The Deformable Workspace: *a Membrane between Real and Virtual Space*

Yoshihiro Watanabe, Alvaro Cassinelli, Takashi Komuro, Masatoshi Ishikawa
Graduate School of Information Science and Technology,
University of Tokyo
Yoshihiro_Watanabe@ipc.i.u-tokyo.ac.jp

Abstract

We propose a variant of the multi-touch display technology that introduces an original way of manipulating three-dimensional data. The underlying metaphor is that of a deformable screen that acts as a boundary surface between the real and the virtual worlds. By doing so, the interface can create the illusion of continuity between the user's real space and the virtual three-dimensional space. The prototype system presented here enables this by employing three key technologies: a tangible and deformable projection screen, a real-time three-dimensional sensing mechanism, and an algorithm for dynamic compensation for anamorphic projection. This paper introduces the concept of the deformable tangible workspace, and describes the required technologies for implementing it. Also, several applications developed on a prototype system are detailed and demonstrated.

1. Introduction

We manipulate real three-dimensional (3D) objects easily with our hands, but this is difficult to achieve when the manipulated objects are represented in a virtual space on the screen. However, the importance of such manipulations is increasing with the progress of computers capable of rendering complex scenes in real time. Indeed, when creating a workspace to manipulate 3D data, we are confronted with at least two problems described below.

A typical virtual object manipulation interface (such as CAD system) consists on a fixed flat display and one or more input devices such as a mouse and a keyboard [1, 2]. In this type of configuration, ”touching” virtual objects in virtual space is mediated through those input devices. This constrains users to learn more or less counterintuitive manipulation methods which may be very different from the strategies employed in the real world. As a result, this classical configuration generates a mismatch between the perceived position/orientation of bodies in virtual and in real space - both for the manipulating hand and for the manipulated object. In particular, the virtual object appears separated from the user's hand and its posture may also be rotated or scaled with respect to the hand coordinate system. Users are obliged to bear with this mismatched perception while manipulating objects, and consequently the virtual manipulation requires a more or less conscious effort from the user in order to compensate for these arbitrary transformations. These problems reduce drastically the efficiency of the 3D-data-manipulation task at hand.

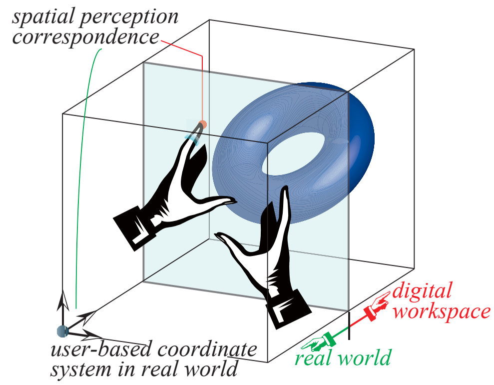


Figure 1. Metaphor diagram

The problem lies in the introduction of input devices that were once conceived mainly for alphanumeric input. In real space, familiar manipulation of 3D objects result from natural body postures and behavior, therefore a properly designed manipulation method should capitalize on this [14, 18]. Such consideration still leaves the problem of the displaying technology: where and how the virtual space should be represented? If we use a fixed, flat display as in present day computer-based workspaces, the mismatch between the perceived manipulated object and the perceived position of the user body (proprioception) is not resolved.

The problem is partially addressed by the introduction of the head-mounted display (HMD) or other 3D vision systems [3, 11, 12]. By wearing such systems, it is possible to reduce the mismatch between the spatial perception in real and virtual worlds. However, this configuration only solves the mismatch in one modality of perception (vision), but does not address the problem of how to reduce the mismatch in another fundamental sensory modality: the sense of touch.

978-1-4244-2897-7/08/$25.00

We would like to introduce here a candidate input/display technology that has the potential to address these problems simultaneously, and if not solve, at least alleviate the perceived mismatch.

The proposed system is inspired from a screen-based interactive-art installation called *The Khronos Projector* [5]. In this installation, the user is able to send parts of the image forward or backwards in time by touching different portions of a deformable projection screen made of spandex. A related strategy developed as part of the *Tangible bits* project [10] relies instead on clay or sand as the projection medium. However, this method introduces several constraints, among which the impossibility to back-project images, which may be indispensable if one wants to implement the real/virtual membrane metaphor.

