

# TZee: Exploiting the Lighting Properties of Multi-touch Tabletops for Tangible 3D Interactions

Cary Williams<sup>1</sup>, Xing-Dong Yang<sup>3</sup>, Grant Partridge<sup>1</sup>, Joshua Usiskin-Miller<sup>1</sup>,  
Arkady Major<sup>2</sup>, Pourang Irani<sup>1</sup>

<sup>1</sup>Dept. of Computer Science  
University of Manitoba, Canada  
{umwill22,umpartr3,ummillaj, irani}@cc.umanitoba.ca

<sup>2</sup>Dept. of Electrical and Computer Engineering  
University of Manitoba, Canada  
a\_major@umanitoba.ca

<sup>3</sup>Dept. of Computing Science  
University of Alberta, Canada  
xingdong@cs.ualberta.ca

## ABSTRACT

Manipulating 3D objects on a tabletop is inherently problematic. Tabletops lack a third degree of freedom and thus require novel solutions to support even the simplest 3D manipulations. Our solution is *TZee* – a passive tangible widget that enables natural interactions with 3D objects by exploiting the lighting properties of diffuse illumination (DI) multi-touch tabletops. *TZee* is assembled from stacked layers of acrylic glass to extend the tabletop’s infrared light slightly above the surface without supplemental power. With *TZee*, users can intuitively scale, translate and rotate objects in all three dimensions, and also perform more sophisticated gestures, like “slicing” a volumetric object, that have not been possible with existing tabletop interaction schemes. *TZee* is built with affordable and accessible materials, and one tabletop surface can easily support multiple *TZees*. Moreover, since *TZee* is transparent, there are numerous possibilities to augment interactions with feedback, helpful hints, or other visual enhancements. We discuss several important design considerations and demonstrate the value of *TZee* with several applications.

## Author Keywords

Input device, TUI, 3D UI, multi-touch surface.

## ACM Classification

H5.2 [Information interfaces and presentation]: User Interfaces. Input devices and strategies.

## General terms

Design.

## INTRODUCTION

Tabletops are ideal devices for direct interaction with 2D virtual objects, such as digital pictures and documents. Interactive surfaces provide easy access to any two dimensions, much like conventional graphical user interfaces or paper-based workspaces. The introduction of a third dimension, however, raises a number of interesting problems [11]. For example, common interaction techniques, like pinching and

flicking, cannot be easily extended to support free three-dimensional manipulations.

Researchers have recently gained interest in displaying and providing interactions with 3D content on tabletop surfaces [5, 11, 13, 26, 30]. The unique affordances of tabletop computers have led to a number of interesting approaches for the manipulation of 3D objects. These include the use of pressure [5], multiple fingers [11], or layers above the surface [26, 13] to interact with the third dimension. These techniques have their own disadvantages, however; for instance, in the case of multiple fingers [5] or specialized hardware [26], users must learn a special gestural syntax.

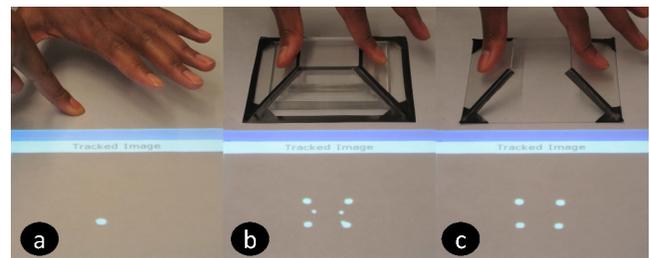


Figure 1 – *TZee* uses the lighting properties of DI tabletops to map gestures on its tilted faces to three-dimensional interactions. (a) Fingers above the surface that could not be detected by the camera are (b) now visible when placed upon *TZee*. (c) Without the acrylic kernel, touch cannot be detected.

To address these weaknesses, we introduce *TZee*, a tangible widget that benefits equally from the characteristics of common touch tabletops and natural interaction strategies. *TZee* – which stands for *tangible z-axis*, referring to the interactions it supports – is a clear, palm-sized block in the shape of a truncated pyramid, or *frustum*. When placed upon a diffused illumination (DI) surface, the widget funnels some light to its five faces without additional power. *TZee* can be manipulated with one hand, leaving the other to create mixed bi-manual interactions on the tabletop surface. *TZee* is simple to create and offers capabilities previously restricted to powered devices like *CubTile* [24]. Since it is easy to build and requires no wires or power source, *TZee* is a cost-effective method for interacting with 3D objects on tabletops that scales to accommodate multiple users. Moreover, it is compatible with common tabletop hardware. This makes it an extremely useful and accessible tool for designers to work with 3D content.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

CHI 2011, May 7–12, 2011, Vancouver, BC, Canada.

Copyright 2011 ACM 978-1-4503-0267-8/11/05...\$10.00.

Our contributions in this paper are as follows. First, we introduce TZee, a novel passive tangible device that allows manipulations of objects in 3D. We outline a set of gestures that facilitate both common and complicated 3D interactions. Finally, we discuss several important factors in the design of TZee, and based on its limitations, we outline some promising approaches to improve this class of device.

### RELATED WORK

Our work builds upon a multitude of prior results for interacting with 3D objects. We review related literature on 3D manipulation with 2D devices, above the surface interactions, and tangible interfaces for 3D control surfaces.

#### Exercising 3D control on 2D interactive surfaces

Researchers have often explored the use of multi-touch systems for interacting with 3D objects [11, 33]. Davidson and Han [5] suggested that motion in the  $z$ -direction could be controlled using pressure. Their approach maps light pressure to lifting the object up and towards the surface and heavy pressure for pushing the object down and away from the surface. Shallow-Depth [10] enables smooth single-finger 2D rotate-and-translate, while 3D operations, such as pitch and roll, can be performed by holding down the first touch and gesturing with a second touch [19]. Hancock et al. [11]’s approach requires users to define a rotational axis using two fingers and then perform the rotation with a third finger. A pinching gesture is commonly proposed [10, 11] for motion along the  $z$ -axis, which is conceptually related to zooming in and out.

A drawback of these sorts of 3D tabletop interaction is that they rely upon explicitly indirect control of at least one dimension (typically the  $z$ -axis). As a result, the techniques cited above can be considered less natural [13], since they do not directly resemble the ways we manipulate objects in the real-world. Additionally, such techniques suffer from ambiguity issues in the interpretation of user input [22].

