ShadowGuides: Visualizations for In-Situ Learning of Multi-Touch and Whole-Hand Gestures

Dustin Freeman[1], Hrvoje Benko[2], Meredith Ringel Morris[2], Daniel Wigdor[3]

[1]Department of Computer Science
University of Toronto
Toronto, ON, M5R 2Y8
dustin@cs.toronto.edu

[2]Microsoft Research
One Microsoft Way,
Redmond, WA 98025
{benko, merrie}@microsoft.com

[3]Microsoft Surface
One Microsoft Way,
Redmond, WA 98025
dwigdor@microsoft.com

ABSTRACT

We present ShadowGuides, a system for in-situ learning of multi-touch and whole-hand gestures on interactive surfaces. ShadowGuides provides on-demand assistance to the user by combining visualizations of the user's current hand posture as interpreted by the system (feedback) and available postures and completion paths necessary to finish the gesture (feedforward). Our experiment compared participants learning gestures with ShadowGuides to those learning with video-based instruction. We found that participants learning with ShadowGuides remembered more gestures and expressed significantly higher preference for the help system.

Author Keywords

Gesture learning, multi-finger, displacement, marking menus.

ACM Classification Keywords

H.5.2 [Information interfaces and presentation]: User Interfaces. – Input devices and strategies; Graphical user interfaces.

INTRODUCTION

Multi-touch and whole-hand gestures have the potential to enable intuitive, efficient interaction on interactive surfaces. Furthermore, the use of multiple fingers and whole-hand shapes increases bandwidth relative to single-touch and pen gestures. However, there is little convergence in user expectation in the mapping of multi-touch gestures to system actions, except for simple gestures (i.e. one finger, one hand) [21].

Complex gestural systems require that users learn physical commands. This may be a worthwhile investment but can present a substantial barrier for infrequent or novice users. Consequently, few manufacturers of commercial devices have implemented multi-touch gestures beyond the basic spatial manipulations described by Schneiderman [13].

Several learning techniques use in-situ visuals that enable the user to *learn while doing* single-touch and pen gestures (e.g.: [1,9]). Such techniques provide a gradual transition from novice to expert use without requiring any drastic change. However, teaching multi-touch and whole-hand gestures is a larger problem than single-touch gestures, primarily because the hand pose and the number of contacts can vary, both in the initial contact posture and throughout performance of the gesture. Since posture is irrelevant in the case of pen and single-touch gestures, systems for teaching such gestures have focused on the movement phase of the gesture, rather than the contact phase [1,9].

In contrast, designers of multi-touch systems must guide users as to which posture they should use to initially contact the display, as well as how that pose should change during the gesture (e.g., the user might start the gesture with two fingers, but end with five). Detecting the initial hand pose has been referred to as the 'registration phase' [23] and 'detection of intention' [2]. In the case of gestures on a touch interface, the user must decide what pose to use to register the gesture *before* making the initial contact with the surface. In most cases, this is before the system is able to sense the location or posture of the hand [5]. Because of this, existing teaching systems for multi-touch gestures have relied on visualizations divorced from user actions, in dedicated gesture-learning areas of the screen (e.g. [4]).

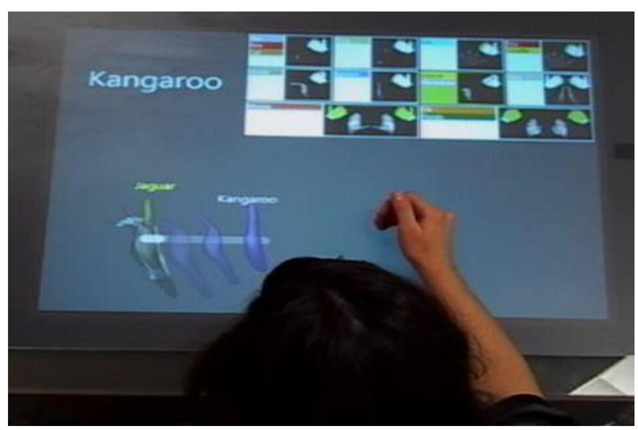

Figure 1. A user learning multi-touch gestures with ShadowGuides. The Registration Pose Guide is seen above the user's hand, and the User Shadow Annotations are to the left.

We present *ShadowGuides*, a multi-touch gesture learning system that displays a dynamic guide within the context of the user's current action (Figure 1). For unobstructed viewing, our visualizations are displayed relative to a displaced raw representation of the user's input, i.e., the *user shadow*. Our visualizations consist of two parts: (a) the *registration pose guide* informs the user what other alternative registration poses are possible in the system and (b) the *user shadow annotations* demonstrate gesture completions available from the current hand pose.

