Exploring Interactions with Smart Windows for Sunlight Control

Patrick Bader
University of Stuttgart
Stuttgart, Germany
patrick.bader@vis.uni-stuttgart.de

Niels Henze
University of Stuttgart
Stuttgart, Germany
niels.henze@vis.uni-stuttgart.de

Huy Viet Le
University of Stuttgart
Stuttgart, Germany
huy.le@vis.uni-stuttgart.de

Julian Strotzer
University of Stuttgart
Stuttgart, Germany

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the Owner/Author. Copyright is held by the owner/author(s).

CHI’17 Extended Abstracts, May 06–11, 2017, Denver, CO, USA.
ACM 978-1-4503-4656-6/17/05.
http://dx.doi.org/10.1145/3027063.3053242

Abstract
Window facades play an increasingly important role in modern architecture. Regular shutters and blinds allow only coarse control over the sunlight coming through windows. Smart windows using see-through displays can be controlled on a per-pixel basis and thereby have the potential of fine-grained control. In this paper, we explore future interaction with such smart windows and conducted an elicitation study with 16 potential users. We provide both a mid-air gesture set and a smartphone interface to define regions for glare protection and brightness control. The study was conducted on a working $1.6 \times 2.6 \text{ m}$ smart window prototype with $130 \times 144$ individually switchable pixels.

Author Keywords
Smart windows; gesture elicitation; guessability; user-defined; mid-air

ACM Classification Keywords
H.5.2 [User Interfaces]: Interaction Styles.

Introduction and Background
Windows play a central role and make up large parts of modern building facades. Thus, sun protection becomes an increasingly important aspect of user comfort in such buildings and influences room climate. Traditional glare protection systems like window blinds or shutters use mechanical
components, which are often controlled manually to block sunlight. From a technical perspective, such systems are prone to failure due to environmental influences, like strong winds, and due to wear-out. Furthermore, users have to decide between brightness in the room and glare protection.

Recently, various applications of see-through displays have gained interest by the research community. See-through displays have been used for layered 3d display devices [3], back-of-device interaction [2] and augmented reality [8]. Also, see-through liquid-crystal (LC) panels can be integrated into window double glazing, see Figure 1. Such smart windows allow users fine-grained control over the shading of arbitrary parts of the window. With short response times compared to e.g. shutters, they can directly react to environmental changes and decouple glare protection from brightness control.

**Smart Windows**

Previous work on smart windows focused on adaptation to environmental changes in the context of climate control and sustainable living. The smart window controls the amount of sunlight and heat which is propagated into a building [4, 9] on a per-tile basis. In this way, user comfort is increased and energy consumption for climate control is reduced.

Rekimoto [11] presented two use cases for interacting with smart windows. In real-world pixelization, users’ location is tracked and sight on specific body parts is blocked. In programmable shadows, users can define regions in the room for which the smart window casts shadows based on the sun’s location. However, previous work does not define how interaction with smart windows is realized. As people typically interact with windows located in the same room and may be spontaneous, interaction can be realized through a gesture interface.

**Gesture Elicitation**

Gesture interfaces enable interaction at short distances and have gained increasing interest since consumer devices like the Kinect or Wii with support for gesture control hit the market. Wobbrock et al. [16] presented a methodology for deriving gestures from users in elicitation studies. In these studies, users are shown the effects of gestures (called referents) and are asked to come up with corresponding gestures that would cause the effect. Gestures are assigned to each referent based on agreement. The notion of agreement was extended and formalized in more recent work [14] which allows statistical tests to be performed on agreement.

Gesture elicitation has been used in various fields including mobile interaction [12], augmented reality [10], smartwatches [1] and music playback [7]. In the context of large displays, gesture sets were elicited for TV control [5, 13]. Wittorf et al. [15] elicited gestures for wall-displays and found that they tend to be more physically-based and larger for large displays whereas hand posture is less important. Referents were largely related to manipulation tasks (13/25) in the context of application windows. Spontaneous interactions with smart windows in homes which may not be perceived as displays by users have not been covered.

