Towards a mobile Everywhere Display based on a depth-camera projector unit

Markus Löchtefeld
German Research Center for Artificial Intelligence (DFKI)
Stuhsatzenhausweg 3, 66123 Saarbrücken, Germany
markus.loechtefeld@dfki.de

Michael Braun
Saarland University
Campus 66123, Saarbrücken, Germany
braun.m@me.com

ABSTRACT
In this paper we discuss the possibility to recreate the classic work of Pinhanez on the Everywhere Display for a mobile setting. While embedded pico-projectors are becoming more and more popular in mobile devices such as Video cameras or mobile phones, they still lag convincing application scenarios. To investigate the possibilities of such devices we created a hand-held prototype for a future mobile device. It is semi-mobile and consists of a depth camera that is mounted on top of a mobile projector. Based on this prototype we discuss how such devices will be able to create ubiquitous personal displays on every suited real world surface in the future.

Author Keywords
Mobile projection, depth camera, Kinect, Everywhere Display

ACM Classification Keywords
H5.2. Information interfaces and presentation – User Interfaces

General Terms
Interaction Techniques, System Design

INTRODUCTION
With recent advances in projection technology, Projector Phones – mobile phones with integrated pico-projector – have become available to the mass market. Even though their capabilities are still limited due to the low brightness and resolution of the build-in projectors, they promise to allow the creation of personal displays everywhere in the environment around the user. Besides the hardware limitations of such devices another problem is visible; these devices lag convincing application scenarios. Besides pure presentational scenarios such as the projecting images, slideshows or videos just a few appealing application scenarios have been explored. A good overview on existing research can be found in [13].

In contrary to that in the past a lot of promising application scenarios for personal displays have been researched with static or semi-static displays. The classic work of Pinhanez – the Everywhere Display [9] – allowed to have situated, personal displays on nearly every surface in an instrumented room. Using a projector with a steerable mirror attached and a digital 3D model of the environment that the projector was placed in, it was possible to create undistorted interactive displays. These displays have been used for a variety of different applications.

In this paper we discuss our approach to adapt the classic work of Pinhanez for a mobile setting. By using a depth camera that is mounted to a projector we allow direct interaction with the projection. We discuss our implementation to detect touches on arbitrary surfaces in depth and lay out our future work on a real time solution that allows creating undistorted projections on different non-planar surfaces. Our current prototype, even though it is connected to a computer, is not fully mobile but we argue that in the near future mobile devices with built in depth sensing capabilities can be build.

The remainder of the paper is structured as follows; first we discuss the related work, followed by a short description on possible application and interaction techniques. Afterwards we will discuss our current prototype and implementation. We will conclude the paper with a discussion about
feasibility of a fully mobile implementation as well as an outlook of future work.

RELATED WORK
With Pinhanez’ Everywhere Display a variety of applications where shown that made use of the possibility to create personal displays in nearly every corner of the users environment [9]. While Pinhanez’ approach used a fixed projector and a mirror, later approaches employed steerable projectors [15]. These approaches have the advantage that they leave the user free to roam but they demand a 3D model of the room as well as an instrumented environment. Initial research in the field of mobile projection has been conducted by Raskar et al. [11]. While their first prototype, the iLamps, mainly focused on creating a distortion free projection or combining several projectors to create a brighter or larger image, their follow-up prototype the RFIG Lamps [12] allowed initial interactions with objects in the environment. For example the RFIG Lamps where able to highlight products in a shelf where the date of expiration was closed to the actual date. The spotlight metaphor has used by Blasko et al. [1], Rapp et al [10] and Cao et al. [3]. While the earlier two adapted the spotlight metaphor with a dynamic peephole, Cao et al. used the spotlight metaphor to explore a workspace around multiple users. Also the effect of multiple displays have been studied [2], showing that the alignment is a key factor. With ShadowPuppets Cowan et al. [4] presented an interaction technique for a multi-user scenario. By casting a shadow, one could interact with the projection. Since this shadow would be to big when the user who is holding the device would cast it, it is not suited for a single user scenario. Direct manipulation using a depth camera of a projection was investigated lately by Harrison et al. [6]. Even though most of the work on mobile projection provides convincing scenarios for applications only few concentrated on undistorted projections [3,11,12,14] and those are either based on a 3D model [3,14] or structured light [11,12]. Our approach on contrary is feasible on nearly every surface and allows direct interaction. With Kinect Fusion Izadi et al. [7] presented an approach to create highly accurate 3D models of scene using a hand-held Kinect in real time. While their approach makes use of a GPU Pipeline extension our approach tries to solve the problem in a more basic manner to enable the fast transportation towards a mobile platform, which normally have limited resources.

INTERACTION TECHNIQUES AND APPLICATIONS
The Everywhere Display had two main innovations; one was the distortion free projection the other one was the interaction technique. Using direct touch interaction allowed to easily and intuitively manipulate the created displays. Even though this is a straightforward interaction technique in a case where all the possible display configurations can be controlled this is not the case in a mobile setting. Social restrictions prevent us from projecting and touching some surfaces. For example one wouldn’t touch the wallpaper while being a guest at somebody else’s place. Therefore we decided to enable direct manipulation by hovering approximately 2cm over the surface instead of needing to touch. In first experiments this was widely accepted as being still intuitive while at the same time more socially acceptable. Our algorithm is flexible enough to also be applicable to foot interaction. This also allows a floor projection whenever a suited floor is available. If the user projects further away then any of his extremities can reach a different input technique is needed. For this we implemented a simple cursor control using a touchpad that is mounted on top of the depth-camera projector unit.