To ensure that our technique is effective for teaching a wide range of gestures, we created a taxonomy that spans the space of multi-touch and whole-hand gestures and used it to select a representative gesture from each area to ensure coverage when testing our system. For evaluation, we compared ShadowGuides to a video instruction help technique; users were tested on the efficiency of using each system and on retention when help was removed.

This paper contributes: (a) a taxonomy for the space of whole-hand and multi-touch gestures, (b) a gesture set spanning this taxonomy that can be used by the community for future comparative tests, (c) a set of in-situ visualizations called ShadowGuides for teaching such gestures, and (d) a formal user study that shows the benefits of ShadowGuides when compared to video-based learning.

BACKGROUND AND RELATED WORK

Wobbrock et al. [21] classified surface gestures along four dimensions: form, nature, binding, and flow. We discuss how our own gesture taxonomy relates to their work in the Taxonomy of Surface Gestures section.

Wu et al. [23] described the process of gesture performance as a finite state machine, with start position (*registration*), a dynamic phase (*continuation*), and end position (*termination*), similar in concept to that described in Charade [2]. In the case of surface gestures, the user registers the gesture with an initial posture when they first touch the device, then continues by performing some actions (possibly further disambiguating the gesture) while maintaining contact, and then terminates by lifting their hand from the device. We describe how ShadowGuides addresses these stages in the ShadowGuides section.

Bau and Mackay [1] discussed the design space of feedforward and feedback mechanisms for single-touch gestures and introduced the concept of dynamic guides, which we extend. One dimension not described in Bau and Mackay's design space is the degree of co-location of the learning space and performance space. Extreme ends of this dimension are learning in-situ, versus learning in a completely separate mode.

Some systems have taught gestures in-situ, such as marking menus [9,16] and dynamic guides [1]; such systems lead the user through the *continuation* portion of gestures. Since these prior systems are for single-point gestures, registration is assumed to be the contact with the surface and is not explicitly taught. Marking menus are restricted to angular gestures, while dynamic guides are generalized to make almost all single-stroke gesture shapes teachable.

Other approaches show available gestures in an area separated from the workspace, essentially a 'cheat sheet' with varying levels of interactivity. Brandl et al. [4] demonstrated a persistent display of available gestures and their function in context, indicating what the effect would be of using pen, single touch, multi-touch, or the whole hand. Bragdon et al.'s GestureBar [3] used a persistent toolbar, including a practice area, to teach pen gestures.

While ink trails are suitable feedback for pen or single-touch gestures, we use a user shadow to show the shape of contact the user is making. This type of feedback has not been used before for teaching symbolic gestures, but "shadows" have been used before for groupware [15] and as a form of input [14].

Vanacken et al. [18] proposed that a system could detect a user performing gestures ineffectively, and suggest a more efficient method after the gesture had been performed. We refer to the most effective way to perform a gesture as *expert style*. Expert style includes factors such as which fingers or body parts to use, the scale of the gesture, and the pressure to apply. Many aspects of expert style may not be detectable by the interface. For ShadowGuides, our goal was to provide guidance for expert style before and during performance of the gesture, heeding the guideline that a teaching system should guide novices to perform the same physical actions as experts [8,24].

Observing demonstration of a gesture is insufficient for learning; it does not convey which elements of the demonstrator's behavior are relevant to the system, and which are not. This has been noted in general for teaching with animated or video demonstrations [17]. There have been two approaches to convey the most relevant aspects of a gesture: use a demonstration or literal representation of the gesture with the important parts emphasized, or use a symbolic notation that the user must learn. In demonstration-based approaches, emphasis can be achieved in two ways: reduction of information content so that only the absolutely relevant parts for the gesture to be recognized remain, or from highlighting important parts of the demonstration. Highlighting emphasis is seen in GestureBar's arrows and short text descriptions [3]. Charade required users to learn a notation along with the gesture set [2]. Notations describing finger, hand and whole body movement sequences are frequent in other fields, such as sheet music and dance notation [10]. Generally, these notations are intended for dense communication between experts, and are not approachable by a novice.

Our goals in developing ShadowGuides were to teach all phases of the gesture (registration, continuation, and termination), to provide feedforward mechanisms to guide novice users, to provide help in-situ to reduce the need for task switching, to provide a modeless gesture system so that novices and experts perform the same gestures, and to teach expert style in performance of gestures.

TAXONOMY OF SURFACE GESTURES

The intention of our work was to produce a gesture learning tool that is useful for teaching a wide variety of multi-touch gestures. Given all the finger and hand pose variations possible on an interactive surface, the space of possible gestures is very large. This makes it difficult to design a visualization technique that comprehensively covers this space. However, to ensure that our technique spans the gestural space, we devised a multi-touch gesture taxonomy. We then used this taxonomy to pick a representative gesture from each category to build a set for evaluation. We based our taxonomy on physical variations in the registration and continuation stages of the gesture, and on the range of variability in hand poses.

