

LightBeam: Interacting with Augmented Real-World Objects in Pico Projections

Jochen Huber¹, Jürgen Steimle², Chunyuan Liao³, Qiong Liu³, Max Mühlhäuser¹

¹ Technische Universität Darmstadt
Germany
{jhuber, max}@tk.informatik.tu-darmstadt.de

² MIT Media Lab
USA
steimle@media.mit.edu

³ FX Palo Alto Laboratory
USA
{liao, liu}@fxpal.com

ABSTRACT

Pico projectors have lately been investigated as mobile display and interaction devices. We propose to use them as ‘light beams’: Everyday objects sojourning in a beam are turned into dedicated projection surfaces and tangible interaction devices. This way, our daily surroundings get populated with interactive objects, each one temporarily chartered with a dedicated sub-issue of pervasive interaction. While interaction with objects has been studied in larger, immersive projection spaces, the affordances of pico projections are fundamentally different: they have a very small, strictly limited field of projection, and they are mobile. This paper contributes the results of an exploratory field study on how people interact with everyday objects in pico projections in nomadic settings. Based upon these results, we present novel interaction techniques that leverage the limited field of projection and trade-off between digitally augmented and traditional uses of everyday objects.

Author Keywords

Pico projectors, handheld projectors, mobile devices, augmented reality, mixed reality, embodied interaction.

ACM Classification Keywords

H5.m. Information interfaces and presentation: Miscellaneous.

General Terms

Design, Human Factors, Theory.

INTRODUCTION

The capabilities of pico projectors have significantly increased lately. In combination with their small form factors, they allow us to dynamically project digital artifacts into the real world. Since pico projectors have been around for some years now, there is a growing body of research on how they could be integrated into everyday workflows and

Permission to make digital or hard copies of all or part 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. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

MUM '12, December 04 - 06 2012, Ulm, Germany
Copyright 2012 ACM 978-1-4503-1815-0/12/12...\$15.00.



Figure 1. Pico projector is placed on a table and uses a nearby espresso cup to show email notifications (concept)

practices. Two major categories of corresponding interaction techniques have evolved [5,18]: (1) using the projector itself for input (either via direct input such as buttons on the projector or by moving the projector like a flashlight); (2) interacting on the projection surface via direct touch or pen-based input. The projection surface is usually supposed to be fixed, large, and flat.

The present paper investigates pico projectors for interaction with real world objects—which is fundamentally different: when we engage with real world objects such as physical paper or a coffee mug, we move the *objects* in three dimensions and engage with them *spatially*: we pass a piece of paper to a colleague, we lift the coffee mug to take a sip, etc. This is particularly interesting considering recent technological developments. Mobile phones with integrated projectors will influence or even determine how projectors are used in our everyday activities. Instead of being held in hand all the time, mobile phones are often placed onto tables, for instance during meetings. Thus physical objects on the table move into the projector’s reach (cf. Figure 1). This enables a novel kind of interactive tabletop: not only the table surface, but the objects on the table become interactive displays. Intuitive handling of such objects has the potential to foster rich, non-obtrusive and tangible UIs.

This paper presents a novel interaction concept for pico projectors and real world objects, which we call *LightBeam*. In *LightBeam*, real world objects act as projection surfaces when brought into the projection beam; spatial manipulation of the objects is interpreted as user input and influences

the projected content. We tend to think of this kind of interaction as a third stage of *pervasive display-centered interaction*, the first stage being ubiquitous availability of interactive displays (Smartphones and touch screens everywhere), the second stage being ordinary flat surfaces combined with pico projectors and direct manipulation input (touch, pen, etc.). In the third stage considered here, arbitrary objects become display surfaces; at the same time, the content displayed and the interaction concepts become object specific. Additional objects brought into the projection ray correspond to additional projection surfaces, adding another degree of freedom, e.g. for tangible interaction.

These observations lead us to the following research questions: How can three-dimensional, physical objects be used for interaction in combination with pico projections in nomadic settings? What type of digital information should be displayed on which kind of objects? How to cope with the very limited field of projection?

The contribution of this paper is two-fold. First, we investigated these questions in an exploratory field study. Our results provide detailed insights into the design space of tangible interactions for real-world objects in pico projections. Second, we conceived and implemented several novel interaction techniques for two application scenarios: mobile awareness and interaction with physical documents. These techniques are specifically designed to (1) turn the drawback of a small projection area into a benefit, (2) trade-off between digitally augmented and traditional uses of everyday objects, and (3) work with almost any object within reach, which is important for nomadic settings.

In the remainder of this paper, we first present the conceptual framework of LightBeam and relate it to prior research on pico projectors. We then report on our exploratory field study and discuss our findings. Next, we illustrate how these findings informed the design of novel interaction techniques. We also give a short system overview of our prototype. In conclusion, we provide early user feedback and discuss our contribution in an integrated way.

LIGHTBEAM: RELATED WORK AND CONCEPTUAL FRAMEWORK

There is already a notable body of knowledge on pico projector interaction. Figure 2 shows the conceptual categories for this kind of interaction. We will discuss both background and conceptual framework of LightBeam in the context of these three categories.

Fixed Projector & Fixed Surface

The small form factor of pico projectors can be leveraged for integrating them virtually anywhere. In *Bonfire* [10], two camera-projector-units are attached to a laptop and therefore extend the display area to the left and right hand sides of the laptop. The projection is used as an interactive surface, allowing users to employ multi-touch gestures on the projected area. Moreover, the system recognizes every-

day objects such as a coffee cup through vision-based methods and can project additional information, however only onto the flat table surface, not onto 3D objects. FACT [11] tracks ordinary paper documents with their natural features and enables word-level augmented reality interaction with the documents. Both projector and paper document need to be placed at a *fixed* position to enable fine-grained document interaction. Other examples are indirect input techniques using gestures [3] or shadows [6].