#### Above the surface interactions

To overcome the aforementioned limitations, developers have combined 3D motion sensing devices with tabletop computers [4, 26]. Cutler et al.’s [4] system accepts users’ hand or finger gestures above the surface via gloves or a stylus equipped with 6DOF sensors. Subramanian et al. [26] suggested dividing the space above the tabletop into horizontal layers and associating each layer with a different set of commands. Benko et al. [2] proposed using muscle sensing techniques to detect mid-air finger gestures. Their system could detect pinch gestures to pick up or drop off a virtual object. However, due to the difficulty of accurately mapping muscle sensing to depth information, complex operations could not be supported in the  $z$ -direction.

Recently, optical depth-sensing techniques have been introduced to resolve some of these concerns [3, 32]. These methods capture 3D hand motions and map them to depth interactions. Hence, 3D interactions on tabletops have become somewhat easier and more natural for users [3, 13]. Nonetheless, optical techniques still suffer from several issues, which

make them unreliable for precise 3D interactions. For instance, current depth-sensing cameras have relatively low sensing resolution, suffer from disruption by finger occlusion, and lack visual and haptic feedback to facilitate 3D interactions. Other problems, such as arm fatigue from constant interaction above the surface, can make such solutions less practical in real-world settings [13].

#### Tangible user interfaces

Tangible user interfaces (TUI) [16] provide intrinsic haptic feedback through a more direct mapping between the user’s input and the digital world. Studies have suggested that TUIs offer improved performance over typical graphical user interfaces for many tasks [15, 12, 6, 27]. Furthermore, using TUIs for 3D interactions can be more intuitive, since they preserve the manner that we manipulate physical objects in the real-world. TUIs have been developed for numerous purposes. For instance, ActiveCubes [18] are cubic TUIs that are used for constructing and interacting with virtual cubes. Magic Cube [35] has a built-in orientation sensor supporting richer interactions with tilt. Graspable UIs [7] are physical gadgets that can be dynamically coupled to virtual objects to serve as physical handles or controls. Many of these TUIs support a few of the most common gestures, but due to limited degrees of input freedom, they lack support for useful operations, including constrained stretching along a fixed axis.

Due to their particular benefits, TUIs have recently appeared as supplementary devices for tabletop computers. Fiebrink et al. [6] augmented a tabletop with physical controls to support precise parameter adjustment. Other developments include using TUIs on a tabletop to control robots [9], to work as musical instruments [17], to augment graphical information displays [23], and to represent parameters inside a software application [21]. Many TUIs need to be powered through a tether or batteries. This can limit the size and the number of the TUIs that can be used on a tabletop, and also increases the development and maintenance costs, making them less accessible for general use.

Because of these reasons, there has been increased attention given to developing unpowered TUIs for tabletop interactions. To date, most of these are built upon computer vision technology. Gallant et al. [8]’s foldable user interface is a flexible physical widget made of construction paper augmented with IR reflectors, the shape of which can be tracked by a video camera. Bending the widget into different shapes manipulates objects in the virtual world. PhotoelasticTouch [25] supports interaction with virtual objects via deformation of its transparent, elastic body. SLAP widgets [31] are physical sliders, buttons and knobs with reflective markers on their undersides that can be seen by the same camera that detects touches upon the table’s surface.

TZee is primarily inspired by two recent TUI developments: Lumino [1] and CubTile [24]. Lumino are building blocks made of glass fiber bundles that pass light from the table to the block’s top and vice versa. This allows several blocks to

be stacked while allowing the system to detect the blocks on top. Lumino was primarily designed for detecting stacked objects. It was not designed for 3D interactions, and since the sides of a Lumino are not sensitive to touch input, it offers no support for  $z$ -axis transformations [1].

CubTile, on the other hand, was designed especially for 3D interactions on large vertical displays. CubTile is a cubic device with five multi-touch faces (all but the bottom). Its four vertical surfaces enable easy motion in the  $z$ -direction. However, it is less than ideal for typical tabletop scenarios. First and foremost, CubTile requires a power source. Also, its size restricts portability and the number of devices that can be used simultaneously in one application. In addition, CubTile requires users to use both hands in order to interact on two faces of the device, which limits its usability. TZee borrows the best of both worlds by enabling a cost-effective solution with an unpowered (passive) tangible interface for multi-touch 3D interaction. TZee blocks can be created in a few hours and operate on DI tabletops without modifying any underlying hardware.

### TZEE

In this section we describe how the design of TZee evolved and explain how it operates. We also demonstrate how to construct a TZee with easily accessible materials.

#### Design of TZee

We designed TZee with three key ideas in mind. First, we decided that one-handed operation is important to keep the second hand free to perform other gestures. We also strived to mimic real-world interactions. Most importantly, the device had to effectively channel light from the table.

*One- vs. two-handed constraints.* Each face of TZee is a multi-touch surface. Therefore, multi-finger interactions can be easily carried out with one or both hands. Two-handed input provides more contact points, providing the potential to support richer interactions. For instance, a natural way to stretch an object is to use both hands to pull in opposite directions, which works well with TZee (see Figure 11). However, we wanted to make sure users could perform the most common gestures with one hand.

*Shape.* Candidate shapes for our tangible widget include a cylinder, a cube, a sphere or a pyramid, since objects of these shapes have vertical or inclined surfaces that can be easily distinguished from input on the  $x$ - $y$  plane. A cube has four such faces at right angles, clearly defining  $x$  and  $y$  axes. A square pyramid has the same properties. This arrangement of symmetric opposing faces also enables useful gestures like grasping (Figure 9). To facilitate input in the  $x$ - $y$  plane, we truncated the pyramid, removing its tip and leaving a flat top. This also has the advantage of providing a convenient surface upon which images can be displayed through the device from the display below (discussed further in the section below labeled *Widget as display*).

To acquire rapid feedback on the ergonomics of the two shapes, we built a preliminary prototype for both design op-

tions. The cube was made by gluing 5 USB touchpads (2 × 2.5 inches each) together on a hard base panel (Figure 2 left). The touchpads were connected to a desktop computer via a USB hub and allowed us to test some basic interaction techniques on each face of the cube. The result was a device much like the CubTile [24], though at a size appropriate for tabletop use. Each of the four inclined faces of the pyramidal device were shaped like a trapezoid. We made the latter prototype 4.5×4.5×1.25 inches in size using a polyurethane mixture (Figure 2 right). At this early stage we were only interested in whether these shapes support easy and comfortable gesturing by using one and two hands.

