3DUI 2010 Contest Grand Prize Winners
Peer-reviewed author version

The First 3DUI Contest Grand Prize Winners

Category: Research

Abstract

The IEEE Symposium on 3D User Interfaces 2010 has run its first 3DUI Grand Prize, a contest for innovative and practical solutions to classic 3DUI problems. This paper describes the rationale of this first contest, an analysis of all submissions, and details about solutions from winners in the two categories.

1 INTRODUCTION

[1] The first 3D user interfaces (3DUIs) were created during the 60s, and since then there has been active research on novel input devices, novel ways to visualize information, or novel ways to manipulate such information [2, 4]. 3DUIs can be found in Virtual Reality applications [5, 10] in a wide variety of domains, from games to health care solutions, from the oil and gas industry to car design. However, there is a rich set of devices that can be used to manipulate 3D information, and endless approaches for mapping input events to actions inside an application. The community of the IEEE Symposium on 3D User Interfaces (3DUI), co-located with the IEEE Virtual Reality Conference, gather annually to discuss novel advances and results in this field. This year, the 3DUI Symposium has created another method to show those advances, the 3DUI Grand Prize.

The goal of the 3DUI Grand Prize is to show common practices and advances in this field, by means of a competition around a problem that reflects our state of the art. We[2] want both to show how much has been done in this field and to encourage contestants to be innovative and passionate about the possibilities of 3DUIs. We want to recognize achievements and common knowledge, while at the same time underline innovations and novel approaches. In general, we believe a contest could be another way to show the

---

[1] original published by IEEE Computer Graphics and Applications

[2] As otherwise stated, *we* is used to refer to the organizers of the contest.

advances and capabilities of our community. Unlike formal papers in which authors come up with few variations for a task, a contest could show a wider variety of approaches that could empirically show what works and what does not.

This paper describes this first experience and reflects on the future of the contest. First, we describe the rationale of the selected problem and an overview of submissions. Later, we describe in detail winner solutions in the two categories of the contest. Finally, we present a discussion on the results and some conclusions.

SIDE NOTE: WEB LINKS

The following are some web links related to this contest:

- The 3DUI 2010 Grand Prize: http://conferences.computer.org/3dui/3dui2010/cfp/contest.html

- The 3DUI 2010's Channel at YouTube, with all the video submissions: http://www.youtube.com/user/3DUI2010#g/c/21451B298805BC09

- Mail archive for the contest: http://listes.lost-oasis.net/?A=LIST&L=3duigrandprize2010_nowan.fr

2 CONTEST AND PROCEDURES

For this first installment of the 3DUI contest, we decided to favor similarity to reality and simplicity in the evaluation process over more abstract scenarios that could lead to more subtle and precise evaluations, but at the same time require more preparation. As one of the organizers put it, we wanted to "mimic real problems that people [in industry] using 3D are facing". We did not know neither how many contestants will accept the challenge, so we had to simplify the work of evaluators in case of overwhelming submissions.

The task we decided after some discussions was to to find some objects in a supermarket that looks as Figure 1, and move them to a table. Copies of the target objects are found on top of the purple table, and objects are hidden below other objects in the aisles marked on top with a red dot board. Users should be able to travel inside the shop, find (select then manipulate) the object that is similar to the one on the purple table and then place it beside its copy in roughly the same orientation (travel then manipulate). Such a task was similar to a commercial solution a company was developing. Since we could not use a commercial

model, we developed a supermarket model that was exported to VRML, one of the most common open formats nowadays, with the hope that such a model could be easily readable from any tool contestants may use.

Figure 1: The Sample Supermarket

Two categories were created: one based on videos and one based on a live competition during the conference. The video category allowed teams to use, integrate, and innovate in terms of devices and hardware setups, any that may be available at their own facilities but may be difficult to transport to the conference's site. The live competition allowed participants of the conference to enjoy some innovative solutions and at the same time allowed contestants to try their solutions with real knowledgeable users.

For the evaluation of videos we used a scheme that counted on jury's experience in order to subjectively mark each demo. We asked juries to give a mark between 1 (worst) to 7 (best) in each one of the following

3

dimensions: navigation, selection, and manipulation techniques; innovation, generalizability, fitness of the solution, ease of use, and enjoyability. Jury members could also add some personal comments and a final mark, either subjective or computed from previous dimensions. In order to facilitate the work of the jury, a YouTube channel with all videos was created. Moreover, we also took advantage of Youtube's rating scheme, and allowed early registrants to vote for their favorite video. All juries and organizers met through Skype in order to make a decision for the video competition. We prepared for this meeting a summary with YouTube's ratings, juries' comments, and juries' final marks, which showed clear winners and therefore made detailed marks from all dimensions unnecessary.

We also encouraged contestants to optionally submit average time to complete the task and average frame rate from a set of 10 samples with users. 6 Teams provided us with such information, with values between 74 and 427 seconds for task completion time, and between 29 and 112 frames per second. There are confounding factors related to this data which were not measured, such as type of graphics card, operating system in use, and user's expertise, which could also affect these results and therefore precluded further analysis.

