Embodied lenses for collaborative visual queries on tabletop displays

KyungTae Kim

Niklas Elmqvist*

Abstract

We introduce embodied lenses for visual queries on tabletop surfaces using physical interaction. The lenses are simply thin sheets of paper or transparent foil decorated with fiducial markers, allowing them to be tracked by a diffuse illumination tabletop display. The physical affordance of these embodied lenses allow them to be overlapped, causing composition in the underlying virtual space. We perform a formative evaluation to study users' conceptual models for overlapping physical lenses. This is followed by a quantitative user study comparing performance for embodied versus purely virtual lenses. Results show that embodied lenses are equally efficient compared to purely virtual lenses, and also support tactile and eyes-free interaction. We then present several examples of the technique, including image layers, map layers, image manipulation, and multidimensional data visualization. The technique is simple, cheap, and can be integrated into many existing tabletop displays.

Keywords: Magic Lenses, focus+context, embodied interaction, tangibles.

1 Introduction

Why are visual representations effective? Cognitive science suggests that such visualizations off-load working memory, rerepresent data in a form amenable to analysis, and constrain reasoning in a process known as *external cognition* [43]. However, external cognition alone is not sufficient to account for other phenomena commonly employed in sensemaking [42] such as interaction [58, 38], physical manifestations [22], spatial and temporal arrangements [21, 40], and social and collaborative aspects [20]. In this paper, we propose the notion of *embodied lenses* that begin to capture these aspects of *embodied interaction* [9] for interactive queries on digital tabletop displays [8, 16].

Our approach is similar to Spindler et al. [49] and Koike et al. [28] in that we bridge virtual objects on the tabletop display with lens manifestations in physical space. However, we continue the physical metaphor further by allowing users to overlap the lenses in physical space to control object composition in virtual space. The concept is simple: we use sheets of normal paper or transparent foil decorated with fiducial markers that can be recognized and tracked by the tabletop. These sheets can now serve as embodied lenses that can be moved around on the tabletop, composed with other physical lenses, and decorated with interface controls (Figure 1). The lenses are truly *embodied*: (i) they have physical form; (ii) they support natural physical affordances [31] for moving, rotating, and overlapping; and (iii) they serve as social indicators of presence and current activity for other collaborators.

Imagine a user exploring real estate using a Google Maps application on a tabletop display. An embodied lens would allow the user to easily filter the available real estate according to cost, square footage, acreage, and number of bedrooms without affecting the global display, thus not interfering with other concurrent users of the real estate application. While not necessarily faster than an entirely virtual lens, the embodied lens has a physical form with many and well-understood benefits [22, 54], including haptic feedback, eyes-free operation, and clear physical affordances. Furthermore, our technique for embodied lenses is exceedingly simple, cheap, and flexible: the sheets can be easily printed on a normal laser printer, cut into desired shape with a pair of scissors, and registered for use on the tabletop surface. In addition, transparent foil has the added advantage of allowing touches to be recognized inside its extents as well as allowing lenses to be nested inside other lenses.

To inform our design, we conducted a formative evaluation with human subjects on spatial filtering to derive their conceptual model for composing physical lenses. Based on these findings, we then explore the design space of embodied lenses made out of thin sheets of paper or transparent foil. We then conducted a formal user study comparing user performance for placing different configurations of virtual versus embodied lenses on a multitouch tabletop. Our results indicate that embodied lenses have comparable performance with virtual lenses with no difference in accuracy, and several potential benefits in relation to the tangible and eyes-free nature of the lenses.

In the second part of the paper, we showcase the technique in several examples, starting with simple focus+context lenses followed by filtering in Google Maps, as in the example above. Image manipulation applications, while not strictly within the

^{*}School of Electrical & Computer Engineering, Purdue University, 465 Northwestern Ave, West Lafayette, IN 47907, USA, E-mail: elm@purdue.edu.



Figure 1: Tangible map application with two users exploring the U.S. Midwest region on a tabletop display using embodied lenses made out of transparent foil. Amoeba fiducial markers [3] placed under rubber handles (repurposed mouse mats) at the corners of each embodied lens allow the tabletop to track the position and rotation of each lens. Overlapping the embodied lenses in physical space combines queries in virtual space.

concept of visualization, is another example where composable filters are common, and we show how our technique generalizes for this purpose. Finally, we demonstrate how our lenses can be used as movable filters in core visualization applications, such as in 2D scatterplots [7] and in parallel coordinate plots [18].

2 Related Work

Lenses are a common interface tool in graphical applications, but are traditionally virtual in nature. However, recent work is starting to make it possible to integrate tangible objects on horizontal displays. Below we survey the literature on these topics.

2.1 Lenses in Information Visualization

Lenses have a long history of use in visual applications, dating back to Furnas original fisheye views in 1986 [12], and work is still ongoing in this domain. Some examples of work in this domain include magic lenses [4], movable filters [50], and high-precision magnification lenses [37]. However, all of these lenses are virtual in scope and have no tangible component.

2.2 Tangible Interaction

A tangible user interface employs physical items for controlling digital information [22]. Research has shown that tangible objects increase the intuitiveness of interactions because of their physical affordance, haptic feedback, and potential for eyes-free operation. Much work has been performed since the original metaDESK [54] platform for tangible interfaces, such as [14, 24, 35, 56]. The work by Ullmer et al. [55] on tangible tokens for performing dynamic database queries is particularly relevant to our work.

2.3 Tangible Interaction on Tabletops

Digital tabletop displays are becoming available to a wide audience, but direct touch interaction in general is plagued by a lack of haptic feedback. To address this issue, recent work has looked at integrating tangible interaction with tabletops. One of the pioneering works in this domain is Sensetable [35], which uses wireless technology to integrate physical pucks with a front-projected tabletop display.

SLAP widgets [57] are silicone-based physical items—such as keyboards, sliders, and knobs—that are tracked by the tabletop using visual markers. They can be used to interact with the tabletop, and, by virtue of being transparent, can also be relabeled on the fly. The SLAP system uses a combination of DI [6] and FTIR [16] to make this hybrid virtual/physical approach possible.

Spindler et al. [47] use tangible props to detach user interface elements from the digital content on the tabletop display. Their solution is based on an additional overhead projector displaying these controls on a tracked sheet of cardboard.

Most recently, Luminos [1] allow for constructing physical 3D structures on the tabletop surface. This is achieved using clever arrangements of fiber optic bundles in tangible blocks that transmit light from the tabletop surface up into the block structure itself.

2.4 Tracking Objects on Tabletop Displays

One of the core problems of tangible integration on tabletops is recognizing and tracking physical objects placed on the horizontal surface. One solution is to use external components, such as wireless sensing [35], RFID [36], or magnetic tags [34]. These techniques are capable of capturing the location of objects, but cannot detect an object's rotation or whether it is touching the surface (although two sensors can be attached to an object to detect rotation [35]).

Computer vision is another promising solution for tracking objects, and it has been used with great success in the augmented reality field. ARTag [11] is capable of tracking unique and rotation-asymmetric visual markers using low-cost cameras—typically webcams—and has also been adapted to tabletop displays for tracking objects using an overhead camera. Another approach makes use of the camera already present in diffuse illumination [6] tabletops to track fiducial markers placed on the undersides of objects resting on the tabletop surface. Both of these approaches detect not only position, but also rotation information.

2.5 Tangible Lenses

Among the many uses of tangible objects, there have been a number of attempts to use tangible objects as lenses. Ullmer and Ishii [54] present tangible lenses as displays on articulated arms in their original metaDESK. SecondLight [23] uses a special diffused film to select the desired state of the display, such as choosing between one of two projected images. This technique allows a physical lens to acquire a different image than the surrounding area, but the system requires separate projectors for each display.

The Paper Windows system [17] uses a camera- and projection-based system to track sheets of tangible paper in a physical environment. Their approach allows for projecting arbitrary digital content on different sheets, and they exhaustively explore the design space of interaction techniques for such digital/physical uses of paper.

Lee et al. [29] utilize four light sensors to identify the corners of a rectangular surface. The technique uses relatively cheap light sensors, and requires only one projector. However, the tangible object needs power, and must thus be wired or battery-powered, and only one region can be active at a time.

