Modeling On and Above a Stereoscopic Multitouch Display

Bruno R. De Araùjo
INESC-ID, DEI IST, Technical
University of Lisbon, Portugal
brar@vimmi.inesc-id.pt

Joaquim A. Jorge
INESC-ID, DEI IST, Technical
University of Lisbon, Portugal
jorgej@acm.org

Géry Casiez
LIFL, INRIA Lille, University of
Lille, Villeneuve d'Ascq, France
gery.casiez@lifl.fr

Martin Hachet
INRIA Bordeaux - LaBRI,
Talence, France
Martin.Hachet@inria.fr

Figure 1: 3D object modeling by interacting on and above a multi-touch stereoscopic surface.

Abstract

We present a semi-immersive environment for conceptual design where virtual mockups are obtained from gestures we aim to get closer to the way people conceive, create and manipulate three-dimensional shapes. We developed on-and-above-the-surface interaction techniques based on asymmetric bimanual interaction for creating and editing 3D models in a stereoscopic environment. Our approach combines hand and finger tracking in the space on and above a multitouch surface. This combination brings forth an alternative design environment where users can seamlessly switch between interacting on the surface or in the space above it to leverage the benefit of both interaction spaces.

Author Keywords

3D Modeling; 3D User Interface

ACM Classification Keywords

H.5.2 [User Interfaces]: Graphical user interfaces (GUI). Input devices and strategies (e.g., mouse, touchscreen), Interaction styles (e.g. commands, menus, forms, direct manipulation)

Introduction

Despite the growing popularity of Virtual Environments, they have not yet replaced desktop CAD systems when it

comes to modeling 3D scenes. Traditional VR idioms are still umbilically connected to the desktop metaphor they aim to replace, by leveraging on the familiar "Windows, Icons, Menus, Pointing" (WIMP) metaphors. Worse, the command languages underlying many of these systems also do not map well to the way people learn to conceive, reason about and manipulate three-dimensional shapes.

In this paper, we explore 3D interaction metaphors to yield direct modeling techniques in stereoscopic multitouch virtual environment. Combined with user posture tracking based on a depth camera and three-dimensional finger tracking, this rich environment allows us to seamlessly pick and choose the sensing technique(s) most appropriate to each modeling task. Based on this groundwork, we have developed an expressive set of modeling operations which build on user's abilities at creating and manipulating spatial objects. Indeed, from a small set of simple, yet powerful functions users are able to create moderately complex scenes with simple dialogues via direct manipulation of shapes in a less cumbersome way.

Related Work

With the widespread adoption of multitouch devices and less expensive and intrusive tracking solutions such as the Microsoft Kinect, academic research on tabletop has refocused on "on" and "above" surface interaction techniques. Müller-Tomfelde *et al.* proposed different methods to use the space above the surface to provide ways of interacting with 2D tabletop content closer to reality [13]. While tangible devices complement the surface physically with a direct mapping to the GUI such as in the Photohelix system and StereoBlocks [9], gestures above the surface mimic physical interaction with real objects. Wilson *et al.* proposed several metaphors to interact with different displays while capturing full body

posture [16]. In this way, users can interact on or above the surface with 2D content or even between surfaces using the body to transfer virtual content to the hand or to another surface while moving their bodies in space. Users can also interact physically in space with projected GUI. In our system, we prefer to use the surface for GUI since it is more adequate for discrete selection and explore space gesture for modeling actions.

