ORIENTATION ON TABLETOP DISPLAYS

by

G. Daryn Mitchell

B.Sc., Queen’s University, 2000

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE
in the School
of
Computing Science

© G. Daryn Mitchell 2003
SIMON FRASER UNIVERSITY
October 2003

All rights reserved. This work may not be
reproduced in whole or in part, by photocopy
or other means, without the permission of the author.
APPROVAL

Name: G. Daryn Mitchell
Degree: Master of Science
Title of thesis: Orientation on Tabletop Displays

Examining Committee: Dr. Torsten Möller
Chair
Dr. Arthur (Ted) Kirkpatrick, Senior Supervisor
Dr. M. Steen Atkins, Supervisor
Dr. Kori Inkpen, External Examiner,
Associate Professor of Computing Science,
Dalhousie University, Nova Scotia

Date Approved: Oct. 30, 2003
PARTIAL COPYRIGHT LICENCE

I hereby grant to Simon Fraser University the right to lend my thesis, project or extended essay (the title of which is shown below) to users of the Simon Fraser University Library, and to make partial or single copies only for such users or in response to a request from the library of any other university, or other educational institution, on its own behalf or for one of its users. I further agree that permission for multiple copying of this work for scholarly purposes may be granted by me or the Dean of Graduate Studies. It is understood that copying or publication of this work for financial gain shall not be allowed without my written permission.

Title of Thesis/Project/Extended Essay:

Orientation on Tabletop Displays

Author: 

(Signature)

G. Daryn Mitchell

31 Oct. 03

(Date Signed)
Abstract

Tabletop computer displays suffer from an orientation problem. Solutions based on competing approaches have been implemented, but there remain unanswered questions about the ideal solution.

Within a context of interacting with documents on tabletop displays, requirements for orientation control were found by examining literature on how people use paper documents and what manipulations they perform. It was determined that control must provide quick, either-handed, low-attention manipulation of individual objects on the display. An evaluation of existing interaction techniques for rotation was performed in light of these criteria. The trade-off between two desirable characteristics, integral translation and rotation, and direct input, was examined. An evaluation of mouse-based techniques confirmed that integral input has the potential to be faster than the established sequential-manipulation techniques due to the time saved by overlapping manipulation of different degrees of freedom. For a single task combining translation and rotation the separable technique took twice as long compared to two separate tasks of the same translation or rotation, whereas the integral techniques took less time for the combined action than the two separate actions. However, the integral mouse-based techniques were too slow in their actual manipulation, with the scroll wheel taking four times as long to rotate than the separable technique, and the new drag technique not being used in an integral manner. It was concluded that the current generally available technology such as the mouse and single-point touch-sensitive overlays are inadequate. Acceptable tabletop orientation control will be dependent on the maturation of newer technologies based on tangible interfaces or possibly multi-finger input. Until then, interaction techniques for manipulating documents on tabletop displays should be chosen according to their ability to concurrently control position and orientation, such as three degree-of-freedom input devices on digitizing tablets.
Acknowledgments

I’ve never written an acknowledgement before. I’ve always thought it seemed funny to write it at the front of the text, when it really belongs at the end... you know, it’s the very last thing after making it all the way through.

Making it all the way through is exactly what this acknowledgement is about.

The most important person, without question, is Brenda, my wife and best friend. I may be the one who gets a few letters to add after my name (at least, when I’m trying to impress someone), but she suffered through it with me every step of the way. Helping provide sticktoitiveness were my parents David & Daphne, without whom I would never have kept aiming for the goal.

Many friends at school helped make the walk enjoyable. Too many to name everyone, especially the many friends from the computer science grad lunchroom / ping pong room. but allow me to mention a few by name: Keith Shu, Felix Lau, Ken Chidlow, Steve Kilstau, Christina Lee, and Jason Sze.

Thanks to Ted Kirkpatrick for all the help with my work. Thanks also to my previous senior supervisors, Stella Atkins and Kori Inkpen.

I am also indebted to the support and encouragement from my church, you know who you are, including those who prayed me through without me even knowing it.

...which leads naturally into my real acknowledgement. Those who know me know that my dependence is on God. All though my school career since becoming a Christian, I have put SDG at the start of my tests and essays, copying the Soli Deo Gloria that I’ve heard Bach put on his work. That’s what needs to go here: a little more explicit and less Hollywood-cool, maybe, but the same notion: all glory to Jesus the Messiah.
Not that I think there’s likely to be “glory” from this thesis. If this thesis is any good, great! Glory to God. Even if it’s not very good, though, I still want to give glory to God, because I wouldn’t have even close to made it through without his encouragement and strength, and my motivation has been to work as though I’m working for him.

Soli Deo Gloria.
Contents

Approval
Abstract
Acknowledgments
Contents
List of Tables
List of Figures

1 Tabletop Displays and the Problem of Orientation
  1.1 Introduction to Tabletop Displays
  1.1.1 DigitalDesk
  1.1.2 ClearBoard
  1.2 Orientation on Tabletop Displays
  1.2.1 Characteristics of Orientation Solutions

2 Requirements for Orientation Control
  2.1 Context of Use: Paper Documents
  2.1.1 The Significance of Paper Documents
  2.1.2 Why Document Orientation is a Problem
  2.2 Requirements for Orientation Control
  2.2.1 How People Read
  2.2.2 How People Write
  2.2.3 Document Actions
2.3 Implications for Orientation Control ........................................ 15

3 Techniques for Manipulation on Tabletop Displays 18

3.1 Techniques for Solving the Orientation Problem ...................... 18
  3.1.1 Non-oriented Interfaces ........................................ 18
  3.1.2 Automatic Orientation ......................................... 19
  3.1.3 Orientation Control .......................................... 20

3.2 Evaluation of Techniques Meeting Basic Requirements ............. 25

3.3 Distinguishing Characteristics of Manipulation ITs ............... 26
  3.3.1 Integral Rotation and Translation ................................ 26
  3.3.2 Direct Input .................................................. 28
  3.3.3 Exploration Space of Tabletop Manipulation ITs ............... 30

3.4 Practical Roadblocks ................................................. 32
  3.4.1 Integrality and Directness of Input ............................ 33
  3.4.2 Pilot Study .................................................... 34
  3.4.3 Touchscreen Unusable ......................................... 37
  3.4.4 Proposed Integral Techniques for Existing Input Devices ....... 38
  3.4.5 Wheel ......................................................... 38
  3.4.6 Drag ........................................................ 39
  3.4.7 Proposed Experiment .......................................... 41

4 Experimental Design ..................................................... 43

4.1 Goal .............................................................. 43

4.2 How to Evaluate the Techniques .................................... 43
  4.2.1 Common Manipulations ....................................... 44
  4.2.2 Measuring Effectiveness ..................................... 45

4.3 Experimental Method ................................................ 48
  4.3.1 Design ......................................................... 48
  4.3.2 Apparatus ..................................................... 49
  4.3.3 Task ........................................................ 51
  4.3.4 Interaction Techniques ...................................... 53
  4.3.5 Task Configurations ......................................... 54
  4.3.6 Subjects ...................................................... 56
  4.3.7 Procedure .................................................... 56
4.3.8 Dependent Variables ........................................... 57
4.3.9 Hypotheses ...................................................... 59

5 Experimental Results ........................................... 63
  5.1 Results and Discussion ........................................ 63
  5.1.1 Preparation for Analysis .................................... 63
  5.1.2 Main Results .................................................. 65
  5.2 Physical Manipulation of Paper ................................ 88
  5.3 Final Comment .................................................. 91
  5.4 Conclusions ..................................................... 91

6 Conclusions ....................................................... 93

A Appendix .......................................................... 97
  A.1 Approval of the SFU Research Ethics Board .................... 97

Bibliography ........................................................ 99
List of Tables

1.1 Differing approaches to Kruger and Carpendale's questions regarding orientation solutions on tabletop displays ........................................ 5

2.1 Workplace Reading Activities ............................................. 11

3.1 Integrality of tabletop manipulation ITs .................................. 28
3.2 Directness of tabletop manipulation ITs .................................. 32

4.1 Categories in the TLX Questionnaire .................................... 47
4.2 Task configurations used in the experiment ........................... 55
4.3 User # designation of IT Ordering ........................................ 57

5.1 Table of results (Min, Max, Mean, and Median dockTimes) ........ 66
5.2 Calculation of summed \((T + R)\) time with 95\% CI ....................... 71
5.3 Comments about wheel reveal support for click-once pickup over mouse dragging. ................................................................. 87
List of Figures

2.1 Examples of rotated documents from the literature on paper-based work habits 13

3.1 Integrality and Directness of ITs for Manipulation on Tabletops ............ 31
3.2 Integrality and Directness examined in Related Multi-DOF Research .......... 35
3.3 Illustration of the problems with the touchscreen tabletop display ............ 36
3.4 Illustration of Drag .............................................................................. 41
3.5 Proposed experiment in relation to the IT exploration space .................... 42

4.1 Physical setup of experiment ................................................................. 50
4.2 Screenshots from the experiment ........................................................... 52
4.3 Altered damperFunction (including central translate-only radius) .......... 55
4.4 Screenshot of TLX questionnaire ............................................................ 60

5.1 Mean Completion Times by Block for each IT ........................................ 64
5.2 Completion times for each task configuration (Boxplot) ......................... 65
5.3 Measures of integrality for each IT ......................................................... 69
5.4 Burstiness of Scroll Wheel During a TR90 Trial ..................................... 69
5.5 Comparison of Time Saved Due to Integral DOF Overlap ........................ 71
5.6 dockTime vs. Integrality for TR task conditions ..................................... 73
5.7 dockTime vs. M-metric for TR task conditions ....................................... 74
5.8 Durations (real and subjective) and RSD for each IT .............................. 75
5.9 Individual time bias is the cause of high RSD variance .......................... 75
5.10 Rankings of ITs within each user’s RSD ............................................... 76
5.11 Normalized RSD for each IT ................................................................. 76
5.12 TLX Workload Scores for each IT ......................................................... 77
5.13 TLX scores for individual categories .................................................... 78
5.14 Weightings for TLX categories ........................................... 79
5.15 Order effect on Average TLX Workload ................................. 80
5.16 Average TLX Workload vs. RSD ........................................... 81
5.17 Locations of the first click for the Drag IT in TR30 ............... 82
5.18 Corner-like approach to using Drag (During a TR90 Trial) .... 83
5.19 Integral approach to using Drag (During a TR90 Trial) .......... 84
5.20 Integrality for each user (Drag IT, TR30) ............................ 84
5.21 Acquisition Time for each IT (all 5 blocks shown) ................. 86
5.22 Docking Time vs. # Mouse Clicks ....................................... 87
5.23 Completion Time for Physical Manipulation ............................ 88
5.24 Completion Times for Paper-based Recreation of the Study .... 89
Chapter 1

Tabletop Displays and the Problem of Orientation

1.1 Introduction to Tabletop Displays

A tabletop display is a computer display that makes up the surface of a table or a desk. Whether it is a horizontal surface or a slanted drafting table, the idea is simply that the surface on which we normally work — spreading out, organizing, and sharing information on paper documents — can now have digital information too.

Most tabletop display research has been based on the same motivations that were behind the two earliest examples of tabletop displays, DigitalDesk and ClearBoard.

1.1.1 DigitalDesk

The first approach to tabletop displays is based on the vision expressed by Pierre Wellner in his seminal work on the Digital Desk (1993), a desire to bring together the best of paper and digital documents. The way we interact with paper documents is considerably different than the way we interact with digital documents. Computers offer powerful tools and tremendous flexibility, but in order to take advantage of them we have to sacrifice the tremendous affordances of paper. As he expressed it,

"We use our fingers, arms, 3D vision, ears and kinesthetic memory to manipulate multiple objects simultaneously, and we hardly think about how we do this because the skills are embedded so deeply into our minds and bodies."
CHAPTER 1. TABLETOP DISPLAYS AND THE PROBLEM OF ORIENTATION

Tabletop display research following the DigitalDesk theme has attempted enhance computers with the skilled paper-document interaction people are used to on real desks or in some cases, to express it slightly differently, to enhance paper with computer functionality (Arai et al., 1995; Robinson et al., 1997; Robertson and Robinson, 1999; Aliakseyeu et al., 2002a, b).

1.1.2 ClearBoard

The second approach to research on tabletop displays has been motivated by the potential they have to support collaboration. People rarely work together on a desktop computer: typically when people meet they gather around a table, which allows them to see each other and to share objects as they work together.

Ishii, Kobayashi, and Grudin (1992) designed ClearBoard with this desire in mind, that people ought to be able to see each other while working on the computer. To explore this notion, they analyzed the gaze patterns of a pair of people sitting across a table from one another, as one taught the other how to play backgammon. The results showed that the student looked at the game board almost the entire duration of the session. This behaviour confirms the work of Argyle and Graham (1976), who studied the gaze patterns of pairs of strangers sitting across from one another at a table with a very low-detail map of Europe on the table. When participants were instructed to "get to know each other", the cumulative duration of mutual gaze was 65% of the session time, while less than 15% was spent looking at the map. When participants were instructed to make travel plans to a particular location, mutual gaze was reduced to 30% and map-gaze increased to 70% of the session time. When the very-low detail map was replaced by a detailed map of the particular destination, total mutual gaze was only 5% of the session, and object gaze was 92%. However, they noted that it was mostly just the duration of glances that changed; partners still made eye contact almost as frequently, even though those glances were much shorter.

However, the results from Ishii et al. (1992) revealed an interesting difference in gaze behaviour: the backgammon teacher had very different gaze patterns than the student, constantly looking back and forth between the game board and his pupil. The data presented in the paper showed that on a number of occasions his gaze shifted as quickly as 8–10 times in 11 seconds. The instructor received constant feedback about comprehension from both the game board and the student’s facial expression. In a different physical configuration, such as the side-by-side configuration required at desktop computers, it would have been
difficult for the instructor to maintain such awareness of the pupil’s reactions.¹

Tabletop displays for collaboration can also bring together the best of both worlds, the power of computers and the interpersonal interaction of face to face work with shared information. A significant amount of research has focused on collaborative tabletop displays, such as tabletop displays designed intentionally for collaboration (Agrawala et al., 1997; Rauterberg et al., 1997; Tandler et al., 2001), some even providing ways to deal with the problem of orientation that occurs with multiple users (Streitz et al., 1999; Vernier et al., 2002; Hancock et al., 2001; Inkpen et al., 2002; Mandryk et al., 2002).

1.2 Orientation on Tabletop Displays

Tabletop displays create a problem not found on desktop computers. Computers have generally been used in a physical arrangement with the screen in the vertical plane. This is true for also laptops and projection screens as well as for desktop computers. On vertical screens, the perceived upright orientation is always the same, independent of viewer location. However, for a horizontal display the user’s perspective is tied to his or her location: looking at the screen from a different location gives the information a different perceived orientation. Consequently, simply placing a screen horizontally and using desktop interaction techniques does not guarantee effective interaction.

The DigitalDesk vision was of natural and easy manipulation of electronic documents on the tabletop display. On desktop computers, mouse manipulation might not exploit all the richness of physical interaction, but at least it is effective at controlling the position of documents on the screen. On tabletop displays, objects may vary not only in position but also in orientation, whereas mice only detect changes in position. The effectiveness of a 2D input device like the mouse may be drastically reduced on tabletop displays.

The problem of orientation is even more critical for collaboration. For a single user, if all the information is at one common orientation, there is still a location from which all the information will appear upright. For multiple users, the information would not appear upright for at least some of the users, unless they positioned themselves side-by-side (possibly quite closely together). Yet in collaborative activities users should arrange themselves in

¹The Clearboard design resulting from this work was actually for remote collaboration rather than co-located. Thus only one user used each display, with a video image of their partner superimposed on their digitally common work surface. However, since the displays had a 30° slope, they were still best categorized as tabletop displays.
relation to each other, not have their location determined by the display, and in particular they should not be required to be too physically close to their partners (Scott et al., 2003). If information could only be displayed at one orientation, collaboration on tabletop displays could be quite awkward and in many instances might not be feasible at all.

For the gaze analysis of collaboration reported in Section 1.1.2, the only reason partners were able to sit across from one another was that, in both cases, the information on the table was not orientation-dependent (a map and a backgammon game). In general, however, limiting tabletop displays to non-oriented information would prevent them from fulfilling their original purpose: supporting the way people already work together at tables, including common oriented-information such as documents.

1.2.1 Characteristics of Orientation Solutions

Since orientation is a problem on tabletop displays, an orientation solution must be provided. Yet, other than recognizing that there is a need for some sort of rotation, no clear requirements have been established for such a solution.

In the most detailed examination thus far of orientation at tabletop displays, Kruger and Carpendale (2002) identified five unresolved questions about rotation of information on tabletop displays:

Q1) “Would rotation of individual objects on the display be more preferable than rotation of entire windows or even the entire screen?”

Q2) “What sort of mechanism could be used to allow for quick and intuitive rotation of objects?”

Q3) “Would having multiple copies of the same information in different orientations be a feasible solution?”

Q4) “If change of orientation was supported, would this result in decreases in productivity or efficiency because of the constant requirement to re-orient objects?”

Q5) “Is supporting freckhand rotation as opposed to orthogonal rotation desirable?”

Their claim that these questions have not been clearly resolved is evident in the variety of orientation solutions that have been implemented on tabletop systems, a sampling of which is shown in Table 1.1.
CHAPTER 1. TABLETOP DISPLAYS AND THE PROBLEM OF ORIENTATION

<table>
<thead>
<tr>
<th>Tabletop System</th>
<th>Q1 Rotate What?</th>
<th>Q5 What Angle?</th>
<th>Q2 What mechanism?</th>
</tr>
</thead>
<tbody>
<tr>
<td>MetaDesk [Ull97]</td>
<td>Whole Surface</td>
<td>Freehand</td>
<td>Tangible (TUI handle)</td>
</tr>
<tr>
<td>BUILD-it [Rau97]</td>
<td>Whole surface</td>
<td>Freehand</td>
<td>Graspable (GrUI handle)</td>
</tr>
<tr>
<td>i-Land [Str99]</td>
<td>Objects</td>
<td>Freehand</td>
<td>Gesture to enter mode</td>
</tr>
<tr>
<td>Augmented Surfaces [Rek99]</td>
<td>Objects</td>
<td>Orthogonal</td>
<td>Automatic (obj. location)</td>
</tr>
<tr>
<td>MERL (method A) [Ver02]</td>
<td>Whole surface</td>
<td>Freehand</td>
<td>Move control point</td>
</tr>
<tr>
<td>MERL (method B) [Ver02]</td>
<td>Objects</td>
<td>Radial</td>
<td>Automatic (obj. location)</td>
</tr>
<tr>
<td>Edgelab [Han01]</td>
<td>Objects</td>
<td>Orthogonal</td>
<td>Automatic (user location)</td>
</tr>
<tr>
<td>VIP [Ali02]</td>
<td>Objects</td>
<td>Freehand</td>
<td>Graspable (AR paper)</td>
</tr>
</tbody>
</table>

Table 1.1: Differing approaches to Kruger and Carpendale's questions regarding orientation solutions on tabletop displays.

This thesis does not address Q3 because the complexities of the interpersonal interaction are beyond the scope of this work. Interested readers might refer to Olson, Olson, Storrosten and Carter (1992), who had groups of people collaboratively author a document; the participants each had their own view of the document on workstations that were "recessed in the table top so as not to interfere with normal eye contact", a similar physical configuration to tabletop displays.

