

Flexpad: Highly Flexible Bending Interactions for Projected Handheld Displays

Jürgen Steimle
MIT Media Lab
Cambridge, MA
steimle@media.mit.edu

Andreas Jordt
Christian-Albrechts-Universität
Kiel, Germany
jordt@mip.informatik.uni-kiel.de

Pattie Maes
MIT Media Lab
Cambridge, MA
pattie@media.mit.edu

ABSTRACT

Flexpad is an interactive system that combines a depth camera and a projector to transform sheets of plain paper or foam into flexible, highly deformable, and spatially aware handheld displays. We present a novel approach for tracking deformed surfaces from depth images in real time. It captures deformations in high detail, is very robust to occlusions created by the user's hands and fingers, and does not require any kind of markers or visible texture. As a result, the display is considerably more deformable than in previous work on flexible handheld displays, enabling novel applications that leverage the high expressiveness of detailed deformation. We illustrate these unique capabilities through three application examples: curved cross-cuts in volumetric images, deforming virtual paper characters, and slicing through time in videos. Results from two user studies show that our system is capable of detecting complex deformations and that users are able to perform them quickly and precisely.

Author Keywords

Flexible display; handheld display; tracking; projection; depth camera; deformation; bending; volumetric data.

ACM Classification Keywords

H5.2. User interfaces: Graphical user interfaces (GUI), Input devices and strategies, Interaction styles.

INTRODUCTION

Projecting visual interfaces onto movable real-world objects has been an ongoing area of research, e.g. [3, 15, 28, 35, 11]. By closely integrating physical and digital information spaces, they leverage people's intuitive understanding of how to manipulate real-world objects for interaction with computer systems. Based on inexpensive depth sensors, a stream of recent research presented elegant tracking-projection approaches for transforming real-world objects into displays, without requiring any instrumentation of these objects [37, 11]. None of these approaches, however, interpret the deformation of flexible objects.

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.

CHI 2013, April 27–May 2, 2013, Paris, France.

Copyright © 2013 ACM 978-1-4503-1899-0/13/04...\$15.00.

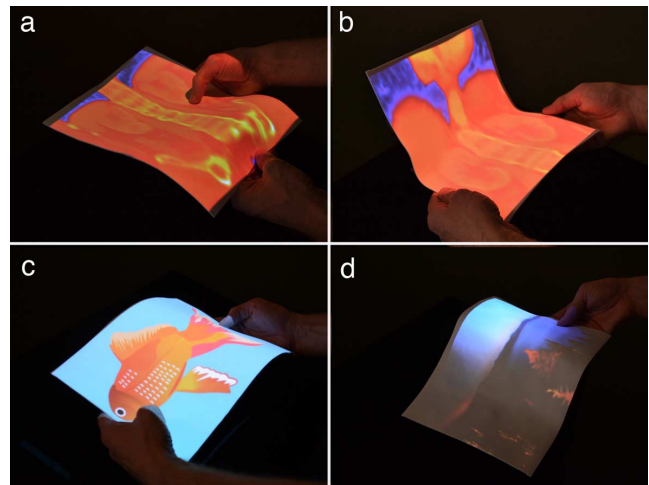


Figure 1: Application examples of Flexpad: Curved cross-cuts in volumetric images (a, b); animation of characters (c); slicing through time in videos (d)

Flexible deformation can expand the potential of projected interfaces. Deformation of everyday objects allows for a surprisingly rich set of interaction possibilities, involving many degrees of freedom, yet with very intuitive interaction: people bend pages in books, squeeze balls, model clay, and fold origami, to cite only a few examples. Adding deformation as another degree of freedom to handheld displays has great potential to add to the richness and expressiveness of interaction.

We present Flexpad, a system that supports highly flexible bending interactions for projected handheld displays. A Kinect depth camera and a projector form a camera-projection unit that lets people use blank sheets of paper, foam or acrylic of different sizes and shapes as flexible displays. Flexpad has two main technical contributions: first, we contribute an algorithm for capturing even complex deformations in high detail and in real time. It does not require any instrumentation of the deformable handheld material. Hence, unlike in previous work that requires markers, visible texture, or embedded electronics [15, 22, 12, 35, 18, 19], virtually any sheet at hand can be used as a deformable projection surface for interactive applications. Second, we contribute a novel robust method for detecting hands and fingers with a Kinect camera using optical analysis of the surface material. This is crucial for robust captur-

ing of deformable surfaces in realistic conditions, where the user occludes significant parts of the surface while deforming it. Since the solution is inexpensive, requiring only standard hardware components, it can be envisioned that deformable handheld displays become widespread and common.

The highly flexible handheld display creates unique novel application possibilities. Prior work on *immobile* flexible displays has shown the effectiveness of highly detailed deformations, e.g. for the purposes of 2.5D modeling and multi-dimensional navigation [30, 7]. Yet, the existing inventory of interactions with flexible *handheld* displays [32, 23, 20, 18] is restricted to only low detail deformations, mostly due to the very limited flexibility of the displays and limited resolution of deformation capturing. Flexpad significantly extends the existing inventory by adding highly flexible and multi-dimensional deformations.

We present three application examples that each leverage rich and expressive, highly flexible deformations. The first application supports exploration and analysis of volumetric data sets (Fig. 1 a, b). We show how highly flexible displays support intuitive and effective exploration of the volumetric data set by allowing the user to easily define curved cross-sections, to get an overview, and to compare contents. The second application enables children to deform and animate 2D characters (Fig 1 c). A third application shows how detailed deformation can be used for slicing through time in videos (Fig. 1 d). These applications demonstrate the utility of Flexpad and introduce transferrable ideas for future handheld devices that use active flexible displays.

To evaluate the feasibility of our approach, both with respect to technology and human factors, we conducted two evaluation studies. They show that the tracking provides accurate results even for complex deformations and confirm that people can perform them fast and precisely.

The remainder of this paper is structured as follows. After reviewing related work, we present the concept and implementation of Flexpad. This is followed by an overview of applications. Finally, we present the results of two evaluation studies and conclude with an outlook on future work.

RELATED WORK

Deformation Capturing

Our work relates to a body of research that realizes handheld displays using optical capturing of the projection surface and spatially aligned projection, either with a static or a handheld projector, e.g. [3, 15, 24, 28, 35, 11, 18].