Similarly, Looser et al. [30] propose a flexible and transparent tangible prop for use as a lens in augmented reality. Once again, however, the proposed lens is only intended to be used as a single lens, but their approach of marker-based tracking is similar to the technique proposed in this paper.

Another related system is UlteriorScape [25] where physical sheets are integrated with the digital content of a tabletop system.

Schmalstieg et al. [44] present two transparent props—a pad and a pen—that can be used for interaction with a 3D world on a stereoscopic tabletop display. The props can be used for a wide range of interactions, including tool palettes, dynamic controls, and magic lenses, but the system supports only a single transparent prop of each type at the same time.

On the topic of tangible lenses for visualization, Spindler et al. [48, 49] study the use of the 3D space *above* a tabletop, projecting dynamic imagery onto a tracked piece of cardboard using a projector. They propose to use this for volumetric data, layers, sampling, zooming, and slicing, etc, and their approach includes an interaction vocabulary for how these cardboard lenses can be used to explore the data. Similar to our work here, they also study scatterplot and parallel coordinate applications as well, and their work is therefore the closest to ours in terms of the visualization application area.

Finally, in terms of technology, the work that is closest to ours are the 2D transparent markers for tabletops proposed by Koike et al. [28]. These transparent markers are invisible to the human eye, yet can be detected by the tabletop using polarization of the underlying LCD display. The authors apply this to a magic lens application and show some examples on its use. Our implementation does not use an LCD tabletop and so must rely on visible markers, but this would be an interesting feature to add to our work. The primary difference with our work is that we focus on the physical and virtual affordances of lens composition, and we also study how to perform touch interaction inside a lens. Furthermore, our treatment includes examples of the utility of applying these lenses to visualization applications.

3 Background: Embodied, Distributed, and Situated Cognition

External cognition [43], where physical representations in the world are used to off-load, re-represent, and constrain internal cognitive processing, is often named as the main cognitive mechanism for perceptualization in general, and visualization in particular [5]. Donald Norman, who has long distinguished between knowledge in the head and in the world [32], draws parallels between graphical representations and the use of external representations in our physical world such as post-it notes, calendars, and pen and paper to extend our cognitive abilities, noting that "it is [physical] things that make us smart." [33].

However, the concept of external cognition is abstract and general in nature, and does not fully capture the situated and embodied aspects of physical items; for example, it is not clear how interaction fits into this model, and much work has been devoted to remediating this discrepancy [58, 38]. More specifically, pens, paper, post-it notes, and calendars all exist in the world as physical artifacts, can be interacted with, and can interact with their environment. For example, a pen can be used to write on paper, post-it notes, or calendars, and post-it notes can be affixed to the world (such as a wall or a monitor) or to other artifacts (such as on the pen). These physical arrangements also utilize our inherent spatial abilities as humans. Furthermore, all physical artifacts exist within a social context that is shared among several individuals [9]; for example, a post-it on a pen might be a label denoting ownership, while a post-it on a monitor might be a reminder about a future event.

Given this reasoning, it seems clear that external cognition alone is insufficient to account for the rich vocabulary of embodied and situated action that we should be able to draw upon for sensemaking [42]. In fact, ethnographic studies [21, 40] of both casual [39] and expert users alike working in both laboratory settings and "in the wild" indicate that the sensemaking process is a highly iterative and polymorphic workflow characterized by a plethora of external visual, tangible, and cognitive aids such as physical artifacts [22], annotations, multiple views [13], environmental cues [40], and arrangements in time and in space [21, 40].

In this work, we draw upon the nascent fields of *embodied cognition* and *embodied interaction* [9] as reasoning tools for discussing, designing, and evaluating tools for sensemaking. Embodied cognition takes the view that the human mind is shaped by the form and function of the body, and embodied interaction applies this to interaction design where the above physical and social aspects come together when building interactive systems situated in the world. In practice, this translates to designing sensemaking tools and representations that harness our innate human abilities to reduce the cognitive load and to facilitate the explicit processing from external to internal representations. More specifically, we derive the following three high-level requirements for such *embodied sensemaking* techniques:

- Physicality. The technique should be tangible and thus inhabit the world as an artifact in itself;
- Embodied. The appearance of the technique should give a indication of its affordances [31], i.e. its purpose and use; and
- Social context. The presence and usage of the technique should communicate the activity to other participants.

4 Design Framework: Embodied Sensemaking

In order to begin to explore these aspects of embodied sensemaking, we adopt interactive tabletop displays as an enabling platform. Tabletop displays are particularly suited for collaboration [41], are relatively cheap and easy to build (even by hobbyists) [16, 6], and are starting to be used for a wide array of co-located collaborative tasks such as tree comparison [19], data analysis [19, 52], and command and control [53]. Furthermore, tabletops based on diffuse illumination [6] can detect and track fiducial markers placed on the bottom of physical objects resting on the display surface.

However, supporting embodied sensemaking on tabletop displays gives rise to a number of design constraints. One of the most important is the need to *separate the displays of individual users* [52] to support both coupled and decoupled collaboration [51]. Furthermore, many applications utilize *global visual representations* that cover the entire display—such as maps, image collections, and graphs—making independent views in the style of standard GUI windows impractical and unsuitable.

Tabletop displays, for all their popularity, sport only virtual controls drawn graphically on the display surface. The primary benefit of such controls is *dynamic visual appearance*, allowing them to be easily moved and modified. However, by virtue of only having a visual appearance and no physical form, these virtual controls have no haptic feedback. Recent work [1, 57] highlights the benefits of combining the dynamic nature of virtual controls with the *physical affordances* of tangible controls integrated with the tabletop display.

4.1 Formative Evaluation

Our research question in this paper is to study novel methods for supporting this combination of physical and virtual spaces in the context of tabletop displays. Our particular approach is to overlap physical lenses to create *composition* in virtual space, a common operation in tasks such as image manipulation, data filtering, and map queries. To guide our inquiries, we conducted a formative evaluation on the physical affordances and conceptual model of these physical lenses for controlling object composition in virtual space.



Figure 2: Real estate interface for the formative evaluation implemented using JavaScript and Google Maps.

4.1.1 Participants

We recruited 6 unpaid participants (1 female, 5 male) from the student population at our university (average age 25). All were regular users of computers, all had experience of using touch devices, and two had used horizontal displays in the past.

We motivate the choice of university students as a representative population by the fact that the focus of our study was to elicit basic affordances and conceptual models for tangible objects on horizontal display, and therefore no particular expertise or knowledge was needed on behalf of the participants.

4.1.2 Methods

The evaluation platform was a 60-inch rear-projected tabletop display. The display showed a maximized Google Maps view of the city of Chicago (Figure 2). Because of the formative nature of the evaluation, we had not yet built a functioning embodied lens system to use as a testing platform. Instead, we used a Wizard of Oz protocol where an experimenter controlled a faceted search interface (right side in Figure 2) using a mouse in response to the participant physical actions. This also meant that touch input was disabled during the experiment.

Participants were given a sheet of trials for building queries in a fictional house hunting task. The map view showed a set of 100 icons for a randomly generated database of available houses, including data on their price, taxes, square footage, number of bedrooms, number of bathrooms, and acreage. The trials all consisted of forming queries to satisfy various constraints on these six dimensions: for example, a typical trial may have the form "Price \leq \$ 220,000 AND 2 \leq Bedrooms \leq 4." We organized trials in increasing order of difficulty as measured by the number of predicates involved. With 6 levels of number of predicates (1 through 6) and 3 repetitions per level, participants performed a total of 18 trials.

Each trial started with participants having access to six physical lenses (overhead transparencies) stacked in one corner of the table. Each lens was clearly labeled for which of the six house dimensions it filtered, and had a range slider (made out of

paper) for manipulating the selected range. Starting from the current trial on the task sheet, participants selected the relevant physical lenses, configured the range sliders, and overlapped them on the tabletop to search for candidate houses. In accordance with the Wizard of Oz protocol, the experimenter changed the dynamic query filters based on the participant's physical actions.

After completing a full session, we asked participants to fill out a post-test questionnaire consisting of both Likert-scale ratings (Table 1) as well as free-form interview questions (Table 2).