For the Everywhere Display the normal Windows applications where used. We followed this example and implemented the touch detection and cursor control using the Windows 7 multi-touch driver. Furthermore we developed a few special applications such as an image browser and a mapping app.

IMPLEMENTATION
Our current hand-held prototype consists of a Microsoft Kinect for Windows, a Microvision ShowWX+ and a Samsung Galaxy S. The Galaxy S was added for cases where the user is too far away from the projection to reach it. Using the touch screen he can control the mouse cursor in the projection and by double tapping he can click on the items. We choose a Kinect for Windows since it allows depth-sensing beginning at a distance of approximately 40cm. To reduce weight, the housing and pedestal of the Kinect where removed (compare Figure 2). All three devices where mounted onto each other in such a way that it is possible to hold them in one hand. With an overall weight of 376g it is still light enough to handle it with ease. The Kinect and projector are connected to a Notebook by

Figure 2: Prototype consisting of a stripped down Kinect, a Microvision SHOWWX+ and a Samsung Galaxy S.
cable while the communication between the phone and the notebook is handled via WiFi.

The software is fully implemented in C# using the Microsoft Kinect for Windows SDK and EmguCV for the image processing. For the creation of the 3D model we use OpenTK, a low-level OpenGL wrapper for C#. The touches of the mobile phone are sent to the application using a socket connection.

**Touch detection**
The detection of the touches is split in to two steps, one is the detection of the hand and the other is the decision whether or not it is touching. The hand detection is done using the following operations based on the depth image (compare Figure 3):

1. Canny-edge detection to determine the hand
2. Dilation and Erosion (10x10 pixel box)
3. Determine contours (findContours with > 1000pixel)
4. Fill contours

The touch detection is then done as follows:

1. Determine the convex hull and its center
2. Determine the edge of the finger
3. Determine the distance to the background (Ray from center to edge of hull extended by 20%)
4. Threshold of distance

(All thresholds where experimentally determined.) Even though is a rather naïve approach it is working quite sufficiently even on curved surfaces and corners, the projection correction is still to slow to be really applicable. Furthermore we want to investigate different application scenarios, for example one could imagine registering the applications in space using a 3D SLAM algorithm and access the applications by pointing in the direction they were registered as it has been done by Feiner et al. [5]. Additionally Augmented Reality applications could be taken into account; e.g. a real paper-based calendar could be merged with the digital one on the device [14]. Also the surface that is detected by the Kinect could be used as a playground and a physics engine could be applied to it. With that a recreation of [8] in three dimensions could be investigated.

**Projection correction**
The correction of the projection is basically done the same way as Pinhanez did it. We use a 3D model and a virtual camera that has the inverted optical configuration as the Microvision projector. The only difference to Pinhanez work is that our 3D model is created from the depth data of the Kinect on the fly. Therefore we down sample the 640x480 pixel point-cloud to a 100x75 3D mesh and import this in our 3D scene using OpenGL respective OpenTK. On top of this mesh the maximum projection area is determined and a texture of the application that should be projected is placed inside it keeping a rectangular shape. The screenshot of this texture is afterwards projected. Currently our implementation is running at approximately 14FPS (while at the same time the touch detection is done) and still needs some improvements to allow real time correction.

**DISCUSSION & FUTURE WORK**
Our current hand-held prototype is not fully mobile but provides enough mobility to explore the design space of such a future device. One might think that due to physical limitations the integration of a Kinect-like depth sensor might not be possible since it needs a certain distance between the infrared projector and the infrared detector. This distance is just 9cm, for comparison a Samsung Galaxy Nexus is 12.5cm long so that an integration would be possible. Of course the infrared projector and detector would need to undergo a certain miniaturization since they are approximately 2cm long, but still it could be made possible with technological advances. The most problematic thing will be the power supply for such a device, but for example the Asus Xtion Sensor, which is comparable to the Kinect is powered only by USB. Also other forms of depth sensing such as stereo cameras could be used. Different form factors instead of just handheld are also imaginable, for example a necklace or integrated in a helmet.

For future work we first want to improve the algorithms. While the touch detection is working quite sufficiently even on curved surfaces and corners, the projection correction is still to slow to be really applicable. Furthermore we want to investigate different application scenarios, for example one could imagine registering the applications in space using a 3D SLAM algorithm and access the applications by pointing in the direction they were registered as it has been done by Feiner et al. [5]. Additionally Augmented Reality applications could be taken into account; e.g. a real paper-based calendar could be merged with the digital one on the device [14]. Also the surface that is detected by the Kinect could be used as a playground and a physics engine could be applied to it. With that a recreation of [8] in three dimensions could be investigated.
CONCLUSION
In this paper we presented our approach towards a mobile adaption of Pinhanez’ Everywhere Display. Using a depth-camera mounted to a projector we are able to interact and correctly project on nearly every surface. Besides a short discussion about the technical feasibility of such a device we presented different suited application scenarios as well.

REFERENCES
15. Spassova, M., Fluid Beam - A Steerable Projector and Camera Unit In ISWC'04