The live competition took place at the 3DUI Symposium. It was open to the public for about 4 hours and it required as much time to setup, especially for teams with special equipment. Each jury member was asked to run the task in each demo, record his time of completion, and rate the solution with an overall mark. The public was asked which demo they preferred when they were leaving the room, although there was no control about how many demos each individual tried or in which order. At the end of the live competition jury members met for about 30 minutes and took a decision based on their own marks and the public rates. Four teams from the 12 in the video category showed up at the conference, with different and interesting 3D interfaces.

3 VIDEO SUBMISSIONS IN SUMMARY

Table 1 shows teams in the video submissions. The numbers in such table are used in Tables 2 and 3, which show the variety of devices and interaction techniques they use, respectively[3]. Names for Interaction Tecniques try to follow the ones in [4] in order to facilitate comparison, and readers are invited to see more

---

[3]These classifications are subjective, based on the organizer's opinions.

Submission	Name	Institution
1	Fighting Gobblers	Virginia Tech, USA
2	Galaxy Rats	University of Hasselt, Belgium
3	GVU Twin Space	Georgia Tech, USA
4	Hank Lab	University of Iowa, USA
5	im.ve	University of Hamburg, Germany
6	WBS_UniBi	University of Bielefeld, Germany
7	Three guys and three girls	University of Hasselt, Belgium
8	UFGRS	Universidad Federal do Rio Grande do Sul, Brazil
9	UHasselt Student Team	University of Hasselt, Belgium
10	U of Minnesota Duluth	*idem*, USA
11	BU Weimar	University of Weimar, Virtual Reality Systems Group, Germany
12	VRM	VRM, Ukraine

Table 1: Team Names

details of these techniques there. Differences in referred names are shown in italics.

As we can notice, contestants used a wide variety of devices, with the Wiimote remote and several types of 6DOF trackers as the most popular. Although winner solutions mostly use commercial off-the-shelf (COTS) devices, the most appealing and interesting demos came from the integration of custom built ones. However, those interesting solutions presented issues in performance or interaction, which may show that these solutions require more time in order to find an optimum balance. Two groups tried 5 or 2 devices, and the rest used 3 or 4. About half of the contestants use 3D display technologies, but there is a wide variety of devices and image qualities. It is also noticeable that half of the contestants used 2D screens, despite the fact that most labs have now access to 3D displays. Finally, it is clear that is very unlikely we could bring all these solutions to just one place, due to the wide diversity of devices in use in labs around the world.

5

Devices	1	2	3	4	5	6	7	8	9	10	11	12
Input Devices												
Wii Balance Board	X					X						
Dance Pad	X											
Wii Remote	X			X	X	X		X		X		
Wii Sensor Bar	X											
Spacemouse		X							X			
Keyboard												X
Mouse									X			X
Gamepad												X
Tangible Props			X				X				X	
Tangible Cart			X									
Chair IO				X								
6DOF Tracker				X	X	X	X			X	X	
ART Progloves						X						
Haptic Master							X					
Output Devices												
Stereoscopic Screen							X				X	
Monoscopic Screen	X	X						X	X			X
Tabletop Display			X									
Touch Enabled Screen			X									
HMD				X						X		
CAVE					X	X						

Table 2: Devices in Video Submission.

Interaction Techniques	1	2	3	4	5	6	7	8	9	10	11	12
Travel Techniques												
Pointing Based Steering		X[4]			X	X	X		X		X	X
Gaze Based Steering *Continuous Input*	X					X						
Gaze Based Steering *Discrete Input*	X			X				X				
Walking			X							X		
Walking in Place						X						
Exocentric View		X							X			
Marking Points along a Path		X				X		X				
Travel by Grabbing Air						X						
Travel by World in Miniature						X						
Selection and Manipulation Techniques												
Ray Casting	X[5]	X	X		X	X	X[6]	X	X	X		X
Absolute Rotation	X		X							X		
Relative Rotation		X		X	X	X	X					X
Snapping to a Surface	X			X								
Constrained Manipulation Along an Axis		X										
Physical Props			X									
Image Plane Pointing						X						
Fishing Reel				X		X						
Gesture Based Techniques						X[7]						
Go-Go Selection											X	
Control Techniques												
Adapted 2D Menus	X						X		X			
Ring Menu		X										
3D Widgets												X

Table 3: Interaction Techniques in Video Submissions.

In terms of interaction techniques, we can see an even wider set of techniques in the solutions, although Pointing Based Steering was the most common travel technique and Ray Casting the most common selection technique. Some PC based solutions used a combination of travel techniques, in order to accelerate such a task. Although it does not show in this table, some immersive solutions also tried to accelerate movement, by means of different movement rates, for example. There are wide differences in visual feedback from similar interaction techniques, which also can not be seen from this table. Of particular interest are the menu adaptations. Mostly they were used in order to further select objects from a group, i.e. in the shopping cart, but there are wide differences in look and feel and visual appearance. A special mention is also deserved for Team 6, who implemented the most complete set of interaction techniques and made comparisons between several solutions. Although their results are very impressive, it was very difficult for juries to compare their submission with others.

4 THE LIVE CATEGORY WINNERS

This section presents the solution that won the live category of the 3DUI contest, designed and implemented by the "Fighting Gobblers"[8] team from Virginia Tech. Our goal was to use accessible hardware to provide a complete 3DUI that compares to complex VR systems.

