

Pervasive Information through Constant Personal Projection: The Ambient Mobile Pervasive Display (AMP-D)

Christian Winkler, Julian Seifert, David Dobbelstein, Enrico Rukzio
Ulm University, Ulm, Germany
{firstname.name}@uni-ulm.de

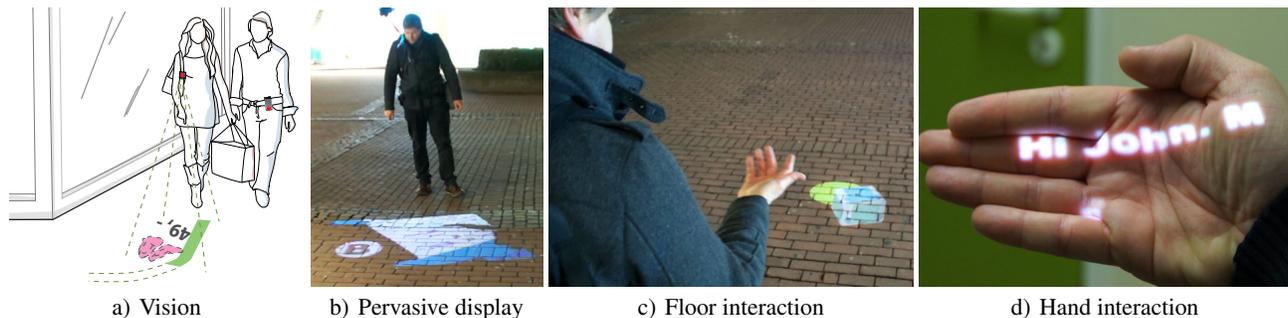


Figure 1. a). The vision: small wearable projectors reveal serendipitous information. b) A user discovers information with the AMP-D prototype. c) The user has received a text message and picks it up from the floor. d) The scrolling message text is read in the user's hand.

ABSTRACT

The vision of pervasive ambient information displays which show relevant information has not yet come true. One of the main reasons is the limited number of available displays in the environment which is a fundamental requirement of the original vision. We introduce the concept of an Ambient Mobile Pervasive Display (AMP-D) which is a wearable projector system that constantly projects an ambient information display in front of the user. The floor display provides serendipitous access to public and personal information. The display is combined with a projected display on the user's hand, forming a continuous interaction space that is controlled by hand gestures. The paper introduces this novel device concept, discusses its interaction design, and explores its advantages through various implemented application examples. Furthermore, we present the AMP-D prototype which illustrates the involved challenges concerning hardware, sensing, and visualization.

Author Keywords

Pervasive displays; Augmented reality; Personal projection

ACM Classification Keywords

H.5.2 [Information interfaces and presentation]: User Interfaces. - Graphical user interfaces.

INTRODUCTION

Today, we observe an ever increasing interest of mobile users towards pervasive information access and serendipitously discovering new information. This is currently achieved by

means of smartphones and public displays in the environment. Public displays will likely never be widespread enough to fulfill this desire alone. Inversely, smartphones can only contribute to this vision when they are held in hand and are actively operated. This becomes a challenge especially when the user is on-the-go as the device distracts the user's focus and connection to the environment. The idea of wearing a location-aware pervasive display that alerts to new relevant information and provides quick access to deal with it, is very compelling. It could provide access to a great variety of information, ranging from *public* content (e.g. stationary pictures from Flickr) to *personalized* content (e.g. personalized advertisements in front of shopping windows (Figure 1a)), and *personal* content (e.g. notifications about new text messages). Previous works on mobile speech interfaces, head-mounted displays, public-display networks, and mobile projectors attended to this vision one way or the other. One crucial requirement for such a wearable pervasive display is that the display is always available. Another is that the display is located in the user's periphery and uses ambient presentation to minimize the risk of annoying users and distracting them from their primary tasks.

In this paper we propose to use constant personal projection as a novel pervasive display. The Ambient Mobile Pervasive Display (AMP-D) integrates a wearable projected display that accompanies users on the floor in front of them with a projected hand display and a smartphone to a continuous interaction space. The floor display is a constant pervasive window into the user's digital world, lying in the user's visual periphery (see Figure 1a). As such it allows for subtle user alerts without occupying the user's hands or field of view (FOV). The hand display allows to deal with information instantly without having to reach to a physical device. The smartphone supports exploring and sharing content from and to the virtual world. Users interact with AMP-D entirely through hand

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. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
CHI 2014, April 26–May 1, 2014, Toronto, Ontario, Canada.
Copyright © 2014 ACM 978-1-4503-2473-1/14/04..\$15.00.
<http://dx.doi.org/10.1145/2556288.2557365>

gestures. For instance, when a new text message is received, a message notification box rolls into the user's view on the floor (Figure 1c). Users can then pick the box up and read the message instantly in their hand (Figure 1d). Optionally, they can answer the message using their smartphone.

In addition to the wearable multi-display, AMP-D uses a consistent information space for typical public and personal mobile content that augments users' virtual world through spatial augmented reality, giving users a natural means of discovering new information.

The contributions of this paper are the AMP-D concept of constant personal projection and its interaction techniques, its prototype implementation, and various implemented application examples that explore and highlight the applicability of AMP-D to typical mobile scenarios. Further, we contribute our lessons learned from the evolution of the prototype and a small user study.

RELATED WORK

Mobile Projection

The recent emergence of pico-projector technology allows for the integration of projectors into mobile and wearable devices [22].

Karitsuka et al. envisioned the concept of a wearable projector that provides an interactive display on any plain surface [11]. SixthSense [17] advances this concept and shows a pendant device and several context-aware application scenarios (e.g., recommendations projected on books or delays on flight tickets). Further, simple pointing interactions in mid-air between the projector and the projection are supported, which have been studied in greater detail by Winkler et al. [28]. OmniTouch [9] employs a shoulder-mounted projector and advances the interaction by using a depth camera to support multitouch interaction on planar surfaces within reach. The SLAMProjector of Molyneaux et al. [18] advances touch interaction even further to arbitrarily shaped surfaces and also supports geometrically correct projections on such. These works focus on a (touchable) display within reach, similar to the hand display of AMP-D. LightGuide [23] demonstrates navigational help on the user's hand. Different to these works, AMP-D also supports passive usage of projected augmentations and multiple projected displays.

