Intuitive Gestures on Multi-touch Displays for Reading Radiological Images

Susanne Bay[2], Philipp Brauner[1], Thomas Gossler[2], and Martina Ziefle[1]

[1] Human-Computer Interaction Center, RWTH Aachen University, Germany
[2] Siemens Healthcare, Erlangen, Germany
susanne.bay@siemens.com

Abstract. Touch-based user interfaces are increasingly used in private and professional domains. While touch interfaces have a high practicability for general daily applications, it is a central question if touch based interfaces also meet requirements of specific professional domains. In this paper we explore the applicability of touch gestures for the domain of medical imaging. We developed a set of intuitively usable gestures, applicable to different screen sizes. The development was entirely user-centered and followed a three-step procedure. (1) The gesture set was developed by asking novices to propose possible gestures for different actions in medical imaging. (2) The gesture set was implemented in a commercial medical imaging solution and (3) evaluated by professional radiologists. The evaluation shows that the user-centered procedure was successful: The gestures did not only work equally well on different screen sizes, but revealed to be intuitive to use or easy to learn.

Keywords: Multi-touch, gestures, medical imaging, radiology, intuitiveness.

1 Motivation

Multi-touch displays are a widespread technology for consumer products like mobile phones and tablet PCs. These devices host a variety of applications which are primarily used in common, everyday scenarios, such as internet browsing, messaging, photo viewing, etc., and are widely accepted and appreciated. The usage of multi-touch for highly specialized professional applications is not trivial but for each specific application field the most frequently performed interactions in the specific scenario need to be translated into common multi-touch gestures. Also, it is not clear whether multi-touch interactions are appropriate for performing highly specialized tasks which may have different requirements on efficiency, precision, and accuracy than the above mentioned "everyday" tasks. In the field of radiological imaging there is a high interest of professionals in accessing their cases from anywhere in order to be able to provide expert feedback in all types of situations, e.g., when being asked for advice by a colleague via telephone, in a clinical conference (tumor board) or when explaining the diagnosis to patients. Therefore, the usage of tablet PCs or smart phones seems to be a valuable option. However, no standards exist on how to translate the most important

S. Yamamoto (Ed.): HIMI/HCII 2013, Part II, LNCS 8017, pp. 22–31, 2013.

functionalities for the interaction with radiological images to multi-touch gestures. Since radiologists often use different software from different vendors to read their cases, it would be a great benefit for this user group if medical vendors agreed on a standard for the multi-touch gestures because this would enable them to use different devices and applications without transition costs. Also, from a cognitive ergonomic point of view it is not clear whether it is possible to identify a uniquely prototypic gesture set that meets medical professionals' needs regarding the expressiveness of gestures in form and content, and is also intuitive to use and easy to learn. This paper presents an empirical study that evaluates multi-touch gestures for the interactions needed when reading radiological images.

2 Method

To develop and test an intuitive gesture set for interacting with medical images we used an iterative empirical-experimental approach: First, we identified intuitive gestures by letting non-radiologists perform possible gestures on a paper prototype. Second, we identified common features among the gestures and compiled these into a complete gesture set. Third, we asked two medical professionals for applicability of the gesture set. Fourth, the gesture set was implemented into a professional imaging solution and radiologists as well as non-radiologists evaluated the gesture set on three different display sizes. The functions required for interacting with medical images are closely related to the physical form of data and the requirements of the radiologists carrying out the diagnoses. Hence, we will briefly introduce the very basics of medical imaging before we detail the empirical procedure.

2.1 Radiological Imaging and Frequently Used Functions

Medical imaging is the technique and process used to create images of the human body (or parts and function thereof) for clinical purposes (medical procedures seeking to reveal, diagnose, or examine disease) or medical science [4]. Radiologists have the task to review and interpret 2D, 3D or 4D (3D data acquired over a period of time) images. Due to the high amount of imaging data produced by state-of-the-art radiological imaging technologies like computed tomography and magnetic resonance imaging, radiologists need efficient techniques to visualize (e.g., in different planes or as volume), manipulate (e.g., change contrast and brightness) and navigate (e.g., scroll through stacked images or rotate volumes) the image data provided. There are uncountable functions in professional medical imaging solutions. For this work we focused on the most commonly used functions and operations that professional radiologists use in their daily work. Radiologists typically work with both two-dimensional and three-dimensional image material.

