

WIPE IT! A DIRECT MANIPULATION TECHNIQUE FOR UBIQUITOUS INFORMATION ITEMS

M. Schneider¹ and A. Butz²

¹) German Research Center for Artificial Intelligence (DFKI) ²) University of Munich, Germany

ABSTRACT

We present an interaction technique called “wiping”, which allows to move information units beyond screen boundaries onto different devices or displays in the environment. The technique uses a physical metaphor and makes the information units behave similar to physical objects. The technique can be used with mice or pens on the corresponding devices, or with bare hands under a camera. We give a technical definition of the wiping gesture, describe a prototypical implementation and demonstration setup, and discuss some properties of the technique.

INTRODUCTION

In his ubiquitous computing vision Weiser (6) predicts a shift away from the single personal computer towards a multitude of smart objects and devices embedded in our every day lives. Within this idea of “everywhere computing” not only applications but also virtual information items will become ubiquitous and need to be generated, handled, processed and presented efficiently in our daily environments.

Significant changes in the way we deal with information items could already be observed over the last years: Instead of manually exchanging and entering phone numbers, virtual contact cards are beamed between PDAs and smart phones. Bar codes and radio frequency identifications (RFID) tags allow to easily link printed information to related virtual information items. Music and films no longer need physical containers, such as tapes or disks. Instead, they are accessed and streamed over the network. Video projectors are used to give presentations and show slide shows instead of using hardcopy presentation slides and diasscopes.

With the growing importance of virtual information items, which can already be seen today, the need for efficient and intuitive manipulation techniques arises.

RELATED WORK

Several interaction techniques have been proposed to deal with the problem of moving information units over larger distances. For the virtual reality community this problem already came up a decade earlier, and Poupyrev (4) and Bowman (1) present an interesting

overview of techniques for virtual reality. In ubiquitous computing scenarios, two basic classes of approaches exist. In Geißler (2) two interaction techniques for a single large display are presented. On a large display, which is potentially used by multiple users simultaneously, not all areas of the display are within reach for all users. The author proposes to use a physical analogy and “shuffle” or “throw” objects across the screen. This effectively extends the user's reach on the screen. The third technique described in Geißler (2) allows users to “take” an information object from one screen position and “put” it in another position. This is in analogy to the “pick-and-drop” technique described for pen input devices in Rekimoto (5). Technically, both “take and put” and “pick and drop” can be used to move information units from one display to a second one, as long as the identity of the hand or pen can be determined. However, this will only work as long as the second display can be physically reached.

If the second display is physically out of reach, as for example a large wall projection at a distance would be, there is no way to “put” or “drop” an object on it. If there are spectators seated between the screen and the presenter, it might be cumbersome or infeasible to physically walk up to the large display in order to move information units there. For these situations, an interaction technique similar to the distant techniques described in Poupyrev (4) is needed to allow users to send objects from one screen to another in an intuitive way.

One way to move physical objects from one place to another is the act of wiping them. An example are crumbs strewn on a table, which can be wiped onto the other hand or onto the ground to clean the table. Another example, in which a very direct analogy is used, is a pen stroke gesture built into Palm OS, in which the pen is moved rapidly from the graffiti area on the bottom of the screen to the top, i.e. away from the user. This gesture can be used for triggering the “beam item” action, and in this case provides a perfect metaphor for wiping the unit of information currently seen on the PDA screen (an address, a note or an appointment) away from the user, through the infrared port on the far end of the device, onto another device in that direction.

“WIPING” INTERACTION TECHNIQUE

A wiping gesture can be defined as an accelerated straight-line motion of a hand, a pen, or more general some arbitrary interaction artifact. A wiping gesture addressing a (virtual) object always starts next to the object and leads across its area towards the opposite side (figure 1). A single wiping gesture can span multiple close-by objects as long as they are aligned in a line along the wiping path. The parameters of a wiping gesture are its timing diagram and its spatial path. A detailed discussion of both is given in the following sections.

