Formalization and Combination of Touch and Point Interaction

Dietrich Kammer, Dana Henkens, Jan Wojdziak
Technische Universität Dresden
dietrich.kammer@tu-dresden.de

Georg Freitag
Hochschule für Technik und Wirtschaft Dresden
freitag@htw-dresden.de

ABSTRACT
Gestural interfaces have been the subject of research in the HCI community for a long time. In contrast to existing examples of multimodal interfaces, which combine free-hand gestures with speech, this paper proposes the combination of both touch and point gestures. Touch requires users to actually initiate contact with the interaction surface, whereas point interaction allows them to use hand movements in the proximity of the user interface. Formalization is deemed necessary to subsequently combine the two gestural interaction styles to realize engineering patterns for multi-touch interfaces. A prototype is presented, showing the potential of combined touch and point interaction in map navigation.

ACM Classification Keywords
H.5.2 Information Interfaces and Presentation: User Interfaces – Input devices and strategies

Keywords
Touch and point interaction, natural user interface, gestures, combination, formalization

INTRODUCTION
Initially triggered by companies like Apple and Microsoft®, multi-touch is integrated in an increasing number of devices. Smart phones, music players, tablets, tabletops, PCs, and more enable users to interact in a direct and more natural manner using touch input. However, multi-touch is not practicable and applicable for interactive systems in all contexts. On the one hand, touch input requires minimal distance between user and input device. However, multi-touch is not practicable and applicable for interactive systems in all contexts. On the one hand, touch input requires minimal distance between user and input device. This is a challenge that is even greater with regard to large-scale surfaces and walls, where exterior regions are not reachable without considerable effort. On the other hand, there is the potential contamination of the input field with dirt or germs, or possible damages as a result of careless and frequent usage. Moreover, media installations, museums, terminals in stores or banks, and other public spaces present the risk of disease transmission. Another challenge of touch interfaces in the context of three-dimensional computer graphics is the cognitive gap between interacting on a two-dimensional space and a three-dimensional visualization. Several cases exist, where users do not want or are not able to contact a sensible electronic device. In consequence, alternative interaction techniques are necessary. By offering both modalities, users are given the freedom to choose their preferred interaction style. The combination of both modalities introduces yet another level of interaction, which promises to enable even more “natural” interaction styles.

The Fraunhofer Institute for Telecommunications, Heinrich Hertz Institute HHI, introduced the term multi-pointing, which refers to touchless interaction with fingers pointing to the screen [6]. In this paper, point is used as a general term for touchless (contactless), gesture-based interaction in three-dimensional space in contrast to touch. The idea presented in this paper is to extend touch interfaces with the modality of touchless hand gestures in the proximity of the touch screen. The goal is to support users in various situations, offering alternative ways of interacting with natural user interfaces. Beside the increased flexibility by providing touch and point gestures, combined interaction techniques using both hands are feasible. As a result, the full range of system functions can be accessed by a novel kind of combined touch and point gestures.

The proposed formalization and subsequent combination of touch and point interaction has two purposes. First, understanding and designing gestural interaction should be supported by identifying the crucial units of composition, e.g. hand shapes and degrees of freedom. Second, such formalization should be precise enough to be integrated into a software engineering process.

OUTLINE OF THE RESEARCH CONTEXT
Development of systems that combine gesture interaction with other input modalities already started in 1982 with Richard A. Bolt’s “Put-that-there” system [3]. Multimodal applications such as XTRA [14], GEORAL [12], SmartKom [8], and AMI [1] followed, but their approach was still focused on supporting speech as additional modality. Streitz et al. differentiate interaction zones, but do not consider a combination of touch and point interaction [13].

Schick et al. present a large-scale wall system allowing users to switch between touch and point interaction “without changing the way they interact” [10]. In addition, Parker et al. compare these interaction techniques in the
context of a large-scale tabletop system in different user
studies [9]. In both solutions, simple deictic gestures are
provided to select and move elements on the surface.
Although accuracy declines when selecting smaller objects
with increasing distance, users prefer pointing gestures and
a combination of touch and point, since the effort is not as
high as using touch only. Tasks are solved more efficiently
on average. Besides pointing gestures, humans use further
hand movements to interact and communicate in the real
world. The challenge is to adapt them to interact with
virtuality in a familiar way. Epps et al. evaluated user
behavior in desktop interfaces determining some
commonly used gestures like flat, angled, and fist hand
shapes as well as 'C' and ‘L’ gesture [4].

