ScaleGest

Surface Gestures for Advanced Graphical Interfaces: Which Gesture for What

Contact

Caroline Appert (LRI) appert@lri.fr http://www.lri.fr/~appert
Gilles Bailly (Telecom ParisTech) gilles.bailly@telecom-paristech.fr www.gillesbailly.fr/
Emmanuel Pietriga (INRIA) emmanuel.pietriga@inria.fr http://pages.saclay.inria.fr/emmanuel.pietriga/cv/cv_en.html

Topic

Make gesture-based interaction scale with graphical applications complexity: which gesture works for which control and how to transition between controls.

Summary

While most gesture-oriented applications target novice users, we also aim at designing gesture-based interaction for expert users who navigate and manipulate large datasets. In this context of *advanced graphical applications*, the number of gestures should be large-enough to cover the set of controls (i.e., commands and parameters settings) but remain simple so as to avoid exceeding human abilities. Making gesture-based interaction scale with graphical applications' growing *complexity* can be achieved only by understanding the foundational aspects of this input modality, that will be investigated during this PhD. The candidate will work on characterizing and structuring both the *space of application controls* and the *space of surface gestures* in order to establish guidelines for appropriate *control-gesture mappings.* The student will define a sound and systematic *evaluation methodology* that will serve as a reference benchmark for evaluating these mappings. The resulting control-gesture mappings will be demonstrated in the specific application domains of cartography and astronomy.

Detailed description

Objective: Understand the foundations of using multi-touch gestures for interacting with advanced graphical applications.

Multi-finger input offers a very expressive channel to interact with devices equipped with a touchscreen, by associating a given human gesture with a system control (i.e., command invocation and parameters setting). While the channel size of gesture-based communication is theoretically very large, it is actually reduced drastically in practice by human and system limitations. On the one hand, cognitive and motor resources limit the number of associations humans can memorize and the complexity of gestures they can perform. On the other hand, recognizing humans' gestures that involve a large number of muscles and joints from a sample of contact points on a tactile surface is difficult.

As a consequence, current interactive devices make an intensive use of single-finger slides and two-finger pinches for viewport navigation. The use of other multi-touch gestures remains anecdotal. In the end, interaction still heavily relies on graphical widgets for many manipulations. However, widgets on touchscreens not only reduce the size that can be dedicated to the content of interest, but also raise usability issues. For example, acquiring and precisely manipulating targets with fingers can be tedious on both small and large surfaces.

The objective of this thesis is to understand the foundations of using multi-touch gestures for interacting with *advanced* graphical applications. While most gesture-oriented applications target novice users, we aim at designing gesture-based interaction for both novice and expert users. In the context of expert users who navigate and manipulate large datasets, gestures should not only remain simple to execute; they should also be easy to input sequentially, so as to provide fluid transitions between the different navigation and manipulation actions that have to be chained in many tasks. This investigation of foundational aspects is indispensable to make gesture-based interaction scale with graphical applications' *complexity*.

The originality of our approach is twofold. First, we will characterize and structure both the *space of controls* and the space of multi-touch gestures in order to define the best mappings between application controls and surface gestures. We will study which gestures can be performed from a human perspective and evaluate their expressive power by conducting empirical studies. Second, we not only focus on commands, but also on transitions between commands and parameter settings.

The outcome of the thesis will be 1) *guidelines* available to practitioners (interface designers and developers) to assist the making of graphical applications that are controlled through multi-touch gestures drawn on a tactile surface; 2) the definition of a sound and systematic *evaluation methodology* that will serve as a reference benchmark for gesture-based interaction.

We will validate the results of this thesis with different applications. We foresee the use of multi-touch gestures for navigating and manipulating large datasets on devices that range from tablet- to wall-sized displays. We target not only applications that run either on a tablet or on a multi-touch wall alone, but combinations thereof as well, such as, e.g., applications that are displayed on a wall and remotely controlled with a tablet. Application domains will include astronomy and geographical information systems, as users in these application domains are faced with large datasets that they navigate and manipulate through numerous interactive controls.

