participants with: (a) *awareness of shoulder-surfing moments*, which in turn helps mediate their social interactions, and (b) *protection of information* when shoulder-surfing is detected.

After briefly summarizing related work, we raise several considerations and design challenges concerning shoulder surfing. We touch upon privacy, self-regulation and social interaction, and people's perceptions of territories and social distancing (proxemics). We then describe a number of solutions, each illustrating how participants can become aware of shoulder-surfing episodes, and how systems can afford some protection. We close by discussing our solutions in a broader context.

2. RELATED WORK

We are not the first to consider shoulder surfing over displays. Several somewhat specialized methods have already been disclosed, each offering some level of protection, as listed below.

Methods tuned to highly specific data. Security professionals are particularly concerned with shoulder surfing attacks of passwords, and have developed various password entry methods to protect against such attacks (e.g., [12]). Our work differs, as we do not know in advance what information may be deemed sensitive.

Limiting what people see based on viewing angles. Another approach physically limits what onlookers can see. As mentioned, this can be done by strategically locating displays in the environment to restrict how passers-by can view the display and its content or using special display techniques [11]. More generally, commercially-available privacy filters – screens attached atop of displays – cause the display to appear increasingly dark as the onlooker's viewing angle increases. Thus people looking at the screen from the side will not see anything. Because privacy filters do not stop a shoulder-surfer from seeing the screen from a straight-on position, strategic positioning of the display and body shielding must still be done. To our knowledge, privacy filters are usually restricted to relatively small displays (e.g., tablet to desktop displays), likely because it would compromise how a person could look around a very large screen.

Offloading private information to a trusted handheld device. Another approach considers how people can symbiotically use both a personal handheld device and a public display to help them perform a task more easily while still protecting privacy. Displaying and entering sensitive information only on a handheld mobile device rather than the public display, provides protection from shoulder surfers [1,9,19]. Another solution censors sensitive information on the public display, while leaving it uncensored on the handheld device. For example, Sharp et al. blurs the content surrounding the pointing device, while Berger et al. [5] blacks out sensitive words in an otherwise viewable document.

Proxemic interactions. Edward Hall's proxemics theory describes how one's social distance is correlated to one's physical distance from another person [10]. Hall defined four *proxemic zones* surrounding a person, beginning with the intimate zone at the center, and moving through the personal, social, and finally the public zone at the perimeter. As the names imply, social engagement is expected to increase as one approaches the other. This theory was exploited by Vogel et al. [22] to define four proxemic zones for large display interaction. His system was an event calendar that shows both public and personal information. From afar, the display presents ambient, undetailed public information. As one moves closer, the information presented and interaction allowed become increasingly detailed and personal. If a second person enters the area, the display is split to provide each with an area to view their own personal information. To safeguard privacy, a person can perform certain gestures, or simply step back away from the display to hide or mute personal information. Building on this work, Ballendat et al. [3] describe the Proxemic Media Player, which incorporates people's position, orientation, movement and identity in order to control a media player. While they do not address privacy per se, they illustrate how the system can balance the needs of particular people in front of them.

3. DESIGN CONSIDERATIONS

The above methods are based on simple assumptions of privacy. They treat all passers-by as suspected security threats. They have a strong notion that some information is clearly 'private', while others are clearly 'public'. This section raises other considerations, which in turn provides a more nuanced design perspective of how one can mitigate shoulder-surfing on public displays.

Privacy is a boundary regulation process. In many cases, privacy is respected through a mutual negotiation of the parties involved [2,6]. People generally respect other's territories, and will do so if they are aware that they may be intruding across another person's privacy boundary [6]. For example, if they glance at a public display and see someone reading email, they may self-regulate their behaviour by looking away, or negotiate permission through social protocol [6]. Similarly, if a person realizes that another may be intruding, they will signal that. Altman [2] describes how people use several different behavioral mechanisms to signal a desired level of privacy: verbal (speech content); para-verbal, (structure of their speech); nonverbal (body language, gaze); environmental behaviors (adjusting one's personal space); and cultural expectations. Thus revealing *more* information rather than less can be a good privacy-preserving strategy: it enables mutual awareness of the situation, which in turn allows people to regulate their behaviours. If awareness is not sufficient, then privacy violations may occur inadvertently [4].

