are not sufficiently precise for selection tasks. As reported by Benko *et al.* [1], the lack of precision for selection tasks has been addressed in various ways including avoidance of target occlusion and alteration of the control-to-display ratio. Nevertheless, all these techniques are limited by the fact that most touch-screen device performs extra tracking during finger press and lift movements. This results in the system acquiring extra motion events, misplacing *press* and *release* events. This imprecision is negligible for sufficiently large targets. It is not acceptable for accurate pointing tasks like setting a land mark on a map, or moving the anchors of a Bezier curve in a drawing editor.

In this article, we propose *TouchOver* that combines screen touch input with device orientation sensed by accelerometers to enhance interaction while increasing touch-screen selection accuracy. In the next section, we describe *TouchOver* and motivate the design options. We then present several examples of interactions that have been enhanced with *TouchOver* followed by the description of an experimental evaluation and its results.

2. THE DESIGN OF TOUCHOVER

Informed by Buxton's early work on graphical inputs [2], the behavior of *TouchOver* is modeled as a three-state finite state automaton that includes two finger *tracking* states. Positioning is achieved through the absolute position of the finger on the touch-screen. Transition between the two finger *tracking* states, (*State 1* and *State 2*), is performed with a gentle tilt forward/backward of the device (Figure 1). Therefore, positioning is decoupled from selection by the way of two distinct modalities, touch and tilt, that are combined in a complementary way [3] to trigger a system meaningful function. The rationale for touch and tilt is the following.

The use of the finger as a pointing technique to denote a point in space is naturally justified by its native deictic property [6] as well as by the widely adopted direct manipulation paradigm. The choice for tilt is based on the following observations. (1) Tilt gestures can be captured with sensors, e.g., accelerometers and gyroscopes, that are now widely available in commercial products. Even if tilt input may suffer from recognition ambiguities, it has been proven useful for a large variety of tasks [8]. (2) Tilt gestures are easy to learn and provide a kinesthetic feedback about the current mode [9]. (3) Tilt requires only one hand and can be performed while a finger is in contact with the screen. (4) By using a smooth gesture of rather small amplitude, tilt allows users to keep their field of vision focused on the screen while performing the gesture. Through informal testing, we found an *11 degrees* angle offset suitable to satisfy this visual requirement. Due to the small amplitude of the *TouchOver* gesture, the ease of finger motion is similar in the two states (*State 1* and *State 2*), thus allowing delicate interaction such as dragging. As a smooth gesture, it also generates less strain than an impulsive shock. (5) The wrist rotation along the ulnar/radial flexion axis goes with the finger press in a natural manner. In addition, along this axis, the amplitude of the gesture is naturally constrained by the wrist capabilities, thus encouraging small amplitude

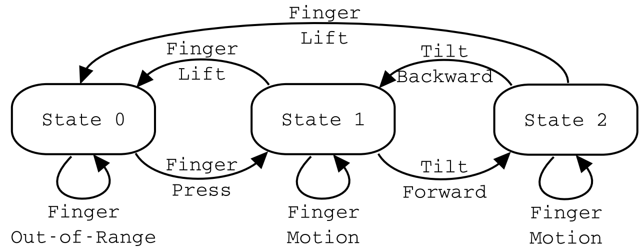

Figure 2: The *TouchOver* three-state input model. In *State 0*, the finger does not touch the screen. *State 1* can be used to model the *mouseover* state, and *State 2*, the selection state.

tilt. Tilting the device along this axis has been demonstrated to be a usable gesture [7]. Nevertheless, *TouchOver* leaves open rotations along the other axes if larger amplitudes are needed.

3. TOUCHOVER IN USE

TouchOver provides the system with user inputs as a three-state model (Figure 2) similar to the graphics tablet with stylus [2]. This automaton includes the original state machine for touch-screens that consists of *State 0* and *State 1* only. Thus, *TouchOver* is able to support existing interaction techniques. With the additional state, *TouchOver* provides an extra transition between *State 2* and *State 0* triggered by a *finger lift* movement. This transition can be used to extend existing interaction techniques. For example, if *State 2* supports dragging, the user can either confirm the drag operation with a tilt backward or cancel it by lifting the finger. In summary, *TouchOver*, not only supports existing techniques, but can also be used to extend and enhance them.