Mobile Projector & Fixed Surface

The aforementioned research conceptually focuses on techniques, where both projector and projection surface are required to be *fixed* in space (cf. Figure 2a). A larger body of research is motivated by the mobility of pico projectors [26]: they can be easily carried around, held in hand and used to project onto fixed surfaces such as walls (cf. Figure 2b). Prominent work has been carried out by Cao et al. [1,2]. They developed various handheld interaction techniques, as well as pen-based techniques for direct input on the projection surface. In both cases, they chose large flat and fixed surfaces, such as walls, as their projection targets. Most of the techniques rely on the so-called flashlight metaphor. Here, the projector only projects a cutout of the virtual information space. By moving the projector, further parts of the information space are being revealed. The flashlight metaphor is also used in other projects such as *Map torchlight* [19], *iLamps* [17], *RFIG Lamps* [16] and *Mouse-Light* [20] to augment static surfaces with digital information. The latter also allows for direct pen input on the projection surface. Most recently, Molyneaux et al. [14] have presented two camera-projector systems, which support direct touch and mid-air gestures on arbitrary surfaces. However, once registered, these surfaces must remain at a fixed location, which impedes tangible interaction. *Motion-Beam*, a concept by Willis et al. [23], also uses a fixed surface as projection target. It allows users to steer a projected virtual character through virtual worlds. The character is bound to the projection; the projector is handheld and reveals only a part of the game world. Willis et al. [24] have also investigated ad-hoc multi-user interaction with handheld projectors on fixed surfaces.

A few projects also investigated wearable projection, where the pico projector is attached to clothes or worn like an accessory. A prominent example here is *Sixth Sense* [12]. A

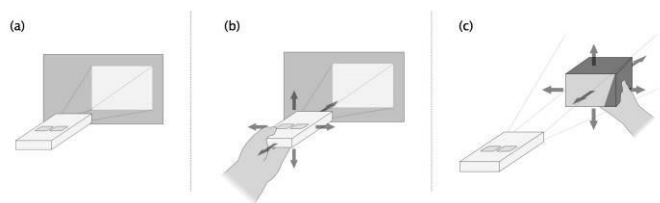


Figure 2. Conceptual levels for pico projector interaction: (a) fixed projector, fixed surface; (b) mobile projector, fixed surface; (c) fixed projector, mobile surface (LightBeam)

camera-projector unit is worn as a necklace. Physical surfaces such as walls, but also parts of the body can then be used as a projection surface. Users are able to interact with the projection using in-the-air gestures in front of the camera. *Skinput* [8] also leverages body parts as projection surfaces but allows for touch input directly on the body. This effort has been further refined in *OmniTouch*, where Harrison et al. [7] enabled touch input on arbitrary surfaces using a depth-camera and a pico projector. Although these three projects support projection onto essentially mobile objects such as a human arm, these objects are only used as hosts for the projection, not for tangible interaction. Hence, from a conceptual viewpoint, they can also be regarded as fixed projection surfaces. A slightly different approach is pursued in *Cobra* [27] by Ye and Khalid. They use a flexible cardboard interface in combination with a shoulder-mounted projector. The cardboard can be bent as a tangible input for mobile gaming. However, the cardboard needs to be held steady at a fixed position.

Fixed Projector & Mobile Surface

In summary, previous work on pico projector interaction emphasized on fixed and flat projection surfaces in physical space. It is worthwhile to note that there is a larger body of knowledge on interaction with objects in larger projection spaces. Prior work in this field dates back to the early 1980s, when Michael Naimark investigated immersive projection environments in art installations [15]. More recently, physical objects such as paper have been used as projection surfaces in *PaperWindows* [9]. This idea has been developed further in *LightSpace* [25], where basically any fixed surface in a small room installation is being recognized. Within this scope, Wilson et al. have investigated interaction on, above and between surfaces—but not using the surfaces themselves as tangible interaction devices. Most related to our work is Molyneaux’s work on smart objects [13]. They have investigated how physical objects can be turned into interactive projected displays. The main focus of the work was on orchestrating a technical infrastructure, allowing for reliable and robust object detection through model-based approaches. In addition to relying on larger projectors, they have not investigated the tangible character of physical objects, but used the projections to display additional object-specific information directly on the objects.

However, compared to larger projectors, the affordances of pico projectors are fundamentally different: they are mobile and have a very small and strictly limited projection ray. Thus we tend to think of pico projectors more like personal devices, which are carried around and used in a plethora of situations and places, such as workplaces or cafés. And as opposed to immersive projection spaces, pico projectors provide only a highly limited projection ray. To the best of our knowledge, the impact of these characteristics have not been systematically explored for tangible interaction with real world objects. Moreover, it is unclear what kind of

projected information actually matches the affordances of physical objects (cf. Figure 2c). *LightBeam* aims at filling this void.

Our LightBeam Concept

In *LightBeam*, the pico projector is fixed in the vicinity of the user and not constantly held in hand. It can be attached to physical objects (e.g. walls, desks or cupboards) and its tilting angle can be adjusted. This way, projection onto the physical space can be supported from flexible perspectives. Figure 2c) illustrates the *LightBeam* concept. The projection is regarded as a constant, but limited ray of light into the physical space. The projection is “always-on”, as long as the user wants. The projector itself is augmented with a depth camera unit and can track objects within its ray in three-dimensional space. Thus the projection provides output as well as input functionality: on the one hand it can augment physical objects with digital artifacts; on the other hand, deliberately moving an object into the ray and manipulating it there can also serve as input. For instance, a physical document held into the ray could get automatically recognized and contextually relevant information could be displayed on the physical document. Moreover, physical interaction with the objects such as movement, rotation or other embodied gestures can be used as tangible control. For instance by gradually bringing the document into the ray, the level of detail of the contents is continuously increased.