Therefore, we concentrate on the deformable screen technology for the development of an interface relying on the metaphor of a physical membrane between the real and virtual world. We will call the implemented system *a deformable workspace*.

2. The deformable workspace

This section describes in detail the membrane metaphor and list the set of required technologies to effectively build such a deformable workspace.

Figure 1 illustrates the metaphor of our proposed deformable workspace. The virtual object ”exists” and is represented in virtual space, while the user exists in real space, but is not represented (as a whole or in part) in the virtual space. The idea is to maintain a useful and simple relationship between virtual and real space by using a unique coordinate system that is shared by both spaces. Between these spaces lies a ”transparent” and tangible membrane. Users can manipulate the objects in virtual space by deforming the membrane and observing the effects on the virtual object (much like a surgeon operating on a patient with gloves).

Of course, the kind of haptic feedback provided by the proposed deformable workspace is extremely limited: the membrane alone cannot generate different textures, nor dynamic patterns of pressure relating to virtual object motion. Although we have been considering the use of embedded actuators on the membrane (such as a textile interwoven with shape memory alloys fibers), we think that such complex haptic feedback may not be crucial in most 3D manipulation tasks. In this paper we focus instead on resolving what we think is a more immediate and fundamental problem, i.e., the perceived visual and proprioceptive mismatch between real position of hand and the represented manipulated virtual objects in virtual space. We believe that a workspace based on the metaphor of the membrane between real and virtual space do alleviate this perceived mismatch (without the need of wearable devices nor head mounted displays); this, coupled with the fact that the real deformable screen has a natural (although passive) tactile feedback may facilitate a more natural interaction with virtual objects.

However, no matter how appealing this metaphor may be, the actual usefulness of the proposed interface lies entirely on an adequate choice of the displaying, sensing and processing technologies. For displaying, we considered a translucid deformable screen. This screen serves three functions: displaying, sensing and passive haptic feedback. Users will feel as if they penetrate the virtual space by pushing on this surface. Real-time sensing of the screen deformation is needed in order to generate meaningful input commands without significant latency. Moreover, the displayed image should not be affected by the deformation of the screen surface in a way that destroys the illusion supporting the interaction metaphor. Therefore, the projected image must be pre-warped before projection to compensate for this deformation. Such considerations impose stringent requirements on the sensing, projecting and processing hardware. The prototype presented in this paper meets these technology challenges. The next section describes the technologies used in our prototype.

3. Working principle and setup

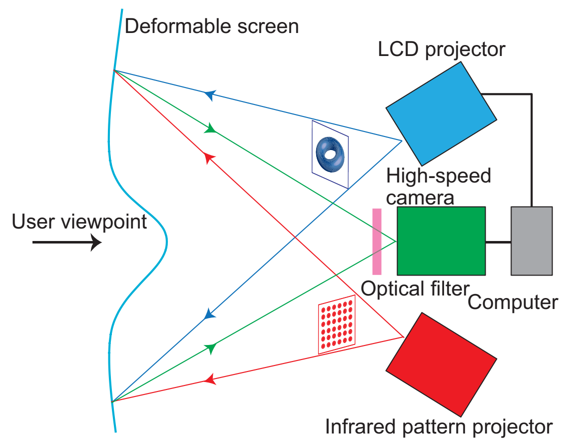


Figure 2. System configuration

This section describes a prototype of the deformable workspace, and details each of its parts. The system is build upon three key technologies, namely a tangible and deformable projection screen, a real-time 3D sensing mechanism, and a dynamic compensation for anamorphic projection.

The setup is illustrated in Figure 2. It comprises a screen, a camera, and two light projectors (one for displaying images and the other to perform active-lighting based surface scanning). The projectors and the camera are set behind the screen. Images are retro-projected on a thin, translucent latex sheet that has good enough optical characteristics (lambertian-type diffusion gives a large viewing angle, and the relatively low transmission gain is not a concern in our prototype).