Many methods exist for capturing the pose and deformation of a mobile projection surface – information required for accurate texture mapping. Existing methods include passive or active markers, stereo cameras, structured light, time of flight, analysis of visible texture via features [36], color distances [6] or optical flow [14] as well as analysis of the object’s contour [8]. The commercial availability of inexpensive depth cameras has recently stimulated a lot of re-

search that produces a detailed geometrical model of a real-world scene for enabling novel forms of HCI. LightSpace [37] uses several depth cameras and projectors to augment walls and tables for touch and gesture-based interaction with projected interfaces. KinectFusion [16] automatically creates a very detailed model of a static scene, enabling physics-based interactions on and above surfaces and objects. Omnitouch [11] uses a depth sensor to support multi-touch input on projected interfaces on virtually any non-instrumented surface. However, none of these approaches model real-time deformation of objects. Very recent work [13, 5] uses proxy particles to model deformable objects, allowing for a range of physics-based interactions that are based on collision and friction forces. However, tracking particles over a sequence of frames requires optical flow, which is not possible with untextured surfaces. Other work [17] captures deformable surfaces from Kinect depth data, but the approach does not support capturing detailed deformations in real time, which is critical for interactive applications. We present an approach that is capable of capturing the pose and detailed deformation of a deformable surface from depth data in real-time, to support very fine-grained interactions and natural deformations. Due to a novel detection of hands and fingers, our approach is very robust to occlusions, which is crucial when users deform objects in natural settings.

Flexible Display Interfaces

Our research is also influenced by existing work on deformation-based handheld interfaces. Here, two streams of work can be distinguished:

First, prior work proposed using a deformable tape or sheet as an *external input device*, to control GUI applications by deformation. Output is given on a separate, inflexible display. An early and influential work is ShapeTape [2] that allows users to create and modify 3D models by bending a tape that is augmented by a deformation sensor. Other applications use a deformable sheet for controlling windows on the computer desktop [10] and for AR purposes [26].

A second type of research, more directly related to our work, studies *deformation of handheld displays*. Prior work has investigated interactions with interfaces of flexible smart phones and e-readers [32, 23, 19, 18, 1, 20] and for volumetric datasets [22]. Most of this work uses projection or simulates flexible displays with existing inflexible displays; PaperPhone [23] and Kinetic [20] use an active flexible display. Lee et al. [25] elicited user-defined gestures for flexible displays. A recent study [21] examined how device stiffness and the extent of deformation influenced deformation precision. While no pronounced influence was identified, users preferred softer rather than rigid materials.

This stream of work represents an important step towards fully flexible displays, yet existing work is limited by quite restricted deformability. It supports only slight bending and captures only simple deformations, which ultimately limits the expressiveness of deformation to mostly one-

dimensional interactions. We contribute novel interactions that extend the inventory of interactions by leveraging more flexible deformations.

FLEXPAD OVERVIEW

Flexpad enables users to interact with highly flexible projected displays in interactive real-time applications. The setup consists of a Kinect camera, a projector, and a sheet of paper, foam or acrylic that is used as projection surface.

Our current setup is illustrated in Fig. 2. A Kinect camera and a full HD projector (ViewSonic Pro 8200) are mounted to the ceiling above the user, next to each other, and are calibrated to define one coordinate space. The setup creates an interaction volume of approximately 110 x 55 x 43 cm within which users can freely move, rotate, and deform the projection surface.

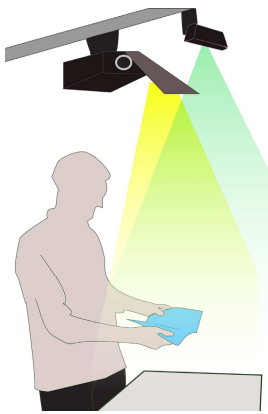


Fig. 2: Physical setup

The key contribution of Flexpad is in its approach for markerless capture of a high-resolution 3D model of a deformable surface, including its pose and deformation (see section about the model below). The system is further capable of grabbing 2D image contents from any window in Microsoft Windows and projecting it correctly warped onto the deformable display. Alternatively, it can use 3D volume data as input for the projection. The detailed model of deformation enables novel applications that make use of varied deformations of the display. We will present some example applications below.

Flexible display materials

Sheets of many different sizes and materials can be used as passive displays, including standard office paper. In our experiments, we used letter-sized sheets of two materials that have demonstrated good haptic characteristics:

The *fully flexible display* is 2 mm thick white foam (Darice Foamies crafting foam). It allows the user to change between different deformations very easily and quickly. The bending stiffness index S_b^* of this material was identified as $1.12 \text{ Nm}^7/\text{kg}^3$, which is comparable to the stiffness of $160\text{g}/\text{m}^2$ paper. ($S_b^* = S_b/w^3$ is the bending stiffness nor-

malized w.r.t. the sheet's basis weight w in kg/m^2 [27].) The display's weight is 12 grams.

The *shape-retaining display* retains its deformation even when the user is not holding it. It consists of 2 mm thick white foam with Amalog 1/16" armature wire attached to its backside. It allows for more detailed modeling of the deformation, for easy modifications, as well as for easy holding. To characterize its deformability, we measured the minimal force required to permanently deform the material. We affixed the sheet at one side; a torque as small as 0.1Nm is enough to permanently deform the material, which shows that the material is very easy to deform. The sheets' weight is 25g.

FLEXPAD IMPLEMENTATION

Existing image based tracking algorithms of deformable surfaces [36, 6, 14] rely on visible features. In this section, we present an approach that requires only image data from a Kinect depth sensor. This allows for tracking an entirely blank (or arbitrarily textured) deformable projection surface in high detail. Moreover, since the depth sensor operates in the infrared spectrum, it does not interfere with the visible projection, no matter what content is projected.

A challenge of depth data is that it does not provide reliable image cues on local movement, which renders traditional tracking methods useless. The solution to this problem is a more global, model-based approach: We introduce a parameterized, deformable object model that is fit into the depth image data by a semi-global search algorithm, accounting for occlusion by the user's hands and fingers.

In the following, we explain 1) the removal of hands and fingers from the input data, 2) the global deformation model and 3) how a global optimization method is used to fit the deformation model to the input data in real time.