The remaining four questions form the foundation of this thesis. In Chapter 2, after establishing a context for text documents as the key "oriented" object, I review the literature about how people work with documents to investigate the role of paper-based orientation control. The non-linear manner in which people read and write depends heavily on document manipulation. An analysis of those manipulation actions provides a set of requirements for document manipulation techniques on tabletop displays.

Existing approaches to orientation control are considered in Chapter 3, and evaluated in light of the requirements. Two characteristics of paper document manipulation are introduced to help distinguish the attributes of orientation control techniques. Many of the most promising techniques depend on technology that is generally unavailable, so an experimental evaluation was performed to compare techniques that use available technology, the results of which are presented in Chapters 4 and 5. In Chapter 6 I consider the implications for tabletop displays.
Chapter 2

Requirements for Orientation Control

2.1 Context of Use: Paper Documents

Requirements for orientation control must be developed within a specific context. The task being performed on the tabletop can result in very different orientation requirements, for instance, non-oriented content such as a map or a board game may not have any orientation problem at all.

The context of the five questions in Section 1.2.1 was a user study performed by Kruger and Carpendale (2002) in which pairs of people put together a puzzle of text or abstract shapes. Participants oriented objects towards themselves to create personal space, and rotated them to a different orientation to designate them as “group objects” or to introduce them into the discussion. Similarly, participants would show interest in an object that was not their own by tilting their head to match its orientation. On a physical level, changes in orientation were often partial and temporary — not completely rotated to the other orientation, only enough for the partner to understand the intent. Participants were so comfortable with rotation that for personal puzzle-solving, orientations changed at times so frequently that it seemed almost continuous.

The results of the study indicate that within a context of puzzle-making, clear guidelines for orientation control can be discerned. The question is to what extent the puzzle-making context can be applied to a more practical activity. In particular, following
the original vision of the DigitalDesk, the application I consider to be essential is support for document interaction. There is no doubt that people would be as adept at manipulating paper documents as puzzle pieces, even though the physical dimensions differ (2cm square vs. a typical piece of paper, 10 times larger), because people are highly skilled at manipulating physical objects (Fitzmaurice, Ishii, and Buxton, 1995). However, puzzle-making is a highly orientation-specific task. The pieces start at random orientations and must be rotated to fit. It is certainly not clear that we could assume that the type of orientation control participants exhibited in Kruger and Carpendale's study is typical of document activity, so it is not possible to extrapolate any of the requirements they observed to apply them to general document manipulation.

2.1.1 The Significance of Paper Documents

Documents — paper or electronic — are an essential part of most work. From studies in real workplaces, 82% (Adler, Gujar, Harrison, O’Hara, and Sellen, 1998) and 97% (Sellen and Harper, 1997) of the work week was spent reading or writing. Predictions of a “paperless office” have not materialized; in fact, paper use continues to skyrocket (Frohlich and Perry, 1994) (Sellen and Harper, 2001). A study of knowledge workers at the International Monetary Fund (Sellen and Harper, 1997) found that while 50% of their work time involved interacting with computer documents, 86% of their work time involved paper documents. Paper use was even more significant for collaborative work: nearly half of all collaboration was by meeting face-to-face, during which paper documents were used 100% of the time. A similar study of a broader selection of professions (e.g., doctor, nurse, social worker, pilot, architect) (Adler et al., 1998) found that on average only 14% of “document time” involved computer documents, while 85% of the time paper documents were involved.

2.1.2 Why Document Orientation is a Problem

Before considering a solution to the problem of orientation for document interaction on tabletop displays, the first question should be whether there really is an orientation problem. Is reading at different orientations truly problematic, or does it simply feel difficult? This is important because if orientation is not truly problem, then no effort need be wasted finding a solution.

There has been little research about reading text at different orientations, presumably
CHAPTER 2. REQUIREMENTS FOR ORIENTATION CONTROL

because paper-based orientation was always adequate. Most of the research reported in this section was from attempts to describe the details of the reading processes, but not explicitly to evaluate the effects of reading rotated text. Thus most of the studies considered reading text upside-down, while comparing it to other transformed forms of text, such as mirrored, or mirrored-and-upside-down.

It turns out that for recognizing a single character, it may not make much difference whether it is upright or inverted: An early study by Kolers and Perkins (1969a) found that naming a single letter was on average 44% slower inverted than upright, but later studies found no significant time difference for recognizing a two-digit number (Taylor, 1972) or a single letter from a set of 6 letters (Corballis et al., 1978). Surprisingly, however, when random letters are grouped together, inverted reading speed decreases. Even when random letters are grouped as "pseudowords", reading speed compared to ungrouped (space-separated) letters is no different for upright text but 13% slower for inverted text, suggesting that reading processes other than character recognition are less effective when reading inverted text (Kolers and Perkins, 1969b). A variety of other studies confirmed that it is slower to read upside down. The speed of reading inverted text was reduced by 40% (six-word sentences) (Graf, 1981), 69% (individual letters in pseudo-word groupings) (Kolers and Perkins, 1969b), over 200% (350-word passages) (Graf and Levy, 1984), and 400% (310-word passages) (Kolers, 1968). Not only is inverted reading slower, but errors increase. As might have been expected, letter recognition errors increase, especially for letters such as b, d, p, and q (Kolers and Perkins, 1969a). Moreover, participants proofreading inverted 350-word passages detected 50% fewer typographical errors than those reading upright text, despite taking twice as long to do the task (Graf and Levy, 1984).\(^1\) Note that there is a trend for greater speed penalties with more continuous text, letter recognition < word recognition < sentences or continuous passages. This could be because recognizing single words does not involve other key mechanisms of reading, such as eye fixations and saccades, which may be even more orientation-dependent.

Only two studies were found that reported the effects of reading at orientations other than 180° (i.e., other than inverted text).

---

\(^1\) A point of interest is that reading the inverted text actually improved scores on follow-up memory tests. Before we consider rotating text to improve retention, however, it should be noted that the similar improvements were seen for reading mirrored text, which similarly slowed reading. The authors suggested that text transformations which make the letters and words less visually recognizable force the reader to depend more on cognitive processes and less on visual processes (Graf and Levy, 1984).
Koriat and Norman performed a series of experiments measuring the time to recognize a single word\(^2\), at every 20° of rotation. They found no effect for rotations up to 40°–60°, but performance degraded rapidly beyond that, reaching 70%–80% slower recognition time at rotations beyond 120°. They also found that significantly greater speed reduction occurred for longer words, e.g., at large rotations (greater than 120°), 2-letter words were only 50% slower, whereas 5-letter words were 220% slower (Koriat and Norman, 1985).

Since Koriat and Norman's stimuli were short words, their results represent a minimum performance penalty. Only one study tested reading a continuous block of text at orientations other than inverted. Tinker, a pioneer of reading research, had participants read a paragraph of text upright, rotated 45°, or rotated 90° (Tinker, 1956). He found that reading times that were increased by 50% for 45° rotated text, and 200% for 90°.

Together these findings confirm that non-upright orientations are unacceptable for reading text. For use of a tabletop display when a user cannot position himself or herself in line with the single upright perspective of the display (a circumstance that will occur at least in many multi-user situations), some kind of orientation control is needed to present upright text.

### 2.2 Requirements for Orientation Control

Unfortunately, the studies presented in Section 2.1.2 involved experimental situations that were unrealistic, generally requiring participants to be strapped in place to prevent them from tilting their head. Thus they confirm that text is highly orientation-dependent, but give us no indication about how it should be controlled. In a comprehensive survey of research on reading from paper and computer screens, Dillon observed that most research on reading controls so many variables that the interaction is no longer representative of the way we actually read (Dillon, 1992). Thus most research on reading would seem to imply, unintentionally, that if there were only one user at a tabletop display, the only necessary orientation control would be to control which side of the table was upright (i.e. answers to Q1 and Q5).

Yet from our own work habits it is clear that we practice much more detailed orientation control. When you read, do you just place a document in front of you on a desk and never

---

\(^2\)The words were Hebrew, not English, but the same processes were involved
CHAPTER 2. REQUIREMENTS FOR ORIENTATION CONTROL

touch it? Are all documents aligned squarely with the side of the table? The answer of course is that we interact dynamically with documents, moving them around as we work on them. Even the fact that a participant’s head was strapped in place reveals something about orientation, that ‘upright’ is not dependent on the side of the table, but on the orientation of the reader’s head. If the orientation of a document could not change, we would not be able to shift to a more comfortable position, which would be unacceptable from an ergonomic standpoint. For this reason alone, automatic orientation on tabletop displays would not be satisfactory unless the user’s head was tracked – even if the user only used a single document and never changed its location.

Of course, realistic interaction with documents is much more complex than that. To design interaction techniques for computer documents on tabletop displays, we need to know more about realistic interaction. Fortunately, three realistic studies of reading were performed in the late 1990’s by Xerox researchers. They examined “real-world” behaviour with documents, through field studies in actual workplaces, combining observation, extensive interviews, and diary-keeping (Sellen and Harper, 1997; Adler et al., 1998), and through a laboratory study of natural, realistic tasks (O’Hara and Sellen, 1997). This research brings to light how people interact with documents, providing clear guidelines for designing orientation solutions for tabletop displays.

2.2.1 How People Read

It is common to think of “reading” as sitting down and reading through a book. The only orientation requirement for such a static activity would be a method for picking the upright direction. Yet in general, people do not read in such a simple linear manner, and even when linear reading occurs there is great variety.

Linear reading can be broken down into different reading processes with different goals, each of which has its own reading rate. For example, Carver (1990) reports that college students reading to know the information had a reading rate of 200 words per minute (wpm), reading to understand the thoughts of the author was 300 wpm, and skimming for a particular word occurred at 600 wpm. The variety of reading activities was more clearly developed by Kenton O’Hara (1996), who created a categorization of twelve practical reading goals describing real activities, which was adapted by Alder et al. (1998) into ten categories of work-specific reading activities, shown in Table 2.2.1 (p. 11). In the last column I have summarized the proportion of reading time represented by each collection of reading
activities based on the findings reported in the paper.

<table>
<thead>
<tr>
<th>Reading Activity</th>
<th>Description</th>
<th>% Work Reading</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reading to self-inform</td>
<td>General knowledge, no specific application</td>
<td></td>
</tr>
<tr>
<td>Reading to learn</td>
<td>Intending to apply knowledge gained</td>
<td>&lt;6%</td>
</tr>
<tr>
<td>Reading in order to identify</td>
<td>Glancing to see which document it is</td>
<td></td>
</tr>
<tr>
<td>Skimming</td>
<td>Establish rough idea of what is written</td>
<td></td>
</tr>
<tr>
<td>Reading own text to remind</td>
<td>Reminders to oneself (e.g., notes, to-do lists)</td>
<td>27%</td>
</tr>
<tr>
<td>Reading to edit or critically review</td>
<td>One’s own text or others’</td>
<td></td>
</tr>
<tr>
<td>Reading to support listening</td>
<td>e.g., looking at slides in a presentation</td>
<td></td>
</tr>
<tr>
<td>Reading to search/answer questions</td>
<td>Searching by sampling information in the text</td>
<td>24%</td>
</tr>
<tr>
<td>Reading to support discussion</td>
<td>To establish a mutual frame of reference in discussion</td>
<td>21%</td>
</tr>
<tr>
<td>Reading for cross-referencing</td>
<td>Integrating information (e.g., supporting writing, reading from multiple sources)</td>
<td>27%</td>
</tr>
</tbody>
</table>

Table 2.1: Workplace Reading Activities (Table by the author, based on the work of Adler et al. (1998))

The surprising finding is that the activities that might be thought of as “sitting and reading straight through” (the first two categories) are so minimal. More varied reading activities can be seen in the next four categories: skimming, “glancing”, and editing. The last three reading categories, representing over 70% of reported reading time, consist of activities that require going back-and-forth between multiple documents or multiple collaborators. This confirms that collaborative reading — with the accompanying orientation problem — cannot be ignored, and that the control solution will need to accommodate multiple documents.

2.2.2 How People Write

Writing, just like reading, is not as linear as one might have thought. Document authoring is more complex than a simplistic notion of “sitting at the computer and writing”. Sellen and
Harper's research at the IMF (Sellen and Harper, 1997) found that for individual document authoring (creation and editing), nearly 100% of the time a computer was used, but that 89% of the time paper documents were also used simultaneously. This finding was confirmed by Adler et al. (1998), who found that participants used at least two documents simultaneously nearly 50% of all time spent writing on paper or computer. They further found that what we normally think of as creating or modifying a document only accounted for 16% of writing time. The vast majority of writing time (on paper or digital mediums), was spent making annotations\(^3\) and notes (48%), or filling out forms\(^4\) (34%).

Habits for collaborative writing — group authoring and editing, or reviewing another person's text — were even further removed from a simple concept of document creation at a computer. For 18% of the collaborative authoring process at the IMF, both computers and paper documents were used; the rest of the time only paper was used (Sellen and Harper, 1997).

In theory, tabletop displays should be able to provide users with the rich interaction they are used to with paper documents. However, it is clear that the way people write is just as non-linear as the way they read. The varied reading and writing activities have implications for the document interaction that tabletop displays need to support, and the method of orientation control. The question for us is, what exactly do people do with paper, and what are the implications for orientation control?

### 2.2.3 Document Actions

O'Hara and Sellen (1997) specifically studied the activities that people perform with paper documents, attempting to answer the question of why so many people still prefer paper over computer documents. They described the actions that people perform with paper, and identified three activities that are most in need of computer support: navigation, annotation and note-taking, and spatial layout.

Navigation involves activities such as fixed page layout, flipping through pages at variable rates, temporary page-marking with fingers, and intrinsic information such as the reader's location in a document. It does not relate to orientation, except perhaps that fixed page

---

\(^3\)Annotation refers to marking directly on a document, such as writing in the margin, highlighting or underlining

\(^4\)The form-filling primarily reflected the medical professionals included in the study. Due to the mobile environment in which they work, form filling may not apply to the context of desk and table interaction
layout, which primarily aids spatial memory, allows for multiple pages in a document to be spread out on the surface. This behaviour helps people organize and understand information. The one aspect of navigation that relates to orientation is that multiple pages from a single document can be spread out, so that if those pages were at different orientations, the "single document" might actually not have a single orientation. The act of spreading out pages on a surface is an example of the most important document activity relating to orientation: spatial layout.

![Spatial layout](image)

Figure 2.1: Examples of rotated documents from the literature on paper-based work habits

Spatial layout has immediate repercussions for tabletop orientation control.

The use of multiple documents involves arrangement of those documents on the surface, an examples of which are shown in Figure 2.1 (p. 13). The spatial layout of documents is used in many different ways. O'Hara and Sellen (1997) observed that people use a central area of focus where the active documents are located, with other documents "stored" on the periphery of the surface, around the area of focus. A document in the periphery is visible and readily accessible, and in an instant the user can move it to the focal area to work with it. In addition, people intersperse different reading and writing activities, e.g., reading something while making notes, or interrupting the writing of a document to refer to other documents (Adler et al., 1998). People use their work surface to organize information, with different categories of information distinguished by the location of the document **cite
They form piles that allow them to manage information without requiring preconceived structure, and in a pile the orientation of documents often marks different sections (Mander et al., 1992).

The observations by O’Hara and Sellen (1997) identified a number of characteristics of document manipulation. The manipulation must be fast: it must be possible to change the layout quickly and easily in the flow of the activity, even “several times within the space of seconds”. In addition, since users often moved one document without stopping another reading or writing activity, it must be possible to manipulate the document without much attention. Manipulation of documents must be possible independently with either hand.

The key is that the layout of documents involved not just their position, but also their orientation. People consistently chose to rotate individual documents. Observations in the Gruvi Lab at SFU of collaborators seated side-by-side and at adjacent sides of a table confirmed that the people did rotate documents to show their partners information, just like the behaviour observed by Kruger and Carpendal (Kruger and Carpendale, 2002). However, O’Hara and Sellen’s (1997) findings reveal that document rotation is just as vital for individual work. The reason users rotated documents so freely could not always be clearly established, though it had meaning for the individual. In some cases the motivation was evident, for example, documents being read were rotated 0°–20° off of vertical for legibility or comfort, since (O’Hara and Sellen, 1997), or documents could be placed at a particular rotation so that it could be quickly identified based on orientation (Mander et al., 1992).

A special case of rotation was the ergonomics of handwriting. Recall that annotations and notes accounted for nearly half of workplace writing (Adler et al., 1998). O’Hara and Sellen (1997) claimed that annotation was one of the main reasons people chose paper over computer documents, since annotations and notes provide a record for later review, help readers to understand what they are reading (O’Hara and Sellen, 1997; Marshall et al., 1999), communicate to other readers (Luff et al., 1992; Posner and Baecker, 1993), and are the primary form of writing for editing and collaboration (Adler et al., 1998). Annotations and highlighting can be supported by software (e.g., (Ginsburg et al., 1996)), and such functions are now found in common applications such as Microsoft Word and Adobe Acrobat. However, handwriting is not effective on desktop computers because direct input handwriting on a vertical desktop monitor causes unacceptable arm fatigue due to the lack of physical support (Greenstein and Arnaout, 1988), and indirect input such as a drawing tablet requires hand-eye coordination so difficult that only highly practiced users are comfortable
CHAPTER 2. REQUIREMENTS FOR ORIENTATION CONTROL

with it (Fitzmaurice, Balakrishnan, Kurtenbach, and Buxton, 1999). Other computer form
factors, such as tablets, allow physical arm support for direct input; as tabletop systems
such as the DigitalDesk have demonstrated, tabletop displays provide such arm support and
make direct input handwriting possible. Yet even support for annotations may be insuffi-
cient. O'Hara and Sellen (1997) observed that people don't just make text annotations:
they handwrite, sketching, employing subtle cues such as line thickness (Luff et al., 1992),
and making “idiosyncratic” marks that have meaning for them and act as memory aids
(Marshall et al., 1999). Because people are so skilled at handwriting, they can control
these subtle cues without being distracted from what they're reading (Schilit et al., 1999), which
is not currently the case with digital annotations.

For my work, it is not important to establish whether current computer annotation
systems are successful. The point is that for tabletop displays to support handwriting,
there are implications for orientation and position control. Whereas reading requires gen-
erally upright orientations, handwriting requires non-upright orientations. Guiard (1987)
demonstrated that during handwriting, the paper is rotated significantly with respect to the
writer's body in order for the writing hand to move horizontally with respect to the page.
Typical angles are 30°–45° off-vertical for writing (O'Hara and Sellen, 1997), and even larger
rotations for drawing (Fitzmaurice, Balakrishnan, Kurtenbach, and Buxton, 1999). People
constantly make small adjustments to the document's position and orientation with their
non-preferred hand, so that handwriting occurred in a small area, to maintain a comfortable
position as well as orientation\(^5\) (Guiard, 1987; O'Hara and Sellen, 1997).

Thus people manipulate the orientation of individual documents for spatial organization,
reading, and handwriting, moving them quickly and frequently, with either hand, without
having to think about it.

2.3 Implications for Orientation Control

The way people interact with documents allows us to answer some of the questions of Kruger
and Carpendale (2002) for document interaction on tabletop displays.

Q1 asked whether rotation of individual objects on the display would be preferable to
rotation of entire windows or the entire screen. Rotation of individual documents — or

\(^5\) Guiard (1987) stated that writing speed was reduced by 20% when people are prevented from using their
non-preferred hand to manipulate the page ((Athènes, 1984), quoted in (Guiard, 1987))
actually, individual "sheets" within the page layout of a document — is necessary. Rotation must not be limited to the entire screen or to collections of objects.