Ota et al. explored different positions of the projector on the body in regard to wearability, and the viewability of a floor projection [19]. McFarlane et al. [15] also explored floor-projected support in the context of military tasks with varying interaction techniques. Both found out that the shoulder position is best suited for floor projections, especially if the keystone effect and swinging of the projection can be neglected (as in the world-fixed projection of AMP-D). These works use mobile floor projections, yet do not consider its ambient properties, nor do they support a world-fixed spatial augmented reality as does AMP-D.

The only works that consider combining a *distant* (floor) and a close display are the GuidingLight prototype by Chung et al. [4], and the steerable projection prototype by Cauchard et al. [3]. Although both show working application scenarios

for each of these display locations, they do not define a continuous interaction space between the displays, and do not support interaction in both display configurations.

In contrast, AMP-D enables a mobile multi-display environment with floor and hand projection in a continuous interaction and information space. It further leverages the peripheral characteristics of the floor projection.

Smart spaces

In contrast to mobile setups, static projector setups, belonging to the field of smart spaces, have the advantage that the environment is known in advance or can be instrumented to be analyzable in real-time, e.g. by introducing cameras at known positions. For instance, the LightSpace [27] and Beamatron [26] by Wilson et al. implement a smart space through projection. By dynamically building a 3D model of the room and present actors by means of multiple depth cameras, the steerable projector system can augment any surface in the room with perspective-correct projections for one user. It thereby defines a continuous interaction space that, for instance, allows users to carry virtual objects in their hand. The concept of RoomProjector [18] uses spatially-aware handheld projectors to reveal parts of the 3D scenery instead of fixed projectors. Further, it supports geometrically corrected projections and interactions on arbitrary surfaces. WorldKit provides a method for dynamically creating projected interfaces in the environment [29].

AMP-D aims to bring the underlying concept to the mobile user on the go who has very different information interests (public or smartphone content) and options for interaction, which leads to new conceptual and technical challenges.

Mobile Peripheral and Head-Mounted Displays

Wearable augmented reality displays date back to the works of Feiner et al. [6, 7] who developed head-mounted see-through displays for spatial augmented reality. This display type constantly overlays the user's foveal view making it less suitable for everyday scenarios. Apart from that, head-mounted displays to date entail unsolved challenges such as perceptual issues (e.g. different focus planes and narrow FOV); security issues (e.g. they may distract from or obstruct an approaching danger), and social issues (e.g. collocutors having no means of knowing what I'm looking at).

A mobile *peripheral* display is the eye-q system [5] that uses a low-resolution LED display embedded in a pair of eyeglasses to provide a small set of notifications. An advanced display version is promised by Google Glass [8], whose display lies slightly to the top of the foveal field of view, also qualifying it for ambient display. Unfortunately, at the same time, its position and small size make it less suitable for complex visual output, augmented reality, or direct interaction.

Leung et al. [13] demonstrate a system that constantly projects a representation of the wearer's online social identity to the ceiling above. As its intended audience are spectators it is not designed to be interactive, though.

Discussion

Works on mobile displays so far dealt with close and distant displays separately whereas AMP-D presents a continuous interaction and information space between these display types. Similar continuous interaction spaces have only been presented in static smart-space setups. In this work, we aim to bring this compelling vision to the mobile user in everyday use cases. In these mobile scenarios, ambient display properties are much more important, which so far have been neglected in works on mobile wearable or handheld displays.

CONCEPT AND DESIGN OF AMP-D

The AMP-D is a wearable multi-display system that provides a pervasive window into the user's virtual world on the floor. Unlike smartphones which have to be taken out to be operated, the AMP-D display is constantly available. Therefore it is suited for ambient alerting to many kinds of public or personal information that is available via the user's connected smartphone. Among others, these information types include location-aware notifications, communication, reminders, and navigational instructions. Additionally, information is not only visualized, but can be handled through gestures in the user's hand which is used as on-demand secondary display.

We illustrate the concept of AMP-D by first discussing each of its basic concepts. Following on that we present various use-cases that highlight the applicability of AMP-D. All of these use cases have been implemented in the AMP-D prototype which is presented later.

Basic Concepts

Pervasive and Ambient Floor Display

To provide an always-available display, we think the floor is well suited since it is the only space that is always existent in our current life. Further, it "must" be looked at regularly, at least while moving which is a benefit for notifications that cannot be missed for too long. Besides, it is easy to glance at quickly. Thus AMP-D projects the permanently available display on the floor, yet content is only displayed when required. The floor display lies outside the foveal field of view (FOV) of the user, therefore it is qualified for peripheral display. Research on peripheral vision and cognitive psychology offers evidence that peripheral vision supports a separate cognitive channel, thereby reducing overall cognitive load [25]. More importantly, the effect of tunnel vision supports users in effectively blending out unnecessary information in the periphery when their cognitive load is high [5]. Inversely, when users' cognitive load is low, the display supports the serendipitous discovery of new information.

As the peripheral vision is much more sensitive to movement than to color or detail [5], we adapt the degree of animation on the display to the priority of the shown content and let the user's cognition outbalance load and priority. We have tested this with AMP-D in a pilot study and the effect can be described as similar to the sudden recognition of a nearby small animal such as a mouse on the ground that is only detected when it starts moving, even though it was present before.

Pousman et al. provide a thorough definition of key characteristics of ambient systems [20]. Based on this definition,

to make AMP-D more environmentally appropriate, we "focus on tangible representations in the environment" [20] by refraining from including any typical GUI elements such as windows or buttons on the display. Instead, the projection only shows a projected window into the user's virtual world, i.e. invariably, all projected content is clearly located in the worldwide coordinate system. This concept builds on Spatial Augmented Reality (SAR) [2] and world-fixed presentation [6], as opposed to the standard display-fixed presentation. In the context of projections, it feels like uncovering the virtual world with a spotlight — or in case of AMP-D, a wearable lantern — which is why it is referred to as the spotlight metaphor [21]. The system tracks users' movement and orientation to provide the corresponding illusion (Figure 2).

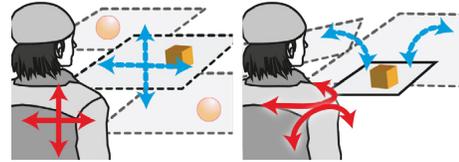


Figure 2. The virtual window follows the user's movement (left) and orientation (right).