Our approach explores the continuous space as presented by Marquardt *et al.* [12]; however we enrich their approach by combining it with the bimanual asymmetric model proposed by Guiard [6]. This model proposes guidelines for designing bimanual operations based on observations of users sketching on paper. For these tasks, Guiard identified different rules and actions for the preferred (also dominant-hand or DH) and non-preferred (also non-dominant hand, or NDH) hand. While the DH performs fine movements and manipulates tools, the NDH is used to set the spatial frame of reference and issue coarse movements. This approach has been explored by several systems [1, 8, 10, 11] by combining finger- or hand- gestures with pen devices. Brandl *et al.* proposed a sketching system where the user selects options through touches using the NDH on a WIMP–based graphical interface, while the DH is used to sketch using a pen device [1]. Such a configuration allows to better explore hand gestures proposing richer interaction concepts to represent 2D editing operations such as demonstrated by Hinckley *et al.* [8]. Indeed, this makes switching between modalities easier and allows users to perform a wide range of 2D editing tasks without relying on gestures or GUI invocations. Lee combined hand gestures while sketching using a collapsible pen to define curve depth on a tabletop [10]. The NDH is tracked allowing users to seamlessly specify 3D modeling commands or modes such
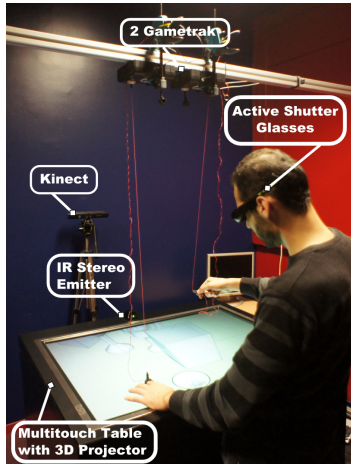
Figure 2: Our stereoscopic multitouch modeling setup showing the multitouch steroscopic display, the Gametraks used to track the fingers above the surface and a Kinect to track user's head.

as the normal direction of an extrusion while specifying the displacement by interacting with the pen on the virtual scene. Contrary to their approach, we preferred to keep the surface for fast and accurate 2D drawing, while benefiting from the 3D input space for controlling depth directly. Lopes *et al.* adapted the ShapeShop sketch based free-form modeler to use both pen and multitouch simultaneously [11]. They found out that the asymmetric bimanual model allows users to perform more manipulations in less time than conventional single interaction point interfaces, which increased the percentage of time spent on sketching and modeling tasks. By tracking the hands of the user, we adopt the asymmetric bimanual model to easily switch between sketching, model editing, navigation and spatial manipulation of objects. In addition, we do not need to rely on special input devices nor extra modalities to assign different roles to each hand.

We rely on a stereoscopic visualization setup for architectural model visualization similar to [3]. While this system allows navigating or annotating the 3D scene mainly as if it was inside the table and use fingers as proxies over the scene, our interaction techniques focus on modeling and direct manipulation since 3D models are rendered as if they were lying atop the table. To avoid hands occlusions over the visualization, Toucheo [7] proposed a fish-tank like setup using a multitouch surface and a stereoscopic display. However such as other setups relying on semi-transparent mirrors to create holographic illusion,it both reduces the working space and constrains the usage of the above surface space to hand gestures. Our stereoscopic visualization setup provides more freedom of movement allowing a continuous space of interaction. In addition, adopting a bimanual asymmetric model makes possible new interaction techniques which

could benefit interaction with holographic display technologies when they become available.

Hardware Modeling Setup

Our setup consists in a semi-immersive environment based on a stereoscopic multitouch display 96×72 cm (42") combined with a Kinect depth camera and two Gametraks[1]. Head tracking is achieved in a non-intrusive way thanks to the Kinect using its skeleton detection algorithm. The skeleton is also used to track user hands allowing to locate the dominant hand according to the handedness of the user. Finger tracking is done using the Gametraks with a good precision over the working space above the table while reducing occlusion problems and providing a higher framerate (125 Hz) compared to techniques based on the Kinect device alone (Figure 2). The visualization relies on a back-projection based system located under the table running at 120 Hz with a XGA pixels resolution. It is coupled with active shutter glasses from 3D Vision NVIDIA for the stereoscopic visualization. The 3D scene is rendered on top of the surface and the point of view is updated according to the position and orientation of the user's head to take into account motion parallax. The IR transmitter for the glasses uses an IR wavelength different from the multitouch table which is based on the Diffuse Illumination technique. It is set at a position to cover the working volume around the table where the user interacts. We use the iLight[2] framework version 1.6 for fingers detection and tracking. Fingers data are then sent using TUIO messages to our custom built application. The two Gametraks are used to track the 3D position of the index and thumb of each hand when they are no longer in contact with the multitouch surface.