Q5 asked whether it was better to support freehand rotation than orthogonal rotation. Documents are rotated to specific angles for comfortable reading and writing, relative to the user rather than the table (thus the required orientation may change when the user simply shifts to a different position). In addition, the orientation of non-active documents on the table can have meaning for organizing information, regardless of the user's position. Orientation should not be limited orthogonal angles or any other scheme that does not allow fine angle variations.

Q4 asked whether orientation control would result in decreases in productivity or efficiency because of the constant requirement to re-orient objects. The evidence from work habits showed that current productivity already involves paper documents the majority of the time. Since paper documents are themselves orient-able objects, there is no reason why providing orientation control on tabletop displays should inherently result in a decrease in productivity. However, the efficiency of paper-based orientation control is closely tied to the speed and ease of our natural physical abilities with paper (Schilit et al., 1998), so the actual effectiveness of orientation control will depend on the particular mechanism providing the control. If it is less efficient than interacting with paper, then productivity could certainly be adversely affected. The issue of the orientation control mechanism relates to the final question:

Q2 asked what orientation control techniques could be used to provide "quick and intuitive rotation of objects". I identified a number of guidelines based on research about how people use paper documents. Tabletop document manipulation should:

- be quick: the user should be able to perform many adjustments in a few seconds.
- control both position and orientation with the same action. Thus we should not only provide orientation control but manipulation, i.e. orientation and position control.
- be possible with the non-dominant hand as well as the dominant hand.
- be allowed on many documents simultaneously. A user should be able to perform simultaneous actions on different documents, and multiple users should be able to work at the same time.
- be easily intermixed with other actions, such as writing/sketching
• be basically ‘invisible’. The user should not have to explicitly think about manipulating the document, or be distracted from the goal they are trying to achieve, such as organizing information, or communicating with a colleague.

The question of what techniques can meet these requirements is the subject of Chapter 3.
Chapter 3

Techniques for Manipulation on Tabletop Displays

In Chapter 2, requirements were identified for freehand, dynamic control of the orientation and position of individual objects on tabletop displays. It was also observed that the performance will be closely tied to the effectiveness of the particular technique that is used. In this chapter, I consider what techniques have been used to solve the orientation problem, and how well they satisfy the needs of document manipulation.

3.1 Techniques for Solving the Orientation Problem

A variety of strategies exist for dealing with orientation, either by attempting to remove the problem or by providing orientation control. A review of such techniques are presented to allow an evaluation in light of the requirements identified in Chapter 2.

3.1.1 Non-oriented Interfaces

One approach on tabletop displays has been to completely remove the need to change orientation, displaying only content that is orientation-independent. This was demonstrated on a tabletop display in an intentionally non-oriented board game (Mandryk et al., 2002).

Another nearly-non-oriented approach that has been suggested is to display repeated copies of orientation-dependent objects at multiple orientations. This can be as simple as a compound object with identical copies at different orientations so you can understand...
it from anywhere around the table, such as an “omniviewable” menu that is really a few identical menus pointing in outwards directions (Balakrishnan et al., 2001). It can also be more complex, where multiple copies of the same information do not necessarily share identical appearance, e.g., in a multi-user situation, each user could have their own view of the same shared document (Kruger and Carpendale, 2002).

3.1.2 Automatic Orientation

Automatic orientation solutions are based on the idea that if you know where users are, then you can display objects upright for them.

One approach has been to apply a rule, inferring the desired orientation from the object’s location on the display. Thus, in Augmented Surfaces, the orientation was assumed to be towards the nearest edge of the table (Rekimoto and Saitoh, 1999). Similarly, each individual display in ConnectTables had a defined orientation; objects would be oriented according to the particular ConnectTable on which they were located (Tandler et al., 2001). MERL’s tabletop display was motivated by a collaborative assumption, so each object was aligned radially based on its location, i.e. pointing outwards on the round table (Vernier et al., 2002).

The other strategy is to actually know where users are. However, that is not necessarily an easy task. A trivial implementation is simply to ask the user to indicate their location. The ‘magnet’ metaphor demonstrated by Vernier et al., a control point towards which all the objects on the table would orient, could be interpreted in this way.

Instead of the user having to explicitly enter their position, it might be desirable for that to be determined automatically. Balkrishnan, Fitzmaurice and Kurtenbach (2001) suggested that referential GPS could be used to track an individual’s location, but generally in interface design it is inadvisable to require users to wear a tracking device like that. Other possible tracking techniques would be pressure panels in the floor (Balakrishnan et al., 2001), or tracking the chair in which the user is sitting. In addition, it is probable that computer vision technology could successfully track user locations at tabletop displays. With multiple users, the tracking becomes more complex since you would need to know not only users’ locations, but also to be able to connect each action performed on the table with the person who performed the action (so that the correct orientation could be determined). MERL’s recently developed DiamondTouch touchscreen display technology does exactly that, identifying the user by passing modulated electric field through the user via their
chair, a somewhat shocking\textsuperscript{1} approach (Dietz and Leigh, 2001).

A completely different approach to automatic orientation was developed by Hancock (2001), who showed that a neural network could be used to reliably determine on which side of a rectangular table a user is located, using only the angle at which they held the stylus when clicking or writing. This system was demonstrated on a tabletop with a pair of users sitting across from one another, sharing a single tracked stylus (Hancock et al., 2001). It was an elegant solution since there was no need to track users' locations or identify who performed an action: the action itself contained all the information needed to infer the location. However, their tabletop system used a stylus equipped with a 6DOF magnetic tracker, a very expensive feature not common in stylus input.

3.1.3 Orientation Control

Another general approach is to provide users with means for controlling orientation. Some techniques have already been used with desktop computers, while others have been used specifically on tabletop displays.

Explicit Selection of Rotation Angle

A trivial method of user control is to abstractly specify the angle of rotation. This method is actually used on desktop computers. For instance, Microsoft Paint (the Windows 2000 default bitmap program), provides a dialogue box that lets you rotate the image by 90°, 180°, or 270°. In less restrictive programs, you would be able to use the keypad to freely specify the numerical angle to rotate. As rudimentary as this technique might appear, in a layout program with a number of objects to be lined up precisely it is the easiest way to accomplish the task.

Rotation Mode

There is a well-established, interactive technique for rotation on desktop computers. In it, the mouse drags a control point to rotate the object around a fixed centre of rotation. Many photo and drawing programs use this technique, where interactive rotation is combined with other interactive manipulations, such as position and size, in a 'free transform' state.

\textsuperscript{1}Get it? : )
Typically, the object is overlaid with a bounding box and control indicators that show how to initiate the spatial modes: e.g., you rotate the object around its centre by dragging the corners of the box (e.g., Corel Draw), or outside the box (e.g., Adobe Photoshop), and you translate the object by dragging within the box.

The control point is necessary because a 2D pointing device is insufficient to control both the 2D position and the 1D rotation of the object. The region in which a pointer drag action is initiated determines the mode of interaction, e.g., dragging the centre initiates translation, while dragging the corner initiates rotation. This is call a mode switch. In the case of control points, the mode switch is spatial, i.e. the mode is determined by where you click. Mode switches can also be temporal, where the mode is determined by the order of a sequence of events. A common temporal mode switch is the ‘tool’ used in paint programs, where all actions on the canvas are interpreted according to the selected tool, until the mode is changed again. Such modes are commonly used on desktop computers: in particular, the ‘corner to rotate’ spatial mode is unquestionably the standard method for orientation control on computers.

Some tabletop researchers have used variations on desktop modes switches to induce the rotation. The magnet technique on the MERL table mentioned in Section 3.1.2 (p. 19) was an example of a different spatial mode, with the magnet icon as the control point instead of the corners of a rectangle.

Another increasingly popular mode-based approach uses a temporal mode: gestures. The i-Land table used a stylus or pen gesture to initiate rotation, though no details were given about the technique or its success (Streitz et al., 1999). Guimbretière and Winograd (2000) designed special flow menus for rotation control, though they did not indicate whether this technique was actually used on the tabletop display in the Stanford iRoom.

Two-handed interaction

One way to get around the limitations of a 2D pointing device is to have a second device in the other hand. Buxton and Myers (1986) found that users were able to manipulate the position of an onscreen rectangle with a pointing device in their preferred hand, while scaling the size of the rectangle with a slider in their non-preferred hand. A similar device could control orientation instead, or it could be replaced with more appropriate control such as a dial instead of a slider. Two handed position and orientation control can also be achieved with a pointing device in both hands. There are two possible methods, symmetric
(the two hands control the position of opposite corners of a rectangle, allowing the user to manipulate its position, orientation, and size), or asymmetrical (one hand controls the position and the other hand controlling the orientation, relative to that position), both of which were demonstrated on a tabletop display by Fitzmaurice (1996). The same techniques were also used for drawing objects on the EnhancedDesk (Koike et al., 2002).

**Multi-point Touch / Hand Input**

Just as two points controlled by two hands can provide 3DOF control, so can two points controlled by a single hand. Unfortunately, the nature of today’s touchscreens — capacitive or resistive touch-sensing membranes — is that they detect only one input point (Quinnell, 1995). Thus if you touch the membrane in two locations, the computer will recognize just one input point, perhaps halfway between the two points of contact. In theory, better transducers of hand posture could exist. The technology exists in theory to make multi-touch, pressure sensitive touch screens based on ultrasonic detection of screen deformations (Quinnell, 1995), but apparently such screens are not available; similarly, multi-touch video-based ‘touchscreens’ exist, but only in small formats that are unsuitable for tabletop displays (e.g., 5”x7” gesture pads from a company called Fingerworks).

A number of tabletop displays for Virtual Reality applications have used sensing gloves with embedded 6DOF sensors to get hand input, e.g., (Krüger et al., 1995). However, glove input is not recommended because it is generally undesirable to require users to wear obtrusive input hardware. Besides, the glove-based input systems demonstrated on tabletops required the use of 6DOF trackers, which can be more easily used directly as graspable input rather than attached to a glove, since the tracker can sit on the tabletop surface when it is not being used.

The bulk of tabletop hand input has been achieved through computer vision techniques. Even the Digital Desk used vision-based finger pointing and rudimentary two-handed gestures. The University of Washington HITLab ‘Hi-Space’ tabletop display allowed users to hold their hands above the tabletop, while the shadow formed by top projection was detected as input, but it appeared to have registered only as a coarse input point.² Multi-point hand input was demonstrated on a tabletop display using a new sensing technology, SmartSkin, that detected the capacitive field caused by fingers or different hand shapes (Rekimoto, ²Based on a demonstration of the system in April 2001)
CHAPTER 3. TECHNIQUES FOR MANIPULATION ON TABLETOP DISPLAYS

2002). In theory, this detection of hand shapes should allow one-handed orientation control, though the SmartSkin input appeared to be so coarse that it was unclear whether it could effectively detect the rotational DOF. The most successful hand input on a tabletop display has been developed for the EnhancedDesk tabletop project (Sato et al., 2000). Using an infrared camera, they were able to precisely track a fingertip and later multiple fingertips, as well recognizing a variety of hand postures, with very high accuracy (Oka et al., 2002). It was with this system that they demonstrated two-handed rotation (Koike et al., 2002). The ability to reliably recognize hand shapes and multiple fingertips in real time should support one-handed orientation and position control.

Rotation of the Physical Display

Just as the MERL magnet simulated rotation of the entire screen, orientation can be controlled by actually rotating the physical display. An advantage of rotating the display itself is that the computer system does not need to address the issue of orientation at all, including rotating the graphics (a potentially costly step of computation). However, tabletop displays are by definition large and often heavy. The technique can be simulated using a tangible interface consisting of a lazy-suzan in the centre of the table, which spins the virtual workspace projected on it, similar to the collaborative augmented reality application demonstrated by Regenbrecht and Wagner (2002). For actual rotation of tabletop displays, only the ConnecTables in GMD’s i-Room were small enough that they could potentially be manipulated this way (Tandler et al., 2001). It is likely that in general only smaller displays are suited to this sort of interaction, e.g., portable reading devices. Reading devices, as developed by Schilit, Price, and Golovchinsky (1998), are essentially tablet PCs with document-related features like fixed page layout, page turning, and free-form annotation. In addition to being rotatable through physical manipulation of the device, they do not suffer from the legibility problems that tabletop displays do. Tabletop displays must provide coverage for the whole work surface, but at any instant the document being worked on only uses a small proportion of the pixels on the table. Furthermore, graphical rotation on the display grid can be computationally expensive (Fitzmaurice et al., 1999) and degrades the legibility of text, since fonts are optimized at a pixel level in a manner that depends on the rectangular grid (Biggs, 2003). However, such small devices lack support for some of the document interaction that tables do provide, such as spatial layout of many documents, and shared public display space for collaboration.
CHAPTER 3. TECHNIQUES FOR MANIPULATION ON TABLETOP DISPLAYS

Multiple-DOF input

If a 2D pointing device is insufficient, it can be replaced with an input device that provides an additional DOF. Hinckley, Pausch, Goble, and Kassell (1994) demonstrated that a physical object tracked in 6DOF space could act as an input device, and it was very effective for performing a 3D position-and-orientation control task. Multi-DOF input has been demonstrated on desktop computers with the Rockin’ Mouse, a tablet-mouse that can tilt side to side (Balakrishnan et al., 1997), and with a mouse that can tilt front and back, and even lift up and down off the surface (Hinckley et al., 1999). Ideally, in a device for controlling orientation, the additional DOF would be rotation. Two techniques using a rotational DOF have been demonstrated on desktop computers. The first added a rotational DOF to the typical mouse, which was achieved by adding a second mouse ball (MacKenzie et al., 1997) or by using an optical mouse with special video processing (Hinckley et al., 1999). The second was achieved by modifying the input devices from a normal absolute position tablet, so that two 2D position points are detected within a single physical input object (Fitzmaurice, 1996), or even more DOF within a close range of the tablet surface (Rekimoto and Sciammarella, 2000), allowing the user to control the onscreen position and orientation simply by moving a physical object.

Many tabletop displays have used direct, multi-DOF input in the form of graspable user interfaces (GrUIs), which were first introduced on a tabletop display by Fitzmaurice, Ishii, and Buxton (1995). Physical object manipulation has been used on tabletops to manipulate the whole display (Ullmer and Ishii, 1997; Rauterberg et al., 1997; Fjeld et al., 2001), text labels within the display (Patten et al., 2001), or more complex interactions where the position and orientation are tracked (Underkoffler and Ishii, 1998, 1999; Patten et al., 2001). In all of these examples the physical input occurred directly on the display rather than on a tablet, which was the case with desktop interaction. Tracking of the objects was achieved through magnetic 6DOF tracking (e.g., (Fitzmaurice et al., 1995)), computer vision tracking from underneath a translucent table surface (e.g., (Ullmer and Ishii, 1997)) or from above the table (e.g., (Fjeld et al., 2001)), or using custom tracking hardware (e.g., (Patten et al., 2001).

One graspable / tangible interface that is particularly associated with tabletop displays is the use of a sheet of paper as the physical “handle”. This research falls under the area of augmented reality as well as tabletop research, since it involves projecting the computer
image onto the piece of paper (i.e. the computer image ‘augments’ the blank piece paper). Projecting onto paper on tabletops dates back to Wellner’s Digital Desk (Wellner, 1993). There have been increasingly successful implementations of paper tracking using computer vision, such as the later DigitalDesk work at Cambridge (Robinson et al., 1997; Robertson and Robinson, 1999), EnhancedDesk at University of Tokyo (Koike et al., 2001), and the Visual Interaction Platform at Eindhoven University of Technology (TUe) (Aliakseyeu et al., 2002a,b). The Visual Interaction Platform used a particularly elegant method, tracking an invisible pattern printed on the paper with ink that reflects infrared light, rather than the typical large black and white registration tags.

3.2 Evaluation of Techniques Meeting Basic Requirements

This overview of orientation solutions has covered a wide variety of interaction techniques (ITs). Evaluating them in light of the guidelines from Chapter 2 will allow us to eliminate many of the techniques in consideration.

Non-oriented interfaces are not usable for document interaction. Text cannot be orientation-independent, so an un-oriented interface is impossible. Omniviewable interfaces are not a sufficient solution either; even if multiple copies of documents for collaboration were to be considered, such a solution only provides multiple fixed orientations, whereas document interaction requires user control over orientation.

For the same reason, automatic orientation will not provide the necessary control. Preset, location-dependent orientation restricts free manipulation of individual documents. Similarly, even if users’ locations were successfully tracked with great detail (such as the tilt of their head), automatic orientation would still preclude user control over the orientation of individual documents. Automatic orientation might turn out to be a helpful feature in a hybrid solution — for instance you could provide low-key automatic orientation yet still allow user control (Kruger and Carpendale, 2002) — but automatic orientation itself does not provide the orientation control users need.

A number of active user-control methods can also be eliminated. Numerical angle entry does not support the necessary dynamic interaction. Rotating the whole display does not provide control of individual objects. Two-handed rotation techniques contradict the requirement for independent either-handed manipulation.

Eliminating all these orientation techniques leaves three main categories from the ones
proposed to date: mode-based rotation with a 2D pointer, multi-point finger or hand shape input, and 3DOF physical object input.

3.3 Distinguishing Characteristics of Manipulation ITs

Even within the three remaining categories (Section 3.2), there are many possible variations in technique. It would be helpful to be able to distinguish between the ITs, and to identify ITs that are likely to be superior.

3.3.1 Integral Rotation and Translation

The observations in Chapter 2 indicated that the position of objects was changed together with moving the orientation.

Using the standard computer rotation control technique, "corner-rotation" with a mouse, my subjective reaction is that it does not feel "natural". It is frustrating because there is a discrepancy between intention — thinking 'move the object over there' — and the actual technique having to separate translation and rotation into sequential actions.

Jacob, Sibert, McFarlane, and Mullen (1994) examined the relationship between attributes we perceive about objects, and manipulation of those attributes. Certain attributes of objects were shown to be integral, meaning they combine perceptually, e.g., the vertical and horizontal position of a dot on a page. Other attributes are separable, e.g., object colour and shape are perceived as distinct. They showed that input devices can also be described in a similar way: an IT is integral if it is natural to control all of its dimensions simultaneously, otherwise it is separable. They results of their study confirmed that when users manipulate integral attributes, an integral IT is more appropriate, but when the attributes they manipulate are separable, a separable IT is better.

The reason integrality is an issue is that currently, standard computers use a 2D pointing input such as mouse or touchscreen. With just two degrees of freedom (x,y movement), controlling three degrees of freedom -- position (x, y) and orientation (rotation in the plane of the display) — requires creative interactions such as the use of two different manipulation modes. The corner-rotation technique using manipulation modes is an example of a separable IT, and therefore it is most effective only for a control of separable attributes.

\[\footnote{I will later refer to this as \textit{separated}, indicating that the manipulation is restricted to separable interaction.}\]
CHAPTER 3. TECHNIQUES FOR MANIPULATION ON TABLETOP DISPLAYS

Jacob et al. (1994) did not determine whether object position and orientation were integral or separable. It seems plausible that an object’s position and orientation are integral, especially in light of the frustration of using a separated IT, but one cannot be sure without an empirical demonstration.

Fortunately, this very issue was the focus of a 1998 kinesiology-HCI paper by Wang et al. (Wang et al., 1998). Participants aligned a small wooden block with a target at a different position and orientation on the same horizontal surface. There were two conditions, one normal and the other in which participants could not see the block they were moving, just the target\(^4\). In both conditions, moving and rotating the block were integral. They concluded that position and orientation are integral based on the perceptual structure of the object, and furthermore that position and orientation are even integral based on motor control, distinct from the visual perceptual characteristics. For the purposes of interaction techniques, the fact that position and orientation are integral means that appropriate techniques for manipulating documents on tabletop displays should provide integral control.

