# Sublimate: State-Changing Virtual and Physical Rendering to Augment Interaction with Shape Displays

**Daniel Leithinger   Sean Follmer   Alex Olwal   Samuel Luescher   Akimitsu Hogge   Jinha Lee   Hiroshi Ishii**
MIT Media Lab
75 Amherst Street, Cambridge, MA 02139, USA
{daniell, sean, olwal, luescher, jinhalee, ishii}@media.mit.edu

## ABSTRACT

Recent research in 3D user interfaces pushes towards immersive graphics and actuated shape displays. Our work explores the hybrid of these directions, and we introduce *sublimation* and *deposition*, as metaphors for the transitions between physical and virtual states. We discuss how digital models, handles and controls can be interacted with as virtual 3D graphics or dynamic physical shapes, and how user interfaces can rapidly and fluidly switch between those representations. To explore this space, we developed two systems that integrate actuated shape displays and augmented reality (AR) for co-located physical shapes and 3D graphics. Our spatial optical see-through display provides a single user with head-tracked stereoscopic augmentation, whereas our handheld devices enable multi-user interaction through video see-through AR. We describe interaction techniques and applications that explore 3D interaction for these new modalities. We conclude by discussing the results from a user study that show how freehand interaction with physical shape displays and co-located graphics can outperform wand-based interaction with virtual 3D graphics.

## Author Keywords

Shape Display, Actuated Tangibles, Spatial Augmented Reality, 3D Interaction.

## ACM Classification Keywords

H.5.2 User Interfaces: Graphical user interfaces, Input devices and strategies, Interaction styles.

## INTRODUCTION

Since Ivan Sutherland's vision of the "Ultimate Display" [30], researchers have aimed to create an immersive environment with the ability to render virtual and physical elements anywhere in 3D space. Although there has been much research in rendering immersive 3D graphics spatially co-located with the user, from Virtual Reality (VR) to Augmented Reality (AR), fewer research projects focus on rendering physical forms. The most common systems render a haptic sensation of objects through articulated arms [18], or require the user to be instrumented with gloves or cables
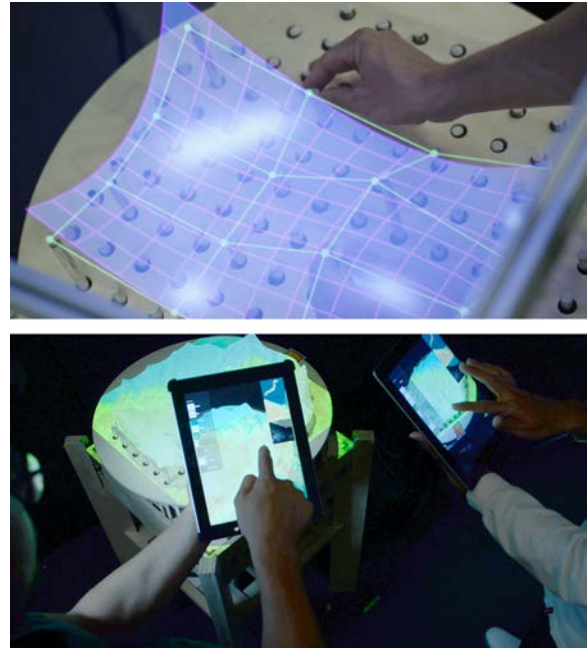
Figure 1: Sublimate combines augmented graphics with actuated shape output. (Top) A user manipulates a virtual mesh through physical deformation of an actuated pin array using optical see-through AR. (Bottom) Multiple users collaborate to edit geospatial data on a shape display augmented with handheld video-See through AR.

[32]. While such devices have been combined with spatially co-located 3D graphics through VR and AR [27, 21], we believe that they fall short of the vision of the "Ultimate Display", as the haptic sensation is limited to discrete points. For this reason, users are commonly aware that the represented object is not real. Another approach is to render the actual shape of physical objects, as proposed by research visions like "Claytronics" [1] and "Radical Atoms" [14]. Systems following this approach include shape displays, which utilize actuators to render objects that users can see, touch and manipulate with bare hands [17]. Current generation shape displays trade the advantages of real objects for the flexibility and realism of high-resolution graphics in VR interfaces. We propose that a combination of these two modalities can open up a rich area of research.

Our vision is that 3D information can be rendered in space as physical objects or virtual graphics. We believe that the most interesting aspect may not be either state alone, but rather the

combination and fast transition from virtual to physical, and vice versa. This approach is different from common AR applications, where elements are either physical or virtual, but do not switch between states. Thus we are not only interested in augmenting shape displays with graphics, or adding haptic feedback to AR, but also how the transition between physical and virtual can enable new user interactions (see Figure 1). Physical models can be partially replaced by floating graphics, allowing the user to physically manipulate a part inside. Virtual interface elements become physical when they need to be touched or modified.

In order to explore this space of virtual/physical state transitions, we designed two implementations of a system called Sublimate, which combines spatial AR with actuated shape displays. The first combines a optical see-through AR display, utilizing a stereo display, acrylic beam-splitter, and head tracking, with a shape display to co-locate 3D virtual graphics and a physical 2.5D surface. The second uses tablet-based video see-through AR displays to add virtual graphics to the scene. Both systems allow for direct interaction from the user, through mid-air interaction with a wand and through physical manipulation of the shape display.

We begin the paper with an overview of related work. Next, we introduce the Sublimate concept and discuss interactions. We describe prototype applications to demonstrate the concept and document the implementation details of our two systems for augmenting shape displays. We then report on a formal evaluation of our system that investigates different input styles with 3D content on a spatial optical see-through display combined with shape output. We discuss these results, which indicate that interacting through direct touch on a shape display can be faster than mid-air manipulation with a wand, and present user feedback on the Sublimate system.

### CONTRIBUTIONS

- Exploration of transitions between physical rendering and virtual 3D graphics, and interaction techniques leveraging such state changes.

- Practical implementations to prototype interactions combining actuated shape display with co-located 3D graphics, using optical see-through spatial AR displays and handheld video see-through AR devices.