4.1 Overall Description

We challenged ourselves to use consumer-level devices. Although many useful and usable 3DUIs have been designed in research labs, the hardware to enable interaction techniques for 3DUIs can be very expensive and inaccessible [13]. However, in recent years, novel spatial input devices have become available in video game systems such as the Nintendo Wii. These low-cost controllers offer new possibilities for the design of more accessible 3DUIs, but they lack the precision and technological advancement of specialized hardware [13].

With that in mind, our main motivation for using commodity hardware was to explore the feasibility

---

[4]With a zoom lense.

[5]Plus a sphere for multiple selection.

[6]Plus a sphere for multiple selection.

[7]Move and scale based on proprioception.

[8]"Fighting Gobblers" was the nickname given to the athletic teams at Virginia Tech prior to 1978.

of creating a complete, general-purpose 3DUI from these consumer-level devices. We also wanted to determine how established 3D interaction techniques needed to change in order to accommodate such input devices while maintaining overall usability.

Figure 2: Overview of the setup used in the design of the live contest winner.

We chose the following devices, designed to be used with the Wii, to implement our interface (see Figure 2).

- Wii Remote: provides a three-axis accelerometer for motion and orientation tracking, in addition to a variety of buttons

- Wii Sensor Bar: offers infrared (IR) tracking for pointing when paired with the Wii Remote

- Wii MotionPlus: adds a 3D rotational sensor to the Wii Remote for more precise orientation tracking

- Nunchuk: includes accelerometers and a 2D analog joystick, plus two buttons

9

- Wii Balance Board: has four weight sensors that can be used to detect leaning

- Dance Mat: provides discrete buttons meant to be stepped on

Although commercial Wii games often allow some form of 3D interaction using the devices listed above, the tasks are usually very specialized for each game. For example, controlling a tennis racket or a golf club is done through gestures with the Wii Remote, but these gestures are not meaningful as general-purpose 3D interaction techniques. We wanted to use Wii input devices to build generic, context-insensitive 3D interactions for the universal tasks of travel, selection, and manipulation [4].

Supporting this sort of generic 3D interaction with the devices we chose is challenging, because unlike the 6-DOF tracking systems typically used in research labs, these controllers provide limited interaction space and limited tracking. The limited interaction space constrained the possibilities for different interaction tasks. For example, we could not design a navigation technique that required physical locomotion, because the Balance Board and Dance Mat require users to be standing on or within a step of them. The Wii Remote's IR tracking system requires users to be within a certain range of the Sensor Bar and to be pointing in its general direction. The accelerometers combined with the 2D IR tracking did not allow us to determine the position of the user or the hand, but only to determine where on the screen the user is pointing. These limitations make it difficult to use standard ray-casting or virtual hand techniques. Further, the accelerometers alone do not provide enough information about orientation, failing to offer absolute rotation values for the Nunchuk and the Wii Remote. Even though the MotionPlus provides better orientation information with its gyroscopic tracking, it still has significant drift issues. This relative tracking (instead of the absolute tracking found in typical 6-DOF devices) makes it even more difficult to use direct manipulation techniques.

4.2 Design Process

We began the design process as soon as the contest and the initial rules were announced. The five members of the team met once a week from that time until the week prior to the contest. In our weekly meetings, we brainstormed about interaction metaphors and devices that we had available. We also mapped the devices to each of the tasks (travel, selection and manipulation). This brainstorming brought up many questions about the contest rules and restrictions, and helped us have a better understanding of how the

contest would work. By the time registration was open and the initial 3D model was released, we already had several ideas laid out, giving us time to decide which ones to choose and which to throw away based on the example model and the specific tasks. It was during these first meetings that we settled on the use of commodity hardware.

An illustrative example of a design decision we made in this phase came from our discussion of the selection task. From the example supermarket model, we noticed that selecting an item in a single step would be a difficult task, since the 3D model was designed to be very cluttered, with a high degree of occlusion. Common 3D selection techniques such as Go-Go [9] or ray-casting [8] would not work well in this environment, because they require high precision from both the input device and the user to select a specific desired object. Since no selectable items were too small, users could still be precise if they travelled very close to the desired item, but that would add in a long navigation phase that we thought wasn't necessary. That's when we decided that selection would be better as a two-phase process, with a low precision, multiple-object selection followed by a refinement phase.

After narrowing down the techniques and devices we would use, we started prototyping, using the Vizard Virtual Reality Toolkit[9]. Vizard is a Python-based toolkit that allows for rapid prototyping of virtual reality applications. We wrote a dll extension in C++, which integrated the WiiYourself[10] library and allowed access to the Wii tracking data directly from Vizard. Our weekly meetings were then split into two parts: new design ideas and refinements, and prototype evaluation. As soon as we had a working prototype that supported all tasks, we presented it to our research group once a month to assess its current state. This process helped us identify major usability issues and gave us insights on how to improve the design of our techniques to solve those issues.

4.3 Designed Techniques

In order to overcome the limitations introduced by the input devices, we followed the basic principle of decomposing the interaction tasks (navigation through the supermarket, selection of objects, and object placement) into several simpler components.