- ✓ State of the art

Research about multi-touch gestures is still in its infancy, especially regarding how to design usable gesture-control mappings for graphical applications. Most research on this topic is conducted in North America by companies like Google or Microsoft and by universities like the University of Washington or the University of Berkeley.

Gesture Recognition. Some of their contributions focus on technical aspects of multi-touch gesture recognition. Tools such as Proton (Kin et al., 12) and Gesture Coder (Lü and Li, 12) allow developers to implement multi-touch gestures. They internally represent gestures as basic touch events organized into a state machine, where a callback can be associated with any state so as to implement continuous control. **These projects focused on addressing technical issues related to the implementation of arbitrary multi-touch gestures, but neither considered which sets of gestures make sense for users, nor enabled transitions between different gestures within the same continuous stream of touch events**.

Gesture-control Mapping. A few studies have been conducted to understand which gestures are guessable and easy to memorize. In a “guessability” experiment, Wobbrock et al. observed that user-defined gestures are easier to memorize than pre-defined gestures (Wobbrock et al., 09). However, the commands for which participants had to define gesture mappings were inspired by simple mouse-based interfaces. This may have swayed participants towards defining gestures that mimic mouse use. In their study, they also found that some gestures elicit little inter-user agreement, suggesting the need for on-screen widgets or pre-defined gestures. Nacenta et al. also suggest that there is a need for pre-defined gestures to complement user-defined ones (Nacenta et al., 13). Pre-defined gestures are usually better recognized by the system; their mapping with controls are more consistent across applications, and they can be transferable among users in a collaborative setting. **ScaleGest is intended to help designers define such pre-defined multi-touch gestures for advanced graphical applications.**

Control Properties. Achieving this goal requires to have a clear representation of the different types of controls users may need in interactive graphical applications in order to classify usable gestures along the required control properties. In that respect, it relates to the work of Card et al. (Card et al., 91) that organizes physical input devices into their control properties like linear vs circular, or absolute vs relative, and then relates them to basic elementary interaction tasks. This taxonomy was proposed in the early 90s. Interactive systems have evolved a lot since then, and are now far from the six elementary tasks Foley et al. identified (e.g., select an object, move an object, draw, etc.) (Foley et al., 84). In particular, we manipulate larger and larger datasets on a wide variety of devices, meaning that navigation plays an increasingly important role in interactive systems. **With ScaleGest, we want to capture the nature of current advanced interactive systems by identifying the different types of control they expose to users and what gestures properties would best support them.**

Human anatomy. Designing multi-touch gestures requires taking anatomical properties and constraints into account. In particular, fingers cannot be considered as independent entities. Studies have shown that, to some extent, one finger will inevitably move when another one does (Wei et al., 10; Zatsiorsky et al., 00). This enslaving is caused both by peripheral factors such as muscles shared between fingers, and by central factors such as overlapping cortical representations (Schieber, 01). The human hand is also not a simple set of identical fingers. For example, the thumb is different from the other fingers as it sits in a different plane and is controlled by its own separate muscles (Moore et al., 13). All these physiological and anatomical knowledge has not yet been really considered in the HCI community for guiding the design of multi-touch gestures. **ScaleGest will exploit the results of these studies to inform the way of characterizing multi-touch gestures so as to relate them to required control properties.**

Finally, ScaleGest is also related to research on gesture discoverability (e.g., Appert and Zhai, 09; Walter et al., 13), gesture learning (e.g., Appert and Bau, 10; Bailly et al., 10) or gesture implementation (e.g., Appert and Zhai, 09). The PhD work proposed here does not target specific contributions in these particular research areas, but will benefit from the advisors' earlier experience to efficiently integrate the proposed gestures into the envisioned demonstrator applications.

✓ concepts, tools and challenges

The overall goal of this proposal is to make guidelines available to interface developers in order to assist the making of graphical applications that are controlled through multi-touch gestures drawn on a tactile surface. In theoretical terms, this means defining the best mappings between application controls (commands and parameter settings) and human gestures (finger traces on the surface). ScaleGest aims at making gesture interaction rich and mature-enough to enable the control of complex graphical applications with a focus on navigation and manipulation of large datasets.