The effect on performance of integral vs. separated devices is not clearly established, and may be very dependent on the application. There has not been a lot of research on multiple DOF input. Some of the non-tabletop 3DOF techniques listed earlier were demonstrated, but not empirically evaluated, e.g., the two-ball mouse (MacKenzie et al., 1997) and video-mouse (Hinckley et al., 1999). Similarly many examples of integral input have been demonstrated on tabletop displays, but I am not aware of any experimental examination of their effectiveness. A number of studies evaluated 6DOF integral control but did not compare it with separated control (e.g., (Hinckley et al., 1994; Zhai and Milgram, 1998; Masliah and Milgram, 2000)).

One experiment that compared manipulation techniques of different integrality was the introduction of the Rockin’ Mouse (Balakrishnan et al., 1997). Participants controlled the 3D position of an object using a spatial mode switching interface with normal mouse input, or the mostly-3DOF Rockin’Mouse. The Rockin’ Mouse was shown to provide integral control of the three dimensional position, performing 30% faster than the mouse with mode-switching. A different study by Hinckley et al. (Hinckley et al., 1997) used a three dimensional rotation-only task, and found that an integral IT (handheld 6DOF tracker)

\(^4\)Showing the target while blocking the object they moved was accomplished using head-tracked augmented reality to create a virtual target that appeared to the user to be physically stationary at the target location.
performed up to 36% faster than two separated, mouse-based ITs.

The only study closely related to tabletop object manipulation was Experiment 1 in Fitzmaurice's Ph.D. dissertation (1996). The task was to line up an onscreen rectangle with a target by controlling one or more of three integral attributes: position, orientation, and size\(^5\). The ITs used were a stylus on a digitizing table using the separable corner-drag technique, or integral 3DOF tablet pucks. The results confirmed that for the integral ITs there was only a nominal increase in completion time for higher DOFs task conditions, while the separated mode-switching was over 100% slower when an additional DOF was added.

### Integrality of Tabletop Manipulation ITs

We can estimate the integrality of the techniques, as shown in Table 3.3.1.

<table>
<thead>
<tr>
<th>Integrality</th>
<th>Interaction Technique(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Separated</td>
<td>Corner-Mode, Gesture-Mode</td>
</tr>
<tr>
<td>3DOF (but not rotational)</td>
<td>Rockin' Mouse</td>
</tr>
<tr>
<td>3DOF</td>
<td>Rotateable mouse, 3DOF tablet, multi-finger input, TUI handle, Augmented paper</td>
</tr>
</tbody>
</table>

Table 3.1: Integrality of tabletop manipulation ITs

#### 3.3.2 Direct Input

A second characteristic for distinguishing orientation control techniques can be drawn from one of the principles behind Wellner's development of the DigitalDesk (1991), direct input. He highlighted the fundamental difference between mouse-pointing interaction that was labelled *direct manipulation* in the 1980s because it provided a visual metaphor and feedback (Schneiderman, 1998, p.71 & p.229), and real direct manipulation of reaching out with your hand and pointing with your finger.\(^6\)

Direct input refers to an IT in which the input space and the display space coincide, such as touchscreen or stylus input on the tabletop display, as opposed to the indirect input of

---

\(^5\)Position and size were also shown to be integral by Jacob et al. (1994).

\(^6\)In an ironic turn of events, the label he applied to direct touch input, "tangible" manipulation, has since been redefined in the same way that he had reclaimed direct input.
CHAPTER 3. TECHNIQUES FOR MANIPULATION ON TABLETOP DISPLAYS

mice and digitizing tablets. Direct input has long been technologically possible on desktop computers using light pens, but it was unusable on desktop computers due to the fatigue associated with its use on vertical monitors, caused by users being unable to rest their arm on a support surface (Greenstein and Arnaut, 1988). Different direct input technologies have found niche uses, e.g., high-wear, walk-up input such as touchscreens registers in restaurants and public information kiosks. In the past decade, small palmtop computers (and more recently, tablet PCs) have used stylus or touch input, which was usable because the portability of the devices allows physical support through other means, e.g., using the device in your lap or on a table. One of the advantages of tabletop displays is that the table surface can provide the necessary physical support for direct input, so the severe fatigue associated with direct input on vertical displays should no longer be a problem on tabletop displays.

In some cases, direct input can provide better manipulation than indirect input. Ware and Rose (1999) had users perform a 3D orientation task in a VR environment, using either direct input, or similar input offset by 60 cm from the displayed object. The direct input condition was 35% faster than the indirect condition. Direct input does not necessarily improve performance in all circumstances: on a large surface like a tabletop display, it requires large arm movements. Indirect input, such as a mouse, can be “multiplied” (i.e. small movements in the input space result in a large movements in the display space), and may therefore be faster than direct input for certain actions (Zhai and Milgram, 1998). Nevertheless, direct input taps into an often-ignored natural sense: kinesthesia, the sense of the location and movements your limbs. People seem to have a better memory of where they placed objects when using a touchscreen than using a mouse, a sort of kinesthetic memory ‘for free’, i.e. without needing to think about it (Tan et al., 2002). Patten and Ishii (2000) had users organize newspaper articles, with either indirect input (moving icons on a vertical screen), or direct input (moving tangible pieces on the surface of a table). Generic icons and TUI blocks were used to remove any visual indication of which article each represented. After the task was completed, a recall task was performed, and participants using direct input had better recall of which objects corresponded to which articles, often employing a strategy based on a reference frame relative to their bodies to help organize the articles on the table surface (Patten and Ishii, 2000).

Direct input may also benefit collaboration, due its implicit non-verbal communication. Inkpen, Hancock, Mandryk, Scott (2002) reported that a pair of participants performing a
collaborative task at a tabletop display gestured significantly more when using direct input (two styli) than indirect input (two mice). This sort of finding led Mandryk et al. (2002) to identify direct input as one of the key factors influencing collaboration, since it improves communication associated with actions, peripheral awareness of partners’ activities, gestures, and shared-knowledge references to common objects.

There is a further, less tangible reason why direct input may be preferable: users express frustration with tabletop systems that do not support direct input. In a preliminary user study of the Edgelab tabletop display, users reported that they felt using mouse was unnatural, and they “felt compelled to manipulate objects on the screen using their hands, as they would when manipulating objects on a table” (Andrews et al., 2000). Very similar feedback was received from users of an earlier tabletop system, who expressed a strong desire to “use their hands and fingers to directly interact with objects on the display surface” (Forsberg et al., 1998). The main thrust of tangible interface work was undergirded with the notion that using physical artefacts on tabletop displays “[makes] interface elements more ‘direct’” (Fitzmaurice et al., 1995), and a recent design guideline for interaction devices identified direct input as one of the five key goals, making it a key feature of their tabletop display implementation (Aliakseyeu et al., 2002b).

**Directness of Tabletop Manipulation ITs**

A second aspect of interaction that allows us to distinguish the techniques is the directness of the input, as shown in Table 3.3.2.

**3.3.3 Exploration Space of Tabletop Manipulation ITs**

The factors of integrality and directness may be interrelated. Zhai and Milgram (1998) compared 6DOF ITs of differing directness, presenting a table with the different overlapping directness IT factors (e.g. absolute vs. relative input, multiplied vs. isospace); in the same table, they listed some 2DOF examples. Using the same approach, but extending it to 3DOF manipulation and using integrality as a second axis, an IT design space can be formed, as shown in Figure 3.1 (p. 31). The visual representation of the available ITs enables easy comparisons of the tradeoffs between different approaches. In general, ITs on the bottom right are most desirable.
Figure 3.1: Integrality and Directness of ITs for Manipulation on Tabletops
CHAPTER 3. TECHNIQUES FOR MANIPULATION ON TABLETOP DISPLAYS

<table>
<thead>
<tr>
<th>Directness</th>
<th>Interaction Technique(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indirect (relative)</td>
<td>Mouse, Rockin' Mouse, Rotateable mouse</td>
</tr>
<tr>
<td>Indirect (absolute)</td>
<td>Puck or stylus on digitizing tablet, Small multi-point touch panel, 3DOF puck on tablet, Indirect multi-DOF input (e.g., Sensable Phantom)</td>
</tr>
<tr>
<td>Nearly-Direct</td>
<td>Touchscreen overlay offset from the surface of the display, 3D tracked input with indistinct on/off boundary</td>
</tr>
<tr>
<td>Direct input</td>
<td>Touchscreen, Stylus, Multi-finger tracking, Magnetic- or Optical-tracked tangible handle (i.e. one tool for manipulating many objects)</td>
</tr>
<tr>
<td>Direct manipulation of physical objects</td>
<td>Graspable UI / TUI, Augmented Paper</td>
</tr>
</tbody>
</table>

Table 3.2: Directness of tabletop manipulation ITs

3.4 Practical Roadblocks

It is easy to say that integral, direct ITs are the most desirable; it is not so easy to implement them. All of them require multi-DOF tracking on the tabletop display, and all current tracking technologies have serious drawbacks.

A commonly-used prototyping technology, 6DOF magnetic trackers, have input components that must be tethered (i.e. they have a wire), which affects manipulation; they can suffer from calibration and drift errors, particularly around metal which may prevent them from being used near LCD or Plasma displays; and without a touch sensor on the input component there may be an indistinct ‘on/off’ boundary that does not correspond exactly with the physical surface of the screen. More importantly, though, the cost of the hardware starts at $10,000. Other tracking hardware is either custom-developed and unavailable (e.g., (Patten et al., 2001; Rekimoto, 2002)), or as prohibitively expensive as magnetic trackers (e.g., NDI’s Optotrak).

Computer vision tracking of finger and tangible interface components may suffer from some of the same problems, such as the indistinct ‘on/off’ boundary and calibration problems. While vision tracking avoids tethered components and errors around metal, it has an equivalent problem: occlusion. Losing track of the object is not just inconvenient, it affects interaction (Aliakseyeu et al., 2002b), and may be very difficult to avoid even if complex systems of multiple cameras are used. The hardware for vision tracking is cheaper than
magnetic trackers, though still possibly costing thousands of dollars (e.g., infrared cameras), but the important factor is the fact that the techniques are at the forefront of computer vision technology and thus are all custom-developed by teams of researchers. Such systems are simply not available, and great resources would be required to develop your own system, making them possibly even more costly than commercial trackers.

It is also worth noting that multi-finger touch screens, promised to be available soon, will face a big challenge even when they do become available: distinguishing unintentional touches from intentional actions. If natural non-control actions, such as leaning on the table, or placing an object on the surface, were interpreted as input, the tabletop would become almost unusable.

All these factors reveal an unspoken requirement: low cost and availability. It is technologically possible to provide integral, direct input — as many tabletop display projects have demonstrated — but the expense may be unacceptable for many. It is not possible to buy a tabletop display system equipped with direct, integral input: the ITs would have to be custom-developed. Realistically, organizations interested in a tabletop display can probably only purchase a system with single-point touchscreen (non-integral) or an indirect input device (potentially integral).

### 3.4.1 Integrality and Directness of Input

Thus from an abstract perspective, there seem to be reasons to prefer direct input on the tabletop display because of its advantages for collaboration and physical manipulation skills. Yet it is unclear whether it would be worth using direct input if the IT were non-integral, whether the benefits of directness would make up for the loss of integrality, in comparison to the benefits of superior rotation control offered by an indirect integral IT such as a rotatable mouse. In addition, touchscreen overlays cost 1–2 magnitudes greater than indirect input devices. An evaluation of the tradeoff between integrality and directness is thus of great interest.

A search of the literature did not reveal any comparisons of ITs that differed in both integrality and directness. Some studies compared indirect ITs of different integrality and found the integral devices to be faster: 50% faster for 3DOF object manipulation ( Fitzmaurice, 1996); 35% faster for 3D position ( Balakrishnan et al., 1997); 35% faster for 3D rotation ( Hinckley et al., 1997). Others compared integral ITs of different directness. Zhai and Milgram (1998) and Masliah and Milgram (2000) performed studies of 6DOF position
and orientation. In both cases the free 6DOF tracker (C:D ratio approximately equal, but positional offset and variable rotational offset due to the non-immersive display) was faster than the 6DOF rate control (indirect). Zhai and Milgram (1998, Figure 6, block 5) stated that the time difference was significant but did not give the means, though the more direct input can be estimated to have been approximately 45% faster. Masliah and Milgram (2000, Figure 2) mentioned that the mean completion time for the more direct tracked input was 1.8 seconds faster than for the rate control, but also did not give the means. The mean can be roughly estimated at approximately 7.5 seconds, which would imply 24% faster completion times. Ware and Rose (Ware and Rose, 1999) compared 3D rotation in an immersive 3D environment, using tracked 6DOF input with and without positional offset from the displayed object: the direct input was 45% faster than the offset input. I did not find any research that compared ITs that differed in both integrality and directness. The comparisons are shown in relation to the IT space in Figure 3.2 (p. 35) — a comparison differing in both integrality and directness would be diagonal in the space.

3.4.2 Pilot Study

I designed and performed pilot testing for a study that explored the effects of integrality and directness, comparing an IT that was integral but indirect with an IT that was direct but separated. The tabletop display used during piloting was a 50" (1365x768) plasma display, located in the CoLab at the Department of Mathematics at Simon Fraser University (Figure 3.3(a), p. 36). The dimensions of the display were 61 cm \times 110 cm, with an additional 7.5 cm border. A touch sensitive overlay was installed above the screen, adding a 5 cm plastic border, and the whole unit was embedded in a custom table with an 17.5 cm wooden border.

Because of the edge of the table was 30.5 cm from the screen, and the display surface was 5–6 cm below the level of the table, it was not possible to comfortably reach or see the screen from a seated position (Figure 3.3(b), p. 36). The wooden border, designed to let people work with documents and laptops, meant that the very same interaction was not possible with digital documents: the screen so far away removed the focal work-area where reading and writing typically occurs. Consequently, for the pilot study subjects had to stand at the table, which created discomfort using the mouse, so large textbooks had to be used to raise the level of the mouse by approximately 25 cm. Another problem was that the touch overlay sat approximately 2–3 cm above the surface of the display screen,
CHAPTER 3. TECHNIQUES FOR MANIPULATION ON TABLETOP DISPLAYS

Figure 3.2: Integrality and Directness examined in Related Multi-DOF Research
(a) The Colab tabletop display  
(b) Maximum reach of the author while seated  
(c) Normal position for reaching (hand support)  
(d) Touchscreen: hand support under centre of gravity is not possible

Figure 3.3: Illustration of the problems with the touchscreen tabletop display
resulting in a disturbing parallax effect that made it difficult to aim accurately.

3.4.3 Touchscreen Unusable

The surprising result was that due to the physical configuration, touchscreen input was unusable for the experiment. Using the touchscreen meant leaning over the table, with no way to adequately support oneself with the other hand, resulting in back pain. I tried piloting the experiment on the touchscreen (30 minutes), and experienced strong lower back discomfort — despite 30 sec breaks every 2 minutes — to the extent that the only way I was able to complete the trials was by alternating between right and left hands, and finally by actually sitting on the wooden border of the table. Later my supervisor, Dr. Kirkpatrick, used the touchscreen and confirmed similar discomfort after only 5–10 minutes of use.

One reason why there were problems with the reach is due to the dimensions of the CoLab table. With all the borders around the screen, the touchscreen, and the additional 17.5 cm wooden border, the far side of the screen was 92.5 cm from the edge of the table. But Das and Behara (1998) found that the maximum possible reach by 95% of the population at a desk work surface is 85 cm. Even that maximum value is only for reaching straight across the surface: the reachable space follows an arc based on the user’s arm length, so the depth of a user’s reach falls off even more on either side thus a large portion of the far side of the screen was beyond comfortable reach.

However, the problem would still have been experienced without the added table border, because of the touchscreen technology. Like all current touchscreen technology, only one contact point could register; pressure on the screen in another location, even unintentional, confused the input. Participants therefore had to be careful not to touch the screen other than with their pointing finger. On a normal table it is easy to support oneself with the other arm when reaching (Figure 3.3(c), p. 36). On the single-point touchscreen, users could not support themselves by leaning on the table (Figure 3.3(d)). Even if the table were small enough that the user could sit and not have to reach beyond an arm’s length, the system would still be basically unusable: not being able to rest one’s hands or elbows on the surface would result in unacceptable fatigue, just as was found for direct input on vertical screens. In fact, the fatigue on a tabletop would probably be worse, since with a vertical screen on a desktop computer it is at least possible to pause and rest. On a touchscreen tabletop display that is reachable from a sitting position (i.e., no border), there would be nothing to lean on.
Unfortunately, touchscreen was the only direct input technique available to me. With the touchscreen configuration causing back pain, the direct input condition had to be eliminated from the experiment, and the experimental exploration space was changed slightly. Since a tabletop display equipped with the standard direct input device was unusable, I decided to explore the manipulation that is possible on the other standard input device, the mouse.

3.4.4 Proposed Integral Techniques for Existing Input Devices

At first it would appear that the standard mouse is restricted to traditional mode-based rotation control. However, I wanted to explore whether there was any way that an integral technique could be possible.

I proposed two potentially-integral techniques for integral manipulation, one using the mouse scroll wheel for an additional DOF, and the other a new manipulation technique inspired by the physical properties of paper.

3.4.5 Wheel

In Section 3.1.3 (p. 24) it was seen that alterations on 2D pointer devices could provide a third DOF, such as the Rockin’ Mouse’s custom tilt detection. Only one additional DOF is needed for orientation control.

The reader is almost certainly familiar with the mouse “scroll wheel”, a standard component on today’s PC mice. The scroll wheel is a roller between the left and right mouse buttons, which can be rolled with the index or middle finger while grasping the mouse. It is called a scroll wheel because in windowed GUIs the default action is to scroll the window, up when the wheel is stroked away from the user, and down when stroked towards oneself. However, despite its common name, the scroll wheel need not be restricted to scrolling. In some application it already has also other functions. For example, in Microsoft Office applications holding the CTRL key will cause mouse wheel rotations to increase or decrease the level of zoom on the document.

In a study of mouse that rotates, MacKenzie and Zhai observed that the mouse wheel could potentially be used for a rotational DOF. However, they cautioned that its effectiveness might be limited due to the fact that moving the mouse (with a grasped hand) uses very different muscle groups than stroking the mouse wheel (with a finger). Consequently they categorized it as “2 + 1 DOF” instead of 3 DOF (MacKenzie et al., 1997). The mouse
CHAPTER 3. TECHNIQUES FOR MANIPULATION ON TABLETOP DISPLAYS

wheel was evaluated by Zhai, Smith and Selker (1998) in a study of scrolling techniques in a browser search task. The wheel mouse performed poorly not only compared to a mouse with isometric joystick, but also worse than a normal mouse forced to use the scrollbar. In a follow-up paper based on the same study, Zhai and MacKenzie (1998) explained this unexpected result by the forced trade-off between either gain and resolution. Scroll Wheel rotations are discreet rather than continuous: the wheel jumps from one angle to the next with a click, with slight resistance between clicks to keep it from rotating unintentionally. This significantly reduces the resolution of the rotation control. The Logitech Cordless MouseMan Wheel, for instance, has only 24 steps for a complete rotation of the wheel, corresponding to 15° per click. Furthermore, actions to rotate the wheel are limited to short bursts with each finger stroke. Larger rotations require repeated actions. Thus there is a forced trade-off between either high-gain but low resolution --- moving a lot with little wheel rotation, but each click is a large jump --- or high resolution but low gain --- fine control, but many repeated strokes required. However, the results reported by Zhai et al. (1998) only included task completion time, so it is not possible to know whether participants used the mouse wheel in an integral manner, which is where its performance benefits would arise. It is not known whether such “mixed-resistance mode” mice can be used integrally.