Metric	Question
Accuracy	Physical lenses are <i>accurate</i> for this task.
Analogies	This task is analogous to common physical actions I perform in the real world.
Collaboration	Physical lenses would be particularly useful when <i>collaborating</i> with others.
Composition	Combining physical objects in this way in the real world is can be used to achieve combined effects.
Efficiency	Physical lenses are <i>efficient</i> for this task.
Enjoyability	Physical lenses are <i>enjoyable</i> for this task.
Naturalness	Physical lenses are <i>natural</i> for this task.
Power	Physical lenses are <i>powerful</i> for this task.
Preference	I would <i>prefer</i> to use physical lenses compared to virtual lenses in my daily work.
Settings	There are some <i>settings</i> for which physical lenses work better than virtual lenses.
Tangible	The <i>physical form</i> of a lens is helpful in some situations.
Tasks	There are some <i>tasks</i> for which physical lenses work better than virtual lenses.

Table 1: Likert-scale (1-5) post-test questionnaire for the formative evaluation (strongly disagree to strongly agree).

Question

Q1	What does the physical	action of	overlapping	or stacking	two physical	objects ((such as	overhead
	transparencies) mean to	you?						

- Q2 Does the model used in this experiment—that overlapping sheets means combining filters—make sense to you?
- Q3 Can you think of any real-world situations where you find yourself overlapping physical objects (sheets of paper, playing cards, books, etc)?
- Q4 The traditional design alternative to physical lenses are virtual lenses that only exist on the computer screen and have no physical form. What are the strengths and weaknesses of each approach? Are there particular tasks, situations, or users for which one is better than the other?
- Q5 Can you think of any particular strengths or weaknesses when using physical lenses for collaboration with several people gathered around the tabletop display?

Q6 Do you have any particular thoughts or suggestions on how to design a user interface that mixes both physical objects (the lenses) and virtual objects (the graphical map on the screen)?

Table 2: Interview questions for the formative evaluation.

4.1.3 Results

Although we timed each trial, the completion time was not the primary outcome of our study (average completion ranged from 7.3 (1 predicate) to 51 (6 predicates) seconds and was largely linear, and a one-way RM-ANOVA analysis revealed a significant effect of predicate number on completion time: F(5, 25) = 286.3, p < .001). Rather, we were interested in qualitative results on the physical affordances of embodied lenses.

For the Likert-scale ratings, depicted in Figure 3, the results were uniformly high for the embodied lenses approach. Perceived accuracy was rated as lowest at 3.83 (s.d. .75), indicating that participants were concerned with the precision of a embodied lens. They also rated the power of this approach at 4.33 (s.d. .52), and preference to use it in their daily work was 4.16 (s.d. .41). On the other hand, all participants rated the utility of physical lenses for collaborating with others at 5, and its enjoyability received identical ratings. In particular, the naturalness of the approach was rated at a 4.83 (s.d. .41) average.

4.1.4 Interview Questions

Summarizing the free-text responses to our interview questions, all participants indicated that the physical action of overlapping two transparent objects means combining data, such as for two overhead transparencies, or retrieving more information, such



Figure 3: Mean responses to 1-5 Likert-scale questions (error bars show standard deviations) listed in Table 1.

as for a magnifying glass. On the other hand, one participant remarked that overlapping opaque physical objects is used in the real world for organizing information, such as when stacking books or sheets of paper, or even for hiding information, such as masking out pieces of an overhead transparency using a piece of paper. However, they all thought that stacking physical objects was a natural and intuitive physical operation for this task. Furthermore, the physical form meant that participants found the lenses easy to see, and memorizing their spatial location was also perceived as easy.

Some of the drawbacks mentioned include confusion about the logical operation resulting from composing lenses (AND vs. OR), scalability concerns when too many physical objects are being overlapped, and some issues with limited precision over virtual lenses. Participants also thought that a physical border, for example out of cardboard, would make the lenses less flimsy and more robust. Finally, as hinted at by the low Accuracy rating, some participants felt that the physical lenses may not be as precise as an entirely virtual lens, although this effect may well arise from the imprecision inherent in the Wizard of Oz protocol.

4.1.5 Observations

Participants required very little guidance on how to operate the physical lenses on the tabletop, and with no exception, all participants immediately grasped the general concept. Practically within minutes of encountering the experimental setup, all participants were constructing visual queries by configuring range sliders on the lens sheets and overlapping them to achieve combined effects. In general, this particular use of physical lenses seemed straightforward and intuitive to our users.

The physical lenses also provided a natural way to efficiently parse visual query tasks. When reading tasks descriptions, participants would immediately grasp the affected sheets corresponding to the query to form, pushing other sheets aside for the moment. The participant would configure the filters of each sheet separately and then instinctively arrange them in a stack that would be moved together as a unit on the tabletop. All of these observations were consistent with our hypotheses in regards to physical interaction on tabletop displays.

4.2 Design Constraints

We summarize the above design constraints as follows:

- C1 Multiple individual views. Effective collaboration requires individual views for each participant.
- C2 Global visual representations. Many canonical tabletop applications cover the entire display surface.
- C3 Physical affordances. Physical controls on tabletops provide intrinsic haptic feedback and eyes-free operation.
- C4 Dynamic visual appearance. Virtual controls on tabletops can be relabeled, moved, and changed on the fly.
- C5 Composition. Views should be composable to support advanced functionality (queries, filtering, selection).

To accommodate all of these constraints, we propose embodied lenses (C1, C2) that are tangible (C3) and composable (C5). These are made out of thin sheets of paper or transparent foil that can be easily tracked, and may thus also have a virtual representation (C4).

5 Embodied Lenses

Our embodied lenses for tabletop surfaces are physical sheets of paper or transparent foil decorated with fiducial markers. The physical nature of the lens allows it to be manipulated by the user through operations such as moving, rotating, overlapping, and even reshaping the lens (C3). The fiducial marker allows a diffuse illumination (DI) tabletop surface to track the lens and translate physical operations into virtual ones (C4). Together, these two features provide a way to bridge the physical and virtual space on the tabletop.

In this section, we explore the design space of embodied lenses. We start by discussing basic physical interaction and its virtual counterpart. We move on to discuss the visual representation of a lens—its interior—as well as visual decorations that can be added to its exterior. Overlapping lenses in physical space leads to lens composition (C5) in virtual space, a topic that renders a section of its own. We then discuss practical issues regarding physical materials.

5.1 Physical Interaction

The basic setup for our embodied lens technique is a diffuse illumination (DI) [6] tabletop display. These devices use computer vision in the infrared spectrum to detect objects on (or above) the horizontal surface that are being illuminated with invisible infrared light. The technology allows for tracking fiducial markers that are placed on the undersides of objects and are resting on the tabletop surface.



Figure 4: Physical interaction operations for embodied lenses.

We summarize the physical interaction of our lenses in Figure 4. Interaction is initiated by placing an embodied lens onto the tabletop surface so that its fiducial marker can be detected and identified (Figure 4(a)). Each lens must have a unique marker so that they can be reliably identified. Depending on the lens material (see below), the lens may now have to be *registered*, i.e., its extents must be defined. We support this by allowing the user to trace a finger around the perimeter of the desired lens extents, at the same time giving visual feedback of the boundary (Figure 4(b)). By lifting the lens from the surface, the lens is reset and its extents can be redefined (Figure 4(d)); however, depending on the software configuration, lenses can be toggled to remember their shapes. This functionality even allows the user to lift the lens, use a pair of scissors to cut the lens to a new shape, and then re-register the lens onto the surface.

Note that lens shape can easily be encoded into the marker-based lens identity, eliminating the need for registration altogether. However, registration does allow for changing lens geometry for situations where this is useful.

With a lens placed and registered on the display, it can now be moved and rotated by sliding it on the surface (Figure 4(c)). The tabletop tracks the position and orientation of the lens and updates the display, including the visual representation within the lens as well as its decorations. When several embodied lenses are registered on the display, they can be overlapped in physical space, causing application-dependent composition (C5) in virtual space (Figure 5). The only constraint is that fiducial markers must not be occluded by other markers from the point of view of the tabletop tracking system; we describe ways to avoid this below.