FORMALIZATION AND COMBINATION
The authors argue that only formalization and combination
of gestural interaction modes can lead to coherent
engineering patterns for multi-touch interfaces. The
approach of GeForMT (Gesture Formalization for Multi-
touch) in [7] shows how custom multi-touch gestures can
be described in a general and comprehensible way, using
functions, operations, and atomic units of composition,
which are easy to remember. By identifying the similarities
of touch and point interaction, it is possible to extend these
concepts to formalize point gestures as well. Similar
frameworks with a strong focus on formalization exist, for
instance Midas by Scholliers et al. [11]. However, a more
intricate approach is chosen, “transforming the low level
input data into logical facts”.

![Figure 1: Some standard gestures and their
formalization – touch gestures (top) and point
gestures (bottom).](image)

The identification of atomic gesture units is essential to the
presented approach of formalizing and combining different
gesture sets and modalities. De-facto standard gestures
need to be considered (cp. Figure 1). In multi-touch
interfaces gestures like SPREAD or ROTATE have gained
acceptance, whereas systems supporting point interaction
provide various solutions. According to [4], the commonly
used gesture to select is TAP or HOLD. The iPonto
Presenter of the Fraunhofer Institute for Applied
Information Technology FIT [5] shows that touch gestures
can be adapted in this way. GRAB and RELEASE used in
[6] or the more accurate PINCER GRASP [1] are
alternative gestures to select objects. Hand-based gestures
like WAVE and SLAP are suited to activate a system or to
navigate in views. Towards a formalization of gestures, a
differentiation between finger and hand-based interaction
(FINGER, HAND), hand shapes (FLAT, SPREAD, FIST,
CURVED, ANGLED) and hand postures (IN, OUT,
FRONT, REAR, TOP, BOTTOM) is proposed by the
authors. These characteristics can be combined to complex
atomic hand gestures.

Having described atomic gestures, their parameters, such as
movement, have to be defined and formalized as well. This
task is increasingly intricate with more degrees of freedom
that have to be regarded (cp. Figure 2). Whereas touch
gestures are performed in two dimensions (four directions)
in a plane according to the specification in [7], three planes
exist in which point gestures can be performed (transversal,
frontal, sagittal). Six directions of movement can be
identified: UP, DOWN, RIGHT, LEFT, AHEAD and
ABACK. Defining directions in the frontal plane, similar to
touch, is conceivable for point interaction as well.

![Figure 2: Degrees of freedom in touch (left) and point
interaction (right).](image)

With a formal, text based specification of gestures for both
touch and point interaction as proposed here, the
combination of interaction modalities is feasible.
Combining gestures can be achieved by an appropriate
grammar, using dedicated namespaces. To this end, the
grammar in [7] should be extended to include point
interaction. Because the concept of composing atomic
gestures and gestures of different interaction techniques are
similar, composition operators can be reused for this task
on a higher level. Beyond the use of a comma for sequence,
a plus-sign (asynchronous) and an asterisk (synchronous)
describing simultaneous executed gestures. The following
example formalizes the gesture used to tilt the map in the
prototype, which is presented in the next section:

TOUCH: FINGER (HOLD) +
POINT: HAND_SPREAD (AHEAD)

PROTOTYPE
In order to show the concept of formalization and
combination of touch and point interaction, a basic
prototype has been implemented. The showcase visualizes
a globe (Virtual Earth) that provides principal navigation
functions like shifting, scaling, rotating, and tilting. Tilting
allows changing the angle of view on the globe.

Technologies and Frameworks
Microsoft Surface® technology is combined with Wiimote-
Tracking, using a custom-built gesture glove (cp. Figure 3,
left). A Wiimote device is used as infrared camera, which
is fixed above the table (cp. Figure 3, right). In order to be
able to concentrate on the interaction concept, detection of
simple hand shapes is achieved by four infrared markers
(IR-LEDs) attached to thumb, fingers, and wrist. Gestures
are recognized depending on the number of recognized markers (cp. Figure 3, right). This approach has advantages over a video-based solution, particularly with regard to performance benefits in recognizing online gestures, since it keeps the recognition process simple. However, only basic hand shapes can be recognized (FINGER, HAND_FIST, HAND_SPREAD).