For two-dimensional images the most frequently used operations are: *Zoom and Pan, Scrolling through a Stack*, and *Windowing (changing brightness and contrast).* The zoom operation allows radiologists to magnify a specific area of an image, whereas the pan operation allows changing the viewport of the given image. The scrolling through a stack operation is used to display different layers of the current image. With this operation radiologists are able to scroll along the axis orthogonal to the display. Radiologists require two types of scrolling: exact / step-wise scrolling (e.g. next/previous layer) and quick scrolling (e.g., to quickly scan the abdominal area for malign tissue).

For three-dimensional material the most often used operations are *Rotating a Volume* and also *Zoom* and *Pan*. *Pan Zooming* and *Panning* 3D images is equivalent to the 2D case. For both two- and three-dimensional images the material is usually displayed in a grid of multiple windows (e.g., one window showing a 3D image, and one window for 2D views in different orientations (e.g., sagittal, coronal and axial plane). The *Blow-up* and the *Blow-down* operations are used to display one of these image segments maximized or to restore the previous grid view.

2.2 Generation of a Gesture Set

To extract intuitive gestures for interacting with medical images we recruited 14 unpaid participants (8 male, 6 female) for a user study. None of them had any experience in medical imaging or medicine. Also, some of the participants had little expertise with touch displays, such as smart phones or tablet devices.

We first gave a brief overview about radiology and the frequently used functions as described above. After that we also presented videos of the effect of each function, in order to support the understanding of the functions and their effects on the displayed images. The participants had the opportunity to ask questions or review the videos at any time. The participants were then asked to perform each gesture on a paper prototype of a medical imaging solution. We monitored the hand and finger movements of the participants with a camera attached to the participant's chest. The approach of presenting the desired outcome of a gesture and letting users perform possible actions is similar to the one used by Wobbrock et al. [3].

All participants had to perform the gestures in the same order (first 2D gestures, then 3D gestures). Each participant performed the gestures twice: once on a small size display (phone-size or tablet-size) and once on a wall-size display (*24"* or *48"*). The size and the order of the paper prototypes were randomized across the participants. After the experiment the performed gestures were classified. Hereto, we first developed a categorization scheme by viewing the video recordings, discussing common features, and defining a set of gesture categories for each gesture. The categorization scheme includes multiple dimensions such as the number of fingers or hands involved or the type of gesture performed.

After that two researchers independently reviewed the material and classified the gesture executions accordingly. We classified the proposed gestures according to the classification scheme. Only rough estimates of the number of mentions will be reported as not all proposed gestures fit in exactly one category. In the following we use the terms *frequently*, *commonly,* and *rarely* for propositions that were made respectively by over *2/3*, *1/3* to *2/3* or less than *1/3* of the participants.

Scrolling through a Stack: Frequently, participants proposed a gesture that utilizes an imaginary scrollbar at the side of the screen (similar to a finger on a telephone book page). Also *frequently* suggested was a swipe gesture in which a finger (small screen) or hand (large screen) was slowly moved across the surface. *Commonly* suggested was a flick gesture in which a finger/hand was rapidly moved across the surface.

Zoom: For zooming participants *frequently* proposed a pinch to zoom gesture. It was either performed with two hands on large screens or with two fingers on small screens. Other *rare* suggestions were opening and closing the hand (all fingers involved) and using a button instead of a gesture.

Pan: Participants *frequently* suggested a tap-drag gesture. However, they disagreed regarding the number of fingers/hands to use. Roughly half of them suggested using one finger/hand, whereas the other half suggested using two. *Rarely* suggested was a gesture that uses the whole flat hand to pan an image on large screens.

Windowing: A variety of gestures were proposed for this function. *Commonly* suggested were a set of two sliders, either visible on demand or permanently on screen (comparable to the set of scrollbars on desktop systems). Another gesture also *commonly* suggested was opening and closing the hand (described as rising and sinking sun). However, this gesture offers only 1 instead of the required 2 degrees-of-freedom. A tap-drag gesture on a 2 dimensional plane was *rarely* suggested: Dragging along the horizontal axis changed the window width and dragging along the vertical axis changed the window height. Again *commonly* proposed was the use of a menu button instead of a gesture.