- Extension of a shape display's resolution, size and scale through co-located virtual graphics.

- Extension of spatial AR with physical shape rendering.

- User evaluation of interaction styles for these systems; tangible manipulation and mid-air interaction.

### RELATED WORK

To support individual and decoupled control over an object's visual appearance and physicality, we require two different techniques. First, we need a display technology that can show graphics both floating in mid-air, as well as overlaid and registered with a physical object. Second, we need techniques that allow us to control the presence of an object's physical parts or components.

### Situated see-through displays for spatial AR
Numerous projects explore techniques where a partially transparent, see-through display augments real objects or environments with superimposed information or graphics [5, 4, 19]. These spatial AR systems can also be combined with passive tangible input in the real world [11].

### See-through displays with co-located manual input
Schmandt [28] describes an early setup, which emphasizes the perceptual advantages of co-locating stereoscopic imagery with the user's hand and input device. A half-silvered mirror reflects 3D graphics that is optically merged with the user's hands underneath, registered using a 3D input device. Yoshida et al. [33] use an LCD and lens array to also provide parallax through retro-reflective projection off an arbitrarily shaped bottom surface. Toucheo [7] demonstrates how these configurations can be combined with multi-touch surfaces and on-surface interaction techniques for 3D manipulations, while HoloDesk [12] uses depth cameras to explore whole-hand interactions, object tracking, motion parallax and physics simulations for enhanced realism.

### 3D displays with co-located tactile feedback
Co-located Spatial AR displays can also be extended to incorporate tactile feedback through haptics. The Haptic Workbench [29] adds single-point force feedback through a PHANTOM device [18], a configuration also explored by Scharver et al. [27] in an immersive interface for tangible design of cranial implants. Plesniak et al. [21] describe the Computational Plastic concept, which envisions the future of real-time programmable material properties through haptics and real-time holography. They demonstrate a number of proof-of-concept systems based on single-point haptics and holograms. Touchable Holography [13] enables force feedback without mechanical devices by using ultrasound for a tracked finger in a 3D display. Other work, such as the SPIDAR-8 [32], has explored precise multi-finger haptic interaction using cables attached to individual fingers. Physical objects can also be used to provide passive haptic feedback, and allow for augmented tangible interaction, so long as their locations are tracked [26].

### Projection-based AR
Projection-based AR approaches have been explored in many projects to alter the visual properties of physical objects [23], particles [20], surfaces [3], or the user's body [8]. One motivation for such systems is that they can modify the appearance of everyday objects without requiring an additional display surface.

### AR interfaces for control of physical objects
AR is also well-suited for visual support and feedback during control, manipulation and actuation of devices and objects. Tani et al. [31] describe a user interface for manipulating physical controls on remote machinery through an augmented video interface. TouchMe [10] applies direct-manipulation techniques for remote robot control using video see-through AR and a touch-screen interface. Ishii et al. [15] enable real-world pointing and gesturing for robot control, using a tracked laser with projected visual feedback.

**Shape-changing interfaces**
Various projects exploit techniques for moving or displacing physical matter as a means to control and affect physical shapes [24]. Lumen [22] provides individual control of shape and graphics by varying the height of LED rods using shape-memory alloys, whereas FEELEX [16] employs a flexible screen overlaid on the actuators for a continuous surface, and top-down projection for graphics. Relief [17] investigates direct manipulation and gestural input to enable interaction techniques that match the capabilities and potential of 2.5D shape displays. AR-Jig [2] is a 3D-tracked handheld device with a 1D arrangement of linear actuators, for shape deformation and display of virtual geometry, viewable through an AR display.

A common motivation for both AR systems and shape-changing interfaces is to unify virtual and physical representations to enable richer interfaces for viewing and interaction. Projects like AR-Jig have explored how to co-locate haptic feedback with AR. With this paper, we introduce additional expressiveness, by enabling dynamic variation of the amount of graphics and physical matter used to represent elements in the interface, and exploring state change between rendering either as virtual or physical output.

## SUBLIMATE VISION AND CONCEPT
Our vision of Sublimate is a human—computer interface with the ability to computationally control both virtual graphics and physical matter. An object rendered through this system can rapidly change its visual appearance, physical shape, position, and material properties, such as density. While such a system does not currently exist and might be physically impossible to build even in the future, we aim at creating interfaces that appear perceptually similar to the user through a mix of actuated shape displays and spatially co-located 3D graphics. We focus on computationally controlling a specific parameter: the physical density of objects. Objects rendered through our system can rapidly switch between a solid physical state and a gas-like floating state. With programmable affordances, objects can be physically rendered when needed and are still visible when not. We call this concept "Sublimate", as it is inspired by the phase transition from solid to gaseous in a thermodynamic system.

The most commonly encountered thermodynamic phases of physical materials are solid, liquid and gaseous. Material properties, such as density, rapidly change between these phases, as one can observe in ice, water and steam. We apply this metaphor to the relationship between physical and virtual output in a "Sublimate" interface (see Figure 2). Similar to the iceberg metaphor describing Tangible Bits [14], the liquid state of water represents the data model, while the solid state represents tangible physical objects for the user to interact with. We extend this metaphor with a gaseous state to represent spatially co-located 3D graphics. A Sublimate interface renders data as a solid object through a shape display or as spatial 3D graphics through an AR display. We refer to the transitions from shape output to 3D graphics as "sublimation," and the transition from 3D graphics to shape output as "deposition". The system can also render partially sub-
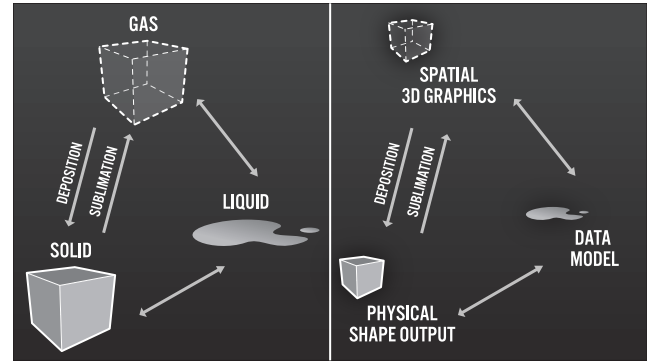