To achieve this goal, the first challenge consists in characterizing and structuring both the *space of application controls* and the *space of surface gestures* into taxonomies. The second challenge consists in establishing guidelines for appropriate *control-gesture mappings*. Finally, the third challenge consists in generalizing these results by considering the *transitions between controls*.

Taxonomies for application controls and surface gestures

Space of application controls: Graphical applications mostly rely on view navigation, object selections (single or multiple) and object manipulations. Identifying the relevant dimensions to describe this very large space in a concise and comprehensible manner is indispensable to understand the object (i.e. the digital content+tools) users want to manipulate. Many dimensions can be listed and explored to build a taxonomy that will help structure this design space. We give here a few examples of such dimensions that can be identified by a study of the HCI literature. Users may want to navigate at different *control order*s. As an example, we can consider the Acrobat hand tool (0-order), the scrollbar (1-order) or even advanced techniques like SDAZ (Igarashi and Hinckley, 00) where users define a speed vector for quick navigation (2-order). Object or value selections depend on the *definition interval* in which the selection is made. Invoking a command or selecting a value in a list (discrete) is different from selecting a value in a slider (continuous). Objects can be manipulated differently according to their number of *degrees of freedom*. For example, a slider knob can only be moved along a segment (one degree of freedom) while a 3D object in a graphical scene may be rotated, translated and scaled along the x, y and z axes (3 + 3 + 3 degrees of freedom). When there are more than one degree of freedom, it is also important to consider what are the degrees of freedom that can be controlled in *parallel*. For example, panning and zooming can potentially be controlled at the same time (*integrable*), while translating and changing the color of an object need to be done in sequence (*separable*) (Jacob et al., 94).

Space of surface gestures: Even though the space of gestures performed on a surface is more constrained than the set of all possible gestures humans can accomplish, it remains very large. However, as mentioned above, cognitive and motor resources obviously reduce the space that can be used as a vocabulary for interaction. To have a clear understanding of the input channel we are considering, we need to identify dimensions for structuring this space of usable surface gestures. First, gestures differ in their *range of motion*. For example, a 2-finger pinch gesture finds its limits as soon as the thumb touches the opposite finger (typically the index or the middle finger); while a gesture where one finger moves circularly can be repeated an arbitrary number of times. Second, gestures can be more or less *fatiguing*. Some gestures can impose finger postures that are hard to maintain during specific motions. In addition, gestures do not all imply the same number of joints and muscles. In the specific context of multi-touch gestures, it is also important to consider *finger constraints*. As mentioned earlier, studies have shown that fingers cannot be considered as independent entities and that, to some extent, one finger will inevitably move when another one does (Häger-Ross and Schieber, 00; Zatsiorsky et al., 00). Finally, we are especially interested in *screen characteristics* that influence the type of movements, as we target interactive systems such as tabletops (horizontal screen), tablets (handheld screen) and walls (vertical screen).

<u>Control-Gesture mappings</u>

We want to establish a design rationale for implementing gesture-based interfaces by answering the question of which gestures work for which controls when expert users navigate and manipulate large datasets. This means identifying functions that take their input in the space of application controls and their output in the space of surface gestures. By relying on the two taxonomies mentioned above, we could systematically study how respective dimensions from both spaces can be associated, under which conditions. For example, gestures that have a limited *range of motion* demand clutching actions for selections within a large *definition interval*. Another example would be assigning different fingers to *parallel* controls under the condition that they are not under *finger constraints* that enslave their respective movements together. While this study of associations is presented as a follow-up work to the taxonomies for the two spaces in this document, it actually has to be done as a background task as it will cause refinements and changes within the individual taxonomies. For example, considering the *cost of accidental activation* of a control can also be interesting to encourage the use of gestures that remain usable but are far from the natural postures and movements (e.g., gestures that reach the bounds of their range of motion) for controls that should not be accidentally activated.