The following two applications illustrate the use of *TouchOver* to extend existing touch-screen based select and scroll interactions with a cancelable drag interaction triggered by a tilt gesture. The first application (Figure 3) extends a scrollable list of selectable items with dragging capability, thus allowing to reorder the items. The second one extends a scrollable map with selectable landmarks with dragging capability, thus allowing to move the landmarks.

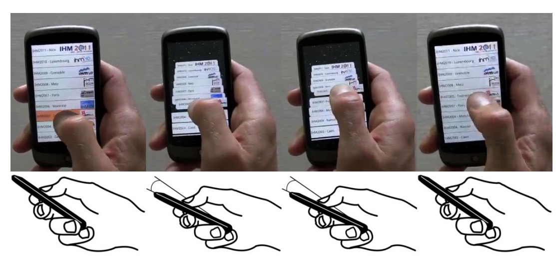

Figure 3: *TouchOver*-based list item dragging sequence.

Another demonstrator illustrates the use of *TouchOver* for *feed-forward interaction*. It is an image processing application where the user can preview the effect of image filters while touching the corresponding filter button. Yet, no action is triggered, and the preview stops as soon as the user moves the finger out of the button or lifts it from the screen. To actually apply a filter, the user needs to tilt the device.

Without *TouchOver*, the Foley's two elementary tasks [4], positioning and selection, are bundled together. *TouchOver* offers the possibility to decouple the selection from the positioning task since selection can be triggered by tilt rather than by finger press or lift. In addition, as demonstrated by the following experimental evaluation, which uses currently available hardware, pointing precision benefits from such decoupling.

4. EXPERIMENTAL EVALUATION

We have conducted two controlled experiments with three different validation techniques. For all three techniques, positioning was performed with the finger in contact with the touch-screen. The three validation techniques were:

Long-Press where validation is performed when the user keeps their finger still for 1 second. Audio and graphical feedback signals the validation;

TouchOver where validation is performed by a tilt forward of at least 11 degrees along the axis parallel to the width of the device while the finger is touching the screen. Audio and graphical feedback signals the validation;

Take-Off where validation is performed on finger lift from the screen. This is used, for example, for soft keyboards.

Based on our rationale and pilot testing, we hypothesized that *Long-Press* outperforms *TouchOver* which in turn outperforms *Take-Off* to a greater extent with regard to precision. Our goal was also to estimate the impact of the extra user's action that *TouchOver* requires in terms of strain and validation time.

4.1 Design and Apparatus

Eighteen right-handed, unpaid, volunteers (2 females), ranging in age from 21 to 33 years participated in the experiments. All but one had prior experience with touch-screen based handheld device among whom 11 used it on a daily basis.

For each of the techniques, participants were explained the technique with a sample application. Then, they performed a first experiment focusing on validation precision, followed by a second experiment focusing on validation time. Both experiments were performed with the dominant hand while standing-up still.

A repeated measures within-participants design was used. Each of the 18 participants performed the experiments for the three *Techniques*. Presentation order of *Techniques* was counter-balanced across participants.

The experiments were conducted on the iPod Touch 4th generation 8GB running iOS4.2.1. The screen is 3.5 inches wide with 960 x 640 pixels (resolution 326 dpi). With one finger, the touch input resolution is 0.5 point (0.18 mm for a resolution of 144 dpi) or 2.25 pixels. We used such a device for its high screen and touch sensor resolutions. As opposed to usual sub-pixel accuracy of pointing devices, here the touch input is less precise than the screen.

4.2 Precision Experiment

For the first experiment, participants were asked to reach and validate a one-dimension position figured by a horizontal dashed line (Figure 4). A second horizontal dashed line figured the position of the finger, thus avoiding the occlusion problem. As we aimed at testing the limitations due to input, not to output, we found necessary during pilot testing to zoom the dashed line by 4 compared to the motor space and add a control-to-display ratio of 5.

The experimental variables were: 3 *Techniques* (*Take-Off*, *TouchOver*, *Long-Press*) x 3 *Blocks* x 15 trials per *Block* = 135 data points per participant. The measured variables were *Distance* and *Errors*.

We considered the error rate while acquiring the target position of size 0.18mm, or 0.5 point (resolution of the touch sensor).

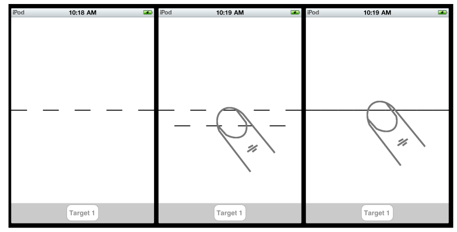

Figure 4: The user interface for the precision experiment: the target line (left); the thumb approaching the target (center); the thumb on the target (right).