Navigation was divided into three subtasks based on different degrees of freedom users have to control:

---

[9]http://www.worldviz.com/products/vizard/index.html

[10]http://wiiyourself.gl.tter.org/

horizontal translation, vertical translation, and orientation. This choice was made to overcome the limited control space of our input devices [7]. For translation on the horizontal plane, we used the Balance Board and the Dance Mat (see Figure 2). The Balance Board is used for slow, precise movements by leaning in the direction of travel. The Dance Mat is used for faster movements, and works by stepping in the desired direction of travel. We also implemented discrete speed control for fast navigation. When one foot was on the Balance Board and one on the Dance Mat, travel was at half speed, whereas to travel full speed, the user would step with both feet in the desired direction. The Nunchuk's front buttons (C and Z) were used for vertical translation at a constant speed; this is a natural mapping since the buttons are on a vertical surface of the device. The combination of these techniques provides a balance of precision and control for the three degrees-of-freedom of translation. We used the 2D joystick on the Nunchuk to control viewpoint orientation (pitch and yaw). We did not implement a control for viewpoint roll, as it would add complexity and confusion without enhancing navigation.

3D object selection is traditionally performed in a single step. In a cluttered environment such as the supermarket, however, this can require high precision (for pointing or touching), significant navigation (to reach a location where the object is reachable, visible, or large enough to select), or both. To achieve rapid yet precise selection with our gaming devices, we divided selection into two phases rough selection and refinement. Unlike the subtasks we used in navigation, which can be performed simultaneously, these selection phases must be performed sequentially. Figure 3 illustrates the two phases.

The rough selection phase uses a modified version of ray-casting, based on IR tracking of the Wii Remote, which casts a sphere onto the nearest intersecting surface to determine which objects will be selectable. We call this subtask sphere-casting. The user simply has to ensure that the desired object is inside or touching the sphere, so that it can be picked from among the other objects in the next phase. Items that will be made selectable are highlighted. In order to improve confidence that the desired object will be available, the sphere's radius increases the farther the user is from the nearest intersecting surface, thus increasing the overall number of objects available in the second phase. For extreme cases, the user also has control over the size of the sphere, by pressing the Wii Remote's '+'and '-'buttons. Note that this technique can be used at any distance from the desired object, and even when the object is invisible.

Clicking the front button (A) on the Wii Remote causes all objects that were inside or touching the

sphere to be distributed among four quadrants on the screen, which we call the quad-menu. In the quad-menu phase, users refine the selection by repeatedly pointing anywhere in the quadrant that contains the item they are looking for, each time reducing the number of objects per quadrant until the desired object is the only one left. The maximum number of selections necessary in the quad menu is $\lceil \log_4 n \rceil$, where n is the number of items inside the sphere. Items in the quad menu rotate so that the desired item is more easily distinguished. Once the item is selected, it is moved to an item inventory, which is shown in the bottom right corner of the display. In the 3DUI contest task, this inventory could be thought of as a shopping cart.
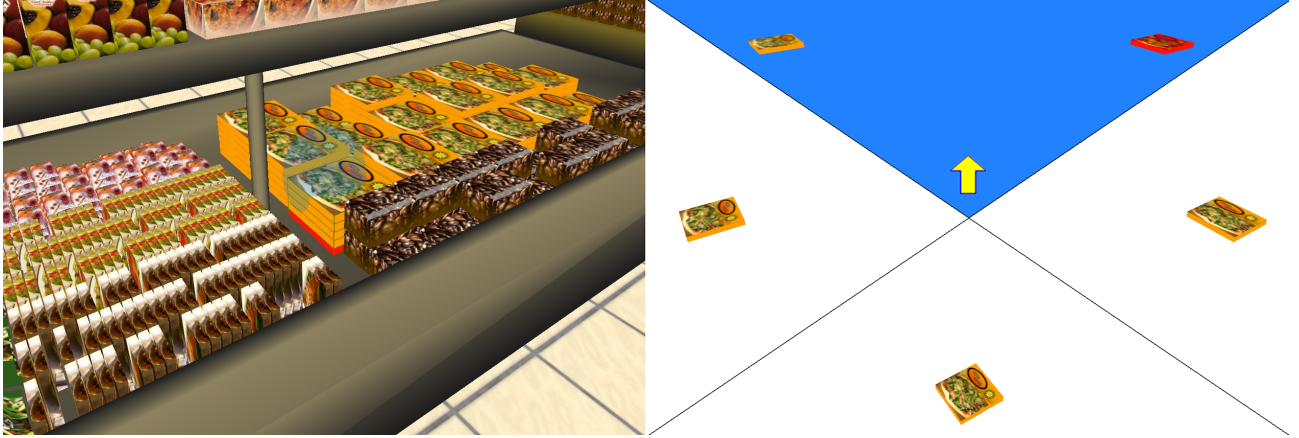

Figure 3: The sphere-casting and quad-menu techniques used for selection.

We also divided the manipulation task into three smaller subtasks: selecting an item from the inventory, setting its position, and setting its orientation. The first phase of manipulation is actually an instance of the quad-menu phase of our selection technique. The user points to the inventory to bring up a quad-menu of all the items that have previously been selected. The item placement phase uses ray-casting to position the selected object at the first intersecting point in the environment. In the last manipulation phase, the user sets the object's orientation using a direct one-to-one mapping of the Wii Remote's orientation, which we achieve using the rotational sensors in the MotionPlus attachment. We maintain contact between a face or edge of the object and surface selected in the previous phase, avoiding the possibilities that part of the object penetrates the surface or that the object floats above the surface. Setting object orientation was the most difficult task in our interface, due to the drift in the MotionPlus sensors and the difficulty of holding the Wii Remote in certain orientations. Still, our choice to decompose 6-DOF manipulation into two phases provided the user with moderately fine control of both the position and the orientation of the

object, even with the inherent limitations of our input devices.