Rotating a Volume: *Frequently,* the participants fixated a point with a finger (small screen) or hand (large screen) on an imagined sphere and rotated that sphere by dragging the finger/hand over the surface. Thus, novices proposed a gesture that resembles the popular *ARCBALL* technique by Shoemake [2]. Separate buttons for rotating the object instead of a gesture were proposed only *rarely*.

Blow-up and *Blow-down*: *Frequently,* a double tap gesture was proposed that either expands the segment in which it was executed or reverts from full screen to the previous state. A *rarely* made suggestion was a tap-drag gesture that moves the segment to be maximized to the center of the screen.

Overall, the proposed gestures were basically the same for small and large displays, showing that radiological gestures are generally prototypic. The only notable difference is that gestures were performed with the whole hand on large displays whereas the fingers were used on small size displays. With the exception of the *Windowing* function, on the whole, the participants proposed the same gestures for each of the different functions in medical imaging. Thus, we can assume that we have found a gesture set that is intuitive and universal for different display sizes.

2.3 Cross-Validation of the Generated Gesture Set

Users who had no knowledge of medical imaging proposed the gesture set. Thus, before the gestures were implemented into a functional prototype and before a formal user study with radiologists was carried out, we gathered professional feedback from two radiologists that was, in general, positive. However, they criticized the lack of a common gesture for the *Windowing* function and they coherently suggested a tap-drag gesture that adjusts the window width and height by dragging the finger along the axis. We therefore continued with the implementation and formal evaluation of the gestures and their universality across different display sizes.

The gestures for *Windowing*, *Scrolling through a Stack*, *Blow-up* and *Blow-down*, *Pan* and *Zoom* and *Rotating a Volume* were implemented as shown in Figure 1.

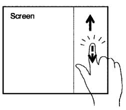
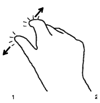
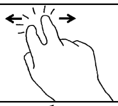
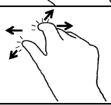
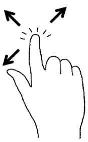
	*Windowing:* To change the window (brightness and contrast), a tap-drag gesture has to be performed. Moving the finger vertically changes the image's brightness. Moving the finger horizontally changes the contrast.
	*Scrolling through a Stack: Flick:* To scroll though the images layer by layer, the *flick* gesture can be used. A finger has to be placed on the surface and quickly moved up- or downwards and released immediately from the surface. In contrast to the scrollbar (see below) this gesture can be performed anywhere on the image.
	*Scrolling through a Stack: Scrollbar:* The scrollbar can be used to quickly skim through the image stack. For this the finger has to be placed on the right hand side of the image and moved up- or downwards.
	*Blow-up / Blow-down:* To perform a Blow-up or Blow-down (activate/deactivate full screen-mode for an image segment), a double tap has to be performed.
	*Zoom:* To enlarge the image, two fingers must be placed on the surface. Then the increasing or decreasing the distance of the fingers will enlarge or shrink the displayed image.
	*Pan:* To pan an image, two fingers have to be placed on the surface. If the fingers are moved the image is panned in the same direction and speed.
	*Zoom and Pan:* It is possible to perform the gestures for zoom and pan simultaneously.
	*Rotating a Volume:* To rotate a 3D-model, a single finger has to be placed on the surface. On an imaginary sphere, the finger fixates a point. By moving the finger, the sphere is rotated with the point and the finger linked together.

Fig. 1. Generated multi-touch gesture set

2.4 Evaluation of the Gesture Set

A development release of a medical imaging software[1] was modified to support touch events. One research goal was to identify gestures that are universal to different display sizes. Therefore we tested the gesture set on multiple display sizes: a 4” mobile phone display, a 10” tablet display, and a 60” wall-sized display. In the following the three sizes will be referred to as *phone-size*, *tablet-size* and *wall-size*.

2.4.1 Experimental Setup

In the experiment, we evaluated the gestures as well as three different display sizes. The order of the display sizes was randomized across the participants. The gestures had to be performed in a fixed order: First 2D gestures and then 3D gestures. The gestures were performed as part of a mock medical diagnosis. For example, to evaluate the *Windowing* gesture, the radiologists were asked to modify the window setting to investigate first the lung, then soft tissue. In addition to the study of isolated gestures, participants also had to perform two complex tasks (one 2D, one 3D) in which all gestures had to be used. Participants had to rate each gesture according to its intuitiveness, perceived ease of use, learnability, precision, and efficiency. In addition to the gesture ratings by participants, a post-hoc video analysis of the gesture executions was accomplished as external validation in which an expert evaluated the intuitiveness, ease of use of the gestures, and the kinds of errors that occurred.