Nevertheless, the Rockin’ Mouse (Section 3.3.1, p. 26) was shown to be integral, despite the fact that one of the DOFs (tilting the mouse, that is, rotating around the y-axis) was very different from the control that it was mapped to (z-position of onscreen object). Since mouse wheel rotation is potentially “closer” to object rotation, differing only in the axis of rotation (around the x-axis for the mouse wheel vs. around the z-axis for the object), it is reasonable to expect that be participants will be able to ‘map’ mouse wheel rotation to tabletop object rotation. It remains to be seen, however, whether the action can even be performed integrally on a mechanical level, despite the different muscle groups used to rotate the mouse wheel, and how well the IT will perform.

3.4.6 Drag

In Section 3.1.3 (p. 20) it was observed that a single 2D input device is unable to control position and orientation simultaneously. However, it turns out that this observation may have been not quite accurate.

I propose a new method for position and orientation, inspired by real paper. The reader may see the inspiration by placing a sheet of paper on a smooth, unobstructed desk, and
moving it with a point such as the back end of a pen (do not use your finger, which makes contact over an area and therefore has rotational force). Moving the paper with the pen located at the centre of the sheet will move it around. With the pen contact position at the edge of the paper, tangential movement will rotate the paper around a point approximately 3/4 of the way across the sheet. As the *drag point* (the pen contact position) is moved in, the centre of rotation will move farther out (beyond the sheet). With the drag point at the centre of the paper, the centre of rotation will be at infinity, i.e. no rotation. The radial component of motion (towards or away from the centre) will always result in translation of the paper. You should find that with a little practice you can usually drag the sheet to a particular location and orientation, albeit with some awkwardness and frustration.

The ability to control both position and orientation with just a 2D point occurs because the friction between the desk surface and the paper provides another force, acting on the paper at a different net point of application any time your pen is not exactly in the centre of the sheet. The direction of the friction force is controlled by you, because it acts opposite to the direction that you drag. Thus if you move in towards or away from the net friction point of application (the centre of the rectangular sheet), there will be no rotation. But if you move tangentially with respect to the point of net friction, the friction will apply a torque and rotate the paper at the same time as it moves! The magnitude of the torque increases the farther your control point (the pen) is from the centre of the paper; with the control point at the centre of the paper there is no torque, and when the drag point is at the corner the torque is maximal.

We can approximate Drag on a computer for use with 2D pointing devices. The question is how to implement it, especially if it involves the centre of rotation moving away to infinity. Looking at the physics helps. The point of the document being dragged will always be translated to follow the pointer. That motion results in friction, which occurs over the entire area of the sheet, but for simplicity let us consider it equivalent to a single frictional force at the centre of the sheet, acting in the opposite direction to the translation. Let us divide that translation into radial and tangential components. The radial component does not contribute any rotation, because the friction is in line with the translation. The tangential component, however, provides a torque that will result in rotation around the drag point, as illustrated in Figure 3.4(b) (p. 41). The magnitude of the tangential component is found by the cross product of the drag \( \vec{d} \) with a vector from the centre to the drag point \( \vec{r} \), where \( |\vec{d} \times \vec{r}| = d.x \times r.y - d.y \times r.x \). For friction of magnitude \( F \) and a drag point at
radius $r$, the magnitude of the rotation is simply $\arctan(F/r^2)$.

(a) Drag force $\vec{d}$, and opposing (b) $\vec{F}_{\text{Tangential}}$ exerts a friction $\vec{F}^f$ torque around drag point $\vec{r}$

(c) The result of the drag

Figure 3.4: Illustration of Drag

However, the magnitude of the friction changes with the location of the drag point: when the drag point is moved in from the edge of the paper, there are also frictional forces applying a torque in the opposite direction. The modeling of the physics is incompletely implemented: instead the effect is approximated by applying a ‘damping’ function to the rotation based on the radius of the drag point. On a computer there is also no reason why we could not make changes to the physics, if desired. For instance, with paper it is frustrating that even at the corner of the sheet a tangential movement cannot just rotate: translation always occurs. With the computer version, a tangential drag at the edge of the object does not have to cause any translation.

It would seem possible to make a physically-inspired interaction that can simultaneously control position and orientation with just a 2D pointer. However, it is not clear whether the IT will actually be effective. It may be possible to perform object manipulation, but it will need to be less frustrating than real paper if it is to be of any use.

3.4.7 Proposed Experiment

Both these new ITs, Wheel and Drag, have potential for integral manipulation of object position and orientation. An empirical study (Figure 3.5, p. 42) was necessary to test whether they are integral. The study also investigated the effects of varying integrality on factors that are important for document manipulation, such as the performance of 3DOF
manipulation actions, and the extent to which the interaction technique caused distraction. The design of the study is given in Chapter 4 and the results are reported in Chapter 5.

Figure 3.5: Proposed experiment in relation to the IT exploration space
Chapter 4

Experimental Design

Many existing techniques for orientation control are not appropriate for interaction on tabletop displays. Other techniques are questionable. In particular, the most prevalent IT for orientation and position control uses a 2-degree of freedom (DOF) mouse to control the 3 DOFs by requiring separate, sequential rotational and positional control. Yet it has been shown that people perceive and manipulate the orientation and position of objects in an integral manner. Traditionally, integral 3-DOF control has not been possible with a mouse.

4.1 Goal

The goal of the user study was to evaluate the effect of integrality on 3-DOF manipulation for document interaction. At the same time, two proposed potentially-integral ITs, Wheel and Drag, were evaluated for integrality and compared to the traditional mode-based method of rotation control.

4.2 How to Evaluate the Techniques

Ideally, a study motivated by document interaction would involve users performing realistic tasks with documents on a tabletop display. However, realistic document interaction would likely be dominated by the cognitive aspects of the task rather than the physical manipulation, making it very hard to distinguish any impact by the manipulation ITs. Since Wheel and Drag had not yet been shown to be integral, or even usable, a simpler first step was needed to test their effectiveness at simple manipulations.
Previous research on Multi-DOF input and integrality used a *docking task* to evaluate and compare ITs in 3-DOF (manipulation in a 2D plane) (Fitzmaurice, 1996) and in 6DOF (manipulation in 3D space) (Zhai et al., 1996; Wang et al., 1998; Masliah and Milgram, 2000). The participant manipulated an object to line it up with a target displayed onscreen. For symmetric objects such as a rectangle, a colour cue was provided to avoid 180° misalignment with the target. By varying the starting positions and orientations of the object and the target, one can reproduce motions of interest, such as common motions expected for documents on tabletop displays.

### 4.2.1 Common Manipulations

Common actions on documents can be condensed from the interactions described in Chapter 2:

**Position** People place one or two active documents in the area of focus. The periphery around the work area serves as visible, quickly accessible ‘storage’ for other documents. Documents may be moved very frequently between the centre and the edge of the desk/table. Documents may be passed to collaborators.

**Orientation** Documents are read at +/- 20° rotation. Notes and sketches are made with the paper rotated 30-45°. Documents stored in the periphery may be oriented radially, facing the readers, or have some other meaningful orientation. Collaborators sitting at adjacent sides of a table will rotate documents 90° to show the documents the partner next to them.

**Position and orientation** Combinations of any of the above can require both translation and rotation. For instance, a document may be moved from one orientation in the focal area, to a different orientation in the periphery, or a document shared with or passed to a collaborator will be rotated as well as translated, with the angle dependent on the collaborators’ respective locations.

Other interesting considerations (these were not examined in this study, however):

- The accuracy of the manipulation. For a document moved towards the side of the table the manipulation may be highly approximate, whereas careful document layout may require great precision. Pilot testing indicated that the level of accuracy (Target
corner radius of 30 pixels vs. 5 pixels) had an effect on performance, and that the Drag method in particular was slowed more than Corner or Wheel. I decided that accuracy should not be an additional variable in this first study. The low accuracy condition better captured the informal manipulations I was attempting to model, interaction that might be expressed in general terms such as “Put this document on the right side of the table”, or “Turn this paper so your partner can read what you’re pointing at”.

- Manipulation by the non-dominant hand.
- Repeated fine adjustments of position and orientation while handwriting

4.2.2 Measuring Effectiveness

Performance and Integrality

The standard measure of effectiveness is response time, how long participants take to complete the docking task. Faster basic actions generally correspond to a large difference in overall effectiveness.

In addition to docking time, measures can be taken on the path used to get there. Of particular interest are measures related to integrality:

*Integraly, introduced by Jacob, Sibert, McFarlane, and Mullen (1994), measures the percentage of the time that all DOF change simultaneously. The path data is separated into regular time segments, and a threshold is chosen. If for all the DOFs the change in a time segment exceeds the threshold, that segment is deemed Euclidean, or integral, otherwise the segment is city-block\(^1\), or separated.*

*Index of Coordination (Zhai et al., 1996), gets at the same notion of Euclidean vs. city-block movement, but in this case it is the ratio between the length of the actual path taken, and shortest possible path to the target cutting across all dimensions.*

*M-metric was introduced by Masliah and Milgram (2000) in an attempt to combine the attributes of the two previous metrics: the temporal aspect of Integrality and the efficiency aspect of the Index of Coordination. Furthermore, the temporal Integrality-like component only counts movement that brings the object closer to the target, not simply any movement

--

\(^{1}\)City-block refers to the way you move from one point to another in a city with streets arranged in a grid: you can’t cut diagonally across city blocks even if it would be the shortest path to your destination. Instead you have to follow the streets, moving in just one dimension (i.e., street direction) at a any given time.
in all DOFs. M-metric does not correspond directly to a physical basis, as the other two do, so it is difficult to describe in more detail here. The reader is referred to the original paper for details, but the basic notion is

\[ M\text{-metric} = SOC \times EFF \]

Where \( SOC \) represents Simultaneity of Control and \( EFF \) represents Efficiency of Control. \( SOC \) is a sum over time segments of how much all DOF moved; it is similar to Integrality but instead of a binary all-DOF-moved sum, SOC is based on what percentage of each DOF’s total movement occurred in the time segment. \( EFF \) is based on the fraction of the length of the optimal path in each DOF, and the length of the actual path, similar to Zhai’s Index of Coordination.

Jacob-Sibert integrality is not strictly a measure of performance. However, it is related, since integral manipulation of multiple DOFs by definition can be faster than manipulating each DOF sequentially. Index of Coordination and M-metric similarly are not directly measures of performance, but they do incorporate a measure of the efficiency of the manipulation in multiple DOFs. These measures will help to understand more about how the ITs were used, particularly whether the Wheel and Drag techniques can properly be called integral.

**Attention and Distraction**

In addition to the actual physical manipulation, a further guideline from Chapter 2 was that for effective interaction with information such as documents on a tabletop display, the user should not be distracted by the technique itself. With paper documents, we do not have to think about the fact that we are manipulating the paper, we just do it.

It is difficult to evaluate the *distraction* that an IT causes. One way is to ask directly, but even the act of asking can change a user’s perspective. For instance, in a within-subjects design where a participant uses multiple ITs sequentially, when should they evaluate the distraction? If you wait until all ITs are finished before asking, the participant’s memory of the first technique will have been affected by the passage of time and the experience of other ITs; but if you ask between each technique, the awareness of a possibility of distraction may affect how they experience the later ITs. It would be extremely difficult to distinguish whether a difference in reported distraction was caused by a genuine difference between ITs, or by the participant expecting to be distracted.
CHAPTER 4. EXPERIMENTAL DESIGN

<table>
<thead>
<tr>
<th>Category</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mental Demand</td>
<td>MD  How much mental and perceptual activity was required (e.g., thinking, deciding, calculating, remembering, looking, searching, etc.)? Was the task easy or demanding, simple or complex, exacting or forgiving?</td>
</tr>
<tr>
<td>Physical Demand</td>
<td>PD  How much physical activity was required (e.g., pushing, pulling, turning, controlling, activating, etc.)? Was the task easy or demanding, slow or brisk, slack or strenuous, restful or laborious?</td>
</tr>
<tr>
<td>Temporal Demand</td>
<td>TD  How much time pressure did you feel due to the rate or pace at which the tasks or task elements occurred? Was the pace slow and leisurely or rapid and frantic?</td>
</tr>
<tr>
<td>Effort</td>
<td>EF  How hard did you have to work (mentally and physically) to accomplish your level of performance?</td>
</tr>
<tr>
<td>Performance</td>
<td>OP  How successful do you think you were in accomplishing the goals of the task set by the experimenter (or yourself)? How satisfied were you with your performance in accomplishing these goals?</td>
</tr>
<tr>
<td>Frustration Level</td>
<td>FR  How insecure, discouraged, irritated, stressed and annoyed versus secure, gratified, content, relaxed and complacent did you feel during the task?</td>
</tr>
</tbody>
</table>

Table 4.1: Categories in the TLX Questionnaire

Related to the amount of focus required by an IT is the concept of workload, which has been explored by the human-factors literature. Of particular interest for my purposes is that ‘mental demand’ was one of six categories of workload incorporated into a general workload questionnaire developed by NASA, Task Load Index (TLX) (Hart and Staveland, 1988). TLX was used to determine the workload in usability studies by asking participants to rate their answer to six different workload scales, shown in Table 4.1. TLX produces a score from 0-100 that is the average of the six categories.

For different applications, the six categories may have varying relative importance. Instead of setting the importance through some arbitrary normative standard such as an expert’s decision, Hart and Staveland had users weight the relative importance by comparing all fifteen pairings of the six categories to produces a weighting for each category, and then a weighted workload. This has the advantage that even if users did not understand identical definitions of the “overall workload”, the weighted workload inherently reflects
overall workload.

A different method for evaluating the same notion of attention and distraction is based on user's perception of the passage of time. For a brief action such as a docking task, it is not possible to monitor the participant's perception of time while they use the IT (e.g., counting aloud while performing the task). Fortunately the principle was been extended to HCI usability evaluation by Czerwinski, Horvitz, and Cutrell (2001), who proposed a measure, relative subjective duration (RSD), adapted from the Ziegarnick effect, an area of 1920’s work-efficiency research. The claim is that the degree to which participants over- or under-estimate the time that passed while using an interface corresponds to the difficulty of the interaction — basically, “time flies when you’re having fun”. RSD is simply a percentage calculated from the ratio of the user’s guess (i.e. subjective duration, SD) and the real duration (RD),

\[
RSD = \frac{SD}{RD}
\]

Czerwinski et al. performed a study on the usability of three websites and the RSD hypothesis was supported: there was a correlation between RSD and the number of mistakes made while performing the required tasks. It was not just that mistakes meant it took longer, but that users overestimated how much longer, and for easier websites they underestimated the time.

It remains to be seen whether TLX and RSD will be useful for our purposes. TLX was developed through thorough research, but it has had limited application to the field of HCI, with non-expert evaluators. RSD was introduced by Czerwinski et al. in 2001 and has not yet been replicated in any conference or journal publication. Yet if the measures perform as intended they will be useful to distinguish between ITs, even if the timing performance is similar.

### 4.3 Experimental Method

#### 4.3.1 Design

A within-subjects design with repeated measures was used. Twelve subjects used all three ITs in sequence, counter-balanced for IT order (two participants for each of the six permutations). For each IT, there were five blocks of twenty trials. Each block contained four trials of each of the five task configurations (Section 4.3.5, p. 54), with the order of trials
randomized within the block. A session consisted of 300 trials, as follows:

3 ITs per participant  X
5 Blocks per IT  X
5 Task Configurations  X
4 Repetitions of each configuration per block

4.3.2 Apparatus

The experiment was run in the Gruvi Lab at Simon Fraser University. The interaction techniques and experiment software were programmed in Java using the Java2D graphics API (1.4.1.01), and were run on a PC running Windows 2000, with an AMD 650 MHz Duron CPU and 3DFX Voodoo 3000 graphics card. The system update and graphics redraw loop cycled every 10-11 ms. Time stamped data revealed that there were update pauses of approximately 100-200 ms every few seconds, apparently caused by Java virtual machine, though these were not generally noticeable in the interaction. The computer was not attached to a network.

The mouse input condition used a standard USB optical mouse (Microsoft Intellimouse).

The tabletop display used in the experiment was a 40" (87 cm x 52 cm) LCD display (1280x768) laid flat on a desk. The video card did not support the wide screen resolution, so it was run at 1024x768 with blank bands on the both sides of the screen. To avoid confusion, strips of white paper were used to cover the blank areas so only the active part of the screen was visible, resulting in screen dimensions of 70 cm x 52 cm. Participants stood at the table, with the display at a height of 89 cm, and the mouse surface elevated on a stiff, heavy box to 102 cm high, with a thin mouse pad fixed on top. One of the taller participant asked for the level to be raised, so for that participant the height was raised an additional 12 cm. The setup is illustrated in Figure 4.3.2 (p. 50).

Paper questionnaires were filled out at a desk in the same small room as the tabletop display. The TLX questionnaire was implemented in Java and filled out on a PC running Linux, located just outside the room where the tabletop display was located, in the Gruvi Lab open space. The clock on the wall was covered for the duration of each session.
Figure 4.1: Physical setup of experiment
4.3.3 Task

The task was a docking task in 3 DOF: (x,y) position and (Rz) orientation, closely based on the task used by Fitzmaurice in Experiment 1 of his PhD dissertation (Fitzmaurice, 1996), which also involved multi-DOF aligning of rectangles.

At the start of each trial, the participant clicked a large button in a popup window with the label “Click here to Start Trial”. When they did, the popup window disappeared, and a rectangle and target appeared (Figure 4.2(b), p. 52). The object to manipulate was a white rectangle 16.2cm wide and 20.9 cm high. These dimensions were 75% of a letter-size sheet of paper, reduced due to the small size of the tabletop display. A three-pixel border provided a clear colour cue, red on the top and left sides, blue on the bottom and right sides. The target consisted of four grey circles corresponding to the target locations for the four corners of the manipulated rectangle, and a red (top left) and blue (bottom right) fill corresponding to the colour cues on the rectangle.

The participant would move and rotate the rectangle to line it up with the target, within the accuracy determined by the radius of the target corner-circles. Whenever a rectangle corner was over its corresponding target corner circle, that circle would be highlighted (Figure 4.2(c), p. 52). Whenever all four corners were lined up at the same time, the whole target fill colour would also highlight, i.e. red/blue would change to orange/cyan (Figure 4.2(d), p. 52), to provide clear feedback when all four corners were lined up. When the rectangle remained line up continuously for 700ms, the trial was completed.

The similarities of this docking task to (Fitzmaurice, 1996) can be seen in the thresholds of acceptance at the target corners, target corners highlighting to indicate alignment, the red/blue cue to avoid 180° misalignment, and the 700ms on-target time to end the trial. Differences include round target corners instead of square (so that the threshold would be equal in all directions), and the “thick” target; that is, Fitzmaurice’s target had a thin border the same dimensions as the rectangle, but for this experiment the target was filled out to the edges of the corner-circles. This was done to maintain the notion of lower accuracy, to avoid presenting a “correct” rectangle with which participants might feel the need to align precisely.
(a) Location of the start trial button superimposed on the start of a trial to illustrate its position with respect to the starting rectangle position

(b) At the start of a TR30 trial (using Corner)

(c) Highlighting of corners

(d) Rectangle lined-up with target

Figure 4.2: Screenshots from the experiment
4.3.4 Interaction Techniques

Corner

For the Corner IT, 5 cm squares were superimposed (in light green -- see Figure 4.2(b), p. 52) in each corner of the rectangle to designate the control points for rotation. When the pointer is pressed, the location of the pointer is tested. If the pointer drag action starts in the central part of the rectangle, it will drag the rectangle; dragging from one of the corners will rotate the rectangle around its centre. For rotation, the orientation of the rectangle is ‘locked’ to the pointer, so that the original drag point stays on the line between the pointer and the centre of the rectangle.