---

[1]http://en.wikipedia.org/wiki/Gametrak
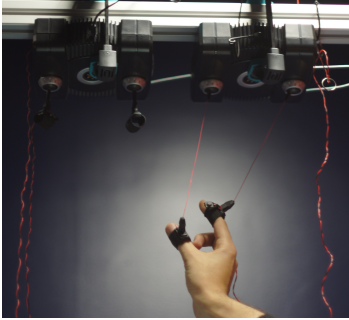[2]iliGHT Tactile Table product page: http://www.immersion.fr

Figure 3: Fingers tracking above the multitouch surface using two Gametrak devices.
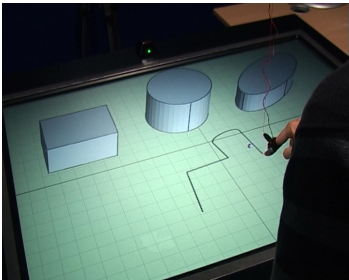

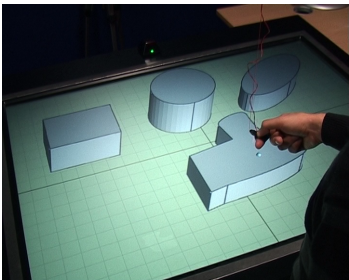**Figure 4:** Sketching on the surface using the DH.

Figure 5: Extruding a shape along its normal in space with a pinch gesture.

These low cost gaming devices are placed in a reverse position centered above the table at a distance of 120 cm. The 3D position of each finger is computed from the two angles of rotation and the length of each cable, digitalized on 16 bits and reported at 125Hz to the host computer. The retractable strings are attached to the fingers through a ring. Although strings introduce some visual clutter, they were not found to distract users from their task. The strings create a minor spring effect which reduces user hand tremor without adding fatigue. We added a 6mm diameter low profile momentary switch button on each index finger to detect pinch gestures without ambiguity (Figure 3). This simple solution provides a good trade-off regarding precision, cost and cumbersomeness compared to using a high end marker based optical tracking system or low sampling frequency device such as the Kinect.

To obtain a continuous interaction space, the coordinates from the different input devices are converted to the Kinect coordinate system. By identifying the four multitouch surface corners in the image captured by the Kinect, we are able to compute the transformation matrix from 2D touch position to 3D space. Regarding the Gametrak devices, a transformation matrix is computed, for each tracked finger, by sampling the multitouch surface screen in one thousand positions. The rigid transformation is then computed using a RANSAC algorithm [4]. Kinect Skeleton tracking data is used to enrich our user model and input data is fused by proximity in a unique reference space. Doing so, we can define the frustum of the off-axis stereo perspective projection to render 3D content on top of the surface from the user point of view. The redundancy of information from the different input devices allows us to identify which finger of which hand is interacting on the surface or in the air or to choose the input source with the best tracking resolution.

Our Modeling Approach

We propose a direct modeling approach to create, edit and manipulate 3D models using a small set of operations. After drawing sketches, users can create 3D models by pushing and pulling existing content of the scene such as [14] or Google Sketchup. Our models are represented using a boundary representation which decomposes the topology of objects into faces, edges and vertices. We adapt our modeling approach to take advantage of the bimanual interaction model while the user is interacting on and above the surface.