Thus, *LightBeam* provides a theoretically motivated conceptual framework, focusing on (1) object-centered interaction, (2) spatial interaction, and (3) a three-dimensional projection space. Central to *LightBeam* is the concept of moving objects in the limited projection space but not the pico projector (except for changing the perspective).

Figure 2 separates the composition of projector and object mobility conceptually. In practice, the boundaries are not rigid and the individual approaches can be combined, leading also to mobile projector interaction with mobile objects as a combination of Figure 2b) and 2c).

EXPLORATORY FIELD STUDY

We conducted an exploratory field study to investigate the aforementioned research questions and to gain a deeper understanding of how pico projectors can be used together with physical objects. Besides exploring the design space, the qualitative results should also inform novel interaction designs. We particularly wanted to explore the following dimensions:

- **Projector placement:** How is the projector positioned in physical space? For instance, is it hand-held or is the projector deliberately placed in the environment?
- **Output:** What kinds of objects are used for mobile projection? What kind of information should be displayed, depending on the target objects? Does mobile projection influence the meaning of objects?

- **Input:** How are real world objects manipulated in 3D space for interaction with mobile projections?

In the following, we outline our study design, the employed methodology and discuss the findings in detail.

Study Design

Setting. We conducted the study in two different places: the subject’s workplace and a café. We selected these two places mainly for three aspects: spatial framing, social framing and the manifold nature of objects contained within these places. In particular, these places allowed us to study personal places, which are thoughtfully arranged by the participant and contain personal objects, and public places, where available objects typically do not have a personal meaning to the participants. Figure 3 shows examples of both places. For the café setting, we ensured that the types of objects present on the coffee table were consistent for all sessions. This was not desired for the office setting, since it was the subject’s personal desk. The participants were seated in both settings. Each session lasted about 1.5 hours in average. The order of the places was counter-balanced.

Participants and Tasks. We recruited 8 interaction design researchers (7m, 1f) between 25 and 33 years of age (mean 28). Their working experience ranged from 1 to 6 years (mean 4). Our main objective was to observe the participants while using the projector for certain interactions in the field. The interactions themselves were embedded in semi-structured interviews, lead by one of the authors. The participant was given an Aaxa L1 laser pico projector and plenty of time for getting familiar with the pico projector.

The participants were told that the projector could be used for the same tasks as they carry out with their mobile phone. The projector was able to display a number of multimedia resources such as photos, videos and digital documents that we had selected and stored on the device before. The content was used during the sessions to simulate typical scenarios for pico projector usage such as photo sharing, video consumption or co-located collaboration with digital documents. The participants were either asked how they would project and interact with certain content or deliberately confronted with a projection. Figure 4 shows the latter case, where the interviewer projected a movie onto a cup on the participant’s personal desk. The interviewer first observed how the participant would react to this and then continued the interview process. The semi-structured inter-



Figure 3. Example photographs from the two settings in the field study; personal desk (left) and café (right).



Figure 4. Projection of a YouTube clip on a coffee mug.

views were highly interactive and had the character of brainstorming sessions.

We used an Aaxa L1 laser pico projector, as a low-fidelity prototype. This was due to two reasons: (1) we did not want to influence the participants by any design and (2) we wanted to explore the aforementioned fundamental dimensions such as projector placement. A high-fidelity prototype would have imposed too many constraints on the interaction space.

Data Gathering and Analysis. We chose a qualitative data gathering and analysis methodology, which we performed iteratively per session. As data gathering methodologies, we used semi-structured interviews, observation and photo documentation. After each session, the interviews and observations were transcribed. Salient quotes were selected and analyzed using an open, axial and selective coding approach [22]. The emerging categories served as direct input for the follow-up session with the next participant. The scope of the session was adapted according to the theoretical saturation of the categories.

In the following three subsections, we present the findings from our study. The coding process yielded various categories, depending on where the projector was placed, which objects were selected as projection targets and how objects actually foster input capabilities.

Results I: Handheld versus Placed Projector

Our observations revealed that the projector was used in a two-step process by all participants in both settings (office and café): initially, the participants used the projector as a handheld device to find a suitable projection area for the beam, which is not physically constrained by objects that cannot be moved. Then, they placed it onto the table and the projector was no longer used in hand throughout the entire session. The only exceptions were rare cases when the projector was moved to another location in its vicinity to slightly readjust the projection space.

Placing the projector instead of using it in hand was mostly due to ergonomic reasons. Once the projector was placed on the table, not the projector, but movable objects were repositioned to serve as projection targets. P8 noted: “When

would I actually make the *effort* of holding the projector? I am constantly looking for objects, which are perfect hosts for the projection, which I can then bring into the beam. I do not want to hold the projector. It constrains me.”

Results II: How to Leverage Objects for Output?

In the interviews, the participants noted that the affordances of objects determine whether and how an object can be used for output or input.

Relationship between Projected Content and Object

We observed a direct correspondence between the cognitive demand required by the projected content and both the size and shape of an object that was chosen as the projection target.

Cognitively demanding content such as presentation slides, where it is crucial to grasp the whole level of detail, was projected onto larger, less mobile and rigid surfaces. Examples comprise larger boxes, tables or the floor. Interestingly, such content was not projected onto walls, since in this case others would have been able to see it. The latter was considered either “impolite and a disturbance to others” (P5) or a privacy issue (mentioned by all participants).

Cognitively less demanding content, such as short YouTube clips or photos, was projected onto rather small and even non-planar objects (e.g. see Figure 4). Participants commented that these are perfectly suitable when only a lower level of detail is required. Moreover, such objects provide the benefits of being easily movable. As a direct consequence, they can be easily replaced by other objects when required. For instance, P8 used the back of his hand as a substitute projection surface, when he viewed a projection together with the interviewer and was required to move the original surface (a rigid paper box) away. He stated: “I considered it impolite to just leave you without the projection. So I figured out that the back of my hand is better than nothing—at least you can see the projection”.