Optical Surface Material Analysis for Hand and Finger Detection

While the user is interacting with the display, hands and fingers partially occlude the surface in the depth image. For a robust deformation tracking, it is important to classify these image pixels as not belonging to the deformable surface, so the model is not fit into wrong depth values. The proposed tracking method is able to handle missing data, so the solution lies in removing those occluded parts from the input images in a preprocessing step.

Due to the low resolution of the Kinect depth sensor it is difficult to perform a shape-based classification of fingers and hands at larger distances. In particular, when the finger or the flat hand is touching the surface, the resolution is insufficient to differentiate it reliably from the underlying surface. Existing approaches use heuristics that work reliably as long as the back of the hand is kept at some distance above the surface [11]. These restrictions do not hold in our more general case, where people use their hands and fingers right *on* the surface for deforming it.

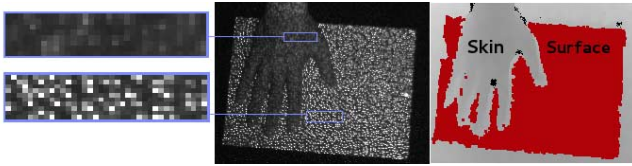


Figure 3: Detection of skin by analyzing the point pattern in the Kinect infrared image. Center: Input. Right: Classification

We introduce *optical surface material analysis* as a novel method of distinguishing between skin and the projection surface using the Kinect sensor. It is based on the observation that different materials have different reflectivity and translucency properties. Some materials, e.g. skin, have translucency properties leading to subsurface scattering of light. This property causes the point pattern, which is projected by the Kinect, to blur. This reduces the image gradients in the raw infrared image, which is given by the Kinect driver (see Fig. 3). The display surface, made out of paper, cardboard, foam or any other material with highly diffuse reflection, varies from skin in reflectivity as well as translucency; so low peak values or low gradient values provide a stable scheme to classify non-display areas.

The projector inside the Kinect has a constant intensity and the infrared camera has a constant shutter time and aperture. Hence, the brightness of the point pattern is constant for constant material properties and object distances and angles. As a consequence, the surface reflectivity of an object with diffuse material can be determined by looking at the local point brightness. However, because the brightness in the infrared image decreases with increasing distance, the local gradients and peak values have to be regarded with respect to the depth value of the corresponding pixel in the depth image. Pixels that cannot be associated to a dot of the pattern are not classified directly, but receive the same classification as the surrounding pixels by iterative filtering. Results from an evaluation, reported below, show that this classification scheme works reliably at distances up to 1.50 m from the camera.

Modeling the Deformable Surface

Without local movement cues in the image, such as feature movement or optical flow, and with incomplete depth data, it is important to formulate very precisely what the tracking method should find in the image, i.e. to define a deformation model. We present a model that is generic enough to approximate a wide variety of deformations, including complex ones, while still being efficient to calculate.

The model is based on a rectangular 25×25 vertex plane at the size of the actual display surface. Using 25 vertices per dimension proved to provide stable tracking results while keeping the computational costs small. We then define a set of 8 basic deformations, 4 diagonal and 4 axis-aligned bends, each deforming half the surface plane (Fig. 4 upper left). It is important to note that all of these basic deformations are *combined* by using weighted factors, such that the model supports complex curvatures.

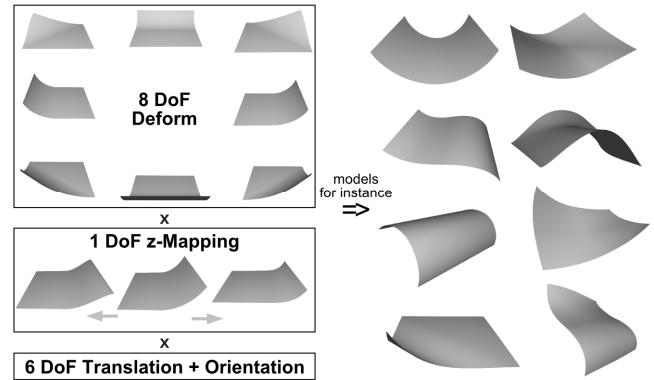


Figure 4: Dimensions of the deformation model (left) and examples of deformations it can express (right).

A drawback of this model is that it lacks the ability to shift the starting point of a deformation, e.g. for distinguishing between a smaller and a larger dog-ear. To add this ability to the deformation model while keeping it computationally efficient and robust through low dimensionality, an additional mapping function for the z components of the vertex position is applied after the deformed vertex positions are calculated. The mapping can change from a square-like function, via the identity function to a square root-like function. This parameterized function is applied to every z -component of each surface vertex to change the shape of the deformation. This largely increases the set of deformations in the model by only adding one additional deformation parameter.

In summary, as shown in Fig. 4, the deformation model is a 15 dimensional vector, containing the angles of the 8 basic deformations, the z mapping parameter as well as 6 variables for the degrees of freedom that are required for the affine 3D transformations. Figure 4 (right) gives some examples of the complex deformations the model supports.

Real-time Calculation

Given this deformation model, the overall tracking goal can be defined as finding the parameters of the model that describe an object that, when synthesized as a depth image, matches the input depth image best. Such approaches are known as Analysis by Synthesis (AbS) or direct tracking methods [17]. The advantage of such methods is the ability to work completely without information on feature movement or optical flow. Such evaluation allows for formulating the tracking as an optimization problem, turning the tracking task into a mathematically well-defined objective of finding the parameter vector that produces the most suitable deformation.

To determine the deformation parameters, the AbS approach synthesizes the depth measurements of the 25×25 pixels to which the model vertices would be projected and compares the vertex-camera distance to the real input depth image. Thereby, areas that were identified as skin by the occlusion handling step are ignored. The average difference over all vertices provides an error value that examines the

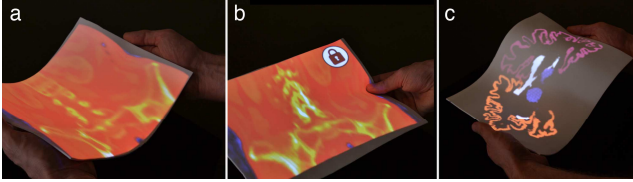


Figure 5: Exploring curved cross-sections (a), flattening the view (b), comparing contents across layers (c)

quality of the parameter guess to model the deformed object in the image. Due to the uneven measurements and the noise in the Kinect image, this error function contains many local minima, making it intractable for common local least squares optimization. Hence, we apply the CMA-ES optimization scheme [29] to find an approximation of the optimum since CMA-ES combines two important properties: it is a global optimization algorithm, which allows for coping with large numbers of local minima, and it applies a smart distribution scheme minimizing the necessary number of function evaluations to find the optimum.