Figure 2: The physical phase transition between gas and solid material states informs our vision of Sublimation.
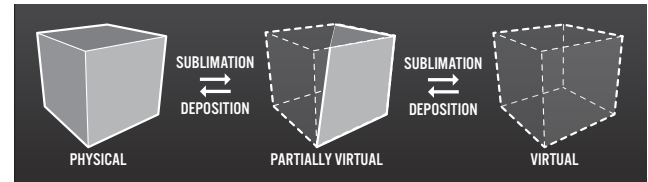


Figure 3: We introduce *sublimation* and *deposition*, metaphors for the transitions between physical and virtual. A Sublimate system can use these capabilities to transform objects representation from physical–partially virtual–virtual.

limated objects, which consist of both tangible physical and intangible graphical portions (See Figure 3).

## DESIGN GUIDELINES
The guiding principles for the design of a Sublimate system are:

- The output should perceptually be as close as possible to real world objects. This means that instead of solely providing a haptic sensation for selected points of interaction, the aim is to render real objects. Users should be able to touch these objects with their bare hands and naturally view them from different directions.

- Synchronized output channels, with the ability to rapidly switch the rendering of an object between them. The system can represent information as graphics, physical objects or both.

- User input through multiple modalities. Users can interact with the system through symbolic commands, gestures and direct touch, based on the physical channel they currently interact with.

## INTERACTIONS
The Sublimate system can render both physical and virtual representations of an object's shape. As the modalities of shape output and virtual graphics are synchronized, the system can render an object in either one of them independently, or in both modalities at the same time. This flexibility in representation and transitioning between modalities enables new interactions capabilities.

**Physical to Virtual: Sublimation**
There are many scenarios where it is advantageous to transition to virtual object representation.

*Transcending physical boundaries*
Parts of physical objects can be rapidly switched to a virtual representation. This allows the system to relax physical constraints, with maintained visual consistency. A user could thus, for example, reach through previously physical parts, for manipulation inside or behind an object.

*Flexible scale and appropriate representation*
Rendering objects with graphics allows the system to overcome the constraints of physical shapes, for example, when a physical representation would be too large or impractical. It also makes it possible to avoid potential safety issues for certain types of physical output.

*Unconstrained manipulation*
Sublimation can be used to enable virtual controls that are not constrained by the shape display's degrees of freedom. This allows the switching from precise constrained interaction with physical control to unconstrained mid-air interaction.

*Visualizing impending materialization*
Graphical previews are an effective way to inform users of impending shape actuation and can, for example, allow them to cancel or confirm the output in progress. This can be particularly useful if the generated shape would interact with other physical objects in the space.

**Virtual to Physical: Deposition**
Deposition provides new ways in which rapid materialization can enhance user interaction.

*Dynamic physical affordances*
Deposition makes it possible to have physical affordances appear dynamically. User interface elements, such as buttons, knobs, sliders and handles, can be rendered graphically and appear physically only when they need to be manipulated.

*Adaptation to contexts and scenarios*
Objects can morph between representations to adapt to changing interaction constraints. A virtually rendered surface can, e.g., materialize when a user approaches it with a finger, and upon proximity with a stylus tool, morph to a flattened shape to better support annotation.

*Guiding interaction with physical constraints*
Physical shapes can be used to restrict movement and interaction to permitted areas, or to provide guiding lines for manipulation. This could enable a user to "freeze" parts of a volumetric data set, or help a user edit a texture map, for example.

**Mixing Virtual and Physical**
Graphics can help to compensate some of the limitations of current generation shape displays. They enhance the visual resolution, size and scale of shape output, and augment features a particular type of shape display might not be able to render, such as overhangs.
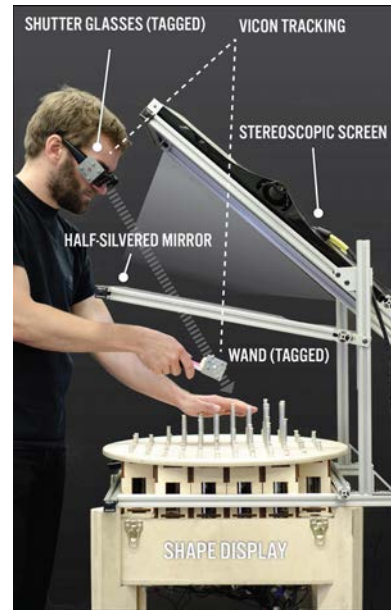


Figure 4: Single-user setup using head tracking, a stereoscopic display and a beamsplitter, to overlay transparent 3D graphics on a shape display.

In addition to transitions between states, many interactions can benefit from the combination of shape output and virtual graphics. We extend classic AR applications where floating graphics augment physical objects, by also introducing dynamic shape change. An example is to visualize the wind flow around moving physical objects. Another application is to overlay alternate versions of an object onto its physical shape in CAD scenarios, similar to "onion skinning" in animation software.

**SUBLIMATE SYSTEM**
We built two proof-of-concept setups to prototype the envisioned interactions of the Sublimate concept. Each setup consists of two main components, a system to render the physical shape output and a display for the spatially co-located 3D graphics. Physical shapes are rendered through a 2.5D shape display, based on our previously introduced "Relief" system [17]. To view the spatially co-located 3D graphics, we utilize display arrangements well-known in AR: a stereoscopic spatial optical see-through display for single users (Figure 4) and handheld video see-through displays for multiple users (Figure 5).