Wobbrock et al. [21] provided a taxonomy of user-defined surface gestures divided into the categories of *form, nature, binding, and flow*. While their categories are important in describing both the appearance and consequence of a gesture, our work focuses on teaching the user *how* to perform a gesture and we therefore focus only on the *form* of the gesture. We expand Wobbrock et al.'s form category along the following three dimensions: *registration pose*, *continuation pose*, and *movement*, as seen in Table 1.

We make a distinction between gestures performed with fingertips and those performed with other parts of the fingers or hand (e.g., the palm). Finger-tip based interactions are usually treated as point-based manipulations (similar to mouse actions), while hand shape interactions usually require detection and classification of the shape's contour, orientation, and/or area (e.g., [6]). We define movement by whether the user's entire hand follows a path along the surface. If the hand remains stationary, but the fingers move relative to each other (e.g., in a "pinch" gesture), we classify that gesture as having no path, but a dynamic continuation pose since the posture of the hand changes, but not its location. Gestures where fingers come in contact with (or leave) the surface, or where the hand in contact changes shape, are also classified as having a dynamic continuation pose.

Table 1. Taxonomy of multi-touch and whole-hand surface gestures.

Registration Pose	*Single Finger*	Initial touch with a single finger
	Multi-Finger	Initial touch with multiple fingers
	Single Shape	Initial touch with a single hand shape ('blob') (e.g., a palm down)
	Multi-Shape	Initial touch with multiple hand shapes (typically bimanual)
Continuation Pose	*Static*	Hand pose remains the same after registration; no relative movement
	Dynamic	Hand pose changes after registration (e.g., new fingers come in contact with the surface)
Movement	*No path*	Hand stays in place
	Path	Hand moves along a surface path

Features that cannot be reliably detected by most multi-touch hardware are excluded from our taxonomy (e.g., those requiring sensing in the 3D space above the surface [19] or pressure changes and disambiguating between multiple users [12]). Additionally, many gestures can be distinguished by varying the timing of their performance. However, teaching these features for a particular gesture is beyond the scope of this work and we defer this to future work. Consequently, these features are not included in our taxonomy, and we make no distinction if a gesture is performed with a particular timing, pressure, 3D posture, or multi-user configuration.

Representative Gesture Set

To create a representative set of gestures that covers all 16 cells (4 x 2 x 2) of our taxonomy, we incorporated some existing gestures found in literature, and created new ones when no applicable gesture was found. The entire set is shown in Table 2 and viewable in the video accompanying

Table 2. Our taxonomy of multi-touch and whole-hand gestures. The table contains the spanning set of gestures used to test ShadowGuides. Gestures drawn from prior work cited as appropriate.

Registration Pose				Movement	Continuation Pose
Single Finger	**Multi Finger**	**Single Shape**	**Multi-Shape**		
1-finger down + tap other [21]	5-finger crumple	Opening fist	2 hands flat + stretch fingers	**No Path**	Dynamic
1-finger right, add 2nd finger & pull down	2-finger pull together + back	Opening fist + move left to right	2 palms + thumbs stretch, move up	**Path**	Dynamic
1-finger tap†	4-finger tap	Corner hand side [6]	2 palms down, thumbs out	**No Path**	Static
Pigtail (left to right) [11]	5-finger rotate (45 degrees ccw) [12]	1-hand edge swipe [22]	2 hand edges outward swipe	**Path**	Static

† Reserved for invoking the help system.

this paper. The gestures themselves do not have any particular iconic or symbolic meaning, and were intentionally selected without consideration of any particular command or use case. Instead, gestures were chosen with the assumption that they were representative of their taxonomic category. In some cases, we intentionally chose multiple gestures having the same registration pose so we could demonstrate the overlaying of multiple user shadow annotations, discussed later.

While the cell for "single finger, static continuation pose, and no path" (i.e., a single-finger touch gesture) is included in the taxonomy for completeness, that gesture is the most basic interaction on any touch interface and thus we did not attempt to teach it in our study. Instead, we used the single-finger tap as a means of invoking our help system without executing a command (similar to the behavior in [1] and [9]). Consequently, we used 15, rather than 16, gestures in our evaluation of ShadowGuides.

It is our intention that this taxonomy and spanning set becomes a useful tool for the community a basis of comparison between different gesture teaching systems.