The optimization process is initialized at the first frame with a vertex set describing a planar surface of the display size and shape, placed at the center in the depth image. In the beginning of each following frame, the optimization is initialized with the optimization result of the preceding frame. This allows for initially finding the sheet not only at this position, but also at different locations and orientation within the instrumented volume. Should the depth error of the optimization result rise above a given limit, it is most likely that the tracking failed, e.g. because the user has removed the deformable sheet from the camera’s viewport. In this case, the initialization step is automatically executed again, until the depth error falls below this limit.

To further increase the performance, our approach leverages the fact that the user cannot make high precision deformations while moving the object quickly, but only when the object is held steadily or slowly moved. The number of iterative optimization steps per frame is automatically adjusted by the pose difference of the preceding frames. During large movements, the number of iterations is reduced (on our hardware, 50 iterations yield 25 fps). During very slow movements, the number is increased, smoothly reducing the frame rate (200 iterations yield 8 fps), allowing for the highest possible accuracy of deformation capturing.

Limitations

Our approach allows for robust real-time tracking of a large variety of deformations with a variety of deformable surfaces. However, due to the real-time constraint and restrictions of the depth sensor, several limitations apply. Obviously, it is only possible to track deformations that can be observed directly by the Kinect camera. Folding, which occludes large parts of an object, as well as very steep bending angles that occlude parts of the object to the camera cannot be handled by the presented system. For the cost of installing and calibrating additional Kinect sensors and

projectors, this limitation can be overcome. Moreover, the presented deformation model makes a trade-off between tracking stability and the set of detectable deformations. Hence, very sharp bends, in particular folds, and stretchable materials cannot be reliably captured.

APPLICATION EXAMPLES

Flexpad supports many application and interaction possibilities. Its unique strength is the high flexibility of the display and the detailed capturing of its deformation. A direct 1:1 mapping between the deformed physical shape and virtual model space enables novel applications with expressive interactions that were not possible with rigid or slightly flexible handheld displays. It is worth emphasizing that all proposed interactions also transfer to future active displays.

Exploring and Analyzing Volumetric Datasets

Analysis of volumetric images and datasets is important in many fields, such as medicine (CT and MRI scans), geology and earth sciences (atmospheric and submarine datasets). Prior research has shown that curved cross-sections are required for analyzing important medical phenomena [31]. We contribute interactions that significantly go beyond existing work on visualization of volumetric datasets, which support only planar cross-cuts [35], only very restricted curvatures [22] or make curvature very hard to control [31].

Our application maps the volumetric dataset to a 3D volume in physical space and lets the user slice through that volume with the display. By deforming the highly flexible display, the user can easily create *non-planar cross-sections* to analyze a large variety of curved structures that would not be visible on planar or only slightly deformed cross-sections. Figures 1 and 5 illustrate several examples: the spine (Fig. 1a) and the pelvis (Fig. 5a). For detailed analysis of a cross-section, the user can *lock the view* by pressing a foot button. When locked, the display can be moved and deformed without affecting the visualization (Fig. 5b). For instance, this allows for *flattening the view* and easy measurement of distances that were on a curve in the original view. It also allows for handing the view over to a colleague. It is worth noting that this application is not only helpful for medical experts, but also supports non-experts who are interested in exploring the inner workings of the human body in an intuitive way. In addition to following a layer, deformation also proves powerful in supporting *curved cuts across layers*. Consider the example of a reservoir engineer exploring the best locations for a new well in an oil field. The engineer can select the desired curvature of the well (nowadays many wells are curved) by bending the display and then skim through the entire volume.

Deformation also supports *better orientation and overview* in the dataset. For instance, one part of the display remains at a visually salient feature while the user curves the other part to explore the surrounding area in the dataset (Fig. 1b). Furthermore, deformation of the highly flexible display provides an intuitive way of *comparing data across layers*. By deforming the display into a wave bend, the user sepa-

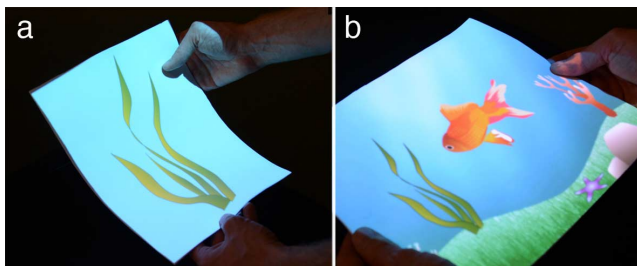


Figure 6: Animating virtual paper characters

rates the display into two slices that are located at different depth levels (Fig 5c). By varying the bend, the user can smoothly select the distance between both slices.

Our proposed interactions can be easily integrated with functionality for cross-cuts of volumetric datasets that have been presented in prior work, such as additional views on external displays [35], panning and zooming, or saving of 2D snapshots for subsequent analysis and presentation.

Animating Virtual Paper Characters

To demonstrate the wide applicability of highly deformable handheld displays, we next present an application for children that leverages deformation as a simple and intuitive means for animating paper characters in an interactive picture. By deforming the display and moving it in space (Fig. 1c and 6a), the creature becomes animated. High-resolution deformation allows very individualized and varied animation patterns. Once animated, the creature is released into the virtual world of the animated picture (Fig. 6 b). For example, fish may move with different speeds and move their fins. A sea star may lift some of its tentacles. A sea worm may creep on the ground. A flatfish may meander with sophisticated wave-form movements and seaweed may slowly bend in the water.

Such rich deformation capabilities can be easily combined with concepts from previous research on animating paper cut-outs [4] that address skeleton animation, eye movement, and scene lighting. Also 3D models could get animated by incorporating as-rigid-as-possible deformation [33]. It is worth noting that our concept can be readily deployed with standard hardware.