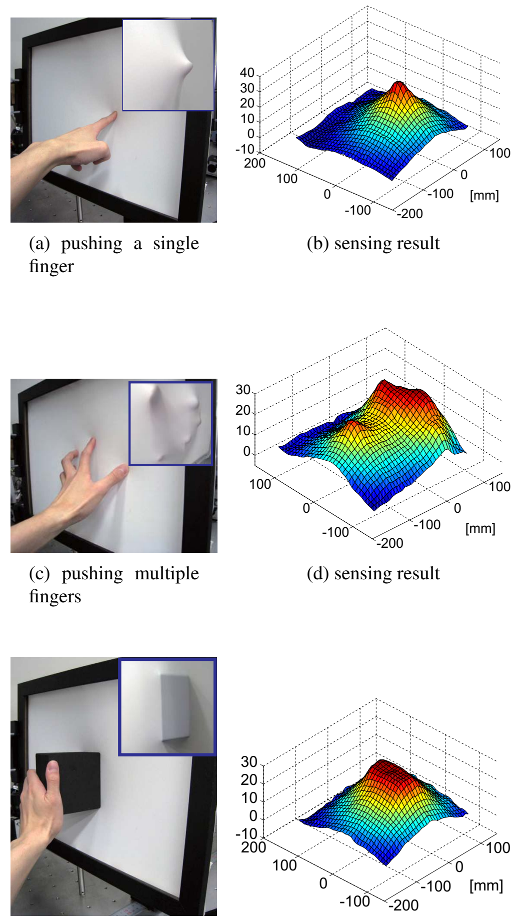
In the remainder of this section, we describe the 3D sensing hardware and the graphic pre-processing for compensating the warping introduced when projecting on a deformable surface.

The sensing mechanism is based on triangulation using structured light. Since at any moment the screen may change its shape, we are obliged to reject 3D scanning methods based on sweeping patterns of light. Indeed, it is required that only a single snapshot of the illuminated surface should suffice to grasp the complete screen deformation. In our prototype, a projected pattern consisting of an array of about 1,100 spots of light constantly illuminates the surface; a high-speed camera acquires a single image, and a dedicated co-processor for high-speed image processing calculates the coordinates of a 3D point for each spot of light in the pattern. This dedicated hardware enables the computation of the deformation in real-time, with a throughput of 955fps and a latency of 4.5ms [20]

In our prototype, the structured-light pattern is projected onto an area of about 30cm $\times$ 30cm using an infrared laser with diffractive optics. An optical filter in front of the high-speed camera sensor enables only infrared light to be detected. This configuration simplifies image processing because the pattern of spots can be easily extracted without additional processing. Note that this infrared light is not visible by the human user, so it does not interfere with the displayed images.

In addition, the throughput of the 3D sensing hardware is much higher than the display frame rate. This enables very low interaction latency, but also reduction of measurement errors (by averaging measures), as well as other interesting possibilities such as real-time computation of the speed of the deformation (which can be used as interaction input). Examples of measured deformations are shown in Figure 3.

Without compensation, the deformation of the screen causes a distortion of the projected image. In order to compensate for this distortion, it is of course necessary to know the shape of the deformed screen, but also the projector calibration parameters. Calibrating the projector means obtaining the intrinsic parameter including focal length, etc. and the extrinsic parameters including its position and orientation. There are various calibration methods for cameras [8], and those can be used in case of a projector calibration as well. In this setup, we used Faugeras's method [6]. This method finds parameters accounting for a simple perspective projection. Though there are other calibration methods which can account for a more refined model of the projector (including optical aberrations), we employed this method in order to simplify the computations required to generate the projected image. In order to compensate the distortion, observation model in human vision is also required. In this setup, we assumed that the model is solely based on perspective projection. This can also simplify the computation in the graphic processing pipeline.



(a) pushing a single finger

(b) sensing result

(c) pushing multiple fingers

(d) sensing result

(e) pushing a box

(f) sensing result

Figure 3. Examples of 3D sensing

Figure 4 shows a projected ray as observed by the user eye on the deformable screen. Consider the light ray projected from the position $\boldsymbol{p}_a$ in the projection plane; when the screen is plane, it will be observed at the position $\boldsymbol{p}_d$ in the eye-image plane. However, when the screen is deformed as shown in the illustration, it will be observed at the position $\boldsymbol{p}_c$. Therefore, if the system does not modify the projected image, the observed image will be warped according to the screen deformation. To compensate for this, one needs to properly and dynamically pre-warp the image by shifting each projected point. In case of a deformed screen, the observed light at $\boldsymbol{p}_d$ is projected from $\boldsymbol{p}_b$ and the image point needs to be translated to this position. Based on these simple geometrical considerations, it is possibility to always provide a seemingly undeformed image to the user.