After the experiment, the participants rated each display size for its suitability for medical diagnoses, the overall quality of the display, the precision, and the intention to use touch-based displays in medical imaging. Figure 2 shows a user performing a *Zoom* gesture on the wall-sized display.

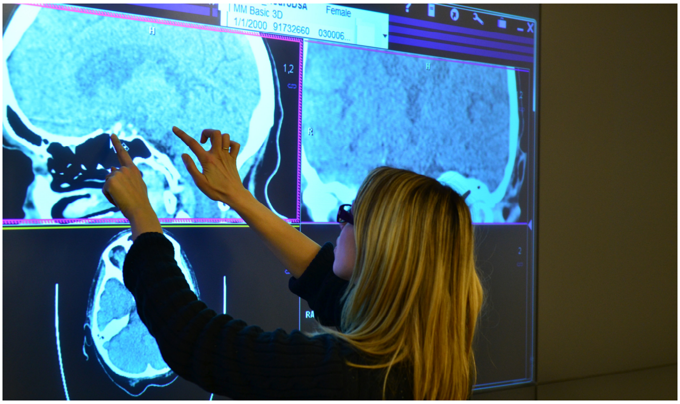


Fig. 2. A user performing a *Zoom* gesture on the wall-sized display

---

[1] *syngo*.via from Siemens Healthcare was used for evaluation. The software is a medical imaging product for radiologists offering routine and advanced reading functionality for multiple modalities like MRI, CT and PET-CT.

2.5 Results

Due to the comparably small number of participants, we report on descriptive outcomes rather than inference statistics.

In total, 24 participants (50% women, 50% men) took part in this study. 13 were professional radiologists (in the following called *experts*) and 11 subjects had no experience in radiology or medicine (called *novices*). As experts were the main target group for the application to be developed, we concentrate on the insights gained from observing the experts. Findings from the novices will be reported where appropriate. On average, experts had 13 years of work experience (5 had more, 8 had less than 10 years of professional experience). 6 participants stated that they have made more than 100,000 diagnoses, with another 3 reporting over 10,000 diagnoses so far.

Gesture Set. The gesture executions were assessed by a post-hoc video analysis. For each gesture the number of help cues was counted and the perceived ease of use was rated.

The observed intuitiveness was in general high for all but the two *Scrolling through a Stack* gestures. The participants executed over 90% of the requested gestures without additional cues from the examiners. Especially the combination tasks were completed without significant help. Yet both gestures for *Scrolling through a Stack* show room for improvement. The *Flick* gesture was used intuitively in only 43% of the trials and in 14% of the trials more than one cue from the examiners was needed. The *Scrollbar* performed better: 74% of the trials were done correctly without any cues. Still, in 11% of the cases more than one cue was necessary (see Figure 3). These findings conform to the observed ease of use during the gesture execution that was also high for all but the two *Scrolling through a Stack* gestures. Additionally, we observed that the participants frequently performed *Windowing* instead *Panning*; both gestures were designed as a tap-drag gesture (the former with one finger, the latter with two fingers). Both combination tasks (diagnoses with multiple gestures) were performed without additional help by almost all participants (see Figure 3).