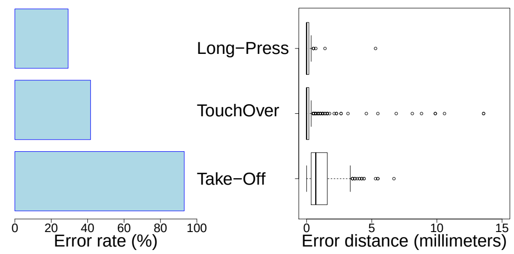

Figure 5: Percentage of errors by validation techniques (left), and boxplot of error distance in millimeters by validation technique (right).

Pearson's Chi-squared independence test between success of target acquisition and the 3 *Blocks* did not show any significant dependency. Thus, there is no evidence of either learning or tiring effects for any of the techniques.

Pearson's Chi-squared independence test between success of target acquisition and the 3 *Techniques* shows a significant dependence (X-squared = 749.16, $p < .0001$, Cramer's V = 0.5). The overall error rate percentage is 55%. This percentage is higher for *Take-Off* (93%) than for *TouchOver* (42%) and *Long-Press* (29%) (Figure 5). This supports our hypothesis and also indicates that *Take-Off* is inappropriate to reach such a touch-screen resolution precision.

We also measured the distance of selected position to the target position (Figure 5). We performed a 3 x 3 x 6 (*Technique* x *Block* x *Presentation order*) within subjects analysis of variance on median absolute error distance. Significant main effect was found for *Technique* ($F_{2,161} = 97.8102$, $p < .0001$). Other effects and interactions were not found significant. Post hoc Tukey multiple means comparison test confirmed that *TouchOver* and *Long-Press* are more precise than *Take-Off* and did not found any significant difference between *TouchOver* and *Long-Press*. These results support our hypothesis.

4.3 Validation Duration Experiment

For the second experiment, participants were asked to select alternatively one among two buttons as quickly as possible with the presented validation technique. We measured validation *Duration*, the time spent between the finger press event and the target validation. *Long-Press* was excluded from this experiment analysis as this measure is a measure of the timeout.

The experimental design was: 2 *Techniques* (*Take-Off*, *TouchOver*) x 2 *Targets* x 5 trials per *Target* = 10 data points per participant. The measured variable was validation *Duration*.

We performed a 2 x 2 x 6 (*Technique* x *Target* x *Presentation order*) within subjects analysis of variance on median validation time of aggregated repetitions for each participant. Significant main effect was found for *Technique* ($F_{1,71}$ = 47.8942, p < .0001). *Presentation order* effect was found significant, though with a p-value of 0.067 ($F_{5,71}$ = 2.1953). *Target* effect and interactions were not found significant. This indicates that the design of the two-target experiment was appropriate.

Post hoc means comparison test confirmed that *Take-Off* is faster than *TouchOver* (T = -11.7233, p < .0001). The mean validation time was 106 ms for *Take-Off* and 213 ms for *TouchOver* with a 95% confidence interval means difference ranging from 76 ms to 138 ms. This provides a measure of the cost of the extra action required by *TouchOver* in terms of task completion time.

5. DISCUSSION

Our experimental results support our hypothesis. Like *Long-Press*, *TouchOver* improves positioning validation precision compared to *Take-Off*. Yet, *TouchOver* is not as precise as *Long-Press*. During the experiments, we observed that in some cases, participants used their thumb and fingers along with their wrist to perform the tilt offset of the device. In such cases, user's thumb footprint would change during gesture introducing erroneous validation. As shown by the experimental results, even with this limitation, *TouchOver* still remains of interest for precise positioning validation.

During the experiment, when users were explained *TouchOver*, they understood it and learnt it quickly. This is comforted by the fact that we did not find any significant indication of the learning effect. They also had no difficulties to reach the 11 degrees angle offset that triggers the validation. Nevertheless, some smartphone-experienced users were tempted to shake the device instead of tilting the device gently to perform the validation. This can be explained by the recent introduction of shake-controlled commands in commercial products. Indeed, previous user experience influences expectation about physical gesture based interaction techniques with handheld devices.