5.2 Visual Representation

The key feature for any focus+context framework is that the visual representation inside the lens extents can be different than the global view for the whole display (C2). This supports the local focus views of the data necessary for shared collaborative displays (C1), for example by filtering or changing the color scale of data inside the lens. From the perspective of the framework, a *lens* is a two-dimensional region identified by a unique fiducial marker. The marker conveys the position and rotation of the lens region. The lens then accesses the global visual representation within its extents and either transforms it (such as for



Figure 5: Composition of three lenses in virtual space. Overlapping areas represent those where a combination of lenses have effect.

an image processing filter), or replaces it with new visual content. Each lens propagates its contents upwards in the hierarchy, allowing for chains of lenses.

Exactly *which* visual representation to use depends on the application. Some applications will even use many different types of lenses, each with different visual representations. We will explore some examples of different lens types later in this paper.

5.3 Visual Decoration

Beyond the visual representation of the lens region, a lens can also choose to incorporate visual decorations, typically outside the perimeter of the lens. Again, the actual decorations depend on the application, but standard decorations include a title, a visual border, and an enable/disable toggle. If the lens is representing a filter, it is customary to add controls for changing parameters of the filter; for example, a range slider to select the data interval of items to show inside the lens [46], or a button to invert the selection.

Figure 5 shows three lenses with decorations similar to an interface window, including a title, a minimize button (to temporarily disable a lens), and a close button (to remove the lens).

Concrete embodied lens implementations may also use visual decorations such as buttons or tabs to control the composition operation (see the next section) that will be used when combining two lenses.

5.4 Lens Composition

Overlapping embodied lenses in physical space will cause the lens regions to be composed in virtual space (Figure 5). This is a key feature of the lenses, and enables a user to, for example, combine real estate filters from our earlier example on house hunting by simply overlapping the lenses in physical space on the table.

Definition: We define *lens composition* simply as basic function composition $f \circ g$, i.e., an associative application of one lens onto the output of another.

In an implementation, this amounts to simply finding the non-empty intersection of each combination of 2D regions for all lenses that are active on the tabletop. These combinations should now be applied to each of the intersections of the display space and transform that portion of the display space.

Depending on the capabilities of the hardware, an additional constraint may be that the lens composition must be commutative since it is often impossible to determine the stacking order (and thus application order) of individual lenses. One approach is to track the order that lenses were added to the tabletop, and then give explicit visual feedback and controls to change this order. Another is to track the dynamics—if a lens A is slid onto a lens B, we can reasonably assume that lens A is on top of B.

Finally, a complication arises if the fiducial markers (shown in Figure 5) of one lens happens to overlap the fiducial marker of another. In such situations, the tabletop computer will lose tracking of the occluded lens. We suggest two ways to deal with this situation:

- When tracking is lost on a lens whose fiducial marker is nested inside another lens, assume that the sheets representing the two lenses were collated into a neat pile and simply merge the functionality of the lenses; or
- Add tangible non-flat handles to each lens on top of the fiducial marker—the physical affordance of these handles suggest that they should not be overlapped with other handles.

5.5 Physical Lens Type

There are several different approaches we can use for material and physical design of the physical lenses. We outline the three major types of physical lenses below (Figure 6).



Figure 6: A tangible lens with a handle and dashed extents.

5.5.1 Paper Sheet

The most straightforward lens type is to use a sheet of paper cut into an arbitrary shape and decorated with a fiducial marker. Paper sheets provide physical and visual feedback on shape, position, and extents. Furthermore, because the lens surface appears opaque to the tabletop system, lens shape registration can be done automatically without user guidance.

We recommend using a slightly heavier paper stock to avoid some of the flimsiness of standard office paper. However, too thick paper may result in loss of tracking when overlapping lenses, and the rear-projected image may also not be able to penetrate the paper. Also, because paper is opaque, touch interaction inside its extents cannot be detected, and overlapping lenses may cause fiducial markers to be partially or completely occluded from the point of view of the tabletop's vision-based tracking system. Figure 7 shows an example of a back-projected image on a paper lens.



Figure 7: A paper lens in action for a Magic Lens [4] application, where the visual representation inside the lens is changed to a wireframe mesh.

5.5.2 Transparent Foil

Using transparent foil, i.e., transparencies used for overhead projectors, is perhaps the most flexible approach because transparencies, just like normal paper, can typically have its fiducial marker printed using a standard laser printer, yet its transparency allows for easily overlapping and nesting several lenses. This will only be problematic when the actual markers of two overlapped lenses happen to overlap, causing one to become occluded. Furthermore, the sheet still gives physical and visual feedback to the user.

Significantly, because the lens is transparent, users can still interact with the interior of the lens using standard multitouch gestures; finger touches are generally visible to the tracking system even through several overlapped lenses. On the other hand, because the lens extents are invisible to the tracking system, transparent lenses must be registered manually. Furthermore, it is often necessary to add a white background to the printed fiducial marker, i.e., by taping a piece of paper behind it, to facilitate tracking. A minor point is also that transparencies are typically made of cellulose acetate and thus prone to buildups of static charges, especially when rubbed against each other (such as when overlapping lenses).

5.5.3 Handle Only

Finally, the most straightforward approach is to just provide a physical handle with no spatial extents and with only a fiducial marker on its underside. The handle controls an entirely virtual lens, so there is no physical feedback on its extents. If the marker does not encode the lens geometry, manual registration of the virtual lens shape is naturally also necessary. On the other hand, this approach avoids issues in physical overlap of lenses on the tabletop surface, particularly if the handles are given some physical form beyond a thin piece of paper (e.g., a wooden brick).

6 User Study

While the purpose of our embodied lens framework is primarily to provide additional scaffolding for embodied sensemaking using visual queries, it would still be interesting to know how these embodied lenses compare to purely virtual lenses (i.e., lenses with only a graphical representation). Such findings would provide another data point for designers who are deciding on which visual query technique to utilize in a particular project. To this end, we conducted a quantitative user study comparing time performance using these techniques in a canonical visual query task. Since our goal was to compare the low-level affordances of these virtual and physical lenses, we chose to use an abstract scenario involving covering geometric regions on the visual space using different lenses.

6.1 Hypotheses

We advance the following three simple hypotheses:

- H1 *Fitting a new virtual lens will be faster than moving an existing virtual lens.* Given participants the capability to create, resize, and rotate a virtual lens out of thin air should yield faster performance than having to acquire and manipulate an existing virtual lens.
- H2 *Embodied lenses will be significantly faster than existing virtual lenses.* We think that the physicality of the embodied lenses will cause users to outperform existing virtual lenses in terms of completion time in the coverage task. However, we will not make any prediction on fitting new virtual lenses.
- H3 *Fitting a new virtual lens will have significantly less accuracy than the other two techniques.* As stated above, because fitting a new lens lets users explicitly select the size and not just the position and orientation of each lens, it will yield either significantly lower coverage **or** significantly higher error than the other two.

6.2 Participants

We recruited 12 (8 male, 4 female) unpaid participants from the general student population at our university (median age 24). Participants were self-selected, had significant computer experience (11 reported significant touch interaction experience), and had normal or corrected-to-normal vision with no color-deficiency (self-reported).

6.3 Apparatus

We conducted the experiment on an $1.2m \times 0.9m$ (approximately 60 inch) DSI [6] multitouch tabletop display equipped with with two DLP projectors, each with 1280×800 resolution (for a total of 1600×1280). The projectors were powered by a computer running Microsoft Windows 7.

6.4 Task

For the purposes of simplicity, we reduced the task studied in this experiment to its simplest components: an empty (white) canvas upon which a constellation of *query regions* are placed (Figure 8). The task entailed placing lenses (embodied or virtual) so that the query regions are fully covered. The objective was to minimize the time to complete the task while maximizing the covered query regions and minimizing the empty area of the space incorrectly covered by a lens.



Figure 8: Experimental setup for our quantitative user study. Participants were asked to maximize the coverage of a constellation of query regions using a set of pre-defined virtual or physical lenses, while minimizing the empty space incorrectly covered by these lenses.

There are several additional design decisions to consider for this simple experimental task. To ease virtual resizing of a lens, each of the query regions were rectangular in shape. In fact, because embodied lenses are not readily resizable (we did not allow participants to cut the lenses to size using scissors), we chose to restrict each of the individual query regions to four specific sizes: small (5×5 cm), medium (10×10 cm), large (15×15 cm), and extra-large (20×20 cm). In other words, participants were provided with exactly one lens corresponding to each query region size. We also restricted the number of query regions used in each trial as a factor N, giving it a value range of one, two, or three. Since there existed only one instantiation of each region size, this meant that for N = 3, all but one of the query region sizes were placed on the canvas.