As all content is only revealed on a fixed location on top of the real world, the projection blends with the environment, for the user as well as for spectators. The publicity of the projection might also lead to interesting new behaviors. For instance, seeing a person uncover public content such as a sign or advertisement with the projection may lead spectators to explore the item themselves with their own AMP-D (or similar AR) devices. Thus the public floor projection also provides a new way of blending real and virtual worlds between users and spectators.

Information Space: World Graffiti, Boxes, and Spheres

The virtual world of AMP-D consists of only two distinct types of visualizations: two-dimensional *World Graffiti* and two three-dimensional objects; *boxes* or *spheres* (Figure 3).

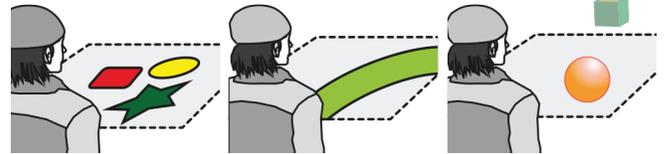


Figure 3. 2D World Graffiti (left), navigation paths (middle), and spheres and boxes as interactive elements (right)

The two-dimensional graffiti is a stationary texture on the ground, such as a navigation path or personalized advertisement. Its flatness indicates that it is not meant to be interacted with. In contrast, the three-dimensional box and sphere items indicate that they are supposed to be interacted with. We choose and limit the system to these two shapes, as it enforces consistency for the user who can approach items from arbitrary angles and they still look familiar. Of course, both visualizations can be combined to create interactive World Graffiti by placing virtual items on top of it.

Spheres are always rolling, accompanying the user wherever they go until they interact with it, or until the sphere is no longer required. For instance, an incoming phone call is represented as sphere item that accompanies the user until the call is taken, declined, or eventually missed. As opposed to

spheres, boxes typically lie at static places. However if they are new and supposed to have an ambient alerting impact on the user (e.g. a notification), they roll into the user’s field of view. If the user is currently moving, they further accompany the user for several seconds before coming to rest.

Boxes and spheres have defined content types which the user can quickly recognize from their different textures. Additionally, new boxes the user has not yet interacted with, carry a yellow border. In this manner, unlike with the use of ambient vibration alerts in smartphones, the user can quickly discern the type and novelty of new notifications by just glancing at the projection. To further interact with the box or sphere, users use their bare hands which are tracked by the system. By reaching out with their splayed hand towards the object, a green selection disk appears in the projection. It acts as hand extension that can be moved beneath the object of interest. The pre-selected item begins slightly bouncing. By closing their fingers, the user selects the object (picks it up) and the object performs a jump animation into the user’s hand (see Figure 4). Initially we tried with a lot of different hand

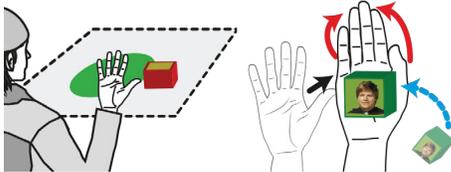


Figure 4. Object selection (left) and pick-up of objects by moving the fingers of the hand together (right).

gestures, especially for picking up objects and releasing them again. Up and down gestures come to mind quickly, but, as other gestures, do not work as they inhibit movement themselves. As the user is moving anyway, we found that gestures based on hand postures work best, followed after gestures that only inhibit horizontal movement.

Private Hand Display

Previous works [10, 9, 23] have demonstrated that various interactions can be performed in people’s hands. The hand display perfectly fits our envisioned interaction scenarios, as many actions can be performed without a separate device. In contrast to the floor display, AMP-D’s hand display supports two-dimensional, display-fixed content.

As soon as content has been picked up to the user’s hand, the focus of the projection changes to follow the user’s hand, showing a focused image within the hand’s boundaries. Consequently, the floor projection becomes blurry. This provides the user with a very suitable means of knowing which interaction zone is currently active.

The hand provides a more private display than the public floor projection, comparable to a phone display. When picked up, many objects can disclose more sensitive information. Message boxes, for example, can show a picture of the sender of the message. Hand gestures allow the user to interact further with the content. By turning the hand 90 degrees towards the center of the body, the user switches to *reading mode*. The box held in the hand shrinks and a preview of the content is displayed. For instance, a text message or the subject of an email as scrolling text, or a preview of a picture can be displayed (see Figure 5).

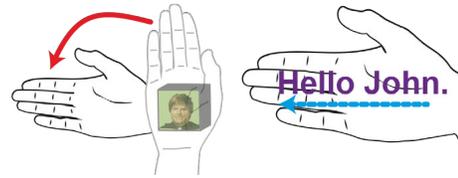


Figure 5. By turning the hand 90° degrees a preview (scrolling text, thumbnail, etc.) of the content in hand is displayed.

Furthermore, binary decisions (e.g. accept a call with or without loudspeakers activated) are supported. Users toggle between binary options by flipping their hands so that the thumb points inwards or outwards and select the option by performing a click gesture by briefly moving the thumb to the palm and back (see Figure 6).

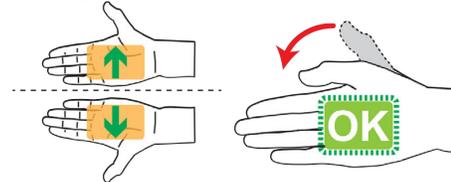


Figure 6. Binary decision gestures.

As long as the user holds the box in hand, it moves with them. This way, users can reposition any three-dimensional item in their virtual world. Finally, users have two options how to proceed with the object: By splaying out their fingers and/or moving their hand outside of the projected area, the item falls down back to the floor in front of them. Or, by performing a gesture as if to throw the item over one’s own shoulder, the item is removed from the virtual world (see Figure 7). The meaning of these gestures depends on the content type (use case) and is explained later.

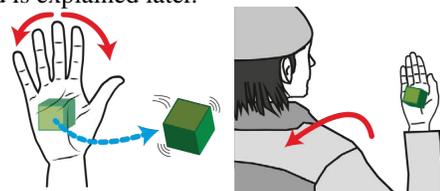


Figure 7. By unfolding the fingers the object falls back to the ground (for notifications this is equivalent to “snoothing” the alert). By throwing the object over the shoulder it is removed from the virtual world.