SHADOWGUIDES

ShadowGuides is a gesture learning tool, implemented on a Microsoft Surface, which provides on-demand assistance to the user by combining visualizations of the user's current hand pose as interpreted by the system (*feedback*) and hand poses and completion paths necessary to complete the gesture (*feedforward*). We do so by annotating the *user shadow*, the raw hand image captured by the interactive surface, which we display at an offset from the user's hand (Figure 2). Due to the nature of the Microsoft Surface hardware, the user shadow appears even when the user is slightly above the interface, yet not touching. We started using the user shadow to increase the user's awareness of how their contact with the interface was interpreted, and this aided them to understand and diagnose problems.

Placing annotations on the user shadow rather than directly beneath the user's hand has several advantages. The most important is the elimination of occlusion problems, noted previously by Vogel et al. [20], where an annotation might be hidden beneath the hand. The user shadow also provides direct feedback on what the system is "observing," giving the user feedback on what parts of their hand are in contact with the surface and therefore what parts are interacting with the system. This is particularly important as users are often unaware of additional fingers or parts of their hand touching the surface [21], but few systems have any means of discarding such undesired contacts. We use the user shadow concept as the basic building block for all our visualizations.

ShadowGuides consists of two parts: a *registration pose guide* for teaching various registration poses and *user shadow annotations* that guide the user from the current hand pose through completion of the gesture. The separation of these two parts follows directly from our taxonomy, where the registration pose is classified separately from the continuation pose and movement.

Figure 2: The user shadow (left) is a contact image captured by our surface that we display at an offset from the user's hands

ShadowGuides is not always visible, but can be invoked on demand by touching the surface and dwelling for 1 second. This dwell (or hesitation) can also occur in the middle of the gesture, in which case ShadowGuides will appear and offer guidance that takes into account the hand poses and movement prior to the dwell (i.e., ShadowGuides will try to guide the user to successful completion of the gesture from that point on). Upon lift up of all contacts, the user shadow annotations disappear, but the registration pose guide persists, allowing users to re-pose.

Registration Pose Guide

The registration pose guide is a pop-up panel that shows diagrams of available registration poses, each accompanied by a list of gestures that can be completed from that pose (Figure 3). This guide is the only aspect of our system that deviates from our overall in-situ design, and that is for a good reason: it doesn't make sense to try to *correct* the current hand pose if a different registration pose is desired. Instead, the user should simply lift-off and re-pose to restart the gesture.

Therefore, the purpose of the registration pose guide is to teach the user how to put their hand in contact with the surface in order to start the desired gesture. We accomplish that through a novel visualization scheme that shows the hand from two perspectives.

First, we show the user shadow expected by the system (Figure 3A), which shows the number and the shape of the contacts with the surface. This allows the user to pattern-match between the registration pose guide and the live user shadow feedback. Second, we show which parts of the hand should be used to make that user shadow (Figure 3B) depicted as highlighted regions of the users hand(s).

This dual visualization is important as it not only highlights the number and shape of the contacts, but it also conveys the expert style for starting each gesture by providing a mapping of desired contacts on the surface to portions of the hand. For example, if a two-finger gesture requires a large separating movement, it will probably be difficult to perform it with a single hand; it is important to show that

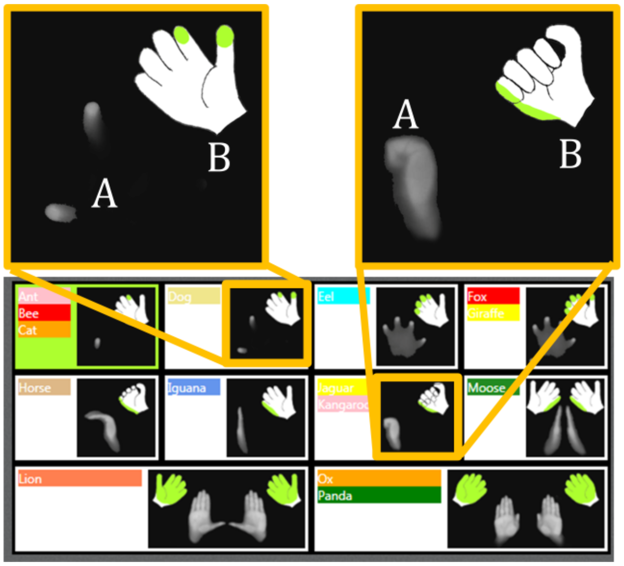

Figure 3. The registration pose guide (bottom) shows available gestures, grouped by registration pose. Insets above show examples of the visualizations used for each pose: the user shadow expected by the system (A) and the parts of the hand used to make that shadow (highlighted, B).

the gesture should be started with contacts from two different hands. Showing user shadows alone does not convey this information, hence our dual-visualization approach.

User Shadow Annotations

Once the user begins a gesture with a specific registration pose, we show user shadow annotations to guide her to completion of the desired gesture. Since multiple gestures might be performable from the same registration pose, we show all continuation possibilities, using color-coding to distinguish which annotations belong with which gesture.

