Direct and gestural interaction with relief: A 2.5D shape display

The MIT Faculty has made this article openly available. Please share how this access benefits you. Your story matters.
Direct and Gestural Interaction with Relief: A 2.5D Shape Display

Daniel Leithinger, David Lakatos, Anthony DeVincenzi, Matthew Blackshaw, Hiroshi Ishii
MIT Media Lab
75 Amherst Street, Cambridge, MA 02139, USA
{daniell, dlakatos, tonyd, mab, ishii}@media.mit.edu

ABSTRACT
Actuated shape output provides novel opportunities for experiencing, creating and manipulating 3D content in the physical world. While various shape displays have been proposed, a common approach utilizes an array of linear actuators to form 2.5D surfaces. Through identifying a set of common interactions for viewing and manipulating content on shape displays, we argue why input modalities beyond direct touch are required. The combination of free-hand gestures and direct touch provides additional degrees of freedom and resolves input ambiguities, while keeping the locus of interaction on the shape output. To demonstrate the proposed combination of input modalities and explore applications for 2.5D shape displays, two example scenarios are implemented on a prototype system.

ACM Classification: H5.2 [Information interfaces and presentation]: User Interfaces: Input devices and strategies; Haptic I/O; I.3.6 [Methodology and techniques]: Interaction techniques.

General terms: Design, Human Factors

Keywords: Gestural Input, Tangible Input, Direct Manipulation, Actuated Surface, Shape Display

INTRODUCTION
Shape displays generate dynamic physical forms controlled by a computer. While today’s shape displays are unable to render arbitrary objects in space, current implementations succeed in generating approximations of an object’s shape. As computer interfaces, they inherit the benefits of Tangible User Interfaces (TUI) [11], allowing multiple users to simultaneously view, feel and manipulate a physical shape instead of an abstract graphical representation. This enables novel, currently unexplored uses in areas such as computer-aided design (CAD). Review of related work, reveals that most shape output mechanisms are comprised of a 2D array of linear actuators. This arrangement forms a 2.5D approximation of an object’s shape, similar to a relief sculpture (Fig. 2). The term 2.5D in this context follows a common definition used for digital elevation models [26]. Current 2.5D shape displays allow user input by touching and pressing into the interface surface. We argue that additional input modalities enable 2.5D shape displays to perform a wider range of applications: by extending touching and pressing with freehand gestural input, we keep the locus of interaction on the shape output. We define a set of gestures for selecting, translating, and rotating the shape of an object. Through our implementation we demonstrate the use of these gestures in two applications. The first application explores the rendering of arbitrary 3D models augmented with graphics. In the second application, we investigate the mixed input capabilities of direct touch and freehand ges-

Figure 1: We propose to extend direct touch interaction with 2.5D shape displays by a set of free-hand gestures

Figure 2: Relief sculpture (image courtesy of Chez Cäsver)
tured for CAD.

RELATED WORK

Shape Displays: Actuated Tangible Interfaces
Shape displays are haptic interfaces with the ability to create and deform physical shapes in real time. The idea of computationally controlled matter to create an immersive visual and haptic sensation was first described by Sutherland in 1965 as the ultimate display [24]. Most shape displays are not only output devices, but also allow user input. The approach of interacting with computers through physical embodiments of information is similar to the concept of Tangible User Interfaces (TUI) introduced by Ishii et al. [10]. Poupyrev et al. identify shape displays as an emerging area of research [23].

Future shape displays might be able to generate arbitrary shapes through self-arranging independent modules [6]. However, current approaches consist of fixed arrangements of actuators forming a surface. These actuators can be embedded into the surface, such as the Nitinol actuators of Surflet [3] or the hinge elements of the Formable Crust project [22]. Other arrangements include volumetric actuators attached to a column, as proposed by VoFlex [13]. However, the majority of shape displays utilize a 2D array of linear actuators, which form 2.5D shapes similar to relief sculptures.

2.5D Shape Displays
While 2.5D shape displays vary in size, type of actuation, and speed of shape output they all follow the same principle of shape generation. An array of actuated pins forms a surface, similar to pin screen toys [25]. This layout of actuators allows simple mechanical design, since the actuators can be hidden under the surface, compared to the complexity of embedding them into the surface.