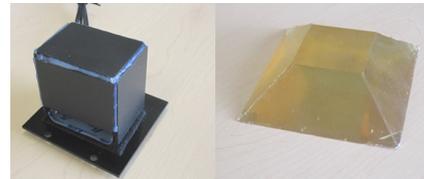


Figure 2 – Left: cube made from touchpads. Right: polyurethane pyramid allowing some light to pass through the object.

After five people evaluated both prototypes, we found that gestures can be easily carried out on both the cube and pyramid with two hands. However, when used with one hand, the pyramid was easier to use than the cube. This effect was particularly obvious with gestures that required two fingers to move on the two opposite faces (Figure 3). Since these gestures require users to apply pressure on the two symmetric faces of the cube, this would accidentally cause the cube to be rotated or otherwise displaced from its resting position. Gesturing upon opposing faces of the pyramid, on the other hand, transferred force downward and prevented accidental motion. Certain gestures that were tricky with the cube like sliding two fingers up and down in the  $z$  direction (Figure 3 right), were easier with the pyramid because of its slope. Overall, our observations suggested that a pyramid offered both better ergonomics, and was thus selected for the physical shape of our device.

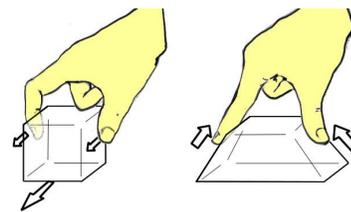


Figure 2 – Left: carrying out gestures on two parallel faces often accidentally moves the cube. Right: the tilted faces make it easy to perform a pinch gesture. Force applied to the tilted faces keeps the pyramid in place upon the tabletop.

#### Lighting optics and construction

Our choice of shape also simplified the device's optics. Recognition of finger gestures on TZee relies on infrared light being transferred through the tabletop surface and into the widget. A passive cube TUI would require a solution to redirect light to its vertical faces. A pyramid differs because the tabletop's camera can directly detect touches upon the device's angled faces.

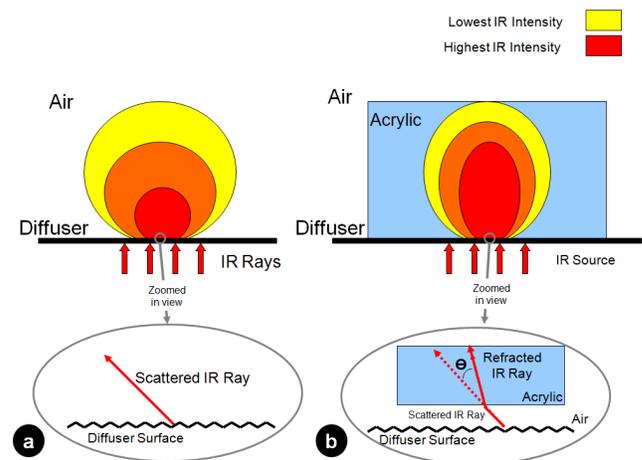
For this purpose, we required a system based on Diffused Illumination (DI) similar to [20]. Much like Lumino [1], TZee allows the image of objects above the surface to be clearly seen by the table's built-in camera. To accomplish this, it must be made from an appropriate material to minimize light loss in transmission. One of our main challenges was finding the most suitable material to construct TZee to meet our requirements. In short, our mission was to combine a suitable material with an appropriate design to reliably transmit finger motions above the surface to the tabletop's built-in camera. Loss of signal during a gesture may cause undesired behaviors (e.g. inadvertently finishing a gesture), impairing the system's performance.

*First Attempt: fiber bundles.* Our initial solution was based on a fiber optic bundle as proposed in [1]. Fiber bundles can fit in blocks of different shapes. In addition, the bundle can be banded or cut into different angles to accommodate the slope of TZee's tilted faces. However, due to the round cross-section of optical fibers, it is difficult to bind them seamlessly, as this creates tiny gaps between the thin fibers. These gaps impair the resolution of gesture images. In a preliminary system test, we were not able to reliably transmit finger touches using a 1.2×1.2×1.2 inch bundle. This implied that if we chose to use fiber bundles, large parts of TZee's surface, especially those parts closer to the edges and the corners, could not detect gestures accurately.

*Second attempt: solid silicone block.* We built our second prototype by pouring a silicon-based mixture – polydimethylsiloxane, or PDMS – into a frustum-shaped mould. PDMS is transparent, light and flexible. The PDMS prototype was created in an air-tight chamber to avoid the intrusion of air bubbles upon solidification that could result in additional loss of light via scattering. Interestingly, upon testing this prototype by shining a red laser pointer beneath it, we found that much of the light was reflected back toward the light source rather than passing through the widget, owing to the phenomenon of total internal reflection (TIR). As a result, finger touches upon the angled faces of the device could not be registered. This encouraged us to further investigate the physical and material properties demanded by our objectives, leading to the third prototype.

*TZee: stacked acrylic.* Enhanced light transfer is possible using a piece of acrylic, a light-weight transparent material that is often used in DI and frustrated total internal reflection (FTIR) tabletop systems. In our experiments we found that the intensity of light at a fixed height above the tabletop surface could be increased by simply placing a piece of flat acrylic panel on the tabletop surface. This effect can be explained by considering the characteristics of the light radiation profile above the tabletop surface and its refraction in materials with relatively high (>1) refractive index (please note that the refractive index of air is ~1). Each point upon the surface of the tabletop (at the interface defined by the diffuser placed atop it) acts like a light source with a Lambertian radiation profile [29]. Figure 4.a schematically represents each point upon the table and, as defined by Lam-

bert's cosine law, the spherical pattern that describes the light intensity at each angle. As can be seen, the intensity of light is gradually reduced above the table and at a certain height the image of the object above the surface will be blurred and eventually become undetectable. Placing an acrylic panel on the tabletop surface modifies the Lambertian profile, reducing the spread inside the acrylic panel and therefore effectively increasing the transmitted light intensity versus an equivalent amount of air. This variation in transmissibility is a result of light refraction at the surface of the acrylic. When passing through two media of different refractive indices, light refracts at the interface between them according to Snell's law [29] (see Figure 4.b). The ray in the higher-index material (in our case acrylic with ~1.49) is closer to the normal of the interface. This translates into a more narrow radiation profile, corresponding with an increase of the transmitted light intensity. Therefore, for a given intensity, light can be brought to a slightly higher position above the surface of the DI table using this simple mechanism (Figure 4.b).