Our annotations extend the idea of dynamic guides, implemented by Bau and Mackay for single-point paths [1], a continually-updating combination of onscreen feedforward and feedback, where the instructions to complete gestures are only shown for those available from the current state. Our implementation of user shadow annotations extends this idea for all gestures in the discussed taxonomy. To do so, we add three new elements: *arrows*, *shape deformation keyframes*, and *dynamic markers*, alongside single-point dynamic guides (Figure 4A). The information contained in these annotations contains only the information required to perform the gestures precisely, and thus emphasizes these aspects.

Arrows

Arrows are added to multi-finger gestures to help users visually parse the direction of the gesture, and to distinguish them from single-finger gestures. These are left out for single-finger or shape gestures for simplicity. (Figure 4B).

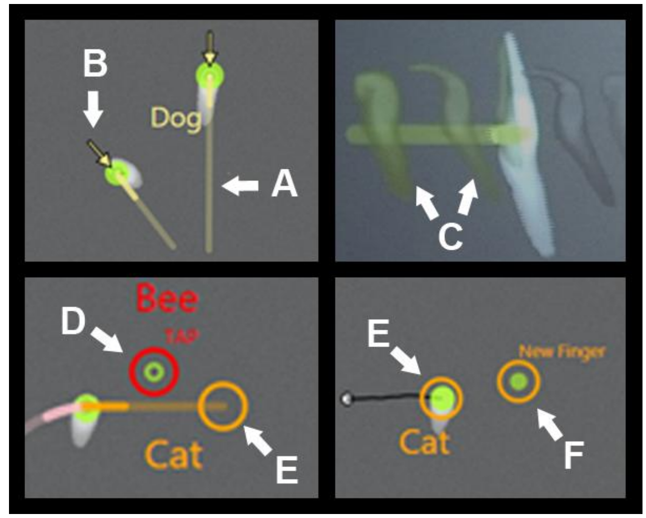

Figure 4. User Shadow Annotations. A: dynamic guides [1]. B: arrows. C: shape deformation key frames. D,E,F: dynamic markers for continuation pose changes.

Shape Deformation

To convey shape deformation (e.g., opening of the user's fist), we show keyframes highlighting the most salient aspects of the gesture. If there is a path in addition to a shape deformation, then a dynamic guide is overlaid on top of the keyframes. (Figure 4C).

Dynamic Markers

If a gesture has a dynamic continuation pose that requires actions other than simple movement across the interface, we use dynamic markers. These markers are annotated with text, and may show that a gesture needs another contact ("new finger", Figure 4F), needs a contact tapped ("tap", Figure 4D), or that a contact needs to stop moving because an action is expected elsewhere (no text, Figure 4E). Dynamic markers only become visible at the relevant stage of the gesture.

Finally, when the user reaches the end position of a gesture, the words "Lift Off" appear. This notification is to let the user know the next step in the gesture is to remove their contacts, causing gesture *termination*. This is especially important for static pose, no-path gestures.

EVALUATION

Our goal was to compare retention, learning speed, and preference of learning with ShadowGuides against the control condition of learning with video instructions (Video condition). To explore these topics, we conducted a between-subjects experiment with 22 participants (8 female) between the ages of 18 and 50, evenly divided between the ShadowGuides and Video conditions. Participants had limited experience with gestural interfaces (nothing more advanced than the iPhone) and had not previously used a Microsoft Surface.

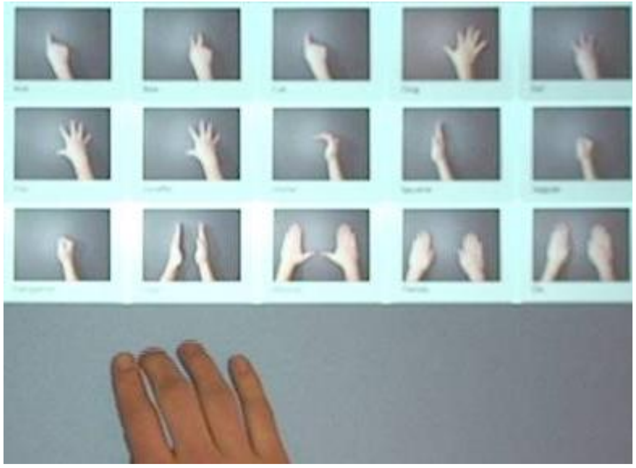

Figure 5. Video menu used to invoke video instructions in our control condition.