Our interface also included a few other tools that helped users complete the tasks. The inventory allowed users to carry more than one item at a time, so that it was not necessary to go back to the table every time an item was found. This inventory is always visible in the lower right portion of the screen, as shown in Figure 3. Another useful tool was the ability to take a screen shot and display it in the upper left corner of the screen. By taking a screen shot of the table, users could see which objects they had to find at any time. Finally, we allowed teleportation back to the starting position when the user pressed the Home button on the Wii Remote.

4.4 Lessons Learned

The design process we used was quite valuable to our team. The weekly meetings allowed us to improve the design and implementation of our prototypes. Listing devices and metaphors helped us to narrow down which ones were appropriate for the given scenario and tasks. Moreover, informally evaluating the different prototypes allowed us to quickly determine how to improve the techniques and the overall design of our application. For example, by adding background music and sound effects, the task of searching for the items became more fun and engaging.

By decomposing the tasks into several simpler subtasks, we were able to use low-cost commodity devices to perform general-purpose 3D interaction. Using multiple rapid coarse-grained selections instead of a single precise selection, we made this interface well-suited for cluttered environments and imprecise pointing devices. By splitting the manipulation task into position and orientation phases, our technique achieves the same expressiveness as 6-DOF manipulation techniques based on expensive, specialized hardware.

However, decomposing the tasks may also increase their complexity. During the live contest, our navigation interface, in which different degrees of freedom are controlled simultaneously by different devices and actions, proved to have a steeper learning curve than our selection and manipulation techniques, which were broken down into sequential phases. In addition, some design choices limited the extent to which our techniques can be used. For example, our selection technique was designed for selecting objects that can be uniquely identified based on their shape and color. It is not suited for selecting objects based on their

spatial location or context, because the quad-menu phase provides no spatial or contextual information. Another example of a minor limitation is that our manipulation technique only allows objects to be placed on surfaces, although this is a reasonable constraint for many applications.

Perhaps the most novel and successful aspect of our interface is the two-phase selection technique. The sphere-casting + quad-menu design illuminates some interesting tradeoffs – primarily the tradeoff between speed and precision. If a selection task is difficult enough (due to small or distant objects, occlusion, or poor tracking performance), then multiple actions with low required precision may be faster than a single action requiring high precision. Traditional ray-casting selection would certainly perform better in environments with large objects and little occlusion. However, the power of our selection technique lies in its ability to use imprecise (even careless) pointing behavior when selecting an object that is visually small and/or presented in a cluttered environment.

Another fair comparison is between sphere-casting and cone-casting [8]. With cone-casting, a cone is projected from the input device onto the environment, and all objects that fall inside the cone are eligible for selection. Sphere-casting reduces the selection volume and is more useful in environments containing objects laid out roughly in a plane. Had we used cone-casting as our selection technique, objects that were on shelves behind the intended one would be selected, and in much greater quantities, since the cone's base area expands quadratically as it becomes longer. This could result in a quad-menu with an immense number of items, rendering them too small to be distinguished.

Finally, the quad-menu selection has some interesting characteristics. The growth in the number of clicks necessary to complete a selection increases logarithmically, with 64 items needing three clicks, 256 items requiring four clicks, and 1024 items needing only five clicks to complete a selection. There is, of course, a practical limit, based on the display size and resolution, on the number of items displayed in the quad-menu. As noted earlier, the desired item must also be visually distinct in order for this technique to work at all.

Certainly, there are improvements that can be made to our sphere-casting + quad-menu technique. For one thing, if only one item is inside the sphere, the quad-menu could be suppressed, and the selection could be performed with a single click. The design of the technique could also change in order to account for spatial context, rather than object uniqueness. For example, instead of showing items in a quad-menu,

the items that were inside the sphere could be displayed in place using different shades, keeping them in their spatial context.

Although the sphere-casting + quad-menu selection technique is promising, we need to verify its effectiveness empirically. We are planning a formal evaluation to demonstrate the actual benefits of our technique and to explore the speed-precision tradeoff.

The 3DUI contest provided a great stimulus for our creativity. We feel it was very successful in pushing us to think outside the box. We were able to develop techniques that allowed general-purpose 3D interaction to be performed with high levels of usability based on consumer-level technologies. We look forward to future editions of the contest.

5 THE VIDEO CATEGORY WINNERS

The team which won the video category are the "Three girls and three guys", a team from the Expertise Centre for Digital Media, a research institute of Hasselt University in Belgium. In the following sections the members of the team will describe their solution as well as the development process which was used and the lessons learned from participating in the contest.

5.1 Overall Description

The starting viewpoint from our solution is derived from primarily two observations with regard to shopping in daily life:

- In a grocery store a shopping cart is typically used.

- Many couples perform their grocery shopping together where one person handles the shopping cart and the other person gathers the items from the shelves.