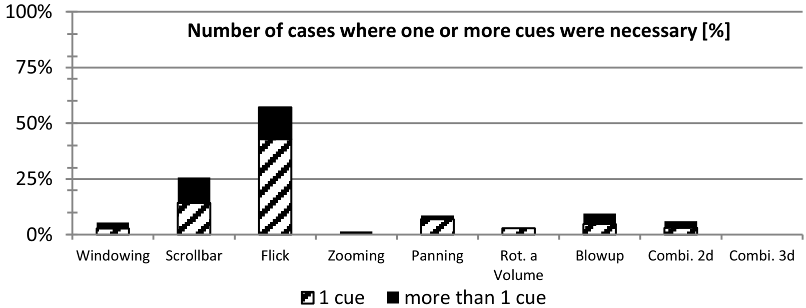


Fig. 3. Observed intuitiveness for gesture executions

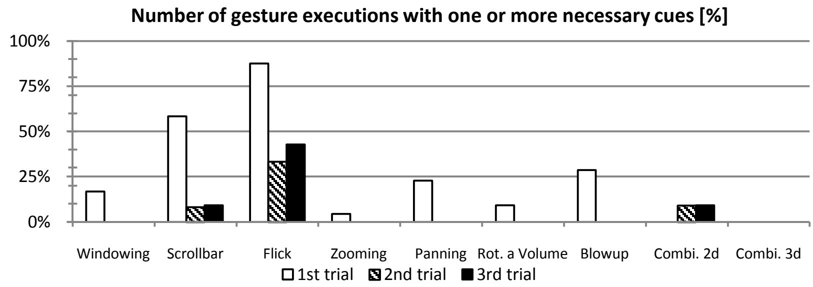


Fig. 4. Learnability of the gesture set

Based on these findings we investigated the learnability of the gesture set, i.e., we studied for each gesture whether the number of cues needed decreases with the number of trials. Indeed, for *Windowing*, *Zooming*, *Panning, Rotating a Volume* and *Blow-up* gestures, cues were only necessary during the 1st trial. In later trials, all participants executed the gestures without additional support from the examiners.

In the 1st trial, the *Scrollbar* gesture required external cues in 58% of the cases. This drops to 9% for the 3rd trial. The number of necessary cues for the *Flick* gestures drops by factor two between the 1st and the 3rd trial. Although this proves a tremendous learning effect, there are still 43% gesture executions that were not performed autonomously by the participants (see Figure 4).

Display Sizes. The rating of tablet-size outperformed the rating of phone-size and wall-size in every dimension (see Figure 5). The intention to use a touch-based medical imaging solution was highest for the tablet (on average +33 points on a scale from -100 to +100), followed by the phone (-20 points) and the wall (-22 points). Likewise the expected usage frequency was highest for the tablet (+61 points); in contrast, phone-size (0 points) or wall-size (-27 points) were rated rather low. *Novices*, however, rated the wall-sized display highest, followed by tablet and then phone. We argue that they might have judged from the patient's perspective and that they might prefer the large display for doctor-patient-communication as they did in other studies [1].

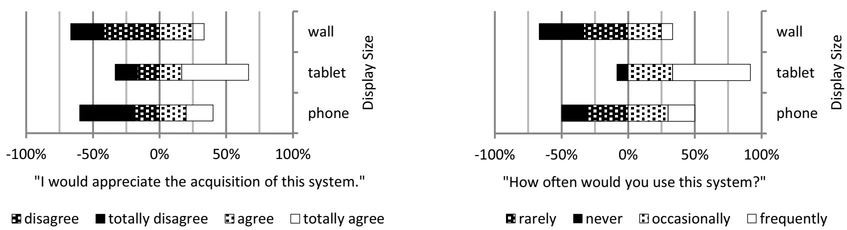


Fig. 5. Desire to use and expected usage frequency dependent on display size

The evaluation of the phone-sized display dominated the wall-size in all but one dimension: the adequacy of screen size. On the phone the available screen space is regarded as insufficient for diagnoses. Radiologists stated that a tablet might be more useful for discussing the findings with patients than doing the actual diagnosis. They dislike using a phone for this purpose as they consider the displays too small.

Touch-Based Interaction in Radiology. Participants had to indicate before and after the experiment whether they would use touch-based interaction for their daily routine and whether they judge touch-based interaction useful in the domain of radiology. At the beginning, the desire to use touch interaction for diagnoses was high (M = 48 points on a scale from -100 to 100%) and increased by 59% to 76 points after the experiment. The perceived usefulness of touch interaction was equally high (48 points) and grew by 31% to 63 points (see Figure 6).