Sketching On the Surface

The multi-touch surface is primarily used as a sketching canvas where the user interacts using fingers as depicted by Figure 4. User can sketch on the surface creating planar shapes from close contours using the DH. Contours might use lines, curves or both and can be sketched using multiple strokes. Open strokes whose extremities are close to each other are merged into a single stroke. While sketching, input data is fitted incrementally to the best fit of lines and cubic Bézier curves. Our incremental fitting algorithm based on curve fitting tries to guarantee the continuity between curves and segments by adding tangency constraints during the fitting process. When a closed contour is created on the surface, simple planar polygons can be created by the user. We perform a simple stroke beautification based on constraints detected from sketches. These constraints rely on line segments to detect parallel and perpendicular line pairs and segment pairs with equal length. We use a threshold on angles between segments for parallelism and perpendicularity and a threshold ratio relationship between segments with similar length. An energy function is specified for each type of constraint and we perform an error minimization method to beautify user sketches. Thanks to this process,
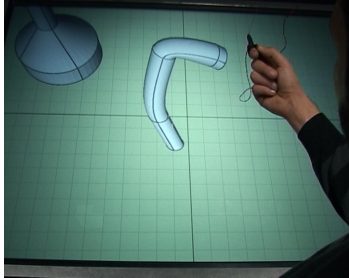
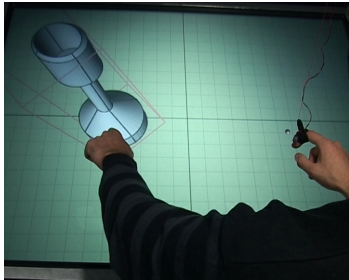**Figure 6:** Extruding a profile along a curvilinear trajectory with the DH.

Figure 7: Moving a shape on the surface using NDH.
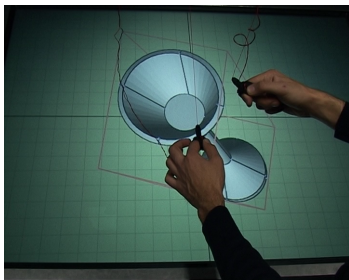

Figure 8: Scaling an object in space using both hands.

regular shapes can be created using line drawing. Regarding closed conic sections, we use a 2D shape recognizer [5] to detect circles and ellipses which are approximated by a closed piecewise curve using four cubic Bézier segments. We also use the 2D shape recognizer to detect simple gestures such as an erasing command by drawing a scribble. When an erasing gesture is recognized, if it overlaps open strokes, they are erased. However, if it overlaps only shapes and not open strokes, overlapped shapes are erased. This solution allows to use open strokes as construction lines while modeling.

Creating 3D Shapes by Pushing and Pulling operations
Gestures with the DH above the surface are interpreted as 3D object creation or edition. Creation of 3D shapes consists in extruding a planar shape previously sketched on the surface following the push and pull modeling metaphor. The user first approaches the DH index finger near a planar shape on the surface to highlight it. He then performs a pinch gesture, pressing the button located on the index finger, to extrude the shape along the normal of the surface (Figure 5). The height of the extruded object is then continuously updated and co-located with the position of the finger until the button is released. Planar shapes can also be extruded along the trajectory defined in the air after the user has selected this operation in a menu displayed on the NDH (Figure 6). While the user is defining the trajectory, the path is continuously re-evaluated and fitted into line segments and curve pieces similarly to what is done for strokes on the surface. Segments and curve pieces are then used to create smooth free form extrusion of the profile offsetting the gesture from the centroid of the face to its vertices as presented by [2]. This method enables to extrude both poly-line and curvilinear profiles along linear or curvilinear paths.

Additionally, topological features of the shape (vertices, edges and faces) can be selected and displaced along a normal direction updating the geometry of the object but not changing its topology as done by the extrusion operation. It offers edition by pushing and pulling any topological feature of our boundary representation. Selection of features is done implicitly by touching a geometrical feature on the surface and explicitly using a pinch gesture in space. Since edges and vertices can be shared by more than one face or edge respectively, a continuous selection mechanism is provided for selection disambiguation analyzing the previously highlighted entity. For example, it is possible to highlight a particular edge of a face shared by two faces by selecting it from the face the user is interested in. If no geometrical feature is selected while doing the pinch gesture with the DH, the user can sketch 3D lines or curves in space.

