

A Model of Embodied Dynamic Peephole Pointing for Hidden Targets

Jochen Huber, Jürgen Steimle, Max Mühlhäuser
Technische Universität Darmstadt
Hochschulstraße 10
64289 Darmstadt, Germany
{jhuber,steimle,max}@tk.informatik.tu-darmstadt.de

Embodied interaction with spatially-aware displays allows users to explore virtual information spaces situated in the real world. However, users are only able to see a limited part of the information space through the rather small display window. Targets are thus often hidden. Optimizing the layout of the information space by considering navigation times to targets therefore becomes essential to increasing the efficiency of a user. We contribute a novel model for the embodied navigation of a-priori unknown information spaces with spatially-aware displays. The model is inspired by physiological aspects of the human body. We have empirically validated the model in a controlled experiment with 32 participants.

Peephole, pointing, Fitts law, theoretical model, user study, spatially aware displays

1. INTRODUCTION

In recent years, the capabilities of mobile devices have considerably increased. They can support users in browsing large multimedia information spaces such as knowledge networks or video collections while being on the move. Recently, two interaction techniques have gained particular attention for this application: peephole interaction (Yee 2003) and flashlight interaction. Both are based on a similar idea. They assume the mobile device to be situated in physical space (as a so-called spatially aware display). The device can then be moved in space to explore a virtual information space in the real world. In case of peephole interaction, the screen of the mobile device is used as a window that overlays the physical information space with additional virtual information. The flashlight metaphor utilizes mobile projectors to display the information space. However, both techniques reveal only a part of the virtual information space to the user. In order to fully explore the information space, the device has to be sequentially moved over the entire space.

Since the user is only able to see a limited part of the information space through the small window, loss of orientation is clearly an issue (Jul and Furnas 1998). Besides addressing this with visual cues (Cockburn, Karlson, and Bederson 2008), optimizing the layout of the information space is essential to increasing the efficiency of a user. The layout can be optimized by considering navigation times to targets depending i.a. on their distance.

The present paper addresses the modeling of navigation times in information spaces with a-priori unknown target locations. Since both interaction

techniques mentioned above utilize embodied interaction through pointing to a target, the problem can be seen as a pointing task. Pointing tasks have been most commonly modeled through Fitts' law. However, in his experiments, Fitts' measured the pointing time between two visible targets, hence an *aimed* movement. In our case, the target is not visible a-priori. Previous research proposed several models for peephole pointing (Rohs and Oulasvirta 2008; Cao, Li, and Balakrishnan 2008). However, in all experiments, the participants were given directional or visual hints for the target's location. As a consequence to this, the empirical data showed a high correlation with variations of Fitts' formula.

We contribute a novel, nonlinear model for the embodied navigation of a-priori unknown information spaces with spatially aware displays. The exploration of unknown information spaces is relevant for many novel location-aware applications such as handheld augmented reality browsers, where users are constantly confronted with new information spaces. Our model focuses on one-dimensional pointing as a basis for future research in the area of multi-dimensional embodied peephole pointing. The model is inspired by physiological aspects of the human body. We conducted a controlled experiment with 32 participants which provides empirical evidence that this type of navigation does not follow a Fitts' law—as claimed in previous experiments.

The remainder of this paper is structured as follows. We first discuss related modeling approaches. Next, we describe our theoretical model and illustrate its relevancy for embodied interaction. We then present the results of our experiments. We conclude the

paper with a discussion of the results and an outlook upon future work.

2. RELATED WORK

There is a large body of knowledge for movement time models. Every model is heavily influenced by the employed input technique and physiological aspects like vision and the human motor system. There exist models i.a. for pointing (Cao et al. 2008; Rohs and Oulasvirta 2008), scrolling (Andersen 2005) and aimed movement (Fitts 1992).

Fitts' original work (Fitts 1992) focused on one-dimensional aimed movement, where subjects were asked to tap two visible targets consecutively. His model predicts the movement time T to a target of width W in dependency of the target distance D . The model he derived is typically formulated as $T = a + b \log_2(1 + D/W)$, where the logarithmic term defines the index of difficulty (ID) and both a and b are empirically determined constants. There exist various different interpretations of Fitts' law (Drewes 2010) and it has also been extended to higher dimensions (MacKenzie and Buxton 1992). This model of aimed movement inspired movement time models for input techniques such as peephole interaction, e.g. by Cao et al. (2008) or Rohs and Oulasvirta (2008).