The setup designed for single users renders 3D graphics on a stereoscopic display with a beam splitter, mounted on top of the shape display. When viewing the physical shape through the beam-splitter with tracked shutter glasses, the graphics appear co-located. To explore co-located multi-user interactions, we also propose a version in which the 3D graphics are rendered on handheld video see-through displays. While the graphics are not co-located in physical space, they are aligned with the video view of a camera mounted on the back of the tablet screen. As the display is handheld, it limits user interactions with the physical shape display to a single hand. In

Figure 5: Multi-user setup, using handheld tablets to augment the shape display through video see-through AR.

future work, we plan to explore head-worn displays to overcome some of these limitations in scenarios where face-to-face interaction is less important and instrumentation of the user's head is acceptable. Another advantage of the handheld display is the built-in touchscreen, which provides an additional input modality for interacting with the content.

## PROTOTYPE APPLICATIONS

In order to highlight features of the Sublimate system we created a number of example applications in different domains, such as computer aided design (CAD), geospatial data visualization and volumetric rendering of medical data. These different applications demonstrate how objects and interaction elements can transition between physical and digital states, as well as showing how augmented graphics can increase the resolution, fidelity and scale of shape displays, and provide augmented feedback to the user.

### Single User Applications

*NURBS Surface Modeling*
Manipulation of 3D meshes is challenging with traditional 2D input devices, such as mice, and therefore alternatives input devices are being developed. Gestural input has advantages due to more degrees of freedom, but lacks the material feedback of deforming real objects. We propose a basic application that combines physical control for mesh manipulation with an overlaid graphical view of the resulting surface. The control points of a NURBS (Non-Uniform Rational Basis Spline) surface are represented by individual pins on the shape display. Grabbing and moving the pins up and down affects the resulting surface, which is displayed through co-located 3D graphics. The control points are simultaneously highlighted through graphical feedback. The user can press
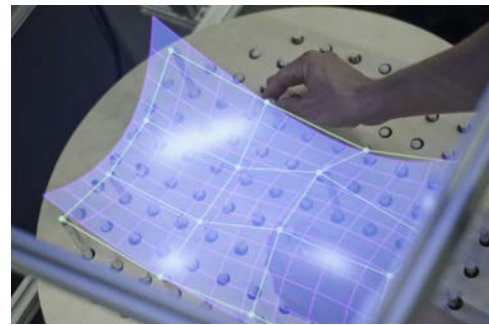


Figure 6: NURBS Surface Modeling. The user can switch between rendering a) the surface physically and control points virtually, or b) control points physically and the surface virtually.
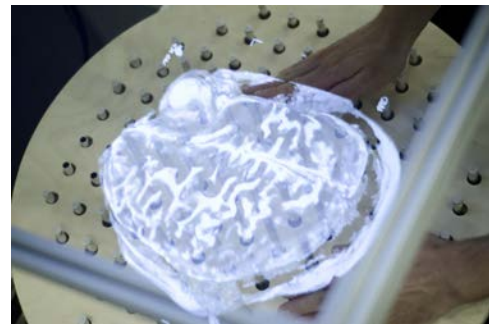


Figure 7: Volumetric Medical Data Viewing. Users can modify cross sections through the volume by physically deforming the shape with their hands. They can switch between defined cross sections through sublimation.

a button to toggle the NURBS surface rendering from graphical to physical. In that case, the shape display outputs the geometry of the modeled surface instead of the control points and the user can feel the physical deformation. This application highlights the ability to dynamically sublimate control widgets, to allow for more precise control, or to provide more degrees of freedom.

*Volumetric Medical Data Viewing*
Volumetric data sets are rendered as 3D graphics that are spatially co-located with a physical shape in this application. The physical shape represents the bounds of the volume ray casting algorithm and can be reshaped by the user to create a non-planar cross section through the volume. This interaction is similar to Phoxel Space [25], but has the advantages of an actuated shape display, such as being able to save and load cross sections, or to define parametric shapes. The cross section can be conveniently flattened and moved computationally, while the user can intervene at any time by modifying its shape by hand. The location of the 3D graphics is not restricted to the surface of the cross section, as volumetric data underneath or above the surface can be rendered to get a better understanding of the data set. This application demonstrates how the system can quickly sublimate data to expose contextually meaningful areas.
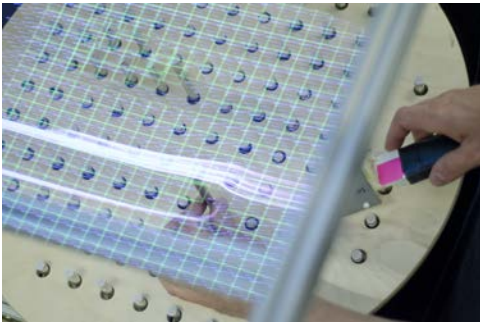
Figure 8: Virtual Wind Tunnel. The wind flows around a dynamic physical model formed by the user, and is visualized through overlaid 3D graphics.



Figure 9: Multi-user geospatial data exploration. Handheld tablets augment shape display, add layers of data, and extend the active workspace. Interaction can be done through the tablet or through the shape display.

*Virtual Wind Tunnel Simulation*
The virtual wind tunnel application renders different materials in their appropriate modality. While solid models are rendered on the physical shape display and can be touched and manipulated by the user, wind flow is displayed through spatially co-located 3D graphics. When the user deforms the physical model, a cellular fluid dynamics wind simulation updates accordingly. The wind flow around the model is visualized as transparent white lines floating in mid-air. To get a better view of the wind flow at a particular location, a tracked wand can be placed in the space around the model to disperse color into the simulation. The virtual wind tunnel shows the advantages of augmenting shape displays with virtual graphics, and having bi-directional control of the output.

**Multi-User Applications**
*Physical Terrain Model with Superimposed Virtual Information*
To explore multi-user interactions, we developed an application for collaborative discussion of geospatial data. In this application scenario, the shape display renders physical terrain, while several tablet computers can be used to simultaneously interact and augment the physical surface. Seen through the camera of the tablets, we can expand the horizon of the map and display the terrain as it extends far beyond the edges of its physical manifestation. Users can adjust the region of interest of the map rendered on the shape display by using pan and zoom touch gestures on the tablet interface. Moreover, individual users may display additional data overlays visible

through their tablets, which align with the captured image of the shape display taken from the tablet's camera. In our scenario, we provide a map showing radioactive contamination levels in Japan, as well as other geospatial map layers. The advantage of this configuration is that users can refer to the physical model during discussion with each other, while controlling a personal high-resolution view that allows them to switch between different perspectives of surrounding terrain or additional data layers.