Continuous Interaction Space: Floor/Hand/Smartphone

In situations when the hand display is not sufficient to adequately deal with content (e.g. to answer a text message), interaction can continue on the smartphone. When a user picks up an object, the smartphone is notified of the selected item. For as long as the user holds the item in hand and a short grace period after that, the smartphone is automatically unlocked and presents the full representation (e.g. the full email view) either immediately, or as soon as it is taken out from where it was stowed (see Figure 8). When users are in a private environment or to support collaboration, they may also want to show full content representations on the large floor display. AMP-D could easily support this through an additional gesture. Also more complex interactions such as text-entry could be supported on the floor projection.

Another usage of the smartphone is to add content from the smartphone to the virtual world. Despite auto-generated

items, the user may want to share content at specific locations. For a personal purpose, for example, a reminder such as “don’t forget to put the bins out” can be placed on the threshold. Moreover, pictures, for instance, can be dropped to the world where they were taken to share them with friends or the public (explained in a moment). The smartphone provides an always available “share service” that allows supported content on the smartphone to be dropped into the environment as virtual boxes of the corresponding type (see Figure 8).

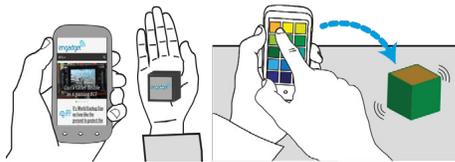


Figure 8. Content in hand can be continued exploring on the phone (left) and arbitrary content from the phone can be placed as box item in the virtual world (right).

Thus, the three interconnected displays form a continuous interaction space providing different levels of details to the user’s requirements at that time.

Public Content And Collaboration

Besides personal content of the user, the constant availability of the projection invites friends, advertisers, or even the local administration to create public content in the virtual world similar to physical signs, banners, etc. The virtual content has the advantage that it is much cheaper to create and can be personalized for the user. Their disadvantage is that the projection is smaller than a person’s physical FOV, therefore they may not reach the same audience. The intrinsic publicity of the projection also invites many co-located multi-user scenarios. For instance, co-located AMP-D users can overlap and “merge” their floor projections to drop content from the smartphone on one side, to then be picked up on the other side by the second user.

Privacy

When using AMP-D, users neither exclusively select the content to project, nor do they monitor the projection all the time as is the case with existing projected displays. Furthermore, the surrounding audience of the projection is rather random. Projected content may be appropriate in the current context, but not in another. Thus, users require effective means to protect sensitive information and ensure that only appropriate information is disclosed. A first means is already given through the concept of SAR. When a user wants to hide information on the floor display quickly and for a short moment only, a small movement or rotation along any axis is often enough to move the window in order to hide previously disclosed items. If this is not sufficient, AMP-D supports a simple horizontal swipe gesture in front of the projector to enable/disable the projection entirely (see Figure 9).

History and Overview

The SAR concept entails another advantage within the context of AMP-D. The implicit revealing or hiding of information using body motion can also be used to look up upcoming content or to revisit past content. For instance, when a user recognizes content on the floor projection too late, walking a few steps back or just turning around will bring the item back

into the projected window. Similarly, when users share their foot trails as World Graffiti, they can revisit them later, e.g. to find their way back to their car. As opposed to that, for instance, tilting the projection far ahead during navigation tasks allows users to preview directions further ahead. Results from a study by Billinghurst et al. [2] indicate that people can easily navigate and relocate spatially augmented information as they are used to the interaction from real life. Therefore, the SAR concept should be able to provide a natural interaction with the information space of AMP-D.

Despite AMP-D’s support in changing the FOV in all directions, no contemporary display technology can compete with the overview (FOV and resolution) of a person’s real view into the distance. Thus, searching for virtual content on the ground can require significantly more effort than searching for real objects. Therefore, the system provides vertical swipe gestures to change between AMP-D’s standard view and an elevated third-person perspective. This acts like a map on a scale of 1:10 to spot nearby objects of interest without having to actually go there or search for them (see Figure 9).

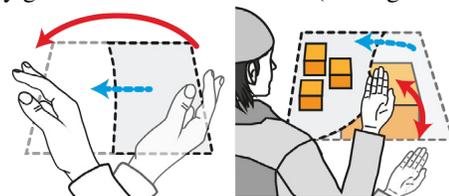


Figure 9. Enable/disable the projection entirely with a horizontal gesture (left). Change from a real to an elevated perspective to get an overview of content around (right).

AMP-D Use Cases

We understand AMP-D as a general mobile companion and the following implemented use cases (cf. video figure) highlight how AMP-D supports typical mobile scenarios.

Context- and Location Awareness

Context- and especially location-aware information such as friends being nearby, or interesting offers in the vicinity are increasingly available to users. With AMP-D being constantly active and capable of displaying visual context information, it is well-suited to provide such information to users on the go. People on a shopping stroll, for instance, see additional offers as World Graffiti and box items (textured with a t-shirt icon) on the ground in front of the shopping windows they pass. By picking up the box and reading its contents, a personalized advertisement appears in the user’s hand (Figure 10a).

We also implemented the system to support persuasive hints. They have been shown to be able to motivate users to change their behavior in a positive way [24]. For instance, when users walk close to an elevator, the system reminds them of their activated fitness plan by showing a red line leading to the elevator and a green line leading to the staircase as World Graffiti beneath the users’ feet (Figure 10b).

Data sharing

Data sharing via the smartphone is used to create new virtual items in the virtual world. Currently, we support two data sharing applications on the smartphone. The first allows the user to create note boxes with text content at world-fixed

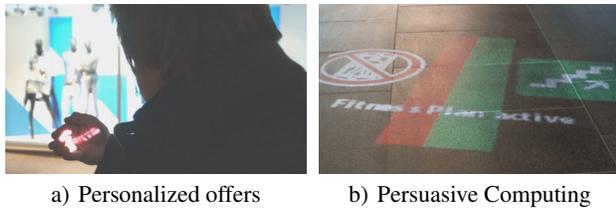
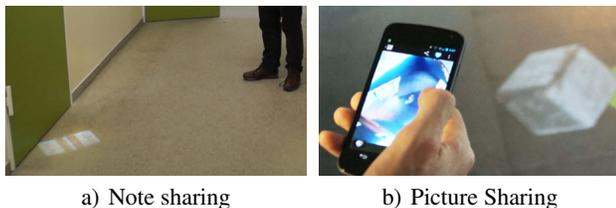


Figure 10. Two location-aware use cases

positions. This can be used to place location-dependent reminders or messages for oneself, or, for example, a colleague or family member in the own environment who will literally stumble over the information (Figure 11a), and can read the contained message in their hand.