The help system in the Video condition consisted of a menu of videos of the available gestures, seen in Figure 5. The video menu consisted of a series of buttons with thumbnails showing each gesture's name and registration pose. When tapped, a life-size video demonstration of the gesture was played on the tabletop. No video was longer than 3.5 seconds. The video menu had the same dwell-to-invoke behaviour as ShadowGuides, i.e., the user had to dwell for one second, before the video menu was made visible. As in ShadowGuides, in the Video condition, users were not required to invoke help and could perform a gesture if they had memorized it.

While contact motion was recognized automatically by our software recognizer, we employed a Wizard-of-Oz recognizer for hand shapes, where the experimenter, looking at the camera image from Microsoft Surface, applied a series of scripted rules to determine successful recognition of the 15 gestures in our set. The same experimenter scored all our trials using a consistent recipe for each gesture.

Participants began with a four gesture tutorial session that showcased the complete vocabulary of ShadowGuides. Video participants used the same four gestures. These tutorial gestures did not overlap with the 15 gestures used in the main experiment. During the tutorial, the experimenter also demonstrated that gestures could be performed without the help system, and that help could also be invoked mid-gesture.

To begin each trial, the user pressed a "Ready" button on the interface, and the name of the gesture to perform appeared in the upper left corner of the display. We assigned each gesture a neutral animal name (e.g., "kangaroo", "cat", and "panda") to keep our evaluation focused on a gesture's form and agnostic to a gesture's meaning. Users were instructed to "…perform each gesture as quickly as you can. You can invoke the help if you'd like, but if you feel you don't need to, then try the gesture without it." Invoking help was always optional.

The experiment consisted of 7 blocks of 15 gestures from our taxonomy with the help system available (the *learning phase*), and a final 8th block when the help was unavailable (the *memory phase*). The ordering of the gestures was random within a 15-trial block, so that in each block each gesture would appear once. Participants had 2 minutes to perform each gesture, with no restriction on number of attempts. If the participant did not perform the gesture within 2 minutes, the experimenter would skip to the next trial. The participants were not instructed to memorize the gestures and did not know there would be a memory test at the end, as we wanted to test how well they would learn simply by using each system.

We recorded video of the participants and logged the raw input from the Surface. We also administered a post-study questionnaire. The total experiment time for each participant was between 60 and 90 minutes.

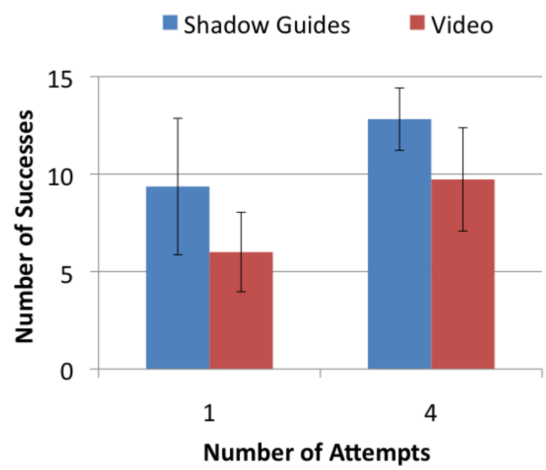

Figure 6. Participants using ShadowGuides were significantly more successful in performing gestures during the memory test.

RESULTS

Learning Phase

We found no significant difference between ShadowGuides and Video during the learning phase, either in number of attempts or speed. This is likely due to the variety of participants' experiences with their respective systems. Intra-group variation greatly exceeded inter-group variation, as learner-learner variation was quite high.

Memory Phase

Participants in the ShadowGuides condition were significantly more successful in recalling gestures on their first attempt than those in the Video condition, correctly performing an average of 9.4 gestures on the first try after learning with ShadowGuides compared to 6.0 with Video ($t_{16.1} = 2.750$, $p = 0.014$) (Figure 6).

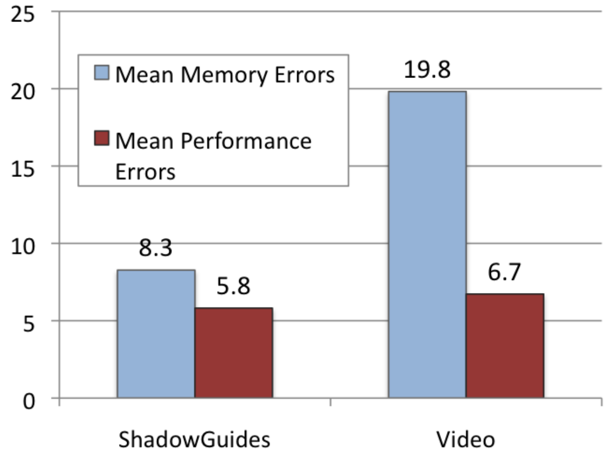

Figure 7. Comparison of mean amounts of error types in the Memory Phase.