## IMPLEMENTATION

### Shape Output with Optical See-Through Display
Our single-user setup consists of a 2.5D shape display and a co-located semi-transparent 3D display. The shape display is based on a hardware setup similar to Relief [17], consisting of a table with 120 motorized pins extruding from the tabletop. The pins have a vertical travel of 100 mm and are arranged in a $12\times12$ array with 38.6 mm spacing. The 3D graphics are rendered in $1920\times1080$ pixel resolution at 120Hz on a $27''$ LCD screen, mounted on top of a semi-transparent acrylic beam splitter, and viewed with NVIDIA 3D Vision Pro shutter glasses (60Hz per eye). In addition to stereoscopic output, the user's head position is tracked by a Vicon motion capturing setup consisting of 10 cameras. This system creates a $425\times425\times100$ mm$^3$ space in which physical and graphical output are co-located for a single user. The shape display is controlled by a 2010 Mac Mini, which communicates with the application PC though OpenSoundControl (OSC). Applications and graphics rendering are running on a Dell Precision T3500 PC with a 2.53 GHz Xeon W3505, 8GB RAM and a NVIDIA Quadro FX 4600 running Windows 7. All applications, as well as the hardware control software, are implemented in OpenFrameworks (OF). The system runs at 60fps.

### Shape Output with Handheld Video See-Through Display
To explore co-located multi-user interactions, we also built a version in which the co-located 3D graphics are displayed on handheld video see-through displays. We utilize 3rd generation iPads, which display a fullscreen video captured by their rear-mounted cameras. A custom OF application tracks visual markers placed around the shape display using the Qualcomm Vuforia API. After computing the screen position relative to the shape output, the video view is overlaid with adjusted 3D graphics. User input is synchronized between multiple iPads over WLAN using OSC. The shape display is augmented with projection onto the object surface to enhance appearance and provide graphical feedback when viewing the shape without the iPad. The projector displays XGA graphics, which are rendered by a custom OF application running on a 2011 Macbook Air. The shape display is controlled by a 2010 Mac Mini, which communicates with the application computer though OSC. The system runs at 60fps.

## USER STUDY
To evaluate the Sublimate system, we conducted a user study to measure the advantages of shape output combined with spatial graphics. We investigate how interacting with spatial AR without haptic feedback compares to spatial AR with

co-located shape output, and to spatial AR with single-point haptic interaction. Our reasoning was that if there were no advantage to physical feedback, then virtual rendering would suffice, and state transitions would be unnecessary. In the study, we tested the following hypotheses:

$H_1$: Physical input is easier and faster than mid-air gestural input for spatial manipulation tasks when interacting with co-located spatial graphics. Haptic feedback provided by shape output is advantageous compared to mid-air interaction with only virtual graphics.

$H_2$: Multi-point, two-handed manipulation of a 3D surface is easier and faster than single-point haptic interaction. Whole-hand interaction is more effective than finger- or single-point interaction.

We collected informal and anecdotal data from users on how well they felt that the virtual graphics aligned with the shape display, the perceived effective difference between virtual or physical rendering when viewed, and general ease of use.

As highlighted in [24], few user evaluations of shape displays exist and we believe that an important first step is to quantify the advantages of direct interaction with shape displays coupled with virtual graphics. In future work, we plan to follow up with investigations of the dynamic transition between physical and virtual states.

**Experiment**
To investigate these hypotheses we chose 2.5D mesh manipulation for CAD, a task domain in the area of actual use that we imagine for the Sublimate system, and that allows for bimanual interaction.

We ran our study using the see-through AR version of Sublimate as it provides for higher accuracy matching of graphics and shape output, while leaving two hands free for input. We used the same 3D scene in all conditions and rendered it stereoscopically at a 1920×1080 resolution on a 27″ LCD in portrait mode, which the participants viewed with active shutter glasses.

To ensure accurate tracking in our study, we used a Vicon motion capture system for both head tracking and 3D input, as opposed to, e.g., a depth camera. We rendered view-dependent graphics based on head position, by tracking a tag on the stereo glasses. A pointing wand was used for 3D input, and the participant used a separate button with the nondominant hand to trigger selections, to avoid potential errors from the shaking that could be induced by a wand-mounted button. For physical input and output we made use of the shape display's physical pins. The pins were 10 mm in diameter, and had a vertical travel of 100 mm.

**Participants**
10 participants (4 female, age 23–40, one left-handed) were recruited through a department e-mail list. All participants were regular computer users, 8 had used some type of 3D display before (including 3D movies), and 4 were at least monthly users of 3D input devices such as a Nintendo Wiimote or Sony PlayStation Move.
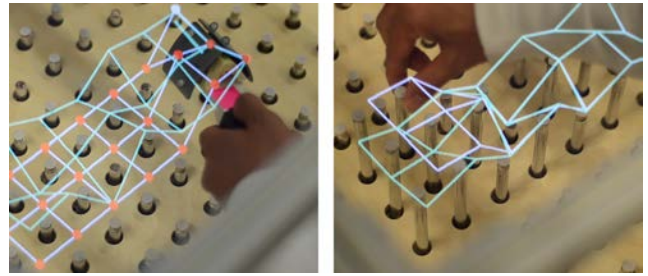


Figure 10: 3D Surface manipulation task and two of the experimental conditions. (Left) Wand interaction with virtual graphics. (Right) Single-handed shape display manipulation.

**3D surface manipulation task**
In the 3D surface manipulation task, the participant is asked to match a target surface with a co-located input surface. Both the input surface and the target surface are displayed as a wire-mesh rendering. In order to test our 2 hypotheses, we developed the following conditions:

- *Wand*. Single-point manipulation of virtual graphics (Wand with Vicon marker, pressing button with non-dominant hand)

- *Single-push*. Single-point manipulation. Physical pins starting *up*.