The second application supports the sharing of pictures from the smartphone's gallery. The boxes are created right in front of the user (Figure 11b) and are textured with the contained image (in a gray inset, identifying them as boxes of the type "picture"). Given the small size of the boxes, it is not possible to recognize the actual image on the floor projection, but it is already useful to distinguish between different items in a pile of pictures. Once users pick up an image box they are interested in, the image is revealed in the user's hand when entering reading mode. This presentation already delivers a much better impression of the picture than the floor projection. As with other content types, the picture can further be viewed on the phone.

Once boxes are created, they can also be easily repositioned by taking them by hand, moving with the box to the desired location, and releasing them again.



a) Note sharing b) Picture Sharing

Figure 11. a) A note is found in front of an office door ("Meet me at coffee place"). b) A picture box transferred from phone to virtual world.

Notifications: Communication, Reminders, News Feeds

The most frequent tasks performed on smartphones - especially while the user is on the go - are related to communication, and management of notifications. Calendar and task reminders, for instance, have become very popular on smartphones. The most important aspect is to actually read them, be reminded at regular intervals if the notification was missed in the first place, and perhaps perform some simple interaction such as snoozing or dismissing the notification. For the user on the go reading the notification on a smartphone often involves a significant overhead. The user must stop or at least slow down, take out the device, possibly unlock it, only to read a few words of text.

AMP-D uses its box items to visualize new notifications regarding text messages, emails, calendar reminders, and news feed updates. As described earlier, the user can quickly discern the type of the notification from their appearance prior to any interaction (see Figure 1c). When a message box is

picked up, a picture of the sender is displayed on the upper face of the box. The first 160 characters of the item's content are displayed as scrolling text in the user's hand when turned to reading mode (see Figure 1d). Otherwise, only a teaser is displayed, or the subject in case of an email, and the whole message can then be read, for instance, on the smartphone. Similarly, news feed updates appear as feed boxes that show their source (publisher) as a texture on the box, reveal their subject in the user's hand, and can be continued to be read on the user's smartphone. They particularly demonstrate the usefulness of serendipitously stumbling over new information when the cognitive load is low.

The visualization of dynamic notifications using the world-fixed SAR concept is not straightforward as the information is time- and context-dependent instead of location-dependent. Our solution is to multiplex the information in the time and location domain. For instance, when users receive a new notification, it is created at the user's current location and rolled into their projected window. Shortly after the user passes by the notification without picking it up, it is removed from the old position and inserted in the same animated way at the user's current location. Once the notification box has been picked up, users decide whether they want to return the box to their world to either "snooze" the notification or dismiss it by throwing it over their shoulder. In the former case, the box will continue to regularly appear across the user's way but without any type of animation (these intervals can be defined on the corresponding smartphone app). In the latter case, the notification - not the content - is deleted.

Incoming calls, in contrast, are presented as a sphere that accompanies the user for the time of the call. It can be picked up to show the caller's picture - if available - as texture on the sphere and to reveal the name or phone number in the reading mode. In this scenario, taking out the smartphone after having picked up the sphere will automatically unlock the smartphone and accept the call; releasing it to the world will keep it ringing; and throwing the sphere over the shoulder will decline the call.



a) Path navigation b) Signs

Figure 12. Path navigation with additional turn-by-turn instructions and signs that help ahead-way planning

Navigation

AMP-D supports augmented reality navigation where the user follows a path as a virtual line overlaying the ground. The path approach fits the concept of World Graffiti on the floor display, since it gets by without time-based instructions

the user could miss. Instead, whenever the user pays attention to the projected navigation, directions can be grasped at a glance (Figure 12a). Additionally, to provide the necessary context for users to plan their way upfront, turn-by-turn navigational instructions (e.g. turn left in 50 meters) and location-dependent help (e.g. take the stairs up, not down) are overlaid right next to the path similarly to road signs (Figure 12b).

PROTOTYPE

We implemented a prototype of the AMP-D concept in order to investigate the technical challenges involved in building such a system and explore possible interactions with aforementioned use cases.

Hardware design

Our AMP-D prototype (see Figure 13) consists of a procam (projector and camera) unit, a backpack, a laptop, and two batteries (one for the projector, one to increase battery time of the laptop). Part of the overall system is also a wirelessly connected Android smartphone running a corresponding service software.

The procam unit (Figure 13b) is attached to the backpack (Figure 13a) that positions it approximately 25cm to the side and 15cm to the top away from the user's eyes (no offset in the forward direction). We found this position to yield a good trade-off between the large size of the projection on the ground (appr. size $2m \times 1.25m$, 2 meters away) and the positioning of the hand in order for the user to comfortably reach into the projection for the selection of objects and the hand display.

On top of the projector sits an ASUS Xtion Pro Live depth and RGB camera ($640 \times 480px$ @ 30 FPS), which we chose for its very small size, low power consumption, and well-suited depth range (0.5m to 3.5m). Finally, an inertial measurement unit (IMU) by x-io technologies is part of the procam unit and delivers precise and robust 9DOF orientation and acceleration data of the user. The system is controlled by a Dell M11x laptop with i7 CPU, 1.7 GHz running Windows 7 and the prototype software that performs all computations at a frame rate of 60 Hz.

The projector and the IMU are powered by batteries and the rest of the components are powered by the laptop. The system's power lasts for 5 hours of continuous usage. The system can be worn to both sides, depending on the primary hand of the user, which should be on the same side as the projector to be within the projection path.

Software

The software components (except smartphone service software) are responsible for computing the graphics and physics of the 3D world augmentation, the user's movement, and the hand and gesture recognition. The software is written in C# and runs on the laptop.

3D World Augmentation

The virtual user and all projected content are modeled in a 3D scene using Microsoft's XNA, the JigLibX [14] physics engine, and a graphics framework [16]. The rendering of this scene delivers the 3D world augmentation to the projector.