<u>Continuous Control and Transitions</u>

Interacting with graphical applications involves interlacing command invocations and parameter(s) settings. Let's consider the scenario where a user wants to move a fragment of text within a document (the scenario can even become more complex if we consider operations among different documents). To achieve this goal, users must be able to adjust the selection interval (which may involve scrolling operations) and to move selected text over the target location. This scenario is actually rather simple when performed on a desktop computer, but becomes very tedious on tactile screens. The main reason is that mice have buttons and wheels that allow users to easily switch between selection, navigation and manipulation of objects whereas, on tactile screens, the only thing fingers can do is touch the surface, which is usually dedicated to basic view navigation. Raising the expressive power of gesture-based interaction requires enabling users with easy transitions between gestures, without resorting to explicit delimiters like a pause or a software button that would break the fluidity of interaction. Getting rid of delimiters requires to develop incremental (or real time) recognizers, as a change should be detected as soon as possible. Recognition algorithms must be able to classify a gesture (or at least detect a change if the next gesture cannot yet be reliably classified) based on the smallest gesture trace possible. To our knowledge, this is a research and engineering topic that has not been tackled in the HCI literature. A few interaction techniques like Motion-Pointing (Fekete et al., 09) and CycloStar (Malacria et al., 10) rely on local geometrical properties to identify a specific motion, but they focus on single-point oscillatory movements for selecting an item in a set, or for transitioning between pan and zoom controls. The goal here is to generalize this approach by observing movement characteristics like speed and acceleration during the transition phase and deriving criteria to recognize it, in order to stop the current control and enter the next one as fast and as reliably as possible.

- ✓ Expected results

<u>A taxonomy of usable multi-touch gestures</u>

This project will start with a systematic literature review in both gesture controls for graphical applications in HCI and in anatomical and physiological aspects of gestures in experimental psychology. Individual taxonomies are a contribution per se. Several taxonomies have already been proposed (e.g Karam and Schraefel 2005). However, they generally fail to precisely distinguish what are the characteristics of the gestures and the characteristics of the mappings. More relevant to our approach are the analysis of the space of gestures focusing on single-point gestures (Zhai et al., 12) and a very coarse-grained taxonomy of multi-touch gestures that consider a small number of very high-level dimensions for tabletop displays (Wobbrock et al., 09). There is a great window of opportunity to publish a taxonomy of multi-touch gestures guided by their usability for various controls on different types of tactile screens. Validating a taxonomy means demonstrating its wide coverage of existing gestures and, ideally, highlighting new design spaces to explore.

Criteria of success: A comprehensible taxonomy of surface gestures organized along their properties for controlling graphical applications. Showing, based on this taxonomy, that there are some inconsistencies in existing designs, and that current implementations of gesture-based interaction do not take full benefit of human capacities and abilities.

Low-level validation of mappings

As a first validation step, hypotheses regarding the right mappings to consider should be tested on low-level abstract tasks. We will have to define a sound and systematic *evaluation methodology* that could eventually serve as a reference benchmark for gesture-based interaction. The different gestures should be compared for different types of controls and tasks. Collected measures will be quantitative (e.g., time and number of errors) and qualitative (e.g., satisfaction and perceived difficulty). These experimental tasks will serve as a benchmark of testing units in the lab, in the sense that they will consider a control in isolation from any real application domain.

Criteria of success: A sound and systematic way of validating multi-touch gestures. Empirical observations that support or invalidate hypotheses about the efficiency of different mappings.

High-level validation of mappings

As further validation, conclusions from abstract testing units should then be demonstrated in the context of realistic applications to demonstrate the scalability of our findings. We focus on applications for navigating and manipulating large datasets. Emmanuel Pietriga and Caroline Appert have already started working on such applications in the domains of astronomy and geographical information systems. Gilles Bailly also had some experience on multi-touch interaction on geographical information systems (Viard et al., 11). These applications are good candidates as they involve a large number of controls.

Criteria of success: Presenting a set of guidelines that will inform the design of graphical applications that rely on gesture input. Following these guidelines to end with a consistent and efficient set of mappings that enable users with a large variety of controls in at least one of the application domains mentioned above.