Privacy as social distancing and as territories. People expect others to obey cultural expectations of proxemics [10]. In addition, people usually mark their territories through the use of symbols, objects and artifacts [2,18], which serve as further boundaries defining personal space. Most people respect social distance and territories of others. If someone breaks into another's personal territory or space, various protection mechanisms then come into play to negotiate what happens next. The problem is that a large display can change the dynamics of this process. Because the display and contents are visible at a distance, territorial boundaries and the size of proxemic zones can become ambiguous.

Private / Sensitive Information. The meaning of private and sensitive information varies between different people [16]. At one extreme, some people have little concern about their information (e.g., they may only be concerned about banking information). At another extreme some are highly sensitive to any information disclosure (e.g., routine purchasing behaviours). Of course, it also depends on context and what activity people are involved in. Therefore an automated system cannot successfully predict what data should be protected from shoulder surfing, as it is highly personal and context-dependent.

Public Displays: Shoulder Surfing and Honey-Pot Effect. Shoulder-surfing often arises from curiosity rather than malicious intent. The easier it is to shoulder-surf, the more likely it is that someone will do it. Tan et al. [21], for example, observed that people tend to be more voyeuristic with increasing display sizes. Similarly, both Brignull et al. [7] and Peltonen et. al. [17] observed that people are more likely to interact with a display if someone else is already working on it, and noted that the new person usually observes the existing user for a while before beginning to interact

with the display [17]. Thus the large display becomes a 'honey-pot' that attracts others to one's work. This is not necessarily a bad thing, for it could also encourage collaboration and engagement. As Müller et al. [14] state, this honey-pot effect is a very powerful cue to attract attention and increase engagement with public displays.

Summary. We believe that most shoulder-surfing on large displays will not arise from malicious intent, but through voyeurism, the honey-pot effect, and territorial and spatial ambiguities. Inadvertent violations may result from the shoulder-surfer not realizing that he or she is viewing sensitive (*vs.* public) information unless it is too late. Consequently, we argue that systems that mitigate shoulder surfing on public displays must meet two important criteria. First, the system should make the passerby and the user of the display aware that shoulder-surfing could occur or is occurring. If done well, both parties can regulate their behaviours via social protocol. Second, the system should provide some degree of shoulder-surfing protection over broad content (rather than about small units of information) until privacy is negotiated.

4. TECHNICAL FOUNDATIONS

Our technologic approach tracks and exploits the proxemic relations between a passerby, the user of the large display, and the display itself (Figure 2). We use the Proximity Toolkit [13] in conjunction with a Vicon motion tracking system to identify and capture each person's position and head / torso orientations, and to define the exact location of the display. In turn, this information is used to calculate the distance and orientation relationships between shoulder the surfer, the user, and the large display.

We also use meta-data about people and applications to provide context. Because we know identity, we can differentiate between strangers, friends, co-workers, and family. Inferences about privacy can then be driven by the expected social relationship, and appropriate levels of protection provided by the system. For example, when viewing a social media application, the user may be fine with friends and family shoulder-surfing, but not strangers.

At a higher level, the system can use all this information to make decisions based upon the inferred proxemic and social relationship between the user and a passerby, by spotting territorial intrusions [2], and by considering the user's desired level of privacy. In turn, these decisions are used to provide various awareness cues and protection mechanisms, as discussed next.

5. AWARENESS OF SHOULDER SURFERS

We can provide awareness to the user of a public display that someone is nearby and that his screen content might currently be shoulder-surfed. The user can then decide whether his displayed data needs protection, and use social behaviors to regulate privacy. For example, he can ask the passerby (either explicitly or implicitly through body language) to respect his territory by not looking at the display. He can also hide his private data by either closing the application or covering sensitive information with his body.