Finally, region placement turned out to be key to achieving interesting scenarios. Since our tabletop is large, we first ensured that all placed query regions were easily reachable without moving from the participant's position along one of the long sides of the table, i.e., in a 0.75m semi-circle centered in the middle of one long side (this semi-circle was explicitly render on the table as visual feedback). Furthermore, since we wanted to study overlapping effects for the lens techniques, we placed query regions so that they overlapped with at least 25 to 75% of their surface area with one or more other lens. Placing an individual query region involved randomly selecting a center position and orientation so as to fulfill the above criteria.

6.5 Visual Query Techniques

The motivation for this experiment was to study the performance of embodied lenses in comparison to purely virtual lenses that exist only as graphical representations manipulated through multitouch interaction. However, depending on the task setup, we identify two separate ways to instantiate a virtual lens: either by fitting a new lens, or by moving (and rotating) an existing virtual lens to a specific location. For this reason, we include three separate techniques in the experiment:

• Virtual-sized: The most straightforward instantiation method is for a user to simply create a new lens by placing two fingers on an empty part of the space, thereby forming the two diagonal corners of a rectangular lens. Moving the fingers

allows for dynamically resizing and rotating the lens. Existing lenses can be modified—moved, rotated, and resized—by simply tapping and dragging on the borders of a virtual lens.

- Virtual-moved: As opposed to creating a lens out of thin air as in the VIRTUALSIZED technique, this technique would instead have the user move and rotate an existing lens by tapping and dragging on the borders of an existing lens from a *lens palette*. Similar to the pinch and drag operations outlined above, the difference with VIRTUALMOVED is that the lens cannot be resized. We provided virtual lenses in the four basic sizes outlined above.
- **Embodied:** This was the basic embodied lens technique as described in this paper, with four physical lenses in the four sizes outlined above. We chose to use paper lenses to maximize the accuracy of the DSI object tracking.

While the VIRTUALSIZED technique is likely the most natural and straightforward instantiation method for virtual lenses, it is also not able to capitalize on the fixed sizes of the query regions. However, it can certainly be argued that these fixed query regions are an artifact of our user study, but this is an unfortunate side effect of many tangible user interfaces that is outside the scope of this experiment. Nevertheless, we were interested to see what impact, if any, these different instantiation methods would have on task performance.

For the VIRTUALMOVED and EMBODIED techniques, we placed the lens palette in an area in the center of the table, i.e. at a distance of 0.75m directly in front of the participant's position (see Figure 8). For the embodied lens condition, participants were required to return all physical lenses to this area prior to starting a new trial.

6.6 Experimental Design

We used a full factorial within-participants design with the following factors:

- 12 participants
- 3 **Techniques** *T* (VIRTUALSIZED, VIRTUALMOVED, EMBODIED)
- \times 3 Region Numbers N(1, 2, 3)
- \times 4 repetitions (training excluded)
- 432 Total trials (36 per participant)

Trials were organized in blocks for each technique. Block order was balanced using a Latin square across participants to counteract learning effects; other factors were randomized within blocks. The experimental platform collected completion time as well as the accuracy of the placed lenses. Completion time was measured from when the participant initiated a trial by clicking a button until the participant clicked the same button to end the trial. Lens placement accuracy was measured in two ways: *coverage*, i.e. the ratio of the query region area covered by the lenses ([0, 1]), as well as *error*, i.e. the ratio of empty area incorrectly covered by a lens in regards to the cumulative query region area ($[0, \infty]$).

6.7 Procedure

Х

A user study session involved briefing each participant on the background of the study (2 minutes), introducing them to the task (1 minute), and letting them train on practice trials (5 minutes). Participants were instructed to perform each trial as quickly as possible, while still making sure to cover each query region accurately. After the participants indicated that they felt confident in each of the three techniques, they were allowed to proceed to the timed trials.

A trial started with an empty canvas and with any embodied lenses placed in the lens palette. Participants started the trial by pressing and holding a button in the center of the semi-circle for 1 second. The query region configuration was then shown on the visual canvas as grey boxes rendered on the white surface (Figure 8), and the task timer was started. Participants were now free to instantiate, place, and orient lenses to cover the query regions. No visual feedback was given on the coverage or task progress. When participants were satisfied with their coverage, they pressed another button to the side of the canvas semi-circle. This stopped the timer and calculated the coverage and error metrics for the trial. These metrics were silently recorded by the user study platform.

After having finished a full user study session, the participants were given a short exit interview. A typical user study session lasted approximately 30-40 minutes.

6.8 Results

Correctness results for coverage and error were all within 90% correctness, and there were no significant differences depending on technique or (surprisingly) number of lenses (repeated-measures analysis of variance, all assumptions valid). For this reason, we choose to disregard correctness for the remainder of this treatment.



Figure 9: Completion times per technique T and number of regions N for the quantitative user study.

Figure 9 shows a summary of the completion times for the experiment; EMBODIED lenses yielded an average completion of 3.57 (s.d. 1.35) seconds, VIRTUALSIZED yielded 3.81 (s.d. 1.46) seconds, and VIRTUALMOVED 5.65 (s.d. 2.24) seconds. We analyzed the completion time using a repeated-measures analysis of variance (all assumptions met), and found that there was a significant main effect of both technique T(F(2, 22) = 105.75, p < .0001) as well as number of regions N(F(2, 22) = 222.73, p < .0001) on this metric. Furthermore, there was significant interaction effects between both factors (F(4, 88) = 3.02, p < .0191).

We investigated pairwise differences between techniques using Tukey HSD, and found that both EMBODIED and VIRTU-ALSIZED were significantly faster than VIRTUALMOVED (p < .05). However, there was no significant difference between EMBODIED and VIRTUALSIZED.

6.9 Discussion

Our findings confirmed H1: VIRTUALSIZED was significantly faster than VIRTUALMOVED. Furthermore, the results allow us to confirm H2: as hypothesized, the EMBODIED technique was significantly faster than the VIRTUALMOVED technique. However, the results rejected H3: the fact that the VIRTUALSIZED technique required users to explicitly select the size of the lens did not yield any significant differences in accuracy compared to the other two lenses.

These findings yield two conclusions: First of all, that our virtual lenses were sufficiently responsive so that resizing the lens did not slow down the VIRTUALSIZED condition. In fact, our observations and the informal comments offered by participants seem to indicate that the ability to move, scale, and rotate a lens in a single multitouch gesture was a natural way of interacting with the lenses.

The second conclusion that can be drawn from our results is that our embodied lenses appeared to promote faster task performance than the equivalent condition with purely virtual lenses (VIRTUALMOVED), without being slower than the VIRTUALSIZED condition. Again, our observations offer some explanation: for the embodied lenses condition, participants often grabbed lenses with both hands when responding to a new trial, and would almost automatically rotate the lenses while sliding them to their destination. This is in contrast with the VIRTUALMOVED condition, where users would often move a single lens at a time (even if our tabletop had no such limitation), and would only rotate the lens once it had reached is approximate destination (again, despite our tabletop easily supporting both moving and rotating at the same time). We speculate that it is the physical and tangible feedback of the embodied lenses that cause this difference.

These same observations do offer some insight on why EMBODIED lenses may be a better fit than VIRTUALSIZED. First of all, even if we found no significant difference in our statistical analysis, the bimanual, tangible, and virtually eyes-free manipulation of the embodied lenses that we observed seem to suggest that the embodied lenses are more natural to use than virtual lenses. In the virtual-sized condition, each lens was created in a separate operation, whereas embodied lenses could theoretically be grabbed by the handful from the lens palette. Second, many visual query lenses in a real application setting must be persistent because they are decorated with user interface elements for type, filtering, and color scale assignment. In such settings, lenses are almost always pre-existing, and the manipulation is then concerned with acquiring and moving an existing lens.