Cao et al. (2008) conclude that one-dimensional dynamic peephole pointing follows a Fitts' law with

$$T = a + b \left(n \log_2 \left(\frac{D}{S} + 1 \right) + (1 - n) \log_2 \left(\frac{D}{W} + 1 \right) \right), \quad (1)$$

where S designates the window size and n is empirically determined. In their experiment, they simulated peephole pointing tasks using a 22" screen as their information space. The peephole as a window onto the information space was controlled using a stylus and a Wacom tablet. Targets were simple vertical lines of a certain width and "infinite" height. Although the targets in the experiment were truly hidden, the participants were given a directional hint towards the target location. However, the experimental setup does not consider embodied interaction with a physical display (which in turn has a certain friction and acceleration). Moreover, the Wacom tablet only provides a small interaction space which does not allow to cover physical constraints imposed by human physiology, which are crucial for embodied interaction.

The work by Rohs and Oulasvirta (2008) targets embodied interaction using mobile devices in two-dimensional space. They claim that dynamic peephole pointing clearly follows a variation of Fitts' law (also with a logarithmic ID), particularly for

the case when the targets are not visible to the users. However, in their experiment, the participants were a-priori aware of the target's position and their actual task was rather an aimed movement, which consequently implies a high correlation with Fitts' formula.

Andersen (2005) derived a movement time model for one-dimensional scrolling tasks. Scrolling tasks are quite similar to peephole pointing tasks, since users only see a small portion of the information space (e.g. a browser window). Andersen's experiment contained implicit hints by letting participants either begin at the top or at the bottom of a document. Andersen found that scrolling for hidden targets does not follow a Fitts' law, since (1) the targets are obviously not visible as opposed to Fitts' experimental setup and (2) the actual movement time is limited by human perception, namely the maximum rate at which a target can be perceived when scrolling by. Andersen found that 1D scrolling tasks follow a simple linear model:

$$T = a + b D, \quad (2)$$

with D being the distance to the target and a and b being empirically determined constants.

3. MOVEMENT TIME MODEL

In the following, we derive a model for one-dimensional dynamic peephole pointing with spatially aware displays where the target's location is a-priori unknown. Users employ embodied interaction to navigate the information space, therefore moving the whole display (or device) through space. The latter fact implies that the navigation within the information space is highly depending on the human motor system. Users hold the device in hand and move it horizontally along an axis to browse the one dimensional information space (see Figure 1). Without loss of generality, we assume that the width of the information space is limited by one's arm length and that one starts exploring it in the middle. Hence, the user has the possibility to either start navigating towards the right or the left hand side. We moreover assume that the user starts her navigation in the center of the information space.

We claim that the device movement time using embodied interaction is trigonometrically dependent due to (1) a non-linear acceleration when moving ones arm and a slightly slower movement directly in front of ones upper body (see interval 0 to d_0 in Fig. 1), (2) a more flexible movement when the arm is near one's upper body (hence the arm is relaxed and not sprawled, see interval d_0 to d_1 in Fig. 1) and (3) a decrease in movement speed after the arm has been sprawled out to a certain degree (see

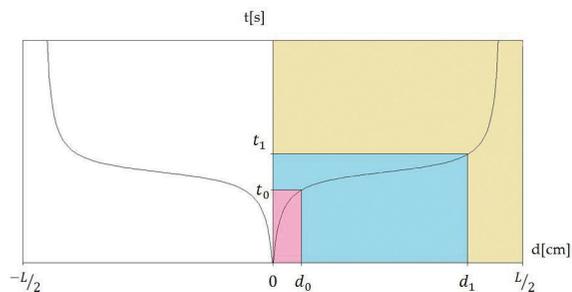


Figure 1: Model for the arm movement of a spatially aware display in one-dimensional space. The vertical axis denotes the time required for the movement to a position. The display is moved along the horizontal axis.

interval d_1 to $\frac{L}{2}$ in Fig. 1. This movement represents a tangent, shifted by half a period and scaled to match the interval of $(-\frac{L}{2}, \frac{L}{2})$.