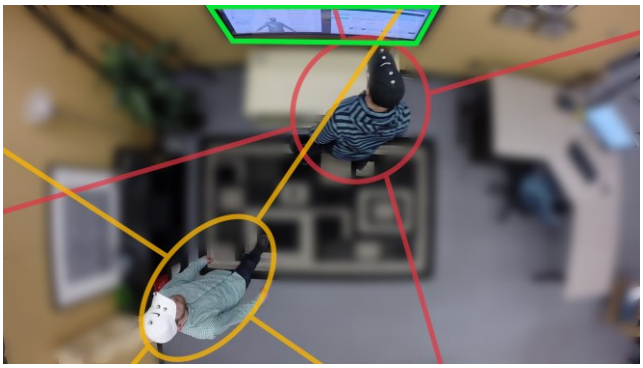

Figure 2. The Proximity Toolkit gives precise information about each person's position and orientation relative to each other and the display.

Our general approach uses visual indicators on the display to provide cues that another is passing by or actually shoulder-surfing. While these indicators primarily inform the user that a passerby is present, they also provide the passerby glancing at the display with an indication that they may be intruding into the user's territory. As our examples below illustrate, cues can range from abstract ones that provide only general awareness information, to literal and very precise cues that give fine-grained awareness of the passerby.

5.1 Flashing Borders

Our simplest cue uses flashing borders. As soon as a passerby enters the visible area around the display, the system notifies the user of the passerby's presence by selectively flashing its borders with meaningful colors. The borders flash green when someone is nearby but not looking at the display (Figure 3a). The color transforms to red as the passerby turns his head towards the display (Figures 3b and 1a). The relative distance of the passerby is coded into the transparency of the border: as the person approaches the display, the border color becomes increasingly opaque. The direction of the passerby can also be indicated by coloring only the sides and center / side border to roughly mirror that person's location (not shown).

Discussion: These cues, while simple, can provide significant awareness information. Because the user knows that someone has entered the scene but is not yet looking at the display (the green border), he can take advanced action to mitigate the potential threat, such as by hiding privacy-sensitive information, or by signaling the other person that privacy is desired. The distance and location cues (transparency and border side), while approximate, also provide the user with a sense of whether the passerby is moving through the area, has stopped, or is approaching the display. When the user knows that someone is actually shoulder-surfing (the red border), his actions can be even more decisive. However, the abstract nature of these cues likely make it inappropriate in walk up and use settings, as neither user nor passerby will know what the flashing borders mean unless they are somehow taught it.

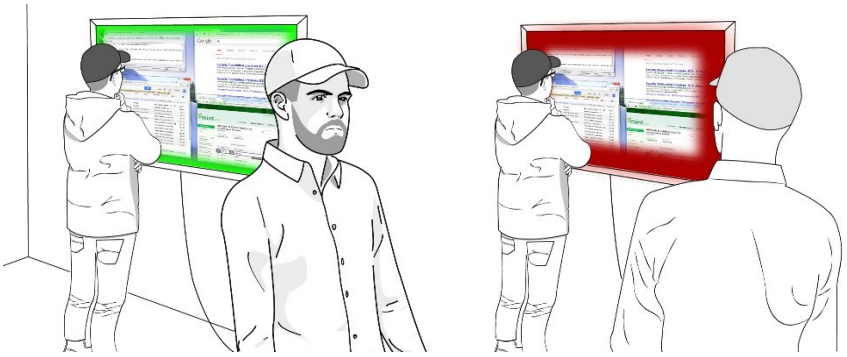

Figure 3. Left: The display border flashes green when a passerby enters a defined area around the display, Right: changing to red as he turns towards the display.

5.2 Mirroring the Passerby as a 3D-Model

Awareness cues can be very precise, where they accurately portray the actual location and

Figure 4. Mirroring a passerby's position and orientation with a 3D-model. The red dot indicates the gaze direction.

orientation of the passerby. We supply this information via a mirror effect, where the passerby's relative location is portrayed as a 3D-model on the screen (Figure 4).