The participants did not mind slightly distorted projections, when they did not want to devote their whole attention to the projection: “I do not care that this projection [a YouTube clip] does not fit onto this object [a small package, 5x3cm] – I still can understand the gist of it”. Moreover, even curved surfaces were used for such a task, e.g. P7 commented in the situation of Figure 4: “Even though it is distorted towards the edges of the cup, I do not mind, since it is not a high quality movie. Moreover, I only focus on the center of the projection and I can understand what is actually happening”.

Objects afford Physical Framing

The natural constraints provided through the boundaries of physical objects were also considered important. P7 noted: “I want to put things into frames. Objects on my desk provide this frame, whereas my table itself is too large—there is no framing”. This is different to just projecting a digital



Figure 5. A participant demonstrates how he would use his hand to quickly skim through a list of pictures and then turn his hand towards the interviewer to present a picture.

frame around the projection, since moving the frame would imply moving the projector. But here, objects are the frames. It was considered crucial that the projection is clearly mapped to the object. P8 elaborates on this by saying: “Objects are like frames for me, they provide space and receive the projection”.

Embodiment of Digital Artifacts

We observed that all of the participants used the mobility of objects and the physical framing of the projections to control who is actually able to see the projected content. P2 stated: “You can easily direct attention by moving it, [turns a menu with the projection on it to herself] and now I can read it.” This leads to a rather object-centric perspective on interaction, as P3 outlines: “It is not the device I care about, it is the object with the projection.” Moreover, P4 argues that “the data is on the object, it is contained within it. The digital artifact is embodied through the physical object.”

Results III: How to Provide Input with Objects?

While larger surfaces provide extensive display area for detailed output, they are likely hard to move and therefore are rather fixed in physical space. Smaller physical objects however afford manipulation in three-dimensional space.

Moving Objects within the Beam

The participants argued that since the data is bound to a physical object, the object itself could be used as a tangible control. P7 described this as “physical shortcuts to certain digital functionality”. He further mentioned that he makes “an abstraction from the actual object towards its Geometry”. He therefore concludes: “For instance, when I look at my coffee mug, I see an object which can be rotated by grabbing its handle; I would want to use this for quickly controlling something like a selection”. Another participant moved his hand forth and back within the projection ray and imagined to quickly skim through a list of pictures (cf. Figure 5). P6 noted that he “would not want to perform a three-dimensional gesture mid-air due to the lack of haptic appeal, but using an object for that as a medium would be perfectly fine”.

Dynamic Modification of Object Shapes

The flexibility that some physical objects exhibit, such as paper, was also used to dynamically modify the projection surface in two ways: (1) to increase and decrease the display size and (2) for (semantic) zooming, comparable to tangible magic lenses [21], but in a mobile situation. Participants used folding gestures with paper to increase or decrease the display size. Folding paper was mapped to decreasing and unfolding paper was mapped to increasing display size.

Participants reported that deformable objects are perfectly suitable for “taking a peek into the beam” (P5). P5 imagined that the projector was constantly projecting into space without a target object and was able to display notifications, like on his Android smart phone. “By lifting a paper and moving it into the beam”, he explained, “I can just take a look at my notifications, you know, to look if something is there”.

Capturing Objects Visually

In the context of document interaction, the projector was also considered as a “scanner”. P7 stated: “If I project onto a document, the projector can also ‘copy’ the physical document to the digital world. I can do this with various documents on the go and share them here.” P2 also noted that the mobile projection can be used to add digital artifacts such as annotations to documents. She exemplified this by lifting an article, grabbing a pen and circling a paragraph.

Overloading Mappings of Physical Objects

Projecting onto an everyday object and mapping digital functionality to it is more than just a visual overlay in physical space. It also redefines the object’s purpose. Moreover, a projection locks objects in physical space, as P7 elaborates: “If I used this coffee mug as a tangible control for an interaction I heavily rely on, I would certainly have to forget its use as a mug. It would have to remain there, at that very place, to allow me to carry out this function at any time.” The consensus across the participants was that overloading the mapping of physical objects is good, for short terms. Physical objects afford casual interaction, as P5 described: “I would want to just put the object within the projector beam, carry out an interaction and remove the object from the beam”.

INTERACTION TECHNIQUES

The results from our exploratory field study show that LightBeam provides a fundamentally different interaction space for tangible interaction than larger immersive projection spaces. Being placed in a user’s vicinity, it provides a dedicated interaction space through its highly limited projection ray. Our results show that moving objects therein is a central theme for interaction in real world settings. Objects provide a physical framing for projections and thereby embody them. Different physical characteristics of objects afford for projecting different digital contents. Furthermore, our results show that LightBeam, as a spatial ray, is not only used for output or tangible interaction, but also for capturing physical objects visually.

Interaction Primitives

Based upon our observations above, we have identified interaction primitives for LightBeam (see Figure 6). These serve as the basis for interaction techniques discussed afterwards.

Move into the beam: Physical objects can be moved into the beam. In addition to moving an object entirely into the beam, the user can vary the degree to which the object resides within the beam. The portion of the object, which is located within the beam can be augmented with digital functionality. Several objects can reside simultaneously within the beam.

Remove from the beam: Removing an object from the beam removes any digital functionality from the physical object.

Move within the beam: Objects can be moved within the beam in three-dimensional space. This can be used to arrange projected contents in 3d space or as tangible control.

Beam captures an object: A visual copy of a physical object in the beam is captured and stored digitally.

Externalizing captured objects: Previously captured copies of objects can be visualized within the beam by projecting them onto physical objects.