Slicing through Time in Videos

A third application of Flexpad enables people to create slices through time in videos, inspired by the ingenious Khronos projector [7]. In contrast to Khronos, which required a very bulky, several cubic meters big, immobile setup with a fixed screen, Flexpad brings similar functionality to virtually any sheet of paper. The user can load a video from YouTube, which is automatically mapped to a virtual volume, whereby time is mapped to the z dimension. By moving and deforming the display within this volume, the user creates ever new combinations of different moments in time that open up new artistic perspectives on the video. Figure 7 depicts an example, in which a tree is “moved” in

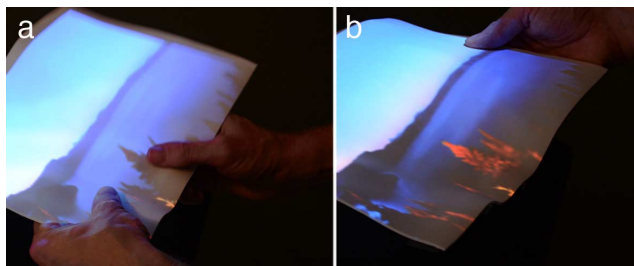


Figure 7: Slicing through time in a video by deformation

time to glow in the sunset. However, this functionality is best illustrated in the video that accompanies this paper.

In contrast to the fixed display of Khronos, the flexible handheld display allows for defining a curvature, i.e. a time gradient, and then move this gradient to different areas within one frame or across frames. Moreover, the handheld display allows for selecting any frame of the video to start with deforming, whereas deformations in Khronos always start at the topmost frame, going in only one direction. This allows for novel, expressive interactions with videos.

EVALUATION

To evaluate the feasibility of the Flexpad approach, we conducted two evaluations, examining both system performance and the users’ ability of precise interaction with highly deformable displays.

Tracking Performance

To evaluate the precision of tracking in a natural real-time scenario, we recruited 10 volunteer participants (5f, 5m, median age 26). Their task was to use the slicing-through-time application and freely create interesting renderings by deforming the display. This application is particularly well suited for a technical evaluation of the system, for the unstructured nature of the interface stimulates users to deform the display in highly varied, complex ways. We recorded the raw video stream from the Kinect camera while the user performed the task, overall 35 minutes of footage. We used this data as an input for our algorithm, calculating the RMS error in each frame from the distance between the model surface to the corresponding depth image values. The average RMS error over all 20,217 frames is 6.47 mm (SD 3.33 mm). This shows that the tracking performs very adequately even in challenging realistic tasks.

To test the abilities of the tracking method in a more distinct way, we additionally picked a set of 8 deformations, ranging from simple bends to complex deformations. Each deformation was recorded 20 times at 90 and 150 cm distance to the Kinect camera. Figure 8 shows the deformations and the average RMS error for each deformation. In addition, we evaluated the preprocessing step, which removes hands and fingers from the input data, in an informal study. The Kinect data of hands from 10 users (aged 25 to 54) occluding three different display materials (paper, card board, plastic) was recorded. Analysis showed that

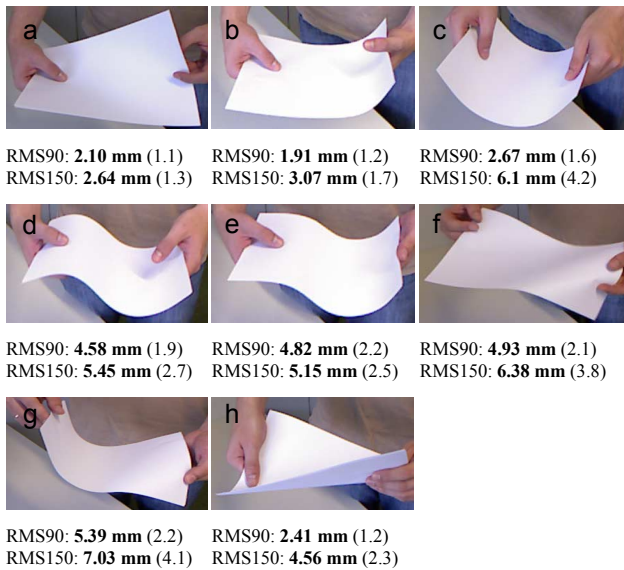


Figure 8: Average RMS error of a set of representative deformations (at 90 cm and 150 cm distance between Kinect and the object; the standard deviation is given in brackets).

independently of skin color, the preprocessing step correctly removed over 99% of the skin pixels without the need to adjust the thresholds for reflection or translucency. Figure 9 shows data from six users and the classification results.

User Performance of Deformation

For the feasibility of the Flexpad approach, it is essential that users are able to perform deformations with the display with ease and sufficiently high precision. To evaluate these human factors, we conducted a controlled experiment with users. Our aim was to evaluate how fast and precise users are able to perform deformations and to examine the influence of flexible vs. shape-retaining display materials. Our hypothesis was that the task completion time increases with the complexity of the deformation and the level of precision required. Moreover, we hypothesized that the shape-retaining material would increase task completion times. We aimed at quantifying this influence. As such, our study adds to an emerging body of experiments on the manipulation of paper-like displays, including target acquisition performance in 3D space with rigid paper displays [34] and performance of deforming slightly flexible displays [21].

We recruited 10 volunteer participants (5f, 5m, all right-handed, median age 26). The two participants who performed the task fastest were offered a \$10 gift card. The experiment consisted of a series of trials that required the participant to deform the display surface as quickly as possible to match a given target shape within a specific precision level (see

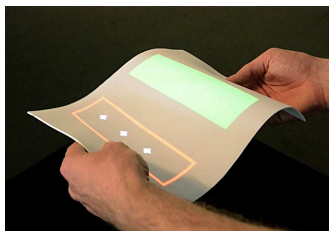


Figure 10: Interactive visualization during a trial

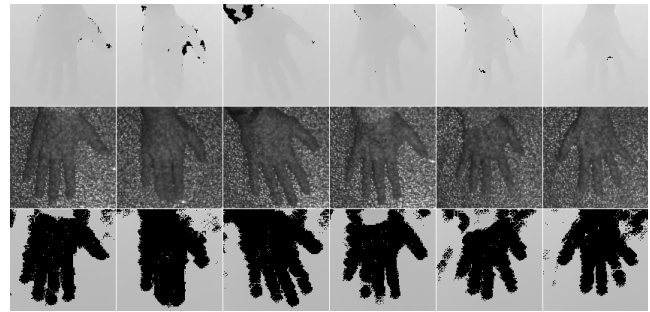


Figure 9: Classification of skin. Top: depth input. Center: infrared input. Bottom: depth image after classification with skin parts removed.

below). A perspective rendering of the target shape was displayed before each trial. During the trial, a real-time visualization on the display guided the participant on how to deform it to match the shape; to match the shape, all red areas had to turn green (Fig. 10). The trial was solved when the target shape was held during 250ms. The completion time for each trial was measured. After each trial, the participant had to rate the perceived difficulty on a five-point Likert scale.