To achieve this dynamic pre-warping, the first required

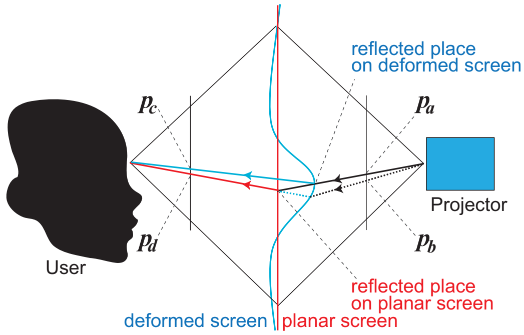


Figure 4. Compensation of image warp caused by screen deformation

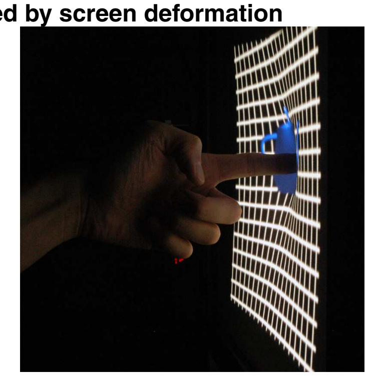


Figure 5. Side view of the user manipulation

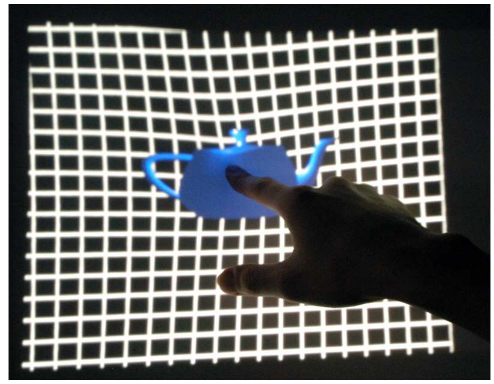


Figure 6. The case where compensation was not applied

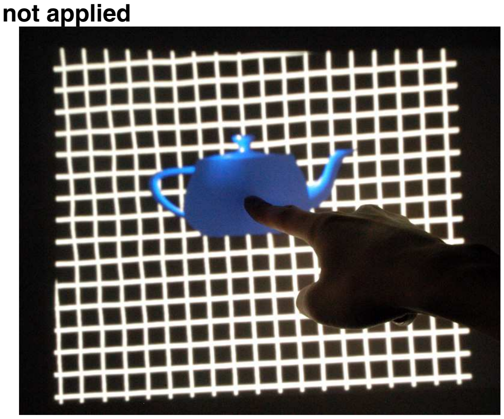


Figure 7. The case where compensation was applied

calculation is to solve the following simultaneous equations for a 3D position (X, Y, Z):

$$\omega \begin{bmatrix} x_e \\ y_e \end{bmatrix} = \boldsymbol{P_e} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \quad (1)$$

$$f(X, Y, Z) = 0 \quad (2)$$

The first equation expresses the perspective projection model. (x_e, y_e) are the observed position projected to human eye, (X, Y, Z) are the 3D position on the screen where the observed point is reflected, $\boldsymbol{P_e}$ is the perspective projection matrix in the model of human vision, and ω is an unknown parameter accounting for the distance from the user. The second equation expresses the deformation of the surface. After this calculation, the position (x_p, y_p) on the projection image from which a light ray reaches to the screen position (X, Y, Z) is obtained based on the calibrated model for the projector. This calculation is also based on the similar equation as shown in Eq. (1).

Using this, the warped image to be projected onto the deformed screen can be calculated. Calculations and rendering are both easily and efficiently implemented on a computer with hardware accelerated graphics using a technique called projective texture mapping. In related works [15], this technique was used for projecting into static non-planar surfaces; in our prototype, the projected image is warped in real-time according to a dynamically deforming screen thanks to high-speed 3D sensing hardware.

Figures 5 and 6 and 7 show the results of the real-time deformation compensation. The user perceives an unwarped image on the screen, despite it being deformed by applied forces. This reinforces the impression that the screen is a transparent membrane between real and virtual space.

A picture of the prototype setup is shown in Figure 8. This setup fulfills most qualitative and quantitative requirements to qualify as a ”deformable workspace” as described in the introduction, thus bringing us close to a real implementation of the membrane metaphor.