Manipulating 3D Shapes
When starting a gesture on the surface with the NDH, it is interpreted as object transformation if it is performed on an object or world manipulation otherwise. Single touch gestures are interpreted as object or world translation. More than one finger gestures are interpreted as translation, rotation and scale operations on objects or world following the well-known RST paradigm. 3D objects are constrained to movements along the plane parallel to the multitouch surface. A gesture started with the NDH can be complemented by the DH allowing translation, rotation and scale with both hands (Figure 7).

The bimanual interaction used on the surface is also valid above the surface allowing to rotate, translate and scale objects using two fingers. As on the surface, the NDH begins the interaction using a pinch gesture. The NDH defines translations only while the DH adds rotation and
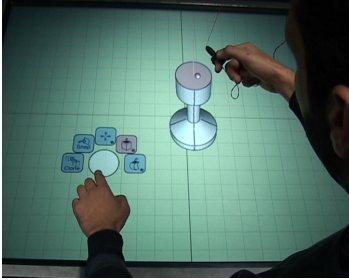
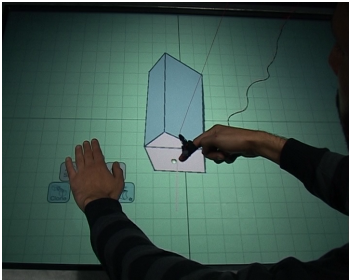**Figure 9:** Contextual menu presented while selecting a face in space.

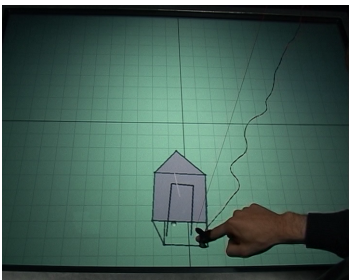**Figure 10:** Selecting a vertical face for snapping.

Figure 11: Cutting a face on the canvas after been snapped.

scale operations using the method proposed by Wang *et al.* [15] as depicted in Figure 8. These direct 3D object manipulations appear much more efficient compared to indirect interactions on the multitouch surface alone.

Menu based Interaction
We rely on menu based graphical user interface to distinguish between modeling modes such as linear and curvilinear extrusion or other operations such as copy. Modes are presented through items shown in a contextual menu presented under the NDH while a shape or part of it is selected with the DH. Modes presented in the contextual menu correspond to the ones available in the current mode associated to the operation performed by the DH (Figure 9). If the operation carried by the DH only supports a single mode, no contextual menu is shown under the NDH. To avoid visual clutter, the contextual menu transparency is adjusted based on the distance between the NDH and the surface. Above 15 cm, the menu is fully transparent and becomes progressively opaque as the NDH approaches the surface. To improve the accessibility, the contextual menu follows the NDH but its location is progressively fixed as the NDH comes closer to the surface to avoid spatial instabilities and reduce errors while selecting an item.

The discrete mode selection includes the extrusion type (normal to a face or along a trajectory), updating the object topology or simply moving it, the cloning operation and the snapping operation described in the following sub section. When a shape is created, we associate to each face the straight extrusion along the normal as the default mode since it is the most likely operation in the push and pull modeling approach. When the straight extrusion starts, we automatically change the mode to the face move operation, updating the shape without adding new

topological changes. Successive extrusions can be done to create stacked like shape parts by interacting with the menu. Since the menu follows the position of the NDH, it can be used to define the location where clones appear when the cloning operation is selected by the user. The cloning is available when any shape is selected and it duplicates the entire shape as illustrated in Figure 12.

Navigating between Surface and Space
Creating 3D planar shapes in space remains an operation difficult to perform due to lack of physical constraints to guide the hand. We propose a snapping operator to easily switch between the surface and space allowing to use sketches on the surface or gestures in 3D space at convenience. Snapping is available through the contextual menu accessible on the NDH to snap on or back on any selected face (Figure 10). It works by computing a transformation matrix to align the 3D scene to the visible grid defined as a representation of the table surface. A simple linear animation between the two orientations is rendered to help the user understand the new orientation of the model. Furthermore, it allows sketching details on existing shapes (Figure 11) or guaranteeing that new shapes are created on top of an existing shape.