**Figure 4 – Left: at each point upon the tabletop surface where the light reaches the diffuser, the light is dispersed according to a Lambertian profile. The microscopic view beneath shows one ray among countless others being scattered. Right: when a piece of acrylic is placed on top of the diffuser, the rays deviate due to refraction. This alters the Lambertian profile, causing it to spread less within the acrylic panel and increasing the light intensity at a certain height above the table.**

Upon light's exit from the acrylic block, it resumes its original propagation direction and continues to spread in a Lambertian fashion. Due to refraction inside the acrylic panel, however, the spot size at which light emerges from the panel will be smaller than if the light had travelled to the same height solely through air. The smaller spot size corresponds to a higher intensity of light (i.e. less spread). It is worth mentioning that since the light intensity at the top of the acrylic panel is higher, a higher intensity of light reflecting off the object will be captured by the IR camera. The resulting light pattern inside the acrylic panel(s) will be a superposition (via interference) of all individual Lambertian sources created by the diffuser under the acrylic panels. Therefore, the net effect of placing the acrylic panel on the tabletop's

diffuser will be similar, i.e. it will result in overall increase of light intensity above the panel.

To support our claim, we performed two separate tests to detect the increase in light intensity once the acrylic panel is placed on the tabletop. In one case, we used a reflective marker above the table for the sake of consistency. When placed at a fixed height without the acrylic panel, the IR camera registered a bright spot corresponding to the image of our marker. Once the acrylic panel was inserted between the marker and the table, the spot became appreciably larger; this is because more pixels in the processed signal had an intensity exceeding the threshold defining surface activity. This indicates that the camera within the table detects more reflected light from the marker, which implies the effectiveness of the acrylic in preserving light intensity. The degree of change in the blob's size depended on the height of the reflective marker and on the number of acrylic panels, though in all cases, we clearly detected the same effect (Figure 5).

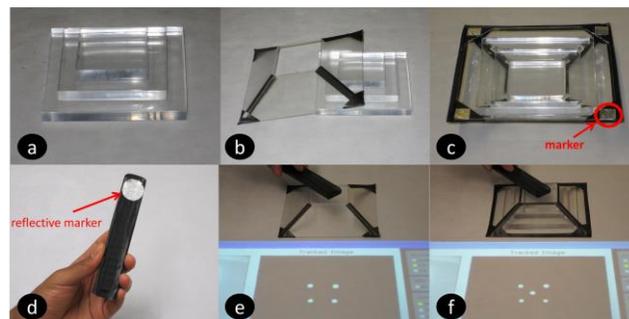
In the second test, we used a silicon (Si) photodiode operating in the photovoltaic regime to quantitatively measure the intensity of light above the table. Similar to the first test, the photodiode was placed at various heights and its signal was registered on an oscilloscope. Note that, when the detector is above the surface, its sensing area is larger than its physical size. Therefore, some light rays from the DI surface are refracted towards the detector while others are refracted outside of it. This is particularly the case for rays from point sources that lie on the periphery of the detector's sensing area. To overcome this unwanted effect we masked a portion of the DI table with a dark tape and left a small opening that was similar in size to the detector's physical size. In this case the detector sees no peripheral light. When the acrylic panel was introduced between the photodiode and the table's diffuser surface, the signal became stronger. For example, with a single panel and a photodiode placed just above it, we detected a 10–15% increase in transmitted light intensity. As expected, intensity gains (versus equivalent amounts of air) diminished as new panels were added. This was caused by spreading of light as well as additional losses introduced by multiple Fresnel reflections at the edges of the panels.

We made our TZee prototype using a stack of acrylic pieces (Figure 5.a), which can be easily found in hardware stores. The prototype is approximately 1.2 inches high, consisting of 3 layers of square acrylic pieces, each 0.4 inches thick (Figure 5, top left). The acrylic pieces measured 4×4 inches, 3×3 inches and 2×2 inches from bottom to top. We polished the cut edges of each acrylic piece. A thin plastic cover (<0.05 inches thick) covers the slope of the steps to provide four tilted surfaces for fingertip contact (Figure 5.b).

We noticed a glowing halo along the edges and joining seams of the device's outer casing, so we covered these joints with black electrical tape. This helped to prevent false blob detections while the user's hand was hovering above the device. The tape also served the purpose of a handle, so that

users could hold onto that part of the device if needed and not have the touch detected.

TZee has four reflective markers upon its bottom corners (Figure 5.c), which allows the tabletop's vision system to detect the region occupied by the tangible block, as well as its orientation. Our system thus interprets the gestures and assigns them appropriately to actions, regardless of TZee's orientation or position on the tabletop. Any actions detected within the boundary of the reflective markers are registered as touches upon TZee.



**Figure 5 – (a) TZee's kernel consists of a stacked pile of 3 acrylic pieces. (b) a thin plastic cover allows fingers to slide upon the stepped sides. (c) reflective markers and their detection allows us to gauge the space for interacting with TZee. (d) a reflective marker for the test. (e) with the reflective marker above the shell without the device's internals, light is not reflected strongly, (f) but when the marker is at the same height above the acrylic blocks, the camera can detect the reflection.**

#### SYSTEM TEST

To assess TZee's ability to transmit signals, we conducted a system evaluation. Since prior work has already demonstrated advantages for tangibles versus virtual controls [6, 12, 15], we were mainly interested in knowing whether continuous gestures could be reliably carried out at all points upon the surface of TZee. We also collected informal feedback from participants on the characteristics of TZee's gestural support. However, additional studies (beyond the scope of this paper) are required to ascertain under which conditions a TUI such as TZee works better than a virtual control for 3D tasks.

#### Apparatus

We used a custom-built diffuse illumination (DI) tabletop system with dimensions 20 in. (length) × 26 in. (width) × 36 in. (height). The diffuser placed upon the tabletop was made from vellum and had a thin silicone coating on the side facing the acrylic tabletop surface. The tabletop used infrared LED lamps emitting light at a wavelength of 850 nm using a 12 volt power supply. The table's built-in IR camera captured the input image at a resolution of 320×240 pixels and at a rate of 15 fps. The experimental platform was implemented using the TUIO protocol with the Community Core Vision tracker [28], and ran upon a 1.86 GHz Core 2 Duo PC running Windows XP.