The corner control points are much larger than in many applications. The Welford formulation Fitts’ Law Index of Difficulty (ID) quantifies the difficulty of clicking in a target area (MacKenzie and Buxton, 1992):

\[ ID = \log_2 \left( \frac{\text{mean movement distance}}{\text{corner size}} + 1 \right) \]

Moving from the centre of the rectangle (distance 13.2 cm) to click in the 7 cm (diagonal) corner represents an ID of 1.5, whereas for the same size rectangle, Corel PhotoPaint (version 8.232) has only a 1.2 cm (diagonal) corner control-point, an ID of 3.6.

Wheel

The Wheel method is described in Section 3.4.5. For the Wheel IT, each click of the wheel resulted in a 5° rotation. For typical wheel ‘strokes’, this would require at least one stroke for 30° rotation, and at least three strokes for 90°. The centre of rotation was the location of the mouse pointer. Counter-Clockwise (CCW) rotation occurred when the wheel was stroked away from the user.

In early testing some users found it difficult to rotate the wheel while holding down the mouse button. With Wheel, holding the button is not necessary: unlike Corner and Drag, there is no need to let go of the object and quickly re-acquire it again in a different location (e.g., to make an adjustment). Wheel was free for users to not have to hold the mouse button: users would press once to pick up the rectangle for moving and rotating, and press again to let go if desired. Thus, for moving the rectangle, with Wheel, the user was actually performing a pointing action rather than a drag.
CHAPTER 4. EXPERIMENTAL DESIGN

Drag

The damping function chosen for the Drag method (Section 3.4.6), was \((|\vec{r}|/R)^2\), where \(|\vec{r}|/R\) is the proportion of the distance to the edge of the rectangle, i.e., the distance from the centre of the rectangle to the drag point \((|\vec{r}|)\), divided by the distance from the centre of the rectangle to edge of the rectangle for a line passing through the drag point \((R)\). Thus the final rotation can be expressed as

\[
Rotation(\text{radians}) = \left(\frac{|\vec{r}|}{R}\right)^2 \cdot \arctan\left(\frac{d \times |\vec{r}|}{r^2}\right)
\]

Early testing showed that instead of being limited to a physical simulation, some changes were desirable. Since translating with rotation is a common task, it was not desirable for users to have to aim precisely in the centre of the rectangle to achieve it, but instead have some leeway. In addition, it was desirable that at the edge of the rectangle a tangential motion would result in just rotation without any translation, and that the region of this action should not be just ‘at the edge’ (with no area), but within a certain distance of the edge. Consequently, translation-only manipulation was initiated by a drag point in the inner 15% of the rectangle and rotation-only at the outer 10% of the rectangle. The damper function of the rotation was thus adjusted to be a quadratic between 0.15 and 0.9, as shown in Figure 4.3.4 (p. 55). No visual indicator of these regions was provided since for document manipulation they would obscure the content of the document.

4.3.5 Task Configurations

Five different task configurations were used to represent the common manipulations discussed in Section 4.2.1, based on three rotation angles (0°, 30° CCW, 90° CCW) and two translation magnitudes (0 cm, 37 cm), shown in Table 4.2. The 0° x 0 cm configuration was not used.

It would have been desirable to reproduce the conditions of (Wang et al., 1998) to compare the ITS with natural manipulation of real physical objects; however, their task configurations (2 rotations (22.5°, 45°) x 3 translation magnitudes (2 cm, 10 cm, 20 cm) did not match well with the document manipulation I was trying to approximate, and more importantly would not have allowed a contrast between integral tasks and separated tasks (manipulating just one mode, either position or orientation).

Many additional variations would also have been interesting, such as
CHAPTER 4. EXPERIMENTAL DESIGN

CHAPTER 3. EXPERIMENTAL DESIGN

Calculation of Damper Function

Figure 4.3: Altered damperFunction (including central translate-only radius)

<table>
<thead>
<tr>
<th>Task</th>
<th>Starting Location (x,y) cm</th>
<th>(Rz) from vertical</th>
<th>Target Location (x,y) cm</th>
<th>(Rz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>“T”</td>
<td>(0,0)</td>
<td>0°</td>
<td>(33.8,14.9)</td>
<td>0°</td>
</tr>
<tr>
<td>“R30”</td>
<td>(0,0)</td>
<td>0°</td>
<td>(0,0)</td>
<td>30° CCW</td>
</tr>
<tr>
<td>“R90”</td>
<td>(0,0)</td>
<td>0°</td>
<td>(0,0)</td>
<td>90° CCW</td>
</tr>
<tr>
<td>“TR30”</td>
<td>(0,0)</td>
<td>0°</td>
<td>(33.8,14.9)</td>
<td>30° CCW</td>
</tr>
<tr>
<td>“TR90”</td>
<td>(0,0)</td>
<td>0°</td>
<td>(33.8,14.9)</td>
<td>90° CCW</td>
</tr>
</tbody>
</table>

Table 4.2: Task configurations used in the experiment
CHAPTER 4. EXPERIMENTAL DESIGN

- Direction of rotation (2)
- Magnitude of rotation (continuous)
- Magnitude of translation (continuous)
- Direction of translation (continuous)
- Starting location and orientation of rectangle (continuous)

However, it was necessary to simplify the design to keep the number of conditions from exploding. Even simply having both directions of rotation (CCW and CW) and translation (in the opposite direction than the current one) would increase the number of task configurations from five to 27. I felt that the five conditions chosen would capture some of the key actions, without adding unnecessary complexity.

For all conditions, the required accuracy represented by the radius of the target corners was set at 30 pixels (2 cm), a relatively high value (max. 25% of rectangle width or 29% of a 30° rotation). Thus the results from the experiment should not be extrapolated to applications requiring high precision.

4.3.6 Subjects

Twelve participants (9 male and 3 female, 11 right-handed and 1 left-handed, and aged 21–33) were recruited by sending an invitation to the Gruvi Lab and the Computer Science graduate students’ email lists. All were regular computer users (more than 14 hours per week). They were informed that participation would take one hour. Participants were not paid, but they were offered cookies at the end of the study, while supplies lasted.

4.3.7 Procedure

At the start of the session participants were asked to remove their watch if they had one, and were shown the questionnaires they would answer after each IT (Subjective Duration and TLX). Participants were assigned a user number ending in a digit (1–6) indicating the IT ordering they would be using, shown in Table 4.3.

Before using each IT, participants were shown how the IT worked, and were given a chance to practice. All participants were required to perform at least ten practice trials (2 per Task Configuration), and could then perform more until they felt comfortable to start.
Timing for the real duration began when the first block began, and continued until the end of the fifth block. Participants were requested to not stop during a trial. Between trials, a popup window would appear with the label, “Begin next trial”. When the block of 20 trials were completed, the screen would indicate that it was a chance to rest, along with information about how many blocks had been completed. After 30 seconds a message would appear indicating that it was time to start the next block. Participants could skip to the end of the rest break at any time by pressing an onscreen button. They were not told how long the breaks were, in order to avoid external timing information that might alter the duration questionnaire. After five blocks were completed, participants would immediately fill out the Subjective Duration and TLX questionnaires.

The duration of each IT condition was approximately 10-15 minutes, consisting of a few minutes for practice, 7-10 minutes to complete the five blocks, and a few minutes to answer the TLX questionnaire. After all three ITs and corresponding questionnaires were completed, subjects were given the TLX weighting, the preference ranking, and general feedback questionnaires to fill out. Total session time was approximately one hour.

4.3.8 Dependent Variables

Six of the dependent variables presented in Section 4.2.2 were used: four objective measures (time, integrality, M-metric, and real duration), and two subjective measures relating to attention and workload (subjective duration and TLX). Additional feedback was sought from participants in the form of a preference ranking, and freeform comments about each IT.

<table>
<thead>
<tr>
<th>Last digit of user #</th>
<th>IT Ordering</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>CDW</td>
</tr>
<tr>
<td>2</td>
<td>CWD</td>
</tr>
<tr>
<td>3</td>
<td>DCW</td>
</tr>
<tr>
<td>4</td>
<td>DWC</td>
</tr>
<tr>
<td>5</td>
<td>WCD</td>
</tr>
<tr>
<td>6</td>
<td>WDC</td>
</tr>
</tbody>
</table>

Table 4.3: User # designation of IT Ordering
CHAPTER 4. EXPERIMENTAL DESIGN

The primary measure, dockTime, was calculated as follows:

\[ \text{trialCompletionTime} - \text{objectAcquisitionTime}^2 - 700\text{ms} \]

The 700 ms subtraction is necessary because the trial ends 700 ms after the rectangle was successfully lined up.

Supplementary information related to performance was also recorded in the event that it might provide insights into the manipulation. These were:

- Acquisition Time (from the time the object appears until the first mouse press on the object)
- The number of mouse clicks
- The number of target re-entries (MacKenzie et al., 2001), with counts for two variations: one with target referring to “any individual cor up”, and another referring to “all four corners lined up simultaneously”.

Integrality and M-metric were calculated on object path data after the trial, because M-metric requires post-completion information. The length of the segments was set at 60 ms, a large enough window to capture some of the integrality of Wheel despite the bursty scroll wheel input.

RSD and TLX questionnaires were administered immediately following the end of the last block for each IT. For RSD, subjects were asked

In your best estimation, how much time has passed since the start of this interaction technique? (Note: don’t include the practice time before the start. Consider the “start” to be after the practice trials finished)

_____ minute(s) _____ second(s)

The TLX questionnaire was administered after the RSD question. Participants answered the questionnaire on a nearby computer (not the tabletop display), by adjusting six sliders of length 100 marked with major and minor ticks (every 10 and 5 points, respectively), with category title and the TLX category description next to it, as shown in Figure 4.4 (p. 60). Following the TLX model, no numbers were shown, but “low” and “high” at the ends (or in

\(^2\)i.e. first mouse press on the object
CHAPTER 4. EXPERIMENTAL DESIGN

the case of OP, “good” and “poor”). All sliders began at the midpoint with a red text label saying “Unanswered”; the label disappeared when any action was registered on the slider. The “I’ve finished the questionnaire” button only became active when all six questions were no longer “unanswered”.

The TLX weighting questionnaire was administered after all three ITSs and questionnaires were completed. The fifteen contrasts were answered by making check marks on paper. At the same time, participants were given paper questionnaires for the preference ranking and general feedback. No time guidelines were given for filling out the questionnaires, but subjects were allowed to look at their watches again, so any who did would have known that approximately 40–45 minutes had passed, leaving approximately 15–20 minutes for the questionnaires.

In addition to all the dependent variables, raw data was taken each time the system updated to keep a record of mouse (position and button press) and object (position and orientation) information.

4.3.9 Hypotheses

H1: The Wheel and Drag methods will both be shown to be integral by the Integrality and M-metric measures.

The measures will never indicate more than partial integrality because people do not manipulate in a perfectly simultaneous manner. Even for the experiment by Wang et al. in which participants moved a wooden block by hand, we can infer from the data that for the 20 cm translation, integrality was approximately 70% for 45°, and 56% for 22.5° rotation (Wang et al., 1998, Figure 2).

Furthermore, the integrality measures for Wheel will be lower than expected for a 3-DOF device because of the bursty nature of the mouse wheel. Bursty changes in just one DOF will result in more of a city block path; nevertheless, it is hypothesized that even during the wheel bursts, translation will still be occurring.

The Integrality of Drag will be higher than Wheel, because it is easy to have all 3 DOF changing at the same time. It is harder to predict for the M-metric, because it incorporates a component of efficiency, and it is likely that some of the Integral movement with Drag may be inefficient — i.e., simultaneous but not necessarily moving closer to the target.
Figure 4.4: Screenshot of TLX questionnaire
CHAPTER 4. EXPERIMENTAL DESIGN

H2: for integral ITs (Wheel and Drag): \( \text{time}(T) + \text{time}(R_z) > \text{time}(TR_z) \)
for separable IT (Corner): \( \text{time}(T) + \text{time}(R_z) < \text{time}(TR_z) \)

That is, for integral ITs the translate-and-rotate task will take less time than the sum of two separate translate and rotate tasks of the same magnitude. This hypothesis is implicit in the notion of integral input, that translation and rotation will overlap to some extent, which will save time. For the separated input, there is no overlap so the time for the translate-and-rotate task will be at least as large as the sum of the separate tasks; however, because the separated task also requires an additional amount of time for at least one mode switch (even if the mode switch is quick), the time for the combined action will actually be greater than the sum of the two separate tasks. It is interesting to note that Wheel may score poorly on integrality measures, due to the city-block path associated with bursty mouse wheel strokes, yet still exhibit the simultaneity expected from integral control.

For Corner, some extra time for translating and rotating (TR) comes from moving between the body of the rectangle and the corner. However, this time should not be very large; if the TR time is much greater than T + R time, it could be inferred that the separate control creates difficulties in aligning, i.e. having to readjust more than once would result in more mode switches wasting time. It would be surprising if large time-wasting behaviour were seen for such a target with such low accuracy requirements: it was my expectation that for Corner, \( \text{time}(TR_z) \) would be approximately the same as \( \text{time}(T) + \text{time}(R_z) \)

The logical consequences of these notions of integral and separated input have implications on the time performance, expressed in the next two hypotheses.

H3: Corner will be fastest IT in the R30/R90 conditions

In general, separated ITs should be best for separated tasks because there is no risk of wasted manipulation. For this study, however, Wheel may be as fast as Corner for the T condition, because it is easy to separate the translation action from the rotation action. Drag also has a translate-only action, but since the target is smaller and — significantly — does not have a visible boundary, it is expected that errors will occur in which users intended to just translate but end up rotating by mistake. For the R conditions, both Drag and Wheel have no way of ‘turning off’ translation, which should result in some wasted action.

H4: Drag and Wheel will be the faster ITs in the TR30/TR90 conditions
If Wheel and Drag are used integrally (Hypothesis H1), the time saved by overlapping the 3-DOF should make them perform fastest.

**H5:** RSD and TLX will be lowest (best) for Drag

In my own use of the ITs, I found that Drag has a sensation of being ‘informal’ and feels the fastest.
Chapter 5

Experimental Results

5.1 Results and Discussion

5.1.1 Preparation for Analysis

Validating the Data

The minimum time was an R30 task configuration trial with 10 ms dockTime. This is very suspicious because 10 ms is just one system update. Examination of the raw data showed a 321 ms pause just before the single update that completed the trial. Because the system was updating regularly every 10 ms, except for that pause, it almost certainly indicated a Java system-busy gap during which manipulation was still happening, so that trial was discarded.

A number of other trials (21 trials in the R30 task configuration) were as low as 50-90 ms. However, I decided not to discarded these trials because they were not isolated from the rest of the data, as the 10 ms trial was: there are least two trials every update increment (i.e. every 10 ms) from 50 ms upwards. Furthermore, 19 of the 21 trials were performed by the two participants with the fastest overall performance at the R30 configuration.

In addition to the low outliers, there are a large number of high outliers. Unfortunately, there is no way to be sure if a high outlier was due to an experimental error (e.g., system pauses, disrupted trial) or simply a participant having difficulty completing the task. Higher numbers of corner re-entries and mouse clicks should indicate poorly executed manipulation, but low values for those measures do not necessarily imply that the length of the trial was not due to poor execution.
Learning

Figure 5.1 (p. 64) shows participants’ mean completion times for each block, with the three ITs. I estimated from this graph that the first block should be eliminated due to the large change in mean completion time for Drag and Wheel between the first and second blocks. It also appeared that the trend did not seem consistent between the second block and the remaining blocks for all ITs, so I decided to approach it conservatively and eliminate the second block also. Thus all analysis in the remainder of this chapter reflects blocks 3-5 (60 trials per IT per participant, with 12 repetitions per condition).

![Figure 5.1: Mean Completion Times by Block for each IT](image)

Transorms on the Data

As is typical with response time data, quantile-quantile plots indicated that times had a lognormal distribution. All analysis of variance of response times was done on the log of
5.1.2 Main Results

Figure 5.2 (p. 65) shows a boxplot\(^1\) of the dockTime for the three ITs (clustered) in each task configuration.

---

\(^1\)For readers unfamiliar with boxplots, they are simply a useful way of summarizing the data. The centre line indicates the median, and the box shows the two middle quartiles, i.e. the middle 50% of the data. The whiskers indicated the farthest data point within 1.5 times the length of the interquartile range (IQR) from the median. Points beyond 1.5 x IQR are marked with "o", and those beyond 3 x IQR are marked with "*".
CHAPTER 5. EXPERIMENTAL RESULTS

<table>
<thead>
<tr>
<th>IT</th>
<th>Task Configuration</th>
<th>N</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Median</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corner</td>
<td>T</td>
<td>144</td>
<td>140</td>
<td>2,013</td>
<td>543.2</td>
<td>281.5</td>
<td>481</td>
</tr>
<tr>
<td></td>
<td>R30</td>
<td>144</td>
<td>50</td>
<td>1,452</td>
<td>230.9</td>
<td>206.8</td>
<td>170</td>
</tr>
<tr>
<td></td>
<td>R90</td>
<td>144</td>
<td>110</td>
<td>2,493</td>
<td>511.2</td>
<td>362.6</td>
<td>430</td>
</tr>
<tr>
<td></td>
<td>TR30</td>
<td>144</td>
<td>951</td>
<td>7,481</td>
<td>2083.3</td>
<td>1044.1</td>
<td>1,757</td>
</tr>
<tr>
<td></td>
<td>TR90</td>
<td>144</td>
<td>871</td>
<td>8,602</td>
<td>2095.2</td>
<td>829.4</td>
<td>1,843</td>
</tr>
<tr>
<td>Drag</td>
<td>T</td>
<td>144</td>
<td>300</td>
<td>4,727</td>
<td>1010.5</td>
<td>810.7</td>
<td>661</td>
</tr>
<tr>
<td></td>
<td>R30</td>
<td>144</td>
<td>120</td>
<td>4,907</td>
<td>799.3</td>
<td>748.5</td>
<td>516</td>
</tr>
<tr>
<td></td>
<td>R90</td>
<td>144</td>
<td>291</td>
<td>5,397</td>
<td>1651.7</td>
<td>973.3</td>
<td>1,457</td>
</tr>
<tr>
<td></td>
<td>TR30</td>
<td>144</td>
<td>681</td>
<td>5,729</td>
<td>2091.8</td>
<td>877.0</td>
<td>1,923</td>
</tr>
<tr>
<td></td>
<td>TR90</td>
<td>144</td>
<td>621</td>
<td>7,551</td>
<td>2467.8</td>
<td>1184.8</td>
<td>2,214</td>
</tr>
<tr>
<td>Wheel</td>
<td>T</td>
<td>144</td>
<td>170</td>
<td>3,936</td>
<td>757.0</td>
<td>528.2</td>
<td>611</td>
</tr>
<tr>
<td></td>
<td>R30</td>
<td>144</td>
<td>230</td>
<td>5,918</td>
<td>953.0</td>
<td>635.5</td>
<td>796</td>
</tr>
<tr>
<td></td>
<td>R90</td>
<td>144</td>
<td>701</td>
<td>4,847</td>
<td>2010.7</td>
<td>629.2</td>
<td>1,878</td>
</tr>
<tr>
<td></td>
<td>TR30</td>
<td>144</td>
<td>501</td>
<td>5,197</td>
<td>1483.8</td>
<td>777.2</td>
<td>1,257</td>
</tr>
<tr>
<td></td>
<td>TR90</td>
<td>144</td>
<td>871</td>
<td>4,376</td>
<td>2201.0</td>
<td>648.8</td>
<td>2,133</td>
</tr>
</tbody>
</table>

Table 5.1: Table of results (Min, Max, Mean, and Median dockTimes all in milliseconds)

Separated Tasks: H3

For the rotation tasks, analysis of variance on the log transformed dockTime showed a significant effect for interaction technique ($F(1, 11) = 177, p < .001$) (R30) and ($F(1, 11) = 308, p < .001$) (R90). Post-hoc contrasts showed that Corner was fastest\(^2\), so H3 is confirmed. This is in contrast to the slow rotation of Drag, which is likely caused by its inseparability. That is, Drag can be as fast as Corner — the fastest Drag rotation times are faster than the typical Corner trial, so it is not technically limited in speed. At a gesture level Corner should be slightly faster because the pointer can travel the shortest path (e.g., a straight line to the target corner), whereas Drag has to follow the circle. Yet Drag can be almost as fast as Corner, if that circular path is well-followed. However, because Drag’s rotation cannot be separated from translation, slight errors in the path while rotating can move the rectangle off-target. It is likely that participants moved more slowly with Drag to avoid having to recover from a time-wasting error.