**Summary**

Previous work investigated application scenarios for see-through displays including smart windows. Gestures have been elicited for various devices ranging from smart watches to large displays. However, interfaces for smart window interactions are largely unexplored. We derive a mid-air gesture set for glare protection and a corresponding smartphone interface from an elicitation study. The former for spontaneous and the latter for distant interaction. The study was conducted on a functional 1.6 × 2.6 m smart window prototype with 18720 individually controllable segments.
Elicitation Study
We conducted a study to investigate two complementary explicit interaction modalities for defining and manipulating rectangular sunlight blocking regions on smart windows. First, a gesture interface intended for spontaneous interactions in front of the window. Second, a smartphone interface for interactions where the user does not have to be located directly next to the window. We followed the method for elicitation studies introduced by Wobbrock et al. [16] and used the metrics and AGATe toolkit for data analysis introduced by Vatavu et al. [14] for both interfaces.

Participants
Sixteen participants (14 males, 1 female and 1 unspecified), ranging in age from 21 to 33 years (M=26, SD=2.5) volunteered in our study. One participant was left handed. We obtained informed consent from each participant.

Referents
We defined 21 basic actions which users can perform to manipulate sunlight blocking regions on the window, see Table 1. They represent basic actions that are either creation, size or transparency related. The gesture interface consists of another category called gesture delimiter for enabling and disabling gesture detection to prevent unintended input.

Design & Procedure
We used a within-subjects design and asked participants to perform gestures and make suggestions for the smartphone interface. The study took 45 minutes on average and participants received some sweets. First we briefed participants on the topic and on the procedure and demonstrated the basic operation of the smart window. Then participants filled out a consent form and a background questionnaire.

We counterbalanced the order of the interaction approaches which participants used first. Referents were shown in randomized order. However, referents in the gesture delimiter category were always shown last and only for the gesture interface. This eliminated priming participants on technical limitations for the other gestures.

For each referent, the initial state of the system before the action was shown. Participants were told which action the system will perform. Then, the resulting state was shown, see Table 1. Transitions between the states were not shown to remove bias towards gestures that mimic specific transitions. For the gesture interface, participants were asked to perform the actual gesture. In case of the graphical smartphone interface, participants were asked how they would perform each action. They were provided with pictures of all input controls available on Android on a sheet of paper. Participants could answer verbally (which we audio recorded) or draw sketches on paper. The study closed with a questionnaire about their overall opinion on such systems.

Table 1: Overview of referents for both interfaces. Each referent is depicted in its initial state (left icon) and target state (right icon). Grey arrows and colors (red) are for illustration purposes and were not visible to the participants.

<table>
<thead>
<tr>
<th>Category</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Creation</td>
<td>R1 Create</td>
</tr>
<tr>
<td></td>
<td>R2 Delete</td>
</tr>
<tr>
<td></td>
<td>R3 Select</td>
</tr>
<tr>
<td></td>
<td>R4 Deselect</td>
</tr>
<tr>
<td></td>
<td>R5 Move</td>
</tr>
<tr>
<td>Size</td>
<td>R6 Enlarge Top</td>
</tr>
<tr>
<td></td>
<td>R7 Shrink Top</td>
</tr>
<tr>
<td></td>
<td>R8 Enlarge Bottom</td>
</tr>
<tr>
<td></td>
<td>R9 Shrink Bottom</td>
</tr>
<tr>
<td></td>
<td>R10 Enlarge Left</td>
</tr>
<tr>
<td></td>
<td>R11 Shrink Left</td>
</tr>
<tr>
<td></td>
<td>R12 Enlarge Right</td>
</tr>
<tr>
<td></td>
<td>R13 Shrink Right</td>
</tr>
<tr>
<td></td>
<td>R14 Scale Up</td>
</tr>
<tr>
<td></td>
<td>R15 Scale Down</td>
</tr>
<tr>
<td>Transparency</td>
<td>R16 Increase Transparency</td>
</tr>
<tr>
<td></td>
<td>R17 Decrease Transparency</td>
</tr>
<tr>
<td></td>
<td>R18 Window Opaque</td>
</tr>
<tr>
<td></td>
<td>R19 Window Transparent</td>
</tr>
<tr>
<td>Gesture Delimiter</td>
<td>R20 Start Detection</td>
</tr>
<tr>
<td></td>
<td>R21 Stop Detection</td>
</tr>
</tbody>
</table>

1Individual actions are called referent in gesture elicitation studies.