Figure 3: The custom-built glove to detect Point gestures (left), construction of multi-touch surface and infrared camera above, including infrared markers for gesture recognition (right).

The open source framework InfoStrat.VE (http://bingmapswpf.codeplex.com/) is used as interface to the bing map service and the visualization of the globe. For getting data from the Wiimote device, WiimoteLib v1.7 is integrated (http://wiimotelib.codeplex.com/).

Interaction

Standard manipulation gestures as supplied by the Surface SDK are used in combination with point interaction above the table as tracked by the custom equipment described earlier. Touch gestures realize scaling, rotation, and translation. Alternatively, HAND_SPREAD is performed along the sagittal plane (cp. Figure 2) and affords zooming or scaling. HAND_GRAB with a movement along the transversal plane (cp. Figure 2) realizes rotation or translation as alternative to touch gestures. Combining touch hold and spreading of the hand along the transversal plane allows tilting of the map (cp. Figure 4). As a simple feedback mechanism, all of the tracked markers above the table are visualized as red dots on the interface. With increasing distance between hand and surface, the dots as well as the distance between the dots enlarge.

CLOSING STATEMENT

This contribution addresses the formalization of touch and point interaction, as well as the subsequent combination. As a result, a variety of functionality can be mastered with gestural interaction. Traditional menus and interface structures can be enriched or partly omitted. In the presented prototype, additional buttons or modes needed to tilt the map with touch interaction only are substituted by a combined HOLD and SPREAD gesture. Moreover, an arguably more natural movement of the user is achieved.

Systems working with a combination of touch and point interaction should be more flexible and make use of similar interaction patterns to be comprehensible. The third dimension above or in front of a multi-touch surface is addressed by the presented approach. The goal of using our approach to gesture formalization to specify gestures in a software framework requires further implementation efforts. In addition, the presented concepts need to be extended and evaluated. Moreover, touch and point gestures need to be recognized reliably and interference must be avoided. Substituting the gesture glove with a video-based solution (e.g. Microsoft® Kinect) for gesture recognition would make the system more usable and less invasive.

Figure 4: Multi-touch scaling using two fingers (top left), HAND_SPREAD to zoom (top right), HAND_GRAB gesture to translate the map (bottom left), and combination of a touch HOLD and point HAND_SPREAD gesture to tilt the map.

Interaction design can be supported by the proposed formalization and additional rules, which address gesture collisions. As a result, detection problems and exclusive gestures might be revealed before conducting excessive user tests. For instance, a point gesture is recognized although the user was simply trying to perform a touch interaction. In our prototype, this is tackled by introducing a speed threshold, i.e. point gestures are only recognized if performed slowly enough. By offering an open, extensible, and flexible way to specify gestures, standardization can be achieved more easily. However, formalization must be done carefully enough to make gesture design easier, but not restrict the freedom to pursue new and creative ways for gestural interaction. User studies as well as the ongoing cultural development of gestural interfaces will reveal appropriate standard gestures.

BIOGRAPHIES

Dietrich Kammer is a research associate and doctoral candidate at the University of Technology in Dresden. His interests include component oriented software engineering and computer graphics. Currently his work is centered on multi-touch technology. A key interest is the formalization
of gestural interaction for the use in appropriate frameworks.

Dana Henkens is a student at the University of Technology in Dresden. Her interests are gesture-based interaction design and information visualization, especially in web-based interfaces. The combination of touch and point interaction was part of her advanced studies as a preparation for her diploma thesis.

Jan Wojdziak is a research associate and doctoral candidate at the Technische Universität Dresden in Germany. His research interests include applied visualistic and interaction design in the range of three-dimensional computer graphics.

Georg Freitag is a research associate and doctoral candidate at the University of Applied Sciences in Dresden. His interests are novel human-computer interfaces, interaction design and information visualization. At present, the focus of his work is on programming environments and frameworks for multi-touch application development.

REFERENCES