The coverage task evaluated in this user study is a low-level task, and it could be argued that other emerging factors would manifest themselves in a higher-level visual query task similar to the one tested in our formative evaluation. While this is certainly possible, high-level visual query tasks are also more suspectible to systematic effects due to chance, individual differences among participants, and compound interaction tasks. In other words, the results from such a task would be significantly harder to analyze, not to mention interpret. We feel that the simple coverage task here is the canonical lens interaction task—even for high-level visual query tasks—and that our experiment therefore is more general and widely applicable when designed in this fashion.

One remaining issue to consider is whether to use the VIRTUALMOVED or VIRTUALSIZED lens instantiation approach in a real application. The latter technique is clearly much faster, and it is highly likely that the instantiation technique is the main cause of the difference. The pinch gesture used in VIRTUALSIZED allows the user to create a lens "out of thin air" and position it appropriately with a minimum of effort. For VIRTUALMOVED, the user must spend time acquiring and moving an existing lens from the lens palette at the edge of the screen, thereby consuming valuable time. However, as argued above, many lenses **must** be pre-existing and cannot simply be created anew every time. Therefore, VIRTUALMOVED can be considered to be a more realistic instantiation technique than VIRTUALSIZED. For this reason, we think that the findings from this study do indeed provide a useful recommendation for designers choosing between embodied or purely virtual lenses.

7 Application Examples

Beyond the elementary usage shown earlier in this paper, we implemented four applications to showcase our embodied lenses. In all of these examples, we use a hybrid FTIR [16]/DI [6] multitouch tabletop display with support for marker-based tracking. The FTIR functionality is used for accurate sensing of finger touches, whereas DI is employed for tracking lenses.

7.1 Multidimensional Data Visualization

Collaborative data visualization is a canonical example of the need for multiple individual views [52] to seamlessly support multiple modes of collaborative coupling [51]. We implemented two scenarios for parallel coordinates [18] and scatterplots [7].

Our lenses in each of these applications have additional decorations to allow filtering within a particular lens using dynamic queries [46] on a specified dimension. To allow the users to control the type of composition used, these visualization supports both AND and OR lenses encoded into the fiducial markers themselves. In other words, the user can explicitly choose different physical lenses depending on whether they want to get the union or the intersection when composing a lens with another lens.

7.1.1 Scatterplot

We implemented a scatterplot visualization capable of visualizing items as points using any two dimensions in the dataset by mapping them to the X and Y axes of the plot (Figure 10). We then integrated our embodied lenses as movable filters [50] supporting excentric labeling [10].

Each lens, once registered, is decorated with buttons to step through the dimensions of the dataset to filter on. Two sliders are also added to each side of a lens for changing the minimum and maximum of the range of items to show inside the lens (this is in lieu of a double-ended range slider, which is the optimal way of supporting dynamic queries [46]). Data points in the scatterplot falling inside the extents of an active lens will be filtered out (i.e., temporarily removed) if its value for the selected filter dimension falls outside of this range. Combining lenses by overlapping causes a conjunction of the filters in data space.

7.1.2 Parallel Coordinate Plot

Our second visualization example is a parallel coordinate plot [18] where all of the dimensions of a dataset are stacked in parallel and data items become polylines connecting the positions for each axis corresponding to the item's value for that axis (Figure 11).

Our implementation adds standard dynamic query sliders [46] to each axis to allow for global filtering affecting the whole display, but also supports our embodied lenses with movable filters for local queries more amenable to collaborative work. As with the scatterplot example, lens decorations allow for filtering data items that fall outside the range of a selected data dimension.



Figure 10: Scatterplot visualization [7] for two dimensions—gas mileage (X) and acceleration (Y)—of a car dataset. In this application—implemented using Adobe Flash/Flex—has lens decorations to control which dimension to filter on, and two sliders on each side of the lens area to control the upper and lower ranges of the filter.

7.1.3 Implementation

Our visualizations have been implemented using Adobe Flex and uses the TUIO [26] protocol. We import data into the tools as comma-separated files exported from a spreadsheet.

7.2 Layer Exploration Lenses

The layer exploration application (similar to that of Spindler et al. [48]) allows users to collaborative explore different layers of a visual space. The layers are partially transparent information substrates that are stacked to occupy the same visual space, and can be toggled on and off [2]. Users register a lens for a specific layer, and the lens will henceforth show only that layer in its contents. This can be applied to a wide variety of scenarios:

- For contractors studying building plans with different kinds of blueprints, such as water, electricity, sewage, etc;
- For car designers with different expertise collaboratively working on related systems in a vehicle (Figure 7); and
- For a team of doctors collaboratively diagnosing a patient with many layers of medical data (scenario below).



Figure 11: Parallel coordinate plot [18] for a car dataset. Each axis has range sliders to allow for axis filtering [45], and also visible are the horizontal range sliders decorating the two lenses. The interface allows for cycling through which dimension to filter using these range sliders.

In this fashion, the system allows users to collaborate in the same visual space while avoid conflicts and interference.

7.2.1 Implementation

We implemented the layer lens system in Java using Java 2D for graphics and using the TUIO [26] protocol for touch interaction. We provide a generalized file format for loading several layers of image data. A slider allows for changing lens transparency.

7.2.2 Scenario: Collaborative Medical Diagnosis

Imagine a situation where a team of doctors are diagnosing the victim of a traffic accident, trying to decide which surgery is most critical for the patient. Each doctor has different expertise and needs to look at different test results and scans of the body. However, the doctors must reach a consensus in a timely manner before they can decide on a course of action. In this scenario, we can add each view of the patient—such as X-rays, MRI imaging, and CAT scans—as a layer into the lens system (Figure 12). Each doctor gets their own lens that they can use to reveal interesting dimension of scanned image. The lens sheet gives the user a clear indication on how to use the system—important for non-expert use—as well as implicitly communicates the focus of each doctor to the others. Visual decorations on the lens allow the user to change settings, such as the layer to show, the transparency, rendering order, etc.

For instance, an orthopedic surgeon wants to see fractures on the patient's skeleton, so he takes his lens and registers for the skeletal layer (Figure 12(a)). This is done by simply placing the new lens on the tabletop and communicating the extents of the lens geometry to the system. The surgeon will now only see X-Ray images in his lens without interfering with the view of the other doctors. As he scans through the body, he finds a major fracture of the tibia, and decides that he needs to



(c) Skeletal and nerve lens composition.

(d) Embodied lenses in action for the anatomy application.

Figure 12: Medical scenario for the image layer system. Imaging and test results for a patient forms different layers to explore collaboratively.

discuss the damage with neurosurgeon to determine if it is critical or not. The neurosurgeon overlays her lens (Figure 12(b)) on the orthopedic surgeon's lens to acquire an overlapped image (Figure 12(c)). They both change the transparencies of their respective lenses by moving a slider attached to the lens to have the best view of the composition of the two layers.

7.3 Tangible Map Interaction

The Tangible Map enables users to view geographical regions and query for geospatial features using embodied lenses (Figure 13). In this application, the lens is used to add information to the map rather than filter it out. After placing a lens on the table, the user registers its extents and chooses a data dimension to view. Moving the lens on the table will now reveal the underlying geospatial data.

The operation to use for compositing lenses can also be controlled using the lens. More specifically, this controls whether the result of a lens composition will be the union or the intersection of the data shown in individual lenses.

7.3.1 Scenario: Planning a Biohazard Storage Site

Let us suppose that the state of Indiana is choosing a new site for biohazard storage. A set of experts are collaborating on this task using a tabletop display and our Tangible Map software (Figure 1 and Figure 13). One of them, a public health official, creates a lens for studying the population density. The lens is augmented with visual decorations to control the parameters of the lens. The health official sets the filter range to 0 to 5,000, since he wants to find areas that have a low density in population, thus minimizing the risk of any hazardous spill in the area. The lens shows values of counties satisfying this criterium using excentric labels [10] to efficiently display the names of places in a small region around his physical lens. Meanwhile, a water management engineer creates a lens for highlighting ground water aquifers. Again, the planning team wants to minimize lasting impact of any accidents involving hazardous waste. The new lens will show the position and density of water flow in its focus region. A third official, this one a transporation expert, uses his own lens to filter out roads with too small capacity for the task. Moving biohazard waste requires special heavy trucks escorted by lead and trailing vechicles, and the roads must be capable of supporting these convoys.