The 3D scene includes World Graffiti as 2D floor textures, and 3D boxes and spheres. The skeleton of the virtual user, who moves through this world, consists of a lower and an upper body entity. The correct perspective projection is achieved by attaching the virtual camera to the user's torso entity with the exact offset that it has in reality. The engine will then compute the correct off-axis perspective homography that lets the projection appear as perceived through the user's virtual eyes. Moreover, it lets the virtual camera turn around the center of the user's body instead of turning around itself. In addition, the virtual field of view has to be inversely matched to the actual field of view provided by the projector. Currently, we do not account for lens distortion of the projector which would further improve registration accuracy.

As we use a fixed physical orientation for the projector (instead of adapting it to the user's height), we can calculate the user's height as required by the system automatically based on the floor distance we receive from the depth sensor. Thus the system does not require manual calibration. The accuracy of the optical illusion during tilting or rolling of the torso can be further improved, though, by providing the exact length of the torso to the system in order to accurately determine the center of the body.

Floor/Hand Tracking and Focus Adjustment

Floor and hand tracking is computed on the depth image from the camera. On every cycle, the algorithm first decides whether the user's hand is present in the depth image:

We use computer vision to recognize hand contour, finger gaps and tips, fingertip direction, the direction of the hand, and the centers of palm and hand (see Figure 13c left). The recognition builds on three-dimensional segmentation of hand-sized clusters and simple heuristics based on sizes, distances, and changes in the derivation of the contour. Our particular shoulder-worn setup allows some assumptions that further simplify the recognition procedure: valid hands must not be further away than 1.5m (depth culling); must not span a depth range larger than 0.5m; and the user's arm (the cluster) must always reach into the camera frame from the bottom and/or right edge (for right-handed users). The recognition is fast and accurate in various environments. When more than one finger is recognized, we detect the *unselected* state that allows the user to steer the green selection disk (cf. Figure 4) for object selection. When one or no fingers have been recognized, we detect the *selected* state. Further, we recognize the user's thumb and compute its relation to the center of the hand to distinguish between the two states of binary decisions. Comparing hand positions, hand directions, and finger counts over multiple frames allows us to recognize the remaining gestures such as *reading mode* (cf. Figure 5), click gesture, and the horizontal and vertical *swipe gestures* (cf. Figure 9).

When the user's hand is *not* detected, the surface in front of the user is analyzed to decide whether it is suitable for showing the floor projection. The depth image is sampled at several grid-based points across the image and first averaged individually for each row, then for the whole image. Based on the depth gradient from individual rows we can decide whether the projection falls on a floor-like (vertical), mostly

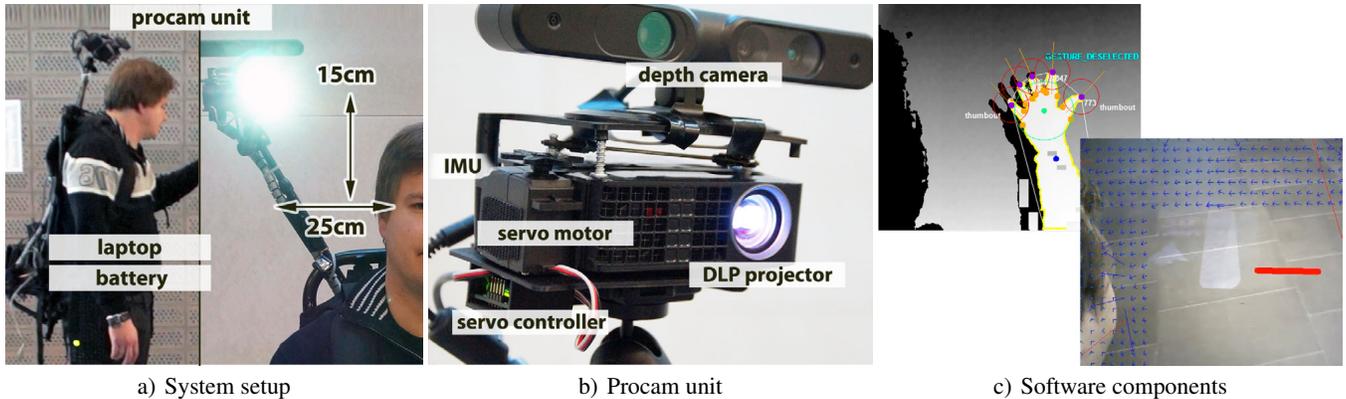


Figure 13. The AMP-D prototype. a) A backpack holds the procam unit, laptop, and battery (left). The procam unit appears as worn on the shoulder (right). b) Close-up of the procam unit. c) Software tracks hands for gesture-based interaction and computes optical flow to enhance step detection.

plain (depth deviation) surface. Additionally, based on the overall depth average, we can then adjust the projector’s focus to show a sharp image on the floor.

Tracking of Orientation and Movement

In parallel, inertial sensor data is received from the IMU. It is used to compute the orientation of the user’s torso in all three dimensions to adjust the virtual user and the attached virtual camera in the 3D world accordingly.

Additionally, we use the acceleration data from the IMU for step detection. As absolute positioning systems are not always available, particularly indoors, AMP-D needs a way of detecting the user’s movement based on dead reckoning. Naturally, this will only work for a short time due to measurement errors and must be regularly corrected by reliable data from an absolute positioning system (e.g. GPS or indoor localization systems increasingly becoming deployed). For the sake of testing the AMP-D concept, we only require short movements for which the dead reckoning approach is sufficient. Algorithms for robust pedestrian navigation usually build on a technique known as zero-velocity-update that relies on data from an accelerometer attached to the user’s foot (e.g. [12]). Following our initial vision of a palm-sized form factor of the system (Figure 1a), we want the system to get by without further user instrumentation. With the IMU unit attached to the procam unit, we cannot rely on the zero-velocity-update method. Instead, we detect steps by finding peaks of vertical and forward acceleration, which are homogeneous in human walking behavior. Step length is approximated based on the automatically calculated height of the user. With the IMU unit alone, we could not reliably detect the user’s walking *direction*, though.

A working solution which increased the reliability of detecting the step direction was found in computing the optical flow of the camera’s RGB image. More precisely, we calculate the optical flow in a 100 px wide border at the top and left side of the RGB image (for right-handed users) wherein the user does not interfere while interacting with the primary hand (see Figure 13c right). Optical flow towards the user indicates forward movement while optical flow away from the user indicates backward movement. Gyroscope data is used to counterbalance the effect on the optical flow generated by

the up and down swings caused by the human gait.