We examined the data to see if this advantage disappeared after more attempts were allowed. Gesture recall in both conditions reached a plateau after about four attempts when ShadowGuides participants correctly recalled an average of 12.8 gestures, and Video participants recalled 9.7, still showing a significant advantage of learning with ShadowGuides (t_{20} = 3.312, p = 0.003).

We observed recorded video of the trials, and manually classified the first four attempts of a gesture in the memory phase as either *correct*, *errors in performance*, or *errors in memory*. Errors in performance were where the user had the correct mental model of the gesture, but performed it clumsily, so that it was not recognized in our system (either the software movement recognition or by not following the set Wizard-of-Oz recipe). An error in memory is where the user performed a completely incorrect gesture.

If ShadowGuides only helped in precise performance, we would expect the difference for errors in Video participants to be mostly errors in performance. However, upon analysis, we found a significant difference in the number of errors in memory, where participants in the Video condition had an average of 19.8, and participants in the ShadowGuides condition averaged 8.3. This indicates that ShadowGuides assisted our participants in memorization, relative to Video (t_{20} = 3.021, p = 0.04). There was no significant difference in errors in performance (t_{20} = 0.703, p = 0.490). Results are shown in Figure 7.

Questionnaire

Participants in the ShadowGuides condition gave significantly better Likert-scale ratings of their help system experience than participants in the Video condition (according to Mann-Whitney U tests), as seen in Figure 8.

Discussion

Based on our observations during the experiment, the feedback and feedforward features of ShadowGuides were key to its success in helping users learn gestures more effectively than video. In the video condition, a gesture

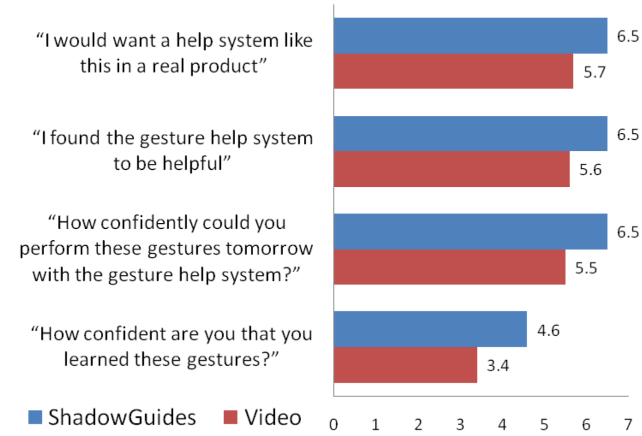

Figure 8. Participants' subjective ratings of the ShadowGuides (SG) and Video (V) learning systems on a 7-point Likert scale, with 7 indicating higher preference. All differences are statistically significant at p<0.05.

would often fail or succeed and the users would express confusion, not knowing what aspect of their performance had caused the failure or success. When we asked video-condition participants why they thought a gesture was finally recognized after several failed attempts, they frequently had no idea. Users of ShadowGuides, on the other hand, were able to learn what aspects of the gesture were most important (due to the emphasized information in the system's feedforward annotations), as well as see how their changing posture affected recognition (due to the feedback of the user shadow).

With video instruction, we observed that users sometimes continued to perform a gesture after they had already made a serious error; however, the continuous feedback in the ShadowGuides mode helped users to identify errors as they occurred, since the user shadow annotations would disappear. ShadowGuides' continuous feedback during gesture performance helped with error correction; however, this continuous feedback was not present when the user tried to find the right registration pose. Participants would often not realize why an initial pose was not recognized. The participant was expected to pattern-match with the registration pose guide, but in practice many had trouble doing so. A future system could improve upon this by highlighting differences between the user's current hand shape and the most similar available registration poses.

While ShadowGuides performed better than Video in our experiment, it is important to acknowledge that video-based instruction has many strengths. For example, video seemed more effective at conveying the general idea, or gist, of the motion, whereas ShadowGuides' strength lay in directing users' attention to specific important features of the gesture. In some ways then, video worked well at conveying general aspects of expert style, but not specifics. Furthermore, some gestural aspects, such as timing or sequencing, are probably better conveyed in video than through ShadowGuides. An potential area for future work is to consider how a gesture-learning system could combine the benefits of both video and ShadowGuides, such as an annotated video, or an interactive notation that suggests movement more directly. Synchronous Objects is a compelling example of modified video of dance choreography [7].

CONCLUSIONS

In this paper, we introduced ShadowGuides, a system for teaching multi-touch and whole hand gestures to the novice user. Additionally, we presented a taxonomy of multi-touch gestures, and representative gestures for each cell of the taxonomy, that can be re-used by other researchers. Finally, we presented an evaluation comparing ShadowGuides to a video-based learning system, and found that users who learned with ShadowGuides performed better on a gesture-memory test and had higher subjective ratings of the system.