We selected five deformation classes: two one-sided deformations, which are oriented with the corner (Fig. 8a) and oriented with the edge (Fig. 8b), and three two-sided ones: center (Fig. 8c), symmetric wave (Fig. 8d) and asymmetric wave (Fig. 8e). Each of these classes contains not only the basic deformation, but all variations of it, on all four sides of the display, and each also flipped vertically. This set of deformations was informed by the deformations made in our application examples, and for reasons of feasibility of this study, focused on deformations along one dimension in landscape format. From each deformation class, we randomly selected 2 deformations for each participant.

Each deformation had to be performed with two precision levels: ± 8 degrees and ± 6 degrees. In a pre-study with 5 participants, we identified these precision levels as challenging, but still feasible, whereas a level of ± 4 degrees turned out not to be reliably reachable. The sequence of trials was randomized. To compare influences of the display material, participants performed the tasks with the fully flexible and with the shape-retaining display, as described in the system overview section. The order was counterbalanced. In summary, the within-subject, multi-factorial experimental design was as follows:

5 deformation classes x 2 instances per class x 2 precision levels x 2 materials x 10 participants = 400 trials

Before the experiment, participants could practice the tasks with both materials until they felt fully confident. After the experiment, participants freely explored a medical dataset in the volumetric application as well as the slicing-through-time application. This was followed by a semi-structured interview. Each session lasted approximately 1 hour.

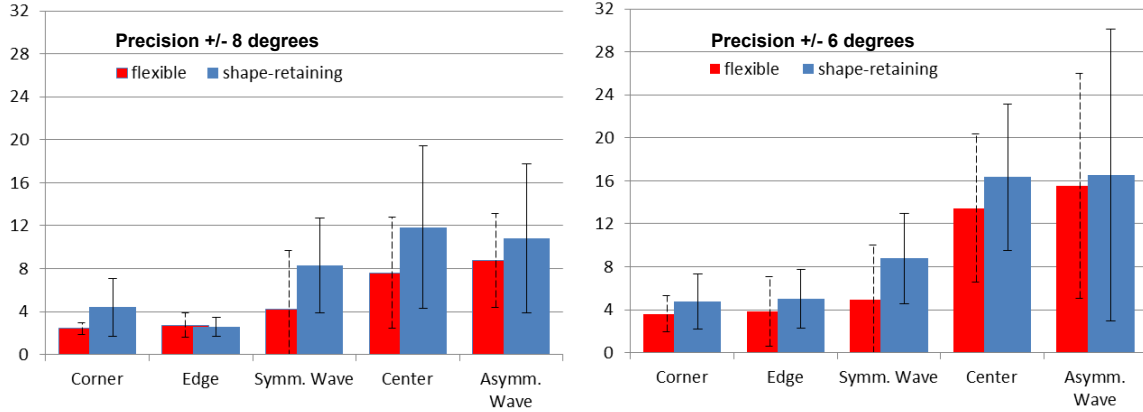


Figure 11: Average trial completion time in seconds. Error bars show the standard deviation.

Results

Figure 11 (left) depicts the average trial completion times for the five basic deformations and both materials for the lower precision level. With the fully flexible material, times range between 2.4 and 8.8 seconds; with the shape-retaining material, values range from 2.6 to 10.8 seconds. Figure 11 (right) depicts those values for the higher precision level. In this case, times range between 3.6 and 15.6 as well as 4.8 and 16.5 seconds, for the two materials, respectively. The rather high standard deviations reflect different levels of manual dexterity of users.

We first performed a multifactorial repeated measures ANOVA with all factors, including rotation and vertical flip of the deformations. We found no significant effect of rotation and vertical flip; therefore, in the following analyses, we group the trials into the five basic deformations, regardless of their rotation and vertical flip.

To identify systematic dependencies, we performed a r-ANOVA with the factors deformation, precision level, and material. Significant main effects were found for deformation ($F_{(4,36)} = 16.13$; $p < 0.001$), precision level ($F_{(1,9)} = 13.17$; $p = 0.005$) and material ($F_{(1,9)} = 8.86$; $p = 0.016$). As expected, more complex deformations, a higher precision level, and the shape-retaining material resulted in higher task performance time. On average, completion time of trials that were performed with the shape-retaining material was 33% longer than of trials with the fully flexible material. When the higher precision level was required, completion time of trials was 46% longer than of trials performed at the lower precision level.

To further analyze differences between the five deformation classes, we conducted post-hoc tests with Bonferroni-adjusted alpha-levels of .01. Results showed highly significant differences between: edge and center (mean difference = 8.77; $p = .004$), edge and asymmetric wave (MD = 9.36; $p = .006$); corner and center (MD = 8.53; $p = .001$), as well as corner and asymmetric wave (MD = 9.12; $p = .004$). This shows a distinctive difference between deformations requiring bending only one versus both sides. The symmetric

wave deformation does not differ statistically from both groups due to a highly material-dependent behavior, which is visible in Fig. 11.

Simple (edge or corner) deformations had an average task completion time of 3.67 seconds (SD = 1.72). Even the most difficult deformations, in high precision level, resulted in an average task performance time of 15.6 seconds (SD = 9.54). The average rating of all trials was 2.06 (SD = 0.88). Analyzed separately for all combinations of deformation, precision level and material, the average ratings range from 1.2 to 3.1, i.e. from very easy to medium. Perceived difficulty and completion times of the trials showed a very high bivariate correlation (Pearson's $r = .75$, $p < .001$). Therefore, our following analyses are based on the completion times, but transfer to the difficulty of the task.