References
We defined 21 basic actions which users can perform to manipulate sunlight blocking regions on the window, see Table 1. They represent basic actions that are either creation, size or transparency related. The gesture interface consists of another category called gesture delimiter for enabling and disabling gesture detection to prevent unintended input.
Results
Participants performed a total of 104 distinct gestures for the gesture interface. For the smartphone interface, 72 distinct interactions were described. None of these were mid-air gestures performed with the phone. We report on agreement between participants and present a taxonomy for both interfaces.

Taxonomy
We categorized actions performed with both interfaces using a unified taxonomy as shown in Table 2. This taxonomy combines and extends two taxonomies from previous work on surface computing [16] and mid-air gestures for people that are blind [5]. It defines four dimensions which apply to both interfaces and three dimensions which apply to either interface. We directly applied the dimensions nature, flow [16] and axes of motion [5] from previous work.

We redefined the binding category which describes how a gesture or action relates to its referent. If the binding is absolute, changes in hand/finger position during the action directly map to changes on the window e.g. the user grabs the right edge of a region with one hand and moves the hand to the right, the edge will move in the same direction as if the user is actually holding the edge. In contrast, if the binding is relative changes in hand/finger position only indirectly map to changes on the window e.g. a user increases transparency by performing a clockwise rotation with her hand. Actions which do not relate hand/finger motion to changes on the window are categorized as arbitrary actions. A user deletes a region by waving his hand, for example.

We added two similar dimensions for the mid-air gestures and smartphone actions regarding number of hands or fingers used. One-handed gestures may be performed in encumbered situations and may also allow the combinations of two gestures at the same time. For the smartphone actions, single finger interactions are easier to perform when holding the phone one-handed.

For the smartphone action set, another dimension regarding the target on the touchscreen was included. Actions can be performed on a preview of the window on the smartphone display. These actions are typically surface gestures. Alternatively, users can use a default UI-component, like a button or slider to perform an action.

Categorization of Mid-Air Gestures
Mid-air gestures were mostly performed with both hands (57.7%) and flow was continuous (67.3%). The nature of the majority of all gestures was physical (65.4%) especially for the size related referents (93.8%). Symbolic gestures were only used to start or stop detection (21.4%). Other gestures for delimiter were either abstract (50.0%) or metaphorical (28.6%). Number of hands used was highly related to whether the binding was relative and absolute ($\chi^2(2, N = 82) = 30.03, p < .001$). Gestures with relative binding were primarily performed with both hands (86.7%) whereas gestures with absolute binding were performed with one hand (73.0%)

Categorization of Smartphone Interaction
In contrast to the number of hands used for mid-air gestures, most actions for the smartphone interface were performed with a single finger (70.8%) and their binding was mostly absolute (58.3%). There was no clear preference in the nature between physical (29.2%), symbolic (22.2%), metaphoric (33.3%) and abstract (15.3%). However, flow of actions was primarily continuous (70.8%). We found statistically significant relations between target and flow ($\chi^2(2, N = 72) = 27.45, p < .001$). Participants used UI-components (66.7%) primarily for discrete actions and the preview (92.2%) for continuous actions.
Agreement Analysis
We analyzed agreement based on the agreement rate $AR$ introduced by Vatavu et al. [14]. The overall agreement was $AR = .203$ for mid-air gestures and $AR = .439$ for the smartphone interface. Individual agreement rates for mid-air gestures are shown in Figure 2 and Figure 4 shows actions for the smartphone interface. Referents which were size related achieved comparatively high agreement rates for the gesture $AR = .340$ and smartphone interface $AR = .535$. Participants had difficulties to come up with transparency related gestures, therefore agreement was low.