Working individually, these experts can use their embodied lenses to find candidate places on the map—potentially even creating a new embodied lens for each potential site they find. They can then start combining their individual lenses to find the small set of candidate sites that satisfy all criteria. The fact that the lenses are physical again helps the officials in instinctively knowing how to use the interface, and also supports implicit communication [15] between participants. They can even leave a lens on a promising location as an indication to other participants, and register a new lens to continue the exploration while



Figure 13: Collaboratively finding a biohazard storage site using the Tangible Map and embodied lenses (also see Figure 1). Each lens has been configured to only show sites that fulfill certain criteria: low population density, ground water aquifers, and high-capacity roads. Combining one or all of the lenses temporarily combines these filters.

awaiting a time to discuss it.

7.3.2 Implementation

We implemented the Tangible Map application in Adobe Flex using the Google Maps API. We again use TUIO [26] to communicate with the tabletop. The application fetches live map data from Google and displays them on the display, overlaying information from a geospatial database using map markers.

We used the public microsample of the U.S. Census 2000 data, containing anonymized data such as population, income, housing units, water area, and land area for geographical regions in the United States. We utilize a county dataset with longitude and latitude data to transform these values to positions on the map, and draw either a marker or a circle scaled by the data on that location.

7.4 Image Manipulation

We have implemented an image manipulation program using our embodied lenses. While image manipulation is not strictly an example of a sensemaking task, this example showcases the general nature of the technique. Just like in any image manipulation program, like Adobe Photoshop, our prototype supports image filters, but with the difference that our filters work inside particular lenses.

Figure 14 shows a photograph of the city of Chicago with three different lenses active—an edge detection lens, a smear lens, and a grayscale lens. Overlapping lenses causes the image filters to be composited in the same order the lenses were applied; in other words, sliding one lens on top of another assumes that the stationary lens is closest to the surface and its filter should be applied first.

7.4.1 Implementation

Our application was implemented in Java and uses the JH Labs Java Image Processing Filters¹ for real-time image processing. The prototype currently only supports the above three lens types, but adapting additional JH Labs image filters is simple. Touch

¹http://www.jhlabs.com/ip/



Figure 14: Image processing using three custom-shaped embodied lenses associated with edge detection, smear, and grayscale filters.

input is received using the TUIO [26] protocol.

8 Implications for Design

A key strength of our proposed embodied lenses is that they are very cheap and straightforward to adopt for any existing collaborative application running on a DI tabletop display. There is no extra equipment needed to detect and track lenses—only extra CPU cycles—and the lenses themselves can be created using a normal laser printer in a matter of minutes. The lenses add a whole new and tangible dimension to the tabletop display.

However, there are a number of potential drawbacks with this approach. First of all, embodied lenses require marker-based tracking and can be imprecise in a way that a fully virtual lens is not. Some of this could be addressed by the transparent visual markers proposed by Koike et al. [28]. In addition, the benefits of an embodied lens are also probably highest for novice users, where the physical affordance will help the user performing tasks, whereas an experienced power-user will likely be much more efficient with a virtual lens. Finally, transparent sheets can build up static charges, and paper sheets result in a less clear, yet still legible (Figure 7), image.

Nevertheless, we think that the drawbacks of embodied lenses are outweighed by their benefits, so we feel confident in recommending this approach be adopted in a wide range of tabletop applications, particularly for novice users and particularly those fulfilling the requirements in our design framework.

One interesting discussion point is the applicability of embodied lenses for visualization, which so far has mostly focused on traditional computing platforms such as desktop and laptop computers. Given this trend, it is not clear how widely applicable these results will be on the field as a whole. On the other hand, recent developments in visualization for tabletop computing, such as [19, 27, 49, 52], suggest that this state of affairs may very well change in the future.

9 Conclusion

We have presented an exploration of the design space of embodied lenses made out of thin sheets of paper or transparent foil for bridging physical and virtual interaction on tabletop displays. The physical affordances of these sheets allow for natural tangible interaction, such as moving, overlapping, and even reshaping the lenses, causing predictable results on the visual display. An empirical user study showed that our physical lenses were equally efficient to use as virtual lenses, but provided additional benefits such as tangible form, eyes-free operation, and instructive physical affordances. Our treatment also included a diverse set of application examples within visual computing, ranging from multiple image and map layers to image manipulation and multidimensional visualization.

Acknowledgements

We thank Raimund Dachselt for advice on the design, implementation, and evaluation of embodied lenses on tabletop displays. This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

References

- [1] Patrick Baudisch, Torsten Becker, and Frederik Rudeck. Lumino: tangible blocks for tabletop computers based on glass fiber bundles. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*, pages 1165–1174, 2010.
- [2] Patrick Baudisch and Carl Gutwin. Multiblending: displaying overlapping windows simultaneously without the drawbacks of alpha blending. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*, pages 367– 374, 2004.
- [3] Ross Bencina, Martin Kaltenbrunner, and Serge Jordà. Improved topological fiducial tracking in the reacTIVision system. In *Proceedings of the IEEE Workshop on Projector-Camera Systems*, pages 99–106, 2005.
- [4] Eric A. Bier, Maureen C. Stone, Ken Pier, William Buxton, and Tony DeRose. Toolglass and Magic Lenses: The seethrough interface. In *Computer Graphics (Proceedings of SIGGRAPH)*, volume 27, pages 73–80, August 1993.
- [5] Stuart K. Card and Jock Mackinlay. The structure of the information visualization design space. In *Proceedings of the IEEE Symposium on Information Visualization*, pages 92–99, 1997.
- [6] Görkem Çetin, Rishi Bedi, and Seth Sandler, editors. *Multitouch Technologies*. NUI Group, 2009. Online book http://www.nuigroup.com/.
- [7] William S. Cleveland and Robert McGill. Graphical perception: Theory, experimentation and application to the development of graphical methods. *Journal of the American Statistical Association*, 79(387):531–554, September 1984.
- [8] Paul H. Dietz and Darren Leigh. DiamondTouch: A multi-user touch technology. In *Proceedings of the ACM Symposium* on User Interface Software and Technology, pages 219–226, 2001.
- [9] Paul Dourish. Where the Action Is: The Foundations of Embodied Interaction. MIT Press, 2001.
- [10] Jean-Daniel Fekete and Catherine Plaisant. Excentric labeling: dynamic neighborhood labeling for data visualization. In Proceedings of ACM Conference on Human Factors in Computing Systems, pages 512–519, 1999.
- [11] Mark Fiala. ARTag, a fiducial marker system using digital techniques. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 590–596, 2005.
- [12] George W. Furnas. Generalized fisheye views. In Proceedings of the ACM Conference on Human Factors in Computer Systems, pages 16–23, 1986.
- [13] Jonathan Grudin. Partitioning digital worlds: focal and peripheral awareness in multiple monitor use. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*, pages 458–465, 2001.
- [14] Cheng Guo and Ehud Sharlin. Exploring the use of tangible user interfaces for human-robot interaction: a comparative study. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*, pages 121–130, 2008.