We incorporated both observations into our solution. One person has to control an input device similar to a shopping cart for navigation while another person has to control a toy gun and aim at the items on the shelves to be able to select and manipulate these and finish the task successfully.

An overview of our solution can be seen in Figure 4. For navigation, our solution consisted of a HapticMaster with a self-made handle resembling the handle of a shopping cart. To provide input for

Figure 4: Overview of the setup of the video contest winner.

selection and manipulation a toy gun with built-in flock of birds tracker and two custom USB interfaced buttons (dominant hand) in combination with another flock of birds tracker attached to a glove (non-dominant hand) were used. For visual feedback a polarisation projection screen (2.4 m x 1.8 m) with passive stereo using two DLP projectors was used.

5.2 The Development Process

The process involved to develop our solution can be split in two parts: the design part and the implementation part. Due to problems with the virtual model (see Section 6) we decided to build our own custom made virtual environment framework based on several open source solutions as our own in-house framework could not reasonably quick be adapted to render and load such a model efficiently. We combined the popular open source graphics rendering engine OGRE [11] with VRPN [11] for tracker support as well as HAPI from H3D [12] for the HapticMaster. We transformed the VRML model we received into

---

[11]The Open Source 3D Graphics Engine is available at `http://www.ogre3d.org/`.

[12]The H3D API is available at `http://www.h3dapi.org/`.

the DotScene format commonly used by OGRE users. In order to make sure that interactive frame rates were achieved throughout the application we also implemented occlusion culling inside OGRE by placing occlusion planes between the backsides of adjacent shelves.

In order to come up with a design of our solution we held weekly meetings in which we started from our viewpoint on shopping as described in the previous section. Firstly, we wanted to simulate the behavior of a shopping cart. Initially we were considering solutions involving a real life shopping cart in combination with tangible objects rather similar to the solution presented by "GVU Twinspace" (see Section 3). After consideration this solution would have a lot of complexities with regard to the realisation and would be rather hard to generalise which was something found important by the team. This lead us to the solution of using the HapticMaster which is a haptic device that can generate force feedback of a higher strength and with a larger workspace than typical haptic devices such as the PHANToM and Novint Falcon. Furthermore, it can also be used while standing which is important when trying to mimic a shopping cart. Another advantage of the HapticMaster is that the end effector attached to the tip of the device can be changed which allowed us to replace it with a self-made handle similar to the one attached to a shopping cart (see Figure 5). Finally, in order to be able to further create the feeling of pushing a shopping cart, realistic force feedback was generated. When the users pushed the handle in any direction (primarily forward or backward) they were pushed back using a virtual spring force. This generated force gives the perception of pushing in the direction in which the users want to navigate. The users are pushed back towards the home zone, in which no movement will occur. In order to be able to steer, the users rotate the handle and implicitly push in to the direction they want to move. Finally, to be able to look up and down it is possible to either pull up or push down the shopping cart handle.

Besides the design of the shopping cart which covered the navigation aspect of our solution we primarily focused on an efficient selection technique. The items which had to be found and picked from the shelves were typically occluded from the viewpoint and closely surrounded by many other shopping items. Therefore, from the beginning the decision was made to focus on a two step selection technique in which first the interesting shopping items would be identified and the final item would be selected using a second step. This meant we needed two buttons to easily and intuitively map all the actions onto respective buttons (e.g. to support cancelling an action). As an input device we embedded a 6DOF magnetic tracker

Figure 5: The Hapticmaster with shopping cart handle used for navigation.

inside a toy gun which we also provided with a trigger button and button reachable with the thumb for less often used actions. We used for our tangible prop a toy gun as this is intuitive for aiming at a screen. The tracker was calibrated such that where the gun pointed on the projection screen it could be used as an aiming point onto the 2D screen projected into the 3D world. This made pointing onto the screen intuitive. This point was visualised as a small blue transparent circle.

In order to support the user in finding the appropriate item we experimented with a similar technique as the BalloonProbe [6] which would scatter dense objects to the boundaries of a selection volume (in our case a sphere). Continuous movement of the BalloonProbe proofed to be too distracting (i.e too much popping of items in and out of the volume), therefore we resorted to keeping the sphere of the BalloonProbe but apply transparency as had been previously successfully used during selection techniques in earlier research performed by members of the team [12]. Besides transparency we also highlighted all objects using a "light up"effect (see Figure 6). This is different from traditionally using a certain color but because items had many different colors and textures we could not pick one color. Furthermore we wanted the ability to resize the selection volume such that the user could control the amount of objects being hit because not all areas in the virtual world have the same density of objects. In order to incorporate this the distance between both hands was used to resize the spherical selection volume. If the non-dominant hand came into a certain range of the dominant hand, a range was defined such that if the non-dominant hand was not used the spherical selection volume would not become too huge, the spherical selection volume would start to resize until the hands meet and the size would be almost a point-in-space.