In the following, we show how combining these primitive interactions creates novel interaction techniques that leverage the limited projection ray of LightBeam. We identified two promising application scenarios: on the one hand, when placing the pico projector on a table (similarly to how many people put their smart phones on a table during a conversa-

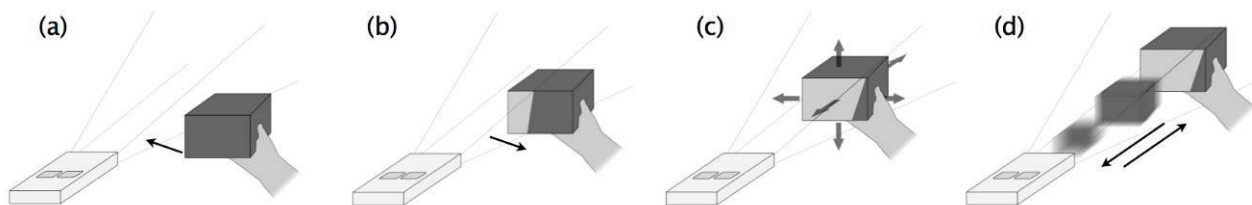


Figure 6. Interaction primitives for LightBeam: (a) Move into the beam, (b) Remove from the beam, (c) Move within the beam, (d1) Beam captures an object (direction toward projector) and (d2) Externalizing captured objects (direction toward object).

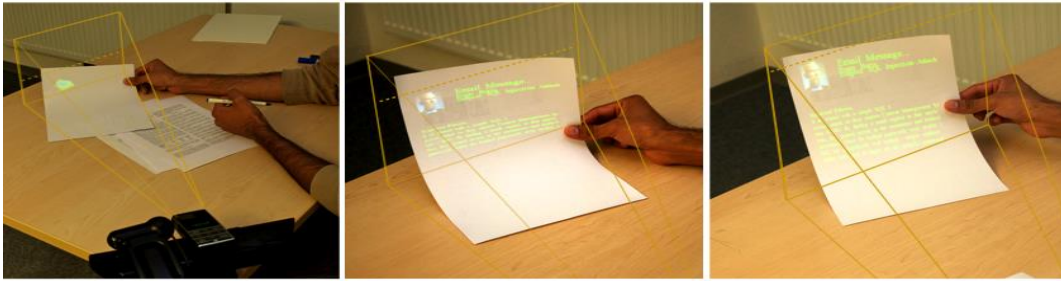


Figure 7. From left to right: the user utilizes the back of one of the papers he is currently working on to take quick look into the projector beam. In the first image, a small envelope is displayed due to the limited projection space. By gradually lifting the paper, the level of detail is adjusted, more text is displayed and automatically wrapped within the boundaries.

tion), it can turn everyday objects in its vicinity into peripheral awareness devices. On the other hand, LightBeam can aid in bridging the digital-physical divide when interacting with paper documents, a class of physical objects that is specific due to its high information content.

Gradual Sneak-Peek Into the Beam

Easily movable objects can be used to display information in-situ by moving them into the beam. Different objects afford different levels of details: while a larger box placed within the beam can show richer information (cf. Figure 7), smaller objects, e.g. a corner of a piece of paper, afford peeking at low-level information notifications.

We leverage the restricted field of projection for quick transitions between different levels of details. As an object is gradually moved into the beam, the projection area increases and more information can be presented. By partially removing the object from the beam, the level of detail of the information presented decreases. While this interaction is possible with any object, we believe that deformable objects lend themselves particularly to this interaction:

Figure 7.1 shows our exemplary interface: the projector is placed on a desk while the user is working with a physical document. The sketched projection ray in figure 7 indicates the highly limited projection area. The dotted line designates the effective projection (EP) area, which is the intersection between the projection area and the object. By slightly lifting the document, the user can take a peek into the beam (small EP) and see if there are any new notifications. Gradually lifting the document further into the beam reveals more details (larger EP, cf. Fig. 7.2 and 7.3). Removing the paper from the beam reduces the EP and displays less information. As a slight variation of this technique, folding and unfolding a piece of paper within the projection beam affords a discrete transition between



Figure 8. A photostream from Flickr is projected onto a box and can be navigated by rotating the coffee mug.

different levels of detail. As a matter of course, objects can also be permanently placed within the beam to immediately receive notifications (push-mode instead of pull-mode of information updates).

Projected contents can be bound to objects of particular shape (e.g. boxes as large displays as in Fig. 8). Alternatively, depending on the application or user preferences, contents can also be displayed on any object that is introduced into the beam. This ensures high usability in mobile contexts where specific objects might not be always at hand.

Using Any Object as Tangible Control

When moved *within* the beam, objects can act as tangible controls. Prior work [4] mapped one particular object to a specific digital functionality. However, in nomadic settings, it cannot be taken for granted that specific objects are always available. Therefore, we advocate mapping a specific function not to one specific object, but to a class of objects that have a certain affordance. For instance, a function could be mapped to physical rotation of a cylinder; hence any cylindrical object that affords rotation can be used to perform that function, e.g. a mug, a bottle, a vase, or a candy box.

Our implementation is shown in Figure 8. We use the rotation of objects, here a mug, to navigate through the displayed pictures.

In particular, a physical object is only mapped to digital functionality while residing within the limited beam. Removing the object from the beam also removes the digital functionality and its original mapping is restored. Putting objects into the beam and removing them from the beam provides a lightweight way for switching between their uses as non-augmented vs. digitally augmented objects. For instance, when the coffee mug is not inside the beam, the user can take a sip from the mug without the system detecting this as tangible input.