- [15] Carl Gutwin and Saul Greenberg. Design for individuals, design for groups: Tradeoffs between power and workspace awareness. In Proceedings of the ACM Conference on Computer-Supported Collaborative Work, pages 207–216, 1998.
- [16] Jefferson Y. Han. Low-cost multi-touch sensing through frustrated total internal reflection. In Proceedings of the ACM Symposium on User Interface Software and Technology, pages 115–118, 2005.
- [17] David Holman, Roel Vertegaal, Mark Altosaar, Nikolaus F. Troje, and Derek Johns. PaperWindows: interaction techniques for digital paper. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*, pages 591–599, 2005.
- [18] Alfred Inselberg. The plane with parallel coordinates. *The Visual Computer*, 1(2):69–91, 1985.
- [19] Petra Isenberg and Sheelagh Carpendale. Interactive tree comparison for co-located collaborative information visualization. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1232–1239, 2007.
- [20] Petra Isenberg, Niklas Elmqvist, Jean Scholtz, Daniel Cernea, Kwan-Liu Ma, and Hans Hagen. Collaborative visualization: definition, challenges, and research agenda. *Information Visualization*, 10(4):310–326, 2011.
- [21] Petra Isenberg, Anthony Tang, and M. Sheelagh T. Carpendale. An exploratory study of visual information analysis. In Proceedings of the ACM Conference on Human Factors in Computing Systems, pages 1217–1226, 2008.
- [22] Hiroshi Ishii and Brygg Ullmer. Tangible bits: Towards seamless interfaces between people, bits and atoms. In Proceedings of the ACM Conference on Human Factors in Computing Systems, pages 234–241, 1997.
- [23] Shahram Izadi, Steve Hodges, Stuart Taylor, Dan Rosenfeld, Nicolas Villar, Alex Butler, and Jonathan Westhues. Going beyond the display: a surface technology with an electronically switchable diffuser. In *Proceedings of the ACM Symposium on User Interface Software and Technology*, pages 269–278, 2008.
- [24] Sergi Jordà, Gunter Geiger, Marcos Alonso, and Martin Kaltenbrunner. The reacTable: exploring the synergy between live music performance and tabletop tangible interfaces. In *Proceedings of the ACM Conference on Tangible and Embedded Interaction*, pages 139–146, 2007.
- [25] Yasuaki Kakehi and Takeshi Naemura. UlteriorScape: Interactive optical superimposition on a view-dependent tabletop display. In *Proceedings of IEEE Tabletop*, pages 189–192, 2008.
- [26] Martin Kaltenbrunner, Till Bovermann, Ross Bencina, and Enrico Costanza. TUIO a protocol for table-top tangible user interfaces. In Proceedings of the Workshop on Gesture in Human-Computer Interaction and Simulation, 2005.
- [27] KyungTae Kim, Waqas Javed, Cary Williams, Niklas Elmqvist, and Pourang Irani. Hugin: A framework for awareness and coordination in mixed-presence collaborative information visualization. In *Proceedings of the ACM Conference on Interactive Tabletops and Surfaces*, pages 231–240, 2010.
- [28] Hideki Koike, Wataru Nishikawa, and Kentaro Fukuchi. Transparent 2-D markers on an LCD tabletop system. In Proceedings of the ACM Conference on Human Factors in Computing Systems, pages 163–172, 2009.
- [29] Johnny C. Lee, Scott E. Hudson, Jay Summet, and Paul H. Dietz. Moveable interactive projected displays using projector based tracking. In *Proceedings of the ACM Symposium on User Interface Software and Technology*, pages 63–72, 2005.
- [30] Julian Looser, Raphael Grasset, and Mark Billinghurst. A 3D flexible and tangible magic lens in augmented reality. In *Proceedings of the International Symposium on Mixed and Augmented Reality*, pages 51–54, 2007.
- [31] Donald A. Norman. The Design of Everyday Things. Basic Books, 1988.
- [32] Donald A. Norman. Cognition in the head and in the world: An introduction to the special issue on situated action. *Cognitive Science*, 17(1):1–6, 1993.
- [33] Donald A. Norman. Things that Make Us Smart: Defending Human Attributes in the Age of the Machine, chapter Visual Representations. Addison-Wesley, 1994.
- [34] Joseph A. Paradiso, Kai yuh Hsiao, and Ari Y. Benbasat. Tangible music interfaces using passive magnetic tags. In *Proceedings of the ACM Workshop on New Interfaces for Musical Expression*, 2001.

- [35] James Patten, Hiroshi Ishii, Jim Hines, and Gian Pangaro. Sensetable: a wireless object tracking platform for tangible user interfaces. In Proceedings of the ACM Conference on Human Factors in Computing Systems, pages 253–260, 2001.
- [36] James Patten, Ben Recht, and Hiroshi Ishii. Audiopad: A tag-based interface for musical performance. In Proceedings of the ACM Workshop on New Interfaces for Musical Expression, pages 11–16, 2002.
- [37] Emmanuel Pietriga and Caroline Appert. Sigma lenses: focus-context transitions combining space, time and translucence. In Proceedings of the ACM Conference on Human Factors in Computing Systems, pages 1343–1352, 2008.
- [38] William A. Pike, John T. Stasko, Remco Chang, and Theresa A. O'Connell. The science of interaction. *Information Visualization*, 8(4):263–274, 2009.
- [39] Zachary Pousman, John T. Stasko, and Michael Mateas. Casual information visualization: Depictions of data in everyday life. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1145–1152, 2007.
- [40] Anthony C. Robinson. Collaborative synthesis of visual analytic results. In Proceedings of the IEEE Symposium on Visual Analytics Science and Technology, pages 67–74, 2008.
- [41] Yvonne Rogers and Siân E. Lindley. Collaborating around vertical and horizontal large interactive displays: which way is best? *Interacting with Computers*, 16(6):1133–1152, 2004.
- [42] Daniel M. Russell, Mark J. Stefik, Peter Pirolli, and Stuart K. Card. The cost structure of sensemaking. Proceedings of the ACM Conference on Human Factors in Computing Systems, pages 269–276, 1993.
- [43] Michael Scaife and Yvonne Rogers. External cognition: how do graphical representations work? *International Journal of Human-Computer Studies*, 45(2):185–213, 1996.
- [44] Dieter Schmalstieg, L. Miguel Encarnacáo, and Zsolt Szalavári. Using transparent props for interaction with the virtual table (color plate S. 232). In *Proceedings of the ACM Symposium on Interactive 3D Graphics*, pages 147–154, 1999.
- [45] Jinwook Seo and Ben Shneiderman. Interactively exploring hierarchical clustering results. *IEEE Computer*, 35(7):80–86, 2002.
- [46] Ben Shneiderman. Dynamic queries for visual information seeking. *IEEE Software*, 11(6):70–77, November 1994.
- [47] Martin Spindler, Michael Hauschild, and Raimund Dachselt. Towards making graphical user interface palettes tangible. In *Proceedings of the ACM Conference on Interactive Tabletops and Surfaces*, pages 291–292, 2010.
- [48] Martin Spindler, Sophie Stellmach, and Raimund Dachselt. PaperLens: Advanced magic interaction above the tabletop. In *Proceedings of the ACM Conference on Interactive Tabletops and Surfaces*, pages 69–76, 2009.
- [49] Martin Spindler, Christian Tominski, Heidrun Schumann, and Raimund Dachselt. Tangible views for information visualization. In Proceedings of the ACM Conference on Interactive Tabletops and Surfaces, pages 157–166, 2010.
- [50] Maureen C. Stone, Ken Fishkin, and Eric A. Bier. The movable filter as a user interface tool. In Proceedings of the ACM Conference on Human Factors in Computing Systems, pages 306–312, 1994.
- [51] Anthony Tang, Melanie Tory, Barry Po, Petra Neumann, and Sheelagh Carpendale. Collaborative coupling over tabletop displays. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*, pages 1181–1190, 2006.
- [52] Matthew Tobiasz, Petra Isenberg, and Sheelagh Carpendale. Lark: Coordinating co-located collaboration with information visualization. *IEEE Transactions on Visualization and Computer Graphics*, 15(6):1065–1072, 2009.
- [53] Philip Tuddenham and Peter Robinson. Distributed tabletops: Supporting remote and mixed-presence tabletop collaboration. In *Proceedings of IEEE Tabletop*, pages 19–26, 2007.
- [54] Brygg Ullmer and Hiroshi Ishii. The metaDESK: Models and prototypes for tangible user interfaces. In Proceedings of the ACM Symposium on User Interface Software and Technology, pages 223–232, 1997.
- [55] Brygg Ullmer, Hiroshi Ishii, and Robert Jacob. Tangible query interfaces: Physically constrained tokens for manipulating database queries. In *Proceedings of INTERACT*, pages 279–286, 2003.

- [56] John Underkoffler and Hiroshi Ishii. URP: A luminous-tangible workbench for urban planning and design. In *Proceedings* of the ACM Conference on Human Factors in Computing Systems, pages 386–393, 1999.
- [57] Malte Weiss, Julie Wagner, Yvonne Jansen, Roger Jennings, Ramsin Khoshabeh, James D. Hollan, and Jan O. Borchers. SLAP widgets: bridging the gap between virtual and physical controls on tabletops. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*, pages 481–490, 2009.
- [58] Ji Soo Yi, Youn ah Kang, John T. Stasko, and Julie A. Jacko. Toward a deeper understanding of the role of interaction in information visualization. *IEEE Transactions of Visualization and Computer Graphics*, 13(6):1224–1231, 2007.