- *Single-pull*. Single-point manipulation. Physical pins starting *down*.

- *Multi-push*. Whole-hand and multi-point manipulation. Physical pins starting *up*.

The two meshes were always co-located and rendered in different color, and the goal was to match the input mesh to the target mesh. In the conditions where the participants manipulated the physical shape display manually, each of the vertices was rendered physically by the height of the pin, and virtual graphics displayed edges connecting the pins, as shown in Figure 10. When using the wand, both meshes were displayed virtually. Each mesh had 7×3 vertices, spaced evenly in the x and z dimensions, 38.1 mm apart. The meshes were randomly generated and vertices were normalized between the upper and lower bounds, 100 mm apart. Because the pin display is limited to one degree of freedom per pin, we constrained the mesh vertices only to y-displacement in all conditions. All interaction was direct.

For the wand condition, participants had to select and move vertices using the end of a virtual cursor that was overlaid on the physical wand. The non-dominant hand was used to press a button to select the closest vertex. The virtual vertices were rendered as spheres, matching the pin size with a 10 mm diameter.

In the single-handed pin manipulation conditions (Single-push and Single-pull), participants were instructed to only manipulate one pin at a time, to be comparable to the wand condition. In the bimanual condition (Multi-push), participants could manipulate as many pins at once as they wanted,
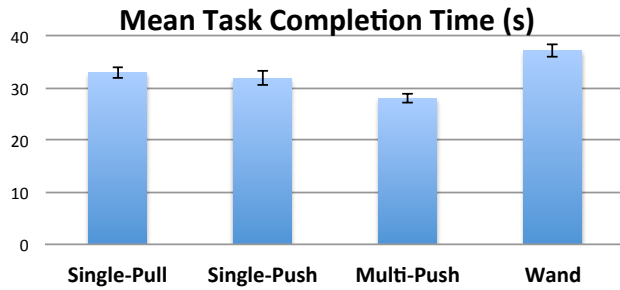
Figure 11: Task completion time between different input conditions. Error bars are +/- SEM.

using their finger, palms or any surface of their two hands. We wanted to also compare the effects of the pins starting down vs starting up, which would require the participant to primarily either pull or push on the pins.

A total of 10 sets of meshes were displayed per trial. As soon as the participant matched all vertices with the two meshes, the current mesh was cleared and a new target mesh was displayed after a 3 second timeout, during which the screen flashes red, yellow, then green, to alert the participant that the next mesh was about to be displayed.

**Procedure**

We used a within-subjects repeated measures design. The order of the 4 different conditions was counterbalanced. Participants were instructed to complete the tasks quickly and were informed that it was a time trial task. After completing each condition, participants would take a 30 second break and fill out a short form based on the NASA Task Load Index [9] to gauge mental and physical demands of the completed task. The experiment lasted 60 minutes. Participants were observed and video recorded for later analysis. Participants filled out a post-test questionnaire and were interviewed about the conditions for qualitative feedback on the system.

**Results**

We present the results of the mesh matching task. The average task completion time of a single $3\times7$ mesh for all conditions was 32.55 seconds. With one-way repeated-measure ANOVA, we found a significant difference between the four input conditions (F(3,27)=8.033, $p < 0.01$, partial $eta^2 = 0.47$)). Figure 11 shows the mean task completion time for all conditions. Multi-push was fastest (28.10s), followed by Single-push (31.97s), Single-pull (32.94s) and the Wand condition (37.20s). Post-hoc pair-wise comparisons (Bonferroni corrected) identified a significant difference in completion time between Multi-push and Wand conditions, and Multi-push and Single-push condition ($p < 0.05$). There was no significant difference in accuracy across conditions.

*Wand versus Pin Manipulation*

Our hypothesis was that physical pin manipulations would be faster than mid-air interaction with the wand. The results show that while task completion for all pin manipulation conditions was faster than when using the wand, only the Multi-push condition was statistically significantly faster. The actuated shape display was designed for two-handed pin manipulation, and that is the dominant method of input using the shape display; therefore we argue that this study validates the hypothesis that the shape display can perform better than a mid-air 3D pointing device. The physical pins provide many benefits in this controlled scenario, such as constrained movement and haptic feedback.

There may be several reasons for the lack of significance in the single-handed pin conditions. Firstly, the wand condition allowed participants to select the closest vertex with a single button press — thus relaxing the accuracy requirement in target acquisition. Secondly, participants mentioned that they sometimes obstructed the interaction with other pins, which could have made the physical pin conditions more challenging. Many participants noted this problem, and even those who did not prefer the wand thought that the lack of obstruction while using it was a clear advantage: "The wand is better at not getting the pins in the way, but it tires you more and it doesn't feel too precise" (P4). Participants developed several strategies to minimize the obstruction of interaction from surrounding pins, which limited this problem: "I had to be careful about strategy and order" (P5). Some participants felt that the bimanual condition alleviated some of this problem. This concern of pin obstruction has been previously discussed [17] and may be one of the key limitations of manipulating and interacting with physical shape displays, which may be addressed through different interaction techniques.

We also wanted to look at pin manipulation task completion times and how these were affected by pin starting location; was it significantly easier to push or pull the pins? We had assumed that pushing would be easier. The results show that it was faster, but not significantly. However, we limited interaction to a single pin at a time in both of these conditions; it is possible that one could push multiple pins more easily than pulling up multiple pins with one hand. Additionally, in the post-test questionnaire, participants preferred pushing (mean 5 out of 7) to pulling (mean 3.5 out of 7) ($p < 0.05$). Participants also reported different strategies for ordering interaction between pushing and pulling; when pulling, many participants started at the back of the mesh, and when pushing, many participants began at the front.