When a passerby enters the display area, a 3D-model of a person appears on-screen, where its position mirrors that of the tracked passerby relative to the display. As the person moves across the room, so does the model. As the person approaches the display, the model increases in size. Additionally, the orientation of the model's head and torso are independently mapped to the tracked head and torso position of the person. For example, if the passerby turns his head (but not his body) towards the display for a quick glance, the model reflects that: its torso remains in its ~90^0 orientation from the display, while its head animates to turn towards the display (Figure 4). The model's transparency offers a further cue indicating how the passerby is attending the display, where the model becomes increasingly solid as a function of both distance and orientation.

Discussion: Unlike the abstract flashing borders, people can quickly comprehend that the model is mirroring passers-by, and understand its spatial relationship. The model informs the user of the passerby's presence, position, distance and orientation in an intuitive manner. It gives a full indication of a passerby's current whereabouts and look direction. Because the model is very responsive and animates in direct correspondence to the passerby's movements, the user can easily tell if someone is moving through the space, or has stopped, or is approaching, or is just giving a quick glance at the display, or is staring at it. Similarly, the passerby will see themselves on the display, and will understand that they have somehow intruded in the user's space by becoming part of it. Both parties can then act on this information as needed.

5.3 Gaze Awareness Indicator

Another visual cue indicates where on the display the shoulder surfer is gazing, i.e., approximately what they are looking at. This cue is realized as a red fuzzy dot, which moves about in a manner somewhat similar to how eye-tracking systems portray eye-gaze direction. Because we do not use eye-trackers, we assume viewing direction from a person's head orientation [8,15,20]. In particular, we consider the passerby's tracked head position and orientation as a vector, calculate its intersection with the display plane, and draw the dot around that intersection point as (Figure 4).

Discussion: The gaze awareness indicator provides reasonably precise information about what screen region a passerby is likely looking at. However, our current head-tracking implementation means that a shoulder surfer can 'game' the system by looking at the display from the corner of one's eye. This is why we see it best used in combination with other cues, such as the 3D-model, that gives additional information about what the passerby is doing.

6. PROVIDING PROTECTION

Awareness is just the first step in helping the user protect his privacy, or in informing the passerby that he or she may be violating the user's territory. This may suffice for many situations. When a territorial violation appears imminent, people normally self-regulate their behaviors to resolve the issue (e.g., where the passerby simply turns away), or enter in some kind of signaling and direct communication to negotiate access [6] (See Section 2). Yet there are times when further protection is needed. For example, even a quick glimpse of the display by the passerby may compromise one's privacy. Or, the user may want to take explicit action to safeguard sensitive information, perhaps because the passerby is just too curious, or because the user does not wish to socially engage with the passerby.

In this section, we show how we can exploit sensed information about people to provide both *explicit protection* (a user can take quick action to gain protection when he or she becomes aware of a potential violation), and *implicit protection* (the system triggers protection when it senses a potential violation).

6.1 Explicit: Moving or Hiding Content

When a person becomes aware of a shoulder surfing risk, he may want to take action to mitigate that risk. Shielding sensitive data with one's body is one such action. Yet because users typically spread application windows over the entire display area, shielding may be difficult or impractical in large display or multiple monitor settings. Alternately, the user may move, resize, hide, or even close windows containing sensitive information. However, conventional interface mechanisms require this to be performed one window at a time, which is a slow and tedious process.

Our solution follows the approach of Vogel et al. [22], in which a user can quickly invoke an action to safeguard privacy. Our particular safeguards allow the user to quickly move all windows to a portion of the screen directly in front of him (thus making shielding possible), or to hide windows until privacy intrusion is no longer a concern. Our first action is based on explicit gestures: the system recognizes a user's hand wave in one direction as a command to move all applications to that side of the display. Our second action is based on user orientation: the system recognizes when the user turns away from the display (for example, turning to face the passerby) and hides all windows by blacking out the screen. Both actions are quickly reversible, e.g., by the user waving his hand in the other direction to spread out the windows, or turning back towards the screen to reveal the windows.

Discussion: These techniques not only protect information, but reinforce how the passerby understands a user's territoriality. The passerby sees information being moved or hidden as a result of a user's action, which feeds into self-regulation and further negotiation. The downside is that explicit action takes extra work, and that the resulting window re-organization (or hiding) can disrupt what one is doing. We should note that easy moving of windows serves a dual purpose, where it can encourage sharing and collaboration rather than protection. For example, if the passerby wishes to use the public display for his own purposes (assuming the

current user invites the passerby to do so), moving windows to one side of the screen frees up space for both to work side by side.