Combining these approaches, our system can detect the user’s forward and backward steps most of the time, although there is room for improvement. By decreasing the form factor of the prototype, for instance, the system can be brought closer to the user’s body which will benefit a more accurate step detection. Nonetheless, a general problem with step detection based on inertial measurements would remain: as the algorithm cannot detect a step until it is close to being finished, a small computational delay is induced. This delay counteracts the optical illusion when walking starts or stops and sometimes leads users to walk one step further than they intended.

Smartphone Service

The Android smartphone is connected to the laptop via Wi-Fi and runs a background service which starts polling the phone’s light sensor whenever the user takes a virtual box into their hand and stops soon after it was released again. Whenever the measured light significantly increases during this time interval, the service wakes the screen, disables the keyguard, and starts the corresponding activity showing the content related to the box. In addition, access to placing information in the world from the smartphone is provided through a notification service. By pulling down Android’s notification center from the phone’s status bar and selecting the AMP-D service, the user can create notes and select pictures from the gallery which are then dropped into the world.

Limitations and Improvements

For most indoor scenarios, the brightness of the displays of the present prototype are already sufficient. For most outdoor scenarios, only the hand display is sufficiently visible as it is very close to the projector and can be shielded against direct sunlight. To the floor projection these reasons do not apply, hence the prototype is currently limited to twilight conditions outdoors. As the brightness of pico-projectors increased four-fold between 2009 and 2011, we hope they will reach sufficient brightness for outdoor conditions in the future.

Another limitation is the current size of the system. With pico projectors advancing quickly, and as depth cameras of the size of a finger becoming available, the procam unit can likely be shrunk considerably in the near future. Additionally,

the software should soon be able to run on an on-board system such as a Raspberry Pi. Power consumption will likely be the most challenging factor for a much smaller version. This could however be mitigated by intelligent power saving which reduces power consumption when no display surface is available.

Finally, the step detection needs to be further improved, e.g., by pointing a second camera towards the user's feet which can immediately decide whether the user starts moving, thereby eliminating the initial detection delay of the current system. Further, GPS can be used outdoors or geo-located image-based pose estimation indoors (cf. [1]) to correct step detection errors.

Preliminary Evaluation

While we are planning on conducting larger user studies using our prototype, in a first initial investigation we wanted to find out if the most important features of AMP-D work for untrained users. Thus we recruited 6 participants between 25 and 30 years (mean 27 years), to identify strengths and weaknesses of the concept or the current implementation. They have been smartphone users for 1.5 years on average (min 0.5, max 2 years) and all used their smartphones at least for messaging, web browsing, calendar, and traffic information. The study lasted between 45 and 60 minutes and was conducted in a public floor (12 x 7 meters) of our building with regular by-passers (to provide users experiencing their attitude towards public usage). First we asked participants (using 5-point Likert-scales from "strong agree" to "strong disagree") about their smartphone usage and all showed *strong agreement* that they receive regular updates (notifications) on their mobile phones. There was further *agreement* to check if notifications have been missed, also while on the go, by all participants. Finally, there was *strong agreement* by all participants that they usually react to new message immediately. These answers show that our participants were in the right target group addressed by AMP-D.

After that we had all participants try out all applications of the prototype. This includes: receiving boxes rolled into their view while walking looking straight ahead, picking up boxes, reading their contents, moving them, releasing and dismissing them. Further continuing reading the contents of a box on the phone as well as taking a picture with the phone and creating a reminder note on the phone and sharing both to the own virtual world. Finally, they also tried to follow a navigation path that led them a parkour around obstacles we had set up.

After having tried the AMP-D prototype, participants showed a generally very positive attitude towards the AMP-D. Again we asked using the same 5-point Likert scale. All participants at least *agreed* that they recognized new notification items on the floor without looking at them. Further, all but one assumed the system would not disturb but enrich their daily life if it was available at a smaller size. Further, all participants at least *agreed* that they think they could react to new information quicker using AMP-D versus a smartphone. Finally, all *agreed* that the prototype worked fine for them, that they enjoyed using it, and that they could handle its complexity. In contrast, users were split in their answers to our questions regarding social acceptance and price/performance ratio – con-

sidering AMP-D would double the costs of the smartphone – both resulting in overall *neutral* feedback.

In response to open ended questions participants criticized, for instance, physical fatigue caused by the high number of interactions tested in the user study. Two participants were concerned with performing large, eye-catching gestures in public space. We also received constructive comments regarding technical challenges like brightness, battery life, and size of the system. One participant, for instance, proposed to show and select between all objects in the vicinity along a virtual string in the hand when the floor display is not bright enough. On the other hand, participants suggested several further application scenarios, among those: using AMP-D for navigation and context-aware instructions for craftsmen on building sites; remotely placing reminder boxes for items to buy across the supermarket at the right locations (like a decentralized shopping list); similarly, using AMP-D as city tour guide with POIs realized as info boxes to stumble over interesting information while keeping connected to the primary interest, the environment.

Naturally, this evaluation is only a first step in evaluating the device. User studies with more users over longer periods and against ordinary smartphone usage, for instance, are required.

CONCLUSION & FUTURE WORK

In this paper we introduced the Ambient Mobile Pervasive Display (AMP-D). For the first time we propose constant personal projection and have shown a working implementation of it, solving technical and conceptual issues. Our prototype provides a wearable multi-display system that combines a pervasive ambient floor display, a private hand display, and a smartphone into a continuous interaction space for mobile everyday scenarios. This goes beyond previous works on wearable augmented reality or wearable multi-display environments as the interaction space between multiple mobile displays have not been considered, yet. We have demonstrated new interaction metaphors and use cases which highlight the applicability of AMP-D in mobile scenarios. Moreover, the presented applications are embedded into a consistent information space that uses spatial augmented reality together with World Graffiti and virtual items to cover a broad range of interaction scenarios. Our concept further presents a new approach to serendipitous access to digital information that can be applied to our physical world, thereby likely reducing the individual's effort to receive and deal with information. As such, it depicts a future direction towards the original vision of pervasive computing.