Using the Beam as a Visual Scanner

In addition to projecting visual output onto objects or leveraging them as tangible controls, the beam can also be interpreted as a visual scanner, which captures objects. Moving an object into the beam selects it for capturing. Figure 9.1 and 9.2 show an example where a physical document is captured, automatically identified and its digital representation (here: a PDF) is stored virtually. With this technique, multiple pages (or documents) can be scanned subsequent-

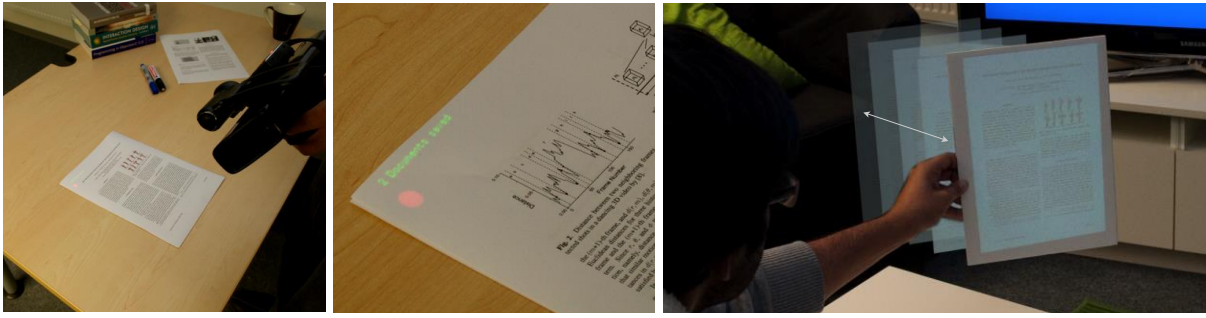


Figure 9. From left to right: (1) and (2) the projector is used to capture a physical document, storing its digital equivalent as a PDF. (3) shows a user skimming through a stack of captured documents by moving a piece of paper forth and back.

ly. We model the process of capturing multiple objects as putting them onto a virtual stack of objects that resides within the beam: each scanned object is put onto the beam's internal stack and is stored digitally. The digital versions can in turn be externalized into the physical space by moving an object into the beam. Moving the object back and forth within the beam (see Figure 9.3) allows for browsing the beam's stack.

Instead of scanning each object in its entirety, we also support more fine-grained selection. Figure 10 shows an example where a physical document is moved into the beam. In addition, a pen is also moved into the beam and can be used for selecting parts of the documents for capturing. Only selected parts are put onto the beam's stack.

In the reverse direction, the pen can be also used for putting a document snippet, which was previously captured by the beam, to a specific location on an object (the same object it was captured from or a different object). This is performed by a flick gesture with the pen towards the object.

As described above for tangible interaction, the mapping of the pen is only temporarily overloaded. Moving the pen into the beam allows using it for copy and paste of document snippets. In turn, removing it from the beam restores its original function: it can be used for writing.

For the sake of focus and clarity, we here concentrate on tangible, ray-based interaction techniques. As a matter of fact, they can be easily combined with touch input, using the approach presented in [7].

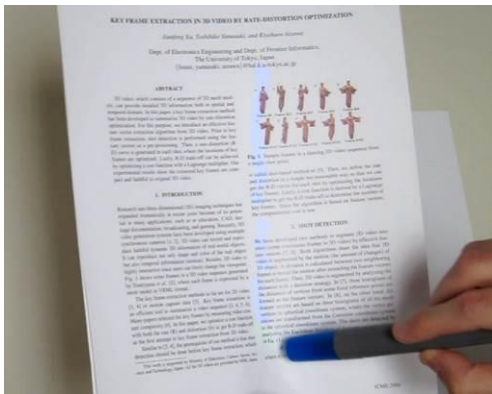


Figure 10. The piece of paper is held in 3D space and a pen is used to select a part of the document (blue line), which is in turn captured and projected into physical space.

SYSTEM OVERVIEW

We have prototypically implemented the interaction techniques. In the following, we describe our hardware setup, as well as our algorithms.

Hardware

Figure 11 shows our prototype. We have attached an Aaxa L1 laser pico projector to a Microsoft Kinect with hook-and-loop tape, which we use as a mobile camera-projector unit. The projector has a resolution of 800x600 pixels. The Microsoft Kinect features a pair of depth-sensing range cameras (320x240 pixels), an infrared structured light source and a regular RGB color camera (640x480 pixels). In order to support hassle free document recognition, we have attached a megapixel webcam with *autofocus* to the unit. Kinect, webcam and pico projector are calibrated and aligned.

The mobile camera-projector unit can be further mounted onto a strong suction cup, which also features a handle. Thus the unit can be easily carried in one hand by using the handle. Moreover, it can be attached to basically any flat surface, even vertical surfaces or ceilings to achieve a top-down projection.

Object Tracking and Interaction Support

As projection surfaces, we currently consider flat surfaces of 3D objects. We model them as 2D planes in 3D space. To support a robust tracking of arbitrary objects, independent of varying lighting conditions, we aimed at using solely the depth image in our tracking algorithm. First, a threshold is applied to the depth image to filter out any background objects. A blob detection for the objects in the scene is carried out. The algorithm then iterates over each object. As a simple example, Figure 12 shows only one object (here: a piece of paper), which is held in hand. We isolate the object from the scene and discard the hand in three steps: (1) we erase thin lines in the input image, connecting larger areas (e.g. the connection between the piece of paper and the arm in Figure 12 right) by applying blur filters, thresholding the image and applying morphological operators. The resulting image of step 1 contains the isolated object. However, due to the image operations, the area and consequently the contour have been reduced. Nevertheless, a further blob detection in step 2 now enables the detection of the reduced area. Then, a rotation invariant bounding rectangle of minimum area is calculated. In step 3, the contour

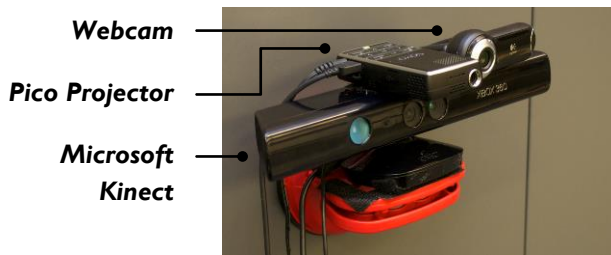


Figure 11. Hardware prototype using a Microsoft Kinect, pico projector placed on top. We have added a webcam on the right hand side for document recognition.

of this bounding rectangle is then mapped to the object's original contour of the image in step 1. In combination with the depth information for the detected contour, we model and track the detected object as a 2D plane in 3D space. The projection is mapped using a homography, correcting any perspective errors. We also analyze the optical flow within the regions of the blobs in the RGB image. This allowed us to detect whether an object has been rotated.