**Participants**

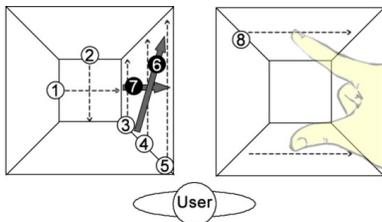
Ten participants (9 male) between the ages of 25 and 35 were recruited from a local university to participate in this study. Every participant was a daily computer user, but only two had previous experience with tabletop systems. All of our participants were right-handed.

**Task and design**

We evaluated TZee using a series of 2D docking tasks. Participants were required to interact with objects displayed upon the device. Users performed sliding gestures upon the TZee using one or two fingers at a time to move objects from an initial position to a marked target. This task allowed us to assess whether all the signals were being captured reliably by our system. A failure in docking could be attributed to a loss of signal. The experiment consisted of two tasks. In the first task, participants were asked to drag the object using their index finger on the top and side surfaces as follows (see Figure 6 left):

<b>TOP SURFACE</b>	(1) left-to-right on the top face;
	(2) back-to-front on the top face;
	(3) front-to-back on the top of the a side;
<b>SIDE SURFACES</b>	(4) front-to-back in the middle of a side;
	(5) left-to-right on the bottom of a side;
	(6) top front-to-bottom back on a side;
	(7) up-to-down on a side.

In the second task, participants were asked to (8) use their thumb and index fingers simultaneously to drag two objects from left to right on the front and back faces (Figure 6 right). This would allow us to evaluate the system’s performance with multiple simultaneous gestures.



**Figure 6 – Left: the direction of movement used in the docking tasks with the index finger. Right: the motion of the docking task with the thumb and index fingers.**

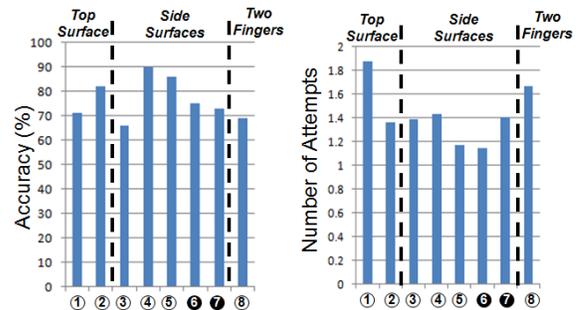
Each condition was repeated 10 times. A trial was completed if participants successfully docked the object. Otherwise, they had to try again until they succeeded. Participants were given an unlimited number of attempts for each trial. The object to be dragged and the target were each large enough (0.6 inches) to ensure that the task was easy to perform. Thus, most of the failure was attributed to the loss of the tracking signal. TZee’s control-display (CD) ratio (only for this experiment) was set to 1:1, so the distance between the object’s initial position and target position was the same as the finger’s dragging distance. Both the object and target dock were

displayed underneath the TZee, clearly visible due to the device’s transparency.

**Results**

Accuracy and number of attempts were recorded for analysis. A failed attempt happened when the drag action was not detected by the system. Therefore the accuracy rate refers to the percentage of time that the system was able to detect the entire gesture (from start to end) without losing the signal. Participants performed the task with an overall average accuracy rate of 76.5%, i.e. in 76.5% of all trials the system detected the entire drag signal. Results show that sliding on the top face attained an average 76% accuracy rate. Accuracy sliding with two fingers (each on opposing faces) had the lowest accuracy (69%). Sliding on the side faces achieved a 78% accuracy rate, except that the performance when sliding on the top (condition 3 in Figure 6) edges of a side face was relatively low (Figure 7 left; around 66%). We found that this usually occurred when users accidentally slid their fingers onto the tape used to hold the faces of the cover together. On the lower ends of each side face, accuracy rates approached 90%.

Participants had an average of 1.43 attempts per trial. The number of attempts was higher with tasks that required sliding along the top and sliding with two fingers (Figure 7 right). This is consistent with the corresponding accuracy rate. Tasks with higher accuracy (e.g. 4, 5 and 6 in Figure 6) also had fewer repeated attempts. Further analysis reveals that most of the attempts were made during the first few trials with the task, suggesting a learning trend.



**Figure 7 – Y-axis shows accuracy rate (left) and average number of attempts (right). X-axis shows different docking conditions described in Figure 6.**

Overall, our results show that TZee can reliably detect several gestures. Furthermore, the results were largely uniform across all of TZee’s regions, albeit without a 100% accuracy rate. It is noteworthy that the results were obtained with minimal amount of training to our participants. The tests were performed using raw camera data. These rates can be improved by applying filters or properly extrapolating points between samples. We did not implement such algorithms as we wanted to obtain a raw performance measure.

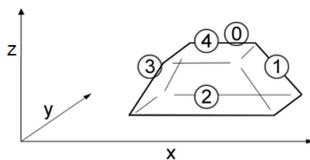
Next, we describe some basic 3D interactions, such as translations and rotations that can be carried out on TZee with gestures exploiting natural metaphors. Additionally, we

show a set of object manipulations that cannot be easily accomplished with existing devices, but are now possible with TZee using simple gestures.

### INTERACTIONS WITH TZEE

Once TZee is placed on the table, a simple calibration routine assigns a unique F-ID (Face Identifier) to each face of the widget (Figure 8). To accommodate input independently of the device's orientation, TZee defines its own local coordinate system. The base square of the TZee pyramid establishes the  $x$  and  $y$  axes. The  $z$ -axis lies perpendicular to the base, extending up and into the table.

TZee can be linked to one or more virtual objects. Once linked, TZee serves as a tangible handle for controlling the object(s). Although we do not explicitly recommend a particular mechanism for linking an object to TZee, we envision two potential methods. The first, Symmetrical Bimanual Synchronous Tapping [31], is similar to that used in SLAP widgets. It pairs an object with the tangible block by double tapping on TZee and the object simultaneously. The second method could allow the user to 'bump' the tangible device against a desired object, as exemplified by [14].