Limitations
2.5D shape displays are limited in the types of shapes they can generate. Most fundamentally they do not allow overhangs (Fig. 4). Another limitation is the resolution of the actuated points. Due to the size and complexity of the actuators, they cannot be packed densely enough to create a resolution capable of outputting a perceived continuous shape [23, 16]. Perceived resolution can be increased by stretching a malleable surface over the pin array producing the illusion of continuous shape.

Actuators have a limited linear range. If the shape formed by the interface exceeds this range, it will be clipped. The generated shape may be augmented with visuals: either through LED’s embedded in the pins or top-down projection on the surface.

History
The “surface display”, created by Hirota and Hirose in 1993 [10] consists of a 4 x 4 linear actuator array. The actuators form a physical surface from the depth buffer of arbitrary 3D geometry.

Iwata et al. developed FEELEX in 1995 to overcome shortcomings they identified in their previously constructed haptic interfaces tethered to the human body [12]. FEELEX consists of a malleable surface deformed by an array of 6 x 6 linear actuators. A top-down projector is used to create visual feedback on the surface. Through embedded pressure sensors the system reacts to the user’s push. An example application renders a moving creature reacting to touch

<table>
<thead>
<tr>
<th>Name</th>
<th>Actuation method</th>
<th>Number of actuators</th>
<th>Display size</th>
<th>Actuator travel</th>
</tr>
</thead>
<tbody>
<tr>
<td>FEELEX</td>
<td>Motor driven screws</td>
<td>6 x 6</td>
<td>240 x 240 mm</td>
<td>8 cm</td>
</tr>
<tr>
<td>FEELEX 2</td>
<td>Piston-crank mechanism</td>
<td>23</td>
<td>50 x 50 mm</td>
<td>1.8 cm</td>
</tr>
<tr>
<td>Lumen</td>
<td>Nitinol actuators</td>
<td>13 x 13</td>
<td>84 x 84 mm</td>
<td>6 mm</td>
</tr>
<tr>
<td>Digital Clay</td>
<td>Hydraulic</td>
<td>5 x 5</td>
<td>25 x 25 mm</td>
<td>48 mm</td>
</tr>
<tr>
<td>Surface display</td>
<td>Slider-crank mechanism</td>
<td>4 x 4</td>
<td>120 x 120 mm</td>
<td>5 cm</td>
</tr>
<tr>
<td>Xenovision Mark III</td>
<td>Electric</td>
<td>7000</td>
<td>91 x 122 cm</td>
<td>15 cm</td>
</tr>
<tr>
<td>Relief</td>
<td>Belt actuation</td>
<td>12 x 12</td>
<td>45 x 45 cm</td>
<td>13 cm</td>
</tr>
</tbody>
</table>

Table 1: Comparison of 2.5D shape displays
input of a single user. A second version, FEELEX 2 decreases the spacing between the actuators for use as a medical palpation simulation device. While FEELEX displays are able to sense user input through touch, they do not allow additional input techniques.

Digital Clay is a research initiative at Georgia Tech investigating various shape rendering devices and their application as human computer interfaces. One of the proposed mechanisms is a 2.5D shape display [22], with applications proposed for 3D modeling by sculpting with fingers. A functional prototype with a 5 x 5 array of hydraulically driven actuators was developed. Potential user interactions for the system were investigated with PHANTOM haptic devices [5]. The proposed interactions are not evaluated on an actual 2.5D shape display, as they would require resolution and sensing capabilities that have not yet been achieved.

Lumen by Poupyrev et al. [20] is a shape rendering apparatus driven by a 13 x 13 nitinol actuator array, similar to the apparatus of Pop-Up [15]. Graphic overlay is integrated into the device by lighting each actuated pin with a monochrome LED. Applications for Lumen include animated shapes, reconfigurable tangible user interface elements, and connected shapes for remote presence. The input mode is similar to that of a touch screen with tactile feedback, thus it does not explore additional interactions, beyond direct touch.

The XenoVision Mark III Dynamic Sand Table trades output speed for a simplified actuation mechanism, which enables a high resolution through 7000 actuators [19, 4]. As rendering a complete shape on the current system takes more than 30 seconds, users cannot interact with the physical shape in real-time.