\(^2\) Post-hoc contrasts showed a significant difference for Corner vs. Drag for R30 ($F(1, 11) = 117.46, p < .001$) and R90 ($F(1, 11) = 92.822, p < .001$), and Corner vs. Wheel for R30 ($F(1, 11) = 177.536, p < .001$) and R90 ($F(1, 11) = 308.585, p < .001$).
Wheel was faster than Drag in the R30 condition, but the slowest in R90. In both conditions it was much slower than Corner. The Wheel rotational DOF is largely separable from the translational DOF, so participants are not likely to slow their rotation for fear of wasting time recovering from an unintentional translation (as occurred with Drag). Instead, the very slow Wheel times demonstrate the gain / resolution problem with the mouse wheel. The slow rotation was caused by the 5° increment, which I believed was a reasonable angle when chosen. Larger increments could be used, but this would reduce the angles that could be produced. Furthermore, even with 15° increments, the six wheel clicks required to rotate 90° would be the same number as were required in the experiment to rotate 30° which still took almost twice as long as rotating 90° with Corner. It is possible that the mouse wheel could be coupled with a multiplier, e.g., pressing down on the mouse wheel button, or holding a key with the other hand, to make the click increments temporarily larger; however, such interaction seems to hold little promise. The performance of the mouse wheel strongly indicates that it is not an effective DOF for rotation. Other indirect input may provide better rotation control, such as rotatable mice or 3-DOF tablet pucks.

For the translation task (T), analysis of variance showed a significant effect for interaction technique \((F(1,11) = 13.1, p = .004)\). As expected, Corner was fastest and Drag was slowest. It is surprising, however that Wheel was slower than Corner. Previous studies had shown that mouse dragging was 36% slower than pointing (MacKenzie et al., 1991), yet Corner (dragging) was on average 46% faster than Wheel (pointing). There is no clear explanation for the difference, except to note that across all five task configurations there were 56% more target re-entries (i.e. all four corners) for Wheel (average 0.25 per trial) than for Corner (0.16 per trial). Analysis of variance showed a significant effect for target re-entries for interaction technique \((F(1,11) = 6.8, p = .024)\) and specifically for the post hoc contrast between Corner and Wheel \((F(1,11) = 6.793, p = .024)\). The reason is unclear, but it suggests that with Wheel participants were more likely to overshoot the target and then have to backtrack, resulting in slower dockTimes.

---

3 Post hoc contrasts showed a significant difference for Drag vs. Wheel for R30 \((F(1,11) = 7.769, p = .018)\) and R90 \((F(1,11) = 7.562, p = .019)\).

4 This illustrates the distinction I make between an IT being separated and separable. The Wheel IT is largely separable in that it possible to spin the mouse wheel separately from moving the mouse, and to move the mouse separately from moving the wheel. This does not necessarily mean it is separated, that is to say that it is not possible to manipulate the DOFs integrally.

5 Post hoc contrasts showed a significant difference for Corner vs. Drag \((F(1,11) = 21.027, p < .001)\) and Corner vs. Wheel \((F(1,11) = 13.107, p = .004)\), though not for Drag vs. Wheel \((F(1,11) = 1.945, p = .191)\).
CHAPTER 5. EXPERIMENTAL RESULTS

**Integrated Tasks: H4**

Analysis of variance showed a significant effect for interaction technique in TR30 ($F(1, 11) = 18.4, p = .001$). Post hoc contrasts showed that Wheel was the fastest at TR30 (Corner vs. Wheel ($F(1, 11) = 18.410, p = .001$) and Drag vs. Wheel ($F(1, 11) = 77.108, p < .001$), but no significant effect for Corner vs. Drag ($F(1, 11) = .067, p = .80$)). Analysis of variance did not show a significant effect for interaction technique in TR90, nor did post hoc contrasts. According to H4, Wheel and Drag were hypothesized to be faster than Corner in both conditions. The integral ITs performed much more poorly than expected. It is particularly surprising for Wheel, since in pilot testing by an experienced user it was consistently the fastest, even in TR90. Its poor performance brings into question whether Wheel was even an integral technique; to interpret these results, the integrality analysis (H1 and H2) will be helpful.

**Integrality: H1 and H2**

H1 stated that Drag and Wheel would be integral, with at least partial integrality on the measures. Figure 5.3 (p. 69) shows that they were indeed used integrally, with average Jacob-Sibert Integrality (in TR30 and TR90) of 50% for Drag and 20% for Wheel. The value for Wheel in particular is very low compared to the average 70% Integrality of the Rockin' Mouse (Balakrishnan et al., 1997), and the implied 56–70% Integrality of manipulating a wooden cube (Wang et al., 1998). The average M-metric values (Drag=.24, Wheel=.10) were similar to Mashiah and Milgram’s (2000) findings for 6DOF position and orientation control, which had M-metric values of approximately 0.13.

As expected, the low integrality of Wheel is partially caused by bursty scroll wheel input. Three scroll wheel bursts can be seen in Figure 5.4 (p. 69), which shows the change in each DOF during a TR90 trial. There are a few interesting observations that can be made about the interaction in this trial. First, notice that during wheel bursts the rate of mouse movement decreases. It is particularly fascinating that the y-direction control is more affected than the x-direction, even moving slightly in the wrong direction on the first and third bursts. This is due to the direction that the finger moves to stroke the wheel --- along the y-axis. Yet, even though the mouse movement decreases, it does not stop entirely. Simultaneous manipulation still does occurs, which explains the average 17% Jacob-Sibert Integrality for this trial.
Figure 5.3: Measures of integrality for each IT

Figure 5.4: Burstiness of Scroll Wheel During a TR90 Trial
Perhaps more important is the structure of the 'timeline', that the rotation was chronologically contained within the translation. That matches the behaviour Wang et al. (1998) observed for object transport, that rotation started after transport, and ended before the transport was completed. This suggests that even though the bursty scroll wheel results in a city-block path, the structure of Wheel was similar to fully integral multi-DOF manipulation.

The non-city-block approach to evaluating integrality that I proposed in H2 is based on this idea, that time-saving DOF overlap can be contained in a motion even if it is not continuously simultaneous. The summed \((T + R_z)\) time was calculated by adding the means of the individual configurations and using the sum of the variances to compute 95% confidence intervals (Table 5.1.2), and the result is illustrated in Figure 5.5 (p. 71). Corner did not take “approximately the same” time for the combined action as the two separate tasks, as was hypothesized: instead, it took twice as long. This indicates that having to manipulate the position and orientation separately introduced more difficulty than just having to switch modes once. For instance, it may have required numerous re-adjustments to line the rectangle up. The same was true of Drag in the TR30 condition. However, the ratios reveal that Drag did exhibit a small amount of time-saving overlap in the TR90 condition. More significantly, in both TR30 and TR90 Wheel was used in the expected manner — the DOF overlapping chronologically — despite the low Integrality resulting from mouse wheel burstiness. In fact, the ranking of the \(TR_z/(T + R_z)\) ratio matches my predictions for the performance ranking: Wheel (best), Drag, Corner (worst). Thus the overlapping-integrality of the manipulation did agree with expectations, but wheel and drag were simply too slow for the overlap to make up the difference with Corner for TR tasks.

In light of the fact that Drag and Wheel exhibited time-saving overlap for TR, despite having low Integrality (Wheel) and M-metric values (Drag and Wheel), I wanted to see whether those integrality measures are at all related with the performance. Figure 5.6 (p. 73) shows that there is indeed a trend towards an relationship between Integrality and dockTime, confirming that the faster trials generally had higher integrality. The low r-squared values indicate that the trend lines do not explain most of the variability; however, visual inspection confirms that the best-fit lines adequately describe the general trend of the data. The trend is more prominent for Wheel than for Drag, which is expected due to the potentially inefficient
CHAPTER 5. EXPERIMENTAL RESULTS

<table>
<thead>
<tr>
<th>Configuration</th>
<th>Corner</th>
<th>Drag</th>
<th>Wheel</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>543</td>
<td>1010</td>
<td>757</td>
</tr>
<tr>
<td>95% CI</td>
<td>46</td>
<td>104</td>
<td>143</td>
</tr>
<tr>
<td>R30</td>
<td>231</td>
<td>799</td>
<td>953</td>
</tr>
<tr>
<td>95% CI</td>
<td>122</td>
<td>59</td>
<td>127</td>
</tr>
<tr>
<td>R90</td>
<td>511</td>
<td>1652</td>
<td>2011</td>
</tr>
<tr>
<td>95% CI</td>
<td>86</td>
<td>159</td>
<td>135</td>
</tr>
<tr>
<td>TR30</td>
<td>2083</td>
<td>2092</td>
<td>1484</td>
</tr>
<tr>
<td>95% CI</td>
<td>34</td>
<td>103</td>
<td>193</td>
</tr>
<tr>
<td>(T + R30)</td>
<td>774</td>
<td>1810</td>
<td>1710</td>
</tr>
<tr>
<td>95% CI</td>
<td>178</td>
<td>163</td>
<td>270</td>
</tr>
<tr>
<td>TR90</td>
<td>2095</td>
<td>2468</td>
<td>2230</td>
</tr>
<tr>
<td>95% CI</td>
<td>122</td>
<td>171</td>
<td>106</td>
</tr>
<tr>
<td>(T + R90)</td>
<td>1054</td>
<td>2662</td>
<td>2768</td>
</tr>
<tr>
<td>95% CI</td>
<td>132</td>
<td>263</td>
<td>279</td>
</tr>
<tr>
<td>TR30/(T+R30)</td>
<td>2.69</td>
<td>1.16</td>
<td>0.87</td>
</tr>
<tr>
<td>TR90/(T+R90)</td>
<td>1.99</td>
<td>0.93</td>
<td>0.80</td>
</tr>
</tbody>
</table>

Table 5.2: Calculation of summed ($T + R_x$) time with 95% CI for Corner, Drag and Wheel. The ratios between the means time for the TR task and the summed means of the T and R tasks is given, to see if there was time-saving overlap ($ratio < 1$), the same time ($ratio = 1$), or extra time waste ($ratio > 1$)

Figure 5.5: Comparison of Time Saved Due to Integral DOF Overlap
Integral movements of Drag\textsuperscript{6}. The M-metric, shown in Figure 5.7 (p. 74), exhibits a similar trend, but the data is less well described (lower r-squared); unfortunately, with the lack of clear physical basis it is hard to say why M-metric is less tightly linked to performance.

**Attention and Workload: H5**

Figure 5.8(a) (p. 75) shows the real times and subjective estimates for each IT, with the corresponding RSD. Analysis of the variance was non-significant for interaction technique. However, Figure 5.9 (p. 75) reveals that variance is largely due to individuals’ tendency to over- or under-estimate. Each individual’s three values RSD were grouped relatively closely together.

This problem was not discussed by (Czerwinski et al., 2001). I am not interested in individuals’ SD bias: the importance lies in each individual’s relative ranking. Figure 5.9(b) (p. 75) suggests a pattern for RSD, that can be drawn out by converting the RSD for each user into and RSD-rank (best = 1, mid = 2, worst = 3), the results of which are shown in Figure 5.10 (p. 76). This seems to offer support for hypothesis H5, that Drag would have lower RSD.

In converting to a ranking, information is lost: the relative magnitudes of a user’s three RSD scores. I propose a method for removing the individual’s bias while maintaining the relative proportions, normalized RSD (N-RSD). Recall that

$$RSD = \frac{\text{subjective duration}}{\text{real duration}}$$

N-RSD is calculated like RSD, but a normalized subjective duration (N-SD) is used in place of the original subjective duration (SD), calculated as follows:

$$N-SD = SD \times \frac{\text{average}(3 \text{ real durations})}{\text{average}(\text{participant’s 3 SD guesses})}$$

The idea of normalizing by each user’s average guess come from Figure 5.9 (p. 75), which shows that each user’s three guesses tend to be clustered together (Fig. 5.9(b)), and that each user’s average of three guesses is highly variable. Normalizing by the user’s average guess clusters the user’s three N-SD values around the average RD, while maintaining the relative proportions of the guesses. Since the N-SD values are generally clustered around the

---

\textsuperscript{6}i.e. Drag may have high Jacob-Sibert Integrality without the movement being effective.
CHAPTER 5. EXPERIMENTAL RESULTS

Figure 5.6: dockTime vs. Integrality for TR task conditions
CHAPTER 5. EXPERIMENTAL RESULTS

Figure 5.7: dockTime vs. M-metric for TR task conditions
CHAPTER 5. EXPERIMENTAL RESULTS

Figure 5.8: Durations (real and subjective) and RSD for each IT. (Recall that lower RSD is better)

Figure 5.9: Individual time bias is the cause of high RSD variance
Figure 5.10: Rankings of ITs within each user’s RSD. For each user, the ITs were ranked 1, 2, 3. The results for each rank for a stacked bar. The vertical length (ticks) of each colour block indicates the number of users for whom the IT corresponding to that colour had that rank.

Figure 5.11: Normalized RSD for each IT
average real duration (RD), the resulting N-RSD values are not all skewed by the individual’s time estimation bias.\(^7\)

Figure 5.11 (p. 76) shows the N-RSD for each IT. Analysis of variance was non-significant for interaction technique ($F(1, 11) = .504, p = .491$). A post-hoc contrast for Corner vs. Drag approached significance ($F(1, 11) = 4.749, p = .052$). Overall, RSD was too weak to support H5, but the trend is in agreement, rating Drag best.

A final note of interest about RSD is that when participants estimated the duration, they avoided precision, leaving the seconds blank or using ‘round’ numbers of seconds such as 15/30/45. The method for collecting the duration might be improved if the fill-in-the-blank were replaced by a slider, which would elicit more precise estimates without the participants having to think in more precise terms.

Figure 5.12 (p. 77) shows the average TLX (Avg._WL) score and weighted workload (WWL), for each IT. Analysis of variance showed a significant effect for interaction technique for Avg._WL ($F(1, 11) = 8.484, p = 0.014$) and WWL ($F(1, 11) = 6.947, p = 0.023$). Corner was best\(^8\), a result that contradicts H5, and disagrees with the suggested RSD results.

---

\(^7\)A thorough analysis of the cognitive processes would be needed to verify if this is a correct model. For instance, it could be that the first time estimate sets the time bias, and subsequent guesses are not arbitrary but are formed by adding to or subtracting from the first guess — if this were the case, N-RSD might need to be calculated by normalization on only the first guess, and subsequent N-RSD values might be found by adding the difference between a guess and the first guess, rather than by multiplicative normalization.

\(^8\)Post hoc contrasts showed a significant difference for Corner vs. Drag ($F(1, 11) = 61.995, p = .001$), Corner vs. Wheel ($F(1, 11) = 8.484, p = .014$), though not for Drag vs. Wheel ($F(1, 11) = .744, p = .407$).
CHAPTER 5. EXPERIMENTAL RESULTS

Drag was predicted to have the lowest workload. Individual TLX categories (Figure 5.13 (p. 78)) do not reveal a particular category that causes Corner to succeed and Drag to fail. Category weightings are shown in Figure 5.14 (p. 79), but they only show that the Frustration category, for most participants, did not count at all towards their WWL.

We can find insight, however, in the participant’s comments. Participants found Drag to be mentally difficult and unpredictable, which highlighted the simplicity and predictability of Corner: one participant said Corner was "Very simple. I never had to think about it". With Drag some users were never sure whether it would behave as they intended, especially when trying to perform a translate-only action, e.g., "not able to see the area where only translation movement occurs". Of course, 10 of the 12 participants indicated that they had previous experience with the corner-rotation technique, whereas Drag was completely new to users, so longer practice might produce better results for Wheel and Drag.

For Wheel, users comments indicated that the significant workload problem was the mouse wheel: overloading the scrolling finger so that there is high physical effort, and yet slow results.

Figure 5.13: TLX scores for individual categories
CHAPTER 5. EXPERIMENTAL RESULTS

Figure 5.14: Weightings for TLX categories
CHAPTER 5. EXPERIMENTAL RESULTS

Order Effects of TLX

The TLX scores seem to show an order effect, as seen in Figure 5.15 (p. 80). The variance is high and analysis of variance was non-significant for IT, but visually the increasing trend seems clear. This could be due to physical strain, but also possibly the mental strain of performing mindless repetitions of tasks.

![Figure 5.15: Order effect on Average TLX Workload]

TLX and RSD

Since Drag was rated very differently by the RSD and TLX measures, it is interesting to know how these two measure of difficulty are related. Figure 5.16 (p. 81) shows that there does not appear to be a strong correlation between the two measures. It was expected that RSD would be closely related to workload, but it is clearly not related to TLX (Pearson Correlation=.131, Sig.(2-tailed)=.445, N=36). This brings into question whether RSD and TLX measure the same construct, and what is meant by ‘workload’.

Performance and Integrality: Approaches used for the Drag IT

The unpredictability that some users experienced with drag is due in part to the particular way they used it. The way I intended it to be used was literally a ‘drag’ technique, where the object would follow the pointer like a trailer pulled by a car. Dragging in that manner requires a drag point on the ‘leading edge’ of the rectangle. When this is done, the results are relatively predictable: the object will be pulled along behind, so the orientation at the
CHAPTER 5. EXPERIMENTAL RESULTS

end of the motion will be determined by the path in the last approach to the object. For larger rotations, the drag point can be moved slightly, or the path can be more curved.

However, the decision was made for the study to show participants the Drag technique without recommending one particular strategy. Many participants did not use the dragging strategy; instead, a number of participants usually attempted to place a ‘perfect’ drag point farther in, so that the constant rate of rotation while they moved would be exactly right to result in the correct orientation when the rectangle reached the target. Figure 5.17 (p. 82) (TR30) shows, for each user, the location on the object of the first button press in each trial (i.e. the drag point), represented by a small circle⁹. The colour of the circle indicates the M-metric value of the trial that started in that location. With this visualization, examples of the different approaches can be seen. Dragging by the leading edge (in TR30, the mid right side of the rectangle) can be seen in user 34 and 195, while users 156 and 164 demonstrated the ‘perfect drag point for rotation-rate’ approach. One other approach is also apparent for user 135, and probably users 111 and 172: using Drag just like Corner. Notice that user

---

⁹Circles allow overlapping points to be more obvious
Location on the rectangle of the first mouse click for each trial

(Drag, TR30 Condition)

Colour indicates M-metric of the trial that started with that click

Figure 5.17: Locations of the first click for the Drag IT in TR30
135 had drag points almost exclusively in the top right corner, and that the M-metric was almost always zero. Figure 5.18 (p. 83) shows a timeline graph of one of user 135’s TR90 trials (like the one seen for Wheel burstiness in Figure Fig:wheelTrial) that reveals that the Corner-like behaviour: rotation was done first (position ended up basically unchanged), then there was a pause while the cursor moved to initiate the translate-only mode, and finally the object was translated. The interaction is completely different than integral use of Drag, such as the TR90 trial illustrated in Figure 5.19 (p. 84).