Constraining 3D Operations
Since most of 3D editing operations are performed using only the DH, we decided to use the free NDH to enrich our 3D operators and constrain both sketching and 3D modeling to create more rigorous and controlled shapes. The simplest constrained operation allows sketching symmetrical shapes on the surface. First, the user sketches a straight line defining a mirroring plane which can be selected by touching it with the NDH. While the mirroring plane is selected, sketches using the DH are automatically mirrored and are considered as additional
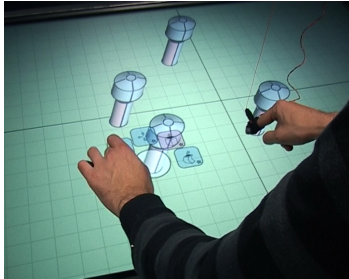
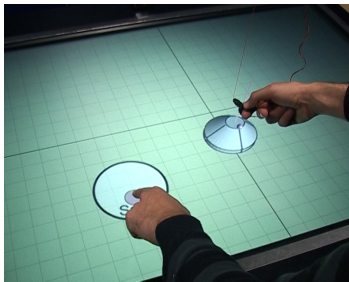**Figure 12:** Cloning an object selected by the DH through the contextual menu.

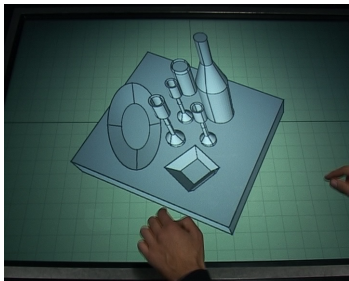**Figure 13:** Scaling a profile with the NDH while extruding a shape.

Figure 14: TableWare scene example.

strokes if the selection remains active at the end of the sketch. By creating a closed shape formed by a stroke and its mirrored version, users can create symmetrical shapes. It can also be used to add symmetrical details to an existing stroke or shape.

3D operations above the surface can also be constrained. For example, while an object is being extruded with the DH, the NDH can select a face of an object to define a maximum or minimum height constraint. Once the constraint is defined, the user continues to move his DH until the maximum or minimum height is reached. Further movements along the preceding direction do not continue to update the height of the object. This allows the user to also define that the height of an object should not be higher or lower that the height of another object.

While the two previous operations illustrate discrete constraints defined by the NDH which can be activated before or during an editing operation, we also explore the usage of dynamic constrains which can be updated continuously during an extrusion operation. This is illustrated with the scale constraint which consists in scaling the profile while extruding a shape (Figure 13). This allows the creation of a cone or a frustum from a circle or a quadrilateral planar face respectively. The scaling factor can be controlled dynamically using a 2D overlay menu accessible by the NDH while extruding the shape.

While traditional modeling interfaces usually require constraints to be defined before performing 3D operations in order to define a problem to be solved by the application, our approach proposes an interactive constraint modeling solution. Doing so, we take advantage of the increase of expressiveness provided by bimanual interaction techniques. Furthermore, we

hypothesis that this definition of constraints on the fly allows to improve the flow of interaction and better fits constraint based modeling in conceptual design stages.

Preliminary Evaluation

Our system was informally tested throughout its development to assess the different design choices and iteratively improve the design of the interface. In total, between 15 and 20 participants tested the interface. Most of the participants were undergraduate and graduate students in Computer Science with variable experience with CAD applications. Our observations suggest that participants quickly learned how to use the interface. However, we noticed that participants can confuse the usage of the two hands at the beginning: sometimes participants wanted to move an object using their dominant hand. As they were first presented an overview of the interface with the basic operations available, further evaluation would help evaluate to determine the learning curve. Figure 14 presents a set of objects created using our system by an expert user in 5'20".