4. Demonstrations

The deformable workspace is conceived for use in applications involving manipulation of virtual 3D objects. In this section, some simple examples are considered and demon-

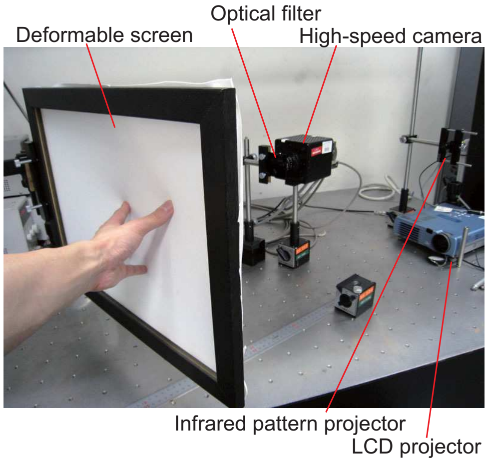


Figure 8. Picture of prototype

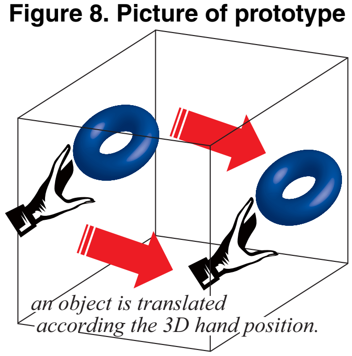


Figure 9. 3D translations on the deformable workspace

strated. All of them exploit the metaphor of a user separated from the virtual world through a transparent deformable membrane, but nonetheless able to touch 3D virtual objects lying on the other side.

4.1. Executing 3D translations

The first demonstrated task consists on simple translating virtual objects both transversally and along the axial dimension (i.e. along the shared axis between real and virtual space). The latter is something that cannot be easily achieved on conventional 3D workspace systems. The conceptual diagram of this task is illustrated in Figure 9. In this demonstration, the position of user's finger is estimated from the deformation of the screen, and the virtual object is moved accordingly. Consecutive snapshots taken during the task are shown in Figure 10. In the center of the image, the axis of the coordinate system are drawn. A torus-like object lying on an horizontal plane (represented by white lines) represents the manipulated target. A red line connects the origin of the coordinate system with the center of the object

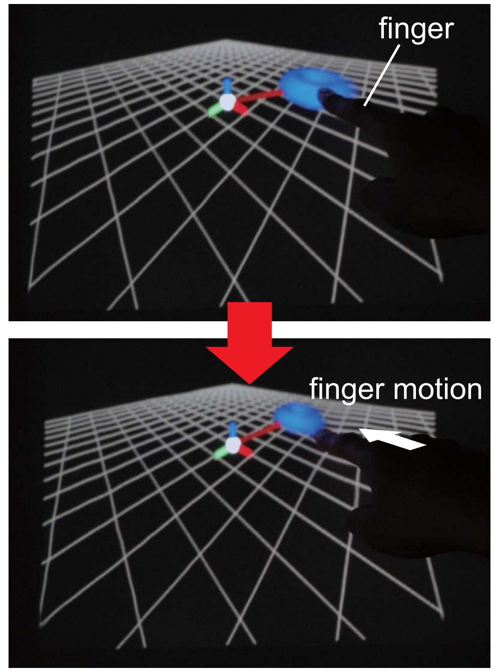


Figure 10. Demonstration of 3D translations

for clarity. The pictures show how the object is translated accordingly with the pressure of the user's finger.

Other strategies can be conceived in order to realize this type of task; for example, the global force distribution on the screen can be used to affect the position of the virtual object in many different ways. However, a clear integration of sensory-motor and spatial perception is realized in the proposed implementation: sensed pressure naturally suggests the object resistance, and axial displacement is the natural consequence of this action. It is interesting to note that axial shifting of virtual objects has been previously described on tangible screens [19], however the screen itself is rigid and the user input is limited to "taps" on the displayed objects, irrespectively of their virtual depth.

4.2. Multi-touch 3D manipulation

Single touch limits the number of degrees of freedom achievable during manipulation. For example, it is difficult with the single-touch configuration to control rotation and translation simultaneously. Multi-touch-based tabletop systems such as [7, 16] provide interesting manipulation strategies for rotations and translations, but when these are not limited to the screen plane, for instance the rotation whose axis is not perpendicular to the screen plane, the command gestures may not be completely intuitive.