The display is of width $S \in \mathbb{R}$ and the information space is of width $L \in \mathbb{R}$. The target's distance is $D \in \mathbb{R}$ and naturally limited by L . This leads to the following formulae:

$$T = a + \frac{b}{S} \tan \left(\frac{D}{L} \pi + \frac{\pi}{2} \right). \quad (3)$$

Both $a, b \in \mathbb{R}$ are empirically determined constants, where a designates the initial reaction time and b depends on the probability a user chooses the correct direction upfront.

4. EXPERIMENT

Prior research on the movement analysis of dynamic peephole pointing has focused either on tasks with prior knowledge of the target location and thus an aimed movement or on other specific interaction techniques, not on embodied interaction. Moreover, directional cues regarding the target location were given in all studies. We therefore wanted to particularly investigate the impact of the human physiology on embodied dynamic peephole pointing and the case without any directional cue. Thus, we verified the following hypotheses with a controlled experiment:

- H1:** The character of the information space (familiar versus unfamiliar contents) affects the search time.
- H2:** Embodied dynamic peephole pointing for hidden targets in an unfamiliar information space is neither sufficiently modeled by Fitts' law or its derivatives (e.g. Cao et al. (2008)), nor by a linear model.

H3: A larger target distance results in a larger search time.

H4: A larger peephole results in shorter search times.

H5: Starting in the wrong direction will add a constant factor to the time taken when starting in the correct direction.

4.1. Experiment Setup and Methodology

Apparatus We have designed a physical apparatus (see Figure 2) for our experiment instead of utilizing a handheld device. A physical apparatus allows for embodied interaction and ensures a reliable in-/output with minimal noise (which can be easily caused by hand jitter when using a handheld device in 3D or tracking errors). We thus are able to abstract from a concrete interaction metaphor such as a flashlight interaction with a mobile projector, without losing general applicability of our results for embodied dynamic peephole interaction.

The apparatus consists of a 1,40 m long and 10 cm wide rail and a belt with an exchangeable plastic window. The participants were apparently able to estimate the length of the rail, but this does not impact our experiment, since we assume that the length of the information space is limited by one's arm length. The window was used as a peephole onto a strip of paper, representing the information space. The targets were printed onto the strip (see Fig. 2). The window was equipped with a handle, enabling the participants to slide the window along the rail, thereby revealing the targets. The physical apparatus was designed such that the acceleration is comparable to that of a handheld device used for embodied interaction. Moreover, once accelerated, there was virtually no friction, comparable to the movement of a handheld device in the air.

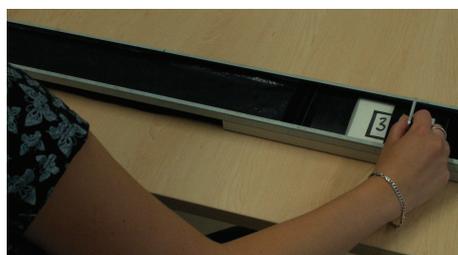


Figure 2: Participant manipulating the window. The targets were printed onto a physical paper strip.

Participants We have conducted a controlled experiment with 32 participants (10 female, 22 male, 29 right- and 2 left-handed). The window handle was positioned according to the handedness (i.e. left for a left-handed participant) to exclude any effects through e.g. occlusion. The age of the participants

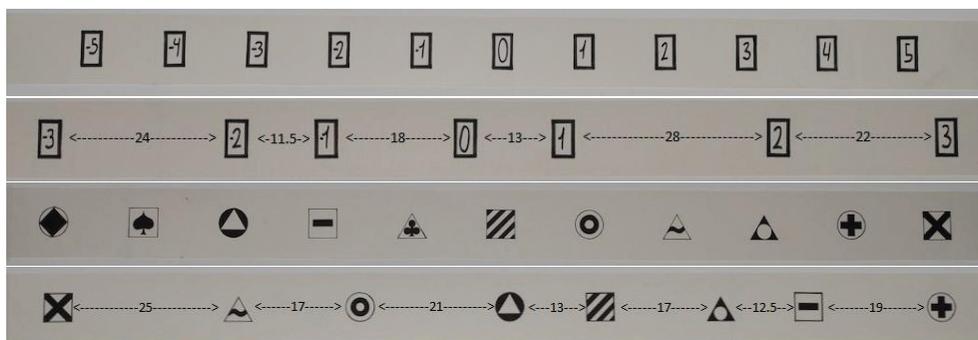


Figure 3: From top to bottom: (1) equidistant numbers, each $4\text{ cm} \times 5\text{ cm}$ large and spaced 8.5 cm apart from each other, (2) non-equidistant numbers, (3) equidistant symbols, with a baseline of 4 cm and spaced 8.5 cm apart from each other, (4) non-equidistant symbols. The distances in the non-equidistant case were not printed onto the paper strip.

ranged from 22 to 30 years. All participants had perfect (natural or corrected) vision. Each session took about 60 minutes.