A Mid-Air Gesture Set for Smart Windows
We derived a gesture set based on users’ agreements, see Figure 3. We selected the gestures with highest agreement for each referent. All but three gestures (create, select, deselect) were performed with two hands and participants held their hands open while performing gestures. Participants “pushed” towards the window (25%) to create new regions and “threw the region away” (25%) to delete it. Most participants (63%) moved both hands diagonally apart or towards each other to scale regions up and down.

To enlarge a region in one directions, participants held both hands in front of them and moved only one hand away from the other. Participants performed inverse gestures to shrink regions, starting with both hands apart and moving one towards the other. We illustrate only the enlarge and shrink gestures for the top edge of a region in Figure 3. Gestures in other directions were performed analogously.

Most agreement (31%) for start and stop detection was achieved with the clap gesture. Referents in the transparency category did not have statistically significant agreement and were thus excluded from the gesture set.

A Smartphone Interface for Smart Windows
We derived a set of actions for a graphical smartphone interface also based on maximum agreement per referent. The action set is depicted in Figure 5. Five participants chose a pinch out gesture to create new regions and most participants (75%) dragged a region out of the preview to delete it. Similar to the gesture set, selection was done by tapping on a region (81%). However deselecting a region was done by tapping on a free area (63%).
Most participants (75%) swiped to either move a region or enlarge/shrink the respective edge of a region. Only enlarging and shrinking a region at the bottom is depicted in Figure 5, other directions were changed analogously. To scale up or down uniformly, participants (63%) pinched opposite corners with two fingers. Create and scale up/down were the only actions which were performed with two fingers simultaneously. Participants chose UI-components for all transparency related referents.

**Discussion**

The smartphone interface achieved a higher overall agreement compared to the gesture interface. There are two reasons for this: Users are accustomed to using smartphones in their daily lives with specific guidelines for the user interface. Mid-air gestures offer more degrees of freedom and thus more variation in actions can be expected. Yet, participants had comparatively high agreement on size related referents with both modalities. This suggests users have a clear mental model of how to change size of 2d objects.

Mid-air gestures for the scale up and down referents received highest agreement whereas agreement for the smartphone interface was only average. One reason could be that users seldom perform two finger gestures as zooming can mostly be done by double tapping.

We had to drop all transparency related referents from the gesture set due to insufficient agreement among participants. One reason for this could be that transparency is a more abstract concept than physically moving objects and users have no clear mental model for it. Furthermore, users typically do not adjust transparency outside graphics software. Results from the smartphone interface support this assumption as all transparency related actions are represented via abstract UI-components like sliders and buttons.

The mid-air gesture set mostly consists of two-handed gestures with relative binding. In contrast, the graphical smartphone interface is usable with one hand and binding is absolute. Participants also did not focus on specific hand postures and some participants performed corresponding gestures with alternating hand postures. This means that an implementation can neglect hand posture for the majority of gestures. Not only does this simplify the implementation it also allows to perform gestures when holding objects.

**Conclusion & Future Work**

In this paper we present the idea of controlling glare protection for smart windows using mid-air gestures or smartphones. We derived a set of mid-air gestures and a graphical smartphone interface for basic smart window interactions from an elicitation study. This work is a first step towards the vision of houses where architecture, especially window locations and shape, may be defined interactively by inhabitants and is not defined statically by an architect.

In a next step, we plan to perform studies to evaluate our gesture set and to receive feedback on users’ perception and usability. Our system may also be extended to allow more precise control over the smart window and increase flexibility by supporting free-form regions and interacting with objects in the room for indirect glare protection.

**Acknowledgements**

This work has partly been funded by the Baden-Württemberg Stiftung gGmbH, Germany in the program “Nachhaltiges Bauen” (sustainable building). The hardware we used has been funded by the Federal Institute for Research on Building, Urban Affairs and Spatial Development, Germany, in the project “TN-Technologie für den Einsatz in Architekturverglasungen” (TN-technologies for use in architecture glazing), grant number: SWD -10.08.18.7-13.26.
References