Document Recognition

The system automatically recognizes paper documents to support the rich interactions described in the mobile document interaction scenario. The recognition is based upon FACT [11], which unitizes local natural features to identify ordinary paper documents without any special markers. Currently, our implementation can operate at ~ 0.5 fps for recognizing a frame of 640×480 pixels on a PC with a quad core 2.8GHz CPU and 4GB RAM. Considering that users usually do not change documents very quickly during their tasks, this recognition speed is acceptable for practical use. The FACT implementation had to deal with various difficulties due to only using data from an RGB camera; e.g. small document tilting angles or interferences of overlaid projections with the original natural features. We leverage the capabilities of the Kinect depth camera to overcome these difficulties. The 3D pose estimation based on the depth image is independent of the document's natural features and thus the system is robust to insufficient feature correspondence. Moreover, a rectification of the color images based on the 3D pose decreases the perspective distortion and allows for greater tilting angles. Last, the pose estimation and the document recognition can be carried out in two separate threads, each updating the world model *asyn-*

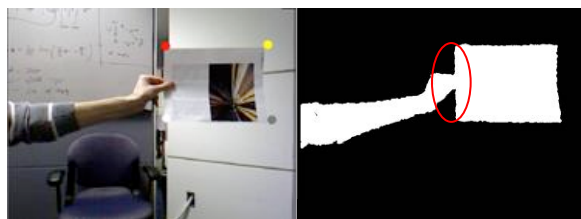


Figure 12. Left: color image. Right: image from the depth camera with a depth-threshold and initial blob detection applied. The red mark designates the thin connection, which the algorithm removes for object detection.

chronously. Therefore, from the aspect of users, the system is able to locate document content in 3D space in real time.

EARLY USER FEEDBACK

We have evaluated the prototypical implementation of LightBeam in an early user feedback session with 6 interaction design researchers. The group session lasted about 3 hours. Our main objective was to get a first impression whether the techniques are conceptually sound and how the participants would actually use them to interact with physical objects. We evaluated the interaction techniques using semi-structured interviews in our living lab. Our lab is an open space, containing desks (to simulate a working environment) and an area comparable to a living room with couches and a large LCD TV.

Method

The participants were asked to familiarize themselves with our hardware prototype. The desk contained typical items such as books, a laptop, pens, etc. They were given the opportunity to explore each technique using the objects in the vicinity. Although our prototype requires to be wired to a PC for data transfer, the participants were able to roam around freely whilst carrying and repositioning the LightBeam. As data sources, we used the semi-structured interviews and also observed the participants. We transcribed the data and analyzed salient quotes.

Results and Discussion

All participants easily understood the interaction techniques. They emphasized the benefit of the tight integration of physical objects and digital information, since "this allows for a direct interaction with the virtual data", as one participant noted.

The participants were focused primarily on the role of physical objects. Throughout the session, the participants repeatedly stressed the significance of using virtually any object to control the projection; in our example the rotation of objects. This also diminished their concerns that objects might lose their original function when being used as tangible controls. One participant commented: "I like this kind of casual functional overlay. Now I am not afraid that I will end up with two coffee mugs on my table, since one might be dedicated to one specific function". However, they noted that they might want to bind certain types of information to special objects on purpose.

Moving any object into the beam to take a peek into the virtual world was considered important for supporting quick information access in-situ. It was considered particularly helpful when already dealing with physical objects, such as paper, on the table, since lifting them further into the beam triggered the seamless transition between different levels of detail. One participant commented: "Projecting onto the table would be good, but actually, the table is too large, there is no frame". The other participants agreed. This further underlines our findings from

the exploratory study: physical objects provide natural frames.

When capturing physical objects within the beam, the participants again considered the casual overloading of physical objects (here: the pen) with digital functionality as useful. They reported that browsing and selecting digitally captured objects using the object movement in the z-direction is beneficial for providing an overview over and quick access to most recently captured objects. For larger collections however, two participants would have preferred to interact on the object itself, e.g. through a gesture-based interface instead of moving it through space.

CONCLUSION

This work has explored using pico projectors as ‘light beams’, adding a novel conceptual dimension to the pico projector design space. LightBeam provides a fundamentally different interaction space for tangible interaction than larger projection spaces. Being placed in a user’s vicinity, it provides a dedicated interaction space through its highly limited projection ray. The results from an exploratory field study show that moving objects therein is a central theme for interaction in real world settings—while moving the projector is not. Objects provide a physical framing for projections and therefore embody them. Projections can be bound to objects of particular shape (e.g. boxes as large displays), but can be also adapted to deformable physical objects, depending on both application and user preferences.