**Figure 8 – Faces of TZee are given a unique F-ID, or face identifier, according to the degree of parallelism between a face and the  $x$ ,  $y$  and  $z$  axes of the table.**

Object manipulations on TZee are aided by the gesture recognition engine. The gesture recognizer requires at least 3-4 detected input samples from a given face to recognize a gesture. Once the gesture is identified, the engine continues the transformation (such as a rotation) until the touch is lifted from the device. This allows for smooth object movements, which compensates for the device's limited input resolution. We assigned a different CD ratio to the sides than the top face of TZee, as our results show that the signal is weaker on the top. In the rest of the paper, we showcase the capability of TZee with a few simple interactions. An investigation of their efficiency and intuitiveness is worthwhile for a future study.

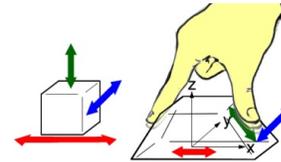
### Translation, scaling, and rotation (TSR)

Basic 3D transformations, i.e. translation, scaling, and rotation on the  $x$ ,  $y$ , and  $z$  axes, can be easily accomplished with TZee in a natural manner. All of the gestures performed upon the device are interpreted in the context of the widget's local coordinate system.

**Translation.** Translation along a particular axis can be accomplished by sliding two fingers across two opposite faces. For instance, translation in the  $z$  direction can be made by sliding straight up or down on faces 1 and 3 (Figure 10). Sliding up moves the object towards the user; sliding down pushes the object away. A horizontal slide on faces 0 and 2

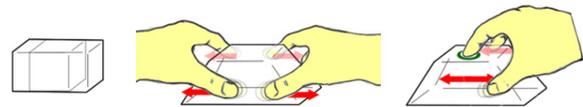
will result in translation in the  $x$  direction. Similarly, sliding horizontally on faces 1 and 3 will translate in the  $y$  direction. Sliding TZee across the tabletop also translates the object along the  $x$ - $y$  plane.

Note that TZee is an indirect input device. One strong benefit of using indirect input for translation is that it allows users to move the object to an area that is unreachable by direct touch. Using indirect input for translation in the  $z$  direction is necessary, since  $z$ -axis motion cannot be performed directly upon a 2D surface.



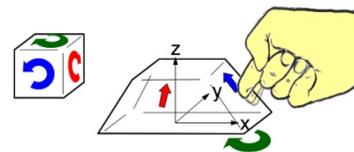
**Figure 10 – Illustration of translation using TZee.**

**Scaling.** Scaling can be achieved with two hands. By placing four fingers on two opposite sides of TZee, the user can stretch the virtual object along that axis. A similar pulling-apart or pushing-together gesture makes stretching in the  $z$  direction possible (Figure 11). Alternatively, this can be achieved using one hand, and a modifier "key". With the index finger resting on the top face as a modifier (face 4), scaling in the  $x$  direction can be accomplished by sliding fingers horizontally on faces 0 and 2. These metaphors replicate our natural actions: scaling requires an origin, represented by the top face touch, plus a spread or contraction, defined by the other two fingers moving in opposition. Scaling is similarly achieved in the  $y$  and  $z$  directions.



**Figure 11 – Illustration of scaling in the  $x$ -axis using two hands (middle) and one hand (right).**

**Rotation.** A horizontal flick gesture on any of the tilted faces (F-IDs 0–3) will result in a rotation about the  $z$ -axis. In addition, turning the widget will also rotate the associated object around the  $z$ -axis. A vertical flick gesture on faces 1 and 3 will rotate the virtual object about TZee's  $y$ -axis. A vertical flick gesture on faces 0 and 2 rotates the object about TZee's  $x$ -axis (details in Figure 9).



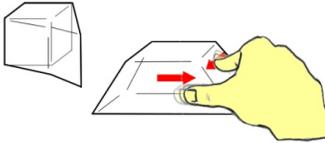
**Figure 9 – Illustration of rotation along each of the three axes using TZee. The interactions map naturally to how one would rotate a physical object along these axes.**

### Beyond TSR

In addition to the basic functions supporting translation, scaling, and rotation, TZee can support a variety of novel inte-

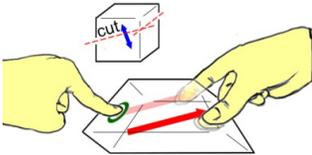
rations. These have not been demonstrated with prior tabletop 3D interaction systems. In this paper, we discuss only a few of them, and leave the design space to be built upon by future application designers.

*Stretching across axes.* With TZee, virtual objects can be stretched across two axes at a time. To do this, the user can slide the thumb and index fingers towards the edge of two adjacent faces. The corresponding part of the object is stretched out (Figure 12).



**Figure 12 - Illustration of stretch on two adjacent faces using thumb and index finger.**

*Modifier key.* As demonstrated by the scaling operation, TZee supports the use of a modifier key metaphor to extend the functionality of fundamental gestures. Alternate modes are triggered by holding one or more fingers on any of the tangible faces. We demonstrate an example of using a modifier key to invoke a *cut* operation. A virtual object can be sliced into two pieces by cutting it. This task is not uncommon when interacting with 3D volumetric data. The slice action is accomplished with a translation gesture with the addition of a finger held on a tilted face as a modifier (Figure 13). In this case, the translation gesture defines the position and the orientation of a cutting plane.



**Figure 13 – Illustration of using a ‘modifier key’ to “cut” an object with different cutting planes.**

*Widget as display.* Since TZee is transparent, images on the surface beneath the widget can be clearly seen through it. This enables a wide range of possibilities. For instance, novice users could benefit from the availability of guides much like our sketches that demonstrate how gestures are performed. Alternatively, to support more sophisticated modes of interaction, various options or parameters for a transformation might be displayed upon the device. At the very least, the usual contents of the tabletop will be visible through the device.

## DISCUSSION

We presented an accessible tangible device capable of supporting numerous useful 3D interactions. Here we summarize our findings. There are also a few limitations in the current design, however, that need to be addressed carefully in future studies.

### Summary of results

The development of TZee was multi-faceted. In the first step, we identified the necessary ergonomics of the device that

would suit our objectives. This in turn led us to create prototypes with materials consisting of fiber, silicone, and finally settle on stacking acrylic blocks and covering them with a thin transparent shell. All of our tests confirmed that this allowed light from the tabletop setup to travel at a height slightly higher than what would be possible in air alone. We leveraged this property to create interactions that were intuitive and that mapped naturally (to the extent possible) to interactions in 3D.