Design Independent variables were the window size and the target distance. Since we varied the window size, we opted for a constant target width. We utilized two different window sizes, a small window with $5 \times 8\text{ cm}$ (resembling the standard display form factor of today’s smartphones like the Apple iPhone) and a larger window with $8 \times 8\text{ cm}$. The dependent variable was the time it took a user to move the window from the center of the strip to the center of a target. The participants were asked to shortly confirm that they had reached the target. The participants used the apparatus horizontally on a table while being seated. They were seated at the center of the rail and the strip respectively. We did not mount the apparatus vertically to a wall to assess the performance while standing, since this only further limits the human motor system in the horizontal axis. We video-recorded the tasks and measured the navigation times manually by analyzing the videoframes.

We chose a within-subjects design, but split the participants into two groups with 16 participants each. Each group was assigned a different set of target distances. This allowed us to get a broader variety of target distances. The order of the tasks, target distances and data types was completely counter-balanced. The participants were introduced to the concepts and were allowed to familiarize themselves with the apparatus upfront.

Tasks and Target Types We particularly wanted to assess the difference between the navigation in a known and an unknown information space. We therefore chose two different target types, *numbers* and *symbols* (see Fig. 3). The numbers were ordered naturally with 0 being at the center of the strip. They resemble a known information space, since users build a mental model of the number ray and map it to the strip easily. The symbols served as an unknown information space. The targets were

distributed both equidistantly and non-equidistantly as shown in Figure 3.

The participants had to fulfill 16 tasks per window size (5 equidistant numbers, 3 non-equidistant numbers, 5 symbols and 3 non-equidistant symbols), resulting in a total of 32 tasks per participant and a total of 1024 data points. In case of the symbols, we showed the participants the symbol they had to look for before each task. The symbol remained visible throughout the whole task.

4.2. Results

Movement Time The average movement times T per task are shown in Table 1. The movement time increased monotonically with the target distance for all data sets. ANOVA tests revealed that this effect is statistically significant ($p < .001$). Moreover, Bonferroni post-hoc tests confirmed that this holds for all distances ($p < .001$). Although the participants took longer in average to find a target using the small window and consequently were faster using the larger window, the speed difference was not significant for any of the data sets.

Table 1: Average movement times per data set

Data	Small		Large	
	T_{avg} [s]	SD_{avg}	T_{avg} [s]	SD_{avg}
Num-Equi	1.11	0.06	0.87	0.07
Num-NEqui	1.39	0.10	1.20	0.07
<i>Correct Direction</i>				
Sym-Equi	1.75	0.09	1.43	0.08
Sym-NEqui	1.82	0.08	1.40	0.06
<i>Wrong Direction</i>				
Sym-Equi	4.84	0.33	4.35	0.32
Sym-NEqui	4.53	0.36	3.86	0.38

When the participants initially moved in the wrong direction, the movement time was significantly higher than when they directly moved towards the correct direction. ANOVA tests and Bonferroni post-hoc tests confirmed the significant effects for both equidistant

symbols (small window: $F_{1,18} = 21.01, p < .001$; large window: $F_{1,18} = 15.39, p < .001$) and non-equidistant symbols (small window: $F_{1,10} = 30.82, p < .001$; large window: $F_{1,10} = 26.66, p < .001$). The statistically significant speed-up ($p < .001$) for equidistant symbols was 2.9 and 2.76 for non-equidistant symbols in average.

Model fitting We fitted the movement times to the linear model from equation (2), Cao's model as in equation (1) and the trigonometric model from equation (3). Table 2 shows the parameter estimates, the respective standard errors for the estimates and the correlation coefficient R^2 for the numbers data sets; Table 3 for the symbols data sets respectively.