Although our implementation of *TouchOver* was based on a simple accelerometers-based gesture recognition, it performed well in a controlled environment. Gesture recognition robustness can still be enhanced by taking advantage of more appropriate sensors like gyroscopes and implementing more sophisticated algorithms. Still, any gesture recognition will suffer from ambiguities.

We believe that an appropriate feedback along with users' habits and expertise can moderate users' thumb movements while tilting the device, discourage users to shake the device instead of a gentle tilt and thus reduce the number of recognition ambiguities.

By introducing an additional *tracking* state to handheld devices, *TouchOver* improves selection precision at the expense of an extra physical action. Then, *TouchOver* offers room for trade-off between precision and speed. With *Long-Press*, users need to wait for a timeout to perform validation, whereas with *TouchOver* they actively control the state transition while involved in a kinesthetic quasi-mode [9]. In addition, experimental results indicate an affordable cost for *TouchOver* in terms of task duration and physical strain.

For common handheld devices interactions like concurrent pointing and scrolling, *TouchOver* is of no particular interest due to the extra user's action it involves. Yet, *TouchOver* can extend such one-finger interactions since it can support pointing, scrolling and dragging concurrently in a modeless manner.

Previous works propose to enrich touch-based interactions. *SimPress* [1] uses a small finger rocking motion to trigger a state transition. Hinckley *et al.* [5] combine touch and motion sensing to extend touch-based interaction. For example, they propose *Tip-to-select* that supports both zoom and 2D selection with two fingers on the screen. To switch mode, the user quickly tips the device away and back. This is in contrast to *TouchOver* which provides a kinesthetic feedback about the current state. In addition, contrary to *TouchOver*, they do not address the precision problem when combining touch and tilt together.

6. CONCLUSION

We have presented *TouchOver*, a complementary multimodal input for one hand interactions on touch-screen based accelerometers-enabled handheld devices. *TouchOver* offers a three-state model input similar to the stylus tablet input with two states where the system tracks fingers motion, thus adding a *tracking* state to touch input. This creates new opportunities for handheld device interaction techniques like *on-over* interactions, feedforward, or visual and eye-free user interface exploration.

When positioning validation is performed on the tilt gesture transition rather than on finger press or lift, positioning tasks gain in precision at the expense of an extra action. Our evaluation of *TouchOver* in a controlled environment shows an encouraging trade-off between *Take-Off* and *Long-Press*. Indeed it improves positioning precision at an affordable cost in terms of task duration and physical strain. Existing precision improvement techniques can benefit from this gain in precision.

7. ACKNOWLEDGMENTS

This work has been supported by the FUI Project NOMAD.

8. REFERENCES

[1] H. Benko, A. D. Wilson, and P. Baudisch. Precise selection techniques for multi-touch screens. In *Proc. CHI'06*, pages 1263–1272. ACM, 2006.

[2] W. Buxton. A three-state model of graphical input. In *Proc. INTERACT'90*, pages 449–456. IFIP, 1990.

[3] J. Coutaz, L. Nigay, D. Salber, A. Blandford, J. May, and R. Young. Four easy pieces for assessing the usability of multimodal interaction: The CARE properties. In *Proc. INTERACT'95*, pages 115–120. IFIP, 1995.

[4] J. D. Foley, V. L. Wallace, and P. Chan. The human factors of computer graphics interaction techniques. *IEEE Comput. Graph. Appl.*, 4(11):13–48, 1984.

[5] K. Hinckley and H. Song. Sensor Synaesthesia : Touch in Motion , and Motion in Touch. In *Proc. CHI'11*, pages 801–810. ACM, 2011.

[6] M. Karam and S. M. C. A taxonomy of gestures in human computer interaction. Technical Report ECSTR-IAM05-009, Electronics and Computer Science, University of Southampton, 2005.

[7] J. Ruiz, Y. Li, and E. Lank. User-Defined Motion Gestures for Mobile Interaction. In *Proc. CHI'11*, pages 197–206. ACM, 2011.

[8] A. Scoditti, J. Coutaz, and R. Blanch. A novel taxonomy for gestural interaction techniques based on accelerometers. In *Proc. IUI'11*. ACM, 2011.

[9] A. Sellen, G. Kurtenbach, and W. Buxton. The role of visual and kinesthetic feedback in the prevention of mode errors. In *Proc. INTERACT'90*, pages 667–673. IFIP, 1990.

[10] D. Vogel and P. Baudisch. Shift: a technique for operating pen-based interfaces using touch. In *Proc. CHI'07*, pages 657–666. ACM, 2007.