Our system test showed that we can reliably detect a complete sliding action, along the sides of TZee, over 75% of the time. This is impressive given that we did not apply any noise reduction algorithms (other than what was available through the TUIO protocol) to the raw camera images. With such algorithms, we expect accuracy rates to improve, whereby lost signals could be extrapolated from samples that were already collected. Since TZee is an indirect input device, it can use the most suitable control-display ratio for smooth input. This can allow designers to extend the physical space available on the sides to create rich interactions.

Despite these favorable properties that make TZee a viable tool for rapid 3D manipulation tasks, certain limitations need to be considered in its construction.

### Limitations

*Touch area.* The ability to detect a signal with TZee is sensitive to the size of users’ touch points. The effectiveness of the device also varies with the pressure exerted by users. The skin tissue at the tips of the fingers flattens out when pressed against the sides of TZee. This forms a larger touch area that reflects more light than a comparable curved surface. If a user placed two fingers together to create a single large touch point on the top surface, for example, signals could sometimes be detected more reliably. The TZee shell was also important in this regard. It not only gave users a smooth surface upon which to glide their fingers on, but also flattened the finger tips as they touched it.

*Resolution.* TZee’s effectiveness of supporting precise gestures is limited by the size of the device, combined with the resolution of the tabletop’s camera. Our system used a 320x240 camera resolution, with which we were able to reliably detect the direction of users’ gestures. However, detecting the precise amplitude of each gesture is difficult with such a low resolution. The effect of this can be mitigated using fixed-rate transformations rather than controlling the rate with the gesture amplitude. We intend on using higher-resolution cameras that support quick frame rates to improve the resolution limitation. In its current design, TZee can only recognize two finger touches per side. There must also be space between each touch point, or else the software will detect a single large blob rather than two smaller ones. A larger surface area for each can resolve this issue, but in all our informal feedback from users, none were concerned about the use of only two touches per face.

*Light loss.* Like other optical tangible widgets [1], TZee is also vulnerable to issues of light loss during the transmission

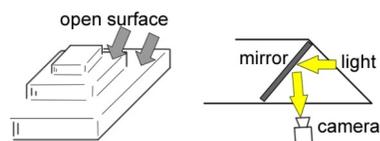
from the LEDs beneath the surface of the tabletop, through the surface itself, and into the device. The current prototype is made of 3 layers of acrylic pieces and a thin acrylic cover. This setup introduces a maximum of 8 reflective surfaces (two per piece) before the light reaches the user's finger. During the transmission of the light from the tabletop surface to the surface of TZee, each ray passing through an acrylic piece gets refracted twice. This increases the amount of loss, and therefore reduces the intensity of the light above the table. Furthermore, the air between the acrylic steps and the cover also introduces a fair amount of light spread. This is a result of change of direction when the light enters or leaves the air (since air has a smaller refractive index than acrylic). These issues limit the vertical distance that light can travel within the TZee. With our prototype, we were unable to detect steps reliably through more than three layers unless the touch points were increased beyond the size of human fingertips (although two finger tips together would provide a large enough area to enable detection with four blocks). On the other hand, the loss at the interfaces can be minimized by manufacturing a monolithic acrylic device or by joining the acrylic panels with optically transparent adhesives with an index of refraction close to that of acrylic. Alternatively, higher light intensity can be achieved by employing panels of transparent materials with higher ( $>1.5$ ) refractive index. This would result in stronger bending of rays, and thus a more pronounced "narrowing" of the light's radiation profile.

*Lighting conditions.* To obtain optimal input, we needed to find the appropriate balance between light intensity, noise and blob size threshold. Like all diffuse illumination systems, TZee requires a certain degree of calibration that could vary slightly based on the lighting conditions available. We are further investigating methods of limiting the noise that can sometimes result from improper light settings.

*Slope of tilted faces.* The slope of the tilted faces is largely dependent upon the size of the open step surfaces in each acrylic layer (Figure 14 left). A steeper slope presents a smaller surface to the camera. To obtain a preliminary understanding of the effects of slope on signal transmission, we built several TZee prototypes, each with a different slope. We tested the performance of these prototypes, and found that finger gestures can be reliably detected on tilted surfaces of slopes up to  $39^\circ$ . Past this limit, the size of each slanted face (when projected upon the tabletop from the perspective of the camera) is not sufficiently large to capture the gestures reliably. Additionally, the number of acrylic layers and their thickness determines the size of the open space on each layer. Therefore, it is important for designers to be aware of the balance between these factors in designing TZees with different slopes.

*Alternative designs.* The limitations of the current prototype can be resolved to some extent by using different methods to construct TZee. For instance, one approach might be to embed a reflective area (mirror) inside a TZee facing the tabletop surface (Figure 14 right). This design allows the slope of tilted faces to be independent of the open space in

each layer; such a device could potentially support the construction of steeper devices or other shapes altogether. These approaches may have potential and more work is needed to verify these possibilities.



**Figure 14 – Left: open surface on each acrylic layer causes some loss of light, but was necessary to avoid total internal reflection; Right: using a reflective surface such as a metallic mirror to bring the image from tilted faces to the camera may also be possible. However the ability to manipulate objects using face #4 would be negatively affected.**

### CONCLUSION AND FUTURE WORK

We introduce TZee, an accessible and cost-effective tangible widget designed for natural 3D interactions on tabletop systems. TZee is an unpowered device that works by exploiting the lighting sources of the tabletop. Its acrylic medium allows IR light to be transferred to its faces above the surface without major intensity loss. This allows finger motions carried out on the faces of TZee to be captured by the tabletop's integral camera. In our system evaluation, we show that even with a low resolution webcam ( $320 \times 240$ , 15 fps), TZee can reliably extend input above the surface of the table. We demonstrate that TZee can be easily built by using cost-effective materials such as acrylic pieces.

Because of TZee's angled faces, 3D object interactions that were previously not natural on tabletop systems can be carried out intuitively. This has the added benefit of avoiding some of the ambiguities that lessen the effectiveness of previous gestural systems. We demonstrated that basic transformations such as translation, rotation, and scaling can be easily accomplished with TZee. In addition, the widget facilitates both one- and two-handed interactions, thus enabling a large design space for novel bi-manual interactions. Future work will be focused on evaluating different hardware designs so that the limitations of our prototypes can be resolved. In addition, we will develop software packages to integrate TZee with real-world applications. We also plan to explore the use of TZee in collaborative settings.

### ACKNOWLEDGEMENTS

We thank Saravanan Nandagopal, Francis Lin and Jing Li for their help with PDMS block fabrication. We also thank NSERC for funds.