6.2 Implicit: Blacking Out Sensitive Content

Because the system can implicitly recognize potential shoulder-surfing moments based on the passerby's relative position, it can take action to shield sensitive information from view. Ideally, the information will remain visible to the user but not to the passerby.

Our implementation does this on a window-level, where particular windows are tagged as public *vs.* personal. For example, the system may know what public windows it has provided (e.g., always-on public weather updates) *vs.* personal windows (e.g., ones the user has created, or has somehow marked as sensitive). Or, the system may keep a list of applications that are privacy-sensitive, such as an email reader, or search for keywords that identify sensitive content (e.g., "bank", "mail", "https"). The system then sets the transparency of each window, where it tries to strike a balance between masking (blacking out) the window's contents from the passerby, while still making it legible to the user. Windows are fully visible when no passers-by are present. As a passerby enters the area at a distance, the transparency levels of private windows are set to make them hard to read from afar but easy to read by the user (who is close to the display). Figure 5 portrays this situation. Transparency increases (and thus window legibility decreases) as a function of the passerby's distance and viewing direction: the closer the passerby gets to the display the more opaque the private windows. Similarly, when the passerby turns his view away from the display, those windows become more transparent.

We also allow the user to override system actions, e.g., by un-hiding windows using an explicit hand-wave gesture as described above. The user may want to do this for various reasons, such as inviting a colleague into collaboration. Thus the overall strategy is one where the system tries to automatically protect sensitive content (to mitigate privacy intrusions), but allows the user to easily override the system.

Discussion: Blacking out of selected content based on inferences of privacy incursion is a somewhat radical approach. Its advantage is that it not only offers protection, but it also clearly marks a potential privacy intrusion to both passerby and user. Another advantage is that public information remains available to the passerby (e.g., a public window showing the time or weather would remain visible). However, our particular implementation is not a sure-fire safeguard of privacy. First, it is difficult to balance occluding personal information from onlookers while still making it visible to the user. Thus this strategy provides only partial protection [22]. Second, it requires that the system somehow 'knows' the difference between sensitive *vs.* public content. As mentioned in Section 2, what is sensitive varies between different people, and an automated system can never predict with 100% certainty whether data needs protection or not.

Figure 5. Blacking out windows

6.3 Implicit: Silhouette Protection

Because the system recognizes the spatial relationship between the passerby, the user and the display, it can roughly calculate what part of the display is shielded from view by the user's body (Figure 6 and 1b). It can then use that calculation to black out (again via appropriate transparency levels) the areas of the screen visible to the passerby, while leaving the area shielded from view visible to the user. That is, if we consider the passerby as an inverse light source, the user working on the display casts a 'shadow of visibility' onto the screen (Figure 6 green line), which we call a *silhouette*. The rest of the screen becomes muted using appropriate transparency levels, where it too tries to strike a balance between hiding content from the passerby's view (Figure 6 red line) while keeping it somewhat accesible to the user. The silhouette disappears entirely when the user turns away from the display, leaving a black screen behind.

The animated silhouette moves when either the user or the passerby moves, reflecting the changes in the area that would otherwise be visible to the passerby (Figure 6 and 1b). The size of the silhouette changes as a function of the passerby's distance: With decreasing distance the size of the visible area decreases, reflecting the smaller inverse shadow of visibility cast by the user on the display.

The silhouette is calculated by creating a vector based on the position of the user and a passerby (Figure 6, green line). We extend the vector to get the intersection point with the display, which we use as the center-point of the silhouette. The silhouette's width (Figure 6, white area) is a function of the distance between the user and a passerby. The vertical position and height of the silhouette on the display is based on the sensed height of the user.