*Bimanual Interaction*

Bimanual pin manipulation, with pins starting up, was significantly faster than both the pin manipulation condition, with pins starting down, and the wand interaction ($p < 0.05$). Participants also often commented in the post-test questionnaire that using two hands was much easier and felt more intuitive than the single hand or wand conditions. "Two-handed interaction felt the most natural. I felt like I was molding the pins into shape" (P1). "It felt more organic" (P5).

There were a number of different strategies with the bimanual condition. One strategy was "to do an unrefined pass with my left hand and a refined pass with my right hand" (P2). However, some participants felt that though they could be faster

with the bimanual condition, it felt more taxing. "I found that I made more errors when using two hands, which I had to later go back and correct — though this method felt faster than using one hand" (P3).

*Alignment, Co-location, and Discomfort*
Participants responded in post-test interviews that they felt that the virtual graphics aligned well with the physical pins, and that the head tracking and view-dependent rendering worked well. "The graphics matched well, and giving you [visual] feedback [was helpful]" (P1).

The overlay of virtual graphics on the physical pins did not seem to have a nauseating effect on the participants. Only one participant reported very slight nausea, and none asked to take a break or stop. 3 participants complained about back pain after using the system for 45 mins. It was also difficult for some participants to reach some pins at the back of the array, although none of these pins were used during the actual trials.

*Hardware Limitations and Universality*
One effect of the shape display is that users were more surprised when the shape display cleared all of the pins, than in the virtual case. Almost all participants appeared surprised at least once, when the pins changed dramatically. It is unclear if this had any effect on performance. This is a possible limitation of sublimation-based interaction techniques, where the physical shape changes quickly.

While our study focused on evaluating direct manipulation on 2.5D shape displays [16, 17] with co-located augmented graphics, we believe that results will be similar using different shape display hardware. Even with very limited shape display hardware, there are positive results that show that these type of interfaces can perform better than freehand gesture in certain cases. We think that future hardware will only improve these results.

In addition, it is worth noting that other interaction techniques could be chosen for the wand condition, as well as, for pin manipulation. Snapping to grid, for example, would change task completion times for this study dramatically. Also, the mesh modification in this case was limited to a 2.5D mesh, constraining vertices' x and z movement. Other interaction techniques would have to be developed to allow a 2.5D shape display to manipulate a 3D mesh, and the wand input clearly has more degrees of freedom, which can easily be mapped to that interaction. However, we believe that there are many new interaction techniques to be developed for shape display interfaces and new hardware configurations that can improve their performance.

## FUTURE WORK
The user evaluation and our analysis of the Sublimate system, point towards numerous interesting challenges to explore in the future – primarily related to hardware technology and interaction techniques.

The current Sublimate system relies on a 2.5D actuated shape display to render the physical objects. Current 2.5D actuated

shape displays have limited spatial resolution, haptic resolution, refresh rate and degrees of freedom, in comparison to other haptic input devices [18, 32]. While MEMS and Soft Robotics will likely play an important role in addressing the scalability issues for actuated pin displays, current interactive systems are limited by their use of at least one actuator per pin. Overcoming the constrained degrees of freedom is a larger challenge; other form factors beyond pin displays, such as modular robotics, could help overcome these issues.

The Sublimate system is also limited by its display capabilities. Ideally, a true volumetric display would be used as opposed to a single-user stereoscopic 3D display with view-dependent rendering. While volumetric displays do not typically allow direct hand interaction in the volume, other optical configurations, such as the one used by Vermeer [6], or by mounting the volumetric display above the beam splitter would allow co-located 3D virtual graphics without requiring view-dependent rendering or stereo glasses. This type of system would also allow for multiple users to view and interact with Sublimate without the need for a handheld AR tablet or HMDs.

Our current implementation relies on a motion capture system to track head position and user input through a wand or glove. Depth cameras could be an interesting alternative as they would enable the tracking of freehand input and potentially provide for denser surface geometry, as opposed to the current marker-based tracking [17, 12].

In our future work, we would like to explore implementing Sublimate interactions with other actuated tangible interfaces and shape displays beyond Relief. We believe that the principles outlined in the Sublimate concept can extend easily to other hardware platforms and provide a basis for work with Spatial AR and the type of physical actuation described by the vision of Radical Atoms [14]. In addition, we are also planning a broader exploration of interaction techniques that leverage the transitions between virtual and physical rendering. In particular, we see interesting potential in applying our concept of dynamic physical affordances to a wide range of user interaction scenarios.

## CONCLUSIONS
We have presented Sublimate, our vision of how 3D spatial graphics and physical shape output can be combined, and we highlight the potential in computational transitions between these states. We described two different implementations of the Sublimate concept. Our single-user system has a spatial optical see-through display for co-located high-resolution graphics and shape output, while our multi-user system employs handheld tablet-based AR. Through demonstration applications in the domains of CAD, medical imaging and geospatial data exploration, we have shown how Sublimate can provide novel interactions for 3D data, allow for switchable control between precise physical manipulation and mid-air gesture, provide physical affordances on demand, and extend the shape display's resolution and scale. A formal user evaluation showed that bimanual interaction with spatial 3D graphics through the shape display can outperform mid-air interaction with a wand. We believe that the intersection

between physical shape output and spatial graphics is a rich area of exploration, and that the state transitions described here can be a valuable avenue for further investigation.

## ACKNOWLEDGMENTS

## REFERENCES

1. Aksak, B., Bhat, P. S., Campbell, J., DeRosa, M., Funiak, S., Gibbons, P. B., Goldstein, S. C., Guestrin, C., Gupta, A., Helfrich, C., Hoburg, J., Kirby, B., Kuffner, J., Lee, P., Mowry, T. C., Pillai, P. S., Ravichandran, R., Rister, B. D., Seshan, S., Sitti, M., and Yu, H. Claytronics: highly scalable communications, sensing, and actuation networks. In *SenSys '05* (2005), 299–299.

2. Anabuki, M., and Ishii, H. Ar-jig: A handheld tangible user interface for modification of 3d digital form via 2d physical curve. In *ISMAR '07* (2007), 1–10.

3. Benko, H., Jota, R., and Wilson, A. Miragetable: freehand interaction on a projected augmented reality tabletop. In *CHI '12* (2012), 199–208.