A new challenge for gesture-based interaction

The study about gesture transitions would be novel in the panorama of HCI research. Proposing a sound experimental methodology to observe phenomena that characterize the transition phase is also challenging. It requires careful control of factors so as to isolate (a) the specific changes in movements that are due to participants' reaction time to start a new gesture in response to a given stimulus from (b) changes that are only due to motor actions that are required to start another gesture.

Criteria of success: Defining an experimental task that succeeds in capturing a transition between two gestures. Deriving criteria for discriminating a transition phase from a regular execution phase.

(Appert and Zhai, 09) **Caroline Appert** and Shumin Zhai. 2009. Using strokes as command shortcuts: cognitive benefits and toolkit support. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '09). ACM, New York, NY, USA, 2289-2298.

(Appert and Bau, 10) **Caroline Appert** and Olivier Bau. 2010. Scale detection for a priori gesture recognition. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10). ACM, New York, NY, USA, 879-882.

(Bailly et al., 10) **Gilles Bailly**, Eric Lecolinet, and Yves Guiard. 2010. Finger-count & radial-stroke shortcuts: 2 techniques for augmenting linear menus on multi-touch surfaces. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10). ACM, New York, NY, USA, 591-594.

(Cao and Zhai, 07) Xiang Cao and Shumin Zhai. 2007. Modeling human performance of pen stroke gestures. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '07). ACM, New York, NY, USA, 1495-1504.

(Card et al., 91) Stuart K. Card, Jock D. Mackinlay, and George G. Robertson. 1991. A morphological analysis of the design space of input devices. ACM Trans. Inf. Syst. 9, 2 (April 1991), 99-122.

(Cohé et Hachet 12) Aurélie Cohé, A. and Martin Hachet, 2012. Understanding user gestures for manipulating 3D objects from touchscreen inputs. In *Proceedings of Graphics Interface 2012* (GI '12). Canadian Information Processing Society, Toronto, Ont., Canada, Canada, 157-164.

(Fekete et al., 09) Jean-Daniel Fekete, Niklas Elmqvist, and Yves Guiard. 2009. Motion-pointing: target selection using elliptical motions. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '09). ACM, New York, NY, USA, 289-298.

(Foley et al., 84) Foley, James D., Victor L. Wallace, and Peggy Chan. The human factors of computer graphics interaction techniques. Computer Graphics and Applications, IEEE 4.11 (1984): 13-48.

(Ghomi et al., 13) Emilien Ghomi, Stéphane Huot, Olivier Bau, Michel Beaudouin-Lafon, and Wendy E. Mackay. 2013. Arpège: learning multitouch chord gestures vocabularies. In Proceedings of the 2013 ACM international conference on Interactive tabletops and surfaces (ITS '13). ACM, New York, NY, USA, 209-218.

(Häger-Ross and Schieber, 00) Charlotte Häger-Ross and Marc H. Schieber. "Quantifying the independence of human finger movements: comparisons of digits, hands, and movement frequencies." The Journal of Neuroscience 20.22 (2000): 8542-8550.

(Igarashi and Hinckley, 00) Takeo Igarashi and Ken Hinckley. 2000. Speed-dependent automatic zooming for browsing large documents. In Proceedings of the 13th annual ACM symposium on User interface software and technology (UIST '00). ACM, New York, NY, USA, 139-148.

(Jacob et al., 94) Robert J. K. Jacob, Linda E. Sibert, Daniel C. McFarlane, and M. Preston Mullen, Jr.. 1994. Integrality and separability of input devices. ACM Trans. Comput.-Hum. Interact. 1, 1 (March 1994), 3-26.

(Karam and Schraefel, 05) Maria Karam, and M. C. Schraefel 2005. A Taxonomy of Gestures in Human Computer Interaction. Technical report, Electronics and Computer Science, University of Southampton (2005)

(Kin et al., 12) Kenrick Kin, Björn Hartmann, Tony DeRose, and Maneesh Agrawala. 2012. Proton++: a customizable declarative multitouch framework. In Proceedings of the 25th annual ACM symposium on User interface software and technology (UIST '12). ACM, New York, NY, USA, 477-486.

(Lü and Li, 12) Hao Lü and Yang Li. 2012. Gesture coder: a tool for programming multi-touch gestures by demonstration. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12).