Discussion: Unlike the 'blacking out of sensitive content' approach, the system does not need to know what content is private *vs.* public. The silhouette acts on a physical metaphor, where it covers only those parts of the screen that can be overseen by a passerby. As with the animated 3D-model, the visuals are easy to understand by both user and passerby, making them both aware of possible intrusions. It also tries to minimize interruption, as the user can continue to work on the visible area (which typically remains in front or close to one's body). Even so, because part of the

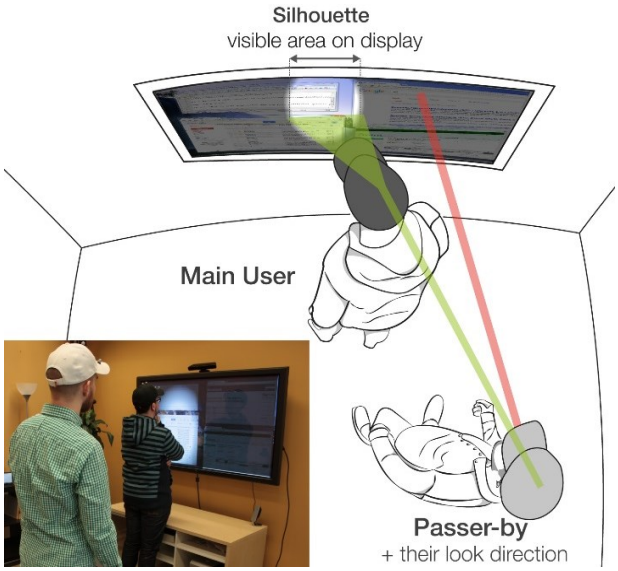

Figure 6. The silhouette reveals only those parts of the display shielded from view by the user's body, allowing him to continue his work; its area is calculated as a function of the vector between the passerby, the user, and the display.

screen's content is muted (especially if the passerby moves close to the display), it becomes more difficult for the user to employ the full display for his work.

7. DISCUSSION AND CONCLUSION

In this paper, we described how we can mitigate shoulder-surfing issues on public displays. All our methods are based on sensing the position, distance, and orientation between people and their environment, which in turn helps us calculate and build upon social notions of proximity and territorial incursions. Our techniques provide varying degrees of mutual awareness to allow user and passerby to engage in social protocol, where awareness helps them self-regulate their behaviors and/or negotiate their consequential actions. They also provide some degree of protection of sensitive information. We do not claim that our protection mechanisms are entirely secure. Rather, they are useful to temporarily protect sensitive information from a passerby who happens to glance at the display, where again we would expect social protocol to stop any serious attempt to breach one's privacy.

Our methods are suggestive of a broader range of other approaches. For example, the notion of indicating the presence and position of a passerby can be realized via many other cues (e.g., different visualizations, 3d sound). Cues can also be constructed to match the fidelity of the sensed information. Similarly, the idea of offering protection by masking information from view on the display can take many visual forms. Design trade-offs will include how understandable the cue is to all parties, the degree of awareness provided by these cues, the distraction caused by the cue or protection mechanism, the degree of security provided, and the amount of effort required by the parties to either explicitly control the system or override the implicit actions taken by the system.

To this point, our explorations have considered only the case of a single passerby and a single user of the display. Thus they are likely appropriate for non-crowd situations where only occasional people pass by. Still, we believe that some of our approaches are somewhat scalable to include a few passers-by. For example, we can include 3D-models and gaze indicators of all people in the scene, or calculate the silhouette size and position as a function of multiple vectors representing each person. Again, there are tradeoffs. For example, the silhouette would shrink considerably or even disappear because there may be no display area that would be completely shielded from at least one person's view by the user's body (especially if passers-by are far apart). This can be remedied somewhat by weighting in the passers-by viewing orientation, where we can leave out of the calculation those passers-by that are currently not looking at the display.

Finally, we recognize that our systems are exploratory prototypes. They currently rely on a high-fidelity (and expensive) motion tracking system that requires people to wear markers. This is clearly not deployable in the wild. However, alternate low-cost technologies can be used instead. For example, the marker-less Kinect 2 can provide almost all the required information, including body motion tracking (via skeletons) and gaze orientation (via facial recognition), and by analyzing skeletal features such as shoulder width).