This section describes 3D rotation and shifting with multi-touch interaction. We consider here the manipulation

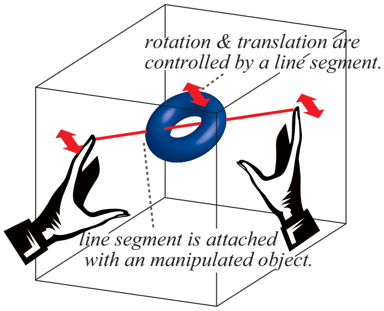


Figure 11. Multi-touch 3D manipulation on the deformable workspace

Figure 12. Demonstration of multi-touch 3D manipulation

of a virtual object thanks to two-points of pressure on the deformable screen - although a larger number of simultaneous pressure sites is possible. The conceptual diagram is shown in Figure 11. As shown in this figure, the orientation and the position of the manipulated object is controlled by shifting and rotating a line segment connecting a pair of pressure centers (generated by fingers on the same or on different hands). A snapshot of the actual task is shown in Figure 12.

4.3. 3D line drawing

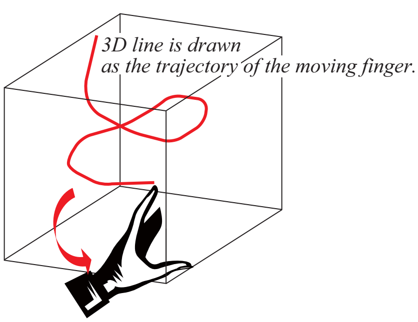


Figure 13. 3D freehand drawing on the deformable workspace

The next demonstration is free-hand drawing in 3D space. The conceptual diagram for this task is illustrated in Figure 13. In this demonstration, the trajectory of the user's fingertip is tracked and rendered as a 3D line in virtual space.

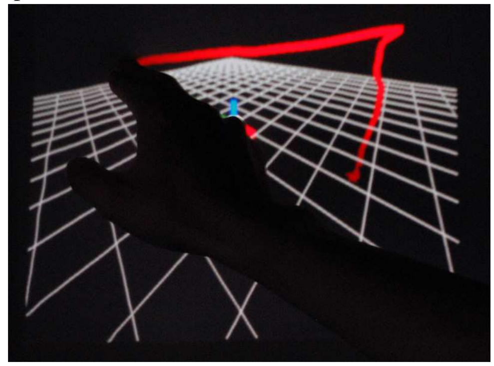


Figure 14. Demonstration of 3D line drawing

Figure 14 shows an actual snapshot of the demonstrated task. Because of the rendered perspective projection, the width of the 3D line changes with depth, providing an additional cue that enhances the drawing task in virtual 3D space.

4.4. 3D surface sculpture

Simple 3D modeling is also possible with the deformable workspace. The user can generate arbitrarily shaped solids by taking advantage of the real deformation of the screen. In this demonstration, as shown in Figure 15, the deformation of the screen is mapped onto a virtual solid surface. The generated solid is rotated and tilted for visibility, but the object could be also rotated in order to sculpt more complex shapes, very much like a potter's wheel. The present demonstration also includes a physical simulation of a multi-particle system evolving over the generated solid surface.

For this type of tasks, it is imperative to generate a realistic perception of space, as well as take into account the body posture and behavior as done in other works [3, 4, 11, 12, 18]. However, the deformable workspace also introduces haptic feedback, which may be passive and elementary, but can be very important when considering manipulation of virtual object with *real* hands.

4.5. Volume Slicing

The final demonstration is virtual *volume slicing*. This is an application displaying an arbitrary cross section of a virtual 3D object. Related works are described in [9, 13]. This application has a great potential in architectural, medical and scientific data visualization, because it allows the