ACKNOWLEDGEMENTS

We would like to thank Jarrod Lombardo, Paul Hoover, Sarah Williams, Xiang Cao, Tony Tang, Andy Wilson, Ken Hinckley, Dennis Wixon, August de los Reyes and Clifton Forlines.

REFERENCES

1. Bau, O. and Mackay, W.E. OctoPocus: a dynamic guide for learning gesture-based command sets. *In Proc. of ACM UIST 2008*. p. 37-46.
2. Baudel, T., & Beaudouin-Lafon, M. Charade: remote control of objects using free-hand gestures. *Communications of the ACM*, Volume 36 Issue 7. 1993. p. 28-35.
3. Bragdon, A., et al. GestureBar: improving the approachability of gesture-based interfaces. *In Proc. of ACM CHI 2009*. p. 2269-2278.
4. Brandl, P., et al. Combining and measuring the benefits of bimanual pen and direct-touch interaction on horizontal interfaces. *In Proc. of ACM AVI 2008*. p. 154-161.
5. Buxton, W. A Three-State Model of Graphical Input. *In Proc. of IFIP INTERACT 1990*. p. 449-456.
6. Cao, X., Wilson, A., Balakrishnan, R., Hinckley, K., and Hudson, S. ShapeTouch: Leveraging Contact Shape on Interactive Surfaces. *In Proc. of IEEE TABLETOP 2008*. p. 139-146.
7. Forsythe, William. Synchronous Objects. http://synchronousobjects.osu.edu/
8. Grossman, T., Dragicevic, P., and Balakrishnan, R. Strategies for accelerating on-line learning of hotkeys. *In Proc. of ACM CHI 2007*. p. 1591-1600.
9. Kurtenbach, G. & Buxton, W. Issues in combining marking and direct manipulation techniques. *In Proc. of ACM UIST 1991*. p. 137-144.
10. Guest, A. H., Labanotation, Theatre Art Book, *New York, 2005*. 4th Ed.
11. Moran, T. P., Chiu, P., and van Melle, W. Pen-based interaction techniques for organizing material on an electronic whiteboard. *In Proc. of ACM UIST 1997*, p. 45-54.
12. Morris, M.R., et al. Cooperative Gestures: Multi-User Gestural Interactions for Co-located Groupware. *In Proc. of ACM CHI 2006*. p. 1201-1210.
13. Schneiderman, B. Direct manipulation: a step beyond programming languages. *IEEE Computer Volume 16 Issue 8 (August 1983)*, p. 57-69.
14. Shoemaker, G., Tang, A., and Booth, K. S. Shadow reaching: a new perspective on interaction for large displays. *In Proc. of ACM UIST 2007*. p. 53–56.
15. Tang, A., Neustaedter, C. & Greenberg, S. VideoArms: embodiments in mixed presence groupware. *In Proc. of BCS HCI 2006*. p 85-102.
16. Tapia, M.A. and Kurtenbach, G. Some design refinements and principles on the appearance and behavior of marking menus. *In Proc. of UIST 1995*. p. 189-195.
17. Tversky, B., Bauer Morrison, J. Animation: can it facilitate? *Int. Journal of Human-Computer Studies Volume 57, 2002*. p. 247–262.
18. Vanacken, D. et al. Ghosts in the Interface: Meta-user Interface Visualizations as Guides for Multi-touch Interaction. *In Proc. of IEEE TABLETOP 2008*. p. 81-84.
19. Vogel, D. and Balakrishnan, R. Interactive public ambient displays: transitioning from implicit to explicit, public to personal, interaction with multiple users. *In Proc. of ACM UIST 2004*, 137-146.
20. Vogel, D., Cudmore, M., Casiez, G., Balakrishnan, R., and Keliher, L. Hand occlusion with tablet-sized direct pen input. *In Proc. of ACM CHI 2009*. p. 557-566.
21. Wobbrock, J. O., Morris, M. R., and Wilson, A. D. User-defined gestures for surface computing. *In Proc. of ACM CHI 2009*. p 1083-1092.
22. Wu, M. and Balakrishnan, R. Multi-finger and whole hand gestural interaction techniques for multi-user tabletop displays. *In Proc. of ACM UIST 2003*. p. 193-202.
23. Wu, M., Shen, C., Ryall, K., Forlines, C., and Balakrishnan, R. Gesture Registration, Relaxation, and Reuse for Multi-Point Direct-Touch Surfaces. *In Proc. of IEEE TABLETOP 2006*. p. 183-190.
24. Zhao, S., Agrawala, M., and Hinckley, K. Zone and polygon menus: using relative position to increase the breadth of multi-stroke marking menus. *In Proc. of ACM CHI 2006*. p. 1077-1086.