Conclusions and Future Work

We have described an approach to model 3D scenes using semi-immersive virtual environments through a synergistic combination of natural modalities afforded by novel input devices. While early experiments and informal assessments of our system show promise and seemingly validate some of these assumptions, we plan to run formal evaluations with both novice and expert users to highlight and explore both the strengths and the weakness of our modeling interface.

Acknowledgements

References

[1] Brandl, P., Forlines, C., Wigdor, D., Haller, M., and Shen, C. Combining and measuring the benefits of bimanual pen and direct-touch interaction on horizontal interfaces. In *Proceedings of AVI '08*, ACM (NY, USA, 2008), 154–161.

[2] Coquillart, S. Computing offsets of b-spline curves. *Comput. Aided Des. 19* (July 1987), 305–309.

[3] De la Rivière, J.-B., Dittlo, N., Orvain, E., Kervégant, C., Courtois, M., and Da Luz, T. ilight 3d touch: a multiview multitouch surface for 3d content visualization and viewpoint sharing. In *Proceedings of ITS '10*, ACM (NY, USA, 2010), 312–312.

[4] Fischler, M. A., and Bolles, R. C. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Commun. ACM 24* (June 1981).

[5] Fonseca, M., and Jorge, J. Using fuzzy logic to recognize geometric shapes interactively. In *Proceedings of IEEE Fuzzy Systems 2000*, vol. 1 (May 2000), 291–296 vol.1.

[6] Guiard, Y. Asymmetric division of labor in human skilled bimanual action: The kinematic chain as a model. *Journal of Motor Behavior 19* (1987), 486–517.

[7] Hachet, M., Bossavit, B., Cohé, A., and De la Rivière, J.-B. Toucheo: multitouch and stereo combined in a seamless workspace. In *Proceedings of UIST '11*, ACM (NY, USA, 2011), 587–592.

[8] Hinckley, K., Yatani, K., Pahud, M., Coddington, N., Rodenhouse, J., Wilson, A., Benko, H., and Buxton, B. Pen + touch = new tools. In *Proceedings of UIST '10*, ACM (NY, USA, 2010), 27–36.

[9] Jota, R., and Benko, H. Constructing virtual 3d models with physical building blocks. In *Proceedings of CHI EA '11*, ACM (NY, USA, 2011), 2173–2178.

[10] Lee, J., and Ishii, H. Beyond: collapsible tools and gestures for computational design. In *Proceedings of CHI EA '10*, ACM (NY, USA, 2010), 3931–3936.

[11] Lopes, P., Mendes, D., Araújo, B., and Jorge, J. A. Combining bimanual manipulation and pen-based input for 3d modelling. In *Proceedings of SBIM '11*, ACM (NY, USA, 2011), 15–22.

[12] Marquardt, N., Jota, R., Greenberg, S., and Jorge, J. A. The continuous interaction space: interaction techniques unifying touch and gesture on and above a digital surface. In *Proceedings of INTERACT'11*, Springer-Verlag (Berlin, Heidelberg, 2011), 461–476.

[13] Müller-Tomfelde, C., Hilliges, O., Butz, A., Izadi, S., and Wilson, A. Interaction on the tabletop: Bringing the physical to the digital. In *Tabletops - Horizontal Interactive Displays*, C. Müller-Tomfelde, Ed., Human-Computer Interaction Series. Springer London, 2010, 189–221.

[14] Oh, J.-Y., Stuerzlinger, W., and Danahy, J. Sesame: towards better 3d conceptual design systems. In *Proceedings of DIS '06*, ACM (NY, USA, 2006), 80–89.

[15] Wang, R., Paris, S., and Popović, J. 6d hands: markerless hand-tracking for computer aided design. In *Proceedings of UIST '11*, ACM (NY, USA, 2011), 549–558.

[16] Wilson, A., and Benko, H. Combining multiple depth cameras and projectors for interactions on, above and between surfaces. In *Proceedings of UIST '10*, ACM (NY, USA, 2010), 273–282.