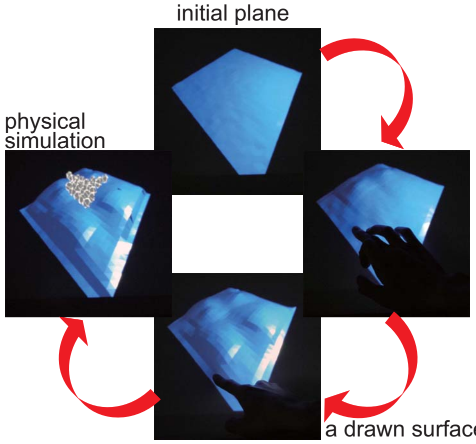


Figure 15. Demonstration of 3D Sculpture

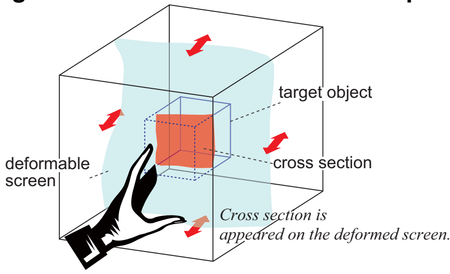
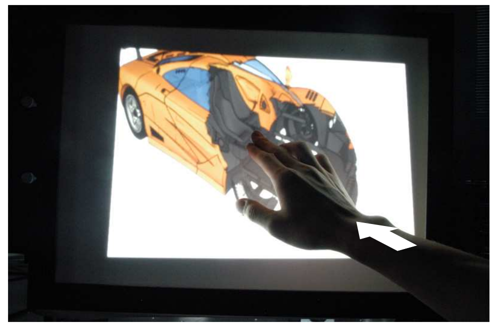


Figure 16. Display of arbitrary volume slicing

(a) Rotating the screen



(b) Deform the screen

Figure 17. Demonstration of volume slicing

user to observe inner structures interactively and intuitively (pushing will break or open a hole in the forefront image layer, revealing hidden image layers). In this case, the deformable screen functions as an arbitrarily shaped ”slicing surface”. Moreover, in order to get more depth range, the whole screen can be moved freely within a certain volume.

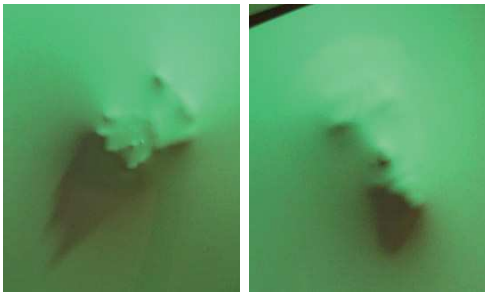
Pictures of the volume slicing application at work are shown in Figure 17. Figure 17(a) was captured when the screen was not deformed, only tilted. The frontal section of the vehicle is cut, and the underlying inner structures are exposed. The snapshot shown in Figure 17(b) was captured when pressure is applied to the center of the screen. This time, the slicing section is large and arbitrarily shaped, and we can see the seats and other parts inside the car.

5. Conclusion

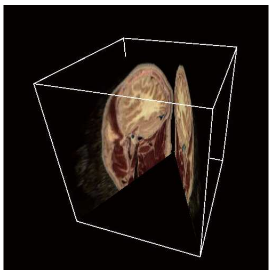
This paper proposes a new paradigm for manipulating 3D virtual objects, namely the deformable workspace. The purpose is to facilitate easy manipulation of 3D virtual objects by keeping a consistent relationship between the real and the virtual space. The efficiency of the interaction method derives from a powerful underlying metaphor: that of a transparent membrane that acts as an invisible, yet tangible boundary between real and virtual spaces. The deformable workspace is expected to solve - at least in some cases - most of the problems afflicting present-day virtual manipulation interfaces, such as the perceived mismatch between spatial and virtual space - and this without having to resort to head-mounted displays or other 3D visualization systems.

The prototype demonstrated here was built upon three key technologies: a tangible and deformable projection screen, a real-time 3D sensing mechanism, and an algorithm and hardware for dynamic compensation for anamorphic projection. The introduced 3D sensing mechanism obtains the shape of the screen at high throughput and with minimal latency. Such performances allows instant feedback for virtual manipulation, but also pre-warping compensation, reduction of measuring noise, and detection of user’s high-speed motion. We successfully demonstrated several applications including dynamic projection onto the deformable surface, 3D translation, 3D manipulation by two hands, 3D freehand drawing, 3D sculpture, and arbitrary volume slicing.

In the future, we plan to address the problems of detecting and interpreting complex body postures and behavior based solely on the deformation of the screen, that can be used to control more complex object manipulations tasks in virtual space. This is possible because the proposed 3D sensing mechanism enables the extraction of complex deformation patterns (for instance, the detailed shape and im-



(a) picture of intricately deformed screen



(b) example simulation for volume slicing. Here separated two planes are set and cross section is projected.