Based on a set of interaction primitives, we contributed several interaction techniques, which leverage this: moving objects into the beam charts them with both output and input functionality. Here, the highly limited projection ray plays an important role. It serves as a dedicated interaction hotspot wherein objects can be deliberately moved, therefore overloading the objects’ original mapping (e.g. using a cup as a tangible control instead of drinking from it). Withdrawing the objects from the beam then removes the overloaded and respectively restores the original mapping. By leveraging physical affordances of objects for tangible controls instead of dedicating specific objects to specific functions, we provide a loose coupling between object and functionality. This is key for object-based interactions in nomadic settings, where it cannot be taken for granted that specific objects are available. We believe that this, in combination with the casual overloading of physical mappings and already existing touch-based interfaces [7], will fundamentally change how we ubiquitously interact with augmented real-world objects in nomadic settings.

ACKNOWLEDGEMENTS

We thank Faheem Nadeem and Fawaz Amjad Malik for their valuable support.

REFERENCES

1. Cao, X., Forlines, C., and Balakrishnan, R. Multi-user interaction using handheld projectors. In *Proc. UIST '07*, ACM, 43-52.
2. Cao, X., and Balakrishnan, R. Interacting with dynamically defined information spaces using a handheld projector and a pen. In *Proc. UIST '06*, ACM, 225-234.
3. Cauchard J.R., Fraser M., Han T. and Subramanian S. Steerable Projection: Exploring Alignment in Interactive Mobile Displays. In *Springer PUC*, 2011.
4. Cheng, K.-Y., Liang, R.-H., Chen, B.-Y., Laing, R.-H., and Kuo, S.-Y. iCon: utilizing everyday objects as additional, auxiliary and instant tabletop controllers. In *Proc. CHI '10*, ACM, 1155-1164.
5. Cowan, L.G., Weibel, N., Griswold, W.G., Pina, L.R., and Hollan, J.D. Projector phone use: practices and social implications. In *PUC 16*, 1 (January 2012), 53-63.
6. Cowan, L. G., and Li, K. A. ShadowPuppets: supporting collocated interaction with mobile projector phones using hand shadows. In *Proc. CHI '11*, ACM, 2707-2716.
7. Harrison, C., Benko, H., and Wilson, A. D. OmniTouch: wearable multitouch interaction everywhere. In *Proc. UIST '11*, ACM, 441-450.
8. Harrison, C., Tan, D., and Morris, D. Skinput: appropriating the body as an input surface. In *Proc. CHI '10*, ACM, 453-462.
9. Holman, D., Vertegaal, R., Altosaar, M., Troje, N., and Johns, D. Paper windows: interaction techniques for digital paper. In *Proc. CHI '05*, ACM, 591-599.
10. Kane, S.K., Avrahami, D., Wobbrock, J.O., et al. Bonfire: a nomadic system for hybrid laptop-tabletop interaction. In *Proc. UIST '09*.
11. Liao, C., Tang, H., Liu, Q., Chiu, P., and Chen, F. FACT: fine-grained cross-media interaction with documents via a portable hybrid paper-laptop interface. In *Proc. ACM MM '10*, ACM, 361-370.
12. Mistry, P., Maes, P., and Chang, L. WUW - wear Ur world: a wearable gestural interface. In *Proc. CHI EA '09*, ACM, 4111-4116.
13. Molyneaux, D., and Gellersen, H. Projected interfaces: enabling serendipitous interaction with smart tangible objects. In *Proc. TEI '09*.
14. Molyneaux, D., Izadi, S., Kim, D., Hilliges, O., Hodges S., Cao, X., Butler, A., and Gellersen, H. Interactive Environment-Aware Handheld Projectors for Pervasive Computing Spaces. In *Proc. Pervasive '12*, Springer LNCS, v 7319/2012, 197-215.
15. Naimark, M. Two unusual projection spaces. *Presence: Teleoper. Virtual Environ.*, 14(5):597-605, October 2005.
16. Raskar, R., Beardsley, P., Baar, J. van, et al. RFIG lamps: interacting with a self-describing world via photosensing wireless tags and projectors. In *Proc. SIGGRAPH '04*, ACM, 406-415.
17. Raskar, R., van Baar, J., Beardsley, P., Willwacher, T., Rao, S., and Forlines, C. iLamps: geometrically aware and self-configuring projectors. In *ACM Trans. Graph.*22, 3, 809-818.
18. Rukzio, E., Holleis, P., and Gellersen, H. Personal Projectors for Pervasive Computing. In *IEEE Pervasive Computing*, (2011).
19. Schöning, J., Rohs, M., Kratz, S., et al. Map torchlight: a mobile augmented reality camera projector unit. In *Proc. CHI EA '09*, ACM.
20. Song, H., Guimbretiere, F., Grossman, T., and Fitzmaurice, G. MouseLight: bimanual interactions on digital paper using a pen and a spatially-aware mobile projector. In *Proc. CHI '10*, ACM.
21. Spindler, M., Tominski, C., Schumann, H., and Dachselt, R. Tangible views for information visualization. In *Proc. ITS '10*, ACM, 157-166.
22. Strauss, A. and Corbin, J. *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*. Sage, 2008.
23. Willis, K.D.D., Poupyrev, I., and Shiratori, T. Motionbeam: a metaphor for character interaction with handheld projectors. In *Proc. CHI '11*, ACM, 1031-1040.
24. Willis, K.D.D., Poupyrev, I., Hudson, S.E., and Mahler, M. SideBySide: ad-hoc multi-user interaction with handheld projectors. In *Proc. UIST '11*, ACM, 431-44.
25. Wilson, A.D., and Benko, H. Combining multiple depth cameras and projectors for interactions on, above and between surfaces. In *Proc. UIST '10*, ACM, 273-282.
26. Wilson, M.L., Robinson, S., Craggs, D., Brimble, K., and Jones, M. Pico-ing into the future of mobile projector phones. In *Proc. CHI EA '10*, ACM, 3997-4002.
27. Ye, Z. and Khalid, H. Cobra: flexible displays for mobile gaming scenarios. In *Proc. CHI EA '10*, ACM, 4363-4368.