Cao's formulae yielded the best fit for all number data sets, whereas our proposed model based on the tangent yielded the best fit for all symbol data sets. The tangent model fit particularly well for the case where the participants initially went in the wrong direction. The results are highly significant ($p < .0001$). Figure 4 shows an example fit for all three formulae.

Table 2: Model fit data for the numbers data set.

	Win.	F	a	E_a	b	E_b	R^2
Equi	S	Lin.	0.37	0.03	0.02	0.00	0.91
		Cao	-0.39	0.04	0.52	0.01	0.94
		Tan	1.12	0.01	0.31	0.01	0.92
	L	Lin.	0.42	0.03	0.01	0.00	0.86
		Cao	-0.62	0.29	0.16	0.08	0.87
		Tan	0.88	0.01	0.19	0.01	0.85
NEqui	S	Lin.	0.07	0.03	0.03	0.00	0.98
		Cao	-1.31	0.08	0.93	0.03	0.98
		Tan	1.29	0.03	0.41	0.02	0.90
	L	Lin.	0.18	0.02	0.03	0.00	0.98
		Cao	-0.65	0.05	0.79	0.02	0.98
		Tan	1.09	0.01	0.33	0.01	0.97

4.3. Discussion

The main goal of our experiment was to investigate how search times in unfamiliar information spaces can be modeled for embodied peephole interaction. We found that the movement times for embodied dynamic peephole navigation significantly depend upon the target distance (H3). While this might seem trivial, it has not been investigated with a physical apparatus before. We found that Cao's model fits best for a familiar information space such as a ray of numbers, as used in our experiment. Users can easily build a mental model thereof and the impact of the physiological constraints of the human body is only minor.

We also found a difference in the search times for the equidistantly and non-equidistantly distributed

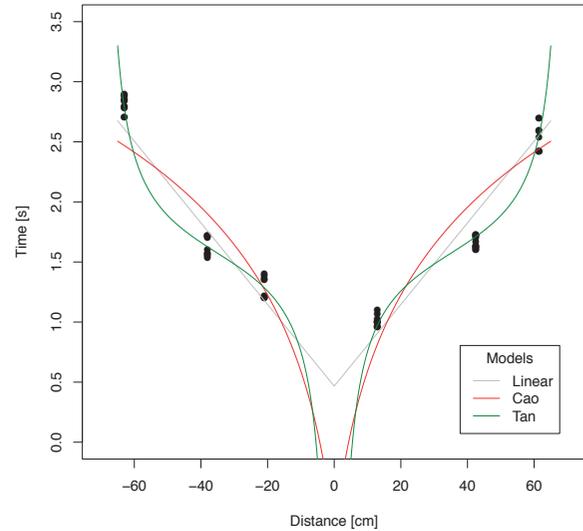


Figure 4: Model fits for non-equidistant symbols, small window. Participants initially chose the correct direction.

number rays. Although this difference is not significant, it gives a hint that the user's familiarity with the information space has an impact on the navigation (which we claimed in H1). We were able to confirm this hypothesis with the results for a completely unfamiliar information space for which we used symbols. In this case, the actual movement of the window played an important role and thus the trigonometric model based on the tangent fitted significantly better (H2). The actual distribution of the symbols, whether aligned equidistantly or not, had no influence on the movement times. However, the average search times for the non-equidistant distribution of the symbols were lower than those for the equidistant distribution. We assume that this attributes to the information density in the information space. The strip with the non-equidistant symbols contained only 8 symbols, whereas the equidistant strip contained 11 symbols.

The size of the utilized window had no significant effect on the navigation times (H4). This leads to very interesting hypothesis with practical relevance: Since we chose window dimensions similar to the displays used in today's handheld devices, this might imply that designers can abstract from the actual display size to a certain extent when designing information spaces which are to be navigated with spatially aware handheld devices. Density and distribution of information elements in the information space can then be determined without having a concrete device in mind during the design phase. However, this remains to be investigated in future experiments.

The initial search direction had a significant impact on the actual navigation time (H5). When the participants initially moved in the wrong direction, the search task was prolonged by a factor of at least 2.5. In this case, the other models fitted even worse

Table 3: Model fit data for the symbols data set.