Figure 18. Examples of complex deformed screen

print of closed or open hands, angles formed by forearms or shapes of other real objects that could be placed over an *horizontal* workspace). The example deformation and simulation of the volume slicing is shown in Figure 18.

Another feature we plan to develop is viewpoint-based projection. The present prototype assumes the position of the user's eyes to be fixed, which is an unrealistic assumption. In the future, the virtual space representation must be rotated and transformed according to the change of viewpoint (as done in the Parallax Augmented Desktop also developed in our lab [17]). This may contribute significantly to the perceived consistence between the real and virtual spaces. In addition, the haptic extension functions described in section 2 would be explored in the next step.

References

[1] Autodesk 3ds Max, http://usa.autodesk.com/.

[2] LightWave, http://www.newtek.com/lightwave/.

[3] blui: Body Language User Interface, http://www.blui.org/.

[4] BlueGrotto, http://sagaweb.csse.muroran-it.ac.jp/English/.

[5] A. Cassinelli and M. Ishikawa. Khronos projector. In *SIGGRAPH 2005, Emerging Technologies*, 2005.

[6] O. Faugeras. *Three-Dimensional Computer Vision: A Geometric Viewpoint (Artificial Intelligence).* MIT Press, 1993.

[7] J. Y. Han. Low-cost multi-touch sensing through frustrated total internal reflection. In *Proceedings of Symposium on User Interface Software and Technology*, pages 115–118, 2005.

[8] E. E. Hemayed. A survey of camera self-calibration. In *Proceedings of 2003 IEEE Conference on Advanced Video and Signal Based Surveillance.*, pages 351–357, 2003.

[9] K. Hirota and Y. Saeki. Cross-section projector: Interactive and intuitive presentation of 3D volume data using a handheld screen. In *Proceedings of 2007 IEEE Symposium on 3D User Interfaces*, pages 55–61, 2007.

[10] H. Ishii, C. Ratti, B. Piper, Y. Wang, A. Biderman, and E. Ben-Joseph. Bringing clay and sand into digital design - continuous tangible user interfaces. *BT Technology Journal*, 22(4):287–299, 2004.

[11] K. Kiyokawa, H. Takemura, and N. Yokoya. Seamlessdesign for 3D object creation. *IEEE Multimedia*, 1(7):22–23, 2000.

[12] G. A. Lee, G. J. Kim, and C. Park. Modeling virtual object behavior within virtual environment. In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology*, pages 41–48, 2002.

[13] T. Nagakura and J. Oishi. Deskrama. In *SIGGRAPH 2006, Emerging technologies*, 2006.

[14] H. Nishino, D. Nariman, and K. Utsumiya. Making 3D objects through bimanual actions. In *Proceedings of 1998 IEEE International Conference on Systems, Man, and Cybernetics*, pages 3590–3595, 1998.

[15] R. Raskar, G. Welch, M. Cutts, A. Lake, L. Stesin, and H. Fuchs. The office of the future: A unified approach to image-based modeling and spatially immersive displays. In *Proceedings of SIGGRAPH 1998*, pages 179–188, 1998.

[16] J. Rekimoto. Smartskin: An infrastructure for freehand manipulation on interactive surfaces. In *Proceedings of Conference on Human Factors in Computing Systems (CHI)*, pages 113–120, 2002.

[17] C. Reynolds, A. Cassinelli, and M. Ishikawa. Meta-perception: reflexes and bodies as part of the interface. In *Proceedings of Conference on Human Factors in Computing Systems*, pages 3669–3674, 2008.

[18] S. Schkolne, M. Pruett, and P. Schroder. Surface drawing: creating organic 3D shapes with the hand and tangible tools. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 261–268, 2001.

[19] Z. B. Simpson. Novel infrared touch-screen technology and associated artwork. In *SIGGRAPH 2004, Emerging technologies*, 2004.

[20] Y. Watanabe, T. Komuro, and M. Ishikawa. 955-fps real-time shape measurement of a moving/deforming object using high-speed vision for numerous-point analysis. In *Proceedings of 2007 IEEE International Conference on Robotics and Automation*, pages 3192–3197, 2007.