We identified one exception to the general finding that the deformations were performed without difficulties: the asymmetric wave deformation in the higher precision level, performed with the shape-retaining material. While average time is in line with those of the other complex deformation, the standard deviation is much higher, as can be seen in Fig. 11 (rightmost column). Statistical testing using Grubbs' Test for outliers showed that this SD is indeed an outlier ($G = 2.68$, $p = .005$.) This high SD shows that some participants were able to perform this deformation rather quickly, whereas other participants could only perform it reliably within a disproportionately long time.

As presented above, and in accordance with our hypothesis, the ANOVA found a significant effect of material. Because of the advantages of the shape-retaining material for complex deformations that were stated above, we will now analyze this effect in more detail for double-sided deformations (symmetric wave, center, and asymmetric wave).

The penalty on completion time in the lower precision level that is due to the shape-retaining material averages out at 95% with symmetric wave, 56% with center and 24% with asymmetric wave. For the higher precision level, the penalties introduced by the shape-retaining material are 77%,

21%, and 6%, for the three deformations respectively. Most notably, the penalty introduced by the shape-retaining *decreases* with increasing difficulty of the deformation and increasing level of precision. This relationship is illustrated by significant medium-sized negative correlations between the difficulty of the deformation (symmetric wave, center, asymmetric wave) and the penalty on completion time, separated for precision level (lower precision level: $r = -.38$, $p = .018$; higher precision level: $r = -.39$, $p = .016$). A similar result was obtained when difficulty of the deformation and level of precision combined were correlated with the penalty on completion time ($r = -.39$, $p = .001$).

DISCUSSION AND CONCLUSIONS

Summary of study findings

The results of the controlled experiment show that users are capable of using highly flexible displays with ease. Users can make various single and dual deformations with both materials quickly and easily, even with a high precision level of +/- 6 degrees. The single exception is the asymmetric dual deformation, which, under high precision level, some participants had difficulties performing reliably with the shape-retaining material. Therefore, we suggest for this deformation a lower bound of +/- 8 degrees as the maximum precision if time is an issue in the application.

The results further showed that use of a shape-retaining material has only a modest average influence on task performance. While the fully flexible material is better suited for applications with only simple deformations, the shape-retaining material is well suited for use in applications with complex deformations. Notably, the time penalty introduced by the shape-retaining material decreases with increasing complexity of the deformation and the level of precision required. For the most complex deformation, it induces a temporal penalty of only 6 %, while it allows for easily holding deformations over time, for easily modifying them, and for modeling curvatures that involve more than two deformations. These quantitative findings are underpinned by qualitative observations. In the interviews, many participants commented that they preferred the fully flexible material for the task of the controlled experiment, whereas they preferred the shape-retaining material for analyzing curved phenomena in the medical dataset. For instance, P3 commented: *“The flexible one is easier to do the first assignment. But this one [the shape-retaining material] is easier to keep. It is more reliable. I can hold it even with just one hand. And I know that I gonna have what I want. It’s way easier.”* P4 stated: *“The less flexible material allows me to freeze a structure. (...) I don’t have to bother with holding the shape.”* P7 commented: *“I prefer the sensation of the flexible material, but the other one gives me more precision.”* In future work, we plan to perform a qualitative analysis of our video recordings to analyze which deformations participants spontaneously made with both materials.

Results from the technical evaluation show that our tracking approach performs with high accuracy to support all of the proposed applications. Even in a technically very challenging task that encouraged participants to make very complex arbitrary deformations, the average error was below 7 mm. These values can be further decreased in a mobile setup, where the Kinect camera is closer to the projection surface. Then, the average RMS error for many of the deformations is close to the random noise of the Kinect sensor.

Touch input on deformable displays

A logical extension of Flexpad is touch input. With respect to this question, future work should address several challenges. Technically, this involves developing solutions for Kinect sensors to detect touch input reliably on real-time deformable surfaces. On a conceptual level and building on Dijkstra et al.’s recent work [9], this also involves an understanding of where users hold the display and where they deform it. This is required to identify areas that are reachable for touch input and to inform techniques that differentiate between touches stemming from desired touch input and false positives that result from the user touching the display while deforming it. An informal analysis from our study data showed that users made single deformations almost exclusively by touching the display close to the edges and on its backside. This also held true for dual deformations with the fully flexible display material. This suggests that touches in the center area of the display can be reliably interpreted as desired touch input. In contrast, the participants touched all over the display for dual deformations with the shape-retaining material. A simple spatial differentiation is not sufficient in this case; more advanced techniques need to be developed, for instance based on the shape of the touch point or on the normal force involved.

Active flexible displays

While this work focuses on projected displays, all of our application examples transfer to future active flexible displays. Currently available prototypes are still very limited in their flexibility, so that they cannot be used to realize our concepts. Given the rapid advancements in display technology, this is very likely to change in the future.

Smart materials: programmable stiffness and stretchability

Future work should investigate smart materials for flexible displays. As discussed above, both fully flexible and shape-retaining displays have unique capabilities. A material that can programmatically switch between both of these states would combine all these advantages. Moreover, future work should examine handheld displays that, in addition to being deformable, are stretchable. This will further increase the expressiveness of interactions with flexible displays.

ACKNOWLEDGMENTS

We gratefully acknowledge Reinhard Koch, James D. Hollan and Simon Olberding for helpful discussions and feedback on earlier versions of this paper. We thank Clemens Winkler and Cassandra Xia for help with the video. The time lapse video used in one application was provided by

courtesy of Dan Heller. This work has partially been funded by the Cluster of Excellence Multimodal Computing and Interaction within the German Federal Excellence Initiative.