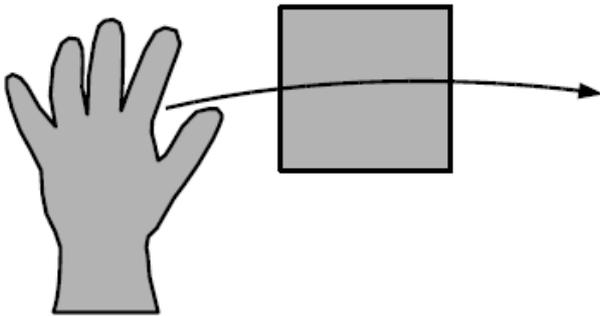


Figure 1: An object is wiped by a continuous motion across its area

The main idea is that wiped objects are accelerated according to the motion vector given by the wiping gesture. As a result, the affected objects start moving with a certain speed into a certain direction. Both parameters can be directly influenced by the user through adjusting speed and direction of the wiping gesture. The movement of the wiped objects proceeds even after the initial wiping impulse has ended, thus enabling the user to move objects to places out of her reach. However, once wiped the pace of the movement decreases over time caused by a virtual slip resistance. Thereby the distance of the movement can be naturally controlled through the intensity respectively the maximum speed of the wiping gesture.

For practical applications, wiped objects should then be “caught” by some target device or display. If there are multiple candidates in different directions, the target device closest to the object's motion direction will catch the object. If there are several candidates aligned in the same direction, the object's traveling speed can be used to disambiguate the gesture, and the target device will be chosen, on which the object would come to a rest if it was physical.

Wiping gestures always take place in a specific plane. For a screen or a projection display, this plane is roughly perpendicular to the ground, for a desk surface it is horizontal, spanning the whole room. This implies for example, that from a desk, all displays on the walls around it can be intuitively reached by wiping, but from

an upright screen, displays behind the user cannot be reached by wiping.

Timing Diagrams

The timing diagram describes the velocity and acceleration characteristics of a wiping gesture. It depicts the velocity of the interaction artifact over time. Periods without motion are used to separate consecutive gestures and to identify potential wiping gestures.

A prototypic timing diagram is shown in figure 2. In the timing diagram three different phases can be identified for every valid wiping gesture:

1. *The acceleration phase p_{acc} .* At the beginning of every wiping gesture the interaction artifact is accelerated to near its maximum speed v_{max} . Objects to wipe may already be “collected” in this phase.
2. *The constant wiping phase p_{const} .* This phase determines the final speed v_{max} of the wiped items. Further items additionally to those already accumulated during p_{acc} may be collected during this phase.
3. *The deceleration phase p_{dec} .* The wiping gesture ends with a deceleration of the wiping motion. This deceleration can be abrupt or smooth. In either case this signals the end of the intended wiping gesture and is rather an appendix to the primary wiping gesture than an intentionally executed motion. Therefore objects crossed in this phase are not affected by the wiping gesture anymore. This conforms to the intuition that the main subjects of the gesture were already accelerated and “wiped away”.

If these three phases can not be discovered in a timing diagram, the corresponding gesture is not considered a valid wiping gesture. However some tolerance should be accepted in the recognition process.

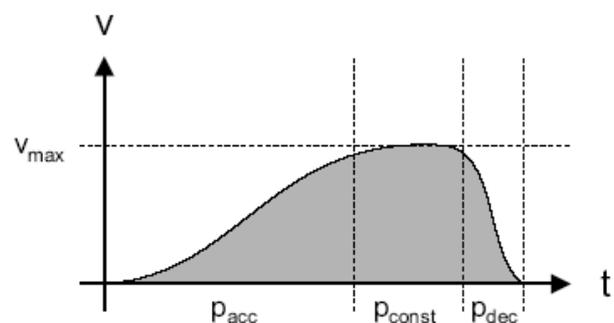


Figure 2: A timing diagram of a prototypic wiping gesture

Wiping Path

The wiping path describes the spatial characteristic of a wiping gesture, especially its location and direction. It is the trail left by the wiping gesture on the interaction plane. We have already seen such a wiping path in figure 1, where the arrow originating from the user's hand is a path for a wiping gesture to the right. The wiping path is relevant for three reasons:

- *To detect whether the wiping gesture is valid.* A wiping gesture is valid if the interaction artifact was moved in a straight line. Because in general it is impossible to move something along a perfectly straight line by hand, some tolerance is permissible here. How much tolerance is acceptable is application dependant. A gesture with an invalid wiping path is not considered a valid wiping gesture. Figure 3.i shows an invalid wiping path.
- *To detect which objects are affected by a wiping gesture.* By definition these are all objects that are *crossed* by the wiping path in phase p_{acc} and p_{const} . Note that an object over which a wiping path originates or terminates is not affected by the according wiping gesture. Figure 3.ii gives an example of a valid wiping gesture affecting objects *B* and *C*. Object *D* is not affected by this gesture because it is crossed in p_{dec} as illustrated by the dashed line.
- *To detect the direction of a wiping gesture.* Because the wiping gesture may not be a perfectly straight line, the wipe direction is averaged over the whole wipe path. The exact approach to calculate the direction of the wiping gesture is implementation specific.

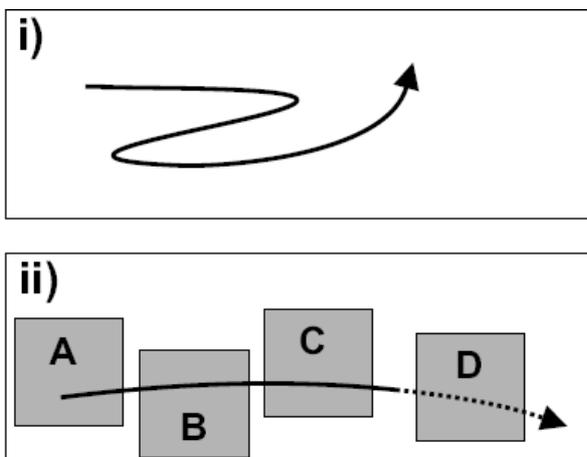


Figure 3: An invalid wiping path (i) and a valid wiping path spanning objects B and C (ii)

Visual Feedback

Visual feedback can be given depending on the application's needs, the interaction status, and the instrumentation of the environment.

The intuition is, that a successfully recognized wiping gesture causes the selected objects to “fly” towards their destination position. Depending on the instrumentation of the environment this may be animated in various ways: If a steerable projector like the everywhere display described in Pinhanez (3) is available, the whole movement to the new display position can be animated if an appropriate projection plane is available along the motion path. This for instance is the case when wiped information items are moved between displays mounted to the same wall. If no such projector or projection plane exists, the animation is limited to show just the disappearance from the source display and the arrival at the destination display.

In case of an unsuccessful wiping interaction two cases have to be distinguished. If the gesture itself was correct, but no matching target display could be identified, the selected objects return to their original position after a short “tremor”. In a case of a gesture that could not be verified as a valid wiping gesture, it depends on the application if and how feedback is given. If the application is such, that other gestures besides wiping might be (consciously or unconsciously) executed by the user, the application may choose to ignore invalid wiping gestures. Otherwise, acoustical or optical feedback (i.e. in form of a question mark displayed in the area where the gesture was observed) could be given.

PROTOTYPICAL IMPLEMENTATION

To test our ideas we have implemented a prototype that uses computer vision techniques to recognize wiping gestures. The setup consists of a steerable projector with an attached video camera, similar to the everywhere display described in Pinhanez (3). Through the projector virtual information items can be displayed onto any adequate surface. Gestures recognized through the camera are used to move these items to other devices or (projected) displays.

We have tested our setup in a home entertainment scenario: Media files, such as images or movies could be wiped from a PC onto the table. The items collected on the table could then either be wiped back onto some PC or onto a projection on the wall, resulting in a slide show of the media files to be displayed. The overall setup as presented on the CeBIT fair 2004 is shown in figure 4.



Figure 4: Experimental setup with a projector and camera above a table. Media files can be wiped from the table back on to the PC or on to the wall in the background

Computer Vision Wipe Recognition

To recognize wiping gestures we use differential picture analysis. The central idea is to calculate the center of gravity of changes between each pair of consecutive video frames and trace the movement of this point over multiple frames. If a sufficiently large interaction artifact is continuously moved, this trace should approximately match the motion path of the artifact. In our implementation we use gray-scaled video snapshots and create a thresholded b/w differential image from each two consecutive frames. After a distance transform on the resulting image the weighted center of gravity is calculated.