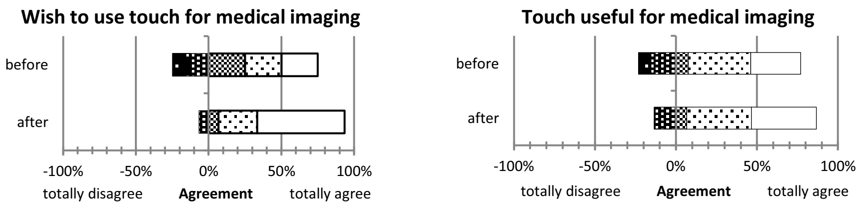


Fig. 6. Intention to use touch displays before and after the experiment

3 Discussion and Future Work

Overall, the study showed that touch-based interactions are a highly promising interaction mode, even in specialized professional areas such as medical imaging. We could reveal that there are prototypic gestures which are perceived as useful by medical professionals. Therefore, the findings represent a promising basis for the development of a standard for multi-touch gestures. A noteworthy finding is that all medical professionals were even more enthusiastic about the usefulness after they had worked with the system. This shows that any evaluation of novel systems profits from real, hands-on experience and confirms the adequacy of user-centered approaches in technical developments. In addition, medical professionals were not only highly willing to contribute to the development in this specific medical domain, but were even glad to provide their professional point of view before a system is marketed.

Display sizes: In general, the large multi-touch wall was evaluated as insufficient: The participants disliked the rather low pixel density and the too large information display. Most criticized was the low precision when interacting with the device. This is caused by our technical setup which is prone to errors due to the use of computer vision, network latency, and the interplay of multiple computers. However, novices liked the large display more than the other sizes. The mismatch between the medical professionals' and the novices' evaluation of the suitability of the wall-sized display might be based on the different perspective (medical professional vs. patient). The tablet-size is evaluated very well and dominated the phone- and wall-sized display in subjective ratings as well as in error metrics. Still, we noticed that participants with long work experience also appreciate the phone-sized display. Interviews revealed that they appreciate the small display for being able to perform diagnoses remotely.

Gesture set: The developed gesture set for interacting with medical data is intuitive and easy to learn. Furthermore, it is suitable for various display sizes, such as smart phones, tablets, or wall-sized displays. Still, two gestures show potential for improvement: Both gestures for *Scrolling through a Stack* were not intuitive as their correct execution required external help. The *Scrollbar* has shown great learnability and is remembered after the first trial. The *Flick* gesture also showed a strong learning effect, but some participants had difficulties recalling this gesture even after the 3rd trial. In addition, we learned that the gestures for *Panning* and *Windowing* are conflicting. Both were implemented as tap-drag: *Windowing* with one finger, *Panning* with two fingers. Participants frequently mixed up both gestures in the beginning.

Thus, the task of creating a *completely* intuitive gesture set could not be achieved in this study. Nevertheless, we have developed a gesture set that was *mostly* intuitive and non-intuitive gestures were easy to learn. Especially, combination tasks, which reflect the work practice of radiologists, were performed without any difficulties.

Overall, the study has shown the high acceptance of multi-touch gestures for interaction with radiological images. The gesture set, however, should be re-evaluated under more stable technical conditions and in a set-up that better reflects the radiologists' work situation. Also, it is planned to evaluate how this gesture set can be extended to non-contact interaction which would be beneficial for interventional radiology and surgery where images need to be manipulated in a sterile environment.

Acknowledgements. Thanks to all participants, but especially the medical professionals, for their time and willingness to share their professional view with us. Thanks also to Luisa Bremen, Tatjana Hamann, Eva Dickmeis, Felix Heidrich, Chantal Lidynia, Oliver Sack, Andreas Schäfer, and Frederic Speicher for research assistance.

References

1. Beul, S., Ziefle, M., Jakobs, E.M.: How to bring your doctor home. Designing a telemedical consultation service in an Ambient Assisted Living Environment. In: Duffy, V. (ed.) Advances in Human Aspects of Healthcare. CRC Press (2012)
2. Shoemake, K.: ARCBALL: a user interface for specifying three-dimensional orientation using a mouse. In: Proceedings of the Conference on Graphics Interface 1992, pp. 151–156. Morgan Kaufmann Publishers Inc., San Francisco (1992)
3. Wobbrock, J.O., Morris, M., Wilson, M.: User-defined gestures for surface computing. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 1083–1092. ACM, New York (2009)
4. Medical imaging - Wikipedia, the free encyclopedia, `http://en.wikipedia.org/w/index.php?title=Medical_imaging&oldid=537453732` (last accesed February 14, 2013)