Photonastic Surface by Oguchi et al. proposes a novel mechanism to address individual actuators of 2.5D shape displays using projected light [18].

Relief, developed by Leitinger and Ishii [14] proposes an actuator apparatus based on commercial hardware and open-source components for low cost and scalability.

Actuator arrangements similar to 2.5D shape displays exist on different scales, from fingertip sized tactile displays [1] to wall-sized displays [7]. However, the focus of this paper is on shape displays at the scale of the users hands.

**Freehand Gestures**

To date, 2.5D shape displays have only utilized touch as an input modality. Based on our investigation of the limitations of direct touch, we propose to extend this modality with freehand gestures. Our proposed gestures have been inspired by previous research using pinching to interact with 3D data.

Wilson proposes a robust computer-vision algorithm for detect freehand pinching gestures [27] to zoom, translate and rotate digital content. This method has been applied for large screen interaction, such as dome-shaped displays [2] and to complement touch on interactive tabletops [9]. The metaphor of pinching a virtual string connected to a 3D object has also been applied by Grossman et al. to interact with volumetric displays [8] and by Zigelbaum et al. to interact with large screens [28].

**INTERACTIONS**

Due to the limited resolution of current 2.5D shape displays, they can only generate an approximation of the actual model. Therefore, only a subset of the total model geometry might be shown at a time (Fig. 5). Thus global operations such as zoom, translation and rotation are essential, in order to view details of the model.

3D modeling software (e.g. AutoCAD), allows modification of model geometry through transformation of individual vertices. These operations require the selection of a subset of pins on the display.

To summarize, the following interactions need to be supported:

- Transformation of scene view
- Selection of model vertices
- Transformation of selected vertices

In order to implement these interactions, first we demonstrate the limitations of direct touch, followed by a discussion of other input modalities.

**Limitations of Touch and Deformation**

Touch input on 2.5D shape displays has been previously limited to buttons [21] or button like behavior [12]. Touching the surface acts like a mouse click in traditional graphical user interfaces (GUI). Interpretation of touch input varies greatly between applications. Proposed interactions have also explored behaviors similar to a touch screen: users can swipe their hand, leaving a shape trace on the interface [21].

Because touch can perform many different functions, its implied use becomes ambiguous. Problems arise when we examine the requirements of a simple 3D modeling application. We define the most basic set of user interactions as the transformation and deformation of an object. Deformation
of object geometry can be achieved by mapping surface changes to the object mesh. This creates a clay-like behavior, allowing the user to push and pull on the surface. While this interaction performs well when pushing, it is practically difficult to pull the surface to a desired height (Fig. 6).

Once object transformations are added to program functionality, input through touch becomes ambiguous. When a user presses onto an object, it is no longer clear whether the object should be moved or deformed. With each additional feature both the complexity of interaction and input ambiguity are increased. In order to keep low input ambiguity, while retaining the locus of interaction at the surface, we propose that a 2.5D shape display requires an additional mode of input.

An interesting design challenge of direct touch interaction is the “reach problem”. The resolution of the 2.5D shape display may hinder users to reach for certain areas due to surface topology (Fig. 7).

**External Controller as Additional Input**

We experimented with interaction modalities beyond direct touch by connecting a standard trackball mouse for global surface navigation. A series of interactions map to the features of the trackball. Rolling the ball in any direction pans the surface accordingly. Similarly, pressing either mouse button activates standard zoom in and out commands across the entire surface. In order to compare the use of an external controller against direct touch, the input features of pan and zoom were implemented directly onto the 2.5D shape display itself. The edge pins of the interface activate panning. Pushing the pins the user can pan the scene according to location of the push along the edge. Zooming in is achieved by pushing on center pins, while zooming out by pushing on opposing edge pins.

Upon comparing the two implementations we argue that while an external controller reduces interaction ambiguity, it also it removes the locus of interaction from the table, thus interrupting the user from performing actions efficiently.