When satisfied with the items currently being highlighted by the selection volume the first step of

the selection technique has to be completed by pressing a button (trigger of the toy gun) which would temporarily remove all the items from the world and place them in a square menu. An item could be selected by pointing at it and again pulling the trigger, if the menu would not contain the correct items the menu could be cancelled with the second thumb button. The possible selected item will be attached to the pointer and the other items are put back into the world. We opted for a square menu because this would most efficiently make use of the effective width of the items, this has recently been investigated and proven to be a very good menu structure [1]. The menu can be seen in Figure 6. To make sure that the visibility of the items inside the menu is optimal they are resized and rotated to nicely fit inside an item.
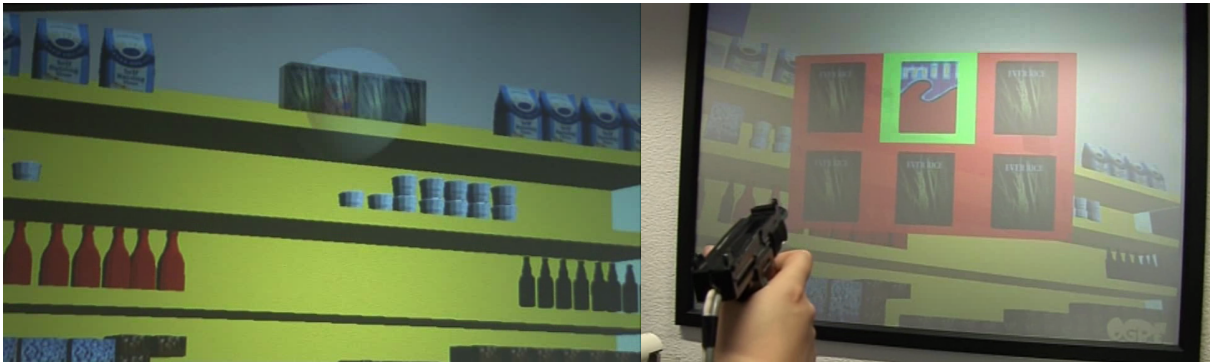

Figure 6: The spherical selection volume on the left hand side and the menu on the right hand side.

Finally when an item is selected, it can either be put back in the shop using the second thumb button or placed (translated) in the world where the user is pointing by pulling the trigger. To make sure that the object would translate intuitively we made sure that the object being translated would be placed on the object which the user was currently pointing at. This way the object would not float in the air and it would also be easy to place it on another item such as the table. Thus for manipulation translation is performed separately from rotation. When placed in the world the rotation is activated. The object closest to where it is placed (unless no object is within 1 meter) is copied as a transparent ghost to the position of the object that has been placed. This transparent copy can then be efficiently used to see the final required rotation. As we have an absolute input device we also provide the possibility to clutch the rotation by using the second thumb button. To finalise the manipulation the user has to pull the trigger.

5.3 Lessons Learned

Several lessons were learned from our participation, first of all taking part in such a contest with the complete virtual environment team from our lab (augmented with one computer graphics researcher) was a new experience for everyone involved. Also a clear definition of a problem that needed to be solved by any means available to us really allowed us to brainstorm about a specific problem at hand. We noticed that a weekly meeting discussing the new interaction mechanisms implemented during the previous week really boosted the iteration speed of solutions and thrived us towards solutions which satisfied our needs. Furthermore, the competition would allow us to compare creative solutions from the same problem with other labs. Finally, the competition also allowed us to create a complete solution for a specific problem, something that typically is not done in a research project. Usually one interaction technique type such as selection, manipulation or navigation are tackled solely and integration with other techniques is of less importance. This integration of all the different techniques is very important to create an intuitive and natural flow during the interaction and thus provided extra challenges.

Besides designing and implementing the solution a video submission needed to be created. We wanted to provide the viewer with a fun overview of our solution while remaining to be informative. As we had used a toy gun for our selection and manipulation thoughts of writing a scenario involving actors from action movies quickly arose. We learned it can be a lot of fun to disseminate 3DUI research using a James Bond scenario applied to the task which had to be solved. To make sure that the video would provide enough information about the solution a mission debriefing was included containing a conversation between both actors explaining our interaction techniques.

Our solution contains potential interesting techniques but is not perfect and contains some drawbacks. On the interaction technique level, primarily the selection technique in two phases is a mechanism that has not been extensively researched. Usually a one phase selection is preferred, but for very dense worlds, or even worlds in which objects are aligned a two phase approach seems very promising and warrants for a deeper investigation of the trade-off between a one and two phase selection technique. A drawback with regard to our chosen interaction techniques is our virtual shopping cart. Even though we mimic the behavior of a real shopping cart we did not add a virtual shopping cart to our solutions and an item currently selected had to be immediately brought back to the table. Unfortunately during the design process of our

21

solution we didn't come up with a good integration of the shopping cart, which would control the items in it, how would we put them in and out of the virtual cart, questions which we on purpose left unanswered. On the other hand our shopping cart with handle was found to be really intuitive, when users informally tested our solution they immediately grasped the concept and were able to navigate freely around the shop.

As solution we immediately opted for a collaborative two person application. A nice side effect is that besides interaction in a 3D world also a lot of fun real-life interaction occurs between two persons co-located in front of the same screen collaborating in the same world and sharing the same experience. While we did not tackle the solution from a single user perspective, it is possible to perform the task as a single user when the toy gun is attached to a belt and can be stowed away by the user when he wants to navigate and vice versa it can be taken when he wants to select and manipulate something. It is possible to release our input device and switch to the shopping cart but the switching between these tasks is not so time efficient with our current solution. Another downside form a single user perspective is that after the selection of an item that item has to be manipulated immediately but a single user would have to switch to navigation which might be very confusing as two tasks are "active"at the same time, while with two users these tasks are nicely divided across both.