A frame dump of a sample run of our recognition system is shown in figure 5. On the left side a sequence of three video frames is given showing an accelerated wiping gesture. On the right side the corresponding differential images are shown. The black dot represents the center of gravity for the actual differential image, the black line shows it's motion path over the last frames.

Wiping Path & Timing Diagram

The wiping path and timing diagram can be directly calculated from the sequence of differential images. The

trace of the center of gravity shown as black line in the differential images presented in figure 5 already corresponds to an approximated wiping path. From this path the timing diagram can easily be derived by considering the distance between two consecutive wiping path vertexes relative to the frame rate.

GENERALIZATION AND DISCUSSION

The wiping interaction technique we have presented has some favorable characteristics, which make it applicable to different application scenarios using a variety of interaction artifacts:

- *Applicable to real and virtual environments.* Wiping can be executed with computer mice, touch screens, hands, real and virtual artifacts, etc. Therefore the wiping interaction technique can be consistently applied to real, virtual, and mixed world applications.
- *Interaction without clicking.* The concept of clicking as known from the computer mouse has no natural equivalent when interacting with the real world. Because wiping does not require clicking at all it is a good candidate for a ubiquitous interaction technique.
- *Implicit addressing of items to manipulate.* By executing the wiping gesture over the information item(s) to manipulate, both target and interaction type are specified at the same time. The ability to specify more than one target with a single gesture makes it a real direct manipulation technique even for larger collections of items.

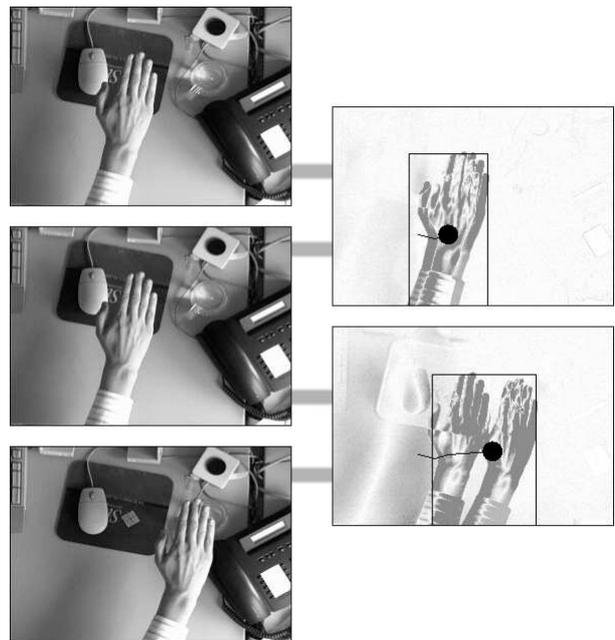


Figure 5: Captured images (left) and differential images (right) of an accelerated wiping gesture to the right

ACKNOWLEDGMENTS

The work presented here was funded by “Deutsche Forschungsgemeinschaft” under contract number BU 1402/1-1. Parts of the software were implemented by Mira Spassova, Christoph Endres and others of the Fluidum team.

REFERENCES

1. Bowman D., and Hodges L., 1997, An evaluation of techniques for grabbing and manipulating remote objects in immersive virtual environments. In Proceedings of 1997 Symposium on Interactive 3D Graphics, 35-38
2. Geißler J., 1998, Shuffle, throw or take it! working efficiently with an interactive wall. In Proceedings of CHI 98, 265-266
3. Pinhanez C., 2001, The everywhere displays projector: A device to create ubiquitous graphical interfaces. In Proceedings of Ubiquitous Computing 2001 (UbiComp'01)
4. Poupyrev I., Weghorst S., Billingham M., and Ichikawa T., 1997, A study of techniques for selecting and positioning objects in immersive VEs: effects of distance, size, and visual feedback. Technical Report R-97-45
5. Rekimoto J., 1997, Pick-and-drop: A direct manipulation technique for multiple computer environments. In Proceedings of ACM UIST '97, 31-39
6. Weiser M., 1991, The computer for the 21st century. Scientific American, 3(265):94-104