4. Bimber, O., Frohlich, B., Schmalsteig, D., and Encarnacao, L. The virtual showcase. *Computer Graphics and Applications, IEEE 21*, 6 (nov/dec 2001), 48 –55.

5. Bimber, O., and Raskar, R. *Spatial augmented reality : merging real and virtual worlds*. A K Peters, 2005.

6. Butler, A., Hilliges, O., Izadi, S., Hodges, S., Molyneaux, D., Kim, D., and Kong, D. Vermeer: direct interaction with a 360 deg viewable 3d display. In *UIST '11* (2011), 569–576.

7. Hachet, M., Bossavit, B., Cohé, A., and de la Rivière, J.-B. Toucheo: multitouch and stereo combined in a seamless workspace. In *UIST '11* (2011), 587–592.

8. Harrison, C., Benko, H., and Wilson, A. D. Omnitouch: wearable multitouch interaction everywhere. In *UIST '11* (2011), 441–450.

9. Hart, S. G., and Staveland, L. E. Development of nasa-tlx (task load index): Results of empirical and theoretical research. In *Human Mental Workload*, P. A. Hancock and N. Meshkati, Eds., vol. 52 of *Advances in Psychology*. North-Holland, 1988, 139 – 183.

10. Hashimoto, S., Ishida, A., Inami, M., and Igarashi, T. Touchme: An augmented reality based remote robot manipulation. In *ICAT '11* (2011).

11. Henderson, S. J., and Feiner, S. Opportunistic controls: leveraging natural affordances as tangible user interfaces for augmented reality. In *VRST '08* (2008), 211–218.

12. Hilliges, O., Kim, D., Izadi, S., Weiss, M., and Wilson, A. Holodesk: direct 3d interactions with a situated see-through display. In *CHI '12* (2012), 2421–2430.

13. Hoshi, T., Takahashi, M., Nakatsuma, K., and Shinoda, H. Touchable holography. In *SIGGRAPH '09* (2009), 23:1–23:1.

14. Ishii, H., Lakatos, D., Bonanni, L., and Labrune, J.-B. Radical atoms: beyond tangible bits, toward transformable materials. *Interactions 19*, 1 (Jan. 2012), 38–51.

15. Ishii, K., Zhao, S., Inami, M., Igarashi, T., and Imai, M. Designing laser gesture interface for robot control. In *INTERACT '09* (2009), 479–492.

16. Iwata, H., Yano, H., Nakaizumi, F., and Kawamura, R. Project feelex: adding haptic surface to graphics. In *SIGGRAPH '01* (2001), 469–476.

17. Leithinger, D., Lakatos, D., DeVincenzi, A., Blackshaw, M., and Ishii, H. Direct and gestural interaction with relief: a 2.5d shape display. In *UIST '11* (2011), 541–548.

18. Massie, T., and Salisbury, K. The phantom haptic interface: A device for probing virtual objects. In *ASME Winter Annual Meeting* (1994).

19. Olwal, A., Lindfors, C., Gustafsson, J., Kjellberg, T., and Mattsson, L. Astor: An autostereoscopic optical see-through augmented reality system. In *ISMAR '05* (2005), 24–27.

20. Piper, B., Ratti, C., and Ishii, H. Illuminating clay: a 3-d tangible interface for landscape analysis. In *CHI '02* (2002), 355–362.

21. Plesniak, W., Pappu, R., and Benton, S. Haptic holography: a primitive computational plastic. *Proceedings of the IEEE 91*, 9 (sept. 2003), 1443 – 1456.

22. Poupyrev, I., Nashida, T., and Okabe, M. Actuation and tangible user interfaces: the vaucanson duck, robots, and shape displays. In *TEI '07* (2007), 205–212.

23. Raskar, R., Welch, G., Low, K.-L., and Bandyopadhyay, D. Shader lamps: Animating real objects with image-based illumination. In *Eurographics*, Springer-Verlag (London, UK, UK, 2001), 89–102.

24. Rasmussen, M. K., Pedersen, E. W., Petersen, M. G., and Hornbaek, K. Shape-changing interfaces: a review of the design space and open research questions. In *CHI '12* (2012), 735–744.

25. Ratti, C., Wang, Y., Piper, B., Ishii, H., and Biderman, A. Phoxel-space: an interface for exploring volumetric data with physical voxels. In *DIS '04* (2004), 289–296.

26. Reed, M. Prototyping digital clay as an active material. In *Proceedings of TEI '09* (2009), 339–342.

27. Scharver, C., Evenhouse, R., Johnson, A., and Leigh, J. Designing cranial implants in a haptic augmented reality environment. *Commun. ACM 47*, 8 (Aug. 2004), 32–38.

28. Schmandt, C. Spatial input/display correspondence in a stereoscopic computer graphic work station. In *SIGGRAPH '83* (1983), 253–261.

29. Stevenson, D. R., Smith, K. A., McLaughlin, J. P., Gunn, C. J., Veldkamp, J. P., and Dixon, M. J. Haptic workbench: a multisensory virtual environment. *SPIE 3639, Stereoscopic Displays and Virtual Reality Systems VI* (1999), 356–366.

30. Sutherland, I. The ultimate display. In *International Federation of Information Processing* (1965).

31. Tani, M., Yamaashi, K., Tanikoshi, K., Futakawa, M., and Tanifuji, S. Object-oriented video: interaction with real-world objects through live video. In *CHI '92* (1992), 593–598.

32. Walairacht, S., Yamada, K., Hasegawa, S., Koike, Y., and Sato, M. 4 + 4 fingers manipulating virtual objects in mixed-reality environment. *Presence: Teleoper. Virtual Environ. 11*, 2 (Apr. 2002), 134–143.

33. Yoshida, T., Kamuro, S., Minamizawa, K., Nii, H., and Tachi, S. Repro3d: full-parallax 3d display using retro-reflective projection technology. In *SIGGRAPH '10* (2010), 20:1–20:1.