We contribute a realistic interaction concept and a complex prototype that demonstrates unique solutions to the specific challenges of the AMP-D concept: most noteworthy, automatically changing the projector's lens focus based on the user's zone of interaction, step detection fusing inertial and optical movement measurements, and tracking of novel hand gestures in a truly mobile environment. We have integrated these components to a standalone mobile system that does not require instrumentation of the environment or the user (despite wearing the system), and runs for several hours.

The positive results of our small evaluation call for larger user

studies that assess how the display can blend into the fabric of everyday life. It would also be interesting to test AMP-D in public spaces, both to study the influence of AMP-D on spectators, and to see how well AMP-D performs in crowded places.

ACKNOWLEDGMENTS

This work has been conducted in the context of the DFG Emmy Noether research group *Mobile Interaction with Pervasive User Interfaces*. We thank Ken Pfeuffert for contributing to the gesture interface, and Florian Schaub and Markus Löchtfeld for very helpful discussions.

REFERENCES

1. Arth, C., Reitmayr, G., and Schmalstieg, D. Full 6dof pose estimation from geo-located images. In *Proc. of, Computer Vision-ACCV 2012*. Springer, 2013, 705–717.
2. Billingham, M., Bowskill, J., Dyer, N., and Morphett, J. Spatial information displays on a wearable computer. *IEEE Computer Graphics and Applications* 18, 6 (1998), 24–31.
3. Cauchard, J. R., Fraser, M., Han, T., and Subramanian, S. Steerable projection: exploring alignment in interactive mobile displays. *Personal Ubiquitous Comput.* 16, 1 (Jan. 2012), 27–37.
4. Chung, J., Kim, I.-J., and Schmandt, C. Guiding light: Navigation assistance system using projection based augmented reality. In *Proc. of, ICCE '11* (Jan. 2011), 881–882.
5. Costanza, E., Inverso, S. A., Pavlov, E., Allen, R., and Maes, P. eye-q: eyeglass peripheral display for subtle intimate notifications. In *Proc. of, MobileHCI '06*, ACM (2006), 211–218.
6. Feiner, S., MacIntyre, B., Haupt, M., and Solomon, E. Windows on the world: 2 d windows for 3 d augmented reality. In *Proc. of, UIST '93*, ACM (1993), 145–155.
7. Feiner, S., MacIntyre, B., Höllerer, T., and Webster, A. A touring machine: Prototyping 3D mobile augmented reality systems for exploring the urban environment. *Personal Technologies* 1, 4 (1997), 208–217.
8. Google Inc. Google glass - home. <http://www.google.com/glass/start/>. Accessed: 2013-03-21.
9. Harrison, C., Benko, H., and Wilson, A. D. OmniTouch: wearable multitouch interaction everywhere. In *Proc. of, UIST '11*, ACM (2011), 441–450.
10. Harrison, C., Tan, D., and Morris, D. Skinput: appropriating the body as an input surface. In *Proc. of, CHI '10*, ACM (2010), 453–462.
11. Karitsuka, T., and Sato, K. A wearable mixed reality with an on-board projector. In *Proc. of, ISMAR '03*, IEEE Computer Society (2003), 321–322.
12. Ladetto, Q., and Merminod, B. Digital magnetic compass and gyroscope integration for pedestrian navigation. In *9th International Conference on Integrated Navigation Systems, St. Petersburg* (2002), 111–120.
13. Leung, M., Tomitsch, M., and Vande Moere, A. Designing a personal visualization projection of online social identity. In *Extended Abstracts, CHI EA '11*, ACM (2011), 1843–1848.
14. Linneweber, T. JigLibX physics engine. <http://jiglibx.codeplex.com/>. 2013-03-27.
15. McFarlane, D. C., and Wilder, S. M. Interactive dirt: increasing mobile work performance with a wearable projector-camera system. In *Proc. of, Ubicomp '09*, ACM (2009), 205–214.
16. Microsoft. Quickstart game engine. <http://quickstartengine.codeplex.com/>, 2013. (Last Access: 2013/03/25).
17. Mistry, P., Maes, P., and Chang, L. WUW - wear ur world: a wearable gestural interface. In *Extended Abstracts, CHI EA '09*, ACM (2009), 4111–4116.
18. 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. of, Pervasive 2012*, Springer (June 2012), 197–215.
19. Ota, S., Takegawa, Y., Terada, T., and Tsukamoto, M. A method for wearable projector selection that considers the viewability of projected images. *Computers in Entertainment (CIE)* 8 (Dec. 2010), 17:1–17:16. ACM ID: 1902596.
20. Pousman, Z., and Stasko, J. A taxonomy of ambient information systems: Four patterns of design. In *Proc. of, AVI '06*, ACM (New York, NY, USA, 2006), 67–74.
21. Rapp, S. Spotlight navigation: a pioneering user interface for mobile projection. In *Ubiprojection Workshop at Pervasive '10* (2010).
22. Rukzio, E., Holleis, P., and Gellersen, H. Personal projectors for pervasive computing. *IEEE Pervasive Computing* 11, 2 (Feb. 2011), 30–37.
23. Sodhi, R., Benko, H., and Wilson, A. LightGuide: projected visualizations for hand movement guidance. In *Proc. of, CHI '12*, ACM (2012), 179–188.
24. Torning, K., and Oinas-Kukkonen, H. Persuasive system design: state of the art and future directions. In *Proc. of, Persuasive '09*, ACM (2009), 30:1–30:8.
25. Treisman, A. Preattentive processing in vision. *Computer Vision, Graphics, and Image Processing* 31, 2 (Aug. 1985), 156–177.
26. Wilson, A., Benko, H., Izadi, S., and Hilliges, O. Steerable augmented reality with the beamatron. In *Proc. of, UIST '12*, ACM (2012), 413–422.
27. Wilson, A. D., and Benko, H. Combining multiple depth cameras and projectors for interactions on, above and between surfaces. In *Proc. of, UIST '10*, ACM (2010), 273–282.
28. Winkler, C., Pfeuffer, K., and Rukzio, E. Investigating mid-air pointing interaction for projector phones. In *Proc. of, ITS '12*, ACM (New York, NY, USA, 2012), 85–94.
29. Xiao, R., Harrison, C., and Hudson, S. E. Worldkit: Rapid and easy creation of ad-hoc interactive applications on everyday surfaces. In *Proc. of, CHI '13*, ACM (New York, NY, USA, 2013), 879–888.