(Malacria et al., 10) Sylvain Malacria, Eric Lecolinet, and Yves Guiard. 2010. Clutch-free panning and integrated pan-zoom control on touch-sensitive surfaces: the cyclostar approach. In Proceedings of the

SIGCHI Conference on Human Factors in Computing Systems (CHI '10). ACM, New York, NY, USA, 2615-2624.

(Martinet et al. 10) Anthony Martinet, Géry Casiez and Laurent Grisoni. 2010. The effect of DOF separation in 3D manipulation tasks with multi-touch displays. In *Proceedings of the 17th ACM Symposium on Virtual Reality Software and Technology* (VRST '10). ACM, New York, NY, USA, 111-118.

(Moore et al., 13) Moore, Keith L., Arthur F. Dalley, and Anne MR Agur. Clinically oriented anatomy. Wolters Kluwer Health, 2013.

(Nacenta et al., 13) Miguel A. Nacenta, Yemliha Kamber, Yizhou Qiang, and Per Ola Kristensson. 2013. Memorability of pre-designed and user-defined gesture sets. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13). ACM, New York, NY, USA, 1099-1108.

(Oulasvirta et al., 13) Antti Oulasvirta, Teemu Roos, Arttu Modig, and Laura Leppänen. 2013. Information capacity of full-body movements. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13). ACM, New York, NY, USA, 1289-1298.

(Pietriga et al., 07) **Emmanuel Pietriga**, **Caroline Appert**, and Michel Beaudouin-Lafon. 2007. Pointing and beyond: an operationalization and preliminary evaluation of multi-scale searching. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '07). ACM, New York, NY, USA, 1215-1224.

(Pietriga et al., 12) **Emmanuel Pietriga**, Pierre Cubaud, Joseph Schwarz, Romain Primet, Marcus Schilling, Denis Barkats, Emilio Barrios, Baltasar Vila Vilaro, Interaction Design Challenges and Solutions for ALMA Operations Monitoring and Control, invited paper, Astronomical Telescopes and Instrumentation, SPIE, July 2012, Amsterdam, the Netherlands

(Nancel et al., 11) Mathieu Nancel, Julie Wagner, **Emmanuel Pietriga**, Olivier Chapuis, and Wendy Mackay. 2011. Mid-air pan-and-zoom on wall-sized displays. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11). ACM, New York, NY, USA, 177-186.

(Schieber, 01) Schieber, Marc H. Constraints on somatotopic organization in the primary motor cortex. Journal of neurophysiology 86.5 (2001): 2125-2143.

(Viard et al., 11) Alexandre Viard, **Gilles Bailly**, Eric Lecolinet, Emmnauel Fritsch, E. 2011. Augmenting Quantum-GIS For Collaborative And Interactive Tabletops. Advances in cartography and GISScience. Vo. 1. Lecture Notes in Geoinformation and Cartography. Part 3. 293-307.

(Wei et al., 10) Yu, Wei Shin, Hiske van Duinen, and Simon C. Gandevia. Limits to the control of the human thumb and fingers in flexion and extension. Journal of neurophysiology 103.1 (2010): 278-289.

(Walter et al., 13) Robert Walter, **Gilles Bailly**, and Jörg Müller. 2013. StrikeAPose: revealing mid-air gestures on public displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '13). ACM, New York, NY, USA, 841-850.

(Wobbrock et al., 09) Jacob O. Wobbrock, Meredith Ringel Morris, and Andrew D. Wilson. 2009. User-defined gestures for surface computing. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '09). ACM, New York, NY, USA, 1083-1092.

(Zatsiorsky et al., 00) Vladimir M. Zatsiorsky, Zong-Ming Li, and Mark L. Latash. Enslaving effects in multi-finger force production. Experimental Brain Research 131.2 (2000): 187-195.

(Zhai et al., 12) S. Zhai, P.O. Kristensson, **Caroline Appert**, T.H. Anderson, X. Cao. (2012). Foundational Issues in Touch-Surface Stroke Gesture Design — An Integrative Review. Foundations and Trends in Human-Computer Interaction. 5, 2 (2012), 97-205.