** Gestures as Additional Input**

Learning from our previous tests, we identify gestural input as the best additional interaction modality with a 2.5D shape display. Gestures can be utilized to perform functional commands over the entire surface without altering the locus of interaction. By expressing gestures directly above the surface, a user can seamlessly switch between selection, manipulation, and translation of objects on a 2.5D shape display. When coupled with direct touch, a user is able to maximize functionality without creating input ambiguity.

A grammar of gestures has been implemented to explore basic functions used to interact with a 2.5D shape display. We have found that the most fundamental set of gestures includes: selection of an area, translation of the selection, rotation of the selection, and scaling of the selection. Further description of these techniques follows:

**Selection**

In order to select a subset of the surface the user forms two parallel vertical planes with their hands (Fig. 9) - a gesture commonly used to indicate the dimension of an object (e.g. “I caught a fish this big!”). The system indicates a selected area with a projected square selection rectangle. A two-finger pinch on either hand locks the selection dimensions, enabling manipulation through a number of gestures.

**Translation**

By adjusting hand position along the X, Y, or Z-axis, the user can simultaneously manipulate the height and position of the selection (Fig. 10). After reaching the desired height and position the user can release the pinch gesture, saving surface state, and resetting the interaction state back to selection mode.

**Rotation**

By rotating the locked hands about the X, Y or Z-axis, the selection rotates accordingly (Fig. 11).

**Scaling**

By changing the distance between the locked hands, the selection scales proportionally (Fig. 12). This gesture is an evolution of the ubiquitously popular pinch-zoom in natural user interfaces.
APPLICATIONS
Through two applications we show how gestural input and direct manipulation can be used to navigate and to manipulate a 2.5D shape display.

Viewing 3D models
As discussed in related work, 2.5D shape displays are unable to create arbitrary shapes. However, shapes without overhangs, such as terrain data, can be formed with reasonable accuracy as previously proposed [4]. To explore the advantages of interactive shape output, we developed a geospatial application, equivalent to a tangible version of Google Maps (Fig. 8). Users can translate, scale and rotate terrain in 2 dimensions through use of the previously described freehand gestures. In addition to terrain data, our application is able to load geometry from a 3DS file and output an approximation of the physical shape on the surface. Similar to the terrain scenario, the object can be translated, rotated and scaled, however these transformations can be applied in all dimensions.

Recompose
Recompose explores the potential as an input device through a system that allows expression of volumetric structures. While designing the system we adhered to the following criteria:

Re-purpose methods from traditional digital design
Methods such as translate, rotate and scale are computational manipulations of data that aid the efficiency of digital design. Previously, as physical matter could not be informed by digital data (it is impossible to scale a mass of clay), such methods could not be applied when designing with physical materials. However, as the surface of the 2.5D shape display is computationally informed, such methods can be utilized freeing the user of constraints imposed by physical matter and speeding up the process of design.

Direct Manipulation
Haptic sensation remains our primary feedback mechanism in traditional surface manipulation (e.g. clay sculpting). We designed a system where direct haptic manipulation is a core component of the interaction. By allowing direct manipulation of pin height, users are afforded such haptic feedback.

Modality
The modality change between direct manipulation and gestural input must be complementary and seamless. In order for direct manipulation and digital design methods to coexist, we need a clear dividing line between each functional mode; a design that allows users to seamlessly change between them.

Metaphors
Functional gestures should be rooted in metaphor to aid intuition. By having strong metaphors to existing interaction paradigms, gestural interaction with the system can be made more intuitive.
SYSTEM DESIGN

Hardware

We utilize a 2.5D shape display similar to the system proposed by Leithinger and Ishii [14]. The table consists of an array of 120 individually addressable pins, whose height can be actuated and read back simultaneously, thus allowing the user to utilize them as both input and output. We have furthered the design by placing a depth camera above the tabletop surface. Through gaining access to the depth information we are able to detect basic gestures from the user (Fig. 13a).

The pins are actuated with ALPS motorized slide potentiometers, widely used in audio mixing boards (Fig. 13b). Each pin has a reach of 13 cm vertically with a resolution of 7 bits, thus the pins can be controlled in 1 mm increments. While trying to achieve minimal spacing between actuators we were limited by the dimensions of the slide potentiometers. Currently, the actuators are arranged in a square grid with 5 cm spacing.