### REFERENCES

1. Baudisch, P., Becker, T., and Rudeck, F. (2010) Lumino: Tangible Blocks for Tabletop Computers Based on Glass Fiber Bundles. *CHI'10*, 1165-1174
2. Benko, H., Saponas, T. S., Morris, D., and Tan, D. (2009). Enhancing Input On and Above the Interactive Surface with Muscle Sensing. *ITS'09*.
3. Benko, H. and Wilson, A. (2009). DepthTouch: Using Depth-Sensing Camera to Enable Freehand Interactions

- On and Above the Interactive Surface. *Microsoft Research Technical Report MSR-TR-2009-23*. 2009.
4. Cutler, L. D., Fröhlich, B., and Hanrahan, P. (1997). Two-handed direct manipulation on the responsive workbench. *i3D'97*, 107-114.
  5. Davidson, P. L. and Han, J. Y. (2008). Extending 2D object arrangement with pressure-sensitive layering cues. *UIST'08*, 87-90.
  6. Fiebrink, R., Morris, D., and Morris, M. R. (2009). Dynamic mapping of physical controls for tabletop groupware. *CHI'09*, 471-480.
  7. Fitzmaurice, G. W., Ishii, H., and Buxton, W. A. S. (1995). Bricks: laying the foundations for graspable user interfaces. *CHI'95*, 442-449.
  8. Gallant, D. T., Seniuk, A. G., and Vertegaal, R. (2008). Towards more paper-like input: flexible input devices for foldable interaction styles. *UIST'08*, 283-286.
  9. Guo, C., Young, J. E., Sharlin, E. (2009). Touch and toys: new techniques for interaction with a remote group of robots. *CHI'09*, 491-500.
  10. Hancock, M., Carpendale, S., and Cockburn, A. (2007). Shallow-depth 3d interaction: design and evaluation of one-, two- and three-touch techniques. *CHI'07*, 1147-1156.
  11. Hancock, M., Cate, T., and Carpendale, S. (2009). Sticky Tools: Full 6 DOF Force-Based Interaction for Multi-Touch Tables. *ITS'09*, 145-152.
  12. Hancock, M., Hilliges, O., Collins, C., Baur, D., and Carpendale, S. (2009). Exploring tangible and direct touch interfaces for manipulating 2D and 3D information on a digital table. *ITS'09*, 85-92.
  13. Hilliges, O., Izadi, S., Wilson, A. D., Hodges, S., Garcia-Mendoza, A., Butz, A. (2009). Interactions in the air: adding further depth to interactive tabletops. *UIST'09*, 139-148.
  14. Hinckley, K. (2003). Synchronous gestures for multiple persons and computers. *UIST '03*, 149-158.
  15. Huang, C. (2004). Not just intuitive: examining the basic manipulation of tangible user interfaces. *CHIEA'04*, 1387-1390.
  16. Ishii, H. and Ullmer, B. (1997). Tangible bits: towards seamless interfaces between people, bits and atoms. *CHI'97*, 234-241.
  17. Jordà, S., Geiger, G., Alonso, M., and Kaltenbrunner, M. (2007). The reacTable: exploring the synergy between live music performance and tabletop tangible interfaces. *TEI'07*, 139-146.
  18. Kitamura, Y., Itoh, Y., Masaki, T., and Kishino, F. (2000). ActiveCube: A Bi-directional User Interface using Cubes. *KES'00*, 99-102.
  19. Kruger, R., Carpendale, S., Scott, S. D., Tang, A. (2005). Fluid integration of rotation and translation. *CHI'05*, 601-610.
  20. Microsoft Surface, <http://www.microsoft.com/surface/>
  21. Patten, J. and Ishii, H. (2007). Mechanical constraints as computational constraints in tabletop tangible interfaces. *CHI'07*, 809-818.
  22. Reisman, J. L., Davidson, P. L., and Han, J. Y. (2009). A screen-space formulation for 2D and 3D direct manipulation. *UIST'09*, 69-78.
  23. Rekimoto, J., Ullmer, B., and Oba, H. (2001). DataTiles: a modular platform for mixed physical and graphical interactions. *CHI'01*, 269-276.
  24. Rivière, J., Kervégant, C., Orvain, E., and Dittlo, N. (2008). CubTile: a multi-touch cubic interface. *VRST'08*, 69-72.
  25. Sato, T., Mamiya, H., Koike, H., and Fukuchi, K. (2009). PhotoelasticTouch: transparent rubbery tangible interface using an LCD and photoelasticity. *UIST'09*, 43-50.
  26. Subramanian, S., Aliakseyeu, D., and Lucero, A. (2006). Multi-Layer Interaction on Digital Tables. *UIST'06*, 269-272.
  27. Tuddenham, P., Kirk, D., Izadi, S. (2010) Graspables revisited: multi-touch vs. tangible input for tabletop displays in acquisition and manipulation tasks. *CHI'10*: 2223-2232.
  28. TUIO (tracker). <http://www.tuio.org/>.
  29. Uiga, E. (1995). Optoelectronics. Prentice-Hall, pp. 352. ISBN 0-02-422170-8.
  30. Vlaming, L., Collins, C., Hancock, M., Isenberg, T., and Carpendale, S. (2010). Integrating 2D mouse emulation with 3D manipulation for visualizations on a multi-touch table. *ITS'10*, 221 – 220.
  31. Weiss, M., Wagner, J., Jansen, Y., Jennings, R., Khosha-beh, R., Hollan, J. D., and Borchers, J. (2009). SLAP widgets: bridging the gap between virtual and physical controls on tabletops. *CHI'09*, 481-490.
  32. Wilson, A. (2007). Depth-Sensing Video Cameras for 3D Tangible Tabletop Interaction. *Tabletop '07*, 201-204.
  33. Wilson, A., Izadi, S., Hilliges, O., Garcia-Mendoza, A., and Kirk, D. (2008). Bringing physics to the surface. *UIST'08*, 67-76.
  34. Yang, X. D., Mak, E., McCallum, D., Irani, P., Cao, X., and Izadi, S. (2010). LensMouse: Augmenting the mouse with an interactive touch display. *CHI'10*, 2431-2440.
  35. Zhou, Z., Cheok, A. D., Li, Y., and Kato, H. (2005). Magic cubes for social and physical family entertainment. *CHIEA'05*, 1156-1157.