REFERENCES

1. Alexander, J., Lucero, A. and Subramanian, S. Tilt displays: designing display surfaces with multi-axis tilting and actuation. In *Proc. MobileHCI'12*, ACM Press, 2012.
2. Balakrishnan, R., Fitzmaurice, G., Kurtenbach, G. and Singh K. Exploring interactive curve and surface manipulation using a bend and twist sensitive input strip. In *Proc. I3D '99*, 1999.
3. Bandyopadhyay, D., Raskar, R. and Fuchs, H. Dynamic shader lamps: Painting on movable objects. In *Intl. Symp. Augmented Reality*, 2001.
4. Barnes C., Jacobs D., Sanders, J. and Goldman B. D., Rusinkiewicz, S., Finkelstein A. and Agrawala, M. Video puppetry: a performative interface for cutout animation. In *Proc. SIGGRAPH Asia '08*, ACM Press, 2008.
5. Benko, H, Jota R., and Wilson, A. MirageTable: freehand interaction on a projected augmented reality tabletop. In *Proc. CHI '12*, ACM Press, 2012.
6. Cagniard, C., Boyer E., and Ilic, S. Probabilistic deformable surface tracking from multiple videos. In *Proc. ECCV'10*.
7. Cassinelli, A and Ishikawa, M. Khronos projector. In *Proc. SIGGRAPH'05*, ACM Press, 2005.
8. Chen, Y., Rui, Y., and Huang, T.S. Multicue HMM-UKF for real-time contour tracking. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2006.
9. Dijkstra, R., Perez, C., and Vertegaal, R. Evaluating effects of structural holds on pointing and dragging performance with flexible displays. In *Proc. CHI '11*, ACM Press, 2011.
10. Gallant, T. D., Seniuk, G.A., and Vertegaal, R. Towards more paper-like input: flexible input devices for foldable interaction styles. In *Proc. UIST '08*, ACM Press, 2008.
11. Harrison, C., Benko, H., and Wilson, A.D. OmniTouch: wearable multitouch interaction everywhere. In *Proc. UIST '11*, ACM Press, 2011.
12. Herkenrath, G., Karrer, T., and Borchers, J. Twend: twisting and bending as new interaction gesture in mobile devices. In *Proc. CHI '08 Extended Abstracts*, ACM Press, 2008.
13. Hilliges, O., Kim, D., Izadi, S., Weiss, M., and Wilson, A. HoloDesk: direct 3d interactions with a situated see-through display. In *Proc. CHI '12*, ACM Press, 2012.
14. Hilsman, A., Schneider, D. C., and Eisert, P. Realistic cloth augmentation in single view video under occlusions. *Comput. Graph.*, 2010.
15. Holman, D., Vertegaal, R., Altosaar, M., Troje, N., and Johns D. Paper windows: interaction techniques for digital paper. In *Proc. CHI '05*, ACM Press, 2005.
16. Izadi, S., Kim, D., Hilliges, O., Molyneaux, D., Newcombe, R., Kohli, P., Shotton, J., Hodges, S., Freeman, D., Davison, A., and Fitzgibbon, A. KinectFusion: real-time 3d reconstruction and interaction using a moving depth camera. In *Proc. UIST '11*, ACM Press, 2011.
17. Jordt, A and Koch, R. Direct model-based tracking of 3d object deformations in depth and color video. *Intl. J. Computer Vision*, 2012.
18. Khalilbeigi, M., Lissermann, R., Kleine, W., and Steimle J. FoldMe: interacting with double-sided foldable displays. In *Proc. TEI '12*, ACM Press, 2012.
19. Khalilbeigi, M., Lissermann, R., Mühlhäuser, M., and Steimle, J. Xpaaand: Interaction techniques for rollable displays. In *Proc. CHI '11*, ACM Press, 2011.
20. Kildal, J., Paasovaara, S., and Aaltonen, V. Kinetic device: designing interactions with a deformable mobile interface. In *Proc. CHI 2012 Extended Abstracts*, ACM Press, 2012.
21. Kildal, J. and Wilson, G. Feeling it: the roles of stiffness, deformation range and feedback in the control of deformable UI. In *Proc. ICMI'12*, 2012
22. Konieczny, J., Shimizu, C., Gary Meyer, and D'ardo Colucci. A handheld flexible display system. In *Proc. IEEE Visualization 2005*, IEEE Press, 2005.
23. Lahey, B., Girouard, A., Burlison, W., and Vertegaal, R. PaperPhone: understanding the use of bend gestures in mobile devices with flexible electronic paper displays. In *Proc. CHI '11*, ACM Press, 2011.
24. Lee J. C., Hudson, S. E., and Tse, E. Foldable interactive displays. In *Proc. UIST '08*, ACM Press, 2008.
25. Lee, S. S., Kim, B., Jin, B., Choi, E., Kim, B., Jia, X., Kim, D., and Lee, K. P. How users manipulate deformable displays as input devices. In *Proc. CHI '10*, ACM Press, 2010.
26. Looser, J., Grasset, R., and Billinghurst, M. A 3D flexible and tangible magic lens in augmented reality. In *Proc. ISMAR '07*
27. Mark, R., Borch, J., Habeger, C., and Lyne, B. *Handbook of Physical Testing of Paper Vol. 1*. CRC Press, 2001.
28. Mistry, P., Maes, P., and Chang, L. WUW - wear ur world: a wearable gestural interface. In *Proc. CHI '09 EA*, 2009.
29. Ostermeier, A. and Hansen, N. An evolution strategy with coordinate system invariant adaptation of arbitrary normal mutation distributions within the concept of mutative strategy parameter control. In *Proc. Genetic and Evolutionary Comp. Conf.*, 1999.
30. Piper, B., Ratti, C., and Ishii, H. Illuminating clay: a 3-d tangible interface for landscape analysis. In *Proc. CHI'02*, 2002.
31. Saroul, L., Gerlach, S., and Hersch, R. D. Exploring curved anatomic structures with surface sections. In *Proc. IEEE Visualization*, 2003.
32. Schwesig, C., Poupyrev, I., and Mori, E. Gummi: a bendable computer. In *Proc. CHI '04*, ACM Press, 2004.
33. Sorkine, O. and Alexa, M. As-rigid-as-possible surface modeling. In *Proc. Eurographics*, 2007.
34. Spindler, M., Martsch, M., and Dachselt, R. Going beyond the surface: studying multi-layer interaction above the tabletop. In *Proc. CHI'12*, ACM Press, 2012.
35. Spindler, M., Stellmach, S., and Dachselt, R. PaperLens: advanced magic lens interaction above the tabletop. In *Proc. ITS'09*, ACM Press, 2009.
36. Varol, A., Salzmann, M., Tola, E., and Fua, P. Template-free monocular reconstruction of deformable surfaces. In *Proc. Computer Vision*, 2009.
37. Wilson, A. D. and Benko, H. Combining multiple depth cameras and projectors for interactions on, above and between surfaces. In *Proc. UIST '10*, ACM Press, 2010.