Arduino Duemilanove boards with attached Ladyada motor driver shields control the actuators. Instead of daisy-chaining the modules as previously proposed, our system utilizes I2C bus configurations for faster, more reliable communication. Our software application communicates over a serial connection with three Arduino master modules, which control up to 12 slave modules over an I2C bus. Using this method, the apparatus achieves a speed of 15 fps for communication between an application and the target positions on the microcontrollers. Though considerably slower than commercial haptic devices like the Phantom, this speed proves sufficient for the interface.

Software

Software on the control machine is written in OpenFrame-works. Rendering on the actuated surface is achieved by processing the depth buffer of the 3D mesh and sending the depth values to the hardware actuators. Direct touch input is detected by computing the difference of the last actuated and currently reported height.

Computer Vision

Freehand gesture input is sensed with a depth camera. The software uses a community contributed set of libraries to access the camera. After acquiring the depth image of the camera, the users contour is computed from the depth values (Fig. 14). For gesture recognition a state-machine is implemented that stores the current gesture state and reacts to changes in hand posture accordingly. After extracting curvature of the blobs corresponding to the hands of the user, our algorithm finds fingers and determines the current height of the hands.

DISCUSSION

Direct manipulation of an actuated surface allows us to precisely affect the material world, where the user is guided throughout the interaction by natural haptic feedback. However, direct manipulation is incapable of affecting large areas of a surface or small sections unreachable due to the reach problem. We demonstrate how free hand gesture
input can solve this problem without requiring the user to shift their focus to an external controller. We designed the system so users can fluidly change between gestural and direct manipulation.

After implementing two applications to demonstrate the interactions presented in this paper, they were tried by over 300 different visitors during various demonstrations in our research laboratory and at Siggraph Emerging Technologies 2011. While we have not conducted formal experiments, we were able to gather interesting feedback and observations:

In the application to view 3D models, many users commented on how they found it easier to understand the shape of the model when shown on the interface compared to a 2D display. While we expected this feedback, it was remarkable how the low resolution of the output apparatus was sufficient to present a perceptual advantage. Projected graphics were key to aiding users in identifying the 3D models. Without projection, the rendered shapes were too low in resolution for the user to perceive the underlying 3D model.

In the Recompose application users found direct manipulation to be intuitive. While users understood selection feedback, it was not apparent that a pinching gesture would be required to transform a shape. Once demonstrated, all users were able to replicate this functionality and many could relate to the interaction metaphor of pinching strings. However, the tracking of small hands was sometimes inconsistent due to the camera resolution.

These observations of visitors using both touch and freehand gestures are encouraging indicators that the proposed interactions are intuitive and efficient for interacting with 2.5D shape displays. The proposed freehand gestures have been utilized previously for spatially transforming data, but our applications demonstrate that they work well in the context of 2.5D shape displays.

CONCLUSION AND FUTURE WORK

In this paper we have identified 2.5D shape interfaces as a distinct subcategory of shape displays. We found properties common to all 2.5D shape displays and listed their advantages and limitations.

We have presented interaction techniques for 2.5D shape displays, which extend manipulation through touch with freehand gestures. After identifying a set of common interactions for viewing and manipulating content on shape displays, we proposed freehand gestures to implement them.

We intend to further the system by implementing a number of additional interaction features beyond the previously explored select, translate, and transform functions. We would like to explore the following set of features:

Copy and Paste
Additional gestures would allow users to quickly duplicate shapes. The duplicated shapes could be both digital models and physical objects captured by the depth-sensing camera.

Timeline Controls
A timeline of all surface manipulation captured during a session allowing for time reversal, state saving, and undo type functions. We believe that implementing a set of gestures to explore the dimension of time grants the user additional power currently impossible by the constraints of the physical world.

Advanced Drafting Techniques
We plan to implement techniques commonly found in drafting and 3D modeling tools, such as spline lofting and Boolean operations.

System design
We intend to pursue both higher fidelity gesture recognition coupled with higher resolution actuated surfaces to allow for a greater range of expression and interpretation.

ACKNOWLEDGEMENTS

We would like to thank our colleagues at the MIT Media Lab, especially Adam Kumpf and Sean Follmer for ideas and insights.

REFERENCES