Acknowledgements. Funding was partially provided by the NSERC / AITF / SMART chair in Interactive Technologies and by the NSERC Surfnet Networks Grant.

REFERENCES

1. Alt, F., Shirazi, A.S., Kubitza, T., and Schmidt, A. Interaction techniques for creating and exchanging content with public displays. *Proc. CHI '13*, 1709–1718.
2. Altman, I. *The environment and social behavior: privacy, personal space, territory, and crowding.* 1975.
3. Ballendat, T., Marquardt, N., and Greenberg, S. Proxemic interaction: designing for a proximity and orientation-aware environment. *Proc. ACM ITS*, (2010), 121–130.
4. Bellotti, V. Design for privacy in multimedia computing and communications environments. *Technology and privacy: The new landscape*, (1997), 63–98.
5. Berger, S., Kjeldsen, R., Narayanaswami, C., Pinhanez, C., Podlaseck, M., and Raghunath, M. Using symbiotic displays to view sensitive information in public. *Proc. IEEE PerCom, 2005*, IEEE (2005), 139–148.
6. Boyle, M. and Greenberg, S. The language of privacy: Learning from video media space analysis and design. *ACM Trans. CHI* , 2 (2005), 328–370.
7. Brignull, H. and Rogers, Y. Enticing people to interact with large public displays in public spaces. *Proc. INTERACT*, (2003), 17–24.
8. Freedman, E.G. and Sparks, D.L. Coordination of the eyes and head: movement kinematics. *Experimental brain research 131*, 1 (2000), 22–32.
9. Greenberg, S., Boyle, M., and Laberge, J. PDAs and shared public displays: Making personal information public, and public information personal. *Personal Technologies 3*, 1 (1999), 54–64.
10. Hall, E.T. *The hidden dimension*. Anchor Books NY, 1969.
11. Harrison, C. and Hudson, S.E. A new angle on cheap LCDs: making positive use of optical distortion. *Proceedings of the 24th Annual ACM Symposium on User interface Software and Technology*, (2011), 537–539.
12. De Luca, A., Von Zezschwitz, E., and Hußmann, H. Vibrapass: secure authentication based on shared lies. *Proc. ACM CHI*, (2009), 913–916.
13. Marquardt, N., Diaz-Marino, R., Boring, S. and Greenberg, S. The proximity toolkit: prototyping proxemic interactions in ubiquitous computing ecologies. *Proc. ACM UIST*, (2011), 315–326.
14. Müller, J., Walter, R., Bailly, G., Nischt, M., and Alt, F. Looking glass: a field study on noticing interactivity of a shop window. *Proc. ACM CHI*, (2012), 297–306.
15. Nickel, K. and Stiefelhagen, R. Pointing Gesture Recognition based on 3D-Tracking of Face, Hands and Head Orientation Categories and Subject Descriptors. *Proc. ICMI*, (2003).
16. Palen, L. Social, individual and technological issues for groupware calendar systems. *Proc. ACM CHI*, (1999), 17–24.
17. Peltonen, P., Kurvinen, E., Salovaara, A., et al. It's Mine, Don't Touch!: interactions at a large multi-touch display in a city centre. *Proc. ACM CHI*, (2008), 1285–1294.
18. Scott, S.D., Carpendale, S., and Inkpen, K.M. Territoriality in collaborative tabletop workspaces. *Proc. ACM CSCW*, (2004), 294–303.
19. Sharp, R., Madhavapeddy, A., Want, R., and Pering, T. Enhancing web browsing security on public terminals using mobile composition. *Proc. MobiSys*, (2008), 94–105.
20. Stahl, J.S. Amplitude of human head movements associated with horizontal saccades. *Experimental brain research 126*, 1 (1999), 41–54.
21. Tan, D.S. and Czerwinski, M. Information voyeurism: Social impact of physically large displays on information privacy. *Extended Abstracts ACM CHI*, (2003), 748–749.
22. Vogel, D. and Balakrishnan, R. Interactive public ambient displays: transitioning from implicit to explicit, public to personal, interaction with multiple users. *Proc. ACM UIST,* (2004), 137–146.