The Drag strategies inferred from the drag point locations can be verified by the measures of integrality for each user, shown in Figure 5.20 (p. 84). The two participants who appeared to use the dragging / leading-edge approach (users 34 and 195), had the highest Integrality; the three users who appeared to use the corner approach (users 111, 135, and 172) had the lowest Integrality. It is worth noting that those three who adopted the corner strategy were all in IT-orderings that used Drag after Corner.

![Movement of DOFs in a Drag TR90 Trial (user 135)](image)

Figure 5.18: Corner-like approach to using Drag (During a TR90 Trial)
Figure 5.19: Integral approach to using Drag (During a TR90 Trial)

Figure 5.20: Integrality for each user (Drag IT, TR30)
CHAPTER 5. EXPERIMENTAL RESULTS

Using the dragging approach does not necessarily guarantee better performance. However, the dragging approach requires less visual attention. The perfect-drag-point approach is overly sensitive to placement of the drag point, since it depends on the rotation rate being perfect. Slight variations in the position of the drag point will have a big effect on the rotation rate, so users have to carefully watch the rectangle to see how much it is rotating. If it rotates too much or too little, the rectangle has to be released and re-acquired from a new drag-point. For the dragging approach, on the other hand, variations in the drag point have only slight impact on the rotation, since the object simply follows the pointer. If the object is acquired somewhere along the desired leading edge, the final orientation is primarily controlled by the path, specifically the final angle of approach to the target — adjustments that do not require the user to re-acquire the object.

This finding reveals that “interaction technique” was too broad a term. Though the Drag interaction technique was identical for all participants, they did not employ a unique “technique”. Ted Kirkpatrick has suggested the term “interaction gesture” be used to capture this distinction (in press).

Acquisition Time

Figure 5.21 (p. 86) shows the average acquisition time in each block. Analysis of variance showed a significant effect for interaction technique ($F(1,59) = 38.3, p < .001$). Drag was particularly longer. This could be related to the behaviour discussed in the previous section, especially for users seeking a perfect drag point, in light of the uncertainty many users expressed about where exactly to click.

It is interesting that Wheel had faster acquisition than Corner, even for the T (translate-only) condition. There are two factors that could caused this. The first is that the Wheel object (the whole rectangle) is a bigger acquisition target than for Corner (the corners take up space). However, since both targets were very large — and from a Fitt’s law perspective, the same width$^{10}$ — it seems unlikely. The second possibility is the different nature of the manipulations: the Wheel action was ‘click-then-point’, whereas Corner was ‘press-then-drag’.

Surprisingly, the most common feedback in the whole study was about this: participants

$^{10}$Figure 4.2(a) (p. 52) shows how a straight line between the starting point (the ‘Start Trial’ button) and the rectangle does not include the corner regions, so for acquisition of the rectangle with Corner and Wheel, both ITs have the same “width” for an application of 2D Fitts’ law (MacKenzie and Buxton, 1992)
responded overwhelmingly to being able to click once to 'pick up' the object (Table 5.3, p. 87), despite the possible mode problems that Wheel could have created (since there was no visual feedback about whether it was picked up). Similar preference for point-and-click over drag-and-drop were observed by Inkpen (2001) in a study of mouse interaction by children, with 27% of participants specifically complaining about having to hold down their finger for drag-and-drop.

**Mouse Clicks**

As one might expect, the longest trials tended to have many clicks (shown in Figure 5.22 p. 87). More mouse clicks represent not just time wasted to reacquire the object, but the very need for reacquisition, i.e. readjustments, indicates that a manipulation error occurred. Note that the figure for Wheel may be unnaturally high, since mouse clicks were counted until the trial ended (700 ms after the target was lined up), and many subjects clicked to release the object so they could move the pointer without waiting for the trial to end.

Figure 5.21: Acquisition Time for each IT (all 5 blocks shown)
Participant Comment on Advantages of Wheel IT

111 “Easier ... since box got selected with single click and didn’t have to hold on to it”
123 “Precision pointing is easier to do”
142 “Click and drag easier than click and hold, simplifies the UI”
164 “No strain from keeping the mouse button pressed”
181 “[With Corner you] have to keep pressing the button while dragging the objects, [but with Wheel] you don’t have to push the button while dragging”
203 “I like how clicking the button picked up the ‘paper’ — it saved energy — although a few times, I forgot that I picked up the paper [it needs a] visual mode indicator”

Table 5.3: Comments about wheel reveal support for click-once pickup over mouse dragging.

Figure 5.22: Docking Time vs. # Mouse Clicks
5.2 Physical Manipulation of Paper

The timing given in Section 5.1.2 allowed a comparison of mouse-based ITSs, but did not provide a context for comparison to other approaches, such as tangible input. Yet we can make an approximate comparison based on the object transportation and orientation times given in (Wang et al., 1998). There, a block was transported 2 cm, 10 cm, or 20 cm, and rotated either 22.5° or 45°. Their completion time was taken from the start of the movement until the end, so it corresponds to the dockTime in my study. Their experiment did not display a specific required accuracy for trial completion; they displayed a precise target and to which the users lined up the wooden block, so it is reasonable to assume that the accuracy of lining up the wooden block was more accurate than was required in my experiment. Thus, while the studies are not appropriate for a precise timing comparison, it gives a ballpark for completing a similar task with physical manipulation. Average completion times by distance are shown in Figure 5.23 (p. 88) and extrapolated to 37 cm, the magnitude of translation in my user study. This suggests an average time of approximately 1100–1200 ms. This stands in contrast to the best means among the three ITSs, 1483 ms for TR30 (Wheel) and 2095 ms for TR90 (Corner). Note that for physical manipulation, the angle or rotation does not seem to increase the completion time. The authors explained was because the rotation takes less time than the translation; rotation is completely contained within the transport time. Thus a 90° rotation, such as the one in my experiment, would not likely take any longer than the 22.5° or 45° rotations. To the extent that these extrapolations hold, moving and rotating a physical block in conditions similar to TR30 and TR90 would be the fastest IT, particularly for TR90.

![Object Transport and Rotation](image)

Figure 5.23: Completion Time for Physical Manipulation

To see what the timing would be like for the experimental setup in my user study, I
performed a follow-up with a paper condition. A 16.2 x 20.9 cm sheet of paper was placed on a desk on which a target had been drawn in pencil, with the same dimensions as the study. Because the target was a pencil drawing, task configurations were not randomly mixed. Performing the same trial repeatedly will generally result in faster times than random trials. However, it should be noted that Balakrishnan et al. (1997) found that for users performing four 3D-movement trials in a row to the same general target location, performance improved significantly between the first trial and the second trial, but not much between the remaining three trials, so these results should be reasonable. I performed forty (40) trials of each configuration sequentially, with the following order of configurations: T, R30, TR30, TR90, R90. A final 20 trials of the T configuration were performed at the end to capture the learning that had occurred.

The same computer system used in the experiment was programmed as a stopwatch to record times between mouse button presses. I manipulated the paper with my left hand, and with my right hand I started the timer just before touching the paper, and stopped it as soon as I saw the paper within the target circles, as an approximation to dockTime.

Figure 5.24: Completion Times for Paper-based Recreation of the Study

The results are shown in Figure 5.24 (p. 89).\footnote{Mean dockTime-equivalent: T\textsubscript{firstblock}=588 ms; T\textsubscript{secondblock}=422 ms; R30=272 ms; R90=498 ms; TR30=451 ms; TR90=553 ms} As with Wang et al. (1998), there
was very little difference in completion time between TR30 and TR90, indicating that the rotation was incorporated into the transport. The lower overall times than Wang et al. can be explained by the much looser accuracy requirements of this study, though the experimenter’s desire for speed should not be discounted. The results provide an initial indication that the mouse techniques for 3-DOF manipulation are 300–400% slower than what is consistently achievable with physical paper manipulation.

A significant observation about paper relates to an interesting question about direct input. At some point in the trials, I realized that I was not looking at the paper, since I was focusing on the target to stop the timer as soon as it was lined up. I knew where the sheet was on the desk and I could place my hand on it and move it without needing to look at it. Rich feedback confirmed that I was moving it, including not only my sense of touch, but also the sound of the paper moving. The ability to manipulate without looking is significant because it highlights one of the benefits that becomes possible with direct input: taking advantage of our sense of kinaesthesia, our awareness of where things are in relation to our body and the location and angle of our limbs. In a realistic situation, instead of looking at a target I could have been doing some other meaningful activity such as reading another document. In fact, that very behaviour was described in (O’Hara and Sellen, 1997). Without direct input it would not be possible.

There are, of course, many factors that may be at work:

1. Direct Input

2. Tactual Feedback (paper on finger, and texture of desk through the paper)

3. Acoustic Feedback (sound of it sliding)

4. Peripheral Vision

5. Repeated fixed target start and end

On the tabletop display, a direct input interface based on tracking physical objects will provide all of these cues, though in general use the start and end of movements will be varied.
5.3 Final Comment

Performance and some of the subjective evaluation showed that Drag was a poor IT for object manipulation. Yet I would like to note that for my personal use, there is no question that I would choose Drag over Corner or Wheel for the informal interaction of manipulating documents on a tabletop display. As an ‘expert user’ I found it to have a sense of ease, and from a subjective perspective it felt best at manipulating position and orientation quickly: my perception was that it was quick, whether or not the timing agreed. I also have a strong sense that direct input would enhance Drag. Similarly, I would prefer Wheel to Corner, as long as the increment were increased, or possibly if a TouchPoint (isometric joystick) mouse were used instead of the wheel.

5.4 Conclusions

Corner, a separable IT, had the best performance on the separated conditions, but the integral ITs did not perform nearly as well as expected on the integral conditions, despite the correlation between Integrality and performance.

Wheel was found by metrics of integrality to be only slightly integral, which was primarily due to the bursty nature of the scroll wheel, which caused the DOFs to follow a non-Euclidean path. Yet further examination revealed that the Wheel movement shared the same basic structure as physical object manipulation, with translation occurring for the whole movement and rotation contained within that translation. Analysis of the “timesaving overlap” expected for integral 3-DOF manipulation confirmed that overlap occurred, in contrast to the large time penalty that the separable IT suffered in TR movements. This confirmed that the primary limiting factor on the Wheel IT was not lack of integrality, but the poor rotation provided by the scroll wheel. The scroll wheel did allow simultaneous control to a certain extent, but limited manipulation in the positional DOFs, particularly the y-direction, when it was spun.

Drag had a much higher Jacob-Sibert Integrality than Wheel, but the low M-metric revealed that much of the simultaneous manipulation of the DOFs was inefficient motion. An examination of the way Drag was used revealed that a variety of approaches were adopted by different users. The notion of drag as simply an “interaction technique” did not account for the possibility of varying gestures, especially not for users to employ Drag like Corner,
which they had used earlier, even though it did not provide the same affordances of Corner.

An attempt was made to gain a sense of how much attention was required by each IT, using two measures, RSD and TLX workload. There was difficulty in establishing what the connection was between the numerical scores, and the abstract concept that a score quantified. In particular, there was no correlation between RSD and TLX, and in fact they appeared to give opposite reports about the Drag IT, which scored worst in the explicit workload questions of TLX, and seemed to score best in RSD relative to the users RSD scores for the other ITS. In the future, I would replace the TLX questions with questions specifically related to the study in question. RSD is still interesting and seems to have potential, but its underlying construct needs to be clarified.

Overall, the mouse control proved to be poor in comparison to the speed and DOF overlap seen in the paper-based recreation of the tasks. It is to be expected that even the most preferred 2D pointer technique still leaves much to be desired in comparison to the richness of direct physical manipulation --- but it is not possible to predict when such technology will become widely available and affordable.
Chapter 6

Conclusions

Tabletop displays suffer from a problem of orientation. Despite the variety of orientation solutions adopted by tabletop display designers, there remained unanswered questions about what kind of orientation control should be provided. Since designers of tabletop displays have often been motivated by a desire to let people work just like they do on real desks and tables, I argued that requirements for orientation on tabletop displays should be based on document interaction. The manner in which people work with documents involves moving them frequently, freely changing the orientation and position to read, write, and organize information. Thus I was able to establish requirements for orientation solutions: they must provide quick, either-handed, low-attention manipulation of individual documents on the display.

I evaluated interaction techniques (ITs) for orientation control in light of the requirements for document interaction, and many of the techniques were eliminated: non-oriented and automatically-oriented tabletop displays, 2-handed interaction, and physical rotation of the whole display. This left only three categories of IT: mode-switching, multi-point touch, and 3-DOF physical manipulation. I arranged the remaining ITs in these three categories in an IT exploration space based on two factors from the affordances of physical document interaction, integrality and directness.

I designed an experiment to explore the interaction of directness and integrality by comparing ITs on the touchscreen to integral, indirect mouse control. However, pilot tests found that the touchscreen tabletop display was unusable due to the physical strain caused by the way it prevented users from leaning on the surface to support themselves, a situation that was exacerbated by the wide border built into the table. Prior to the study, such a system
would likely have been strongly recommended. Now it is clear that current touchscreen technology cannot really be considered for general interaction on tabletop displays, except perhaps for small displays or applications in which only brief interaction is expected.

I performed an experiment to examine indirect control, contrasting a mode-based separated IT with two new potentially-integral ITs.

The separated ‘Corner’ IT was simple and performed well. The new ITs, Wheel and Drag, were weakly-integral and unsatisfactory on TR tasks, where they should have dominated. For Wheel, the slow rotation control of the scroll wheel crippled it as a 3-DOF manipulation technique. Drag was harder to characterize: its overall performance was relatively poor, and some users found it very difficult and unpredictable, but other users had success with it. Different users actually used the IT in different ways. Potential exists for further exploration of the drag IT, such as visual indicators, different drag behaviour, or a Drag-Corner hybrid, as well as an evaluation of the effect of direct input on performance.

The timing results indicated that indirect control (Corner) was capable of providing fast manipulation performance on tabletop displays for separated tasks. However, as a separated IT, more general manipulation (translation and rotation) was slow. The integrality results for the weakly integral ITs indicated that a better form of integral input that performed faster in individual dimensions (translation or rotation), would be much better suited for object manipulation than the separated Corner technique. In particular, physical object manipulation is very fast even for large rotation angles, suggesting that an IT such as graspable 3-DOF input would be best suited for manipulation on tabletop displays.

Yet the effect of direct input still needs to be explored. The paper-based exploration suggested that direct input might enhance other aspects of document manipulation, especially due to the inherent multi-sensory feedback that makes it possible to rely less on visual feedback for acquiring and manipulating the objects. However, the exact effect of directness on a wider variety of ITs needs to be examined.

**Recommendations**

For manipulation on tabletop displays, a separated orientation control IT like the corner-rotation method, is poorly suited to the frequent manipulations expected for documents on tabletop displays.

Tabletop displays could be used in conjunction with digital reading devices. Physically rotating the devices would offer orientation control while maintaining high resolution. The
tabletop display would provide missing attributes of reading devices, such as large public displays and spatial layout.

The best control would be direct, integral input. The Sensetable, built by Pattu, Ishii, Hines, and Pangaro (2001), implemented a method for 3DOF input on a tabletop display by combining the input of two large drawing tablets, and projecting the display onto those tablets. Their system experienced lag, and was smaller than would be desired for general tabletop displays (52 cm x 77 cm) (unlike a display it is not easily scalable), but it did demonstrate how larger 3DOF sensors could provide the desired control.

Another possibility exists in multi-touch sensing technologies. Touchscreens would need to avoid the fatigue problems observed in this thesis. For this to happen the system must ignore unintentional touches, which could be accomplished in two ways. First, the system should use touch input only for document manipulation, and stylus for “information” input. As O’Hara and Sellen (1997) observed, this avoids the drawback of current GUIs which require users to switch into a “navigation mode” by acting in a small control area (e.g., the title bar of a window). Second, the touchscreen should be pressure sensitive. If documents only moved only when the pressure were greater than a certain threshold, it would avoid unintended manipulations just like in real life (e.g., brushing your finger across a document may not move it, but when you press harder it will move). It could also allow richer multi-finger input, since the pressure of different fingers could resolve the issue of which finger should be considered an anchor if the fingers move relative to each other. Furthermore, it would solve the problem of how to interpret objects on surface, since with real paper documents it is the weight of objects on the surface that determines whether documents underneath the object can be moved.

Finally, augmented paper is a promising approach. Further developments must be made to ensure robust tracking and to avoid occlusion of the image, but the ease of manipulation will be highly desirable since the use of paper intrinsically exploits familiar manipulation skills. Projecting onto paper could have an added benefit of bringing a non-horizontal aspect to the tabletop display: if the paper were traced in 6-DOF, the information could be projected on it even if the sheet were raised off the surface or even if it were angled a small amount towards the reader.

However, these ideal control techniques for direct, integral input are not currently available, and it is not possible to predict when the technology will change.

For general manipulation on tabletop displays at this time, I would recommend opting for
CHAPTER 6. CONCLUSIONS

integral input even if the input must be indirect. Specifically, I would recommend 3-DOF input on a digitizing tablet. The customization required to create the tracked graspable object was beyond the scope of this thesis, but it can be accomplished by altering relatively inexpensive, commercially available components. Tablet input is absolute, so unlike a mouse (which as a relative input device can only be used from one side of the table) the input will always retain the same reference frame as the display regardless of the user’s location. Tablet input requires a specific input area (the tablet), which would require installing multiple tablets on the tabletop display, e.g., one on each edge of the table (off-centre, so it would not be in the way of the ‘central focus’ area where documents would be read).

Future Work

The future work stemming from this thesis can be summarized as follows:

- Further development of the Drag technique, refining the model so it behaves as users expect.
- Comparison of Drag on direct input vs. indirect.
- More general examination of the effect of the directness of input for manipulation on tabletop displays.
- Develop interaction techniques for manipulating documents on tabletop displays that also support seamlessly moving documents between the tabletop display and portable reading devices.
- Implement a rotational environment for easier testing, such as a Linux windows manager that would allow normal applications to be rotated.
- In the meantime, it appears that the best realistic, general-use IT for controlling documents must be implemented and evaluated on tabletop displays, 3-DOF manipulation on drawing tablets.

In this thesis I believe I have shown that good techniques for orientation control are a necessity for successful implementation of tabletop displays. The way people interact with paper documents is not currently supported in a satisfactory manner, and more work is needed to provide the required control, if the desired benefits of tabletop displays are to realized.
Appendix A

Appendix

A.1 Approval of the SFU Research Ethics Board

The approval of the SFU Research Ethics Board for research involving human participants is included on page 98.
April 17, 2003

Mr. Daryn Mitchell  
Graduate Student  
School of Computing Science  
Simon Fraser University

Dear Mr. Mitchell:

Re: Moving and Rotating Objects on Tabletop Computer Displays

The above-titled ethics application has been granted approval by the Simon Fraser Research Ethics Board, at its meeting on March 17, 2003 in accordance with Policy R 20.01, "Ethics Review of Research Involving Human Subjects".

Sincerely,

Dr. Hal Weizberg, Director  
Office of Research Ethics

---

For inclusion in thesis/dissertation/extended essays/research project report, as submitted to the university library in fulfillment of final requirements for graduation. Note: correct page number required.
Bibliography