	Data	Win.	F	a	Err _a	b	Err _b	R^2		a	Err _a	b	Err _b	R^2
	Correct Direction	Equi	S	Lin.	0.63	0.03	0.03	0.00		0.96	Wrong Direction	0.91	0.22	0.10
Cao				-0.36	0.09	0.72	0.03	0.92	-3.11	0.54		2.69	0.17	0.89
Tan				1.62	0.01	0.36	0.01	0.98	4.37	0.06		1.28	0.04	0.98
L			Lin.	0.54	0.02	0.02	0.0	0.95	1.67	0.17		0.07	0.00	0.92
			Cao	-0.08	0.06	0.63	0.02	0.93	-0.35	0.38		2.02	0.16	0.87
			Tan	1.33	0.01	0.28	0.01	0.97	4.08	0.05		0.89	0.03	0.97
NEqui		S	Lin.	0.47	.06	0.03	0.00	0.96	2.73	0.18		0.05	0.00	0.88
			Cao	-0.77	0.15	0.86	0.05	0.91	1.02	0.39		1.16	0.13	0.83
			Tan	1.57	0.01	0.39	0.01	0.99	4.19	0.06		0.53	0.04	0.92
		L	Lin.	0.44	0.04	0.02	0.00	0.96	1.91	0.16		0.05	0.00	0.89
			Cao	-0.30	0.08	0.69	0.03	0.95	0.57	0.31		1.35	0.12	0.86
			Tan	1.22	0.01	0.28	0.01	0.98	3.48	0.06		0.57	0.03	0.93

and the trigonometric model still fitted best. Although users typically have to decide where to move first in practice, previous research neglected this fact.

5. CONCLUSION

We have contributed a novel model for the embodied navigation of a-priori unknown information spaces with spatially-aware displays. It models the navigation time to a target in one-dimensional information spaces, depending on the size of the display and the distance to the center of the target. The model is inspired by physiological aspects of the human body and thus particularly addresses the affordances of embodied interaction with such displays. The results of a controlled experiment with 32 participants using a physical apparatus validate our model for the navigation in a-priori unknown information spaces.

We found that a user's familiarity with the information space and her initial search direction has a significant impact on the navigation time to hidden targets. When a user is familiar with the information space, the search times are logarithmically dependent as in Cao's case, whereas in the case of an unfamiliar information space, the search time significantly increases towards the physical end of the information space due to human physiology and is thus better modeled by a trigonometric formulae.

Our results provide hints that the density of the information space should be considered in future experiments. Our results furthermore lead to the hypothesis that the minimal variance in the display size of current handheld devices has no effect on navigation times. This should be investigated in future research, particularly in light of the advent of novel tablet devices such as Apple's iPad with a larger display.

6. REFERENCES

- Tue Haste Andersen. A simple movement time model for scrolling. In *CHI '05: CHI '05 extended abstracts on Human factors in computing systems*, pages 1180–1183, New York, NY, USA, 2005. ACM.
- Xiang Cao, Jacky Jie Li, and Ravin Balakrishnan. Peephole pointing: modeling acquisition of dynamically revealed targets. In *CHI '08: Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*, pages 1699–1708, New York, NY, USA, 2008. ACM.
- Andy Cockburn, Amy Karlson, and Benjamin B. Bederson. A review of overview+detail, zooming, and focus+context interfaces. *ACM Comput. Surv.*, 41(1):1–31, 2008.
- Heiko Drewes. Only one fitts' law formula please! In *CHI EA '10: Proceedings of the 28th of the international conference extended abstracts on Human factors in computing systems*, pages 2813–2822, New York, NY, USA, 2010. ACM.
- P. M. Fitts. The information capacity of the human motor system in controlling the amplitude of movement. 1954. *J Exp Psychol Gen*, 121(3):262–269, September 1992.
- Susanne Jul and George W. Furnas. Critical zones in desert fog: aids to multiscale navigation. In *Proc. UIST '98*, pages 97–106, New York, NY, USA, 1998. ACM.
- I. Scott MacKenzie and William Buxton. Extending fitts' law to two-dimensional tasks. In *CHI '92: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 219–226, New York, NY, USA, 1992. ACM.
- Michael Rohs and Antti Oulasvirta. Target acquisition with camera phones when used as magic lenses. In *Proc. CHI '08*, pages 1409–1418. ACM New York, NY, USA, 2008.
- Ka-Ping Yee. Peephole displays: pen interaction on spatially aware handheld computers. In *Proc. CHI '03*, pages 1–8, New York, NY, USA, 2003. ACM.