6 OVERALL DISCUSSION AND CONCLUDING REMARKS

We are very grateful for the participation of our community in this first 3DUI Grand Prize. Solutions were very interesting and innovative, and we believe we accomplished our goals. We want to congratulate each one of the participants for their effort and willingness to endure the difficulties of this first contest.

There are several things we would like to address for future installments. Comparison of solutions from just a video is very hard, and we have to come up with ways to test solutions at a distance. A more clear set of metrics from the very beginning will also be important, in order to both provide hints about expected solutions and facilitate evaluation. It may be also desirable to have several categories, with different metrics, in order to encourage different approaches. For example, some solutions were very innovative and intuitive, but they were outperformed by simpler, PC based solutions. In general, we would like to aim a scenario in which we could "... provide grounds for direct comparison studies with different measures. We could envision, for example, measures like how efficient, how effective can a task

be performed, are people happy with it, is it intuitive, short learning curve, how fast is travel performed, how well is picking performed, how easy is it to place things at new positions, how precisely can one place them, and so on." [3].

It is also important to find out ways to better present and compare results from several contestants. Tables in Section 3 are an initial attempt to summarize results, but there are too many details left behind: differences in render quality, aids such as shopping carts and moving objects[13], interesting props[14], vibration and audio feedback, among others elements, are not visible from those tables.

A better virtual model should also be provided, one that tools can read and manipulate at ease. In this regard, we found very little support for interaction from the models we exported from commercial tools to VRML. Specifically, exported versions did not keep shared geometries, object definitions, or transform hierarchies, key elements for interactive 3D applications. These issues forced some teams to redo their models from the original ones, or create their own files in other formats. In this regard, 3D creation tools can do a better job at keeping the information of objects, identifiers, transformation hierarchies, and multiple references to a particular geometry. Other issues with the model were related to units, number of polygons generated by the exporting tool, and lack of optimization for textures[3].

Finally, we should aim for a balance of difficulty of tasks for different types of users. Novice users spent most of the time looking for objects, but this difficulty disappeared once users became knowledgeable on this scenario. This issue clearly affects average time of completion. It is also important to take into consideration user's cognitive load in the task design, although this could also be part of the set of metrics for evaluation and therefore part of the solution.

REFERENCES

[1] D. Ahlström, A. Cockburn, C. Gutwin, and P. Irani. Why it's quick to be square: modelling new and existing hierarchical menu designs. In *CHI '10: Proceedings of the 28th international conference on Human factors in computing systems*, pages 1371–1380, New York, NY, USA, 2010.

---

[13] See video from team 8.

[14] See videos from teams 3 and 11

[2] J. Barrilleaux. *3D user interfaces with Java 3D*. Manning Publications Co., Greenwich, CT, USA, 2001.

[3] S. Beckhaus. Personal email, February 2010.

[4] D. A. Bowman, E. Kruijff, J. J. LaViola, and I. Poupyrev. *3D User Interfaces: Theory and Practice*. Addison Wesley Longman Publishing Co., Inc., Redwood City, CA, USA, 2004.

[5] G. C. Burdea and P. Coiffet. *Virtual Reality Technology*. John Wiley & Sons, Inc., New York, NY, USA, 2003.

[6] N. Elmqvist. Balloonprobe: Reducing occlusion in 3d using interactive space distortion. In *VRST '05: Proceedings of the ACM Symposium on Virtual Reality Software and Technology*, pages 134–137, 2005.

[7] R. J. K. Jacob, L. E. Sibert, D. C. McFarlane, and M. P. Mullen, Jr. Integrality and separability of input devices. *ACM Trans. Comput.-Hum. Interact.*, 1(1):3–26, 1994.

[8] M. R. Mine. Isaac: A virtual environment tool for the interactive construction of virtual worlds. Technical report, Chapel Hill, NC, USA, 1995.

[9] I. Poupyrev, M. Billinghurst, S. Weghorst, and T. Ichikawa. The go-go interaction technique: non-linear mapping for direct manipulation in vr. In *UIST '96: Proceedings of the 9th annual ACM symposium on User interface software and technology*, pages 79–80, New York, NY, USA, 1996. ACM.

[10] W. R. Sherman and A. B. Craig. *Understanding Virtual Reality: Interface, Application, and Design*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2002.

[11] R. Taylor II, T. Hudson, A. Seeger, H. Weber, J. Juliano, and A. Helser. VRPN: A device-independent, network-transparent vr peripheral system. In *In Proceedings of the ACM*, pages 55–61, 2001.

[12] L. Vanacken, T. Grossman, and K. Coninx. Multimodal selection techniques for dense and occluded 3d virtual environments. *International Journal on Human Computer Studies*, 67(3):237–255, 2009.

[13] C. A. Wingrave, B. Williamson, P. D. Varcholik, J. Rose, A. Miller, E. Charbonneau, J. Bott, and J. J. LaViola Jr. The wiimote and beyond: Spatially convenient devices for 3d user interfaces. *IEEE Comput. Graph. Appl.*, 30(2):71–85, 2010.